Is Epinephrine Efficacious in the Treatment of Bronchiolitis?

Meta-analyses are increasingly popular for synthesizing multiple trials to determine the “true” effect of a particular therapy. We use the term meta-analysis to refer to a systematic review that applies quantitative methods to combine data from multiple trials. In this issue of the ARCHIVES, Hartling et al1 use this method to examine the efficacy of epinephrine hydrochloride in the treatment of bronchiolitis.

See also page 957

Despite their recent proliferation, meta-analyses are difficult to appraise in a critical fashion. In this article, we have used several resources2-5 with minor modifications as a guide to evaluating the meta-analysis by Hartling and colleagues.

CHOOSING THE QUESTION

Is the Question Important?

Yes. Viral bronchiolitis is a common yet potentially serious illness in young children. Although it places a significant burden on the health care system, no single treatment has proved most effective. A meta-analysis showing that epinephrine is efficacious in the treatment of bronchiolitis would be an important contribution to the literature and could immediately affect patient care.

Is the Question Clearly Defined?

As much as possible. The clarity of the central question in a meta-analysis can be considered from both a methodological standpoint and a clinical or commonsense standpoint.

From a methodological standpoint, a clearly defined question should include 5 components: (1) the type of patient, (2) the type of studies included, (3) the type of exposure (ie, the treatment of interest), (4) the control(s) with which the exposure is being compared, and (5) the outcome(s) of interest.2,3 The authors state, “[T]he objective of this study was to review randomized controlled trials that compared the effects of inhaled or systemic epinephrine vs placebo or other bronchodilators in infants and young children (age, ≤2 years) with bronchiolitis.”1 This statement clearly addresses the types of patients (children ≤2 years), studies (randomized controlled trials [RCTs]), exposures (inhaled or systemic epinephrine), and controls (placebo or other bronchodilators). The outcomes of interest, however, range from clinical scores to pulmonary function tests to admission rates, reflecting the wide range of outcomes assessed in the published trials.

From a clinical standpoint, a clearly defined question allows readers to determine when a meta-analysis might apply to their clinical practice. If it is clear that the question being asked pertains (or does not pertain) to a clinical setting and an outcome that the reader cares about, that question meets the definition of clinical clarity. In this meta-analysis, the authors separate inpatient and outpatient trials, allowing readers to easily distinguish which portions are applicable to their clinical setting. However, the wide variety of outcomes included makes it more difficult for readers to determine how this study’s results compare with the clinical criteria they may use to make decisions about an individual patient’s care.

CHOOSING AND APPRAISING STUDIES

Are the Inclusion Criteria Appropriate?

Yes, for the most part. Inclusion criteria can also be considered from both methodological and clinical points of view. The term methodological inclusion criteria refers to the characteristics of trials, such as experimental design, used by investigators to identify studies appropriate for inclusion in a meta-analysis. These criteria should be the same as those used to judge the validity of an original research study. In this meta-analysis, Hartling and colleagues report, “All RCTs evaluating the efficacy of epinephrine . . . in the treatment of bronchiolitis were considered for inclusion, regardless of language or publication status.”1 This reflects the best standards for meta-analyses addressing therapeutic interventions. From a methodological standpoint, therefore, the authors’ inclusion criteria were appropriate.

Readers should also consider whether the selection criteria used to choose trials make clinical sense. Included studies must be relevant to the central question and similar enough to be grouped together to draw a single conclusion. In this meta-analysis, the authors used relatively broad clinical criteria. They included studies that used both placebos and other bronchodilators as controls, those conducted in both inpatient and outpatient
settings, and studies that used a wide range of outcome measures. Although these choices were necessary given the small number of studies, they limit the degree to which the authors’ conclusions can be applied to individual patient scenarios. The reason for the decision to consider systemic (subcutaneous) epinephrine as equivalent to the inhaled form was not as obvious, and the article may have benefited from more explicit justification.

Is It Likely That Relevant Studies Were Missed?

