
Evidence-based retrieval in evidence-based medicine*

By Timothy B. Patrick, PhD
PatrickT@health.missouri.edu
Assistant Professor, Health Management and Informatics

George Demiris, PhD
DemirisG@health.missouri.edu
Assistant Professor, Health Management and Informatics

Lillian C. Folk, BSBA
FolkL@missouri.edu
Pre-doctoral Fellow, Health Management and Informatics

David E. Moxley, MLS
MoxleyD@health.missouri.edu
Instructor, Health Management and Informatics

Joyce A. Mitchell, PhD
MitchellJo@health.missouri.edu
Professor, Health Management and Informatics

Donghua Tao, BS
dt6b3@missouri.edu
Doctoral Student

Department of Health Management and Informatics
School of Medicine
University of Missouri
324 Clark Hall
Columbia, Missouri 65211

Objective: Clinical decisions based on a meta-analysis that is based on an ineffective retrieval strategy may have serious negative consequences for patients. The study objective was to investigate the extent to which meta-analyses report proof of their retrieval strategies' effectiveness.

Methods: The authors examined a random sample (n = 100) of articles in the 1996 to 2002 full-text subset of Ovid MEDLINE indexed as "meta-analysis." We classified the articles in three ways: the article (A) reported both a retrieval strategy in sufficient detail (such that it could be repeated) and with evidence of the strategy's effectiveness, (B) reported a retrieval strategy in sufficient detail but not with evidence of the strategy's effectiveness, or (C) neither reported a strategy in detail nor evidence of the strategy's effectiveness. Articles classified as (A) were further classified according to the level of evidence reported.

Results: Of the eighty-nine articles in our final analysis, six (6.7%) were classified as category (A), fifty-seven (64%) as (B), and twenty-six (29%) as (C). Articles in category (A) reported a previously validated search, a published strategy, or strategy based on expert opinion.

Conclusion: Peer-review standards must be developed that require authors of meta-analyses to report evidence for the effectiveness of their retrieval strategies.

INTRODUCTION

Evidence-based medicine implies the need to base clinical decisions on the results of prior scientific study. It also implies the need for clear indication of the effectiveness of the strategies and methods used to retrieve that evidence. Such a need is increasingly recognized [1–3]. Typically, according to standards for evidence-based medicine (such as the “Levels of Evidence” developed by the Centre for Evidence-Based Medicine [4]), stronger evidence is based on rigorous scientific study, while weaker evidence is based on expert opinion. By analogy, strong evidence of the effectiveness of a bibliographic retrieval strategy might be based on prior testing of the strategy, while weaker evidence might be an expert searcher’s opinion of its effectiveness. The potential negative consequences of ineffective retrieval strategies in health care are well demonstrated by the death of a healthy research volunteer at Johns Hopkins in 2001 [5]. The US Department of Health and Human Services Office for Human Research Protection (OHRP) Compliance Determination Letter of July 19, 2001, regarding the incident specifically cited the inadequacy of the bibliographic retrieval strategy used, saying

In particular, OHRP notes the following: (a) Prior to the research being approved by the IRB [internal review board], the investigators and the JHBMC [Johns Hopkins Bayview Medical Center] IRB failed to obtain published literature about the known association between hexamethonium and lung toxicity. Such data was readily available via routine MEDLINE and Internet database searches, as well as recent textbooks on pathology of the lung. [6]

Meta-analyses of prior studies reported in the medical literature are an increasingly important method of generating evidence to support clinical decision making. In addition to emphasizing the statistical methods used to compare studies in a meta-analysis, the focus needs to be on the bibliographic strategy used to retrieve reports of these studies. Quality control of a meta-analysis needs to focus on the statistical methods used to compare reports as well as the bibliographic strategy used to retrieve these reports. Clinical decisions based on a meta-analysis that is itself based on an ineffective retrieval strategy can have serious negative consequences for patients.

The authors sought to investigate the extent to which such meta-analyses themselves report proof of the effectiveness of their retrieval strategies. We examined a sample of peer-reviewed journal articles indexed in MEDLINE as publication type “meta-analysis” to determine whether they reported evidence of the effectiveness of their retrieval strategies.

