Special Articles

A COMPARISON OF OBSERVATIONAL STUDIES AND RANDOMIZED, CONTROLLED TRIALS

KJELL BENSON, B.A., AND ARTHUR J. HARTZ, M.D., PH.D.

ABSTRACT

Background For many years it has been claimed that observational studies find stronger treatment effects than randomized, controlled trials. We compared the results of observational studies with those of randomized, controlled trials.

Methods We searched the Abridged Index Medicus and Cochrane data bases to identify observational studies reported between 1985 and 1998 that compared two or more treatments or interventions for the same condition. We then searched the Medline and Cochrane data bases to identify all the randomized, controlled trials and observational studies comparing the same treatments for these conditions. For each treatment, the magnitudes of the effects in the various observational studies were combined by the Mantel–Haenszel or weighted analysis-of-variance procedure and then compared with the combined magnitude of the effects in the randomized, controlled trials that evaluated the same treatment.

Results There were 136 reports about 19 diverse treatments, such as calcium-channel–blocker therapy for coronary artery disease, appendectomy, and interventions for subfertility. In most cases, the estimates of the treatment effects from observational studies and randomized, controlled trials were similar. In only 2 of the 19 analyses of treatment effects did the combined magnitude of the effect in observational studies lie outside the 95 percent confidence interval for the combined magnitude in the randomized, controlled trials.

Conclusions We found little evidence that estimates of treatment effects in observational studies reported after 1984 are either consistently larger than or qualitatively different from those obtained in randomized, controlled trials. (N Engl J Med 2000;342: 1878-86.)

OBSERVATIONAL studies have several advantages over randomized, controlled trials, including lower cost, greater timeliness, and a broader range of patients.[1] Concern about inherent bias in these studies, however, has limited their use in comparing treatments.[2,3] Observational studies are used primarily to identify risk factors and prognostic indicators and in situations in which randomized, controlled trials would be impossible or unethical.[4]

The empirical assessment of observational studies rests largely on a number of influential comparative studies from the 1970s and 1980s.[5-9] These studies suggested that observational studies inflate positive treatment effects, as compared with randomized, controlled trials. In one major study, Chalmers et al.[6] showed that 56 percent of nonrandomized trials yielded favorable treatment effects, as compared with 30 percent of blinded, randomized, controlled trials. Three other studies had similar results.[7-9] According to many experts, these results mean that observational studies should not be used for defining evidence-based medical care: "If you find that [a] study was not randomized, we'd suggest that you stop reading it and go on to the next article."[10]

Evaluations of observational studies have primarily included studies from the 1960s and 1970s. We evaluated observational studies reported between 1985 and 1998, studies which may be methodologically superior to earlier studies. Possible methodologic improvements include a more sophisticated choice of data sets and better statistical methods. Newer methods may have eliminated some systematic bias.

METHODS

Search for Observational Studies

Observational studies were found by systematically searching Medline and the Cochrane Database of Systematic Reviews for studies reported from 1985 through 1998. Although Medline is now indexed for highly sensitive searches for randomized, controlled trials, "observational studies" is not an indexable concept in Medline, and there is no search term for observational studies (Wright N, National Library of Medicine: personal communication). Therefore, we used a text-word strategy to search for "observational," "cohort," "retrospective," "cross-sectional," and "nonrandomized." We limited the search to journals in the Abridged Index Medicus, which indexes the 120 most widely read, prestigious clinical journals. To restrict the search to studies comparing treatments, we added the Medline tag "comparative study/," defined as a comparison of any two or more concepts from any Medical Subject Heading category.

This strategy identified 3868 articles. We reviewed the abstracts

From the Department of Family Medicine, University of Iowa College of Medicine, Iowa City. Address reprint requests to Dr. Hartz at the Department of Family Medicine, University of Iowa College of Medicine, 01292-D PFP, Iowa City, IA 52242-1097, or at arthur-hartz@uiowa.edu.

of these articles and selected only those that met four criteria. First, the study was not experimental — that is, treatments were not assigned for purposes of research. Second, the study assessed the difference between two treatments or between one treatment and no treatment. Third, the treatments were implemented by physicians. Studies of diet, exercise, lifestyle changes, or nonprescription medication were not included, since the type of bias in these studies differs from the type of bias in studies of physician-implemented treatment. Fourth, the study included a control group.

Search for Related Studies

When an article that met all four criteria was identified, we searched the entire Medline data base from 1966 to 1998 for all corresponding randomized, controlled trials and observational studies — that is, those that compared the same two treatments (or the same treatment and no treatment), used the same outcome measure, and used the same inclusion criteria for patients. We included a few studies that did not have the same inclusion criteria or follow-up times as the observational studies; these studies are noted in the Results.

The Cochrane Database was searched by the same strategy used with Medline. This data base is a continuously updated series of reviews by members of the Cochrane Collaboration, an international organization that collects research information on the effects of health care interventions.[11] Search of the Cochrane Database found three additional treatments for which there were both observational studies and randomized, controlled trials.

Additional articles were found by searching the reference lists of the reviewed articles and by searching for pseudo-randomized studies with the Medical Subject Heading "controlled clinical trial." Pseudo-randomized studies are controlled trials that assign treatments in a nonrandomized way, such as by giving the same treatment to every other admitted patient.[12] The results of the pseudo-randomized trials are reported separately from the results of both the observational studies and the randomized, controlled trials.

No ideal criteria were available to evaluate the sensitivity of our search strategy. However, the United Kingdom Health Technology Assessment Group recently completed a systematic review of 22 treatments that were the subject of both randomized, controlled trials and observational studies.[13] For an estimate of the sensitivity of our search strategy, we compared our search results with theirs.

