



The NEW ENGLAND JOURNAL of MEDICINE

[HOME](#) | [SEARCH](#) | [CURRENT ISSUE](#) | [PAST ISSUES](#) | [COLLECTIONS](#) | [HELP](#)

Administration for Institution: JOHN HOPKINS UNIV | [Sign Out](#) | [Manage Account](#) | [FAQ](#)

EDITORIAL

[◀ Previous](#)

Volume 342:1907-1909

June 22, 2000

Number 25

[Next ▶](#)

Randomized Trials or Observational Tribulations?

The role of observational studies in the evaluation of treatments is a long-standing and contentious topic.¹ In this issue of the *Journal*, Concato et al.² and Benson and Hartz³ report that observational studies give results similar to those of randomized, controlled trials. If these claims lead to more observational studies of therapeutic interventions and fewer randomized, controlled trials, we see considerable dangers to clinical research and even to the well-being of patients.

Any systematic review of evidence on a therapeutic topic needs to take into account the quality of the evidence. Any study, whether randomized or observational, may have flaws in design or analysis. Both types of study may have quirks in methods of recruiting patients, in the clinical setting, or in the delivery of the treatment that can cast doubt on the generalizability of the results. And for some studies, the reports are never published at all, especially if the findings are negative. These problems of heterogeneity and publication bias are relevant to all comparisons of evidence from randomized, controlled studies and observational studies. However, all observational studies have one crucial deficiency: the design is not an experimental one. Each patient's treatment is deliberately chosen rather than randomly assigned, so there is an unavoidable risk of selection bias and of systematic differences in outcomes that are not due to the treatment itself. Although in data analysis one can adjust for identifiable differences, it is impossible to be certain that such adjustments are adequate or whether one has documented all the relevant characteristics of the patients. Only randomized treatment assignment can provide a reliably unbiased estimate of treatment effects.

So why does the evidence in the articles by Concato et al. and Benson and Hartz appear to contradict this commonly accepted wisdom? The question is difficult to answer. The findings in the two reports are based on 5 meta-analyses and 19 treatment analyses, respectively, and cover a wide range of therapeutic questions. The descriptions of the analyses are quite condensed, so it is not possible to check the details without going back to the original sources. However, our reading of a few of the examples in the studies

COMMENTARY

▶ [Letters](#)

TOOLS & SERVICES

- ▶ [Add to Personal Archive](#)
- ▶ [Add to Citation Manager](#)
- ▶ [Notify a Friend](#)
- ▶ [E-mail When Cited](#)

MORE INFORMATION

- ▶ [Related Article](#)
by Benson, K.
- ▶ [Related Article](#)
by Concato, J.
- ▶ [Find Similar Articles](#)
- ▶ [PubMed Citation](#)

casts doubt on the validity of the conclusion that observational studies produce the same answers, in general, as randomized, controlled trials.

Screening for breast cancer is one of the five issues examined by Concato et al.² They do not cite individual studies but, instead, refer to a meta-analysis that found that mammography had a strong protective effect against death from breast cancer in both observational studies and randomized, controlled trials of breast-cancer screening. This meta-analysis made no allowance for the quality of the studies it incorporated. Such an approach contrasts sharply with a more recent meta-analysis⁴ of the relevant high-quality randomized, controlled trials on the topic. The authors of that meta-analysis concluded that there was little evidence of benefit of screening for breast cancer. Hence, the estimates of the magnitude of the effect of screening in case-control studies and good randomized, controlled trials are, in fact, very different.

Other medical topics covered in the study by Concato et al.² include the influence of treatment for high blood pressure on the risk of stroke and coronary heart disease. The observational studies they reported were general population studies that included no data on antihypertensive treatment. Thus, the only evidence that antihypertensive drugs reduce the risk of stroke and heart disease comes from randomized, controlled trials.

The study by Benson and Hartz³ summarizes the evidence from observational studies and randomized, controlled trials for 18 separate therapeutic comparisons (7 cardiologic and 11 noncardiologic) in just four paragraphs of text and two figures. Three further topics have then been selected for more detailed presentation (a selection bias?). Such lack of detail regarding most of the topics means that readers have no information on which to agree with or reject the authors' findings.