* This research was supported in part by National Library of Medicine Integrated Advanced Information Management Systems grant 5-G08-LM05415-06 and by National Library of Medicine Biomedical and Health Informatics Research Training grant 2-T15-LM07089-11.

METHODS

According to the Medical Subject Headings scope notes provided by the National Library of Medicine, articles reporting the details of a meta-analysis are indexed with the publication type “meta-analysis” [7]. (Articles indexed with the main subject heading of “meta-analysis” are articles about the topic *meta-analysis*, rather than reports of specific meta-analyses.) Thus, to construct a sample of peer-reviewed articles reporting the methods and results of specific meta-analyses, on November 21, 2002, we searched the 1996 to 2002 subset of the MEDLINE database through OVID for articles with publication type “meta-analysis.”

We used the Ovid MEDLINE search query “meta-analysis.pt,” producing 4,740 results. We limited the results to the years 1997 to 2002, producing 4,258 results. Next, we limited the results to English, producing 4,027 results. Of these results, 664 were available in Ovid full text, covering 52 peer-reviewed general and specialty journals and including 28 journals out of the 120 core clinical journals of the Abridged Index Medicus (AIM) [8]. We limited the results to Ovid full text, and we took a random sample ($n = 100$) of the 664 Ovid full-text results. Four authors (Patrick, Tao, Folk, and Moxley) independently classified each article in the sample in one of three ways.

Category (A) indicates that the article both reported a retrieval strategy in sufficient detail such that it could be repeated *and* reported evidence of the effectiveness of that strategy. We considered that a strategy was reported in sufficient detail to be repeated if the report at least included the databases searched, the years searched, and all of the search terms used. An example of a retrieval strategy reported in sufficient detail is the following statement:

We identified potential English-language sources from the MEDLINE database for the years 1966 through 1997 using the search terms *physical examination, palpation, breast, breast diseases, diagnosis, diagnostic tests, and sensitivity and specificity*. [9]

An example of a report of evidence of the effectiveness of a retrieval strategy, in this case that of a previously tested search strategy, is the following statement:

We searched Medline for 1966 to December 1995 using a combination of the March 1996 update of the optimally sensitive search strategy for trials from the Cochrane Collaboration. [10]

Category (B) indicates that the article reported a retrieval strategy in sufficient detail such that it could be repeated but did not report evidence of the effectiveness of that strategy. Finally, articles classified as category (C) neither reported a retrieval strategy in sufficient detail such that it could be repeated nor reported evidence of the effectiveness of a reported strategy. For example,

We performed searches of the MEDLINE database and re-

viewed the bibliographies of review articles, to identify studies describing treatment for acute myocardial infarction. We used population-based studies that reported at least 10 years of data whenever possible to determine changes in intervention rates for different therapies. Published meta-analyses of randomized controlled trials were used to estimate the average benefit from changes in these various interventions. [11]

In this case, the authors did not report the years searched, all of the search terms used, or any evidence of the effectiveness of the search strategy.

After completing the independent classifications, we compared results. When researchers classified articles differently, the criteria and article were discussed to reach consensus. Such cases were few and typically concerned whether a strategy was reported in sufficient detail such that it could be repeated. Finally, articles assigned to category (A) were further classified according to level of evidence cited. We used a one-dimensional classification consisting of three levels of evidence: (1) *previously validated search strategy* (strongest), (2) *previously published search strategy*, and (3) *expert opinion* (weakest).

RESULTS

Eleven of the 100 articles were excluded from further analysis. One article was excluded from further analysis, because it was a brief consumer-oriented summary. One article was excluded, because it was a letter commenting on another study. Nine articles were excluded from further analysis, because they did not involve a meta-analysis based on retrieval of studies from the medical literature but instead reported a meta-analysis based on reports contained in non-bibliographic databases. In one case, an article in our original sample was a "Patient-Oriented Evidence that Matters" (POEM) article [12] summarizing another article. Miser 1998 [13] was a POEM summary of a specific reference [14], and, while Miser 1998 was included in our original sample, that reference was not. We excluded Miser 1998 from further analysis and included the specific reference. The remaining 89 articles included in the analysis were contained in 25 peer-reviewed general and specialty journals covering at least 15 health care specialties (Table 1).

