Research methodology search filters: are they effective for locating research for evidence-based veterinary medicine in PubMed?

By Sarah Anne Murphy, M.L.S.
murphy.465@osu.edu
Head, Veterinary Medicine Library

Ohio State University
225 Veterinary Medicine Academic Building
1900 Coffey Road
Columbus, Ohio 434210

Objective: The study examined the effectiveness of research methodology search filters developed by Haynes and colleagues and utilized by the Clinical Query feature of PubMed for locating literature for evidence-based veterinary medicine (EBVM).

Methods: A manual review of articles published in 6 commonly read veterinary journals was conducted. Articles were classified by format (original study, review, general article, conference report, decision analysis, and case report) and purpose category (etiology, prognosis, diagnosis, and treatment). Search strategies listed in PubMed's Clinical Query feature were then tested and compared to the manually reviewed data to calculate sensitivity, specificity, and precision.

Results: The author manually reviewed 914 articles to identify 702 original studies. Search #1 included terms determined to have the highest sensitivity and returned acceptable sensitivities over 75% for diagnosis and treatment. Search #2 included terms identified as providing the highest specificity and returned results with specificities over 75% for etiology, prognosis, and treatment.

Discussion: The low precision for each search prompts the question: Are research methodology search filters practical for locating literature for the practice of EBVM? A study examining terms related to appropriate research methodologies for advanced clinical veterinary research is necessary to develop filters designed to locate literature for EBVM.

In the past decade, evidence-based medicine (EBM) has been promoted as a valid mechanism for locating answers to tough questions while fulfilling physicians' continuing education or lifelong learning needs [1]. Focused on the double-blinded, randomized clinical trial and systematic reviews as the best clinical research to answer everyday clinical questions, EBM has recently been advocated for the veterinary medical profession and adapted to veterinary school curricula [2–6]. Research, however, illustrates that physicians have difficulty applying the theory of evidence-based medicine to their actual medical practice [7]. Lack of time, difficulty formulating or translating questions for EBM, difficulty developing an optimal search strategy, and an inability to access the literature at time of need are frequently noted obstacles and discourage physicians from practicing EBM [8, 9]. The small number of doubled-blinded, randomized clinical trials or systematic reviews published in the veterinary medical literature creates additional obstacles for clinicians searching for evidence-based veterinary literature. Veterinary researchers often choose smaller, alternative research methodologies for financial, ethical, and other reasons [10]. Thus, the research base for evidence-based veterinary medicine is not very large. While veterinary medical publishers work to encourage researchers to improve the research base for evidence-based veterinary medicine (EBVM), it is important to examine whether veterinary clinicians will be able to locate such research once it is published.

A recent study of the CAB Abstracts database found the research methodology search filters developed by Haynes and colleagues impractical for locating literature for evidence-based veterinary medicine in two major veterinary journals [11]. While CAB Abstracts indexes the largest subset of veterinary medical liter-

ature in the world, most veterinarians, unfortunately, do not have access to CAB Abstracts after graduation from veterinary school. PubMed offers a useful alternative. The database indexes approximately seventy major veterinary titles, including the *Journal of the Veterinary Medical Association* and *Veterinary Record.* The Haynes study used MEDLINE to develop search terms and phrases that would best identify studies by research design for evidence-based medicine [12]. Since relatively few research methodologies are appropriate for advanced investigation of clinical medicine, the Haynes study tested 134,264 potential combinations of search terms derived from these methodologies in MEDLINE. Results were compared against a manual review of articles published in the 1986 and 1991 issues of ten major internal medicine and general medicine journals. The combination of terms that resulted in the highest sensitivity and specificity were identified and later adapted by PubMed's Clinical Query feature [13]. Haynes' study, however, addresses human medicine, not veterinary medicine. Could veterinarians effectively utilize PubMed's Clinical Query feature to locate literature for evidence-based veterinary practice?

This study examines the effectiveness of research methodology search filters identified by the Haynes study and used by the Clinical Query feature of the PubMed database for locating literature for evidence-based veterinary medicine. A manual review of articles published in the year 2000 issues of six veterinary journals was completed to identify articles appropriate for EBVM and to establish a standard against which to determine the sensitivity, specificity, and precision of the search filters. The intention was to determine whether veterinarians could rely on these research methodology search filters to locate literature for the practice of evidence-based veterinary medicine.

