

EMBASE versus MEDLINE for family medicine searches

Can MEDLINE searches find the forest or a tree?

Thad Wilkins, MD Ralph A. Gillies, PHD Kathy Davies, MLS

ABSTRACT

OBJECTIVE Many physicians access electronic databases to obtain up-to-date and reliable medical information. In North America, physicians typically use MEDLINE as their sole electronic database whereas in Europe, physicians typically use EMBASE. While MEDLINE and EMBASE are similar, their coverage of the published literature differs. Searching a single literature database (eg, MEDLINE or EMBASE) has been shown not to yield all available citations, and using two or more databases yields a greater percentage of these available citations. This difference has been demonstrated in a variety of disciplines and in family medicine using the term “family medicine,” but differences have not been shown using specific diagnostic terms common in family medicine. We sought to determine whether searching EMBASE with terms for common family medicine diagnoses yields additional references beyond those found by using MEDLINE alone.

DESIGN Literature search comparison.

SETTING An academic medical centre in the United States.

INTERVENTIONS Fifteen family medicine topics were selected based on common diagnoses in US primary care health visits as described in a National Health Care Survey on Ambulatory Care Visits. To promote relevance to family medicine physicians and researchers, the qualifiers “family medicine” and “therapy/therapeutics” were added. These topics were searched in EMBASE and MEDLINE. Searches were executed using Ovid search engine and were limited to the years 1992 to 2003, the English language, and human subjects. Total, duplicated, and unique (ie, nonduplicated) citations were recorded for each search in each database.

MAIN OUTCOME MEASURES Number of citations for the 15 topics.

RESULTS EMBASE yielded 2246 (65%) of 3445 total citations, whereas MEDLINE yielded 1199 citations. Of the total citations, only 177 articles were cited in both databases. EMBASE had 2092 unique citations to MEDLINE’s 999 unique citations. EMBASE consistently found more unique citations in 14 of the 15 searches ($P = .0005$).

CONCLUSION Overall, EMBASE provides twice as many citations per search as MEDLINE and provides greater coverage of total retrieved citations. More citations do not necessarily mean higher-quality citations. In a comprehensive search specific to family medicine, combined EMBASE and MEDLINE searches could yield more articles than MEDLINE could alone.

EDITOR’S KEY POINTS

- Family medicine researchers and practitioners in North America traditionally rely on MEDLINE to search for relevant references. This study examined how many additional references would be found if the European EMBASE were also searched.
- For 15 common primary care diagnoses, both EMBASE and MEDLINE were searched for citations dealing with therapy. EMBASE provided more than twice as many citations as MEDLINE. Only about 5% of the references were listed on both databases.
- More citations do not necessarily mean better citations, but EMBASE provided more clinical trials. Thus, to ensure a comprehensive search, family physicians in North America are encouraged to use EMBASE in addition to MEDLINE.

This article has been peer reviewed.

Full text available in English at www.cfpc.ca/cfp

Can Fam Physician 2005;51:848-849.

Many physicians access electronic databases to obtain up-to-date and reliable medical information. In North America, physicians typically use MEDLINE as their sole electronic database, whereas it is common in Europe for physicians to use EMBASE.¹ While EMBASE and MEDLINE are similar electronic databases, there are differences in coverage.²

Prior studies^{1,3-14} comparing the comprehensiveness and relevance of these databases as well as other databases (ie, PsycINFO, BIOSIS) have demonstrated that a single search engine does not capture all of the pertinent and available articles, and using two or more databases provides greater coverage of all possible citations. These findings have been replicated in searches in a variety of specialties including psychiatry, rehabilitation medicine, biomedicine, and nutrition.

Rosser et al¹⁵ demonstrated how terms such as “family practice,” “family medicine,” and “general practice” mapped differently to articles in EMBASE and MEDLINE, but did not draw any conclusions about specific health-related terms. Common diseases encountered in everyday practice of family medicine (FM), such as hypertension, diabetes, or sinusitis, have not been specifically examined. Assessing whether differences in database coverage found in other disciplines are as prominent in FM is a salient research question that needs to be addressed. Does adding an EMBASE search to a MEDLINE search yield significantly more citations for FM-specific searches? The answer to this question could have important implications for FM researchers and clinicians who might want to search more than one database. The objective of this study was to determine the difference in search retrieval between EMBASE and MEDLINE for common FM topics.