Of the eighty-nine randomly selected articles that met the inclusion criteria, only six (6.7%), were assigned to category (A): those that both reported a retrieval strategy in sufficient detail such that it could be repeated and reported evidence of the effectiveness of that strategy. Fifty-seven articles (64%), were assigned to category (B): those that reported a retrieval strategy in sufficient detail such that it could be repeated but did not report evidence of the effectiveness of that strategy. Finally, twenty-six articles (29%), were assigned to category (C): those that neither reported a retrieval strategy in sufficient detail such that it could be repeated nor reported evidence of the effectiveness of a reported strategy.

Of the six articles assigned to category (A), four re-

Table 1

Twenty-five journals included in random full text sample (n = 100)

Journal
<i>American Journal of Medicine</i>
<i>American Journal of Obstetrics & Gynecology</i>
<i>American Journal of Psychiatry</i>
<i>American Journal of Surgery</i>
<i>Annals of Internal Medicine</i>
<i>Annals of Surgery</i>
<i>Archives of General Psychiatry</i>
<i>Archives of Internal Medicine</i>
<i>Archives of Neurology</i>
<i>Archives of Surgery</i>
<i>BMJ</i>
<i>British Journal of Psychiatry</i>
<i>Chest</i>
<i>Fertility & Sterility</i>
<i>JAMA</i>
<i>Journal of Bone & Joint Surgery</i>
<i>Journal of Clinical Endocrinology & Metabolism</i>
<i>Journal of Clinical Oncology</i>
<i>Journal of Clinical Psychopharmacology</i>
<i>Journal of Pediatrics</i>
<i>Journal of the American Academy of Child & Adolescent Psychiatry</i>
<i>Nursing Research</i>
<i>Pediatrics</i>
<i>Psychological Medicine</i>
<i>QJM</i>

ported evidence that we classified as *previously validated search strategy* (level 1), one reported evidence that we classified as *previously published search strategy* (level 2), and one reported evidence that we classified as *expert opinion* (level 3).

DISCUSSION

Of the twenty-five journals containing articles included in our final sample, eighteen (72%) make an explicit reference in their statement of instructions for authors to the "International Committee of Medical Journal Editors, Uniform Requirements for Manuscripts submitted to Biomedical Journals" (ICMJE) [15]. The ICMJE directs authors to

[i]dentify the methods, apparatus (give the manufacturer's name and address in parentheses), and procedures in sufficient detail to allow other workers to reproduce the results,

implying a requirement that a meta-analysis report its retrieval strategy in sufficient detail such that it can be repeated. The ICMJE does not, however, explicitly require authors to report evidence of the effectiveness of their retrieval strategies. That eighteen journals containing articles in our final sample themselves make an explicit reference to the ICMJE statement appears to be consistent with our finding that 64% of the articles fell into category (B), those that report a retrieval strategy in sufficient detail such that it could be repeated but do not report evidence of the effectiveness of that strategy. However, ten journals that include an explicit reference to the ICMJE statement also include articles classified as our category (C), those that neither report a retrieval strategy in sufficient detail such that it could be repeated nor report evidence of the effectiveness of a reported strategy.

Two statements of requirements for authors of meta-analyses, the QUOROM and MOOSE statements, are mentioned in the author instructions of one journal, and the QUOROM statement was mentioned by one additional journal. Both of these journals include articles that we classified as category (C). The QUOROM statement for meta-analyses requires an author to include a description of the search strategy used,

in detail (e.g., databases, registers, personal files, expert informants, agencies, hand-searching), and any restrictions (years considered, publication status, language of publication). [16]

The MOOSE statement requires the author to similarly describe details of the bibliographic retrieval methods used including

- Qualifications of searchers (e.g., librarians and investigators)
- Search strategy, including time period included in the synthesis and keywords
- Databases and registries searched
- Search software used, name and version, including special features used (e.g., explosion) [17]

Because the MOOSE statement requires mention of the qualifications of searchers, it requires an indication of at least indirect evidence for the effectiveness of the retrieval strategy used.