METHODS

The author's previous research examined the application of the research methodology search filters identified in the Haynes study to the CAB Abstracts database and focused on issues of the *Journal of the American Veterinary Medical Association (JAVMA)* and *Veterinary Record* from the year 2000 [14]. Research showed that these titles were commonly perused by British and American veterinarians [15, 16]. Both titles contain content pertaining to both large and small animals. The author added four titles also identified as commonly used by British and American veterinarians to the sample to increase validity of the current investigation: *American Journal of Veterinary Research, Equine Veterinary Journal, Journal of Small Animal Practice,* and *Journal of the American Animal Hospital Association* [17, 18]. As in the previous study, letters to the editor, book reviews, announcements, editorials, news, obituaries, classified advertisements, and continuing education articles were not classified or included in the study. Articles in the Veterinary Medicine Today and Short Communications sections of *Equine Veterinary Journal,*

Table 1
Format of classified articles

Category	Number of articles
Original study	702
Review	1
General article	26
Conference report	0
Decision analysis	0
Case study	185
TOTAL	914

JAVMA, and *Veterinary Record* were omitted since most articles in these sections did not meet the criteria the Haynes study identified as necessary for evidence-based practice.

Following the methodology reported in the Haynes study, a standard form was used to manually record data related to format, purpose category, and methodological rigor for each article included in the study [19]. Articles were first classified by the format categories outlined in the Haynes study: original study, review article, general article, conference report, decision analysis, and case report [20]. Since studies in veterinary medicine often contain a small number of subjects, the Haynes study's criteria for a case report was redefined as ''an article of a descriptive nature, pertaining to a particular event and involving fewer than 10 subjects'' [21]. Purpose categories for articles classified as original studies were then assigned; categories included etiology, prognosis, diagnosis, treatment, and other. Each article was classified for all applicable purposes. Methodologic rigor of the original studies was established next using the lenient criteria established in the Haynes study. This allowed articles to be classified as methodologically sound if they met a minimum of one criterion established as an appropriate research method for each purpose category. Haynes' group acknowledged that few studies actually meet the criteria for methodological soundness but reasoned ''clinicians are likely to be better informed looking at the best available literature even if it falls short of perfection'' [22].

The methodological search filters identified by Haynes and colleagues and utilized by the Clinical Query feature of PubMed were then run on PubMed and compared against the manually recorded data to determine sensitivity, specificity, and precision (Table 5). For each journal title, the author ran a search on PubMed using the filters for etiology, prognosis, diagnosis, and treatment and limiting the search to issues from the year 2000. Terms the Haynes study identified as returning the highest sensitivity were included in search #1. Search #2 consisted of terms the Haynes study identified as returning the highest specificity.

Sensitivity, specificity, and precision were detected using the formula provided in Table 1 of the Haynes study [23]. Sensitivity represents the proportion of relevant, methodologically sound articles detected by the search filter divided by the total number of relevant

Table 2
Original studies, classified for purpose

Category	Number of articles	Percentage meeting methodologic criteria
Etiology	35	68.6%
Prognosis	33	63.6%
Diagnosis	49	53.1%
Treatment	142	27.1%
Other	527	—
Total	786	—

Table 3
Sensitivity, specificity, and precision of evidence-based medicine methodological search filters for detecting veterinary literature in PubMed

Category	Sensitivity	Specificity	Precision
Etiology			
Search #1 (sensitivity search)	62.5%	63.6%	12.9%
Search #2 (specificity search)	54.2%	100%	40.6%
Prognosis			
Search #1 (sensitivity search)	54.2%	41.7%	8.8%
Search #2 (specificity search)	42.8%	83.3%	21.4%
Diagnosis			
Search #1 (sensitivity search)	76.9%	21.7%	3.6%
Search #2 (specificity search)	57.7%	73.9%	51.7%
Treatment			
Search #1 (sensitivity search)	82.0%	47.6%	12.8%
Search #2 (specificity search)	15.4%	95.8%	42.8%

citations in existence. Specificity represents the total number of relevant, but methodologically unsound articles *not* detected by the search filter divided by the total number of relevant, methodologically unsound articles in existence. This examines the search filter's ability to screen out false positives. Precision represents the total number of relevant, methodologically sound citations detected by the search filter divided by the total number of items found.

