Systematic reviews in emergency medicine: Part I. Background and general principles for locating and critically appraising reviews

Peter J. Zed, BSc, BSc (Pharm), PharmD;[*†‡] Brian H. Rowe, MD, MSc;[§]
Peter S. Loewen, BSc (Pharm), PharmD;[*†] Riyad B. Abu-Laban, MD, MHSc, FRCPC[‡¶]

ABSTRACT

Reviews of the medical literature have always been an important resource for physicians. Increasingly, qualitative and quantitative "systematic reviews" have replaced the traditional "narrative review" as a means of capturing and summarizing current evidence on a topic or, when possible, answering a specific clinical question. This paper is part one of a two-part series designed to provide emergency physicians with the background necessary to locate, critically evaluate and interpret systematic reviews. The paper provides a brief background on systematic reviews and general principles on locating and critically appraising them. To facilitate readability, examples from the emergency medicine literature have been included for illustrative purposes and technical details have been kept to a minimum. The references, however, are comprehensive and provide a resource for readers seeking further information.

Key words: systematic reviews; emergency medicine; evidence-based medicine

RÉSUMÉ

Les revues de la littérature médicale ont toujours été une ressource importante pour les médecins. De plus en plus, les «revues systématiques» qualitatives et quantitatives ont remplacé les «revues narratives» comme moyen de saisir et résumer les données courantes sur un sujet précis ou, autant que possible, de répondre à une question clinique spécifique. Le présent article constitue la première de deux parties conçues pour donner aux médecins d'urgence le contexte nécessaire pour trouver, faire une évaluation critique et interpréter les revues systématiques. La première partie offre une brève mise en contexte sur les revues systématiques et les principes généraux pour trouver et évaluer ces revues. Des exemples provenant de la littérature de médecine d'urgence sont inclus pour illustrer le principe et les détails techniques sont maintenus à un strict minimum. Par contre, les références sont exhaustives et offrent des ressources pour les lecteurs à la recherche de plus d'information.

*Clinical Service Unit Pharmaceutical Sciences, Vancouver General Hospital, Vancouver, BC
†Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC
‡Division of Emergency Medicine, Department of Surgery, Faculty of Medicine, University of British Columbia, Vancouver, BC
§Division of Emergency Medicine, Department of Public Health Sciences, University of Alberta, and Capital Health Authority, Edmonton, Alta.
¶Department of Emergency Medicine and Center for Clinical Epidemiology and Evaluation, Vancouver General Hospital, Vancouver, BC

Received: Feb. 1, 2003; final submission: July 9, 2003; accepted: July 18, 2003

This article has been peer reviewed.

Can J Emerg Med 2003;5(5):331-5

Introduction

Reviews of the medical literature have always been an important resource for physicians. Increasingly, qualitative and quantitative "systematic reviews" (SRs) have replaced the traditional "narrative review" as a means of capturing and summarizing current evidence on a topic or, when possible, answering a specific clinical question. The purpose of this 2-part series is to provide emergency physicians with the background necessary to locate, critically evaluate and interpret SRs. This paper will provide background information on SRs and general principles for locating and critically appraising them. Part II will appear in the next issue of the *Canadian Journal of Emergency Medicine* and will broaden the discussion of critical appraisal principles by focusing on quality assessment, data synthesis and interpretation of results. To facilitate readability, examples from the emergency medicine literature have been included for illustrative purposes and technical details have been kept to a minimum. The references, however, are comprehensive and provide a resource for readers seeking further information.