Statistical Analysis

We compared the magnitudes of the effects of treatment on principal outcomes in observational studies and randomized, controlled trials. We used the Cochrane Collaboration's RevMan and MetaView software (version 3.1) to combine the magnitudes of the effects of a treatment in corresponding studies that had the same design. For binary outcomes, we used the Mantel–Haenszel method for estimating the overall odds ratios. For continuous outcome measures, we used a fixed-effects, weighted analysis-of-variance model, with the inverse of the variance of the magnitude of the effect as the weighting factor. The use of fixed-effects models exaggerates differences between the results of observational studies and randomized, controlled trials if the actual treatment effects vary among study populations.

Five studies included in our analysis did not report a confidence interval for the magnitude of the effect. For three of these studies, we estimated the confidence interval from the magnitude of the effect and the P value. Our procedure was as follows: we transformed the magnitude of the effect into a statistic with an approximately normal distribution (e.g., the log of the odds ratio); we transformed the P value into a normal test statistic; we used the transformed magnitude of the effect and the P value to compute the standard error of the transformed magnitude of the effect; we used this information to create a 95 percent confidence interval for the transformed magnitude of the effect; and we used this confidence interval to create a confidence interval for the untransformed magnitude of the effect. Although these confidence intervals may not be identical to those that could be computed by other means, they should provide a qualitative indication of the degree of precision with which the magnitude of the effect was estimated.

RESULTS

We found 19 treatment comparisons that were the subject of at least one observational study and at least one randomized, controlled trial. There were 53 observational studies and 83 randomized, controlled trials. Two additional studies of these treatments were pseudo-randomized.

Our search identified the studies for 13 of the 22 treatment comparisons identified by the United Kingdom Health Technology Assessment Group. Among the other nine treatment comparisons, four had not been the subject of an observational study in a journal listed in the Abridged Index Medicus, two had not been the subject of both a randomized, controlled trial and an observational study, one had not been the subject of an observational study after 1984, one was not a medical treatment, and one had not been the subject of a study indexed under the Medical Subject Heading "comparative study/."

In the selection of corresponding studies, there may have been differences in how some of the treatments were administered (e.g., evaluations by geriatric assessment units) or in how some of the outcomes were assessed (e.g., the incidence of infection, recurrent dysphagia, or retinopathy). Both follow-up times and inclusion criteria were identical for 15 of the 19 treatment comparisons. The results for treatments with fewer than five observational studies or five randomized, controlled trials are summarized in Figures 1 and 2. Results for treatments with more studies are shown in subsequent figures.

Figure 1 summarizes the results of observational studies and corresponding randomized, controlled trials for seven cardiologic treatments. The differences in design between the two types of study were as follows: the dose of nifedipine in the observational study was 30 to 60 mg, as compared with 30 to 50 mg in the randomized, controlled trials. The inclusion criteria and follow-up times varied among the randomized, controlled trials of nifedipine. For the observational study comparing coronary-artery bypass grafting (CABG) with percutaneous transluminal coronary angioplasty (PTCA), low risk was defined by a proprietary Medisgroups scale.[41]

The observational results fell within the confidence intervals of the randomized, controlled trials in every area except for the comparison of CABG with PTCA in patients at low risk. All of the other odds ratios were similar with the two study designs, except for the comparison of CABG and PTCA in diabetic patients. The confidence intervals of the observational studies were slightly narrower than those of the randomized, controlled trials.

Figure 2 summarizes the results of observational studies and randomized, controlled trials of 11 non-

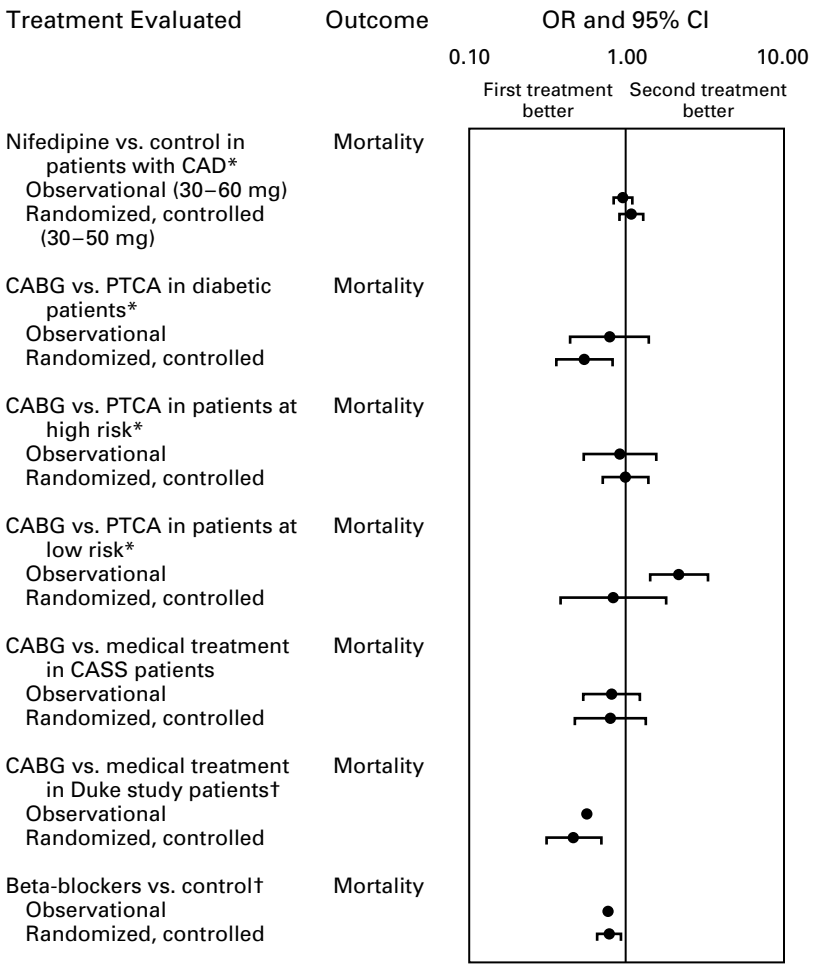

Figure 1. Results of Observational Studies and Randomized, Controlled Trials of Cardiologic Treatments.

