

Circulation

JOURNAL OF THE AMERICAN HEART ASSOCIATION



Lessons Learned From Recent Cardiovascular Clinical Trials: Part II

David L. DeMets and Robert M. Califf

Circulation 2002;106;880-886

DOI: 10.1161/01.CIR.0000023220.26465.89

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75214
Copyright © 2002 American Heart Association. All rights reserved. Print ISSN: 0009-7322. Online
ISSN: 1524-4539

The online version of this article, along with updated information and services, is
located on the World Wide Web at:

<http://circ.ahajournals.org/cgi/content/full/106/7/880>

Subscriptions: Information about subscribing to *Circulation* is online at
<http://circ.ahajournals.org/subscriptions/>

Permissions: Permissions & Rights Desk, Lippincott Williams & Wilkins, a division of Wolters
Kluwer Health, 351 West Camden Street, Baltimore, MD 21202-2436. Phone: 410-528-4050. Fax:
410-528-8550. E-mail:
journalpermissions@lww.com

Reprints: Information about reprints can be found online at
<http://www.lww.com/reprints>

Lessons Learned From Recent Cardiovascular Clinical Trials: Part II

David L. DeMets, PhD; Robert M. Califf, MD

In this second part of a 4-part series about clinical research, we will complete the lessons we have learned during the past 2 decades from our involvement in cardiovascular clinical trials. The final 2 articles will delve into principles we have derived from our lessons that are meant to help practicing clinicians incorporate good clinical trial results into their patient care.

Structural Issues in the Conduct of Trials

The present structure for administering multicenter clinical trials funded by the government was established by the Greenberg Report¹ and first implemented by the Coronary Drug Project.² Referred to as the National Institutes of Health (NIH) clinical trial model (Figure 1), this example became the standard for trials sponsored by the National Heart, Lung and Blood Institute (NHLBI) and by many other NIH institutes.³ Key components of this model include the sponsoring agency, steering committee, data coordinating center, and data monitoring committee (DMC). The steering committee usually is made up of a study chair and other selected (or elected) representatives from the investigators and sponsor. Steering committee members develop the protocol, lead the trial, and publish the results when the trial is completed. The data coordination center is responsible for the management and quality control of the trial data, as well as for interim or final analyses of the baseline, safety, and efficacy data.

The DMC monitors the trial for evidence of relative harm or convincing evidence of benefit. Most DMCs also track the trial's progress, adherence to protocol, and the quality of the data.⁴ Although outright fraudulent data are rare, the responsibility of assuring high-quality trial operations is widely shared but falls directly into the oversight mantle of the DMC and the institutional review board (IRB). Definitive large-scale, randomized trials with an irreversible outcome (death) and a serious morbidity outcome (myocardial infarction or stroke) are most likely to appoint a DMC. The US Secretary of Health and Human Services recently announced that all trials must have a monitoring plan and all supported the use of DMCs when a concern exists about irreversible outcomes.⁵ The US Food and Drug Administration (FDA) also recently published a draft guidance on DMC structure and function.⁶

IRBs are appointed by each research institution to review the ethics, the protocol's scientific soundness, the relevance of the intervention, and the patient consent process. IRBs also provide local oversight of the safety of patients. Recently, the US system of IRBs has been heavily criticized by the Inspector General,⁷ and several individual institutions have been sanctioned by the Office of Health Research Policies.^{8,9} The concerns initially centered on failure to adhere to established standards of review, and attention has recently shifted to redefining these standards so that sound quantitative principles of quality and trial design can be incorporated into ethical review.⁴

Over the past decade, most noncardiovascular trials sponsored by the medical products industry have not used the NIH clinical trial model. Instead, some have appointed company employees to oversee the trial. In other cases, they have hired a contract research organization, with little participation in or influence from representatives of clinical practice or the academic community, and often without an independent DMC. Cardiovascular disease trials have a much stronger tradition of independent input by the steering committee and DMC than trials investigating most other diseases. In fact, a modified version of the NIH clinical trial model is used frequently in cardiovascular trials sponsored by industry (Figure 2).¹⁰

The Swedish metoprolol trial in heart attack patients¹¹ and the Prospective Randomized Milrinone Survival Evaluation (PROMISE) trial of milrinone in congestive heart failure patients¹² were two of the first to use this model. Many other industry-sponsored trials have followed their lead¹³⁻¹⁵ because of the model's many benefits. First, academic investigators can provide, through the steering committee, considerable input into the design of the protocol and the leadership of the trial. Second, an independent DMC is essential for a trial to be kept masked to those involved in its conduct, thereby minimizing bias until convincing evidence for benefit or harm has emerged.

Another benefit of this model is the independent statistical analysis center's role: to provide support to the DMC during the trial and to the steering committee after the trial is completed, allowing the sponsor's statisticians to remain

From the Duke Clinical Research Institute and the Division of Cardiology (R.M.C.), Duke University Medical Center, Durham, NC, and the Department of Biostatistics and Medical Informatics (D.L.D.), University of Wisconsin, Madison, Wis.

