

How to Read a Review Paper

Rachael A Callcut MD and Richard D Branson MSc RRT FAARC

Introduction

Limitations of Narrative Reviews

Systematic Reviews

The GRADE System

How Does the GRADE System Consider Strength of Recommendation?

Meta-analyses

Meta-analyses: What Are They? How to Interpret Them

How to Read a Review Paper

What Is the Purpose of the Review?

Is the Topic Clearly Defined and Is There a Focused, Important Clinical Question?

Is the Literature Search Thorough and Were All Possible Sources of Information Evaluated?

What Criteria Were Used to Select Papers for Inclusion and Were These Criteria Predefined?

Are the Results Sensitive to the Review Methodology?

Other Questions

Summary

Review papers commonly summarize the current knowledge on a selected topic. These types of papers are considered narrative reviews. Narrative reviews rarely detail the methods used to select the literature included, nor do the authors typically report the purpose of the review. Narrative reviews may be biased due to inadequate literature reviews or individual beliefs. A systematic review limits bias by disclosing the purpose of the paper, the assembly of the literature, and the appraisal of study quality. A meta-analysis, a specific style of systematic review, quantitatively pools data from individual studies for re-analysis. Pooling data increases the sample size and improves statistical power. The common representation of a meta-analysis is the forest plot. The forest plot demonstrates the odds ratio of individual studies, the weight each trial contributes to the analysis, and the 95% confidence intervals. Systematic reviews and meta-analyses are not without shortcomings, including issues related to study heterogeneity. Because of their transparency, systematic evaluations of the literature are superior to narrative reviews. *Key words: narrative review, literature review, systematic review, meta-analysis, forest plot.* [Respir Care 2009;54(10):1379–1385. © 2009 Daedalus Enterprises]

Rachael A Callcut MD and Richard D Branson MSc RRT FAARC are affiliated with the Department of Surgery, Division of Trauma/Critical Care, University of Cincinnati, Cincinnati, Ohio.

Dr Callcut has disclosed no conflicts of interest. Mr Branson has disclosed relationships with Ikaria, Cardinal, Newport, and Covidien.

Mr Branson presented a version of this paper at the RESPIRATORY CARE Journal Symposium at the 54th International Respiratory Congress of the American Association for Respiratory Care, held December 13-16, 2008, in Anaheim, California.

Table 1. A Comparison of Narrative and Systematic Reviews

Characteristic	Narrative Review	Systematic Review
Topic or question	Broad overview without a specific question	Narrow scope, with a specific question to be answered
Literature sources	Limited sources, typically not specified	Wide variety of specifically named databases
Literature selection	Unspecified, potentially biased	Pre-determined specific criteria for selection of papers
Literature appraisal	Unspecified, variable, and potentially biased	Critical review using specific criteria
Literature synthesis	Qualitative summary	Quantitative summary
Recommendations	Opinion, potentially biased, not evidence-based	Evidence-based

Introduction

Review papers are a common feature in medical journals and frequently published in the pages of *RESPIRATORY CARE*. Most readers become exposed to reviews in college or in the form of term papers written as a class assignment. In the latter case a literature search provided a list of papers that were generally selected by convenience or because they supported the writer’s hypothesis. Papers with a contrary opinion were typically ignored and not discussed in the final manuscript. This type of journalistic or narrative review is common in all disciplines, with medicine being no exception.¹⁻⁴

Limitations of Narrative Reviews

Narrative reviews are limited by incomplete searches of the literature, intentional or unintentional bias by the authors, and failure to account for the quality of individual publications. Additionally, narrative reviews typically fail to effectively deal with studies having conflicting results. As reviews are supposed to summarize the available literature, and clinicians may rely on this summary as the current state of the art, the entire body of literature should be explored.

Narrative reviews by experts are notoriously biased, simply because experts tend to rely on their own expertise and experience rather than on the available evidence.⁵ Authors may be asked to write a review because of their expertise in the area, or they may contribute in an effort to advance their views, or both. Bias secondary to conflicts of interest are inevitable, but are somewhat alleviated by disclosure prior to publication.

