# DISCREPANCIES BETWEEN META-ANALYSES AND SUBSEQUENT LARGE RANDOMIZED, CONTROLLED TRIALS

Jacques LeLorier, M.D., Ph.D., Geneviève Grégoire, M.D., Abdeltif Benhaddad, M.D., Julie Lapierre, M.D.,
and François Derderian, M.Sc.

## ABSTRACT

*Background*   Meta-analyses are now widely used to provide evidence to support clinical strategies. However, large randomized, controlled trials are considered the gold standard in evaluating the efficacy of clinical interventions.

*Methods*   We compared the results of large randomized, controlled trials (involving 1000 patients or more) that were published in four journals (the *New England Journal of Medicine,* the *Lancet,* the *Annals of Internal Medicine,* and the *Journal of the American Medical Association*) with the results of meta-analyses published earlier on the same topics. Regarding the principal and secondary outcomes, we judged whether the findings of the randomized trials agreed with those of the corresponding meta-analyses, and we determined whether the study results were positive (indicating that treatment improved the outcome) or negative (indicating that the outcome with treatment was the same or worse than without it) at the conventional level of statistical significance ($P < 0.05$).

*Results*   We identified 12 large randomized, controlled trials and 19 meta-analyses addressing the same questions. For a total of 40 primary and secondary outcomes, agreement between the meta-analyses and the large clinical trials was only fair (kappa = 0.35; 95 percent confidence interval, 0.06 to 0.64). The positive predictive value of the meta-analyses was 68 percent, and the negative predictive value 67 percent. However, the difference in point estimates between the randomized trials and the meta-analyses was statistically significant for only 5 of the 40 comparisons (12 percent). Furthermore, in each case of disagreement a statistically significant effect of treatment was found by one method, whereas no statistically significant effect was found by the other.

*Conclusions*   The outcomes of the 12 large randomized, controlled trials that we studied were not predicted accurately 35 percent of the time by the meta-analyses published previously on the same topics. (N Engl J Med 1997;337:536-42.)

©1997, Massachusetts Medical Society.

LARGE randomized, controlled trials are generally considered the gold standard in evaluations of the efficacy of clinical interventions. However, since such trials are not always available, clinicians increasingly rely on meta-analysis to support their choice of clinical strategies. Critics have emphasized the intrinsic weaknesses of meta-analysis.[1-5] Pooled results incorporate the biases of individual studies and embody new sources of bias, mostly because of the selection of studies and the inevitable heterogeneity among them.

Although much has been said about the strengths and weaknesses of meta-analysis, there are limited data systematically comparing the results of meta-analyses of several small trials with those of large randomized, controlled trials. Villar et al.[6] reviewed 30 meta-analyses of various interventions in perinatal medicine from the Cochrane data base. They recalculated the results of each meta-analysis after removing the largest trial from the analysis and then compared the results with those of the large trial that had been removed. They found a kappa of 0.46 to 0.53 and a positive predictive value of 50 to 67 percent. We compared the results of a series of systematically compiled large randomized, controlled trials with those of the relevant meta-analyses that had been published previously.

## METHODS

### Data Base

We searched the *New England Journal of Medicine,* the *Lancet,* the *Annals of Internal Medicine,* and the *Journal of the American Medical Association* and retrieved all large randomized, controlled trials (those in which 1000 patients or more were studied) that were published between January 1, 1991, and December 31, 1994. All the trials had to have adequate statistical power to detect the desired benefit specified by the authors. Adequate power was defined on the basis of the a priori calculations of power reported by the authors in the Methods sections of their articles.

---

We then searched for meta-analyses of similar topics that had been published before the large randomized, controlled trial. Our search included the references listed in the randomized trials and computerized searches of Medline without language restrictions. We then compared each trial with the set of meta-analyses corresponding to it and selected only those meta-analyses that coincided with the trial with regard to the similarity of the populations studied, the therapeutic intervention, and at least one outcome. We studied the principal and secondary outcomes.

For each outcome that was studied in both the large randomized, controlled trial and the meta-analysis, we determined whether the results were positive (indicating that treatment resulted in a better outcome) or negative (indicating that treatment resulted in an equal or worse outcome) at the conventional level of statistical significance (P<0.05). Two investigators working independently of each other reviewed each trial and its corresponding meta-analyses. Discrepancies were resolved by consensus, with the help of a third investigator. To quantify the effect of interobserver variation, we performed a sensitivity analysis; the statistical calculations were performed with the data obtained by consensus and were repeated with the data that corresponded to the opinion of the dissenting investigator.