The issue of the comparative merits of coronary-artery bypass grafting and percutaneous transluminal coronary angioplasty is a case in point. The classification of patients into high-risk and low-risk groups is not explained. The observational studies appear to show that coronary-artery bypass grafting is inferior in low-risk patients, although the randomized, controlled trials do not. Furthermore, the data on diabetic patients are from a post hoc subgroup analysis in the Bypass Angioplasty Revascularization Investigation (BARI). Subgroup analyses often exaggerate differences between treatments in randomized, controlled trials,⁵ so this comparison of the types of studies may suffer from a double dose of selection bias, with sources of bias in each type of study design.

The first example discussed by Benson and Hartz, the use of nifedipine in patients with coronary artery disease, reminds one of the fact that observational research on calcium-channel antagonists has provoked concern about the safety of these drugs. All such claims from observational studies have been hotly disputed, and the available evidence from randomized, controlled trials is as yet insufficient for definitive conclusions to be drawn. Thus, there is all the more reason to require additional evidence from major randomized, controlled trials, such as the ACTION trial,⁶ which compares nifedipine and placebo in coronary heart disease.

Several of Benson and Hartz's examples, such as the Coronary Artery Surgery Study and BARI, nicely

illustrate the merits of enhancing a randomized, controlled trial by the addition of observational data from a concurrent registry of all nonrandomized patients in the same centers. This approach improves the quality of observational research, since the same rigorous attention to detail in defining eligible patients, maintaining follow-up, and recording outcomes is applied in both the randomized and the observational cohorts. The observational cohort may still suffer from selection bias, but there is a greater likelihood that its causes can be identified.

It is likely that the studies used in both reports^{2,3} are a highly selected sample, since it is rarely sensible for a therapeutic question to be equally and simultaneously addressed by both experimentation and observation. This factor may explain why the authors refer to the overall paucity of therapeutic questions evaluated in both randomized, controlled trials and observational studies. Regulatory authorities appropriately require randomized, controlled trials as the prime acceptable evidence for drug licensing, and medical journals have become reluctant to publish claims for treatments based on observational data. Recent examples demonstrate that such concern is not merely theoretical. The Heart and Estrogen/Progestin Replacement Study,⁷ the Heart Outcomes Prevention Evaluation Study,⁸ and the Beta-Carotene and Retinal Efficacy Trial⁹ helped to restrain earlier observational claims that hormone-replacement therapy reduces the risk of coronary heart disease, vitamin E lowers the risk of cardiovascular disease, and beta carotene lowers the risk of lung cancer.

In some areas, such as monitoring for drug toxicity or studying risk factors for disease, we must rely on observational data. Detection of serious but rare side effects requires very large numbers of patients and can be achieved only through analysis of records from routine clinical practice. For some treatments, ethical considerations, practicality, clinical judgment, or unwillingness on the part of patients make randomized trials unrealistic, especially when the alternative treatments are radically different — as in the case of surgery and medical management. Problems also arise if the investigators themselves simply refuse to accept randomization. For instance, a randomized, controlled trial comparing carotid-artery stenting and carotid endarterectomy has been planned and funded by the National Institutes of Health.¹⁰ Its success depends on participating surgeons' and interventionists' allowing randomization of their eligible, consenting patients. If, instead, established patient-referral patterns continue, with each specialty treating whoever comes its way, selection bias may confound real differences in treatment. In general, there is insufficient evaluation by randomized, controlled trials of medical devices and surgical procedures.

When recruitment of patients for a randomized, controlled trial is exceptionally difficult, threatening to make the sample of patients unrepresentative, neither reliance on randomized, controlled trials nor reliance on observational studies is wholly satisfactory.¹¹ Investigators should avoid premature judgments about therapeutic claims and overcome their prejudices against experimentation so that the necessary randomized, controlled trials can be performed. Making the trials as much like routine practice as possible may help to make randomized, controlled trials more feasible and their results more widely generalizable.