RESULTS

Results for the format of manually reviewed articles and for original studies classified for the purpose are reported in Tables 1 and 2. Of the 914 articles manually reviewed, 702 met the criteria for original study. The majority of these articles (n = 527) were categorized under the category ''other.'' This reflects the large scope of the journals examined and of the veterinary profession. Articles in the ''other'' category included reports of pharmacological and epidemiological studies, including examinations of the toxicity of drugs or the temporal and geographic distribution of disease, anatomy, physiology, food safety, animal production, animal welfare, and other areas traditionally addressed by veterinarians. The addition of the *American Journal of Veterinary Research, Equine Veterinary Journal, Journal of the American Animal Hospital Association,* and the *Journal of Small Animal Practice* to the study added only 85 articles pertaining to etiology, prognosis, diagnosis, and treatment. The percentage of articles meeting the methodologic criteria established by Haynes and colleagues was relatively low for articles related to diagnosis (53.1%) and the largest purpose category, treatment (27.1%). Fewer than 50% met the methodologic criteria for etiology, diagnosis, and treatment in the Haynes study [24].

The sensitivity, specificity, and precision of the search filters are reported in Table 3. Results are further broken down by individual journal title in Table 4. The treatment searches were of greatest interest as the largest number of articles was classified under this purpose category (n = 142), and research concerning therapy is often the most sought after information by veterinary clinicians. Search #1, consisting of terms determined to have the highest sensitivity, resulted in an acceptable sensitivity of 82.0%. This, however, was nearly 20% lower than the sensitivity reported in the Haynes study for this purpose category. While the

search filters detected 32 of the 39 possible methodologically sound treatment articles, an additional 217 hits were returned, resulting in a precision of 12.8%. These hits included articles pertaining to diagnosis, etiology, prognosis, case studies, letters to the editor, and news reports. Thus the patron might find approximately 1 out of every 8 articles relevant. A similar, yet perhaps more revealing phenomenon occurred for diagnosis search #1. This search resulted in a sensitivity of 76.9% and a precision of 3.6%. Of the 561 articles detected by the search filter, only 20 were considered relevant and methodologically sound. Thus, after reviewing approximately 28 articles, the patron may find one that is relevant and methodologically sound. Nearly 25% of the total hits for this search were case studies.

The chances of locating a relevant, methodologically sound article were somewhat more promising for search #2, which consisted of terms identified as providing the highest specificity. These searches, as designed, returned a smaller number of hits. Still, for each category, precision was either less than or nearly equal to 50%. Thus, closer to 1 out of every 2 articles could be relevant and methodologically sound. Statistically, a specificity of 100% indicates that all relevant, methodologically unsound articles were excluded from the search results. Etiology search #2, with a specificity of 100% succeeded in excluding all relevant, yet methodologically unsound articles. Still, with a sensitivity of 54.2%, nearly half of the relevant, methodologically sound articles were also excluded. Search #2 for treatment excluded more than three quarters of the methodologically sound articles. The near perfect specificity of 95.8% shows that 92 of the 96 relevant, but methodologically unsound articles were excluded from the study; however, 33 of the 39 relevant and methodologically sound articles were also excluded.

No articles were classified under the etiology purpose category for the *Equine Veterinary Journal,* the *Journal of the American Animal Hospital Association,* and the *Journal of Small Animal Practice.* It is still important to note, however, that the methodological search strate-

Table 4
Sensitivity, specificity, and precision of evidence-based medicine methodological search filters for detecting veterinary literature in PubMed by individual journal title