Probably not. The most commonly accepted strategy for identifying relevant studies for a meta-analysis includes 3 steps: (1) searching bibliographic databases, (2) hand searching the reference lists of retrieved studies, and (3) personal contact with experts in the field. Hartling and colleagues completed each of these steps, although to varying degrees. First, they searched 3 bibliographic databases: MEDLINE (maintained by the US National Library of Medicine, Bethesda, Md), EMBASE (maintained by Elsevier Science Publishers, Amsterdam, the Netherlands), and CENTRAL (the Cochrane Central Register of Controlled Trials). In this way, they took advantage of EMBASE’s extended coverage of the non-US literature and CENTRAL’s extensive hand searches of published literature for RCTs that may not have been appropriately indexed by other large databases. They also manually reviewed the reference lists of trials identified by these database searches for other potentially relevant studies. Finally, “Primary authors of relevant trials were contacted for information on additional trials.”

The language criteria used by the authors merit particular mention. Most meta-analyses include only English-language studies. It has been documented, however, that trials with statistically significant results are more likely to be published in English. The exclusion of studies in languages other than English may therefore bias the results of a meta-analysis. Hartling and colleagues are to be commended for identifying, translating, and incorporating the results of non–English-language trials.

Their efforts to identify unpublished trials are less clearly described. Like non–English-language trials, unpublished trials are less likely to report significant positive results; therefore, their exclusion may result in bias toward positive findings (commonly referred to as publication bias). It has been observed that meta-analyses “based on a small number of small studies with weakly positive results are the most susceptible to publication bias. Because this meta-analysis does base its conclusions on a small number of small studies, readers need to be satisfied that the potential for publication bias was adequately addressed by the authors’ efforts to identify unpublished trials.

Finding unpublished trials can be difficult and time consuming. Ideally, complete reporting of this process includes the number of authors contacted, their response rates, the number of unpublished studies identified, and the number for which data were obtained as well as descriptions of any other strategies used. In this study, the investigators did contact primary authors of included trials. One unpublished study was identified, presumably using this strategy, but no further details were provided.

Was the Quality of Studies Adequately Appraised and Accounted For?

Yes and no. In a study regarding therapy, the methodological quality is typically judged by whether the assignment of patients to various treatments was randomized, whether patients and providers were blinded to these assignments, and whether all patients were properly accounted for at the trial’s conclusion. In this meta-analysis, the authors made these judgments using the Jadad scale, a commonly used instrument that assigns each study a single score based on these 3 attributes. Furthermore, they assessed whether allocation to the treatment arm was adequately concealed, a characteristic that is related to the size of the treatment effect but not included in the Jadad scale. The primary reason for assessing methodological quality is that trials not meeting these standards may overestimate treatment effects, thus introducing another source of bias. Although Hartling and colleagues describe their assessments of the methodological quality of trials in considerable detail, it is not clear whether or how these assessments were used in generating or interpreting the results of this meta-analysis.

Clinical quality is based primarily on disease-specific characteristics of a trial, such as whether the treatment being used was appropriate or the outcomes being measured were reasonable. Because clinical quality is disease specific, it cannot be judged using the Jadad scale or any instrument designed to be applicable across diseases or clinical settings. Readers must rely on authors’ descriptions of selected studies, which are typically provided in Table 1. In this study, Table 1 contains a list of included trials as well as selected characteristics of each trial. The original version of the Table actually contained more information about each trial, but reviewers requested that this and other tables be streamlined prior to publication. Our preference would have been closer to the original version. We would have preferred information regarding the nature of the intervention (eg, whether single or multiple inhaled treatments were used), the severity of illness among participants, and perhaps most important, a summary of the trial’s findings. In our opinion, this information enhances the ability of readers to make their own judgments about the quality and applicability of individual trials.

These decisions are part of the natural give-and-take of the editorial process. One suggestion for future meta-analyses might be to make more complete and complex tables available to readers online.