The figure is based on data from eight articles.[13-20] Some articles contain data from more than one study. OR denotes odds ratio, CI confidence interval, CAD coronary artery disease, CABG coronary-artery bypass graft surgery, PTCA percutaneous transluminal coronary angioplasty, CASS Coronary Artery Surgery Study, and Duke the Duke University Cardiovascular Disease Databank. Asterisks indicate studies that reported relative risks rather than odds ratios. Daggers indicate studies that reported neither a confidence interval nor a P value for the odds ratio.

cardiologic treatments. For the insulin studies, there was variation in inclusion criteria, and the follow-up times varied from 3 to 7.5 years. The results of all the observational studies fell within the confidence intervals of the randomized, controlled trials, except for the comparison of pneumatic retinopexy with scleral buckling. The results of the two types of study also differed qualitatively for three other treatments, although these differences are difficult to interpret because of the wide confidence intervals.

In one pseudo-randomized trial comparing water-soluble with oil-soluble contrast medium for flushing of ovarian tubes, the odds ratio for pregnancy was 2.00, as compared with 1.92 for both the randomized, controlled trials and the observational studies.[40] In one pseudo-randomized trial comparing geriatric assessment units and medical wards, the odds ratio for death was 0.51, as compared with 0.69 for the one observational study and 0.65 for the randomized, controlled trials.[42]

Figure 3 shows the results of studies of the effects of only one treatment, hormone-replacement therapy, on lumbar bone mineral density after one to two years of treatment. The inclusion criteria for the two types of study were identical, except for one randomized, controlled trial that included only women

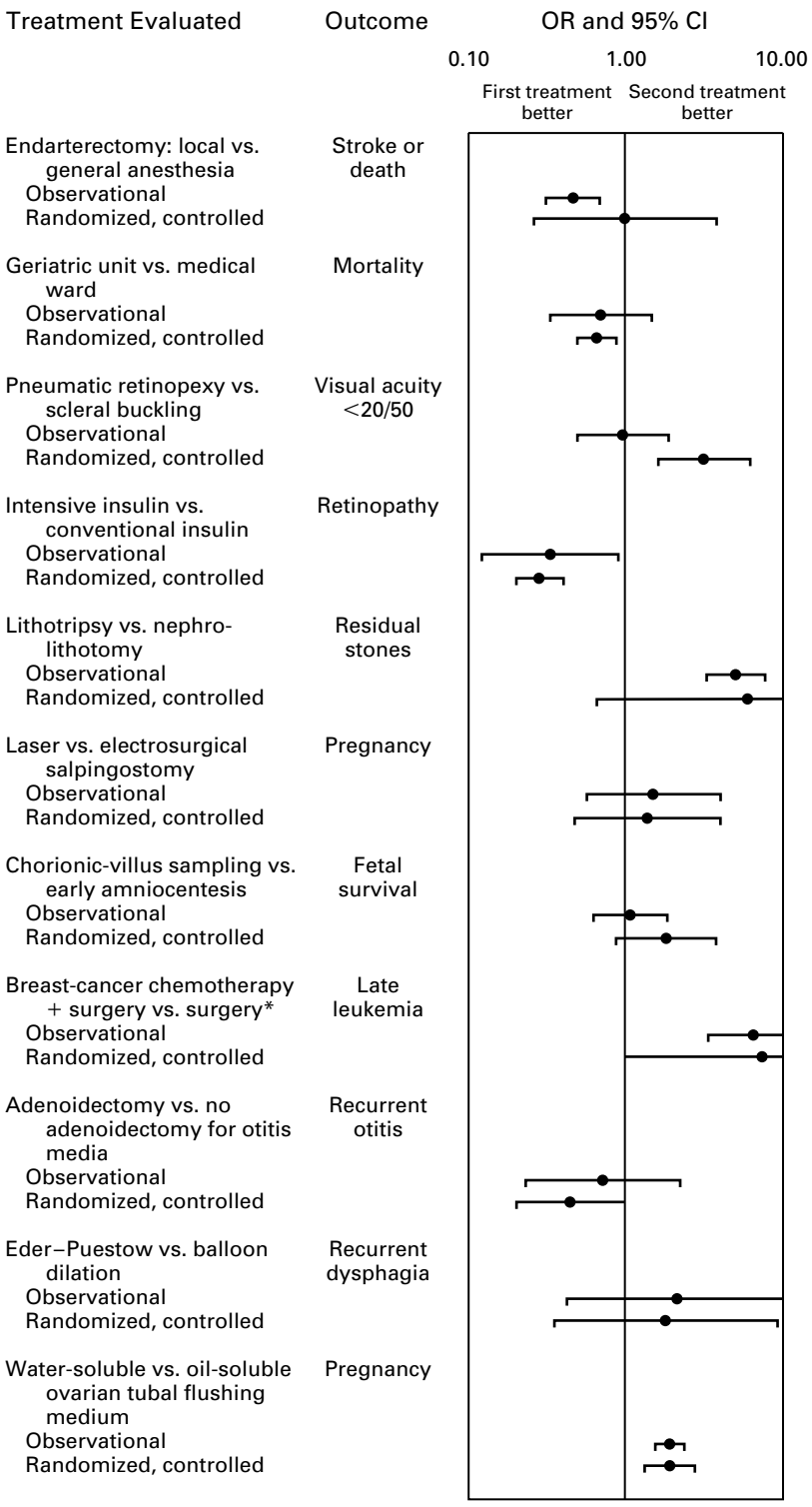

Figure 2. Results of Observational Studies and Randomized, Controlled Trials of Noncardiologic Treatments.