Category	AJVR	EVJ	JAAHA	JAVMA	JSAP	VR
Etiology						
Search #1 (sensitivity search)						
Sensitivity	0%	na	na	82.3%	na	20.0%
Specificity	100%	na	na	50.0%	na	75.0%
Precision	0%	na	na	60.9%	na	7.1%
Search #2 (specificity search)						
Sensitivity	0%	na	na	70.6%	na	20.0%
Specificity	100%	na	na	100%	na	100%
Precision	0%	na	na	57.1%	na	25.0%
Prognosis						
Search #1 (sensitivity search)						
Sensitivity	0%	0%	0%	87.5%	100%	na
Specificity	0%	66.6%	0%	0%	50.0%	na
Precision	0%	0%	0%	24.1%	5.3%	na
Search #2 (specificity search)						
Sensitivity	0%	0%	0%	50.0%	100%	na
Specificity	100%	100%	50.0%	0%	0%	na
Precision	0%	0%	0%	44.4%	11.1%	na
Diagnosis						
Search #1 (sensitivity search)						
Sensitivity	83.3%	0%	100%	37.5%	100%	100%
Specificity	50.0%	0%	0%	22.2%	0%	20.0%
Precision	5.6%	0%	3.5%	2.1%	5.3%	3.6%
Search #2 (specificity search)						
Sensitivity	66.7%	0%	100%	37.5%	25.0%	83.3%
Specificity	75.0%	100%	0%	88.9%	66.7%	40.0%
Precision	57.1%	0%	100%	75.0%	33.3%	41.7%
Treatment						
Search #1 (sensitivity search)						
Sensitivity	100%	0%	100%	81.2%	0%	69.2%
Specificity	42.8%	44.4%	55.5%	53.5%	50.0%	35.0%
Precision	13.3%	0%	50.0%	16.9%	0%	16.4%
Search #2 (specificity search)						
Sensitivity	12.5%	0%	50.0%	12.5%	0%	15.4%
Specificity	100%	88.9%	100%	95.3%	0%	95.0%
Precision	25.0%	0%	100%	50.0%	0%	66.7%

AJVR = *American Journal of Veterinary Research*
EVJ = *Equine Veterinary Journal*
JAAHA = *Journal of the American Animal Hospital Association*
JAVMA = *Journal of the Veterinary Medical Association*
JSAP = *Journal of Small Animal Practice*
VR = *Veterinary Record*

gies for search #1 returned 20 articles published in these titles. None were relevant.

DISCUSSION

Search #1, the sensitivity search, generally provided results with acceptable sensitivities over 75% for diagnosis and treatment while search #2, the specificity search, overall returned results with an acceptable specificity. The inverse relationship between sensitivity and specificity was more pronounced in this study than in the Haynes study, especially for searches in the treatment category. Regardless, results support the importance of searching a broad range of titles when seeking literature appropriate for evidence-based veterinary medicine. Three of the journal titles included in this study did not publish research pertaining to etiology in 2000. Only half of the articles evaluated for

diagnosis and a quarter of the articles evaluated for treatment met the criteria for methodological soundness. This may reflect the small number of randomized clinical trials published in the veterinary literature [25]. Still, the articles pertaining to diagnosis and treatment comprised the majority of the sample.

The author empathizes with clinician frustration when searching for literature to support evidence-based medical decisions, especially as research demonstrates they are already struggling to master basic EBM concepts [26, 27]. PubMed simplifies the definitions for sensitivity and specificity on its *Clinical Queries Using Research Methodology Filters* page stating the emphasis of the filters ''may be more sensitive (i.e., most relevant articles but probably some less relevant ones) or more specific (i.e., mostly relevant articles but probably omitting a few)'' [28]. A study of physicians in Iowa revealed that their lack of knowledge about

Table 5
PubMed's table for clinical queries using research methodology filters
http://www.ncbi.nlm.nih.gov/entrez/query/static/clinicaltable.html

