STATISTICS IN MEDICINE, VOL. 6, 733–744 (1987)

# META-ANALYSIS OF CLINICAL TRIALS AS A SCIENTIFIC DISCIPLINE. II: REPLICATE VARIABILITY AND COMPARISON OF STUDIES THAT AGREE AND DISAGREE

THOMAS C. CHALMERS, JAYNE BERRIER, HENRY S. SACKS, HOWARD LEVIN, DINAH REITMAN, AND RAGURAMAN NAGALINGAM

*Clinical Trials Unit, Mount Sinai School of Medicine of CUNY, 1 Gustave L. Levy Place, New York, NY 10029, U.S.A., and Department of Health Policy and Management, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, U.S.A.*

## SUMMARY

The replicate variability of meta-analyses of controlled clinical trials has been assessed as a measure of scientific precision. 46 of 91 known meta-analysis papers were divided into 20 cohorts of studies of the same therapies. Ten cohorts contained meta-analyses with different statistical conclusions; 14 contained differing clinical conclusions with a wider spread than the statistically differing studies. Possible causes of variability, such as different trials included, different policies regarding the inclusion of non-randomized and unpublished trials, and different statistical methodologies, were not obvious causes of differing conclusions. Further work in this area should include multivariate analyses in order to explore possible interactions in the factors accounting for the variability found in replicate meta-analyses.

KEY WORDS Meta-analysis Clinical trials Replicate variability Randomized control trials

## INTRODUCTION

If meta-analysis is to assume an important place in the transmission of clinical research data from originators to other investigators and practitioners, one must carry it out with as much attention to scientific rigor as possible. It is retrospective research, and as such requires meticulous attention to the possibility that bias may influence the results. In two previous papers we have surveyed all of the meta-analyses of controlled clinical trials we could find in the English language, and have compared the results of meta-analyses of multiple small studies with those of large co-operative studies. We found that the majority of published trials lacked evidence of efforts to control bias, and presented insufficient details of methods employed to engender confidence in their results.[1] In two of the three situations in which we could compare large scale trials with multiple small ones the results were almost identical, but a difference in the third suggests the possible influence of publication bias.[2]

In this paper we take a different approach by comparing the replicate variability when more than one meta-analysis of the same treatment has been carried out. We have found modest differences in the statistical significance of results and greater differences in the interpretations by the authors.

0277–6715/87/070733–12$06.00

*Received October 1986*
*Revised April 1987*

## METHODS

We have found 86 meta-analyses published in the English language by MEDLINE searches, reviews of 'Current Contents', and obtaining references from published review articles and other meta-analyses. The technique of meta-analysis is new enough not to be routinely indexed under the terms 'meta-analysis',[3] 'overviews',[4] or 'pooling'[5] as it is referred to by some groups. Whenever the title or abstract of a review article suggested that the authors might have combined published data, we reviewed the paper itself. We also included unpublished papers from our unit, making a total of 91 papers. The criteria for acceptance of papers were that they concern controlled clinical trials of therapeutic or diagnostic modalities, with the requirement that at least some of the trials employed random assignment, and that the data from the individual trials were combined to give an overall quantitative conclusion. Of these 91 papers, 46 included replicate analyses of 20 different treatments. Because some of the 46 papers analysed more than one treatment, we divided a total of 57 meta-analyses into 20 treatment groups (referred to as 'cohorts' in this study) with from two to six meta-analyses per cohort. (See Appendices I and II for a list of the 57 meta-analyses.)

Two observers independently reviewed the papers in random order, with the methods section separated from the results and without matching the replicate meta-analyses. Observers ironed out their differences in conference. We had differentially photocopied the individual papers making up a cohort of replicate analyses so that the assessor had no clue as to the origin of the paper or the results when he/she reviewed the methods section. Each observer classified the results on two scales (one statistical and one clinical) according to level of agreement. The statistical scale was:

1. experimental therapy significantly better ($P < 0{\cdot}05$);
2. trend in favour of experimental therapy ($P > 0{\cdot}05$);
3. no apparent statistical effect;
4. trend favouring control group ($P > 0{\cdot}05$);
5. control group significantly better ($P < 0{\cdot}05$).

The clinical scale concerned the enthusiasm of authors:

1. strongly favouring experimental therapy;
2. moderately favouring experimental therapy;
3. no difference of clinical interest;
4. moderately favouring control;
5. strongly favouring control.