Bias can be introduced by failing to perform a complete review of the available literature, as well as through choice of the literature included. Authors may fail to find important papers by limiting their search to English-only sources or to a single database. Alternatively, an author may choose

only articles that support his or her view, regardless of quality. As an example, great weight may be placed on case reports and case series published in low-quality journals. Despite the admonition that the plural of anecdote is not data, authors may overemphasize the importance of case reports.

Validity refers to the methodological quality of the review. If papers are selected without regard to quality and without objective methods for interpretation, the validity of the review is in question. In order to assure validity, the author should detail the methods by which papers were selected, assessed, and analyzed. This information allows the reader to evaluate for potential bias in the selected literature.

Simply stated, most narrative reviews are biased by the opinions of the authors, have questionable validity, and often make inappropriate recommendations when compared to the actual evidence.³

Systematic Reviews

A systematic review is different from a narrative review in purpose and process (Table 1). A systematic review typically addresses a specific question about a topic. For example, a narrative review might consider the topic of noninvasive ventilation, with a title of “Noninvasive ventilation in the hospital.” A systematic review would refine the topic and make the topic more specific. A systematic review might then be titled “Success of noninvasive ventilation in preventing reintubation following extubation failure in postoperative coronary bypass grafting patients.” The first broad title allows the author a number of avenues to explore and interject opinion along with facts from the literature. The second title is explicit and requires that the literature be pared to papers specifically dealing with this issue, and the review must focus on answering the questions. Table 2 lists proposed advantages of systematic reviews.

A systematic review also details the methods by which papers were identified in the literature. This includes the search terms and the search engines utilized. As an example, a search using only MEDLINE may miss many papers in the nursing literature, which are easily accessed via the

Correspondence: Richard D Branson RRT FAARC, Department of Surgery, University of Cincinnati, 231 Albert Sabin Way, Cincinnati OH 45267-0558. E mail: richard.branson@uc.edu.

Table 2. Proposed Advantages of Systematic Reviews

Explicit methods for searching and grading the literature limit bias
Conclusions are more reliable and less prone to author opinion and bias, owing to explicit methods of literature identification and review
Large volumes of information are made accessible to stakeholders
Best practices can be identified and disseminated quickly, potentially resulting in an improvement in patient care
A wide variety of studies can be compared to determine generalizability of results and identify inconsistency
Inconsistency in results can be identified and used for asking new research questions to address individual populations

Table 3. Quality of Evidence and Definitions Used Within the GRADE system

Grade	Definition
High	Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low	Any estimate of effect is very uncertain.

Cumulative Index to Nursing and Allied Health Literature (CINAHL). If the search term is related to complications of airway suctioning, the MEDLINE search will miss a number of papers.

A systematic review also uses predetermined criteria for selection of papers to be included in the review. For example, the author may choose to include only randomized controlled trials (RCTs) in the review, and therefore all case reports, case series, and studies with no control group would be eliminated. The important aspect is that these criteria are decided a priori. Explicit methods for appraising the chosen literature must be used for evaluating the quality and validity of individual studies. Systems for grading the evidence include the Cochrane methodology and the Grading of Recommendations Assessment, Development, and Evaluation (GRADE).^{6,7}

The GRADE System

In an effort to demonstrate transparency and simplicity, GRADE classifies the quality of evidence using 4 levels: high, moderate, low, and very low (Table 3). In some instances, authors may choose to combine the low and very low categories. The highest quality evidence is generally considered to arise from RCTs. However, the GRADE system allows even an RCT to be downgraded in strength if there are important study limitations. Examples

would include inconsistency of results, indirectness of evidence, reporting bias, and imprecision. Typically, observational studies (case-control studies) are initially graded as being of “low quality.” However, the GRADE system allows these studies to be upgraded under certain conditions, for example, if the magnitude of the treatment effect is very large, if there is evidence of a dose-response relationship, or if all biases would increase the magnitude of a treatment effect.

How Does the GRADE System Consider Strength of Recommendation?

The GRADE system allows for 2 grades of recommendations: “strong” and “weak.” When an intervention’s desirable effects clearly outweigh its undesirable effects, or clearly do not, strong recommendations can be made for or against the intervention. However, when the tradeoffs are less certain—either because of low-quality evidence or because evidence suggests desirable and undesirable effects are similar—weak recommendations are warranted.