### Statistical Analysis

Two-by-two tables were used to calculate the degree of agreement between the large randomized, controlled trial and its associated meta-analysis as expressed by the kappa statistic and its 95 percent confidence interval, as well as the sensitivity, specificity, positive predictive value, and negative predictive value. The point estimates in each pair were compared by using a test statistic constructed as the difference in the proportions or means divided by the square root of the sum of the variances.

The odds ratios of the randomized, controlled trial and the meta-analysis were represented graphically. When the result of the meta-analysis was not presented as an odds ratio for a dichotomous outcome, we computed the odds ratio and its 95 percent confidence interval by the fixed-effects Mantel–Haenszel method.[7] When no odds ratio could be computed for a meta-analysis that represented the size of the treatment effect, we transformed the odds ratio in the corresponding randomized, controlled trial into an effect size by treating the proportion for each group as the mean of a distribution of 0's and 1's.[8] These transformations were made only to permit graphic representation and did not affect the P values reported in the corresponding papers. Figure 1 shows the odds ratios computed by the fixed-effects method, and Figure 2 shows the effect sizes obtained by transformation of the odds ratios. P values of less than 0.05 were considered to indicate statistical significance. All the calculations and statistical tests were done with the SAS statistical package (SAS Institute, Cary, N.C.).

## RESULTS

We identified 12 large randomized, controlled trials to which 19 meta-analyses corresponded in terms of the populations studied, the therapeutic interventions, and at least one outcome. Since both the primary and the secondary outcomes were considered, a total of 40 outcomes coincided and were included in the analysis.

Table 1 shows the data on which we based our evaluation of the performance of meta-analysis as a predictor of the results of subsequent large randomized, controlled trials. The meta-analysis occupied the role usually assigned to a diagnostic test being assessed, whereas the trial was considered the gold standard. Table 2 shows the results in terms of sensit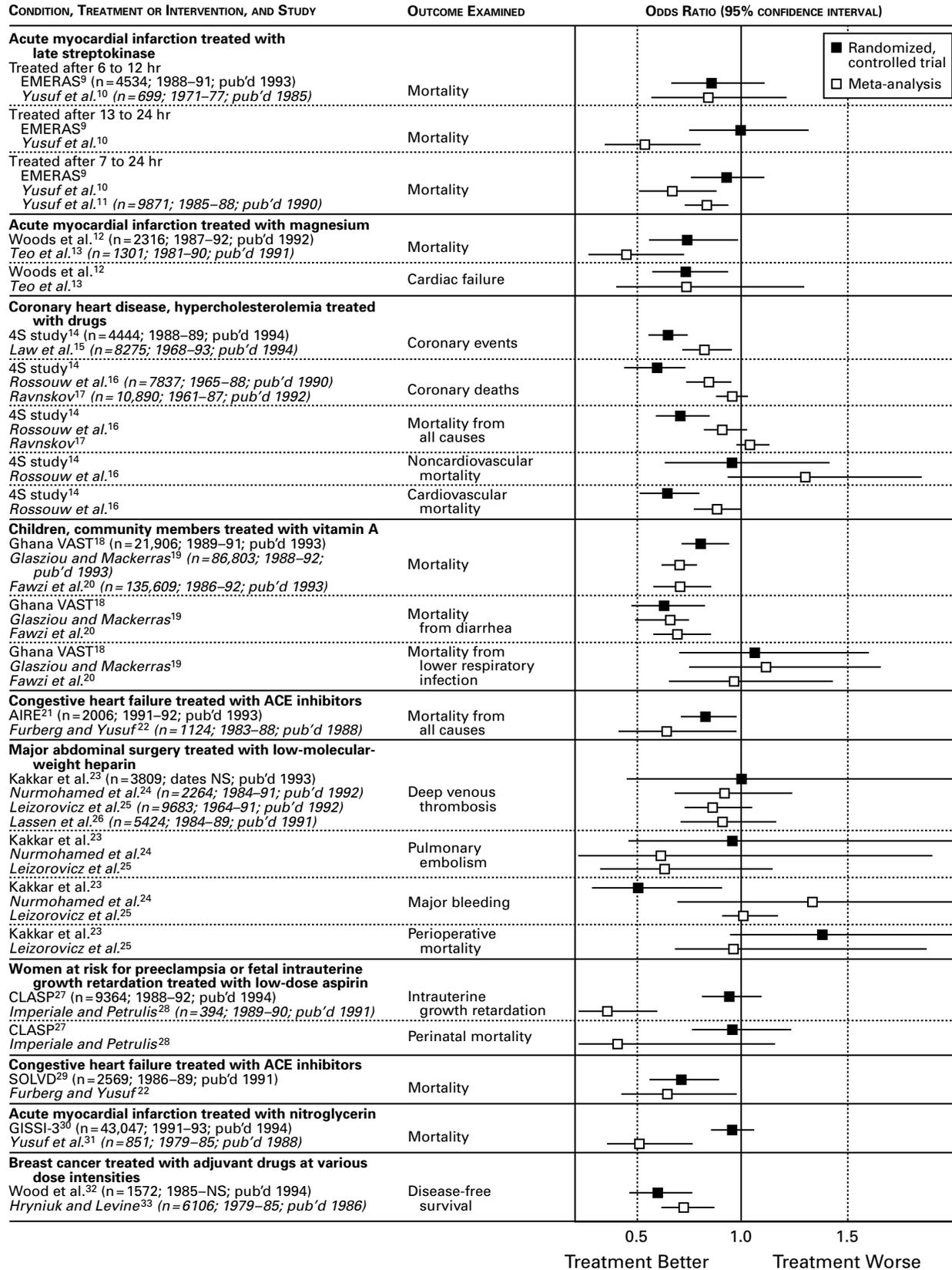ivity, specificity, and negative and positive predictive values. The results for the consensus opinion are all in a range of values (65 to 70 percent) that corresponds to the values usually obtained in average diagnostic tests. The kappa statistic, which measures agreement beyond that due to chance alone, was 0.35 (95 percent confidence interval, 0.06 to 0.64). Kappa values at or below 0.40 are considered to represent fair-to-slight agreement. Table 2 also shows the results of the sensitivity analysis, which compares the results obtained when the calculations were made on the basis of the consensus between investigators with those obtained when the calculations were based on the dissenting investigator's opinion.