We then divided all 20 cohorts into two groups according to the above statistical scale: (1) an agreement group in which all meta-analyses in each cohort agreed, and (2) a disagreement group in which at least one meta-analysis in a cohort disagreed with the others in that cohort. Similarly, we divided the cohorts into two additional groups according to clinical agreement or disagreement. This was a 'first sort' of agree/disagree for descriptive purposes only.

In an attempt to discover the factors that contributed to the differing statistical conclusions, we performed a 'second sort' of the statistically agreeing and disagreeing cohorts for comparison purposes. We re-classified as 'agreeing' all those meta-analyses in the original statistical agreement group as well as those with the majority viewpoint in a cohort with three or more meta-analyses from the original statistical disagreement group. Thus, the re-classified 'disagree' category consisted of all those meta-analyses that remained in the original statistical disagreement group after removal of the majority viewpoint analyses. That is, in statistically disagreeing cohorts of two meta-analyses, we classified both meta-analyses as 'disagree' and in statistically disagreeing cohorts of three or more meta-analyses, we classified the divergent ones as 'disagree' and re-classified those

Table I. Meta-analyses of clinical trials divided into single treatment cohorts and separated as to whether or not any one member of a cohort comes to a different statistical or clinical conclusion from the others

| | |
|---|---|
| Total papers found in English language literature + 5 unpublished meta-analyses | 91 |
| Total papers in cohort study | 46 |
| Total meta-analyses in cohort study | 57 |
| Total cohorts | 20 |
| agreeing cohorts (statistical) | 10 |
| disagreeing cohorts (statistical) | 10 |
| agreeing cohorts (clinical) | 6 |
| disagreeing cohorts (clinical) | 14 |

with the majority viewpoint as 'agree'. The idea was to accept agreement by two or more as approximating the truth so that we could compare those with single ones that did not agree with the majority.

To examine the question of whether or not inclusion or exclusion of specific trials in a meta-analysis might influence the result of that meta-analysis, we listed the trials included in a particular meta-analysis as a percentage of all available trials in that particular cohort. Then we divided the meta-analyses in each cohort into two groups according to whether the authors used both published and unpublished trials or published trials only. We calculated percentages for the 'published and unpublished trials' group by dividing the number of trials analysed in each meta-analysis by the number of all trials analysed in that cohort, corrected for dates of publication. We calculated similar percentages for the 'published trials only' group using the total number of published trials in a given cohort as the denominator.

## RESULTS

Classification of the meta-analyses with regard to agreement and disagreement for descriptive purposes appears in Table I. Within the total of 57 meta-analyses, there were 20 single treatment cohorts. In ten of the 20 cohorts, all meta-analyses agreed statistically. Within the remaining ten cohorts, at least one meta-analysis disagreed statistically. The clinical assessment was somewhat different; only six cohorts contained meta-analyses with clinical agreement, while 14 had some disagreement.

Table II(a) shows the levels of statistical agreement by cohort and numbers of meta-analyses. We found the therapeutic effect significantly positive in eight of the ten cohorts.

When there was statistical disagreement (Table II(b)), it was almost always between two adjacent levels, for example, $P < 0{\cdot}05$ or $P > 0{\cdot}05$, in the same direction. In the case of psychotherapy, four meta-analyses showed statistically significant treatment effects, and one found no effect. In the case of anticoagulants to prevent recurrent myocardial infarction, one meta-analysis found no effect while two found statistically significant treatment effects.

The data on clinical agreement appear separately for those that had statistical agreement and disagreement. Of the ten cohorts that had statistical agreement, six also had clinical agreement (Table III(a)), but the remaining four had clinical disagreement (Table III(b)). Five of the clinically agreeing cohorts were strongly positive and one had no opinion. The clinical disagreements (Table III(b)) were largely differences of one level, except for stimulant treatment of hyperactivity,

Table II. Meta-analyses that agree or disagree statistically, separated by treatment cohort and five levels of statistical conclusions