In addition to the quality of the evidence, a number of other factors determine whether recommendations are strong or weak. Examples include quality of evidence, uncertainty about the balance between desirable and undesirable effects, and uncertainty or variability in values and preferences. Many professional societies have adopted GRADE for writing clinical practice guidelines, as a consequence of the straightforward approach and transparency.

Meta-analyses

A meta-analysis is a quantitative systematic review that uses statistical methods to improve the precision of conclusions in answering a question. Grading individual research studies allows studies with the most rigorous methodology to carry more weight in the final review than those studies that lack important methodological elements. A quality systematic review is capable of compiling the data from individual studies performed across time and geography into a cogent analysis of the question, avoiding bias despite conflicting study results.

Meta-analyses: What Are They? How to Interpret Them

A meta-analysis is a type of systematic review in which data are statistically combined from 2 or more studies to determine the outcome of a specific research question.⁸⁻¹⁰ The major advantage of a meta-analysis is that the statistical power of the analysis can exceed that of the individual primary studies because of the ability to achieve a pooled, larger sample size. This type of study is most often

utilized in the medical community for combining data from RCTs. Observational studies are also amenable to meta-analysis, although the usage of such an analysis is controversial, as observational studies are more prone to bias, and this can significantly impact the outcome of a meta-analysis.^{8,9}

The strength of a meta-analysis is a function of not only the quality and validity of the original studies, but also of the methods utilized for identifying, selecting, and analyzing which original studies to include in the meta-analysis. The authors of meta-analyses should describe in the methods section the search strategy employed for identification of all relevant articles.^{4,10} Ideally, more than one source should be used to ensure all articles have been identified, as single electronic database searches have low sensitivity, even for locating RCTs.^{4,8}

The authors then have to determine which of these articles should be included in the meta-analysis. This can be done by agreement between the lead authors or by consensus of the group. Another method involves calculating the Jadad score.¹¹ Jadad et al published a 3-point questionnaire to judge the quality of RCTs. Each question is answered yes or no. Each yes scores a single point, each no results in a score of zero points. Jadad suggested that this score would allow the reader to evaluate a paper in less than 10 minutes. The questions that compose the Jadad score are shown below.

1. Was the study described as randomized?
2. Was the study described as double-blind?
3. Was there a description of withdrawals and dropouts?

With respect to withdrawals and dropouts, in order to receive a point, the paper must describe the number of withdrawals and dropouts in each group, and the reasons.

Additional points are awarded if:

1. The method of randomization is described and that method is appropriate.
2. The method of blinding is described and appropriate.

Points are deducted if the converse are true: the method of randomization is described but is inappropriate, or the method of blinding is described but is inappropriate.

An individual RCT then can receive a Jadad score of between zero and five, zero being a trial of poor quality and 5 being a high-quality study.

A meta-analysis has to have a specific research question that is as close as possible to the hypothesis that was investigated in the primary studies. Each of the primary studies ideally should have a consistent study group similar for all important characteristics that can affect outcome, have the same intervention, and collect the same end points. Primary publications have to also contain

enough information to allow the relevant data points to be extracted and independently analyzed. The authors should explicitly state the type of studies included, to allow the reader to independently conclude whether such studies can adequately answer the research question.⁸

A well done meta-analysis will include 2 independent reviewers, to avoid bias or error that can occur in the selection of articles to include or when data are extracted from the studies.⁹ If there is conflict with selection or extraction, a third reviewer is commonly utilized to reach consensus. These data extractors have to pay particular attention to the discrepancies that can be found in publications and cross-check the abstract with the text of the articles to avoid erroneous data. Some meta-analyses also report the degree of agreement between the reviewers, as concordance or a kappa statistic (the higher the kappa statistic, the more consistent the data are between extractors).

After data extraction, the results are pooled and a test for homogeneity should be reported. Homogeneity refers to the consistency of results between the included primary studies. Homogeneity is often statistically represented by a type of chi-square test known as the Q test.⁴ However, most often the degree of agreement between the results of the primary studies are graphically represented in forest plots. This allows the reader to more easily understand an individual study's contribution to the overall treatment effect. This is especially important when there is a large degree of heterogeneity (inconsistency) of the treatment effect between primary studies, because a pooled result may be misleading.⁹ Thus, the combined treatment effect does not allow the reader to readily understand the individual study contributions, and forest plots provide a visual way for the reader to identify such contributions.