Figures 1 and 2 show the results graphically and include the most pertinent information about each cluster of comparisons. They show that independently of their statistical significance, the point estimates were on the same side of 1.0 in Figure 1 and on the same side of 0 in Figure 2 in 32 of the 40 comparisons (80 percent). No situation was found in which the point estimates were both statistically significant and on opposite sides of 1.0 or 0. All the disagreements thus occurred because one result showed a statistically significant treatment effect, whereas the other indicated that such an effect was lacking. There was a statistically significant difference between the randomized clinical trial and the meta-analysis in 5 of the 40 comparisons (12 percent).

Five positive outcomes from four meta-analyses[10,28,31,37] that used fixed-effects models were followed by negative randomized clinical trials. We had the information needed to redo the statistical analyses with random-effects models for four of these outcomes,[10,28,31] and the results in all four remained statistically significant.

We found very good agreement between the meta-analysis and the randomized clinical trial with regard to the following six clinical matters: the effect of magnesium on overall mortality in patients with myocardial infarction,[12,13] the effect of treatment for hypercholesterolemia on coronary events and mortality from cardiovascular causes among patients with coronary heart disease,[14-16] the effect of vitamin A supplementation on mortality from all causes and mortality from diarrhea among children in developing countries,[18-20] the effect of angiotensin-converting–enzyme inhibitors on the mortality of patients with congestive heart failure,[21,22] the effect of adjuvant therapy on disease-free survival in patients with breast cancer,[32,33] and the value of multiple interventions as compared with single interventions in smoking cessation.[38,39]

Considerable divergence was evident in several other cases. With regard to the effects of late thrombolysis (thrombolysis performed at least six hours after the first symptoms of myocardial infarction)[9-11] and nitroglycerin on mortality in patients with myocardial infarction, the meta-analyses were positive, where-

| CONDITION, TREATMENT OR INTERVENTION, AND STUDY | OUTCOME EXAMINED | ODDS RATIO (95% CONFIDENCE INTERVAL) |
|---|---|---|

**Acute myocardial infarction treated with late streptokinase**
Treated after 6 to 12 hr
  EMERAS[9] (n=4534; 1988–91; pub'd 1993) — Mortality
  *Yusuf et al.[10] (n=699; 1971–77; pub'd 1985)*

Treated after 13 to 24 hr
  EMERAS[9] — Mortality
  *Yusuf et al.[10]*

Treated after 7 to 24 hr
  EMERAS[9] — Mortality
  *Yusuf et al.[10]*
  *Yusuf et al.[11] (n=9871; 1985–88; pub'd 1990)*

**Acute myocardial infarction treated with magnesium**
Woods et al.[12] (n=2316; 1987–92; pub'd 1992) — Mortality
*Teo et al.[13] (n=1301; 1981–90; pub'd 1991)*

Woods et al.[12] — Cardiac failure
*Teo et al.[13]*

**Coronary heart disease, hypercholesterolemia treated with drugs**
4S study[14] (n=4444; 1988–89; pub'd 1994) — Coronary events
*Law et al.[15] (n=8275; 1968–93; pub'd 1994)*

4S study[14] — Coronary deaths
*Rossouw et al.[16] (n=7837; 1965–88; pub'd 1990)*
*Ravnskov[17] (n=10,890; 1961–87; pub'd 1992)*