| Subject | Total | Favours treatment $P < 0.05$ | Favours treatment $P > 0.05$ | No effect | Favours control $P > 0.05$ | Favours control $P < 0.05$ |
|---|---|---|---|---|---|---|
| (a) Agree | | | | | | |
| Beta-blockers, post-MI | 6 | 6 | | | | |
| Beta-blockers, acute MI | 4 | | 4 | | | |
| IV streptokinase for acute MI | 3 | 3 | | | | |
| Aspirin, post-MI | 3 | 3 | | | | |
| Psychoeducational intervention | 2 | 2 | | | | |
| Patient education | 2 | 2 | | | | |
| Nicotine chewing gum | | | | | | |
| (a) clinics | 2 | 2 | | | | |
| (b) practice | 2 | | 2 | | | |
| Prevention of venous thrombosis | 2 | 2 | | | | |
| Stimulant therapy of hyperactivity | 2 | 2 | | | | |
| (b) Disagree | | | | | | |
| Psychotherapy | 5 | 4 | | 1 | | |
| Steroids in alcoholic hepatitis | 3 | 1 | 2 | | | |
| Anti-depressant drugs | 3 | 2 | 1 | | | |
| Lidocaine, acute MI | 3 | 2* | | | 1* | |
| Single dose versus conventional for UTI | | | | | | |
| (a) TMP-SMZ | 3 | | | 1 | 1 | 1 |
| (b) Amoxicillin | 3 | | | | 1 | 2 |
| Anticoagulants, post-MI | 3 | 2 | | 1 | | |
| Radiotherapy after radical mastectomy | 2 | | | | 1 | 1 |
| Association of steroids and peptic ulcer | 2 | 1 | 1 | | | |
| Diuretics in pregnancy | 2 | (1* | 1*) | 1 | | |

* Different endpoints

where one meta-analysis strongly favoured treatment and another moderately favoured the control. Both authors found statistical evidence of a benefit, but one felt that the drug's side effects clinically outweighed the benefits. All of the ten cohorts with statistical disagreement also had clinical disagreement (Table III(c)).

When compared according to our 'second sort' of statistical agreement/disagreement (see methods section), the question of whether or not the authors of the individual meta-analyses included trials with random assignment of treatment and control groups seemed to have no influence on their agreement status. There were some differences in statistical methodology, however, which were of interest but not significant (Table IV). The Mantel–Haenszel method[6] and effect size[3] were the most common methods used in the meta-analyses that fell in the agreement group, while crude pooling was the most commonly used method in the disagreement group.

Inclusion or exclusion of available trials by the various meta-analysis authors is a possible explanation for differences in conclusions. The percentage of trials included in a given meta-analysis from those available in that cohort appear in Table V.

We found significant differences with regard to publication policy or agree/disagree status. Notably, however, nine of 33 meta-analyses in the published-only category included less than 50 per cent of the trials available to the authors.

We examined the question of inclusion/exclusion criteria more closely by comparing the meta-analyses of beta-blockers[4] and streptokinase[7,8] in the treatment of acute myocardial infarction (in

Table III. Meta-analyses that agree or disagree clinically as well as statistically, separated by treatment cohort and five levels of clinical conclusions

| Subject | Total | Favours treatment strong | moderate | No opinion | Favours control moderate | strong |
|---|---|---|---|---|---|---|
| (a) Agree clinically and statistically | | | | | | |
| IV streptokinase for AMI | 3 | 3 | | | | |
| Psychoeducational interventions | 2 | 2 | | | | |
| Patient education | 2 | 2 | | | | |
| Nicotine chewing gum | | | | | | |
| (a) clinics | 2 | 2 | | | | |
| (b) practice | 2 | | | 2 | | |
| Prevention of venous thrombosis | 2 | 2 | | | | |
| (b) Disagree clinically, agree statistically | | | | | | |
| Beta-blockers, post-MI | 6 | 4 | 2 | | | |
| Beta-blockers, acute MI | 4 | | 3 | 1 | | |
| Aspirin post-MI | 3 | 1 | 2 | | | |
| Stimulant drugs for hyperactivity (efficacy + side effects) | 2 | 1 | | | 1 | |
| (c) Disagree clinically and statistically | | | | | | |
| Psychotherapy | 5 | 3 | 1 | | 1 | |
| Single dose versus conventional for UTI | | | | | | |
| (a) TMP-SMZ | 3 | | 2 | | | 1 |
| (b) Amoxicillin | 3 | | 1 | | 1 | 1 |
| Steroids in alcoholic hepatitis | 3 | 1 | 2 | | | |
| Anti-depressant drugs | 3 | | 2 | 1 | | |
| Lidocaine, acute MI | 3 | 2* | | 1* | | |
| Anticoagulants, post-MI | 3 | | 2 | 1 | | |
| Radiotherapy after radical mastectomy | 2 | 1* | | | | 1* |
| Association of steroids and peptic ulcer | 2 | 1 | | | 1 | |
| Diuretics in pregnancy | | | 1* | | | 1* |

* Different end points

which the authors included all available trials, regardless of their publication or language status) with the meta-analyses we performed during this research project[2] (in which we included only English language trials reported in full length manuscripts, usually in peer-reviewed journals).