Importantly, homogeneity between the results of the primary studies does not necessarily mean that there are no relevant differences in the study populations, methods of the original studies, interventions, or data analysis. However, when a large degree of heterogeneity exists between studies, this is most often a reflection of such differences.⁹ Generally, if a meta-analysis contains substantial heterogeneity, the results are more skeptically viewed. Heterogeneity is quantified by the I^2 statistic (percentage of total variation across studies that is attributed to heterogeneity rather than chance) and a lower percentage is indicative of lower heterogeneity (or more homogeneity).⁹

The I^2 and the Q score evaluate heterogeneity and homogeneity, but are not opposite sides of the same coin. Rarely are both provided in a single meta-analysis, but both can be calculated. There is no evidence to suggest one is better than the other.

The overall result of a meta-analysis is most commonly presented as a forest plot (Fig. 1). Treatment effects can be reported in meta-analyses using odds ratio and/or relative

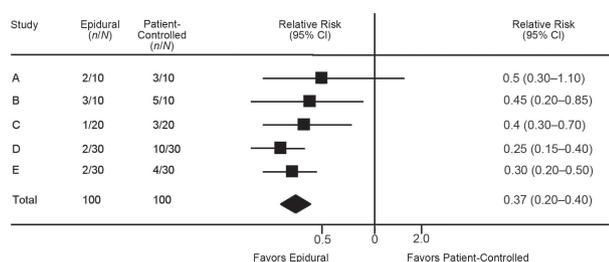


Fig. 1. Example of a forest plot, depicting the important features. (Hypothetical data for illustrative purposes only.)

risk for dichotomous outcomes (for example, disease vs no disease), or as mean differences for continuous outcomes (for example, number of blood products needed).⁸ Odds ratio is the ratio of the odds of an outcome in the treatment group to the odds of an outcome in the control group. Alternatively, risk is the number of patients in a group who have the outcome of interest divided by the entire number of patients in the group of interest. Thus, the relative risk is the risk in the treatment group divided by the risk in the control group.

Generally, if the odds ratio or relative risk exceeds 1, the likelihood of the outcome is greater in the treatment group. Alternatively, if the value is below 1, the outcome is less likely in the treatment group. The closer the value is to 1, the more similar the outcomes are in the treatment and control groups. Therefore, if the confidence interval overlaps 1, the results are not considered to be statistically different from one another. Additionally, the wider a confidence interval is, the less precise the treatment effect is considered and often can signify smaller sample sizes.

In a forest plot the outcome is placed on the X axis, and in Figure 1 this is the relative risk of requiring intubation in patients with rib fractures treated with epidural anesthesia pain control versus patient-controlled analgesia. The vertical line centered at relative risk = 1.0 represents the no-effect line.^{4,8} Confidence intervals that cross this vertical line indicate the study groups have equal risk of the outcome of interest (intubation). The square boxes are the point estimates for each study, and often are presented as different-sized boxes, which correspond to the weight given to the study. A larger box reflects a higher weight assigned to that study. In some instances the weight is simply the number of patients contributing to the overall sum. In others the size of the box is a function of both the number of patients and the quality of the trial. The cumulative treatment effect is represented by the diamond symbol at the bottom of the diagram. The size of the diamond has meaning. The center of the diamond represents the point estimate of the combined result, and the width of the diamond represents the 95% confidence interval of the point estimate. In Figure 1, when the relative risk falls below and does not cross 1, the likelihood of intubation is lower

in the epidural group, compared with the patient-controlled analgesia group.

One of the most challenging aspects of performing, reporting, and interpreting meta-analyses is overcoming publication bias, which can substantially impact the quality of the analysis. Publication bias reflects the increased likelihood of a study being published when the study has a positive result.¹² The effect of this is that if a meta-analysis is based solely upon the published literature, an intrinsic bias toward a positive treatment outcome will be incorporated into the study, because fewer negative or equivocal studies exist in the literature. The authors and readers of meta-analyses often will not know the true extent of publication bias, because it is impossible to know how many equivocal or negative studies were terminated before completion or not published. However, statistical techniques (such as funnel plots) have been developed to aid in the identification of such bias.¹⁰

