Meta-analysis in Neonatal Perinatal Medicine

Roger F. Soll, MD*

Abstract
Systematic overviews provide a comprehensive and thorough review of the available data from clinical trials. When these reviews include meta-analyses, clinicians can synthesize the results of related studies and gain greater precision in their estimates of the effects of therapy. Even when inconclusive, meta-analyses allow for the exploration of differences between studies and may point toward promising areas of future research (or steer clinicians away from further nonproductive areas). In the field of neonatal-perinatal medicine, systematic overviews have provided the basis of several major changes in guidelines with measurable impact on neonatal outcome.

Objectives After completing this article, readers should be able to:
1. Describe the process of conducting a systematic review.
2. Delineate biases inherent in meta-analyses.
3. Review how heterogeneity affects meta-analyses.

Introduction
It is almost impossible to keep up with the scientific evidence on which clinicians base their practice and policies. The simple exercise of searching a common bibliographic database such as PubMed, using the word “neonate” and limiting the search to “all infants” and “randomized, controlled trials” retrieves more than 7,000 citations. No clinician could possibly keep up with this mass of literature. When multiple trials are available, clinicians are confronted with a situation more like taking a Rorschach test than like interpreting the evidence; what we see more reflects our biases than the evidence at hand.

A systematic approach to reviewing the literature is necessary to address this difficulty. Systematic reviews identify, appraise, and synthesize research-based evidence and present it in an accessible format. (1) Systematic overviews apply specific research strategies to data from all relevant studies. If appropriate, these reviews can include meta-analyses, a quantitative statistical method used to combine the results of similar randomized, controlled trials (RCTs) to produce typical estimates of effect size. By combining information from all relevant studies, meta-analysis can provide more precise estimates of the effect of health care than those derived from the individual studies included within the review. (2)(3) They also facilitate evaluation of the consistency of evidence across studies and the exploration of differences between studies.

The following review discusses the approach to creating a systematic overview and how to interpret the results of meta-analyses. A more detailed and technical discussion of creating a systematic review and meta-analysis is given in “The Cochrane Handbook.” (4)

How is a Systematic Review Conducted?
Systematic reviews are essentially “studies of studies.” They use rigorous scientific methods similar to those used in any clinical trial. The specific objective of the review, framed as an answerable question, provides the backbone of the analysis. Like a clinical trial, a protocol for a systematic review should be developed that clearly states the objectives of the review, the population and intervention of interest, and the methods used at each stage of the review.
The next major step is selection of studies for inclusion. As noted previously, the criteria for study inclusion must be specifically defined in the protocol. This includes the specifics regarding the intervention, the population, and the outcomes assessed. If appropriate, exclusion criteria must be explicitly stated.

Searching for trials for inclusion is much simpler in the age of computer databases. Long gone are the days of sitting in the medical library pouring over volumes of *Index Medicus*. Today there are multiple resources for access to the general medical literature and some resources unique to the field of neonatal-perinatal medicine. Searching the medical literature is now widely available through the internet, including several bibliographic databases. These include the Cochrane Library, MEDLINE, EMBASE, and CINAHL.

Once studies are located, rigorous evaluation of whether they meet the criteria for inclusion is necessary. If included, the study must be assessed for the validity of the methodology in study design, conduct, analysis, and reporting. For the purpose of Cochrane reviews, only randomized or quasi-randomized, controlled trials are included. (5)

Outcome data must be extracted and tabulated for each included trial. There are two distinct stages to the analysis. First, a clinically relevant standard statistic is calculated for each study to describe the observed intervention effect. For example, the standard statistic may be the relative risk (RR), the RR reduction, the risk difference if the data are dichotomous, or a difference between means if the data are continuous. Second, an estimate of the summary (pooled) intervention effect is calculated as a weighted average of the intervention effects estimated in the individual studies. The statistical methods for pooling results are similar to the statistical methods used in analyzing the data for multicenter trials. Pooling the results of similar RCTs increases the statistical power lacking in individual smaller trials and enables the clinician to have greater security in accepting or rejecting treatment differences demonstrated by the trials. (2)

Once the analysis has been completed, it is important to assess the importance of the evidence. Clinical trials may use a variety of statistical techniques in reporting the results. A “statistically significant” reported difference does not make the finding clinically relevant. (6) To assess whether the results of a trial are clinically relevant requires calculation of some simple statistics from study findings (Table). The relative risk reduction (RRR) indicates the relative, but not absolute, reduction in the event rate. The absolute risk reduction (ARR) indicates the absolute reduction in the event rate. If the overall incidence of the event is low, the ARR also will be low, even if there is a relatively large difference in the relative risk. Understanding both the RR and ARR is essential to making any clinical judgment. Also of use is the number needed to treat (NNT), which is calculated by dividing 1 by the ARR. For example, if the ARR is 20%, then the NNT is 5. In other words, five infants would need to be treated to prevent one theoretical event. In addition, confidence intervals (CIs) should be reported with each of these statistics. A 95% CI reflects 95% certainty that the true value of the measure lies within the bounds of the interval.