Table VI shows that we excluded a large number of studies of beta-blockers that Yusuf *et al.*[4] included. Figure 1, however, shows little difference in the results of the meta-analyses.

The situation is similar in the case of intravenous streptokinase. We are aware of 24 trials on this question. The first meta-analysis, by Stampfer *et al.*[8] included data from eight published trials. The update by Yusuf *et al.* found twelve additional trials (mostly unpublished or non-English language). Our meta-analysis[2] included eleven published trials, seven of which both Stampfer *et al.* and Yusuf *et al.* used. Although we excluded all unpublished and non-English language trials, the meta-analyses again gave similar results (Figure 2).

A close look at the overlaps shows that most of the data come from a small number of fairly large trials, included in all three meta-analyses, that is, the inclusion of unpublished studies added little data. Obviously, the possibility of a type II error is very large, and this question needs examination in many more instances.

Table IV. Statistical methods used by agreeing and disagreeing meta-analyses

| | Total meta-analyses | |
|---|---|---|
| | Agree | Disagree |
| Mantel–Haenszel | 17 | 5 |
| Effect size | 11 | 1 |
| Crude pooling | 7 | 8 |
| Analysis of variance | 1 | 1 |
| Log linear | 2 | 0 |
| Fisher exact | 1 | 0 |
| Unknown | 3 | 1 |
| Totals | 43* | 16* |

*The total meta-analyses add up to more than 57 because two of the studies used two statistical methods

Table V. Percentage of trials used by members of single treatment cohorts of meta-analyses listed according to policy with regard to inclusion or exclusion of unpublished trials and classified according to statistical agreement or disagreement*

| Percentage available trials | Published and unpublished | | Published only | |
|---|---|---|---|---|
| | Agree | Disagree | Agree | Disagree |
| 100 | 1 | 0 | 5 | 4 |
| 50–99 | 9 | 1 | 10 | 5 |
| 1–49 | 1 | 0 | 6 | 3 |
| Totals | 11 | 1 | 21 | 12 |

*Data were not available for three cohorts including ten meta-analyses (eight agrees and two disagrees)

Table VI. Number of randomized control trials of beta-blockers in treatment of acute myocardial infarction included and excluded in meta-analyses by two different investigators

| Chalmers *et al.* (1987)[2] | Yusuf *et al.* (1985)[4] | | |
|---|---|---|---|
| | Included | Excluded | Totals |
| Included | 24 | 1 | 25 |
| Excluded | 22 | — | 22 |
| Totals | 46 | 1 | 47 |

## DISCUSSION

Our investigation reveals differences in published meta-analyses of the same therapeutic modalities: these differences, however, are almost always of degree rather than direction. We find greater disagreement in authors' interpretations with regard to clinical applications of the therapy

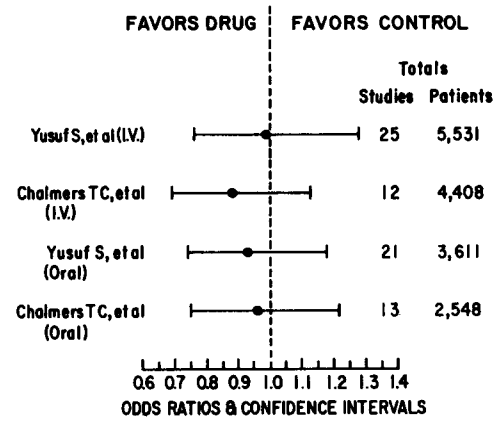


Figure 1. Comparison of two meta-analytic techniques
(1) Restricted to English language published manuscripts (Chalmers *et al.*) and (2) Inclusion of abstracts, non-English papers, unpublished papers and additional data (Yusuf *et al.*), in evaluation of beta-blockers for acute myocardial infarction. Two large co-operative studies, 'MIAMI'[17] and 'ISIS-I'[18] are omitted

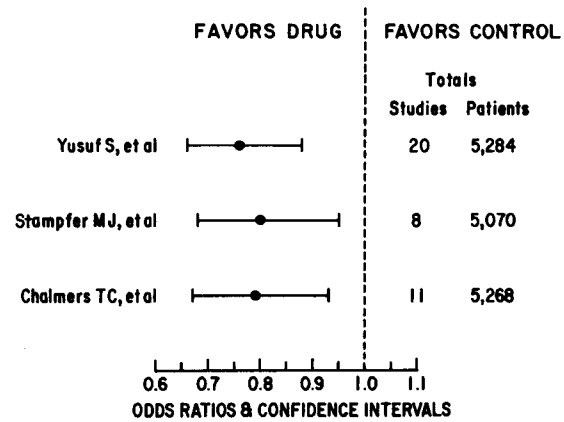


Figure 2. Comparison of two meta-analytic techniques
(1) Restricted to English language published manuscripts (Chalmers *et al.*) and (2) Inclusion of abstracts, non-English papers, unpublished papers and additional data (Yusuf *et al.*; Stampfer *et al.*), in evaluation of I.V. streptokinase for acute myocardial infarction. One large co-operative study 'GISSI'[19] is omitted