COMBINING THE DATA

Were Decisions Regarding Whether to Combine Data, and What Data to Combine, Made Appropriately?

Yes, it appears so. In considering the results of a meta-analysis, it is easy to focus on numeric results generated by sophisticated statistical analyses that combine data from disparate studies. In this context, one may easily overlook what is perhaps the most critical question regarding the validity of a meta-analysis: how did the authors...
decision when it made sense to combine data from different studies and which data to combine?

In this study, the authors were hampered by the many differences in the design and conduct of various studies. They were particularly challenged by the lack of consistency in the measurement of outcomes across studies. They showed good clinical sense in distinguishing inpatient from outpatient studies and separating those measuring the same outcome at different points in time. As a result, however, there were often few data left to combine for any particular analysis.

Did the Authors Estimate a Single Common Effect?

No, but appropriately so. This question addresses the central activity of any meta-analysis: pooling data from different studies to calculate a single composite estimate of the effect of a given intervention. The authors chose not to pool results from studies using different clinical settings, different control groups, or outcomes measured at different points in time. Consequently, no outcome was measured using more than 4 original studies, and many outcomes were measured with only 1 or 2. The authors therefore could not estimate a single common effect, even within subgroups of studies. They opted instead to generate a list of outcomes that represented varying degrees of pooling among studies.

Was Heterogeneity Among Studies Adequately Accounted For?

Yes. When combining data from different studies, authors must account for the variability, or heterogeneity, of results among these studies. To do this, they assess the degree to which the size of treatment effect varies across studies and the likely reasons for such variation. If heterogeneity among studies is considerable, one must conclude that different studies are actually measuring different clinical phenomena. In such situations, composite estimates of clinical effect should be calculated using statistical models that best account for such heterogeneity, known as random-effects models. In contrast, fixed-effects models assume that the different studies are measuring exactly the same phenomenon and thus account only for sampling error.

In this study, Hartling and colleagues did use random-effects models in calculating estimates of effect for their many outcomes. This represents the more conservative approach; estimates calculated using random-effects models have wider confidence intervals than those calculated using fixed-effects models, to account for multiple sources of variation. With random-effects models, differences of the same magnitude are therefore less likely to be statistically significant. Given the variation in clinical score measures (as seen in the authors’ Table 3), this is an appropriate application of the random-effects model.

Were Estimates of Common Effects Robust?

It is impossible to tell. To convince readers that the point estimate of a clinical effect is believable, authors of a meta-analysis often attempt to demonstrate that this value is robust, in other words, that it is not sensitive to changes in underlying assumptions. To do this, authors conduct sensitivity analyses; that is, they repeat their statistical analyses after changing their assumptions and compare these results with those generated by the original model. Hartling and colleagues recalculated their results using fixed-effects models instead of random-effects models. According to these analyses, no significant outcomes remained significant, suggesting that their findings were not robust.

Other sensitivity analyses, however, may have been more appropriate in this meta-analysis. For example, point estimates could have been recalculated after excluding the study that may have used a fundamentally different intervention (subcutaneous epinephrine) from other included trials. Alternatively, studies lacking certain methodological characteristics, such as blinding, could have been excluded. The results from those analyses could then have been compared with findings based on data from all studies regardless of quality.

Were the Results Presented Appropriately?

Not in enough detail. The authors present their results in several ways: in tables, in graphical form, and in the text. Despite these multiple formats, the reporting of results is still incomplete. Each group of studies measured anywhere from 7 to 16 outcomes for a total of 43 outcomes among all 14 studies. Only 26 of these 43 outcomes, however, are listed in Tables 4 and 5. Of note, the Tables in the original version of this manuscript included more complete reporting of results, but this was scaled back at the request of reviewers. We would have preferred not only a complete reporting of outcomes but also detailed referencing of which studies contributed data to which pooled outcomes. When combined with the lack of detail in Table 1 regarding individual studies, these omissions limit readers’ ability to make independent judgments regarding the validity of pooled results. This in turn limits the applicability of this study’s findings to patient care.