This is Part II of a 4-part article. Part I appeared in the August 6, 2002, issue of the journal (*Circulation*. 2002;106:746-751), and Parts III and IV will appear in the August 20 and August 27, 2002, issues, respectively.

Correspondence to Robert M. Califf, MD, Duke Clinical Research Institute, PO Box 17969, Durham, NC 27715. E-mail calif001@mc.duke.edu (*Circulation*. 2002;106:880-886.)

© 2002 American Heart Association, Inc.

Circulation is available at <http://www.circulationaha.org>

DOI: 10.1161/01.CIR.0000023220.26465.89

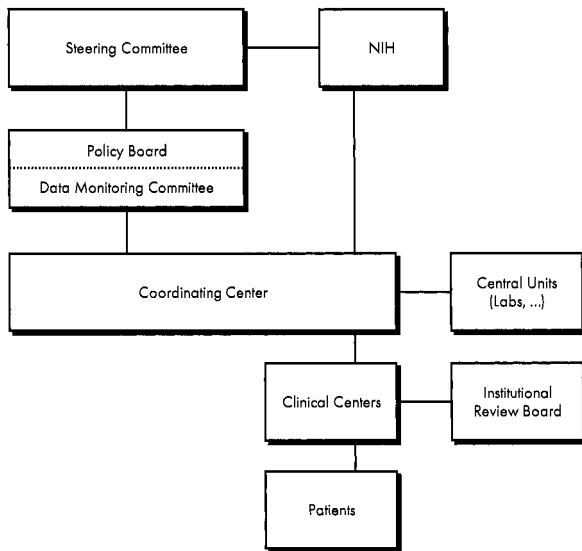


Figure 1. Clinical trial model that evolved through experience at the National Institutes of Health. Reprinted with permission from Fisher M, Roecker E, DeMets D. The role of an independent statistical analysis center in the industry-modified National Institutes of Health model. *Drug Inf J.* 2001;35:115–129.¹⁰

masked until the data have been locked (finalized) and are ready for regulatory submission. Once a trial has been completed, the industry sponsor often focuses resources on preparing registration documents for regulatory review. The independent statistical analysis center can then concentrate on the academic needs of the investigators, although the center may also participate in some of the regulatory documentation relating to the primary analysis. Overall, the success of the NIH model and the industry-modified model provides a supportive structure for conducting clinical trials in cardiovascular medicine on the cutting edge. As public scrutiny of human experimentation continues to heighten, recognition of

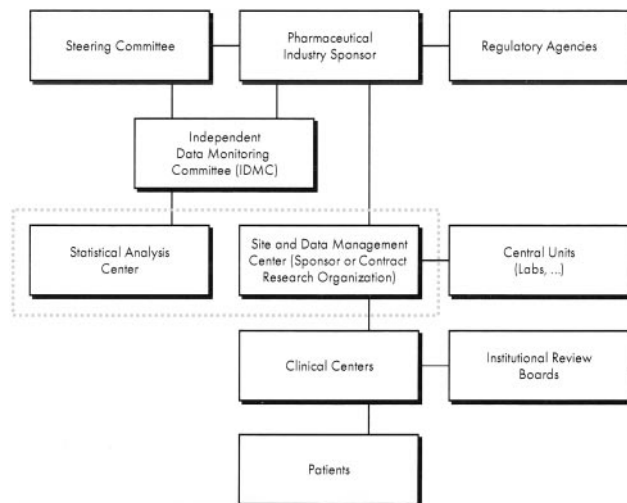


Figure 2. Modified NIH clinical trial model, which includes an independent statistical center. Reprinted with permission from Fisher M, Roecker E, DeMets D. The role of an independent statistical analysis center in the industry-modified National Institutes of Health model. *Drug Inf J.* 2001;35:115–129.¹⁰

the unique and valuable perspectives of all elements, including the sponsor, investigator, IRB, and DMC, will be increasingly critical.

Minimizing Bias

Some recent clinical trials have emphasized the importance of assuring the absence of bias when assessing the effectiveness of therapies. For example, several nonblinded trials investigated transmyocardial laser revascularization of the myocardium and showed a benefit in improving angina status.^{16,17} When the first double-blind trial was done (by use of a sham procedure), no benefit was seen and there was a trend toward excess adverse outcomes (M.B. Leon, unpublished observations). Despite the fact that a series of unblinded trials had shown dramatic improvement in exercise time for patients with angina and peripheral arterial disease after exogenous administration of vascular growth factors, the first substantial double-blind, placebo-controlled trial showed no benefit. Although negative trials do not completely disprove the previously touted benefits of either transmyocardial laser revascularization or growth factors, the superior design of the randomized trials raises significant questions that future trials must address.