4S study[14] — Mortality from all causes
*Rossouw et al.[16]*
*Ravnskov[17]*

4S study[14] — Noncardiovascular mortality
*Rossouw et al.[16]*

4S study[14] — Cardiovascular mortality
*Rossouw et al.[16]*

**Children, community members treated with vitamin A**
Ghana VAST[18] (n=21,906; 1989–91; pub'd 1993) — Mortality
*Glasziou and Mackerras[19] (n=86,803; 1988–92; pub'd 1993)*
*Fawzi et al.[20] (n=135,609; 1986–92; pub'd 1993)*

Ghana VAST[18] — Mortality from diarrhea
*Glasziou and Mackerras[19]*
*Fawzi et al.[20]*

Ghana VAST[18] — Mortality from lower respiratory infection
*Glasziou and Mackerras[19]*
*Fawzi et al.[20]*

**Congestive heart failure treated with ACE inhibitors**
AIRE[21] (n=2006; 1991–92; pub'd 1993) — Mortality from all causes
*Furberg and Yusuf[22] (n=1124; 1983–88; pub'd 1988)*

**Major abdominal surgery treated with low-molecular-weight heparin**
Kakkar et al.[23] (n=3809; dates NS; pub'd 1993) — Deep venous thrombosis
*Nurmohamed et al.[24] (n=2264; 1984–91; pub'd 1992)*
*Leizorovicz et al.[25] (n=9683; 1964–91; pub'd 1992)*
*Lassen et al.[26] (n=5424; 1984–89; pub'd 1991)*

Kakkar et al.[23] — Pulmonary embolism
*Nurmohamed et al.[24]*
*Leizorovicz et al.[25]*

Kakkar et al.[23] — Major bleeding
*Nurmohamed et al.[24]*
*Leizorovicz et al.[25]*

Kakkar et al.[23] — Perioperative mortality
*Leizorovicz et al.[25]*

**Women at risk for preeclampsia or fetal intrauterine growth retardation treated with low-dose aspirin**
CLASP[27] (n=9364; 1988–92; pub'd 1994) — Intrauterine growth retardation
*Imperiale and Petrulis[28] (n=394; 1989–90; pub'd 1991)*

CLASP[27] — Perinatal mortality
*Imperiale and Petrulis[28]*

**Congestive heart failure treated with ACE inhibitors**
SOLVD[29] (n=2569; 1986–89; pub'd 1991) — Mortality
*Furberg and Yusuf[22]*

**Acute myocardial infarction treated with nitroglycerin**
GISSI-3[30] (n=43,047; 1991–93; pub'd 1994) — Mortality
*Yusuf et al.[31] (n=851; 1979–85; pub'd 1988)*

**Breast cancer treated with adjuvant drugs at various dose intensities**
Wood et al.[32] (n=1572; 1985–NS; pub'd 1994) — Disease-free survival
*Hryniuk and Levine[33] (n=6106; 1979–85; pub'd 1986)*

Legend:
■ Randomized, controlled trial
□ Meta-analysis

Axis: 0.5 — 1.0 — 1.5
Treatment Better     Treatment Worse

**Figure 1.** Odds Ratios and 95 Percent Confidence Intervals for Clusters of Studies in Which the Findings of Large Randomized, Controlled Trials Were Compared with the Results of One or More Meta-Analyses on the Same Subject, in Which at Least One Common Outcome Was Studied.

Each randomized trial and its associated meta-analyses are separated from the others by a solid horizontal line. Dashed lines delineate each cluster of trials and meta-analyses in which a single outcome was examined. The solid squares at right are the point estimates (odds ratios) for the randomized trials, and the open squares are the odds ratios for the meta-analyses. The bars on either side of the squares are 95 percent confidence intervals.

The vertical line indicating an odds ratio of 1.0 is the line at which treatment was found to have no effect. Odds ratios to the left of that line (lower than 1.0) indicate that outcome was better with treatment; those to the right (higher than 1.0) indicate that outcome was worse. When the 95 percent confidence interval does not span the "no difference" line at 1.0, the study findings are considered to be significant (P<0.05).