A meta-analysis can be subject to many of its own biases. Any attempt at pooling results from various studies not only incorporates the biases of the primary studies but adds further bias attributable to study selection and the inevitable heterogeneity of the selected studies. (7) Publication bias, the tendency for investigators to choose positive studies for publication, skews the medical literature toward favorable reports of treatment. Unless the authors of the meta-analysis scrupulously research all available resources, these studies are not located and the meta-analysis could report a false-positive finding. This problem is compounded further by the greater chance that such a false-positive finding will be published. Meta-analysis can offer false-negative conclusions because of inappropriate study selection. If the studies selected cannot be grouped (heterogenous), the positive effects observed with one specific treatment or in one specific

Table. **Key Terminology for Estimating the Size of the Treatment Effect**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Positive</th>
<th>Negative</th>
<th>Risk of Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated (Y)</td>
<td>a</td>
<td>b</td>
<td>Y=a/(a+b)</td>
</tr>
<tr>
<td>Control (X)</td>
<td>c</td>
<td>d</td>
<td>X=c/(c+d)</td>
</tr>
</tbody>
</table>

- Relative Risk (RR) is the risk of the outcome in the treated group (Y) compared to the risk in the control group. RR=Y/X
- Relative Risk Reduction (RRR) is the percent reduction in risk in the treated group (Y) compared to the control group (X). RRR=1-Y/X×100% or 1-RR×100%
- Absolute Risk Reduction (ARR) is the difference in risk between the control group (X) and the treatment group (Y). ARR=Y-X
- Number Needed to Treat (NNT) is the number of patients that must be treated over a given period of time to prevent one adverse outcome. NNT=1/(X-Y) or 1/ARR
population may be lost. To minimize bias, the authors of the meta-analysis and the readers of the review must demand the same methodologic quality from these analyses that they would from individual RCTs. It is essential that all meta-analyses include a prospectively designed protocol, a comprehensive and extensive search strategy, strict criteria for inclusion of studies, standard definitions of outcomes, and standard statistical techniques.

What is Heterogeneity?
Invariably, studies brought together in a systematic review have differences. Variability among studies in a systematic review is termed “heterogeneity.” Clinical heterogeneity refers to variability in the participants, interventions, and outcome included in the studies. Methodologic heterogeneity refers to variability in the study design. Clinical or methodologic heterogeneity may contribute to measurable statistical heterogeneity. Methods have been developed for quantifying inconsistency across studies that move the focus away from testing whether heterogeneity is present (which is almost inevitable) to assessing the impact of heterogeneity on the meta-analysis. (4) The I-squared ($I^2$) statistic can be used to evaluate whether substantial heterogeneity is present and may influence the interpretation of the analysis. If noted, heterogeneity can be explored using subgroup analyses.

Sample Meta-analysis: Early Corticosteroids for the Prevention of Chronic Lung Disease
To understand a meta-analysis, it may be useful to examine a systematic review and meta-analysis that has influenced practice in neonatal perinatal medicine, such as early corticosteroids for the prevention of chronic lung disease in preterm infants. (8) This review examines the relative benefits and adverse effects of postnatal corticosteroids administered within the first 7 days of birth to preterm infants at risk of developing chronic lung disease. The review can be found on the National Institute of Child Health and Human Development web site (http://www.nichd.nih.gov/cochrane/hallida3/hallida.htm) or in the Cochrane Library.

Chronic lung disease remains a significant problem among very low-birthweight infants. It carries with it both costs, in terms of longer hospital stay, and risks, in terms of later development. (9) Corticosteroids can reduce lung inflammation in newborns who have chronic lung disease; but there are major short- and long-term adverse effects.

For the purposes of this review, the authors included only RCTs of postnatal corticosteroid therapy. The “study participants” were defined as preterm infants believed to be at risk of developing chronic lung disease who were enrolled within the first 7 days of birth. Infants were not required to be on ventilator support, which is an important distinction. The authors could have limited the population to very low-birthweight or extreme low-birthweight infants and restricted the review only to those receiving ventilator support. The authors chose to cast a wide net for studies, but this could potentially lead to clinical heterogeneity.