Maintaining the blind is also important for managing the data for a clinical trial. Over the past several decades, industry has become a major sponsor of cardiovascular clinical trials. To ensure the absence of bias, many industry-sponsored cardiovascular trials have adopted and modified the NIH clinical trial model^{8,18} (Figure 2). One important modification divides the responsibilities of the data-coordinating center between an independent statistical analysis center and a data management center, the latter often being a contract research organization, an academic research organization, or a data management group internal to the industry sponsor. One site, of course, could still be both the statistical analysis center and the data management center. However, the key components, the steering committee and the independent DMC, remain a part of the model.

Who Is Responsible for Monitoring?

Despite the success of the clinical trial model, the lines of communication and responsibility among the DMCs, IRBs, sponsors, and government regulators have not been well established.¹⁹ For example, many early phase trials do not have a formal DMC. When an early phase trial is a single-institution study, the IRB must serve as the monitor of patient safety. For multicenter trials, problems arise because most IRBs lack the resources needed for the intensity of monitoring patient safety and evaluating evidence of early efficacy.

Current federal regulations require that all serious adverse events be reported to the local IRB. The majority of serious adverse events are defined as adverse events that are life threatening or require hospitalization. For multicenter trials, the large amount of serious adverse event data can overwhelm the resources of most IRB offices. Without true information about the denominator for the events and a balancing view of the potential benefits of therapy, there is little credible action that a local IRB can take on the basis of individual adverse events. For multicenter trials or even local

trials with a properly constituted DMC, the requirement to send all serious adverse event reports to the local IRB seems duplicative and unnecessary. Because of the overwhelming amount of needless work, IRBs find it difficult to focus on trials with no DMCs, although such trials would benefit most from more attentive monitoring by the IRB. Careful reconsideration of IRB responsibilities in multicenter trials that have a DMC seems warranted, and fortunately for large cardiovascular trials, the cardiorenal branch of the FDA has routinely both allowed and encouraged a markedly abbreviated approach to adverse event reporting.²⁰

New Ethical Mandates

Recent controversy has arisen about conflicts of interest in the conduct of human experimentation. Those who are treating patients in clinical trials must be vigilant about avoiding overt financial conflict by not owning equity in the companies of products they are evaluating. Concern is also emerging about financial conflict created by payments to investigators that exceed standard reimbursement levels for the work completed²¹ or that provide a bounty for enrolling specific patients.²² These excess payments are most often encountered in trials under time pressures to meet the financial objectives of commercial sponsors. Consideration is needed for standards of payments for the work involved in conducting human experiments (clinical trials). A major statement recently produced by the Association of American Medical Colleges provides a detailed review of these evolving standards.²³

Negative Trends (or Flexibility With Negative Trends)

Monitoring trials with emerging negative trends presents a major challenge.²⁴ Most trials are designed to compare best available care with or without the experimental intervention. Sometimes the intervention is a new drug or procedure, although in other cases it may be the new use of an existing intervention. Trials also evaluate interventions that are being used to some extent but lack definitive or adequate evidence to support their widespread use. When negative trends favoring the control or standard of care become evident, the DMC must keep in mind the current status of the experimental intervention. If the intervention is a new drug or procedure, then less harmful or negative evidence may be required to terminate the trial. If such a trial is not going to show that the intervention is superior to the current standard of care, then there may be cause to terminate the trial and spare the experimental subjects from continued exposure to an intervention providing no apparent benefit. In other words, the degree of evidence to make this judgment would not be symmetrical to what might be required to judge a new intervention superior to conventional intervention.

However, other situations may call for more substantial evidence before terminating a trial with an emerging negative trend. Examples include when the intervention is already in use (perhaps on the basis of surrogate evidence of benefit or just by opinion) or when data suggest the intervention to be beneficial but not all interested parties are convinced. In the Cardiac Arrhythmia Suppression Trial (CAST),²⁵ the antiar-

rhythmic drugs being evaluated were already in widespread use because conventional medical opinion was based on the effect of these therapies on an invalid surrogate. Thus, although the termination of CAST was very rapid by the DMC, the trial went beyond a point where the results were so negative that a positive result was highly unlikely. Simply proving that the drugs did not reduce mortality may have been insufficient to call for discontinuing their use in patients with cardiac arrhythmias; a clear demonstration of increased mortality was probably necessary. On the other hand, allowing the trial to proceed to this extent would have been pointless for an experimental therapy not already in clinical use.

Another example of this needed flexibility is provided by the PROMISE trial,¹⁵ which assessed milrinone in patients with congestive heart failure. In earlier studies, milrinone had shown improvement in cardiac function for such patients. In PROMISE, patients with heart failure were randomly assigned to receive best available care with or without milrinone to evaluate the effect on total mortality and mortality plus hospitalization. A negative trend began to emerge early for both outcomes. The DMC allowed the trial to continue beyond the point where milrinone was unlikely to show a benefit in order for the investigators to distinguish between a neutral result, which would encourage the use of milrinone, and a harmful effect, which would discourage its use. The investigators later established that orally administered milrinone was significantly harmful compared with the standard of care in this patient population.