Names of randomized, controlled trials are given in roman type, and names of meta-analyses are given in italics. For the randomized, controlled trials, the inclusive dates listed are the years when the first and last patients were enrolled; for the meta-analyses, the dates are the years when the first and last papers were published. Pub'd denotes published, ACE angiotensin-converting enzyme, and NS not specified.

as the results of the subsequent large randomized, controlled trials were on the positive side of 1.0 but were not statistically significant. In these instances statistical power could not have been the issue, because the randomized, controlled trials included more patients than the meta-analyses. With regard to the question of preventing intrauterine growth retardation with low-dose aspirin in women at risk of preeclampsia, a clearly positive meta-analysis[28] with only 394 patients was followed by a very large randomized, controlled trial with 9364 patients that had negative results.[27] Despite a negative meta-analysis,[35] a large randomized, controlled trial[34] showed that sodium reduction decreases diastolic blood pressure, whereas in the case of calcium supplementation the reverse occurred.[34,37]

Since the decision to conduct a large randomized, controlled trial could have been made when clinicians and researchers saw a meta-analysis as inconclusive, we examined whether the meta-analysis had already been published at the time the first patient was randomized in the corresponding clinical trial. Four of the 12 trials[9,21,30,38] had evidently been started and most probably designed after the publication of the corresponding meta-analysis. Of these four trials, two[9,30] (evaluating the merits of thrombolysis and treatment with nitroglycerin) had results that diverged from those of the meta-analysis — that is, a negative randomized, controlled trial did not confirm the findings of a positive meta-analysis.

## DISCUSSION

Few will disagree with the use of the large randomized, controlled trial as the gold standard in the evaluation of the efficacy of therapeutic interventions. All the meta-analyses except one that were found by our process of systematic research had been published in major peer-reviewed journals, where they were in a position to influence clinical practice.

The strategy we used to decide whether a given meta-analysis corresponded to a specific randomized, controlled trial raises certain methodologic issues. For the studies to qualify, the population studied, the therapeutic intervention, and at least one outcome had to be similar. In some cases, such similarity could involve judgment and thus be subject to variation between observers. By having two investigators decide independently on the appropriateness of each match, we could quantify the variation and adjust for it. The sensitivity analysis (Table 2) shows that our findings were essentially the same both when the calculations were based on consensus and when they were based on the opinion of the dissenting investigator. Another methodologic issue is raised by the dichotomous classification of the results as positive or negative. The reason for choosing this approach was that the outcome of interest was whether the results of the meta-analysis should be applied to clinical practice. Clinical decisions tend to be dichotomous in that a treatment is said either to work and be recommended or not to work and not to be recommended.

According to our analysis, if there had been no subsequent randomized, controlled trial, the meta-analysis would have led to the adoption of an ineffective treatment in 32 percent of cases (100 percent minus the positive predictive value) and to the rejection of a useful treatment in 33 percent of cases (100 percent minus the negative predictive value). It is important to recognize that these measures of disagreement, which are constructed from the perspective of medical decision making, tend to overstate the degree of statistical discrepancy. This is evident from the fact that in no case was there a divergence in which the randomized clinical trial and the meta-analysis gave statistically significant and opposite answers. Furthermore, wherever the point estimates were located in relation to the "no difference" line, the difference in results between the meta-analysis and the randomized, controlled trial was statistically significant for only 5 of the 40 comparisons (12 percent); this does not appear to be a large percentage, since a divergence in 5 percent of cases would be expected on the basis of chance alone.

In our study, 46 percent of the divergences in results involved a positive meta-analysis followed by a negative randomized, controlled trial. There are sev-
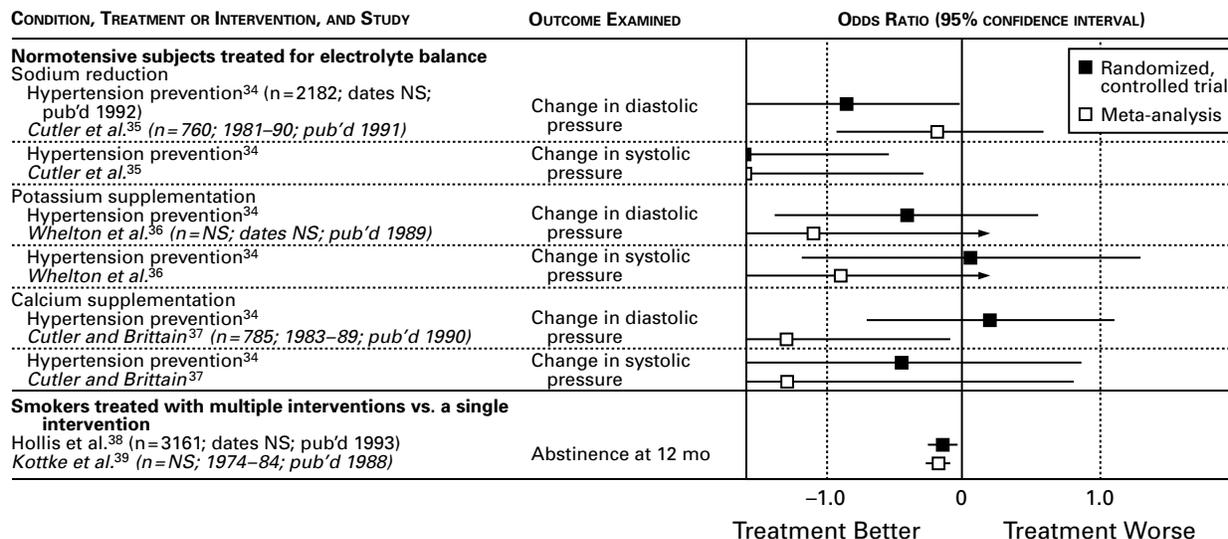
**Figure 2.** Treatment Effects and 95 Percent Confidence Intervals after Transformation of the Odds Ratios in Clusters of Studies in Which the Findings of Large Randomized, Controlled Trials Were Compared with the Results of One or More Meta-Analyses on the Same Subject, in Which at Least One Common Outcome Was Studied.