Background

A review paper is, by definition, both retrospective and observational. As a result, its quality is particularly dependent on the efforts taken to minimize error and bias.[1] The traditional "review" or "overview" in the medical literature is termed a "narrative review." Such papers often lack a clearly described and systematic approach to the identification and interpretation of evidence; and while the conclusions from narrative reviews may be confidently and lucidly stated, they do not arise in a scientifically reproducible manner.[2] For this reason, narrative reviews can reach biased or erroneous conclusions. Despite the increasing awareness of this fact, many papers still fail to meet established criteria for a good review.[3–5] A recent paper on the utility of amylase and lipase testing in the emergency department (ED) is an example of a narrative review.[6] While not necessarily incorrect, the conclusions and recommendations of this otherwise comprehensive paper would have been strengthened through the use of systematic methods.

SRs differ in many ways from narrative reviews, as outlined in Table I. An SR is, in essence, a study that enrolls other studies as its subjects.[1] Analogous to research on human subjects, many factors can undermine the reliability and validity of an SR; thus, the methods must be carefully considered and clearly described. SRs can be either "qualitative" or "quantitative." Both types of SRs employ similar techniques to capture the available evidence, a task that has become increasingly complex and specialized.[7] In a qualitative SR, the results from primary studies are summarized but not statistically combined. An example of a qualitative SR in the emergency medicine literature discussed methemoglobinemia from topical anesthetic spray.[8] This paper reported a single case, then reviewed the relevant literature using qualitative methods. Medical journals increasingly request the inclusion of a brief qualitative SR like this with case reports.

In a quantitative SR, more commonly known as a "meta-analysis," specialized statistical techniques are used to evaluate and, when possible, combine the results of 2 or more studies. The primary purposes of meta-analysis are to (i) increase statistical power; (ii) resolve uncertainty about the presence or absence of an effect; (iii) improve effect size estimates; and (iv) answer questions not posed by primary studies.[9] Since meta-analytic statistical techniques can potentially obscure poor review methodology, readers must first be satisfied with the methods before trusting the pooled results. Moreover, meta-analysis cannot overcome biases or "fatal flaws" in the primary studies included in the review.

A recent quantitative SR assessed the effectiveness of corticosteroids in the treatment of acute asthma after ED discharge. This meta-analysis, which combined the results of 5 trials and documented fewer 7-day relapses with the use of systemic corticosteroids (odds ratio [OR] = 0.35; 95% confidence interval [CI] = 0.17–0.73), illustrates how combining results from multiple smaller trials can clarify a specific clinical question that was not adequately addressed in the primary studies.[10]

Advances in evidence-based medicine have clarified how difficult it is to convincingly establish cause-effect relation-

Table 1. Comparison of narrative and systematic reviews

Feature	Narrative review	Systematic review
Question	Often broad in scope	Usually focused
Sources and search	Usually unspecified and potentially biased	Comprehensive sources and explicit search strategy
Selection	Usually unspecified and potentially biased	Uniformly applied criterion-based selection
Appraisal of evidence	Variable	Rigorous
Synthesis	Usually qualitative	Quantitative or qualitative
Inferences	Sometimes evidence-based	Usually evidence-based

https://doi.org/10.1017/S1481803500015840 Published online by Cambridge University Press

ships in a single trial.[11] Consequently, a well-performed meta-analysis of methodologically sound primary studies represents the highest level of medical evidence (Grade 1a) (www.cebm.net/levels_of_evidence.asp#levels). This is particularly important in emergency medicine, where mega-trials are uncommon. In such a situation, meta-analyses can estimate true effect sizes by combining data from multiple smaller trials, without the expense of a large trial. SRs can also help to clarify when there is insufficient evidence on a given topic.[12] In order to establish that a need exists for proposed research, granting agencies increasingly expect SRs to accompany applications for funding of randomized controlled trials (www.cihr-irsc.gc.ca/services/funding/apply/instr/rct_e.shtml).