The figure is based on data from 20 articles.[21-40] Some articles contain data from more than one study. OR denotes odds ratio, and CI confidence interval. The asterisk indicates a study that reported relative risks rather than odds ratios.

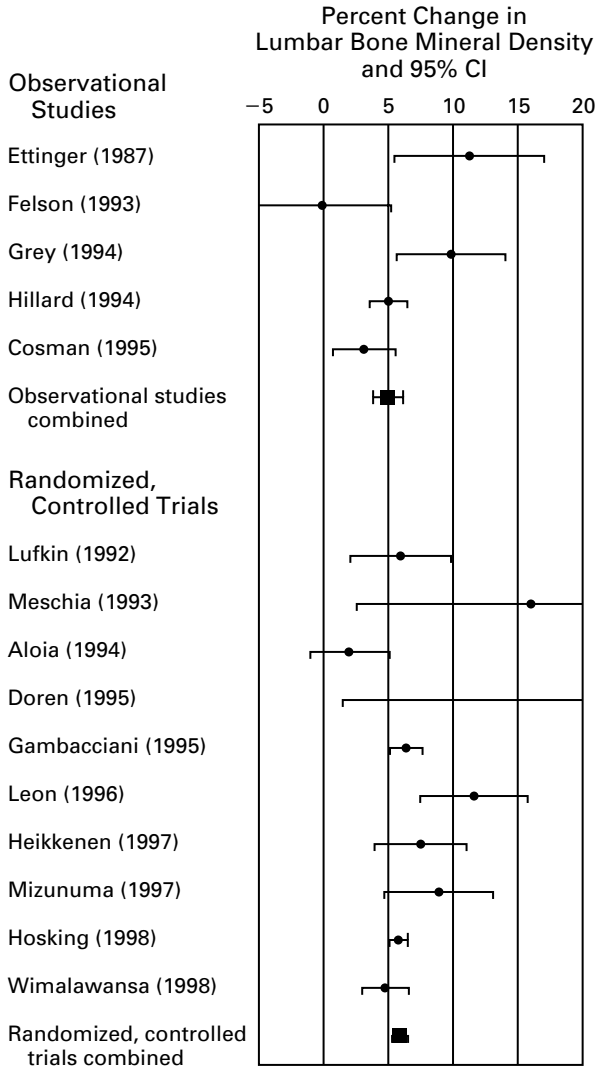

Figure 3. Percent Change in Lumbar Bone Density in Postmenopausal Women Given One to Two Years of Hormone-Replacement Therapy as Compared with Controls.

The figure is based on data from 15 articles.[43-57] CI denotes confidence interval.

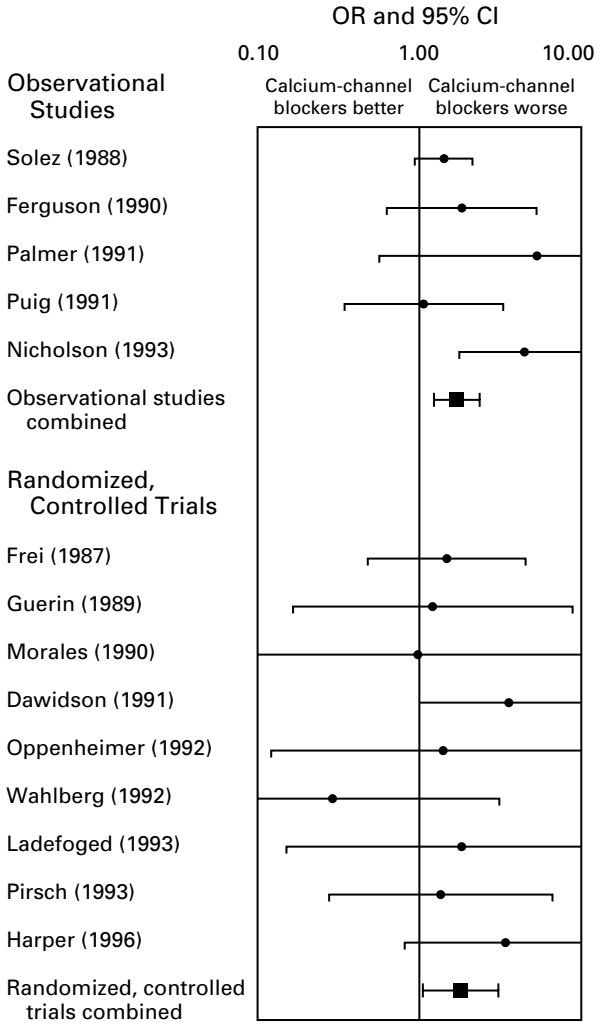

Figure 4. Odds Ratio for Graft Survival after Kidney Transplantation in Patients Receiving Calcium-Channel Blockers as Compared with Controls.

The figure is based on data from six articles.[58-63] The nine randomized, controlled trials were analyzed by Ladefoged and Andersen.[63] OR denotes odds ratio, and CI confidence interval.

with osteoporosis. The randomized, controlled trials of this treatment, particularly the later studies, had larger samples and narrower confidence intervals. The combined result of the observational studies lay just below the lower bound of the confidence interval of the combined randomized, controlled trials, although the results were qualitatively very similar.

Figure 4 shows the results of studies evaluating the use of calcium-channel blockers in patients receiving kidney allografts. There were some differences in follow-up times among these studies: the follow-up time in the Morales study was only 30 days, and the follow-up time in the Wahlberg and Ladefoged studies was 3 months.[63] All other follow-up times were between six months and two years. There were also differences in the immunosuppressive regimens administered. The individual studies did not detect a significant effect of calcium-channel blockers, although the meta-analysis did. The overall results were almost identical with the two study designs.