The Heart and Estrogen/Progestin Replacement Study (HERS) evaluated the benefits of hormone replacement therapy (HRT) on heart disease in postmenopausal women with a definite history of coronary heart disease.¹⁶ Before HERS, no randomized trials had been conducted to provide convincing evidence that HRT was beneficial for heart disease, although large observational studies (vulnerable to bias) found that women on HRT had a lower risk of heart disease than women not taking HRT. Despite this deficit in randomized evidence, the use of HRT continues to be extremely widespread. Women are prescribed HRT to relieve their postmenopausal symptoms and to prevent bone mineral loss, a risk factor for hip and vertebral fractures, although HRT also has not been proved to prevent fractures in a prospective clinical trial.

Against this background, the HERS trial developed an early negative trend (Figure 3). Even though it seemed unlikely that this negative trend would reverse itself and become strongly positive within the designated period of time, the DMC recommended the trial continue in order to determine if the trend would become even stronger or drift back toward neutrality. Because of the widespread use of HRT, definitive evidence was required to evaluate whether it caused harm or just failed to provide a cardiovascular benefit. Lack of benefit for heart disease might allow HRT to remain an attractive treatment because of its other beneficial effects; however, an established harmful effect would substantially alter the risk-to-benefit relationship.

The DMC allowed the investigators to publish a short communication²⁷ before the trial was completed, demonstrat-

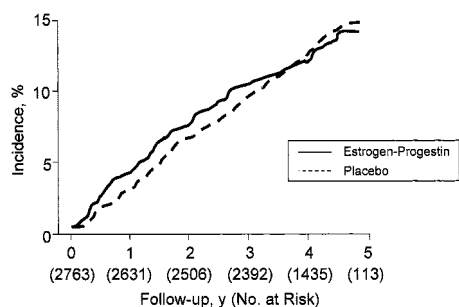


Figure 3. Long-term outcome from the Heart and Estrogen/Progestin Replacement Study. In the first several years, coronary heart disease (CHD) events—a composite of CHD death and nonfatal myocardial infarction—tended toward an adverse impact of hormone replacement therapy, as is demonstrated by the diverging curves; in the third and fourth years, this trend modified as the curves came back together. Reprinted with permission from Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease on postmenopausal women. *JAMA*. 1998;280:605–613.²⁶

ing that HRT caused a significant increase in deep vein thrombosis during the first year of treatment. As HERS continued, the negative trend in heart disease events began to reverse itself, but the results did not establish neutrality definitively. If the DMC had decided to terminate HERS when the results demonstrated lack of benefit and only suggested potential harm, there would have been no information on long-term use. Just this month, the Women's Health Initiative (WHI) reported that a similar excess risk was observed in the setting of primary prevention.²⁸

Before the Blockade of the glycoprotein IIb/IIIa Receptor to Avoid Vascular Occlusion (BRAVO) trial²⁹ began to investigate lotrafiban, an orally administered glycoprotein IIb/IIIa inhibitor, the DMC recognized that several similar compounds had failed because of an excess in mortality.^{30,31} Accordingly, the DMC, in concert with the sponsor and steering committee, developed an asymmetrical stopping rule that eventually led to the trial's early termination, long before it would have stopped with a symmetrical boundary.

Publication of Negative Trials

The previously described trials had negative outcomes, and the results were, in fact, published in leading journals. High-profile trials such as CAST²⁵ and PROMISE¹⁵ are of interest to the cardiology community, and editors of peer-reviewed journals will usually publish them regardless of the results. Less high-profile trials that do not show a significant benefit are less likely to be published, however, which leads to publication bias. Some researchers have argued for a registry of all clinical trials so that at least their existence is known and results can more easily be sought.³² Indeed, recently passed federal legislation calls for such a registry, which is presently being developed by the National Library of Medicine.³³

In some cases, practical issues limit the ability of a trial's results to be published. The Prospective Randomized Flosequinan Longevity Evaluation (PROFILE) trial, which investigated flosequinan in the treatment of chronic heart failure,

was terminated early by its DMC because of a significantly harmful mortality effect. This occurred despite highly favorable short-term effects on both cardiac function and quality of life.³⁴ Soon afterward, the sponsor closed its US facility and severely limited the cleanup of the data and access to it. An abstract by the investigators and a publication by a senior staff member of the US FDA provided the majority of the information about the design and primary outcome. Even though there was an independent statistical analysis center, the amount of data transferred to this center was inadequate for a typical scientific publication.

The Flolan International Randomized Survival Trial (FIRST) evaluated epoprostenol in the treatment of severe heart failure. Despite this prostacyclin analogue's appearing to have a highly beneficial effect on hemodynamics in early phase trials, FIRST showed a definitive adverse effect on mortality. Funding for the project was soon lost, and the sponsor never completed the database. The steering committee was able to obtain the incomplete database and publish the findings.³⁵

An AIDS vaccine trial,³⁶ sponsored by a small biotech company, offers a more extreme case with serious consequences. The DMC recommended the trial be stopped because of harm, but the sponsor was not anxious to have the results published and so delayed finalizing the data file. The steering committee and the study chair decided to publish the incomplete data that they did have. The sponsor then took legal action against the principal investigator, claiming that the publication did harm to the company.³⁷ Although it is not clear how this case will be resolved, it is likely that if it is settled in favor of the sponsor, the publication of all future industry-sponsored negative trial results will remain in doubt. This would be a severely negative blow to a necessary and advantageous partnership between academia and industry. If the modified NIH clinical trial model described earlier was followed, academia, industry and, most importantly, patients would benefit. In addition, this partnership would encourage more therapies to be developed in a proper context that would assure the public and regulators and allow findings to be passed on to patient care more rapidly.

