

Reviews in accident and emergency medicine: the past and the future

R D Hardern, D W Hamer

Abstract

Objective—To determine whether reviews published in the past 10 years are as helpful as they could be to their readers, and to explain why future reviews should meet certain methodological criteria.

Methods—The quality of 16 reviews published in two journals dealing with accident and emergency medicine over 10 years was objectively assessed using 11 currently recommended criteria.

Results—The median number of methodological criteria satisfied was two. Only five of the reviews posed a specific question or problem and then answered this with conclusions based on evidence presented.

Conclusions—If reviews are to enhance patient care, the way in which they are performed, written and used must change.

(*J Accid Emerg Med* 1996;13:169-172)

Key terms: accident and emergency medicine; systematic reviews

“If people are to benefit from the results of trials all the steps between research and practice must be accomplished effectively”.¹

It is impossible for an individual to remain up to date with the literature; over two million articles are published in biomedical journals each year. If stacked, they would reach a height of 500 metres.² The results of studies, even when well conducted, do not always affect the delivery of health care or its outcomes.

One solution to this is for clinicians to read review articles and base clinical practice on their conclusions. Reviews should provide a manageable amount of useful information based on a thorough literature search, and highlight evidence that should influence practice. Such “systematic” reviews are overdue: “It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, updated periodically, of all relevant randomised controlled trials”.³

If conducted appropriately, reviews can provide unbiased guidance for clinicians. Because they are such a crucial link in the chain between research and clinical outcomes, the importance of the science of reviewing research has been recognised.⁴ Until very recently review articles and “authoritative texts” have been subjective: “expert” opinion does not always reflect published data.⁵ Unsystematic reviews lead to biased conclusions.^{6,7} Once a

review is published, any errors may be perpetuated.⁸

To avoid errors arising from bias, “secondary” research (as integrative or review studies are termed) requires the same degree of methodological rigour as is expected in “primary” research. Few people would question the superiority of randomised controlled trials in primary research (studies to assess the value of an intervention). An analogous change seems likely to occur in the field of secondary research (integrative studies or reviews).

The quality of reviews in accident and emergency (A&E) medicine has not been assessed, nor has the “science” of secondary research been discussed in United Kingdom A&E medical journals. The aim of this study was to determine whether reviews published in the past 10 years are as helpful to their readers as they could be. Its other aim is to explain why we feel it important that future reviews meet certain methodological criteria. We hope to promote discussion rather than produce a definitive guide to carrying out systematic reviews.

Methods

There is no gold standard set of criteria to assess the quality of reviews; no single system has proved to be superior.⁹⁻¹² We selected, before retrieving the reviews themselves, 11 criteria common to most of these papers (see Discussion and appendix 2).

A Medline search was carried out. All articles with the classification “review” in Archives of Emergency Medicine (1984 onwards) and the Journal of Accident and Emergency Medicine (up to volume 11, number 3 inclusive) were assessed. These journals were also hand searched. Case reports and reviews of clinical cases were excluded. Sixteen reviews were identified (appendix 1).

The reviews were blinded, so neither author nor institution was visible, then scored (“yes”, “no”, or “not applicable”) for each criterion, independently, by both authors. The two sets of initial answers were compared and the degree of agreement estimated using the κ statistic.¹³ It was decided, before the reviews were assessed, to attempt to resolve disagreement by discussion, but if this was not possible to ask staff at the NHS Centre for Reviews and Dissemination to adjudicate.

Results

There was a very high level of agreement between us. Of 176 criteria (16 reviews with 11 criteria each) the answers were different in 23

**Accident and
Emergency,
The General
Infirmary,
Leeds**

R D Hardern
D W Hamer

Correspondence to:
Dr R D Hardern,
Accident and Emergency,
St James's University
Hospital,
Beckett Street,
Leeds LS9 7TF,
United Kingdom.

instances ($\kappa = 0.85$). It is likely that other assessors would reach the same conclusions. Where the initial answers did not agree, it was possible to reach agreement after discussion.

The median number of criteria satisfied was 2 (of a maximum 11). The range was 0 to 5, and the interquartile range 2 to 3.75. Six criteria (numbered 2 to 7 in the discussion and appendix 2) were not satisfied by any of the reviews. Criterion 8 ("Were conclusions or recommendations made?") was the most frequently satisfied: 13 reviews contained specific conclusions. In four of these, however, the conclusions were not based on the data presented. Six reviews did not make their aim clear. In only five of the reviews was there a clear question followed by a clear answer based on the evidence contained in the review. Directives for future research were made in six reviews.