than we find in the reported statistical analyses, but half of the latter did show differences in one or more members of a cohort with regard to whether or not the observed effects were statistically significant. As one might expect, the clinical interpretations had a broader spread, but again, there were no drastic differences with regard to the applicability of the therapy. In two of the ten instances of statistical disagreement, the explanation most probably lies in the fact that authors chose different primary endpoints for their meta-analyses. In the case of lidocaine for acute myocardial infarction, two positive meta-analyses concerned only arrhythmias[9,10] and one negative meta-analysis concerned only survival.[11] In the case of diuretics and pregnancy, the clinical differences also apparently resulted from the choice of different endpoints.[12,13]

The two meta-analyses of radiotherapy after radical mastectomy merit special comment.[14,15] Authors of the second paper[15] re-analysed the data from five published randomized control trials in an unblinded manner, changed the statistical conclusion of the previous meta-analysis[14] from significantly deleterious to non-significant, and came to a drastically opposite clinical conclusion with emphasis on prevention of local recurrences, rather than the deleterious effects on survival.

Although this paper does not settle the question of whether meta-analyses of clinical trials as now performed have sufficient scientific rigor to reveal reproducible facts, the process must continue in the future; hopefully, disagreements will disappear as meta-analyses methodology becomes more rigorous.[1] The extent of agreement is encouraging, and, taken with the apparent lack of disagreement between results of meta-analyses of small trials compared with large, co-operative studies,[2] suggests that one should not discourage, on the basis of their anticipated size alone, well designed and conducted small trials. The small trial with assignment of patients at random should be judged against the common standards of the uncontrolled practice of innovative medicine, not the rare very large trial.

Since the start of this research another assessment of the replicability of meta-analyses has appeared.[16] The National Institute of Education commissioned six experts to conduct independent meta-analyses of the best 19 of 157 studies of the impact of desegregation on the academic achievement of blacks. These experts also found approximately the same consistency in statistical demonstration of a positive effect and a definite discrepancy in the interpretation of the meaning of that effect.

We recognize that our present method of seeking overall explanations for the individual differences of members of cohorts of meta-analyses lacks sensitivity. Further work should include multivariate analyses to look for additional factors and possible interactions between the factors, but analysis of many more papers will be necessary. The contribution of this method of assessment in establishing a place for meta-analysis will require publication of more RCT's and meta-analyses of common therapies, but that time will come.

## APPENDIX I: AGREEING COHORTS

### I: Beta-Blockers, Post-Myocardial Infarction (MI)

1. Baber, N. S. and Lewis, J. A. 'Confidence in results of beta-blocker postinfarction trials', *British Medical Journal*, **284**, 1749–1750 (1982).
2. Bassan, M. M., Shalev, O. and Eliakim, A. 'Improved prognosis during long-term treatment with beta-blockers after myocardial infarction: Analysis of randomized trials and pooling of results', *Heart and Lung*, **13**, 164–168 (1984).
3. Furberg, C. D. and Bell, R. L. 'Effect of beta-blocker therapy on recurrent nonfatal myocardial infarction', *Circulation*, **67** (Suppl. 1), 183–185 (1983).
4. Furberg, C. D. and May, G. S. 'Effect of long-term prophylactic treatment on survival after myocardial infarction', *American Journal of Medicine*, **76**, 76–83 (1984).
5. 'Long-term and short-term beta-blockade after myocardial infarction', *Lancet*, **1**, 1159–1161 (1982).
6. Yusuf, S., Peto, R., Lewis, J., Collins, R. and Sleight, P. 'Beta blockade during and after myocardial infarction: An overview of the randomized trials', *Progress in Cardiovascular Diseases*, **27**, 335–371 (1985).

**II: Beta-Blockers, Acute MI**

1. Baber, N. S. *et al.* (1982, see above).
2. Chalmers, T. C., Levin, H., Sacks, H. S., Reitman, D., Berrier, J. and Nagalingam, R. 'Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials', *Statistics in Medicine*, **6**, 315–325 (1987).
3. 'Long-term and short-term beta-blockade after myocardial infarction' (1982, see above).
4. Yusuf, S. *et al.* (1985, see above).

