

Editorials

Dueling Meta-Analyses

Robert L. Wears, MD, MS
Department of Emergency Medicine
University of Florida Health Science Center
Jacksonville, FL

Address for reprints: Robert L. Wears, MD, MS, Department of Emergency Medicine, University of Florida Health Science Center, 655 West 8th Street, Jacksonville, FL 32209; 904-549-4124, 904-549-4508; E-mail wears@ufl.edu.

47/1/109692

doi:10.1067/mem.2000.109692

Easier Breathing?

Rita K. Cydulka, MD
Department of Emergency Medicine
MetroHealth Medical Center
Case Western Reserve University
Cleveland, OH

Reprints not available from the author.

Address for correspondence: Rita K. Cydulka, MD, Department of Emergency Medicine, MetroHealth Medical Center, Room S1-203, Cleveland, OH 44109; 216-778-5088; E-mail rcydulka@metrohealth.org.

47/1/109687

doi:10.1067/mem.2000.109687

The Prudent Layperson Definition: Will It Work for Emergency Medicine?

Robert M. Williams, MD, DrPH
Department of Health Management and Policy
University of Michigan School of Public Health
Ann Arbor, MI

Address for reprints: Robert M. Williams MD, DrPH, School of Public Health, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109-2029; 231-938-2494, fax 231-938-0364; E-mail erdoctor@traverse.com.

47/1/109212

doi:10.1067/mem.2000.109212

Emergency Departments and Uninsured Children: An Enrollment Opportunity

Susan M. Nedza, MD
Chair, Legislative and Regulatory Committee
American College of Emergency Physicians
Irving, TX

Deborah Mulligan-Smith, MD
Committee on Pediatrics
American College of Emergency Physicians
Irving, TX

Russell Harris, MD
Immediate Past-President
New Jersey Chapter
American College of Emergency Physicians
Trenton, NJ

Address for reprints: Susan M. Nedza, MD, Department of Emergency Medicine, Elmhurst Memorial Hospital, 200 Berteau, Elmhurst, IL 60126; fax 630-986-9443; E-mail s-nedza@nwu.edu.

47/1/109448

doi:10.1067/mem.2000.109448

Copyright © 2000 by the American College of Emergency Physicians.
0196-0644/2000/\$12.00 + 0

Dueling Meta-Analyses

See related articles, p. 181 and p. 191.

[Wears RL. Dueling meta-analyses. *Ann Emerg Med*. September 2000;36:234-236.]

The replication of others' results is a fundamental principle of the scientific method, yet we rarely see replication even attempted, much less reported in medical investigations. In this issue of *Annals*, the meta-analyses by Rowe et al¹ and Alter et al² on the use of magnesium in the treatment of acute bronchospasm offer a rare opportunity for independent, direct comparison of research methods and results. As is the case with most meta-analyses, the differences between studies are more interesting than the similarities.³ It is illustrative to compare these articles, particularly in light of a recent call for standardized reporting of meta-analyses.⁴ There are 3 general areas of difference: the studies selected for inclusion, the assessment of study quality, and the analysis of the results.

With similar questions, sources, and criteria, we would expect to find largely the same studies included in each meta-analysis. Both groups cast their nets widely to try to retrieve all relevant studies, both offered detail on the articles rejected, but in the end, they came up with slightly different sets of studies. Rowe et al¹ from the Cochrane group, accepted 7 articles, and Alter et al² accepted 9. One article included by Alter et al dealt with chronic obstructive pulmonary disease, so its exclusion from Rowe et al's set, which was limited to asthma only, is appropriate. The other's absence is puzzling, because it appears to be relevant and is not excluded by any of the stated exclusion criteria. It does not appear on the list of excluded studies (provided in the electronic version of the paper in the Cochrane Library), and was discovered by Alter et al's group via author consultation, not by database searching.