Discussion

All the reviews we assessed had methodological shortcomings (as judged by current standards). Our aim was not to denigrate the efforts of the authors of these reviews (scores have not been given for individual papers) but to highlight the scope for improvement. This situation is not unique to A&E medicine⁹; the data should not be used to decry the specialty.

The systematic review should usurp the traditional. The key difference between the two is that systematic reviews are carried out with methods chosen to eliminate bias. Groups such as that responsible for the Effective Care in Pregnancy and Childbirth (ECPC) database¹⁴ led the way. They analysed relevant randomised controlled trials and combined the findings to decide which interventions are beneficial. The Cochrane Collaboration aims to establish systematic reviews of randomised controlled trials in all areas of medicine¹⁵:

"When rigorously performed, these systematic exercises (which involve a good deal more than statistical tactics such as meta-analysis) employ three powerful strategies to protect the truth – currency, scientific rigour, and openness in execution...The Research and Development arm of the British National Health Service, and a group of collaborators around the world, have now recognised the need for a mechanism to expand the power of the up-to-date systematic review to encompass all of health care, and to expand the means of communicating their results."

CRITERIA THAT SHOULD BE SATISFIED BY A REVIEW

The aims and techniques of systematic reviews have been discussed in a recent series of articles that are strongly recommended.^{4 7 16 17} Readers wanting an "authoritative" guide are advised to consult section VI of the *Cochrane collaboration handbook*.¹⁸

The 11 criteria used to evaluate the reviews are not original but have been proposed by others.⁹⁻¹²

(1) The first criterion is that the purpose of the review is made clear. Readers often refer to a review article when faced with an unfamiliar

clinical problem. Unless the purpose of the review is clearly stated at the outset, the reader may have to read the whole review before knowing if it is relevant to them. This information should be in the title or abstract; clinicians using a computerised search rely on the title and abstract when deciding which articles to retrieve. The aims of this review were detailed in the abstract and in the introduction.

(2) A potential source of bias is when reviewers include only those articles that back up their own point of view. Reviews should describe the method used to identify data (the original papers). This section of a review is as important as the methods section of an experimental paper. It should contain enough information for the reader to be able to repeat the search. As most reviewers will now make use of on line or CD-ROM databases such as Medline, this section should include the search strategies used. The search strategy used in this review was detailed in the methods section.

(3) Computer searching identifies around 50% of relevant randomised controlled studies.¹⁶ Reviews that do not use other methods of identifying studies are inadequate as a basis for clinical practice. The reference list of all articles identified from the computer search should be scrutinised and relevant studies retrieved. Further methods of identifying studies include hand searching relevant journals, personal communication with the authors of identified studies, and a search of the "grey" literature. We suggest that unless a reviewer describes the methods used to identify data, and unless these methods are adequate, the conclusion of the review should be disregarded (in the same way as one would disregard the conclusions of a study that did not have appropriate randomisation or blinded assessment of outcome measures). We hand searched the journals being assessed.

(4) Papers with a negative outcome are less likely to be published (publication bias). This hinders attempts to identify all relevant data. Ideally, therefore, a review should include some assessment of the likelihood of publication bias (such as the "funnel plot" method¹⁹). This is not relevant to this review; it was not examining the efficacy, or otherwise, of a specific intervention.

(5, 7) Some primary research is of poor quality: randomisation may be non-existent or inadequate; there may be contamination of the control group, high withdrawal rates, or inappropriate statistical analysis. One would not normally change clinical practice on the basis of one such paper. Many people therefore feel that it is inappropriate to alter behaviour based on pooling of the data from several such studies. Because of this, some reviewers exclude papers that fail to meet certain methodological criteria. No single set of selection criteria has been shown to be superior. It is important, if a bias is to be avoided, to decide upon the criteria to be used before performing the literature search. Enough information about the methodological criteria used should be included to allow the

reader to decide if they have been applied consistently. Because of the uncertainty about the best way of including or excluding studies, reviewers should include a "sensitivity analysis" if they have excluded studies on the grounds of poor methodology. This is an assessment of the impact of applying or not applying methodological criteria. The methodological criteria used in this review were decided upon before the reviews being assessed were retrieved. They are described and referenced; it is possible for readers of this review to judge for themselves whether they have been applied appropriately. No sensitivity analysis was carried out as no reviews were excluded.