In an effort to assure that the results of clinical research reach their intended audience, the editors of some of the world's leading peer-reviewed journals issued new rules in mid-2001 that established stricter standards over the control and publication of trial results.³⁸ Authors of such manuscripts will be required to disclose the details of the role they and the sponsor played in the trial; in addition, most journals will require the primary author to take responsibility, in writing, for the conduct of the trial, and to assert that he or she had full access to the data for independent analyses and made the decision to publish the results.

Noninferiority Trials

Noninferiority trials are designed to determine whether, when 2 treatments are compared, one is not worse than the other by a prespecified amount when given in conjunction with best available care. If the new treatment is not judged to be inferior but either the same or better, then the new intervention might be used because of some other advantage, such as

less toxicity, less invasiveness, easier administration, or lower cost. Formally, this goal requires the design of the trial to have adequate power to detect any prespecified differences judged to be clinically important. The results of these trials are summarized by estimating the treatment differences, or relative risks, with the appropriate confidence intervals. This is a direct comparison and inference.

A second desirable, but more challenging, inference is that the new intervention would have been better than the standard of care without either intervention. For example, would the new drug be better than placebo if a placebo arm had been included? Because a placebo arm was not included, any such inferences about a new treatment to placebo must be indirect by use of other data. This indirect approach is often based on a meta-analysis. But if weaker and weaker treatments are used for the comparison or control, almost any new treatment can be shown to be noninferior.

Despite these challenges, noninferiority trials have been conducted, including the Assessment of the Safety and Efficacy of a New Thrombolytic (ASSENT-2) trial³⁹ and the Bypass Angioplasty Revascularization Investigation (BARI).⁴⁰ ASSENT-2 compared tenecteplase and alteplase as reperfusion therapy for ST-elevation myocardial infarction, and BARI compared bypass surgery and angioplasty in multivessel disease. In the ASSENT-2 trial, the minimally important clinical difference was clearly defined before the trial started, and the criteria were met to declare noninferiority. In contrast, the Global Utilization of Strategies to Open Occluded Coronary Arteries (GUSTO-III) trial⁴¹ was designed as a superiority trial. When reteplase was not found to be superior to alteplase, a controversy arose as to whether the results of the trial could be interpreted as showing noninferiority or whether it was simply a failed superiority trial.⁴²

The recently reported Do Tirofiban and ReoPro Give Similar Efficacy Outcomes Trial (TARGET)⁴³ provides a reminder that a properly designed noninferiority trial should be adequately powered to show superiority if such superiority exists. TARGET was designed to show that tirofiban, a less expensive glycoprotein IIb/IIIa inhibitor, was not inferior to the more expensive abciximab in patients undergoing percutaneous coronary intervention. The final result demonstrated the superiority of abciximab over tirofiban for the primary end point. Although longer-term results have not maintained statistical significance in the TARGET trial, the fundamental point has been made—a well-designed noninferiority trial will reveal superiority of a treatment if it is better for the chosen outcomes.

Despite the difficulty in interpreting noninferiority trials, more such trials are needed. New therapies simply cannot be added to what is already available without concern for cost, compliance, and the possibility for unanticipated negative interactions. Directly comparative trials are needed so therapies that do not provide enough benefit can be discarded. In an increasing number of cases, the decision on which therapy to use may be made on the basis of cost, ease of use, or side effect profile. Decision makers want to be assured that the therapy they choose has not been proved inferior (with regard to effectiveness) to the therapy not chosen for the most important outcomes before they adopt it for ease of use, cost,

or minor side effect differences. The ASSENT-3 trial used a novel approach in which several thousand patients were entered on different complex therapeutic “cocktails” with a goal of accumulating data short of a definitive result so that choices could be made about a larger definitive clinical trial.⁴⁴ By use of a composite end point and confidence intervals, excess mortality can be excluded and insight into benefit can be gained, allowing targeted design of definitive outcome trials.

Confirmation Trials

The FDA and other global regulatory agencies traditionally require >1 adequately well-controlled trial to approve a new drug. Exceptions have been made for very large trials with mortality and serious morbidity outcomes in patient populations with life-threatening diseases. However, the criteria for allowing a single trial to be so influential have been discussed⁴⁵ and are subject to further review. The field of cardiology has had the benefit of learning from confirmatory trials both for specific compounds and devices and within classes of therapies.