Observational data bases can be useful adjuncts to randomized, controlled trials, to see whether efficacy

under controlled conditions in specialist centers translates into effective treatment in routine practice. A recent investigation of antiretroviral therapy for human immunodeficiency virus infection¹² concluded that bias in observational research does exist but is not inevitable, especially if the factors used for adjustment strongly predict the outcome. Furthermore, observational studies are useful guides to the design of new controlled trials.

Surely, society expects us to evaluate new healthcare interventions by the most scientifically sound and rigorous methods available. Although observational studies often are cheaper, quicker, and less difficult to carry out, we should not lose sight of one simple fact: ignorance calls for careful experimentation. This means high-quality randomized, controlled trials, not observations that reflect personal choices and beliefs. The key is to persuade physicians, patients, researchers, and study sponsors to collaborate on major trials, with the necessary patients, treatments, outcome measures, sample sizes, and follow-up to resolve therapeutic uncertainty properly. In some areas this is difficult, but perhaps we have not tried hard enough to convert the skeptics. Evaluation in health care has become much more stringent than it once was. Rather than turn the clock back, we should build on these efforts, so as to serve future patients with the most reliable information possible.

Stuart J. Pocock, Ph.D.

Diana R. Elbourne, Ph.D.

London School of Hygiene and Tropical Medicine

London WC1E 7HT, United Kingdom

References

1. Byar DP. Why data bases should not replace randomized clinical trials. *Biometrics* 1980;36:337-342. [\[Medline\]](#)
2. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000;342:1887-1892. [\[Abstract/Full Text\]](#)
3. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 2000;342:1878-1886. [\[Abstract/Full Text\]](#)
4. Gøtzsche P, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;355:129-134. [\[CrossRef\]](#)[\[Medline\]](#)
5. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355:1064-1069. [\[CrossRef\]](#)[\[Medline\]](#)
6. Lubsen J, Poole-Wilson PA, Pocock SJ, et al. Design and current status of ACTION: a coronary disease trial investigating outcome with nifedipine GITS. *Eur Heart J* 1998;19:Suppl I:120-132. [\[Medline\]](#)
7. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. *JAMA* 1998;280:605-613. [\[Abstract/Full Text\]](#)
8. The Heart Outcomes Prevention Evaluation Study Investigators. Vitamin E supplementation and cardiovascular events in high-risk patients. *N Engl J Med* 2000;342:154-160. [\[Abstract/Full Text\]](#)
9. Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of beta carotene and vitamin A on lung cancer and cardiovascular medicine. *N Engl J Med* 1996;334:1150-1155. [\[Abstract/Full Text\]](#)
10. Hobson RW II, Brott T, Ferguson R, et al. CREST: carotid revascularization endarterectomy

versus stent trial (editorial). *Cardiovasc Surg* 1997;5:457-458. [\[CrossRef\]](#) [\[Medline\]](#)

11. Black N, Brazier J, Fitzpatrick R, Reeves B. Health service research methods: a guide to better practice. London: BMJ Books, 1998.
12. Phillips AN, Grabar S, Tassie J-M, Costagliola D, Lundgren JD, Egger M. Use of observational databases to evaluate effectiveness of antiretroviral therapy for HIV infection: comparison of cohort studies with randomized trials. *AIDS* 1999;13:2075-2082. [\[CrossRef\]](#) [\[Medline\]](#)

Related Letters:

Observational Studies and Randomized Trials

Kunz R., Khan K. S., Neumayer H.-H., Sacks H. S., Liu P.-Y., Anderson G., Crowley J. J., Friedman H. S., Smith R. P., Meier P., Benson K., Hartz A. J., Concato J., Shah N., Horwitz R. I.

[Extract](#) | [Full Text](#)

N Engl J Med 2000; 343:1194-1197, Oct 19, 2000.