The authors chose to study intravenous or oral corti-

![Figure 1. Early (<8 days) postnatal corticosteroids for preventing chronic lung disease in preterm infants. Effect on chronic lung disease or death. Reprinted with permission from Halliday et al. (8)](http://neoreviews.aappublications.org/content/12/1/10.full)
corticosteroids (including dexamethasone and hydrocortisone). Trials of inhaled corticosteroids were not included. The outcome measures included the important short-term measures related to neonatal intensive care (eg, mortality, chronic lung disease) as well as longer-term outcomes, including development at 18 to 24 months. The search strategy was specifically detailed in the review. In addition, a formal plan for data collection and analysis was proposed.

The authors identified 28 trials that qualified for inclusion in the review. This discussion is limited to the primary outcome measures of chronic lung disease or death and cerebral palsy. Figures 1 and 2 are the “forest plots” of the relative risk reported in each study and a “summary” statistic showing the “pooled” relative risk estimated by the meta-analysis. In the review of all studies (using either dexamethasone or hydrocortisone), chronic lung disease (defined as oxygen requirement at 36 weeks adjusted age) and death was significantly reduced (typical RR, 0.89; 95% CI, 0.84 to 0.95 and typical risk difference, 0.06; 95% CI, −0.09 to −0.02).

Figure 2. Early (<8 days) postnatal corticosteroids for preventing chronic lung disease in preterm infants. Effect on cerebral palsy. Reprinted with permission from Halliday et al. (8)

For the studies of hydrocortisone, little heterogeneity is noted (I² 24%). For the studies exclusively using dexamethasone, a moderate degree of heterogeneity persists (I² 49%). Close inspection of these studies shows many differences in the dexamethasone studies, including patient population, length of treatment, timing of treatment, and dosage.

Although the impact of corticosteroids on chronic lung disease seems promising, the results regarding cerebral palsy raise grave concerns. Cerebral palsy was increased with the use of corticosteroids (typical RR, 1.45; 95% CI, 1.06, 1.98 and typical risk difference, 0.03; 95% CI, 0.00, 0.06 for 12 studies and 1,452 infants). Again, there was moderate heterogeneity (I² 66%) in the overall analysis that was particularly notable in the studies that used dexamethasone (I² 34%). The meta-analysis provides a framework to examine individual studies compared with the other included studies. Of interest, Shinwell and associates (10) reported the greatest risk of cerebral palsy and yet represented some of the least exposure to dexamethasone, suggesting to investigators that perhaps timing of treatment is critical.

The understanding of this meta-analysis led to strong statements from the Committee on Fetus and Newborn of the American Academy of Pediatrics. (11) These statements recommend severe restriction in the use of corticosteroids to prevent and treat chronic lung disease, noting that “outside the context of a randomized, controlled trial, the use of corticosteroids should be limited to exceptional clinical circumstances (eg, an infant on maximal ventilator support and oxygen support).”

Conclusions

Systematic overviews and meta-analysis are critical tools in efforts to practice evidence-based medicine. Although flawed by the deficiencies and limitations of the included studies as well as by biases created by the analysis itself, meta-analysis provides a framework to inspect individual trials as well as a method to gain more precise estimates of effects by analyzing the trials in aggregate. Analyses such as the review of early postnatal corticosteroids have provided an improved perspective on the risks and benefits of treatments and have changed practice. (12)
American Board of Pediatrics Neonatal–Perinatal Medicine Content Specifications

- Understand the purpose of a systematic review.
- Understand the advantages of adding a meta-analysis to a systematic review.
- Interpret the results of a meta-analysis.
- Identify the limitations of a systematic review.
- Identify the limitations of a meta-analysis.

References


NeoReviews Quiz

1. You are reviewing the results of a randomized trial to determine the effect of a drug (treatment group) in preventing the occurrence of a specific disease (outcome) as compared with that of a placebo (control group). The results are tabulated below:

<table>
<thead>
<tr>
<th>Study Group</th>
<th>Disease Present</th>
<th>Disease Absent</th>
<th>Risk of Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment (n=100)</td>
<td>41</td>
<td>59</td>
<td>0.41</td>
</tr>
<tr>
<td>Control (n=100)</td>
<td>52</td>
<td>48</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Of the following, the number needed to treat (number of infants needed to be treated to prevent the occurrence of the disease in a single patient) in this trial is closest to:

A. 7.  
B. 9.  
C. 11.  
D. 13.  
E. 15.

2. A systematic review is designed to identify, appraise, and synthesize research-based evidence from all relevant studies. Invariably, studies brought together in the systematic review differ in some aspects, leading to heterogeneity in the clinical, methodological, and statistical domains. Of the following, methodological heterogeneity is most likely to represent variability in the:

A. Data analysis.  
B. Primary outcome.  
C. Study design.  
D. Study participants.  
E. Treatment interventions.