This illustrates 3 important points. First, it is extraordinarily difficult to be certain that all the available evidence has been obtained when evidence on a clinical question is being sought. This has important implications for advocates of evidence-based medicine. In this movement, great emphasis is placed on the teaching and practice of critical appraisal, but relatively little attention has been paid to the problem of finding the relevant evidence to begin with. Second, it emphasizes the importance of prospective registries of randomized trials as a method of obtaining information about this "gray literature." *Annals*

participates in the Controlled Trial Registry (as do many other journals), but trials only become registered by this mechanism after submission of a report. Unsubmitted data, or data initially published as an abstract but never followed by a full article, are generally unregistered. This is a serious ethical issue, since in even the most benign of trials, human subjects are placed at some potential risk (eg, they might be assigned to the study arm with the worse outcome). The justification for this risk is the potential contribution to scientific knowledge, but if the information gathered is not accessible in some form, that justification disappears. Third, automated searching of databases such as MEDLINE or EMBASE is not sufficient to ensure a comprehensive assembly of the relevant evidence. Fortunately, in this case it does not appear that the additional study changed the results, but the potential for one or more inadvertently omitted studies to do so in other settings is apparent. However, it is reassuring that a core subset of 7 studies were selected by 2 groups of analysts working independently.

Even though the 2 analyses shared many of the same studies, they did not read them in exactly the same way. For example, they report different sample sizes for one study. This might be explicable based on their handling of different allocation groups, but emphasizes the need for careful quality control in the data extraction phase of a meta-analysis. There are further differences in quality assessment of the component studies. Both groups used ostensibly the same quality instrument,⁵ but their descriptions of it sound as if 2 different methods were applied. Of the 7 studies they had in common, they disagreed on 4 of the quality scores, often by as much as 2 points on a 5-point scale. However, we may take some relief from a recent study comparing quality scoring schemes. It showed substantial disagreement between scoring instruments, disagreement so great that the results of a meta-analysis could swing from affirming an effect to denying one, based on the quality score used.⁶ In other words, the variation in the quality scores may not be important because quality scoring itself is suspect.

The problem with quality scoring is that there are many dimensions to quality in a scientific investigation, and they are relatively independent. Some of these dimensions (eg, random allocation, blind outcome assessment, intention-to-treat analysis) relate to bias in estimation, whereas others (eg, completeness of reporting, handling of ethical issues) may not. This problem of multidimensionality means that combining these dimensions into a single "quality score" will be problematic, for

2 reasons. First, the appropriate weighting of the different dimensions is unknown, and would differ depending on the purpose for which the scale is used. Second, some dimensions may be irrelevant for some purposes, so their inclusion in a quality system adds noise and obscures relevant information. The difficulties with quality scoring should not be misconstrued as evidence that quality does not matter. There is abundant evidence that poorer performance in certain aspects of clinical trials is associated with inflated (ie, biased) estimates of effect.⁷⁻¹¹ The solution is to abandon attempts to reduce quality to a single number, and instead to incorporate relevant components of quality into the meta-analysis, either through meta-regression models, or through qualitative analysis.^{2,12,13}

Although the procedural differences between the 2 meta-analyses are interesting, the ultimate question is, did they arrive at the same conclusion? It would be a bad day indeed for meta-analysis if 2 analyses using largely the same data came to diametrically opposed conclusions. A meta-analysis produces 2 fundamental results: a measure of the variability of effect among studies (termed heterogeneity) and a pooled estimate of effect magnitude. Heterogeneity is an important outcome, because if the results differ substantively among studies, there is some question about whether it is reasonable to combine their results at all.¹⁴ If heterogeneity is found, it is important that its source be investigated; in fact, this is sometimes more useful than the pooled effect estimate,^{15,16} even though this investigation is often post hoc.