(6) Data from different studies can be combined in different ways. Meta-analysis is used increasingly often. This is a formal statistical technique to combine numerical data quantitatively. Such techniques are not always appropriate or valid. Readers of a review should check that the methods used are appropriate (although a journal's editor and referees should ensure that reviews using inappropriate techniques are returned for revision).

Data from the reviews assessed were not combined in this review because they dealt with different topics.

(8, 9) Having identified and selected relevant studies using methods designed to eliminate bias, and taken pains to combine data appropriately, it is important that the conclusions or recommendations are based on these data, not the reviewer's opinion. Unless the conclusions are the same as those that would be made by somebody else presented with the same evidence, they should be disregarded. We believe the conclusions of this review are clear and have been based on the data contained within it.

(10) In some circumstances there are compelling data to support a particular intervention. One example is the use of streptokinase in patients presenting with acute myocardial infarction. Often, however, data are less compelling. It is probably useful for readers if the reviewer links that the strength of the recommendation to the strength of the data.^{20 21}

(11) Somebody who has performed a systematic review of a particular topic should have an appreciation of "gaps" in the literature or areas for which no conclusions can be made. A good systematic review will identify these and provide directives for future research. We have not identified specific areas for research but have, on the basis of our findings, made recommendations about the way in which secondary research should be carried out.

STRUCTURED ABSTRACTS

Structured abstracts enable readers to establish rapidly if a review is relevant to their needs and may help authors of reviews to address important aspects of methodology. The following headings have been proposed^{22 23}:

Objective: a precise statement of the primary objective of the review, including details of

populations and patient types considered, the interventions considered, and the outcomes considered.

Data identification: a succinct summary of data sources. There should be sufficient detail in the main article for the search to be reproduced. Any limitation, such as English language only, should be mentioned.

Study selection: including the number of studies identified that were actually selected for analysis and how they were selected, that is, inclusion and exclusion criteria and the methods used to apply them.

Data extraction: this section should describe how the data were extracted and how their validity was assessed.

Data synthesis: the main results of the review should be stated with the method used to obtain them. A review of a diagnostic test should include details of sensitivity, specificity, positive and negative likelihood ratios, a ROC curve, and predictive values.

Conclusion: the results and their interpretation should be clearly stated. Relevant future studies may be suggested.

CONCLUSIONS

The reviews that we identified did not meet the criteria we feel should be met by reviews published in the future. Unless the quality of reviews improves they will continue to be subject to bias and, therefore, potentially lead to ineffective (or even harmful) care. The brunt of this will be born by our patients.

The responsibility for high quality reviews appearing in this journal rests with authors, with the editorial staff and referees, and with readers. We share a duty to our patients to demand the highest standards from reviewers and from reviews.

Appendix 1

REVIEWS ASSESSED

- Barnes SM, Milsom PL. Abscesses: an open and shut case. *Archives of Emergency Medicine* 1988;5:200-5.
- Beerhuizen GI, Goris RJ, Kreuzer FJ. Skeletal muscle PO₂ during imminent shock. *Archives of Emergency Medicine* 1989;6:172-82.
- Branfoot T. Motorcyclists, full-face helmets and neck injuries: can you take the helmet off safely, and if so, how? *Journal of Accident and Emergency Medicine* 1994;11:117-20.
- Cumming AD. The role of calcium in intravenous fluid therapy. *Archives of Emergency Medicine* 1993;10:265-70.
- Donnelly SC, Robertson C. Trauma, inflammatory cells and ARDS. *Archives of Emergency Medicine* 1993;10:108-11.
- Fulde GW, Harrison P. Fat embolism - a review. *Archives of Emergency Medicine* 1991;8:233-9.
- Deakin CD, Hicks OR. AB or ABC: pre-hospital fluid management in major trauma. *Journal of Accident and Emergency Medicine* 1994;11:154-7.
- Harborne DJ. Emergency treatment of adder bites: case reports and literature review. *Archives of Emergency Medicine* 1993;10:239-43.
- Mazurek AJ. Epidemiology of paediatric injury. *Journal of Accident and Emergency Medicine* 1994;11:9-16.
- Randall P, Banks J, Little RA. Medical (military) anti-shock trousers - a short review. *Archives of Emergency Medicine* 1984;1:39-51.
- Robertson C. (1985) A review of the use of corticosteroids in the management of pulmonary

injuries and insults. *Archives of Emergency Medicine* 1985;2:59-65.

Robertson C. Open-chest cardiac massage for non-traumatic cardiac arrest. *Archives of Emergency Medicine* 1987;4:207-210.