Locating systematic reviews in the literature

Research suggests that locating SRs can be a difficult task; however, identification of a well-conducted SR is often a useful first step in attempting to answer a clinical question.[13] One of the difficulties is that the nomenclature varies for SRs and includes such terms as "overviews," "state-of-the-art reviews," "EBM reviews" and other hybrid terms. In order to identify true SRs in the medical literature, multiple search strategies are often required. Computerized searches using MEDLINE or EMBASE, although often assumed to be comprehensive, fail to identify many SRs.[13] One readily available solution to this problem is to first search the Cochrane Library (www.update-software.com/Cochrane/default.htm), a resource named after Archie Cochrane, a British physician and epidemiologist who drew attention to the fact that physicians did not have ready access to reliable reviews of the available evidence. Through use of the Cochrane Database of Systematic Reviews (CDSR) (www.cochrane.org/), the most valid and current SRs in diverse therapeutic areas can be located, many of which are relevant to emergency medicine.[14] In addition, the Cochrane Library contains the Database of Abstracts of Reviews of Effects (DARE), (http://nhscrd.york.ac.uk/darehp.htm), a compilation of abstracts of SRs published in the primary literature. Alternatively, search strategies can be used to identify SRs in computerized databases or "gray" literature — literature not readily accessible (e.g., unpublished studies, thesis work and technical reports).

Critically appraising systematic reviews

After locating an SR, readers should critically appraise the review much the way a clinician would appraise a published clinical trial. Checklists like those in Table 2 help focus readers on the key questions they should ask before accepting the conclusions of an SR.[2] As with clinical trials, there are 3 fundamental questions: "Are the results of the review valid?" "What are the results?" and "Will the results help me care for my patients?" Only by answering these can the validity, appropriateness and generalizability of an SR be discerned.

Clinical question

A good SR is based on a well-formulated, answerable clinical question. This question guides the SR by defining the population, intervention, controls and outcomes to be considered, and the design of the studies to be included. The question should be clearly described so that readers can determine the appropriateness of the strategies used to identify relevant primary studies, to select studies for inclusion, and to extract data from each study.[15] A clear question also helps readers rapidly determine whether the SR is relevant to their own clinical practice.

A good clinical question describes 4 elements: (i) the patients; (ii) the intervention; (iii) the control treatment or exposure; and (iv) the primary (and sometimes secondary) outcomes of interest.[15] This is often referred to as a "PICO" formula (Patient/Population, Intervention, Comparison, Outcomes).[16] The clinical question addressed in the SR should ideally be relevant to patients in the ED. An example of a clinical question in emergency medicine practice described using the PICO formula would be: "Does adding inhaled corticosteroids to oral corticosteroids reduce 7-day relapse rates in adult patients with severe acute exacerbations of asthma who are discharged from the ED?"

Table 2. Critically appraising systematic reviews

Are the results of the review valid?
Did the review address a focused clinical question?
Were the criteria used to select studies for inclusion appropriate?
Is it unlikely that important, relevant studies were missed?
Was the validity of the included studies appraised?
Were the assessments of the studies reproducible?
Were the results similar from study to study?
What are the results?
What are the overall results of the review?
How precise are the results?
Will the results help me in caring for my patients?
Can the results be applied to my patients?
Were all clinically important outcomes considered?
Are the benefits worth the harms and risks?

https://doi.org/10.1017/S1481803500015840 Published online by Cambridge University Press

Searching and selecting studies for systematic review

SRs should include the best available evidence. The authors must explicitly state the criteria used to search for and select studies for inclusion, and the reader must be confident that these methods are not likely to result in the omission of relevant studies. Reviews of treatment and prevention should seek the highest level of evidence, namely randomized, controlled trials. Reviews of diagnostic tests should include prospective studies using independent assessment comparing one or more tests with an appropriate "criterion standard" (commonly called a "gold standard"). Reviews of prognosis should include prospective cohort studies in which a representative sample of patients was enrolled at a similar point in their disease course. Finally, reviews of risk factors should include appropriate case and control populations.[17] Authors should search for the best available evidence relevant to the question using the highest quality studies. If high quality studies are not used as the basis for the search then the authors must justify this decision. In summary, the search strategy must be explicitly described and comprehensive to avoid the two main threats to SR validity: *publication bias* and *selection bias*. These are discussed under "search" and "selection" below.