Some confirmatory trials are nearly exact replications in design but may differ in the exact therapy being tested. Three recent trials investigating β -blockers^{16–18,46} were very similar in design except for the mixture of the severity of heart failure and the use of 3 different β -blockers (metoprolol, bisoprolol, and carvedilol). Although each trial addressed a slightly different population, the results were remarkably consistent, each reducing mortality by between 30% and 40%. Consistency was even found in subgroups, such as populations defined by NYHA class. This level of confirmation is fortunate for the cardiology community and for heart failure patients.

However, a major NIH trial, the β -Blocker Evaluation of Survival Trial (BEST),⁴⁷ did not confirm the same results with bucindolol, and the results were somewhat heterogeneous when placed in a systematic overview with the other trials. This result has sparked a controversy as to whether bucindolol is fundamentally different from other β -blockers or if the heterogeneity occurred because the power of study hampered the ability to show a difference, because the large proportion of African-American patients showed no effect on mortality, or because the trial simply had bad luck.

The previously discussed PRAISE I and PRAISE II trials illustrate that not all confirmatory trials will confirm the initial trial.^{48,49} A comparison of baseline risk factors and concomitant medications failed to provide any insight into why the second trial did not confirm the first.

In another example, the drug vesnarinone was evaluated in a small, randomized trial of patients with heart failure. The results suggested a nearly 50% reduction in mortality with this drug.⁵⁰ A second, much larger trial, the VESnarinone Trial (VEST),⁵¹ did not confirm the first trial. In fact, a 30% increase in mortality was observed in the dose that was common to the 2 trials, a stark contrast to the earlier results.

These examples reinforce the concept that the highest level of scientific proof comes from independent confirmation.

Specifying Primary and Secondary End Points

Several important cardiovascular trials have raised the issue of how to interpret the results of a trial or group of trials when the treatment is not proved to have an effect on the primary end point but an apparent result is seen on a secondary end point. In the US carvedilol trials, a series of trials evaluated the effect of carvedilol on exercise capacity.⁵² Although no clear result was observed with regard to this prespecified primary end point, a safety committee terminated the trials because of a striking reduction in mortality.

The FDA Cardiorenal Advisory Committee initially did not recommend approval of carvedilol for heart failure on the basis of these results because the primary end point was not significant. This recommendation sparked a spirited public debate, with some arguing that random selection of post-hoc end points would leave the public unprotected by use of therapies based on chance findings.⁵³ Others argued that the consistency across multiple trials and the importance of mortality provided enough evidence that the result had to be accepted. Eventually, a new panel recommended approval of the drug on the basis of additional evidence from another trial, and the Carvedilol Prospective Randomized Cumulative Survival (COPERNICUS) trial¹⁸ recently confirmed the mortality benefit in a comparison of carvedilol with placebo in severe heart failure patients.

In contrast, the Evaluation of Losartan in the Elderly (ELITE I) trial was designed to demonstrate better preservation of renal function in elderly patients treated with losartan, an angiotensin receptor blocker, compared with the ACE inhibitor captopril.⁵⁴ Although the primary end point was not significant, a nominally significant reduction in mortality was demonstrated with losartan (RR 46%, $P=0.035$). The ELITE II trial, an almost identical but larger trial, was constructed, but it showed a small trend toward higher mortality with losartan.⁵⁵

The major lesson from these experiences is that failure to find an effect in the primary end point of a trial need not dissuade investigators from examining secondary end points. Yet any positive findings must be regarded with suspicion, and confirmation should be sought from independent evidence. Another approach is to allocate the type I error of a trial to multiple end points so that a positive finding for any one end point is considered to be primary evidence.^{56,57}

With the completion of our lessons in cardiovascular medicine garnered from recent clinical research, we will address the application of those lessons to clinical medicine in the next 2 parts of this 4-part series.

Acknowledgments

This project was supported in part by an Agency for Healthcare Research and Quality (AHRQ) Centers for Education and Research on Therapeutics (CERTs) cooperative agreement, grant No. U18HS10548. The authors thank Penny Hodgson and Karen Staman for their help with the multiple revisions of these articles. They also express their appreciation to the thousands of collaborators who have made these observations possible by participating in the clinical trials effort.

References

1. Organization, review, and administration of cooperative studies (Greenberg Report): a report from the Heart Special Project Committee to