Figure 5 shows the results of studies comparing laparoscopic with open appendectomy. This analysis involved 24 studies, the greatest number of individual studies for any comparison. Few of the individual

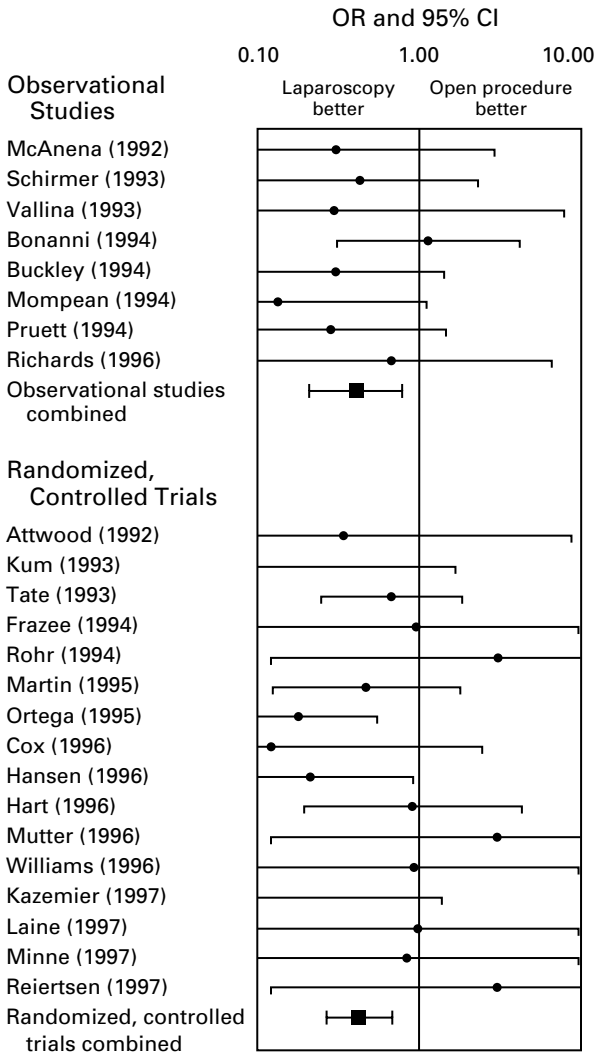

Figure 5. Odds Ratio for Infection after Laparoscopic as Compared with Open Appendectomy.

The figure is based on data from nine articles.[64-72] The 16 randomized, controlled trials were analyzed by Golub et al.[72] Trials that reported rates of all complications rather than rates of infection have been excluded. OR denotes odds ratio, and CI confidence interval.

studies demonstrated a significant benefit of laparoscopic appendectomy. However, the meta-analysis did detect a benefit, which was of the same magnitude in observational studies and randomized, controlled trials.

In general, for any specific treatment, the observational studies were conducted before the results of the randomized, controlled trials became available. However, in some cases the observational studies were conducted after the randomized, controlled trials. Reasons given for conducting later observational studies included a lack of long-term safety data,[14,46] concurrent collection of the observational data from patients who declined to be included in a randomized, controlled trial,[17,38,39] and evaluation of the generalizability of the results of the randomized, controlled trial in a wider population.[16]

DISCUSSION

In this study we compared the results of observational studies and randomized, controlled trials. We found 136 articles in 19 treatment areas. All but six of these articles were published between 1985 and 1998. The estimates of the effects of treatment in observational studies and in randomized, controlled trials were similar in most areas, and for only 2 of the 19 treatments did the magnitude of the effect in the combined observational studies lie outside the 95 percent confidence interval for the combined randomized, controlled trials. For most treatments, however, there were insufficient data to exclude the possibility of clinically important differences between the results of the two types of study.

The small number of suitable articles we found may be due partly to limitations of computerized searches for reports of observational studies and partly to the paucity of treatments that have been evaluated by both randomized, controlled trials and observational studies. Our results may not apply to other treatments. However, because the treatments evaluated were diverse, it is likely that randomized, controlled trials and observational studies (at least those reported since 1985 in journals listed in the Abridged Index Medicus) often produce similar results.

There were discrepancies between the confidence intervals of the observational study and the randomized, controlled trial that compared CABG with PTCA for patients at low risk. The mortality rates were similar for the two treatments in the randomized, controlled trial, but in the observational study the mortality rates were higher for the patients undergoing CABG, particularly during the first 60 days after surgery.[73] In the randomized, controlled trial, patients at low risk undergoing CABG had very low early mortality. This mortality rate may not be representative of the mortality rate associated with CABG in most community hospitals.[74]

The greatest statistical discrepancy between the results of the two types of study was for studies comparing pneumatic retinopexy with scleral buckling for the treatment of retinal detachment. The observational studies and the randomized, controlled trial both found that the two procedures were associated with similar final rates of reattachment after reoperation and similar rates of postoperative proliferative vitreoretinopathy.[24-26] However, the randomized, controlled trial, but not the observational studies, found that the two procedures were associated with similar rates of reattachment after the first operation and

that pneumatic retinopexy had a better visual outcome than scleral buckling. The results for patients undergoing scleral buckling were similar for the two study designs. One possible explanation for these results is that the patients undergoing retinopexy in the observational studies were at higher risk than those undergoing scleral buckling. A more likely explanation, however, is that the outcome of retinopexy was unusually good in the randomized, controlled trial.

We did not select articles to reduce the heterogeneity of the results or to ensure high quality (except that articles from journals listed in the Abridged Index Medicus were included in every treatment comparison). The choice of selection criteria was subjective and may have affected the results.[75] On the other hand, our results may have been influenced by the inclusion of flawed studies.

Our finding that observational studies and randomized, controlled trials usually produce similar results differs from the conclusions of previous authors. A study in 1977 reviewed the evidence of the effectiveness of anticoagulants in the treatment of acute myocardial infarction, using eight observational studies and six randomized, controlled trials.[5] The differences in mortality rates between control and treatment groups were larger in the observational studies than in the randomized, controlled trials. The observational studies reviewed were published before 1975, and the authors did not use current meta-analytic techniques for pooling data. The results of the comparison might have differed if current methods had been used to combine the results of several trials.