Correspondence

This article has been cited by other articles:

- Wieand, S., Murphy, K. (2004). A Commentary on Treatment at Random: The Ultimate Science or the Betrayal of Hippocrates?. *J Clin Oncol* 22: 5009-5011 [\[Full Text\]](#)
- Retsas, S. (2004). Treatment at Random: The Ultimate Science or the Betrayal of Hippocrates?. *J Clin Oncol* 22: 5005-5008 [\[Full Text\]](#)
- Pedone, C., Pahor, M., Carosella, L., Bernabei, R., Carbonin, P., for the GIFA Investigators, (2004). Use of Angiotensin-Converting Enzyme Inhibitors in Elderly People With Heart Failure: Prevalence and Outcomes. *J Gerontol A Biol Sci Med Sci* 59: M716-M721 [\[Abstract\]](#) [\[Full Text\]](#)
- Couvret, C., Laffon, M., Baud, A., Payen, V., Burdin, P., Fusciardi, J. (2004). A Restrictive Use of Both Autologous Donation and Recombinant Human Erythropoietin Is an Efficient Policy for Primary Total Hip or Knee Arthroplasty. *Anesth Analg* 99: 262-271 [\[Abstract\]](#) [\[Full Text\]](#)
- Bernard, G. R., Margolis, B. D., Shannies, H. M., Ely, E. W., Wheeler, A. P., Levy, H., Wong, K., Wright, T. J. (2004). Extended Evaluation of Recombinant Human Activated Protein C United States Trial (ENHANCE US): A Single-Arm, Phase 3B, Multicenter Study of Drotrecogin Alfa (Activated) in Severe Sepsis. *Chest* 125: 2206-2216 [\[Abstract\]](#) [\[Full Text\]](#)
- Siderowf, A. D. (2004). Evidence from Clinical Trials: Can We Do Better?. *neurorx* 1: 363-371 [\[Abstract\]](#) [\[Full Text\]](#)
- Concato, J. (2004). Observational Versus Experimental Studies: What's the Evidence for a Hierarchy?. *neurorx* 1: 341-347 [\[Abstract\]](#) [\[Full Text\]](#)
- Tandon, P., McAlister, F. A., Tsuyuki, R. T., Hervas-Malo, M., Dupuit, R., Ezekowitz, J., Cujec, B., Armstrong, P. W. (2004). The Use of {beta}-Blockers in a Tertiary Care Heart Failure Clinic: Dosing, Tolerance, and Outcomes. *Arch Intern Med* 164: 769-774 [\[Abstract\]](#) [\[Full Text\]](#)
- Dewey, T. M., Mack, M. J. (2003). Myocardial Revascularization Without Cardiopulmonary Bypass. *Card Surg Adult* 2: 609-625 [\[Full Text\]](#)
- Mack, M. J., Pfister, A., Bachand, D., Emery, R., Magee, M. J., Connolly, M., Subramanian, V. (2004). Comparison of coronary bypass surgery with and without cardiopulmonary bypass in patients with multivessel disease. *J Thorac Cardiovasc Surg* 127: 167-173 [\[Abstract\]](#) [\[Full Text\]](#)
- Myrmel, T., Lai, D. T.M., Miller, D. C. (2004). Can the principles of evidence-based medicine be applied to the treatment of aortic dissections?. *Eur J Cardiothorac Surg* 25: 236-242 [\[Abstract\]](#) [\[Full Text\]](#)

COMMENTARY

▶ Letters

TOOLS & SERVICES

- ▶ Add to Personal Archive
- ▶ Add to Citation Manager
- ▶ Notify a Friend
- ▶ E-mail When Cited

MORE INFORMATION

- ▶ Related Article by Benson, K.
- ▶ Related Article by Concato, J.
- ▶ Find Similar Articles
- ▶ PubMed Citation