Here there is some cause for concern in that Rowe et al¹ found considerable heterogeneity among their component studies, whereas Alter et al² did not, even when both groups analyzed changes in pulmonary function tests as outcome measures. Rowe et al commendably spent considerable effort investigating sources of heterogeneity, and found much of the variability could be explained by differences in disease severity. When stratified on severity (in an a priori, planned subgroup analysis), they found a favorable response to magnesium in patients with severe asthma, and no response in patients with mild disease, and further that the response was not heterogeneous within those severity groups.

Statistical tests of heterogeneity are known to suffer from low power, particularly when the number of studies to be combined is small, so a negative test statistic, such as Alter et al² obtained, should be viewed with some caution, particularly if it is close to the traditional "significance" cutoff.¹⁶⁻¹⁸ In most cases, it is reasonable to presume that clinical trials will be somewhat heterogeneous,

so failing to reject a null hypothesis of homogeneity does not necessarily provide much assurance that no heterogeneity is present. An additional factor explaining the difference in heterogeneity might be that Alter et al used standardized effect measures, whereas Rowe et al¹ performed the analysis in each study's "natural units." Standardized effect measures have long been used in the nonmedical meta-analytic world as a means of combining measurements of the same underlying phenomenon that are taken on different scales, such as peak expiratory flow rate (PEFR) and FEV₁.^{19,20} This likely reflects the social science origins of meta-analysis and social scientists' interest in scale development. By using a standardized effect, Alter et al were able to use data from all 9 studies in calculating a pooled estimate of effect, whereas Rowe et al were limited to subsets that used common measures, thus losing some statistical power. However, standardized effect measures have been seriously criticized, at least in the medical realm, because the "standard unit" used varies across studies.²¹ In some situations, this can minimize heterogeneity, while in others, it can increase it. It can even reverse the order of effects.²² For this reason, it seems reasonable to give more weight to the analysis performed in natural units, at least with respect to heterogeneity, even while reserving judgment on the general usefulness of standardized effect measures.

Finally, it is important to note that despite their differences, the general tenor of the results of the 2 meta-analyses is similar. Both agree that magnesium can be beneficial, at least to severe asthmatic patients; both estimates of effect magnitude are reasonably close; and both agree that the effect is modest, not dramatic.

Just as in meta-analysis, exploring the differences between studies can be more enlightening than examining their similarities. We are fortunate to have had the opportunity to make such a comparison. In the future, authors should be encouraged to produce independent replications of work, and editors and journals encouraged to publish them. To do otherwise would be unscientific.

1. Rowe BH, Bretzlaff JA, Bourdon C, et al. Intravenous magnesium sulfate treatment for acute asthma in the emergency department: a systematic review of the literature. *Ann Emerg Med.* 2000;36:181-190.

2. Alter HJ, Koepsell TD, Hilty WM. Intravenous magnesium as an adjuvant in acute bronchospasm: a meta-analysis. *Ann Emerg Med.* 2000;36:191-197.

3. Greenland S. Can meta-analysis be salvaged? *Am J Epidemiol.* 1994;140:783-787.

4. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet.* 1999;354:1896-1900.

5. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996;17:1-12.

6. Juni P, Witschi A, Bloch R, et al. The hazards of scoring the quality of clinical trials for meta-analysis [see comments]. *JAMA*. 1999;282:1054-1060.
7. Chalmers TC, Celano P, Sacks HS, et al. Bias in treatment allocation in controlled clinical trials. *N Engl J Med*. 1983;309:1358-1361.
8. Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. *JAMA*. 1995;273:408-412.
9. Moher D, Pham B, Jones A, et al. Does the quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet*. 1998;352:609-613.
10. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. 1999;282:1061-1066.
11. Coronary Drug Project Research Group. Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project. *N Engl J Med*. 1980;303:1038-1041.
12. Greenland S. Quality scores are useless and potentially misleading. *Am J Epidemiol*. 1994;140:300-301.
13. Berlin JA, Rennie D. Measuring the quality of trials: the quality of quality scales [editorial; comment]. *JAMA*. 1999;282:1083-1085.
14. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods [see comments]. *Am J Epidemiol*. 1994;140:290-296.
15. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ*. 1994;309:1351-1355.
16. Pocock S. Meta-analysis [editorial]. *Stat Methods Med Res*. 1993;2:117-119.
17. Boissel JP, Blanchard J, Panak E, et al. Considerations for the meta-analysis of randomized clinical trials. Summary of a panel discussion. *Control Clin Trials*. 1989;10:254-281.
18. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med*. 1987;107:224-233.
19. Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis*. New York, NY: Academic Press; 1985.
20. Cooper H, Hedges LV, eds. *Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994.
21. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev*. 1987;9:1-30.
22. Greenland S, Schlesselman JJ, Criqui MH. The fallacy of employing standardized regression coefficients and correlations as measures of effect. *Am J Epidemiol*. 1986;123:203-208.