Each randomized trial and its associated meta-analysis are separated from the others by a solid horizontal line. Dashed lines delineate each cluster of trials and meta-analyses in which a single outcome was examined. The solid squares at right are the point estimates (effect sizes) for the randomized trials, and the open squares are the effect sizes for the meta-analyses (calculated as described in the Methods section). The bars on either side of the squares are 95 percent confidence intervals. Arrows mean that the differences were not statistically significant but the 95 percent confidence intervals could not be determined.

The vertical line indicating an effect size of 0 is the line at which treatment was found to have no effect. Odds ratios to the left of that line (lower than 0) indicate that outcome was better with treatment; those to the right (higher than 0) indicate that outcome was worse. When the 95 percent confidence interval does not span the "no difference" line at 0, the study findings are considered to be significant (P<0.05).

Names of randomized, controlled trials are given in roman type, and names of meta-analyses are given in italics. For the randomized, controlled trials, the inclusive dates listed are the years when the first and last patients were enrolled; for the meta-analyses, the dates are the years when the first and last papers were published. NS denotes not specified, and pub'd published.

eral reasons why a meta-analysis might have positive results that would not be confirmed by a subsequent trial. Publication bias refers to the tendency of investigators to preferentially submit studies with positive results for publication, and the tendency of editors to accept them. A meta-analysis that excluded unpublished studies or did not locate and include them would thus be more likely to have a false positive result. The systematic exclusion of papers written in languages other than English (the "Tower of Babel" bias[40]) can add to the publication bias. In our sample, the use of the fixed-effects model, which narrows the confidence interval, does not appear to account for the statistically positive meta-analyses whose findings were not subsequently confirmed by a randomized trial, since the four studies that could be reanalyzed by the random-effects model remained positive and continued to have statistically significant results when that reanalysis was done.

The remaining 54 percent of identified divergences involved a negative meta-analysis followed by a positive randomized, controlled trial. The heteroge-

neity of the trials included in the meta-analysis may partially account for divergence of this type, since meta-analysis assumes that such variation is mostly caused by random error, rather than by differences in the characteristics of the selected studies. A properly done meta-analysis involves the a priori determination of strict standards to ensure that the criteria used for the inclusion of patients, the administration of the principal treatment, and the ascertainment of outcome events are similar in all the trials selected. Although according to these strict criteria the protocols of the selected trials look very similar, their application usually yields very different products. The patients enrolled in comparable trials may belong to the same basic population, but even small differences in the criteria for diagnosis, coexisting conditions, severity of disease, and age will produce very different groups of patients. Differences in doses, time to onset, and duration of therapies can also produce substantial disparity among trials that are included in meta-analyses with the intention of evaluating a therapeutic intervention. The choice of concomitant

**TABLE 1.** AGREEMENT OR DISAGREEMENT BETWEEN RANDOMIZED, CONTROLLED TRIALS AND META-ANALYSES IN 40 CASES IN WHICH THE TWO WERE COMPATIBLE.*

| RESULTS OF META-ANALYSIS | RESULTS OF RANDOMIZED, CONTROLLED TRIAL | | |
|---|---|---|---|
| | POSITIVE | NEGATIVE | TOTAL |
| Positive | 13 | 6 | 19 |
| Negative | 7 | 14 | 21 |
| Total | 20 | 20 | 40 |

*Positive indicates that the outcome of treatment was significantly better (P<0.05) than the outcome of no treatment, and negative indicates that the outcome of treatment was worse or the same.