- Marks, H. M (2003). Rigorous uncertainty: why RA Fisher is important. *Int. J. Epidemiol.* 32: 932-937 [\[Full Text\]](#)
- Lyerly, A. D., Gates, E. A., Cefalo, R. C., Sugarman, J. (2001). Toward the Ethical Evaluation and Use of Maternal-Fetal Surgery. *Obstet Gynecol* 98: 689-697 [\[Abstract\]](#) [\[Full Text\]](#)
- Ray, W. A. (2003). Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. *Am. J. Epidemiol.* 158: 915-920 [\[Abstract\]](#) [\[Full Text\]](#)
- Koch, C. G., Mangano, C. M., Schwann, N., Vaccarino, V. (2003). Is it gender, methodology, or something else?. *J Thorac Cardiovasc Surg* 126: 932-935 [\[Full Text\]](#)
- Morris, C. D., Carson, S. (2003). Routine Vitamin Supplementation To Prevent Cardiovascular Disease: A Summary of the Evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 139: 56-70 [\[Abstract\]](#) [\[Full Text\]](#)
- Kirk, O., Reiss, P., Uberti-Foppa, C., Bickel, M., Gerstoft, J., Pradier, C., Wit, F. W., Ledergerber, B., Lundgren, J. D., Furrer, H. (2002). Safe Interruption of Maintenance Therapy against Previous Infection with Four Common HIV-Associated Opportunistic Pathogens during Potent Antiretroviral Therapy. *Ann Intern Med* 137: 239-250 [\[Abstract\]](#) [\[Full Text\]](#)
- Vandenbroucke, J. P., de Craen, A. J.M. (2001). Alternative Medicine: A "Mirror Image" for Scientific Reasoning in Conventional Medicine. *Ann Intern Med* 135: 507-513 [\[Abstract\]](#) [\[Full Text\]](#)
- Maggioni, A P, Sinagra, G, Opasich, C, Geraci, E, Gorini, M, Gronda, E, Lucci, D, Tognoni, G, Balli, E, Tavazzi, L (2003). Treatment of chronic heart failure with β adrenergic blockade beyond controlled clinical trials: the BRING-UP experience. *Heart* 89: 299-305 [\[Abstract\]](#) [\[Full Text\]](#)
- Gum, P. A., Thamilarasan, M., Watanabe, J., Blackstone, E. H., Lauer, M. S. (2001). Aspirin Use and All-Cause Mortality Among Patients Being Evaluated for Known or Suspected Coronary Artery Disease: A Propensity Analysis. *JAMA* 286: 1187-1194 [\[Abstract\]](#) [\[Full Text\]](#)
- Ioannidis, J. P. A., Haidich, A.-B., Pappa, M., Pantazis, N., Kokori, S. I., Tektonidou, M. G., Contopoulos-Ioannidis, D. G., Lau, J. (2001). Comparison of Evidence of Treatment Effects in Randomized and Nonrandomized Studies. *JAMA* 286: 821-830 [\[Abstract\]](#) [\[Full Text\]](#)
- Bell, I. R., Caspi, O., Schwartz, G. E. R., Grant, K. L., Gaudet, T. W., Rychener, D., Maizes, V., Weil, A. (2002). Integrative Medicine and Systemic Outcomes Research: Issues in the Emergence of a New Model for Primary Health Care. *Arch Intern Med* 162: 133-140 [\[Abstract\]](#) [\[Full Text\]](#)
- Licker, M., Khatchaturian, G., Schweizer, A., Bednarkiewicz, M., Tassaux, D., Chevalley, C. (2002). The Impact of a Cardioprotective Protocol on the Incidence of Cardiac Complications After Aortic Abdominal Surgery. *Anesth Analg* 95: 1525-1533 [\[Abstract\]](#) [\[Full Text\]](#)
- Brott, T. G. (2002). Angioplasty and Stenting Should Be Performed Only in the Setting of a Clinical Trial. *Stroke* 33: 2519-2520 [\[Full Text\]](#)
- Bigby, M. (2001). Challenges to the Hierarchy of Evidence: Does the Emperor Have No Clothes?. *Arch Dermatol* 137: 345-346 [\[Full Text\]](#)
- Bergh, C., Brodin, U., Lindberg, G., Sodersten, P. (2002). Randomized controlled trial of a treatment for anorexia and bulimia nervosa. *Proc. Natl. Acad. Sci. U. S. A.* 99: 9486-9491 [\[Abstract\]](#) [\[Full Text\]](#)
- Licker, M., Spiliopoulos, A., Frey, J.-G., Robert, J., Hohn, L., de Perrot, M., Tschopp, J.-M. (2002). Risk Factors for Early Mortality and Major Complications Following Pneumonectomy for Non-small Cell Carcinoma of the Lung*. *Chest* 121: 1890-1897 [\[Abstract\]](#) [\[Full Text\]](#)
- Petkova, E., Teresi, J. (2002). Some Statistical Issues in the Analyses of Data From Longitudinal Studies of Elderly Chronic Care Populations. *Psychosom Med* 64: 531-547 [\[Abstract\]](#) [\[Full Text\]](#)
- Bleakley, C, MacAuley, D (2002). The quality of research in sports journals. *Br J Sports Med* 36: 124-125 [\[Abstract\]](#) [\[Full Text\]](#)
- Ernst, E (2002). What's the point of rigorous research on complementary/alternative medicine?. *J R Soc Med* 95: 211-213 [\[Full Text\]](#)
- Shulman, R J (2002). Effect of enteral administration of insulin on intestinal development and