Although some have advocated incorporating unpublished studies into a meta-analysis, these unpublished studies have not necessarily been subjected to the same level of scientific scrutiny as published studies. Therefore, bias can inadvertently be introduced by utilizing non-peer-reviewed data as well. In addition, those selecting papers for inclusion into a meta-analysis may also demonstrate a preference for studies published in English, those that are readily available, appear in more prestigious journals, or that support their own view point on a topic. Some reviews have failed to include up to 50% of the reported trials, and therefore these issues are important to understand before using a meta-analysis in clinical practice.¹³

If done correctly, meta-analysis can be a powerful research modality. A well done study requires that a clearly defined research question was identified prior to data collection, the literature review and selection was systematic and reproducible, the quality of the selected studies was assessed, and that the outcome variables were consistent between studies.^{8,14} Additionally, the statistical methods for pooling data must be appropriate, and issues of heterogeneity must be addressed. When incorporating the results of a meta-analysis into clinical practice, the reader should consider the strength of the analysis performed, and, most importantly, the overall relevance of the study to their individual patient population.

The largest repository of meta-analyses can be found in the Cochrane Collaboration, online at <http://www.cochrane.org>. A select list of topics pertinent to respiratory care from the 1,240 available is shown in Table 4.

How to Read a Review Paper

Given the methodology reviewed here and the number of caveats potentially present in a review paper, reading a review requires active participation. That is, rather than

Table 4. A Sample of the 1,240 Respiratory-Care-Related Topics in the Cochrane Collaboration Available Online

Inhaled nitric oxide for acute hypoxemic respiratory failure in children and adults
Animal-derived surfactant extract for treatment of respiratory distress syndrome
Palivizumab for prophylaxis against respiratory syncytial virus infection in children with cystic fibrosis
Digoxin for preventing or treating neonatal respiratory distress syndrome
Lung-protective ventilation strategy for the acute respiratory distress syndrome
Ribavirin for respiratory syncytial virus infection of the lower respiratory tract in infants and young children
Delayed antibiotics for respiratory infections
Advising patients to increase fluid intake for treating acute respiratory infections
Antibiotic prophylaxis to reduce respiratory tract infections and mortality in adults receiving intensive care
Quantitative versus qualitative cultures of respiratory secretions for clinical outcomes in patients with ventilator-associated pneumonia
Nitric oxide for respiratory failure in infants born at or near term
Chest physiotherapy for reducing respiratory morbidity in infants requiring ventilatory support
Extracorporeal membrane oxygenation for severe respiratory failure in newborn infants

just absorbing the information, the reader should ask questions about the purpose and quality of the review. The following questions are among those that should be considered.

What Is the Purpose of the Review?

Do the authors explain the purpose for the review? Most reviews are narrative and can be part of a special issue on a given topic or can stand alone. The reader should consider the author's purpose for the paper. Academic journals require that authors disclose their relationships with industry with respect to the topic of the review. These are the most common potentials for bias. However, disclosure does not imply bias. The reader should judge the author's comments with the knowledge of relationships in mind. Just as importantly, inventors and researchers may bring inherent bias to a given topic, not because of financial interests, but because of status. Academic reputations can be based on the success of therapies and techniques.

Narrative reviews may be persuasive, with the intent of changing practice. Reviews concerning the importance of low-tidal-volume ventilation in acute respiratory distress syndrome are both evidence-based and persuasive. Reviews can be used to encourage changes in practice, specifically movement toward evidence-based "best practices."

Is the Topic Clearly Defined and Is There a Focused, Important Clinical Question?

Narrative reviews may cover broad topics. A systematic review should provide a precise, definitive question. Since a systematic review asks a specific question, the reader should be able to answer yes or no after finishing the paper. As in the example regarding the use of noninvasive ventilation for post-extubation hypoxemia, at the end of the review the reader should know if this therapy should be used. Not all topics are open to such simple answers. Finally, the reader should ask how this information applies to their practice and if the question addresses a current clinical conundrum.

Is the Literature Search Thorough and Were All Possible Sources of Information Evaluated?