Category	Search strategy
Etiology	
Search #1 (sensitivity search)	"cohort studies" [MESH] OR "risk" [MESH] OR ("odds" [WORD] AND "ratio*" [WORD]) OR ("relative" [WORD] AND "risk" [WORD]) OR "case" control*" [WORD] OR case-control studies [MESH]
Search #2 (specificity search)	"case-control studies" [MH:NOEXP] OR "cohort studies" [MH:NOEXP]
Prognosis	
Search #1 (sensitivity search)	"incidence" [MESH] OR "mortality" [MESH] OR "follow-up studies" [MESH] OR "mortality" [SH] OR prognos* [WORD] OR predict* [WORD] OR course [WORD]
Search #2 (specificity search)	prognosis [MH:NOEXP] OR "survival analysis" [MH:NOEXP]
Diagnosis	
Search #1 (sensitivity search)	"sensitivity and specificity" [MESH] OR "sensitivity" [WORD] OR "diagnosis" [SH] OR "diagnostic use" [SH] OR "specificity" [WORD]
Search #2 (specificity search)	"sensitivity and specificity" [MESH] OR ("predictive" [WORD] AND "value*" [WORD])
Treatment of prevention	
Search #1 (sensitivity search)	"randomized controlled trial" [PTYP] OR "drug therapy" [SH] OR "therapeutic use" [SH:NOEXP] OR "random*" [WORD]
Search #2 (specificity search)	(double [WORD] AND blind* [WORD]) OR placebo [WORD]

whether all relevant literature had been located was a major obstacle to implementing evidence-based practice [29]. This is a valid concern when search filters return results with a high specificity and moderate sensitivity and precision, like search #2 for etiology. The low precision of each search strategy hinders clinicians' abilities to locate relevant hits in a timely manner. Searches returning a large volume of hits may overwhelm the clinician at first glance. A large volume of irrelevant results may discourage clinicians from sifting through twenty or more articles to find a relevant hit, especially when they are used to an online culture where search engines like Google list results using a relevance ranking system.

An additional obstacle for veterinarians is whether the search results will return species-specific information. An informal review of results for search #1 found that less than 20% of hits in all-purpose categories were relevant articles pertaining to dogs or cats. Only 4 of the 116 hits, or 3.4% in etiology search #1, were relevant articles addressing disease in cats. Only 7 of the 116 hits, or 6%, pertained to dogs. Articles addressing cats, cows, dogs, a goat, horses, pigs, a rat, reindeer, sheep, a bird and a wolf were all retrieved by treatment search #1.

The low precision for each search prompts this question: are the Haynes study's research methodology search filters practical for evidence-based veterinary medicine? A study examining the reasons certain searches failed to detect relevant, methodologically sound studies and to exclude relevant, but methodologically unsound studies may be helpful. The Haynes group examined this issue in a follow-up study to their 1994 research, and ultimately concluded that in addition to improved assignment of methodologically relevant terms by indexers, authors needed to include more relevant methodological text words in the titles and abstracts of their work [30]. JAVMA's structured abstract requirement for original and retrospective studies may assist researchers in including the termi-

nology required for the search filters to detect their data. JAVMA instructs authors to "describe the experimental design in sufficient details to allow others to reproduce the results" [31]. The wide differences in search sensitivities and specificities, however, for this title raises another question: is it realistic or appropriate to expect authors to use the specific terminology required for EBVM in reports of their research? In a recent study examining search strategies for retrieval of diagnostic studies in MEDLINE, Bachmann and colleagues note that the gold-standard article required for assessment of sensitivity, specificity, and precision is "based on hand searches and complete articles, whereas . . . the filters are assessed on their ability to identify the abstracts" [32]. Specificity is criticized as a "risky" assessment of search success in this article. Since search filters with high specificity often return results with low sensitivity and precision, Bachmann and his colleagues argue clinicians receive a biased overview of the evidence, or what methodologically sound studies are available. A new calculation evaluating the number needed to read (NNR) is recommended. This is defined as 1/precision and indicates to the clinician that 1 out of every x number of articles is relevant and methodologically sound.

Overall, however, reliance on randomized controlled clinical trials (RCT) and systematic reviews may not be possible for evidence-based veterinary medicine because of the current lack of quality RCTs published in the veterinary literature [33–35]. Veterinarians may need to rely on less valid research: uncontrolled clinical trials, case series, case reports, or observational studies [36]. A study examining terminology related to research methodologies appropriate and acceptable for advanced veterinary medical investigations may be required to develop methodological search filters capable of locating literature for evidence-based veterinary medicine more consistently, with higher sensitivity, specificity, and precision.