Dr Wilkins and Dr Gillies are Assistant Professors in the Department of Family Medicine at the Medical College of Georgia, and Dr Gillies is a member of the HamesNet research network. Ms Davies is an Education Services Coordinator and Assistant Professor in the Greenblatt Medical Library at the Medical College of Georgia.

METHODS

Fifteen topics were selected from the National Health Care Survey on Ambulatory Care Visits to physicians' offices, outpatient departments, and emergency departments.¹⁶ Diagnoses labeled as other infections, other disorders, and other diseases were excluded from consideration as being insufficiently described to determine appropriate search strategies. Search topics reflected a broad scope of medical conditions including diabetes, obesity, asthma, otitis media, and depression.

Once topics were selected, researchers devised terms to restrict the search to FM literature available in each database. Subject headings were chosen for each database to retrieve articles containing one of three FM concepts: the practice of FM, the profession of FM, and provision of primary health care. Each topic was represented by the official database term and the addition of the subheading “therapy” to focus the search on treatments used for the conditions. The term “therapy” was used for EMBASE and “therapeutics” for MEDLINE in the search strategies.

All searches were executed using the Ovid search engine. First, the terms for FM were entered and combined using the Boolean “OR” to make one set. Second, each topic term was entered and combined with the subheading “therapy.” Then the topic terms were entered unrestricted, followed by the entry of the term “therapy” or “therapeutics.” These search sets were then combined with the FM set to retrieve articles on therapy in FM on each specific topic. The final set of articles was limited to English language, human subjects, and a publication date of 1992 to 2003 (**Figure 1**).

Results from individual EMBASE and MEDLINE searches were downloaded to EndNote, a bibliographic database organizer for analysis, and stored in separate files. The notes field for each citation was assigned the term EMBASE or MEDLINE to distinguish the database source for each record. The results for EMBASE and MEDLINE for the individual topics were then combined into a joint library or file. Each of the joint topic libraries was examined and duplicate citations were removed. Each of the

Figure 1. Sample search using EMBASE and MEDLINE

DATABASE: EMBASE 1970- PRESENT		DATABASE: MEDLINE 1966- PRESENT	
SEARCH STEP #	SEARCH STRATEGY	SEARCH STEP #	SEARCH STRATEGY
1	exp Family medicine / (2339)	1	exp Family practice/ (40212)
2	exp General practitioner/ (13434)	2	exp Physicians, family/ (8438)
3	exp Primary health care/ (17258)	3	exp Primary health care/ (34257)
4	1 or 2 or 3 (30128)	4	1 or 2 or 3 (77428)
5	exp Urinary tract infections/th [Therapy] (235)	5	exp Urinary tract infections/th [Therapy] (924)
6	exp Urinary tract infections/ (13843)	6	exp Urinary tract infections/ (27387)
7	exp Therapy (1081156)	7	exp THERAPEUTICS/ (1488843)
8	6 and 7 (2750)	8	6 and 7 (5266)
9	5 or 8 (2985)	9	5 or 8 (5855)
10	4 and 9 (73)	10	4 and 9 (82)
11	limit 10 to English language (68)	11	limit 10 to English language (75)
12	limit 11 to human (67)	12	limit 11 to human (74)
13	limit 12 to yr=1992-2003 (65)	13	limit 12 to yr=1992-2003 (30)

exp—explode command selects any record with this specific term, resulting in comprehensive search retrieval.