**TABLE 2.** VARIABLES MEASURING THE ABILITY OF META-ANALYSES TO PREDICT THE RESULTS OF LARGE RANDOMIZED, CONTROLLED TRIALS.

| VARIABLE | BASIS OF CALCULATION* | |
|---|---|---|
| | CONSENSUS OPINION | DISSENTING OPINION ONLY |
| Sensitivity — % | 65 | 60 |
| Specificity — % | 70 | 75 |
| Positive predictive value — % | 68 | 71 |
| Negative predictive value — % | 67 | 65 |
| Kappa value (95 percent confidence interval)† | 0.35 (0.06–0.64) | 0.35 (0.06–0.64) |

*The sensitivity analysis compared the results of two calculations, one based on consensus between two investigators and the other based on the opinion of the dissenting observer. The two observers disagreed 5 percent of the time.

†A positive kappa value indicates that there is more than chance agreement. A value of 0.61 or above denotes substantial agreement, a value between 0.41 and 0.60 moderate agreement, and a value of 0.40 or below fair-to-slight agreement.

therapies and the degree of leeway in their administration can also affect the results. Changes in medical practice over time may also account for important differences in concomitant therapies, since the trials included in a given meta-analysis are often conducted over a period of a decade or more.

How should clinicians use meta-analyses, given that systematic comparison with randomized clinical trials shows that they have poor predictive ability? Most will agree that if a large, well-done randomized trial has been conducted, practice guidelines should be strongly influenced by its results. The question arises when the only available evidence is from a series of small randomized, controlled trials. The simplest solution, and currently the most popular one, has been to rely on the results of a meta-analysis.

Our findings seem to indicate that summarizing all the information contained in a set of trials into a single odds ratio may greatly oversimplify an extremely complex issue. The popularity of meta-analysis may at least partly come from the fact that it makes life simpler and easier for reviewers as well as readers. However, oversimplification may lead to inappropriate conclusions.

The result of this study would appear to encourage readers to go beyond the point estimates and confidence intervals that represent the aggregate findings of a meta-analysis and, as Cook et al. have suggested,[41] look carefully at the studies that were included and evaluate the consistency of their results. When the results are mostly on the same side of the no-difference line, the meta-analysis merits more confidence. Others may consider following the advice of Horwitz[42] and appraising each trial separately. Although such an approach is admittedly more laborious, it has the advantage of allowing pragmatic clinicians to benefit from the diversity of studies by distinguishing the effects of treatment among them.

## REFERENCES

**1.** Feinstein AR. Meta-analysis: statistical alchemy for the 21st century. J Clin Epidemiol 1995;48:71-9.

**2.** Horwitz RI. "Large-scale randomized evidence: large, simple trials and overviews of trials": discussion: a clinician's perspective on meta-analyses. J Clin Epidemiol 1995;48:41-4.

**3.** Bailar JC III. The practice of meta-analysis. J Clin Epidemiol 1995;48: 149-57.

**4.** Meinert CL. Meta-analysis: science or religion? Control Clin Trials 1989;10:Suppl:257S-263S.

**5.** Abrahamson JH. Meta-analysis: a review of pros and cons. Public Health Rev 1990/91;18:1-47.

**6.** Villar J, Carroli G, Belizán JM. Predictive ability of meta-analyses of randomised controlled trials. Lancet 1995;345:772-6.

**7.** Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22:719-48.

**8.** McNemar Q. Psychological statistics. 3rd ed. New York: John Wiley, 1962.

**9.** EMERAS (Estudio Multicéntrico Estreptoquinasa Repúblicas de América del Sur) Collaborative Group. Randomised trial of late thrombolysis in patients with suspected acute myocardial infarction. Lancet 1993; 342:767-72.

**10.** Yusuf S, Collins R, Peto R, et al. Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials. Eur Heart J 1985;6:556-85.

**11.** Yusuf S, Sleight P, Held P, McMahon S. Routine medical management of acute myocardial infarction: lessons from overviews of recent randomized controlled trials. Circulation 1990;82:Suppl II:II-117–II-134.

**12.** Woods KL, Fletcher S, Roffe C, Haider Y. Intravenous magnesium sulphate in suspected acute myocardial infarction: results of the second Leicester Intravenous Magnesium Intervention Trial (LIMIT-2). Lancet 1992;339:1553-8.

**13.** Teo KK, Yusuf S, Collins R, Held PH, Peto R. Effects of intravenous magnesium in suspected acute myocardial infarction: overview of randomised trials. BMJ 1991;303:1499-503.

**14.** Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet 1994;344:1383-9.

**15.** Law MR, Wald NJ, Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? BMJ 1994;308:367-72.

**16.** Rossouw JE, Lewis B, Rifkind BM. The value of lowering cholesterol after myocardial infarction. N Engl J Med 1990;323:1112-9.

**17.** Ravnskov U. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. BMJ 1992;305:15-9. [Erratum, BMJ 1992;305:505.]