**III: I. V. Streptokinase for Acute MI**

1. Stampfer, M. J., Goldhaber, S. Z., Yusuf, S., Peto, R. and Hennekens, C. H. 'Effect of intravenous streptokinase on acute myocardial infarction. Pooled results from randomized trials', *New England Journal of Medicine*, **307**, 1180–1182 (1982).
2. Chalmers, T. C., *et al.* (1987, see above).
3. Yusuf, S., Collins, R., Peto, R., Furberg, C., Stampfer, M. J., Goldhaber, S. Z. and Hennekens, C. H. 'Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: Overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials', *European Heart Journal*, **6**, 556–585 (1985).

**IV: Aspirin Post-MI**

1. 'Aspirin after myocardial infarction', *Lancet*, **1**, 1172–1173 (1980).
2. Canner, P. L. 'Aspirin in coronary heart disease. Comparison of six clinical trials', *Israel Journal of Medical Sciences*, **19**, 413–423 (1983).
3. Furberg, C. D., *et al.* (1984, see above).

**V: Psychoeducational Intervention**

1. Devine, E. C. and Cook, T. D. 'A meta-analytic analysis of effects of psychoeducational interventions on length of postsurgical hospital stay', *Nursing Research*, **32**, 267–274 (1983).
2. Mumford, E., Schlesinger, H. J. and Glass, G. V. 'The effect of psychological intervention on recovery from surgery and heart attacks: An analysis of the literature', *American Journal of Public Health*, **72**, 141–151 (1982).

**VI: Patient Education**

1. Mazzuca, S. A. 'Does patient education in chronic disease have therapeutic value', *Journal of Chronic Diseases*, **35**, 521–529 (1982).
2. Posavac, E. J. 'Evaluations of patient education programs. A meta-analysis', *Evaluation and The Health Professions*, **3**, 47–62 (1980).

**VII and VIII: Nicotine Chewing Gum (in clinics and in private practice)**

1. Lam, W., Sze, P. C., Sacks, H. S. and Chalmers, T. C. 'Meta-analysis of randomized control trials (RCTs) of nicotine chewing gum', *Lancet*, **2**, 27–30 (1987).
2. Raw, M. 'Does nicotine chewing gum work?', *Lancet*, **1**, 1231–1232 (1985).

**IX: Prevention of Venous Thrombosis**

1. Colditz, G. A., Tuden, R. L. and Oster, G. 'Rates of venous thrombosis after general surgery: Combined results of randomised clinical trials', *Lancet*, **2**, 143–146 (1986).

2. Gent, M. and Roberts, R. S. 'A meta-analysis of the studies of dihydroergotamine plus heparin in the prophylaxis of deep vein thrombosis', *Chest*, **89** (Suppl.) 396S–406S (1986).

**X: Stimulant Therapy of Venous Thrombosis**

1. Kavale, K. A. and Forness, S. R. 'Hyperactivity and diet treatment: A meta-analysis of the Feingold hypothesis', *Journal of Learning Disabilities*, **16**, 324–330 (1983).
2. Thurber, S. and Walker, C. E. 'Medication and hyperactivity: A meta-analysis', *Journal of General Psychology*, **108**, 79–86 (1983).

## APPENDIX II: DISAGREEING COHORTS

**I: Psychotherapy**

1. Andrews, G. and Harvey, R. 'Does psychotherapy benefit neurotic patients? A reanalysis of the Smith, Glass, and Miller data', *Archives of General Psychiatry*, **38**, 1203–1208 (1981).
2. Landman, J. T. and Dawes, R. M. 'Psychotherapy outcome. Smith and Glass conclusions stand up under scrutiny', *American Psychologist*, **37**, 504–516 (1982).
3. Prioleau, L., Murdock, M. and Brody, N. 'An analysis of psychotherapy versus placebo studies', *Behavioral Brain Science*, **6**, 275–310 (1983).
4. Shapiro, D. A. and Shapiro, D. 'Meta-analysis of comparative therapy outcome studies: A replication and refinement', *Psychological Bulletin*, **92**, 581–604 (1982).
5. Smith, M. L. and Glass, G. V. 'Meta-analysis of psychotherapy outcome studies', *American Psychologist*, **32**, 752–760 (1977).