REFERENCES

1. GUYANTT G, RENNIE D. Users' guide to the medical literature: a manual for evidence-based clinical practice. Chicago: AMA Press, 2002.

2. BONNETT B, REID-SMITH R. Critical appraisal meets clinical reality: evaluating evidence in the literature using canine hemangiosarcoma as an example. Vet Clin North Am Sm Anim Pract 1996 Jan;26(1):39–61.

3. BONNETT B. Evidence-based medicine: critical evaluation of new and existing therapies. In: Schoen A, Wynn S, eds. Complementary and alternative veterinary medicine. St. Louis: Mosby, 1998:15–20.

4. POLZIN DJ, LUND E, WALTER P, KLAUSNER J. From Journal to patient: evidence-based medicine. In: Bonagura JD, ed. Kirk's current veterinary therapy VIII: small animal practice. Philadelphia: W.B. Saunders, 2000:2–8.

5. KEENE BW. Towards evidence-based veterinary medicine. J Vet Intern Med 2000 Mar/Apr,14(2):118–9.

6. MAIR TS. Evidence-based medicine: can it be applied to equine clinical practice? Equine Vet Educ 2001 Feb;13(1):2–3.

7. YOUNG JM, GLASZIOU P, WARD JE. General practitioners' self rating of skills in evidence-based medicine: validation study. BMJ 2002 Apr 20;324(7343):950–1.

8. CUMBERS B, DONALD A. Evidence-based practice: data day. Health Serv J 1999 Apr 15;109(5650):30–1.

9. ELY JW, OSHEROFF JA, EBELL MH, CHAMBLISS ML, VINSON DC, STEVERMER JJ, PIFER EA. Obstacles to answering doctor's questions about patient care with evidence: qualitative study. BMJ 2002 Mar23;324:710.

10. KEENE, op. cit.

11. MURPHY SA. Applying methodological search filters to CAB Abstracts to identify research for evidence-based veterinary medicine. J Med Libr Assoc 2002 Oct;90(4):406–10.

12. HAYNES RB, WILCZYNSKI N, MCKIBBON KA, WALKER CJ, SINCLAIR JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Med Inform Assoc 1994 Nov–Dec;1(6):447–58.

13. PUBMED. PubMed clinical queries. [Web document]. Bethesda, Maryland: U.S. National Library of Medicine, 2002. <http://www.ncbi.nlm.nib.gov/entrez/query/static/clinical.html>.

14. MURPHY, op. cit.

15. RAW ME. Survey of libraries in veterinary practice. Vet Rec 1987;121(6):129–31.

16. PELZER NL, LEYSEN JM. Use of information resources by veterinary practitioners. Bull Med Libr Assoc 1991;79(1):10–6.

17. RAW, op. cit.

18. PELZER, op. cit.

19. HAYNES, op. cit.

20. IBID.

21. MURPHY, op. cit.

22. HAYNES, op. cit.

23. IBID, p. 448.

24. IBID.

25. POLZIN, op. cit.

26. ELY, op. cit.

27. YOUNG, op. cit.

28. PUBMED, op. cit.

29. ELY, op cit.

30. WILCZYNSKI, NL, WALKER CJ, MCKIBBON KA, HAYNES RB. Reasons for the loss of sensitivity and specificity of methodologic MeSH terms and textwords in MEDLINE. Proc Annu Symp Comput Appl Med Care 1995; 436–40.

31. JOURNAL OF THE AMERICAN VETERINARY MEDICAL ASSOCIATION. Information for authors. [Web document]. Chicago: American Veterinary Medical Association, 2002 [rev 15 December 2002; cited 15 December 2002]. <http://www.avma.org/publications/javma/javma_ifa.asp>

32. BACHMANN LM, CORAY R, ESTERMANN P, TIER RIET G. Identifying diagnostic studies in MEDLINE: reducing the number to read. J Am Med Inform Assoc 2002 Nov/Dec; 9(6):657.

33. POLZIN, op cit.

34. LUND E, JAMES KM, NEATON JD. Veterinary randomized clinical trial reporting: a review of the small animal literature. J Vet Intern Med 1998 Mar/Apr;12(2):57–60.

35. KEENE, op. cit.

36. POLZIN, op. cit.