Some of the same authors later reviewed 160 studies that evaluated six cardiology treatments.[6] They found that the reported outcomes were better for the treatment group than the controls in 60 percent of randomized, controlled trials and 93 percent of observational studies. As pointed out at the time, however, most of their studies of beta-blockers were randomized, controlled trials, whereas most of their studies evaluating treatment in coronary care units were observational.[76] The greater treatment effects in the observational studies might be explained by the greater effectiveness of treatment in coronary care units than of treatment with beta-blockers.

Three other studies commonly cited to show the inadequacy of observational data,[7-9] as well as one that found no bias in observational data,[77] also compared observational studies and randomized, controlled trials that evaluated different treatments. As compared with these previous studies, our study has the advantage that the comparisons were stratified according to treatment. In addition, the studies that we reviewed were more recent and therefore may have used better methods than those in the earlier reviews.

A recent investigation to compare observational studies and randomized, controlled trials was performed by the United Kingdom Health Technology Assessment Group.[13] They found eight treatments, not evaluated by us, that were the subject of a randomized, controlled trial and of an observational study with a control group. In seven of these there were no differences between the results of the observational studies and the results of the randomized, controlled trials, and in the other the effect was greater in the observational studies. For the last treatment comparison (cost savings associated with hospice care),[78] the length of time in the hospice differed between the two types of studies. The study by the United Kingdom Health Technology Assessment Group concluded that there were no systematic biases in observational studies.

Although observational studies may generally give valid results, there are known limitations. In particular, as found by Green and Byar,[79] observational studies cannot be used to evaluate treatments that physicians routinely select for the sickest patients. On the basis of our findings, this misuse of observational studies does not often occur in the recent literature listed in the Abridged Index Medicus.

The fundamental criticism of observational studies is that unrecognized confounding factors may distort the results. According to the conventional wisdom, this distortion is sufficiently common and unpredictable that observational studies are not reliable and should not be funded. Our results suggest that observational studies usually do provide valid information. They could be used to exploit the many recently developed, clinically rich data bases. Only with a greater willingness to analyze these data bases is it possible to achieve a realistic understanding of how observational studies can best be used.

Supported in part by grants from the Health Services and Resources Administration (PD15 PE87007 and 5D32PE10195-02) and the National Heart, Lung, and Blood Institute (2T35HL07485-21).

We are indebted to Dedra Diehl, M.L.S., for help with library research.

REFERENCES

1. Feinstein AR. Epidemiologic analyses of causation: the unlearned scientific lessons of randomized trials. J Clin Epidemiol 1989;42:481-9.
2. Moses LE. Measuring effects without randomized trials? Options, problems, challenges. Med Care 1995;33:Suppl:AS8-AS14.
3. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. BMJ 1998; 317:1185-90.
4. Naylor CD, Guyatt GH. Users' guides to the medical literature. X. How to use an article reporting variations in the outcomes of health services. JAMA 1996;275:554-8.
5. Chalmers TC, Matta RJ, Smith H Jr, Kunzler A-M. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. N Engl J Med 1977;297:1091-6.
6. Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. N Engl J Med 1983;309:1358-61.
7. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. Am J Med 1982;72:233-40.
8. Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I. Medical. Stat Med 1989;8:441-54.
9. Miller JN, Colditz GA, Mosteller F. How study design affects outcomes in comparisons of therapy. II. Surgical. Stat Med 1989;8:455-66.

10. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine: how to practice and teach EBM. New York: Churchill Livingtone, 1997.
11. Jadad AR, Haynes RB. The Cochrane Collaboration — advances and challenges in improving evidence-based decision making. Med Decis Making 1998;18:2-9.
12. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408-12.
13. Britton A, McPherson K, McKee M, Sanderson C, Black N, Bain C. Choosing between randomised and non-randomised studies: a systematic review. Health Technol Assess 1998;2(13):i-iv, 1-124.
14. Braun S, Boyko V, Behar S, et al. Calcium antagonists and mortality in patients with coronary artery disease: a cohort study of 11,575 patients. J Am Coll Cardiol 1996;28:7-11.
15. Furberg CD, Psaty BM, Meyer JV. Nifedipine: dose-related increase in mortality in patients with coronary heart disease. Circulation 1995;92: 1326-31.
16. Detre KM, Guo P, Holubkov R, et al. Coronary revascularization in diabetic patients: a comparison of the randomized and observational components of the Bypass Angioplasty Revascularization Investigation (BARI). Circulation 1999;99:633-40.
17. Coronary Artery Surgery Study (CASS): a randomized trial of coronary artery bypass surgery: comparability of entry characteristics and survival in randomized patients and nonrandomized patients meeting randomization criteria. J Am Coll Cardiol 1984;3:114-28.
18. Hlatky MA, Califf RM, Harrell FE Jr, Lee KL, Mark DB, Pryor DB. Comparison of predictions based on observational data with the results of randomized controlled clinical trials of coronary artery bypass surgery. J Am Coll Cardiol 1988;11:237-45.
19. Horwitz RI, Viscoli CM, Clemens JD, Sadock RT. Developing improved observational methods for evaluating therapeutic effectiveness. Am J Med 1990;89:630-8.
20. A randomized trial of propranolol in patients with acute myocardial infarction. I. Mortality results. JAMA 1982;247:1707-14.
21. Tangkanakul C, Counsell C, Warlow C. Local versus general anaesthesia for carotid endarterectomy (Cochrane review). In: The Cochrane library, 2. Oxford, England: Update Software, 1999.
22. Teasdale TA, Shuman L, Snow E, Luchi RJ. A comparison of placement outcomes of geriatric cohorts receiving care in a geriatric assessment unit and on general medicine floors. J Am Geriatr Soc 1983;31:529-34.
23. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. Lancet 1993;342:1032-6.
24. Han DP, Mohsin NC, Guse CE, Hartz A, Tarkanian CN. Comparison of pneumatic retinopexy and scleral buckling in the management of primary rhegmatogenous retinal detachment. Am J Ophthalmol 1998;126: 658-68.
25. McAllister IL, Meyers SM, Zegarra H, Gutman FA, Zakov ZN, Beck GJ. Comparison of pneumatic retinopexy with alternative surgical techniques. Ophthalmology 1988;95:877-83.
26. Tornambe PE, Hilton GF. Pneumatic retinopexy: a multicenter randomized controlled clinical trial comparing pneumatic retinopexy with scleral buckling. Ophthalmology 1989;96:772-84.
27. Friberg TR, Rosenstock J, Sanborn G, Vaghefi A, Raskin P. The effect of long-term near normal glycemic control on mild diabetic retinopathy. Ophthalmology 1985;92:1051-8.
28. Rosenstock J, Friberg T, Raskin P. Effect of glycemic control on microvascular complications in patients with type I diabetes mellitus. Am J Med 1986;81:1012-8.
29. Reichard P, Nilsson B-Y, Rosenqvist U. The effect of long-term intensified insulin treatment on the development of microvascular complications of diabetes mellitus. N Engl J Med 1993;329:304-9.
30. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin dependent diabetes mellitus. N Engl J Med 1993;329:977-86.
31. Carlsson P, Kinn AC, Tiselius HG, Ohlsen H, Rahmqvist M. Cost effectiveness of extracorporeal shock wave lithotripsy and percutaneous nephrolithotomy for medium-sized kidney stones: a randomised clinical trial. Scand J Urol Nephrol 1992;26:257-63.
32. Mays N, Petruckevitch A, Burney PG. Results of one and two year follow-up in a clinical comparison of extracorporeal shock wave lithotripsy and percutaneous nephrolithotomy in the treatment of renal calculi. Scand J Urol Nephrol 1992;26:43-9.
33. Watson A, Vandekerckhove P, Lilford R. Techniques for pelvic surgery in subfertility (Cochrane review). In: The Cochrane library, 2. Oxford, England: Update Software, 1999.
34. Nicolaides K, de Lourdes Brizot M, Patel F, Snijders R. Comparison of chorionic villus sampling and amniocentesis for fetal karyotyping at 10–13 weeks' gestation. Lancet 1994;344:435-9. [Erratum, Lancet 1994;344: 830.]
35. Curtis RE, Boice JD Jr, Stovall M, et al. Risk of leukemia after chemotherapy and radiation treatment for breast cancer. N Engl J Med 1992; 326:1745-51.
36. Fisher B, Rockette H, Fisher ER, Wickerham DL, Redmond C, Brown A. Leukemia in breast cancer patients following adjuvant chemotherapy or postoperative radiation: the NSABP experience. J Clin Oncol 1985;3:1640-58.
37. Shapiro CL, Recht A. Late effects of adjuvant therapy for breast cancer. In: Breast cancer in younger women. Journal of the National Cancer Institute monographs. No. 16. Washington, D.C.: Government Printing Office, 1994:101-12. (NIH publication no. 93-03559.)
38. Paradise JL, Bluestone CD, Rogers KD, et al. Efficacy of adenoidectomy for recurrent otitis media in children previously treated with tympanostomy-tube placement: results of parallel randomized and nonrandomized trials. JAMA 1990;263:2066-73.
39. Yamamoto H, Hughes RW Jr, Schroeder KW, Viggiano TR, DiMagno EP. Treatment of benign esophageal stricture by Eder-Puestow or balloon dilators: a comparison between randomized and prospective nonrandomized trials. Mayo Clin Proc 1992;67:228-36.
40. Vandekerckhove P, Watson A, Lilford R, Harada T, Hughes E. Oil-soluble versus water-soluble media for assessing tubal patency with hysterosalpingography or laparoscopy in subfertile women (Cochrane review). In: The Cochrane library, 2. Oxford, England: Update Software, 1999.
41. Brewster AC, Karlin BG, Hyde LA, Jacobs CM, Bradbury RC, Chae YM. MEDISGRPS: a clinically based approach to classifying hospital patients at admission. Inquiry 1985;22:377-87.
42. Gayton D, Wood-Dauphinee S, de Lorimer M, Tousignant P, Hanley J. Trial of a geriatric consultation team in an acute care hospital. J Am Geriatr Soc 1987;35:726-36.
43. Ettinger B, Genant HK, Cann CE. Postmenopausal bone loss is prevented by treatment with low-dosage estrogen with calcium. Ann Intern Med 1987;106:40-5.
44. Felson DT, Zhang Y, Hannan MT, Kiel DP, Wilson PWF, Anderson JJ. The effect of postmenopausal estrogen therapy on bone density in elderly women. N Engl J Med 1993;329:1141-6.
45. Grey AB, Cundy TF, Reid IR. Continuous combined oestrogen/progestin therapy is well tolerated and increases bone density at the hip and spine in post-menopausal osteoporosis. Clin Endocrinol (Oxf) 1994;40: 671-7.
46. Hillard TC, Whitcroft SJ, Marsh MS, et al. Long-term effects of transdermal and oral hormone replacement therapy on postmenopausal bone loss. Osteoporos Int 1994;4:341-8.
47. Cosman F, Nieves J, Walliser J, Lindsay R. Postmenopausal osteoporosis: patient choices and outcomes. Maturitas 1995;22:137-43.
48. Lufkin ED, Wahner HW, O'Fallon WM, et al. Treatment of postmenopausal osteoporosis with transdermal estrogen. Ann Intern Med 1992;117: 1-9.
49. Meschia M, Brincat M, Barbacini P, Crossignani PG, Albisetti W. A clinical trial on the effects of a combination of elcatonin (carbocalcitonin) and conjugated estrogens on vertebral bone mass in early postmenopausal women. Calcif Tissue Int 1993;53:17-20.
50. Aloia J, Vaswani A, Yeh JK, Ross PL, Flaster E, Dilmanian FA. Calcium supplementation with and without hormone replacement therapy to prevent postmenopausal bone loss. Ann Intern Med 1994;120:97-103.
51. Doren M, Reuther G, Minne HW, Schneider HP. Superior compliance and efficacy of continuous combined oral estrogen-progestogen replacement therapy in postmenopausal women. Am J Obstet Gynecol 1995;173: 1446-51.
52. Gambacciani M, Spinetti A, Orlandi R, et al. Effects of a new estrogen/progestin combination in the treatment of postmenopausal syndrome. Maturitas 1995;22:115-20.
53. Speroff L, Rowan J, Symons J, Genant H, Wilborn W. The comparative effect on bone density, endometrium, and lipids of continuous hormones as replacement therapy (CHART study): a randomized controlled trial. JAMA 1996;276:1397-403.
54. Heikkenen J, Kyllonen E, Kurttila-Matero E, et al. HRT and exercise: effects on bone density, muscle strength and lipid metabolism: a placebo controlled 2-year prospective trial on two estrogen-progestin regimens in healthy postmenopausal women. Maturitas 1997;26:139-49.
55. Mizunuma H, Okano H, Soda M, et al. Prevention of postmenopausal bone loss with minimal uterine bleeding using low dose continuous estrogen/progestin therapy: a 2-year prospective study. Maturitas 1997;27:69-76.
56. Hosking D, Chilvers CED, Christiansen C, et al. Prevention of bone loss with alendronate in postmenopausal women under 60 years of age. N Engl J Med 1998;338:485-92.
57. Wimalawansa SJ. A four-year randomized controlled trial of hormone replacement and bisphosphonate, alone or in combination, in women with postmenopausal osteoporosis. Am J Med 1998;104:219-26.