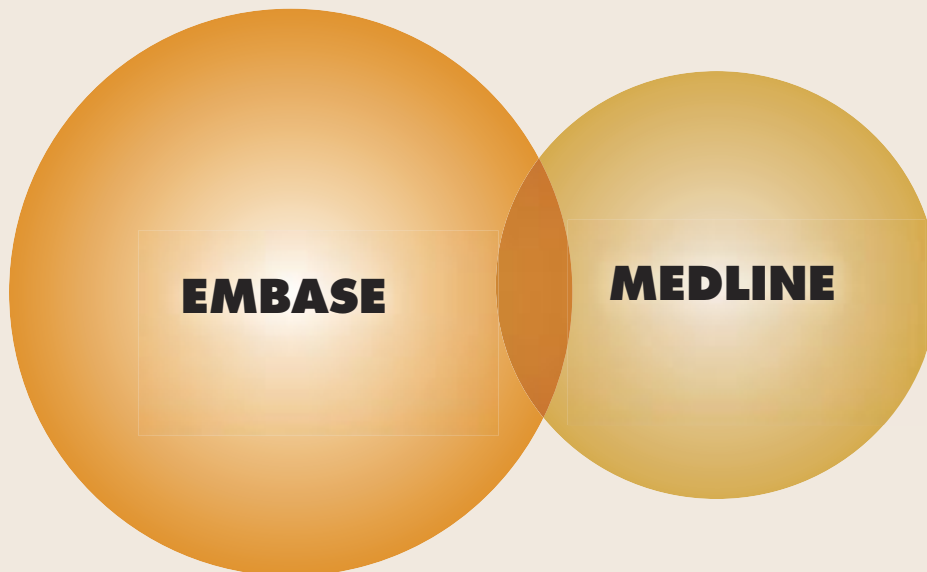
Descriptive analyses were used to evaluate frequencies, percentages, and means of study variables and were completed using SPSS, version 11. The nonparametric sign test was used to determine whether EMBASE consistently outperformed MEDLINE in the number of retrieved citations across the 15 searches using StatXact-5.

RESULTS

A total of 3445 citations were found in EMBASE, MEDLINE, or both. EMBASE contributed 2246 of these citations (65.2%); MEDLINE contributed 1199 (34.8%) (Figure 2). Surprisingly, only 177 citations from the overall total were duplicates (5.1%). EMBASE yielded twice as many unique citations as MEDLINE (2092 versus 999, respectively). In fact, EMBASE consistently referenced more unique citations than MEDLINE in 14 of the 15 searches (nonparametric sign test $P = .0005$) (Table 1 and Figure 2). The

topic libraries was then searched to determine the number of unique records retrieved from each database. The resulting data were then used to analyze search results for the selected FM topics.

Figure 2. Venn diagram showing difference between total citations retrieved from EMBASE and MEDLINE literature databases: Overlap demonstrates duplications.



magnitude of the difference between the databases was also significant using the Wilcoxon signed rank test ($P = .006$).

As an example, the “urinary tract infection” searches yielded 60 unique EMBASE citations and 25 unique MEDLINE citations, with only five duplications noted in this search. The EMBASE set of 60 unique articles on urinary tract infections included 47 articles that were available in MEDLINE but not retrieved in the search results. Therefore, 75% of the EMBASE results were retrieved due to indexing of terms and not database availability. By adding the results of EMBASE to the citations retrieved with MEDLINE alone, 3.5 times as many citations were retrieved. Review of retrieved citations for urinary tract infection showed that 63% of the EMBASE articles retrieved were clinical trials, while only 50% of the MEDLINE articles were clinical trials.

DISCUSSION

In our study, EMBASE and MEDLINE were accessed through the Ovid search engine. In our experience, clinicians with prior experience searching MEDLINE with Ovid will find searching EMBASE with Ovid effortless. As expected from earlier comparison studies, we noted that the two databases at times recognized different terms. For example, MEDLINE used the term “Physicians, Family” whereas EMBASE used “General Practitioner.” Each database has a controlled vocabulary, a unique set of official terms for a specific concept that are assigned as each article is indexed. Assigning specific and unique subject headings to each article accounts for articles being retrieved from an EMBASE search but not from a MEDLINE search. Another reason for the difference in article retrieval is the scope of each database in journal selection and subject coverage.

MEDLINE indexes more than 4600 journals in more than 30 languages with a focus on biomedical research and clinical sciences.¹⁷ EMBASE indexes more than 4000 journals from more than

70 countries on human medicine and related disciplines. EMBASE also offers comprehensive drug-related information, screening an additional 350 journals for this purpose.¹⁸ The articles on urinary tract infections from the MEDLINE search were published under only 21 individual journal titles; the EMBASE search produced 47 unique titles, offering a broader picture of the international medical literature. In addition, 26% of the EMBASE titles were not indexed by MEDLINE and thus not available for searching.

Why would someone want a comprehensive literature search? Grant writers doing background literature searches for grant applications would want to include all relevant citations. Researchers performing a comprehensive search could avoid duplicating a prior study that has been cited in one database but not both. Prevalence rates of certain diseases differ in various countries, which could affect funding opportunities (also research and subsequent publications) in these countries. If studies are non-English and are published in Europe or Asia, they might be more likely to be referenced in EMBASE. As clinicians, we are interested in having access to all studies that could aid in the care of our patients and not just studies that have been referenced in an electronic database that we are accustomed to searching.