Search

Before embarking on a search of the literature to answer a well-defined question, authors of SRs should describe their strategy for addressing publication bias. In addition to the PICO formula and study design, issues regarding language of publication, year of publication, publication status (unpublished v. published) and preliminary results (abstracts) must be addressed.[17] Many SRs restrict their searches to English language articles, but this can bias the results because relevant studies may be published in other languages.[17] Of interest, recent research suggests that studies published in non-English language journals tend to exhibit larger treatment effects.[18] Some researchers feel unpublished research should be excluded from SRs because it has not undergone peer review and therefore may be unreliable; however, it has been demonstrated that many high quality articles are not published for a myriad of unrelated reasons. Moreover, study results can influence publication, and many negative trials are never published — a concept referred to as publication bias.[19] Since studies with negative or neutral results are less likely to be published than studies with larger or positive results, SRs that exclude unpublished work may overestimate the relationship between an intervention and outcome.[12] Authors who wish to conduct high quality SRs can obtain unpublished research and evaluate its relevance and validity, although this is often difficult and adds cost to the review. A "best case, worst case" analysis (or sensitivity analysis) increases the strength of a review's conclusions.

The most common method of locating articles for inclusion in SRs is through electronic database searching. MEDLINE, produced by the National Library of Medicine, is the best-known and most comprehensive electronic database available. MEDLINE indexes nearly 4000 biomedical journals and includes citations for over 11 million journal articles, editorials and letters, dating from 1966.[20] Readers are referred to a series on MEDLINE for further detail about this electronic database.[20,21]

Although MEDLINE is the most common electronic database used in conducting literature searches, it will not identify all published articles. Authors who rely solely on MEDLINE will detect fewer than 50% of randomized controlled trials conducted in an area of interest;[22] therefore, other electronic databases including EMBASE, CINAHL, ERIC, LILACS, and other specialized registries, should be searched. Authors should hand-search relevant journals and conference proceedings and check the references listed in studies and review articles identified in the electronic search.[23] "Gray literature" is often difficult to find and evaluate, but efforts should be made to complete this important aspect of the process. Approaches such as contacting industry representatives, authors of previous research, experts in the area of interest and colleagues can enhance the identification of gray literature.

High quality SRs provide a clear description of the search strategies and terms used to identify possibly relevant studies for inclusion in the review. Moreover, they have a comprehensive approach using multiple search strategies to avoid publication bias. To be confident in the results of SRs, reviewers should report search methods that are impartial, explicit and reproducible.

Selection

Selecting primary studies for inclusion in the SR requires methods that avoid bias in the decision-making process. High quality SRs establish clearly defined inclusion and exclusion criteria *a priori* (i.e., before any studies are reviewed). A comprehensive search of the literature typically has high sensitivity and low specificity for the question of interest and will usually yield a large number of articles. To weed out irrelevant articles, most SRs employ a process of relevance screening that is conducted independently by two or more reviewers.[19] Such a process reduces the biases inherent in study selection that can undermine SR validity.

https://doi.org/10.1017/S1481803500015840 Published online by Cambridge University Press

In addition, it provides an opportunity to evaluate the inter-rater agreement (reliability) for each step. Consequently, SRs should clearly describe these selection processes, and SRs conducted by a single reviewer should be interpreted with caution.

Summary

SRs are an important source of high-level medical evidence and very relevant to the practice of emergency medicine. The first aspects of critical appraisal include a well formulated PICO question, with well-described and comprehensive search strategy and an unbiased process of study selection.

Competing interests: None declared.

Acknowledgments: Dr. Brian Rowe is supported by the Canadian Institute of Health Research (CIHR) as a Canada Research Chair in Emergency Airway Diseases. Dr. Riyad B. Abu-Laban is supported by a Clinical Scholar award from the Michael Smith Foundation for Health Research.