Reports of bibliographic-based meta-analyses that do not report the retrieval strategy in sufficient detail to be repeated run counter to the basic tenets of meta-analysis research and evidence-based medicine. Furthermore, any reporting of a retrieval strategy that does not also report evidence of the effectiveness of that strategy is similarly at odds with the basic tenets of evidence-based medicine. More effort should be directed toward establishing clear standards of evidence of the effectiveness of search retrieval strategies in evidence-based medicine. In addition, editorial and peer-review standards must be developed and applied that require authors of meta-analyses to report evidence for the effectiveness of the retrieval strategies they employ.

ACKNOWLEDGMENT

We thank Rebecca S. Graves, University of Missouri-Columbia J. Otto Lottes Health Sciences Library, for her assistance.

REFERENCES

1. HUNT DL, JAESCHKE R, MCKIBBON KA. Users' guides to the medical literature: XXI. using electronic health information

resources in evidence-based practice. Evidence-Based Medicine Working Group. JAMA 2000 Apr 12;283(14):1875-9.

2. ELDRIDGE JD. Evidence-based librarianship: what might we expect in the years ahead? Health Inf Libr J 2002 Jun; 19(2):71-7.

3. ROBINSON KA, DICKERSIN K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. Int J Epidemiol 2002 Feb;31(1): 150-3.

4. OXFORD CENTRE FOR EVIDENCE-BASED MEDICINE. Levels of evidence and grades of recommendation. [Web document]. Oxford, UK. [cited 16 Dec 2003]. <http://www.cebm.net/levels_of_evidence.asp>.

5. RAMSAY S. Johns Hopkins takes responsibility for volunteer's death. Lancet 2001 Jul 21;358(9277):213.

6. MCNEILLY PJ, CAROME M. Human subject protections under Multiple Project Assurance (MPA) M-1011. Letter to: Edward D. Miller, Chi Van Dang, Gregory F. Schaffer. Rockville, MD: 19 Jul 2001, 11 pages. <<http://ohrp.osophs.dhhs.gov/detrm.lettrs/jul01a.pdf>>.

7. 2003 MeSH, annotated alphabetic list. Bethesda, MD: National Library of Medicine, 2002.

8. NATIONAL LIBRARY OF MEDICINE. Abridged index medicus (AIM) journal titles. [Web document]. Bethesda, MD: The Library. [rev. 22 Jun 2001; cited 4 Feb 2003]. <<http://www.nlm.nih.gov/bsd/aim.html>>.

9. BARTON MB, HARRIS R, FLETCHER SW. The rational clinical examination. does this patient have breast cancer? the screening clinical breast examination: should it be done? how? JAMA 1999 Oct 6;282(13):1270-80.

10. ALDERSON P, ROBERTS I. Corticosteroids in acute traumatic brain injury: systematic review of randomised controlled trials. BMJ 1997 Jun 28;314(7098):1855-9.

11. HEIDENREICH PA, MCCLELLAN M. Trends in treatment and outcomes for acute myocardial infarction: 1975-1995. Am J Med 2001 Feb 15;110(3):165-74.

12. GEYMAN JP. POEMs as a paradigm shift in teaching, learning, and clinical practice. Patient-Oriented Evidence that Matters. J Fam Prac 1999 May;48(5):343-4.

13. MISER WF. Treating depression in older ambulatory patients. J Fam Prac 1998;47(1):16-7.

14. MCCUSKER J, COLE M, KELLER E, BELLAVANCE F, BERARD A. Effectiveness of treatments of depression in older ambulatory patients. Arch Intern Med 1998 Apr 13;158(7):705-12.

15. INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. Uniform requirements for manuscripts submitted to biomedical journals. JAMA 1997 Mar 19;277(11):927-34.

16. MOHER D, COOK DJ, EASTWOOD S, OLKIN I, RENNIE D, STROUP DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. quality of reporting of meta-analyses. Lancet 1999 Nov 27; 354(9193):1896-900.

17. STROUP DF, BERLIN JA, MORTON SC, OLKIN I, WILLIAMSON GD, RENNIE D, MOHER D, BECKER BJ, SIPE TA, THACKER SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-Analysis of Observational Studies in Epidemiology (MOOSE) Group. JAMA 2000 Apr 19;283(15):2008-12.

Received May 2003; accepted October 2003