58. Solez K, Racusen LC, Keown PA, Vaughn WK, Burdick JF. The influence of antihypertensive drug therapy on renal transplant function and outcome. Transplant Proc 1988;20:Suppl 3:618-22.
59. Ferguson CJ, Hillis AN, Williams JD, Griffin PJ, Salaman JR. Calcium-channel blockers and other factors influencing delayed function in renal allografts. Nephrol Dial Transplant 1990;5:816-20.
60. Palmer BF, Dawidson I, Sagalowsky A, Sandor Z, Lu CY. Improved outcome of cadaveric renal transplantation due to calcium channel blockers. Transplantation 1991;52:640-5.
61. Puig JM, Lloveras J, Oliveras A, Costa A, Aubia J, Masramon J. Usefulness of diltiazem in reducing the incidence of acute tubular necrosis in Euro-Collins-preserved cadaveric renal grafts. Transplant Proc 1991;23:2368-9.
62. Nicholson ML, Dennis MJ, Beckingham IJ, Smith SJ. Effect of nifedipine on renal transplant rejection. Br J Surg 1993;80:1318-21.
63. Ladefoged SD, Andersen CB. Calcium channel blockers in kidney transplantation. Clin Transplant 1994;8:128-33.
64. McAnena OJ, Austin O, O'Connell PR, Hederman WP, Gorey TF, Fitzpatrick J. Laparoscopic versus open appendicectomy: a prospective evaluation. Br J Surg 1992;79:818-20.
65. Buckley RC, Hall TJ, Muakkassa FF, Anglin B, Rhodes RS, Scott-Conner CE. Laparoscopic appendectomy: is it worth it? Am Surg 1994;60:30-4.
66. Lujan Mompean JA, Robles Campos R, Parrilla Paricio P, Soria Aledo V, Garcia Ayllon J. Laparoscopic versus open appendicectomy: a prospective assessment. Br J Surg 1994;81:133-5.
67. Schirmer BD, Schmieg RE Jr, Dix J, Edge SB, Hanks JB. Laparoscopic versus traditional appendectomy for suspected appendicitis. Am J Surg 1993;165:670-5.
68. Vallina VL, Velasco JM, McCulloch CS. Laparoscopic versus conventional appendectomy. Ann Surg 1993;218:685-92.
69. Bonanni F, Reed J III, Hartzell G, et al. Laparoscopic versus conventional appendectomy. J Am Coll Surg 1994;179:273-8.
70. Pruett B, Pruett J. Laparoscopic appendectomy: have we found a better way? J Miss State Med Assoc 1994;35:347-51.
71. Richards KF, Fisher KS, Flores JH, Christensen BJ. Laparoscopic appendectomy: comparison with open appendectomy in 720 patients. Surg Laparosc Endosc 1996;6:205-9.
72. Golub R, Siddiqui F, Pohl D. Laparoscopic versus open appendectomy: a metaanalysis. J Am Coll Surg 1998;186:545-53.
73. Hartz AJ, Kuhn EM, Pryor DB, et al. Mortality after coronary angioplasty and coronary artery bypass surgery (the national Medicare experience). Am J Cardiol 1992;70:179-85.
74. Hartz AJ, Kuhn EM. Coronary angioplasty compared with bypass grafting. N Engl J Med 1995;332:888.
75. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999;282:1054-60.
76. Gillman MW, Runyan DK. Bias in treatment assignment in controlled clinical trials. N Engl J Med 1983;310:1610-1.
77. Ottenbacher K. Impact of random assignment on study outcome: an empirical examination. Control Clin Trials 1992;13:50-61.
78. Emanuel EJ. Cost savings at the end of life: what do the data show? JAMA 1996;275:1907-14.
79. Green SB, Byar DP. Using observational data from registries to compare treatments: the fallacy of omnimetrics. Stat Med 1984;3:361-73.