A systematic review should list all databases searched. While PubMed remains the most common search engine, not all important data can be captured with a single search. There are a number of other databases and possible sources of papers. Many searches are English-only, which may cause the authors to ignore important papers from the foreign-language literature. Another method for searching the literature is known as "references of references"—that is, reviewing the reference list of papers identified in the original search. Gray literature refers to non-peer-reviewed journals, industry reports, and unpublished papers. Personal communication with experts may also uncover important sources. While not all sources can be tapped, the reader should understand how extensively the authors have searched the known literature. The reader may also want to consider the likelihood that relevant data were missed.

What Criteria Were Used to Select Papers for Inclusion and Were These Criteria Predefined?

A systematic review should identify the criteria used to both include and exclude papers from consideration. These criteria should be defined prior to the literature review and provide appropriate weight, based on methodological quality. For example, the authors may decide prior to searching the literature only to consider RCTs. If all types of literature are included, RCTs should carry more weight than lesser publications, such as pilot studies with no control group. The reader should be able to tell how the authors have chosen the literature used to make recommendations.

Are the Results Sensitive to the Review Methodology?

A sensitivity analysis measures the impact of the results after adjustment of methodology related to the studies. If

the results are unchanged under varying conditions, the strength of the review is enhanced. An example of a sensitivity analysis would be including results of studies from less rigorous trials to results from RCTs. While few readers will have the ability to make such analyses, the reader should ask how the results might change if the authors had included other sources.

Other Questions

Readers may also wonder about the reproducibility of results and the differences in study results. Heterogeneity of results can occur as the result of differing patient populations and by sample size. In general, larger trials drive conclusions, compared to smaller trials. If studies provide consistently conflicting results, the reader should question the method of review or consider the rigor of individual studies.

Finally, the reader should consider how the results of the review should impact their individual practices. Important questions include: Were all clinically important outcomes considered? How should I apply these results to my patients? Do the potential benefits suggested by the review outweigh the potential adverse events and cost?

Summary

Systematic reviews utilize defined methods to summarize all studies addressing a specific clinical question. This allows the systematic review to consider all the relevant known data on a topic. A meta-analysis is a quantitative systematic review that can improve the precision and power of estimates of treatment effects by pooling data from a number of similar trials. Narrative reviews can be helpful in introducing new ideas and summarizing the current state of the art. However, narrative reviews are rarely evidence-based. Reading a review requires that the reader consider

the purpose of the paper, the methods of literature selection, and the applicability of the results.

REFERENCES

1. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med* 1997;126(5):376-380.
2. Mulrow CD, Cook DJ, Davidoff F. Systematic reviews: critical links in the great chain of evidence. *Ann Intern Med* 1997;126(5):389-391.
3. McAlister FA, Clark HD, van Walraven C, Straus SE, Lawson FM, Moher D, Mulrow CD. The medical review article revisited: has the science improved? *Ann Intern Med* 1999;131(12):947-951.
4. Greenhalgh T. Papers that summarise other papers (systematic reviews and meta-analyses). *BMJ* 1997;315(7109):672-675.
5. Antman EM, Lau J, Kupelnick B, Moesteller F, Chalmers TC. A comparison of the results of meta-analyses of randomized controlled trials and recommendations of clinical experts. *JAMA* 1992;268(2):240-248.
6. The Cochrane collaboration. Cochrane collaboration handbook for systematic reviews of interventions. <http://www.cochranehandbook.org>. Accessed August 9, 2009.
7. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001;323(7308):334-336.
8. Akobeng AK. Understanding systematic reviews and meta-analysis. *Arch Dis Child* 2005;90(8):845-848.
9. Zlowodzki M, Poolman RW, Kerkhoffs GM, Tornetta III P, Bhandari M. How to interpret a meta-analysis and judge its value as a guide for clinical practice. *Acta Orthopaedica* 2007;78(5):598-609.
10. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA* 2000;283(15):2008-2012.
11. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gaghan DJ, McQuay HJ. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17(1):1-12.
12. West RR. Evidence based medicine overviews, bulletins, guidelines, and the new consensus. *Postgrad Med J* 2000;76(897):383-389.
13. Jørgensen AW, Hilden J, Gøtzche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic reviews. *BMJ* 2006;333(7572):782.
14. Delaney A, Bagshaw SM, Ferland A, Laupland K, Manns B, Doig C. The quality of reports of critical care meta-analyses in the Cochrane database of systematic reviews: an independent appraisal. *Crit Care Med* 2007;35(2):589-594.