**18.** Ghana VAST Study Team. Vitamin A supplementation in northern Ghana: effects on clinic attendances, hospital admissions, and child mortality. Lancet 1993;342:7-12. [Erratum, Lancet 1993;342:250.]

**19.** Glasziou PP, Mackerras DE. Vitamin A supplementation in infectious diseases: a meta-analysis. BMJ 1993;306:366-70.

**20.** Fawzi WW, Chalmers TC, Herrera MG, Mosteller F. Vitamin A supplementation and child mortality: a meta-analysis. JAMA 1993;269:898-903.

**21.** The Acute Infarction Ramipril Efficacy (AIRE) Study Investigators. Effect of ramipril on mortality and morbidity of survivors of acute myocardial infarction with clinical evidence of heart failure. Lancet 1993;342:821-8.

**22.** Furberg CD, Yusuf S. Effect of drug therapy on survival in chronic congestive heart failure. Am J Cardiol 1988;62:41A-45A.

**23.** Kakkar VV, Cohen AT, Edmonson RA, et al. Low molecular weight versus standard heparin for prevention of venous thromboembolism after major abdominal surgery: the Thromboprophylaxis Collaborative Group. Lancet 1993;341:259-65.

**24.** Nurmohamed MT, Rosendaal FR, Buller HR, et al. Low-molecular-weight heparin versus standard heparin in general and orthopaedic surgery: a meta-analysis. Lancet 1992;340:152-6.

**25.** Leizorovicz A, Haugh MC, Chapuis FR, Samama MM, Boissel JP. Low molecular weight heparin in prevention of perioperative thrombosis. BMJ 1992;305:913-20.

**26.** Lassen MR, Borris LC, Christiansen HM, et al. Clinical trials with low molecular weight heparins in the prevention of postoperative thromboembolic complications: a meta-analysis. Semin Thromb Hemost 1991;17:Suppl 3:284-90.

**27.** CLASP (Collaborative Low-dose Aspirin Study in Pregnancy) Collaborative Group. CLASP: a randomised trial of low-dose aspirin for the prevention and treatment of pre-eclampsia among 9364 pregnant women. Lancet 1994;343:619-29.

**28.** Imperiale TF, Petrulis AS. A meta-analysis of low-dose aspirin for the prevention of pregnancy-induced hypertensive disease. JAMA 1991;266:260-4.

**29.** The SOLVD Investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. N Engl J Med 1991;325:293-302.

**30.** Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico. GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction. Lancet 1994;343:1115-22.

**31.** Yusuf S, Collins R, MacMahon S, Peto R. Effect of intravenous nitrates on mortality in acute myocardial infarction: an overview of the randomised trials. Lancet 1988;1:1088-92.

**32.** Wood WC, Budman DR, Korzun AH, et al. Dose and dose intensity of adjuvant chemotherapy for stage II, node-positive breast carcinoma. N Engl J Med 1994;330:1253-9. [Erratum, N Engl J Med 1994;331:139.]

**33.** Hryniuk W, Levine MN. Analysis of dose intensity for adjuvant chemotherapy trials in stage II breast cancer. J Clin Oncol 1986;4:1162-70.

**34.** The Trials of Hypertension Prevention Collaborative Research Group. The effects of nonpharmacologic interventions on blood pressure of persons with high normal levels: results of the Trials of Hypertension Prevention, phase I. JAMA 1992;267:1213-20. [Erratum, JAMA 1992;267:2330.]

**35.** Cutler JA, Follmann D, Elliott P, Suh I. An overview of randomized trials of sodium reduction and blood pressure. Hypertension 1991;17:Suppl:I27-I33.

**36.** Whelton PK, Thaker GK, Klag MJ, Seidler AJ. Blood pressure effects of potassium supplementation. Circulation 1989;80:Suppl II:II-301. abstract.

**37.** Cutler JA, Brittain E. Calcium and blood pressure: an epidemiologic perspective. Am J Hypertens 1990;3:137S-146S.

**38.** Hollis JF, Lichtenstein E, Vogt TM, Stevens JV, Biglan A. Nurse-assisted counseling for smokers in primary care. Ann Intern Med 1993;118:521-5.

**39.** Kottke TE, Battista RN, DeFriese GH, Brekke ML. Attributes of successful smoking cessation interventions in medical practice: a meta-analysis of 39 controlled trials. JAMA 1988;259:2883-9.

**40.** Grégoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? J Clin Epidemiol 1995;48:159-63.

**41.** Cook DJ, Guyatt GH, Laupacis A, Sackett DL, Goldberg RJ. Clinical recommendations using levels of evidence for antithrombotic agents. Chest 1995;108:Suppl:227S-230S.

**42.** Horwitz RI. Complexity and contradiction in clinical trial research. Am J Med 1987;82:498-510.