Easier Breathing?

See related article, p. 198.

[Cydulka RK. Easier breathing? *Ann Emerg Med*. September 2000;36:236-238.]

Asthma affects 4% to 5% of the population in the United States and accounts for approximately 2 million visits to the nation's emergency departments.^{1,2} Although asthma is a chronic illness that is usually easily managed with proper ongoing, comprehensive ambulatory care, much of its economic impact is associated with asthma exacerbation, ED use, and hospitalization, which is likely a result of a failure of our medical system to provide the aforementioned ongoing care.² As a result of the increasing prevalence and impact of asthma on both the adult and pediatric population, much time and energy has been devoted to its study. A quick computerized literature search of PubMed yielded 1,311 articles related to "asthma

AND emergency treatment" since 1990.³ The proliferation of asthma-related research and literature has helped promote and publicize the understanding that asthma is largely a chronic inflammatory disorder (or group of disorders) rather than an episodic bronchospastic event and has served to formulate the guidelines that direct our current treatment of asthma.⁴⁻⁹ Despite the recognition of asthma as an inflammatory disorder and the ongoing quest to determine the optimal method of reducing inflammation in asthmatic patients, bronchodilation with β_2 -agonists remains an essential ingredient in the therapeutic armamentarium for emergency physicians, who primarily treat asthmatic patients with acute exacerbation, rather than managing chronic disease.

Many questions still remain regarding the use of β -agonists in the treatment of acute exacerbation. For example, researchers have not determined whether a single β -agonist agent is superior to the others; nor have they resolved whether isomers of β -agonists are superior to the racemic versions of the drug. Also unresolved are the problems of optimal dosing, optimal methods of delivery, timing of delivery, as well as a myriad of other issues regarding both treatment efficacy and effectiveness. The result is that treatment choices recommended by current guidelines are left somewhat open-ended.⁶⁻⁹ In 1997, the National Institutes of Health National Asthma Education and Prevention Program revised its guidelines based on several well-conducted clinical trials to suggest that asthmatic patients experiencing a severe exacerbation can be treated with inhaled high-dose β -agonists by either nebulization every 20 minutes or continuously for 1 hour.^{6,10,11}

The use of continuous albuterol nebulization was first reported in 14 children in 1987, where high-dose continuous nebulization was shown to produce a greater improvement in FEV₁ than intermittent-dose albuterol or bolus-dose albuterol followed by a lower continuous dose.¹² Although 2 early reports indicated that improvement in pulmonary function tests persisted longer in patients treated with continuous nebulization than in those treated with intermittent treatment,^{13,14} the study by Scalabrini et al¹⁵ of 21 children with acute asthma did not support these findings. Papo et al¹⁶ later found that children with severe asthma unresponsive to conventional treatment who were subsequently treated with continuous nebulization demonstrated a more rapid clinical improvement in asthma and had shorter hospital stays than those treated with intermittent β -agonist therapy.

The option of continuous nebulization of albuterol was first described in adults as a case report in 1990.¹⁷