Russel DC. Experimental studies into mechanisms of cardiac arrest. *Archives of Emergency Medicine* 1984;2:79-88.

Schieden V. Violence at work. *Archives of Emergency Medicine* 1993;10:79-85.

Smith JAR. Steroids: a surgeon's view. *Archives of Emergency Medicine* 1985;2:131-6.

Trunkey D. Towards optimal trauma care. *Archives of Emergency Medicine* 1985;2:181-95.

Appendix 2

THE CRITERIA USED TO ASSESS QUALITY

- 1 Does the review specify its purpose? [ie, could the reader tell from the title and abstract the issues addressed?]
 - 2 Was the method used to identify relevant studies identified in sufficient detail for a reader to be able to reproduce the literature search?
 - 3 Was the search method adequate? [Must include computer search (with search strategy) for studies and reviews and all references from studies identified this way together with at least one of: hand searching of relevant journals, personal communication with authors of identified studies, or "grey literature" search.]
 - 4 Was there a formal assessment for evidence of publication bias?
 - 5 Was the study selection method described? [There should be sufficient detail to allow the reader to judge whether other studies (perhaps published after the review) would be included were the review to be repeated.]
 - 6 If data were combined in a quantitative manner, was the method described?
 - 7 Was a sensitivity analysis made? [An assessment of the impact of including/excluding different studies (eg, according to their methodological quality) on the conclusions of the review.]
 - 8 Were conclusions or recommendations made?
 - 9 If conclusions or recommendations were made were they based on the evidence presented? If someone else was presented with the same evidence would they reach the same conclusions?
 - 10 If conclusions or recommendations made were they formally linked to the "strength" of the evidence?
 - 11 Were specific directives for future research made?
- 1 Chalmers I, Dickerson K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992;305:786-8.
 - 2 Ad hoc Working Group for Critical Appraisal of the Medical Literature. Academia and clinic: a proposal for more informative abstracts of clinical articles. *Ann Intern Med* 1987;106:598-604.
 - 3 Cochrane AL. Effectiveness and efficiency. Random reflections on health services. London: Nuffield Provincial Hospitals Trust, 1972.
 - 4 Chalmers I, Haynes B. Reporting, updating, and correcting systematic reviews of the effects of health care. *BMJ* 1994;309:862-5.
 - 5 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analysis of randomised control trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA* 1992;268:240-8.
 - 6 Haynes BH. Clinical review articles. *BMJ* 1992;304:330-1.
 - 7 Knipschild P. Systematic reviews: some examples. *BMJ* 1994;309:719-21.
 - 8 Crichton RR. Complacent reviewing. *Nature* 1989;337:110.
 - 9 Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987;106:485-8.
 - 10 Oxman AD, Guyatt GH. The science of reviewing research. *Ann NY Acad Sci* 1993;703:125-33.
 - 11 Oxman AD. Checklists for review articles. *BMJ* 1994;309:648-51.
 - 12 Oxman AD, Cook DJ, Guyatt GH, for the evidence-based medicine working group. Users' guides to the medical literature. VI. How to use an overview. *JAMA* 1994;272:1367-71.
 - 13 Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ* 1992;304:1491-4.
 - 14 Chalmers I, Enkin M, Keirse MJNC, eds. *Effective care in pregnancy and childbirth*. Oxford: Oxford University Press, 1989.
 - 15 Cochrane's legacy. *Lancet* 1992;340:1131.
 - 16 Dickerson K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994;309:1286-91.
 - 17 Mulrow CD. Rationale for systematic reviews. *BMJ* 1994;309:597-9.
 - 18 Oxman AD. Preparing and maintaining systematic reviews. In: *Cochrane collaboration handbook*, 1994. [e-mail: GENERAL@COCHRANE.CO.UK.]
 - 19 Vandembroucke JP. Passive smoking and lung cancer: a publication bias? *BMJ* 1988;296:391-2.
 - 20 Cooke DJ, Guyatt GH, Laupacis A, Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1992;102:305-11S.
 - 21 Emergency Cardiac Care Committee and Subcommittees, American Heart Association. Guidelines for cardiopulmonary resuscitation and emergency cardiac care. Part I. Introduction. *JAMA* 1992;268:2171-83.
 - 22 Mulrow CD, Thacker SB, Pugh JA. A proposal for more informative abstracts of review articles. *Ann Intern Med* 1988;108:613-5.
 - 23 Squires BP, Keith RG, Meakins JL. Structured abstracts for clinical research manuscripts and reviews. *Can J Surg* 1992;35:473-5.