**II: Steroids in Alcoholic Hepatitis**

1. Conn, H. O. 'Steroid treatment of alcoholic hepatitis. The yeas and the nays', *Gastroenterology*, **74**, 319–326 (1978).
2. Galambos, J. T. and Riepe, S. P. 'Use of colchicine and steroids in the treatment of alcoholic liver disease', *Recent Developments in Alcoholism*, **2**, 181–194 (1984).
3. Kirschner, E., Silverman, B. and Blackburn, B. 'Steroid treatment of acute alcoholic hepatitis. Analysis of the seven randomized control trials', *Gastroenterology*, **75**, 971 (1978).

**III: Anti-Depressant Drugs**

1. The Quality Assurance Project. 'A treatment outline for depressive disorders', *Australian New Zealand Journal of Psychiatry*, **17**, 129–146 (1983).
2. Smith, A., Traganza, E. and Hanison, G. 'Studies on the effectiveness of antidepressant drugs', *Psychopharmacology Bulletin*, Suppl., 1–53 (1969).
3. Wechsler, H., Grosser, G. H. and Greenblatt, M. 'Research evaluating antidepressant medications on hospitalized mental patients: A survey of published reports during a five-year period', *Journal of Nervous and Mental Disease*, **141**, 231–239 (1965).

**IV: Lidocaine, Acute MI**

1. DeSilva, R. A., Hennekens, C. H., Lown, B. and Casscells, W. 'Lignocaine prophylaxis in acute myocardial infarction: An evaluation of randomised trials', *Lancet*, **2**, 855–858 (1981).

2. Goldman, L. and Batsford, W. P. 'Risk-benefit stratification as a guide to lidocaine prophylaxis of primary ventricular fibrillation in acute myocardial infarction: An analytic review', *Yale Journal of Biology and Medicine*, **52**, 455–466 (1979).
3. Hine, K., Laird, N. and Chalmers, T. C. 'Meta-analysis indicates a need for more mortality data on routine lidocaine use in acute myocardial infarction (AMI)', *Clinical Research*, **34**, 368A (1986).

**V and VI: Single Dose versus Conventional for Urinary Tract Infection (TMP/SMZ and Amoxicillin)**

1. Carlson, K. J., Mylley, A. G. 'Management of acute dysuria. A decision-analysis model of alternative strategies', *Annals of Internal Medicine*, **102**, 244–249 (1985).
2. Freire, J. M., Statschenko, S., Fonberg, E., Berlin, J. and Chalmers, T. C. 'Meta-analysis of the evidence comparing single versus conventional treatment in lower urinary tract infections in adult women using amoxicillin and TMP/SMZ', (Unpublished manuscript, 1986).
3. Philbrick, J. T., Bracikowski, J. P. 'Single-dose antibiotic treatment for uncomplicated urinary tract infections. Less for less?', *Archives of Internal Medicine*, **145**, 1672–1678 (1985).

**VII: Anti-Coagulants, Post-MI**

1. 'Collaborative analysis of long-term anticoagulant administration after acute myocardial infarction', An International Anticoagulant Review Group, *Lancet*, **1**, 203–209 (1970).
2. Furberg, C. D., *et al.* (1984, see above).
3. Leizorovicz, A. and Boissel, J. P. 'Oral anticoagulant in patients surviving myocardial infarction. A new approach to old data', *European Journal of Clinical Pharmacology*, **24**, 333–336 (1983).

**VIII: Radiotherapy after Radical Mastectomy**

1. Levitt, S. H., McHugh, R. B. and Song, C. W. 'Radiotherapy in the postoperative treatment of operable cancer of the breast. Part II. A re-examination of Stjernsward's application of the Mantel–Haenszel statistical method. Evaluation of the effect of the radiation on immune response and suggestions for postoperative radiotherapy,' *Cancer*, **39**, 933–940 (1976).
2. Stjernsward, J. 'Decreased survival related to irradiation postoperatively in early operable breast cancer', *Lancet*, **2**, 1285–1286 (1974).

**IX: Association of Steroids and Peptic Ulcer**

1. Conn, H. O. and Blitzer, B. L. 'Nonassociation of adrenocorticosteroid therapy and peptic ulcer', *New England Journal of Medicine*, **294**, 473–479 (1976).
2. Messer, J., Reitman, D., Sacks, H. S., Smith, H. Jr. and Chalmers, T. C. 'Association of adrenocorticosteroid therapy and peptic-ulcer disease', *New England Journal of Medicine*, **309**, 21–24 (1983).

**X: Diuretics in Pregnancy**

1. Collins, R., Yusuf, S. and Peto, R. 'Overview of randomised trials of diuretics in pregnancy', *British Medical Journal*, **290**, 17–23 (1985).
2. Hemminki, E. 'Diuretics in pregnancy: A case study of a worthless therapy', *Social Science and Medicine*, **18**, 1011–1018 (1984).