References

1. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med 1997; 126:376-80.
2. Oxman AD, Cook DJ, Guyatt GH. Users' guide to the medical literature. VI. How to use an overview. JAMA 1994;272:1367-71.
3. McAlister FA, Clark HD, van Walraven C, Straus SE, Lawson FME, Moher D, Mulrow CD. The medical review article revisited: Has the science improved? Ann Intern Med 1999;131:947-51.
4. Olsen O, Middleton P, Ezzo J, Gotzsche PC, Hadhazy V, Herxheimer A, et al. Quality of Cochrane reviews: assessment of sample from 1998. BMJ 2001;323:829-32.
5. Kelly KD, Travers A, Dorgan M, Slater L, Rowe BH. Evaluating the quality of systematic reviews in the emergency medicine literature. Ann Emerg Med 2001;38:518-26.
6. Vissers RJ, Abu-Laban RB, McHugh DF. Amylase and lipase in the emergency department evaluation of acute pancreatitis. J Emerg Med 1999;17:1027-37.
7. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 1992;268:240-8.
8. Abu-Laban RB, Zed PJ, Purssell RA, Evans KG. Severe methemoglobinemia from topical anesthetic spray: case report, discussion and qualitative systematic review. Can J Emerg Med 2001; 3(1):51-6.
9. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. N Engl J Med 1987;316:450-5.
10. Rowe BH, Spooner CH, Ducharme FM, Bretzlaff JA, Bota GW. Corticosteroids for preventing relapse following acute exacerbations of asthma [Cochrane review]. In: The Cochrane Library; Issue 1, 2003. Oxford: Update Software.
11. Peto R. Why do we need systematic overviews of randomized trials? Stat Med 1987;6:233-44.
12. Egger M, Smith GD, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. 2nd ed. BMJ Publishing Group: London (UK); 2001.
13. Hunt DL, McKibbon KA. Locating and appraising systematic reviews. Ann Intern Med 1997;126:532-8.
14. Emond SD, Wyer PC, Brown MD, Cordell WH, Spooner CH, Rowe BH. How relevant are the systematic reviews in the cochrane library to emergency medical practice? Ann Emerg Med 2002;39;153-8.
15. Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews. Ann Intern Med 1997;127: 380-7.
16. Corrall CJ, Wyer PC, Zick LS, Bockrath CR. Evidence-based emergency medicine. How to find evidence when you need it, part 1: databases, search programs, and strategies. Ann Emerg Med 2002;39:302-6.
17. Meade MO, Richardson S. Selecting and appraising studies for systematic review. Ann Intern Med 1997;127:531-7.
18. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? [review]. Health Technol Assess 2003;7(1):1-76.
19. Clarke M, Oxman AD, editors. Cochrane reviewers' handbook 4.1.6 [updated January 2003]. In: The Cochrane Library, Issue 1, 2003. Oxford: Update Software. Updated quarterly.
20. Gallagher PE, Allen TY, Wyer PC. How to find evidence when you need it, part 2: a clinician's guide to MEDLINE: the basics. Ann Emerg Med 2002;39:436-40.
21. Gallagher PE, Allen TY, Wyer PC. How to find evidence when you need it, part 3: a clinician's guide to MEDLINE: tricks and special skills. Ann Emerg Med 2002;39:547-51.
22. Dickerson K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. BMJ 1994;309:1286-91.
23. Langham J, Thompson E, Rowan K. Identification of randomized controlled trials from the emergency medicine literature: comparison of hand searching versus MEDLINE searching. Ann Emerg Med 1999;34:25-34.

Correspondence to: Dr. Peter J. Zed, CSU Pharmaceutical Sciences, Vancouver General Hospital, 855 West 12th Ave., Vancouver BC V5Z 1M9; 604 875-4077, fax 604 875-5267, zed@interchange.ubc.ca