Our study was limited to articles on therapeutics; it is unknown what similarities and differences exist between EMBASE and MEDLINE for other types of articles (eg, diagnostics or prognostics). We also limited our search to articles related to family physicians, family practice, or primary care. While it is true that the total number of citations retrieved could have been limited by the terms related to FM, we believe FM researchers and clinicians similarly limit their searches.

Searches on coronary artery disease, rhinitis, and dermatitis retrieved fewer citations than the other search topics. We speculate that fewer citations were retrieved on these topics because our search was limited to FM, family physicians, and primary care. While EMBASE consistently outperformed MEDLINE searches, searches on diabetes, coronary artery disease, and asthma retrieved about 50% unique citations from each database.

We hypothesize that diseases with high prevalence are indexed with more terms and therefore are more likely to be retrieved (Table 1). We chose to use MeSH headings in our searches because the diseases we used are common diseases with well established MeSH headings. We think this was a reasonable search strategy with common disease states; however, searches involving less common diseases or diseases or conditions with no MeSH headings should employ text words or key words to ensure more complete searches.

Family medicine researchers and physicians who exclusively use MEDLINE as their database could miss many relevant citations. EMBASE provides more citations per search and greater coverage of the total number of citations. When a comprehensive search is necessary, searching both EMBASE and MEDLINE should be more beneficial than searching MEDLINE alone. Further, medical institutions without EMBASE as a database option should consider adding EMBASE.

It took approximately 15 minutes to complete individual searches in EMBASE and 20 minutes in

MEDLINE. MEDLINE required additional time to retrieve citations older than 5 years, as the database is divided into sections by period. This division is to improve retrieval speed when searching a database of more than 12 million references dating from 1966 (as compared with 9 million in EMBASE from 1970). Total search time required for 30 searches was approximately 9 hours; however, actual search times for family physicians and other clinicians could vary depending on technical issues including computer memory, connection speed, and method of accessing EMBASE and MEDLINE.

Several limitations in our study could limit the generalizability of our results. A librarian with a Master’s degree in library science and expertise in searching by MeSH headings across various databases and specialties completed all of the searches. As a result, the search efficacy (eg, accuracy, time commitment) obtained in this study might be higher than a search completed by a researcher with less specialized search experience. Another limitation of this study was that no attempt was made to determine the quality and relevance of the citations retrieved in regard to the specified search terms.

We did, however, examine citations retrieved on “urinary tract infections” in regard to articles involving clinical trials and articles not involving clinical trials. In our review, 63% of EMBASE articles were clinical trials compared with 50% of MEDLINE articles. Clinical trials provide the highest level of evidence for clinicians to apply the conclusions of health care research to patient care. Searching EMBASE provided additional access to these critical publications. Until a qualitative analysis is conducted, an argument could be made that one database cites more irrelevant articles and thus is no more beneficial. Future studies should evaluate the quality and relevance of citations retrieved.

Table 1. Results of EMBASE and MEDLINE searches

SEARCH TERMS	EMBASE			MEDLINE		
	TOTAL CITATIONS	UNIQUE CITATIONS		TOTAL CITATIONS	UNIQUE CITATIONS	
	NS	N	%	NS	N	%
Diabetes	163	147	55	141	119	45
Depression	815	775	81	242	182	19
Otitis media	69	66	68	34	31	32
Coronary artery disease	6	6	55	5	5	45
Hypertension	452	425	73	191	158	27
Pharyngitis	27	26	63	16	15	37
Sinusitis	48	43	73	21	16	27
Rhinitis	11	9	90	3	1	10
Bronchitis	54	54	71	22	22	29
Asthma	281	247	46	330	290	54
Urinary tract infection	65	60	71	30	25	29
Dermatitis	7	7	64	4	4	36
Obesity	149	136	63	98	81	37
Osteoarthritis	43	38	60	32	25	40
Rheumatoid arthritis	56	53	68	30	25	32
TOTAL CITATIONS	2246	2092	68	1199	999	32