- feeding tolerance in preterm infants: a pilot study. *Arch. Dis. Child. Fetal Neonatal Ed.* 86: F131-133 [\[Abstract\]](#) [\[Full Text\]](#)
- Stables, R H (2002). Observational research in the evidence based environment: eclipsed by the randomised controlled trial?. *Heart* 87: 101-102 [\[Full Text\]](#)
 - Pugh, P. J., O'Toole, L., Channer, K. S., Taggart, D. P., Banning, A., Channon, K., Mack, M. J., Murtaza, M., Singh, M., Dharmarajan, L., Serruys, P. W. (2001). Coronary-Artery Bypass Surgery versus Stenting for Multivessel Disease. *N Engl J Med* 345: 1641-1643 [\[Full Text\]](#)
 - Exner, D. V., Klein, G. J., Prystowsky, E. N. (2001). Primary Prevention of Sudden Death With Implantable Defibrillator Therapy in Patients With Cardiac Disease: Can We Afford to Do It? (Can We Afford Not To?). *Circulation* 104: 1564-1570 [\[Full Text\]](#)
 - McMahon, A. D. (2001). Observation and Experiment with the Efficacy of Drugs: A Warning Example from a Cohort of Nonsteroidal Anti-inflammatory and Ulcer-healing Drug Users. *Am. J. Epidemiol.* 154: 557-562 [\[Abstract\]](#) [\[Full Text\]](#)
 - Jacobson, J. S., Troxel, A. B., Evans, J., Klaus, L., Vahdat, L., Kinne, D., Lo, K. M. S., Moore, A., Rosenman, P. J., Kaufman, E. L., Neugut, A. I., Grann, V. R. (2001). Randomized Trial of Black Cohosh for the Treatment of Hot Flashes Among Women With a History of Breast Cancer. *J Clin Oncol* 19: 2739-2745 [\[Abstract\]](#) [\[Full Text\]](#)
 - Ioannidis, J. P A, Haidich, A.-B., Lau, J. (2001). Any casualties in the clash of randomised and observational evidence?. *BMJ* 322: 879-880 [\[Full Text\]](#)
 - Kunz, R., Khan, K. S., Neumayer, H.-H., Sacks, H. S., Liu, P.-Y., Anderson, G., Crowley, J. J., Friedman, H. S., Smith, R. P., Meier, P., Benson, K., Hartz, A. J., Concato, J., Shah, N., Horwitz, R. I. (2000). Observational Studies and Randomized Trials. *N Engl J Med* 343: 1194-1197 [\[Full Text\]](#)
 - Redwine, D, Mann, C H, Wright, J T, Farquhar, C. (2000). Evidence on endometriosis. *BMJ* 321: 1077a-1077 [\[Full Text\]](#)
 - Barton, S. (2000). Which clinical studies provide the best evidence?. *BMJ* 321: 255-256 [\[Full Text\]](#)
-

[HOME](#) | [SEARCH](#) | [CURRENT ISSUE](#) | [PAST ISSUES](#) | [COLLECTIONS](#) | [HELP](#)

Comments and questions? Please [contact us](#).

The New England Journal of Medicine is owned, published, and [copyrighted](#) © 2005 [Massachusetts Medical Society](#). All rights reserved.