Four figures (metagraphs) depict 2 outcomes, admission rates among outpatients and length of stay among inpatients, for each of the 2 control groups (albuterol sulfate and placebo). The reason that these 2 outcomes were chosen is not entirely clear; presumably, the authors felt that they represented the closest approximations to summary measures for each clinical setting. However, none of the 4 outcomes depicted achieved clinical significance. One would therefore conclude from these graphs that epinephrine does not confer any significant advantage compared with either albuterol or placebo in the treatment of bronchiolitis. This differs from the text’s guarded yet positive endorsement of its use among outpatients.

APPLYING THE RESULTS TO CLINICAL PRACTICE

Were All Clinically Important Outcomes Considered?

No, although this may not have been possible. In some centers, the use of inhaled epinephrine obligates physi-
Are the Benefits Worth the Harms and Costs?

This judgment cannot be drawn from the information available. As noted previously, requirements for extended observation or transfer of young children to other clinical settings incur costs that could not be accounted for in this meta-analysis. Harms include adverse events associated with epinephrine use, but no such outcomes were reported by any included trial. This may suggest that the use of epinephrine in bronchiolitis is associated with fewer adverse effects than commonly presumed. Alternatively, it may suggest that adverse events were narrowly defined in these clinical trials and that the use of epinephrine in these children is associated with adverse effects that would be important to the practicing physician but could not be identified in this particular study.

Whatever the reason, no significant costs emerged from this meta-analysis. Therefore, not enough information is provided to judge whether the marginal benefits that may accrue from the use of epinephrine in the outpatient treatment of bronchiolitis are worth the costs of its use, either in monetary or clinical terms. This highlights the importance of uniform adverse event reporting and perhaps expanded adverse effect reporting in future studies.

Can the Results Be Applied to My Patient Care?

This question has no single answer. Instead, readers must draw their own conclusions about how the results of this meta-analysis can be applied to their patients. Unfortunately, few guidelines exist to assist readers in this final and perhaps most important step in evaluating a meta-analysis. In arriving at any conclusion, readers must assimilate the answers from all of the preceding questions to make a series of subjective judgments that include the following:

- Do I believe the results?
- Do they apply to my patient population?
- Can they be applied in my current clinical setting?

Within this context, a few global statements can be made. First, this meta-analysis does not provide evidence so strongly in favor of using epinephrine in treating bronchiolitis that other treatments should be abandoned or that there should be a change in the overall standard of care. Second, there is no evidence that epinephrine is clearly inferior to either albuterol or placebo or that there is an unacceptable risk of adverse consequences associated with its use. Beyond this, the results of the meta-analysis are equivocal, and readers are left to make their own clinical judgments based on their beliefs, the nature of their patient population, and the characteristics of their practice situation. The article by Hartling and colleagues highlights the need for further studies in this area using consistent outcome measures and consistent practices to assess adverse consequences for patients and their families.

Tracy M. King, MD, MPH
Kari Alperovitz-Bichell, MD
Peter C. Rowe, MD
Harold P. Lehmann, MD, PhD
Baltimore, Md

Accepted for publication July 1, 2003.

We are grateful to the other faculty and fellows in the Department of Pediatrics at The Johns Hopkins School of Medicine and the Department of Family Medicine at the University of Maryland School of Medicine (Baltimore) who participated in the journal club session and greatly contributed to our learning process: Carmen Arroyo, PhD; Anne Duggan, ScD; Samer S. El-Kamary, MD, MPH; Duniya Lancaster, MD; Lisa Lowery, MD, MPH; Catherine Nelson, MD; Ryan Pasternak, MD; Patricia Richardson-McKenzie, MD, MPH; Sharon Richter, DO; Janet Serwint, MD; and Karen P. Zimmer, MD.

Corresponding author and reprints: Tracy M. King, MD, MPH, 600 N Wolfe St, Park 364, Baltimore, MD 21287 (e-mail: tking9@jhmi.edu).

REFERENCES