CONCLUSION

EMBASE consistently out-cited MEDLINE almost two to one with little duplication of

citations. This study shows that EMBASE retrieves more citations than MEDLINE. More citations do not necessarily mean higher-quality citations; therefore, future studies should look at the quality and relevance of studies retrieved from both databases. We believe that a comprehensive search that surveys the whole forest of the literature rather than a limited number of trees should employ both EMBASE and MEDLINE to ensure that a greater percentage of all available citations is recovered. ✨

Contributors

Drs Wilkins and Gillies contributed to study conception and design, and to analysis and interpretation of data; drafted and revised the article; and approved the final version to be published. Ms Davies contributed to study design and acquisition and analysis of data, drafted and revised the article, and approved the final version to be published.

Competing interests

None declared

Correspondence to: Dr Thad Wilkins, Medical College of Georgia, HB-4031, Augusta, GA 30912; telephone (706) 721-8018; fax (706) 721-7518; e-mail twilkins@mcg.edu

References

1. Suarez-Almazor ME, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Control Clin Trials* 2000;21(5):476-87.
2. Greenhalgh T. How to read a paper. The Medline database. *BMJ* 1997;315(7101):180-3.
3. Minozzi S, Pistotti V, Forni M. Searching for rehabilitation articles on MEDLINE and EMBASE. An example with cross-over design. *Arch Phys Med Rehabil* 2000;81(6):720-2.
4. McDonald S, Taylor L, Adams C. Searching the right database. A comparison of four databases for psychiatry journals. *Health Libr Rev* 1999;16(3):151-6.
5. Watson RJ, Richardson PH. Identifying randomized controlled trials of cognitive therapy for depression: comparing the efficiency of Embase, Medline and PsycINFO bibliographic databases. *Br J Med Psychol* 1999;72(Pt 4):535-42.
6. Barillot MJ, Sarrut B, Doreau CG. Evaluation of drug interaction document citation in nine on-line bibliographic databases. *Ann Pharmacother* 1997;31(1):45-9.
7. Farioli M, Jorda-Olives M, Padro JB. Bibliographic information retrieval in the field of artificial nutrition. *Clin Nutr* 1998;17(5):217-22.
8. Woods D, Trewheller K. Medline and Embase complement each other in literature searches [comment]. *BMJ* 1998;316(7138):1166.
9. Burnham J, Shearer B. Comparison of CINAHL, EMBASE, and MEDLINE databases for the nurse researcher. *Med Ref Serv Q* 1993;12(3):45-57.
10. Brazier H, Begley CM. Selecting a database for literature searches in nursing: MEDLINE or CINAHL? *J Adv Nurs* 1996;24(4):868-75.
11. Brand-de Heer DL. A comparison of the coverage of clinical medicine provided by PASCAL BIOMED and MEDLINE. *Health Inf Libr J* 2001;18(2):110-6.
12. Gallagher KE, Hulbert LA, Sullivan CP. Full-text and bibliographic database searching in the health sciences: an exploratory study comparing CCML and MEDLINE. *Med Ref Serv Q* 1990;9(4):17-25.
13. Gehanno JF, Paris C, Thirion B, Caillard JF. Assessment of bibliographic databases performance in information retrieval for occupational and environmental toxicology. *Occup Environ Med* 1998;55(8):562-6.
14. Ingram TG. A cross-sectional analysis of family medicine publications in the indexed medical literature. *Fam Med* 1992;24(4):303-6.
15. Rosser WW, Starkey C, Shaughnessy R. The pond is wider than you think! Problems encountered when searching family practice literature. *Can Fam Physician* 2000;46:103-8.
16. Schappert SM. Ambulatory care visits of physician offices, hospital outpatient departments, and emergency departments: United States, 1995. *Vital Health Stat* 13 1997 Jun;(129):1-38.
17. National Library of Medicine. *Fact sheet: MEDLINE* [home page on the Internet]. Bethesda, Md: US National Library of Medicine; updated 2001 Sep 18. Available from <http://www.nlm.nih.gov/pubs/factsheets/medline.html>. Accessed 2005 April 14.
18. Elsevier Science. *EMBASE, Excerpta Medica Database* [homepage on the Internet]. Amsterdam, The Netherlands: Elsevier Bibliographic Databases; 1999-2004. Available from http://www1.elsevier.com/homepage/sah/spd/site/locate_embase.html. Accessed 2005 April 14.