- the National Advisory Heart Council, May 1967. *Control Clin Trials*. 1988;9:137-148.
2. Coronary Drug Project report on clofibrate and niacin. *Atherosclerosis*. 1978;30:239-240.
3. Geller NL, Stylianou M. Practical issues in the data monitoring of clinical trials: summary of responses to a questionnaire at NIH. *Stat Med*. 1993; 2:543-552.
4. Morse MA, Califf RM, Sugarman J. Monitoring and ensuring safety during clinical research. *JAMA*. 2001;285:1201-1205.
5. Shalala D. Protecting research subjects: what must be done. *N Engl J Med*. 2000;343:808-810.
6. Department of Health and Human Services, Food and Drug Administration. Draft "Guidance for Clinical Trial Sponsors on the Establishment and Operation of Clinical Trial Data Monitoring Committees." In: *Federal Register*. November 20, 2001. Vol 66, No. 224, pp 58151-58153. Docket No. 01D-0489, CBER 200130 [FR Doc. 01-28962]. Available at: <http://www.fda.gov/OHRMS/DOCKETS/98fr/112001c.htm>. Accessed July 1, 2002.
7. Kaufman M. Clinical trials sanctions urged. *The Washington Post*. May 24, 2000: A02.
8. Snyderman R, Holmes EW. Oversight mechanisms for clinical research. *Science*. 2000;287:595-597.
9. Compliance Determination Letter from Office of Human Research Protections to Dean, and CEO, Johns Hopkins University School of Medicine. Office of Human Research Protections web site. US Department of Health and Human Services. Available at: http://ohrp.osophs.dhhs.gov/detrm_letters/jul01a.pdf. Accessed July 1, 2002.
10. Fisher M, Roecker EB, DeMets D. The role of an independent statistical analysis center in the industry-modified National Institutes of Health model. *Drug Inf J*. 2001;35:115-129.
11. Hjalmarson A, Goldstein S, Fagerberg B, et al. Effects of controlled-release metoprolol on total mortality, hospitalizations, and well-being in patients with heart failure: the Metoprolol CR/XL Randomized Intervention Trial in congestive heart failure (MERIT-HF). *JAMA*. 2000;283: 1295-1302.
12. Packer M, Carver JR, Rodeheffer RJ, et al. Effect of oral milrinone on mortality in severe chronic heart failure. The PROMISE Study Research Group. *N Engl J Med*. 1991;325:1468-1475.
13. The International Steering Committee. Rationale, design, and organization of the Metoprolol CR/XL Randomized Intervention Trial in Heart Failure (MERIT-HF). *Am J Cardiol*. 1997;80:541-58J.
14. The Cardiac Insufficiency Bisoprolol Study II (CIBIS-II): a randomized trial. *Lancet*. 1999;353:9-13.
15. Packer M, Coats AJS, Fowler MB, et al. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med*. 2001;344:1651-1658.
16. Cooley DA, Frazier OH, Kadipasaoglu KA, et al. Transmyocardial laser revascularization: clinical experience with twelve-month follow-up. *J Thorac Cardiovasc Surg*. 1996;111:791-799.
17. Frazier OH, March RJ, Horvath KA. Transmyocardial revascularization with a carbon dioxide laser in patients with end-stage coronary artery disease. *N Engl J Med*. 1999;341:1021-1028.
18. Fisher LD. Carvedilol and the Food and Drug Administration (FDA) approval process: the FDA paradigm and reflections on hypothesis testing. *Control Clin Trials*. 1999;20:16-39.
19. DeMets DL, Yusuf S. The data and safety monitoring committee: some final thoughts. *Am Heart J*. 2001;141:548-549.
20. Fisher L, Klibaner M. Regulatory issues for Data and Safety Monitoring Committees. *Am Heart J*. 2001;141:536-541.
21. Hillman AL, Pauly MV, Kerslein B. How do financial incentives affect physicians' clinical decisions and the financial performance of health maintenance organizations. *N Engl J Med*. 1989;321:86-92.
22. Smith R. Beyond conflict of interest. *BMJ*. 1998;317:291-292.
23. Association of American Medical Colleges Task Force on Financial Conflicts of Interest in Clinical Research. *Protecting Subjects, Preserving Trust, Promoting Progress: Task Force on Financial Conflicts of Interest in Clinical Research*. Available at: <http://www.aamc.org/members/coitf/firstreport.pdf>. Accessed July 1, 2002.
24. DeMets DL, Pocock SJ, Julian DG. The agonising negative trend in monitoring of clinical trials. *Lancet*. 1999;354:1983-1988.
25. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med*. 1989;321:406-412.

26. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease on postmenopausal women. *JAMA*. 1998;280:605–613.
27. Grady D, Wegner NK, Herrington D, et al. Postmenopausal hormone therapy increases risk for venous thromboembolic disease. The Heart and Estrogen/progestin Replacement Study. *Ann Intern Med*. 2000;132:689–696.
28. Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288:321–333.
29. Topol EJ, Easton JD, Amarenco P, et al. Design of the blockade of the glycoprotein IIb/IIIa receptor to avoid vascular occlusion (BRAVO) trial. *Am Heart J*. 2000;139:927–933.
30. The SYMPHONY Investigators. Comparison of sibrifiban with aspirin for prevention of cardiovascular events after acute coronary syndromes: a randomized trial. *Lancet*. 2000;355:337–345.
31. The PARAGON Investigators. International, randomized, controlled trial of lamifiban (a platelet glycoprotein IIb/IIIa inhibitor), heparin, or both in unstable angina. *Circulation*. 1998;97:2386–2395.
32. Horton R, Smith R. Time to register randomized trials. *BMJ*. 1999;319:865–866.
33. McCray AT, Ide NC. Design and implementation of a national clinical trials registry. *J Am Med Inform Assoc*. 2000;7:313–323.
34. Moe GW, Rouleau JL, Proulx G, et al, on behalf of the Canadian PROFILE investigators. Increased mortality with flosequinan in patients with failure is accompanied by increased plasma norepinephrine. *Circulation*. 1994;90:I-380.
35. Califf RM, Adams KF, McKenna WJ, et al. A randomized controlled trial of epoprostenol therapy for severe congestive heart failure: The Flolan International Randomized Survival Trial (FIRST). *Am Heart J*. 1997;134:44–54.
36. Kahn JO, Cherng DW, Mayer K, et al. Evaluation of HIV-1 Immunogen, an immunologic modifier, administered to patients infected with HIV having 300 to 542×106/L CD4 cell counts: a randomized trial. *JAMA*. 2000;284:2193–2202.
37. Hiltz PJ. Company tried to bar report that H.I.V. vaccine failed. *The New York Times*. November 1, 2000: A20.
38. Davidoff F, DeAngelis CD, Drazen JM, et al. Sponsorship, authorship, and accountability. *JAMA*. 2001;286:1232–1234.
39. Assessment of the Safety, and Efficacy of a New Thrombolytic Investigators. Single-bolus tenecteplase compared with front-loaded alteplase in acute myocardial infarction: the ASSENT-2 double-blind randomised trial. *Lancet*. 1999;354:716–722.
40. Writing Group for the Bypass Angioplasty Revascularization Investigation (BARI) Investigators. Five-year clinical and functional outcome comparing bypass surgery and angioplasty in patients with multivessel coronary disease: a multicenter randomized trial. *JAMA*. 1997;277:715–721.
41. The Global Use of Strategies to Open Occluded Coronary Arteries (GUSTO III) Investigators. A comparison of reteplase with alteplase for acute myocardial infarction. *N Engl J Med*. 1997;337:1118–1123.
42. Ware JH, Antman EM. Equivalence trials. *N Engl J Med*. 1997;337:1118–1123.
43. Moliterno DJ, Topol EJ. A direct comparison of tirofiban and abciximab during percutaneous coronary revascularization and stent placement: rationale and design of the TARGET study. *Am Heart J*. 2000;140:722–726.
44. The ASSENT-3 Investigators. Efficacy and safety of tenecteplase in combination with enoxaparin, abciximab, or unfractionated heparin: the ASSENT-3 randomized trial in acute myocardial infarction. *Lancet*. 2001;358:605–613.
45. Temple R. Current definitions of phases of investigation and the role of the FDA in the conduct of clinical trials. *Am Heart J*. 2000;139:S133–S135.
46. Sleight P. Subgroup analysis in clinical trials: fun to look at—but don't believe them. *Curr Controlled Trials Cardiovasc Med*. 2001;1:25–27.
47. The BEST Trial Investigators. A trial of the beta-blocker bucindolol in patients with advanced chronic heart failure. *N Engl J Med*. 2001;344:1659–1667.
48. Packer M, O'Connor CM, Ghali JK, et al. Effect of amlodipine on morbidity and mortality in severe chronic heart failure. Prospective Randomized Amlodipine Survival Evaluation Study Group. *N Engl J Med*. 1996;335:1107–1114.
49. Carson PA, O'Connor CM, Miller AB, et al. Circadian rhythm and sudden death in heart failure: results from Prospective Randomized Amlodipine Survival Trial. *J Am Coll Cardiol*. 2000;36:541–546.
50. Feldman AM, Bristow MR, Parmley WW, et al. Effects of vesnarinone on morbidity and mortality in patients with heart failure. Vesnarinone Study Group. *N Engl J Med*. 1993;329:149–155.
51. Cohn JN, Goldstein SO, Greenberg BH, et al. A dose-dependent increase in mortality with vesnarinone among patients with severe heart failure. Vesnarinone Trial Investigators. *N Engl J Med*. 1998;339:1810–1816.
52. Packer M, Bristow MR, Cohn JN, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. US Carvedilol Heart Failure Study Group. *N Engl J Med*. 1996;334:1349–1355.
53. Fisher LD, Moye LA. Carvedilol and the Food and Drug Administration approval process: an introduction. *Control Clin Trials*. 1999;20:1–15.
54. Pitt B, Segal R, Martinez FA, et al. Randomized trial of losartan versus captopril in patients over 65 with heart failure (Evaluation of Losartan in the Elderly Study, ELITE). *Lancet*. 1997;349:747–752.
55. Pitt B, Poole-Wilson PA, Segal R, et al. Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: rationale, design, and baseline characteristics of patients in the losartan heart failure survival study: ELITE II. *J Card Fail*. 1999;5:146–154.
56. Moye LA. Alpha calculus in clinical trials: considerations and commentary for the new millennium. *Stat Med*. 2000;19:767–779.
57. Moye LA. Random research. *Circulation*. 2001;103:3150–3153.

KEY WORDS: trials ■ cardiovascular diseases ■ therapy ■ outcome assessment ■ statistics