ACKNOWLEDGEMENT

Supported in part by grant number R01 LM03116 from the National Library of Medicine and by grant number GA-8HS-8204 from the Rockefeller Foundation.

## REFERENCES

1. Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V. A. and Chalmers, T. C. 'Meta-analyses of randomized controlled trials', *New England Journal of Medicine*, **316**, 450–455 (1987).
2. Chalmers, T. C., Levin, H., Sacks, H. S., Reitman, D., Berrier, J. and Nagalingam, R. 'Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials', *Statistics in Medicine*, **6**, 315–325 (1987).
3. Glass, G. V., McGaw, B. and Smith, M. L. *Meta-analysis in Social Research*, Sage Publication, Beverly Hills, 1981, Chapter 1.
4. Yusuf, S., Peto, R., Lewis, J., Collins, R. and Sleight, P. 'Beta blockade during and after myocardial infarction: An overview of the randomized trials', *Progress in Cardiovascular Diseases*, **27**, 335–371 (1985).
5. Goldman, L. and Feinstein, A. R. 'Anticoagulants and myocardial infarction: The problems of pooling, drowning and floating', *Annals of Internal Medicine*, **90**, 92–94 (1979).
6. Mantel, N. and Haenszel, W. 'Statistical aspects of the analysis of data from retrospective studies of disease', *Journal of the National Cancer Institute*, **22**, 719–748 (1959).
7. Yusuf, S., Collins, R., Peto, R., Furberg, C., Stampfer, M. J., Goldhaber, S. Z. and Hennekens, C. H. 'Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: Overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials', *European Heart Journal*, **6**, 556–585 (1985).
8. Stampfer, M. J., Goldhaber, S. Z., Yusuf, S., Peto, R. and Hennekens, C. H. 'Effect of intravenous streptokinase on acute myocardial infarction. Pooled results from randomized trials', *New England Journal of Medicine*, **307**, 1180–1182 (1982).
9. DeSilva, R. A., Hennekens, C. H., Lown, B. and Casscells, W. 'Lignocaine prophylaxis in acute myocardial infarction: an evaluation of randomised trials', *Lancet*, **2**, 855–858 (1981).
10. Goldman, L. and Batsford, W. P. 'Risk–benefit stratification as a guide to lidocaine prophylaxis of primary ventricular fibrillation in acute myocardial infarction: An analytic review', *Yale Journal of Biology and Medicine*, **52**, 455–466 (1979).
11. Hine, K., Laird, N. and Chalmers, T. C. 'Meta-analysis indicates a need for more mortality data on routine lidocaine use in acute myocardial infarction (AMI)', *Clinical Research*, **34**, 368A (1986).
12. Collins, R., Yusuf, S. and Peto, R. 'Overview of randomised trials of diuretics in pregnancy', *British Medical Journal*, **290**, 17–23 (1985).
13. Hemminki, E. 'Diuretics in pregnancy: A case study of a worthless therapy', *Social Science and Medicine*, **18**, 1011–1018 (1984).
14. Stjernsward, J. 'Decreased survival related to irradiation postoperatively in early operable breast cancer', *Lancet*, **2**, 1285–1286 (1974).
15. Levitt, S. H., McHugh, R. B. and Song, C. W. 'Radiotherapy in the postoperative treatment of operable cancer of the breast. Part II. A re-examination of Stjernsward's application of the Mantel–Haenszel statistical method. Evaluation of the effect of the radiation of immune response and suggestions for postoperative radiotherapy', *Cancer*, **39**, 933–940 (1976).
16. Cooper, H. 'On the social psychology of using research review: the case of desegregation and the black achiever'. in Feldman, R. S. (ed.) *Social Psychology of Education*, Cambridge Press, Cambridge, 1986, pp. 341–363, Chapter 14.
17. The MIAMI Trial Group. 'MIAMI: Metoprolol in acute myocardial infarction', *American Journal of Cardiology*, **56**, 1G–57G (1985).
18. ISIS-I (First International study of Infarct Survival) Collaborative Group. 'Randomised trial of intravenous atenolol among 16,027 cases of suspected acute myocardial infarction: ISIS-I', *Lancet*, **2**, 57–66 (1986).
19. Gruppo Italiano Per Lo Studio Della Streptochinasi Neli' infarto Miocardico (GISSI). 'Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction', *Lancet*, **1**, 397–402 (1986).