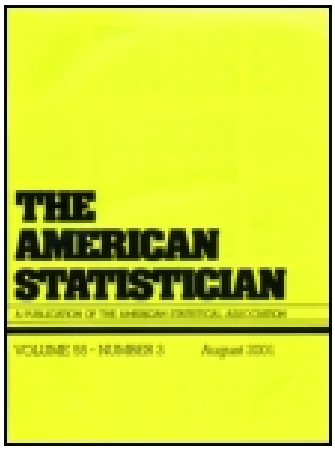

The American Statistician
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/utas20

The Influence of Biostatistics at the National Heart, Lung, and Blood Institute

David L. DeMets, Janet Turk Wittes & Nancy L. Geller

Published online: 29 Jun 2015.

To cite this article: David L. DeMets, Janet Turk Wittes & Nancy L. Geller (2015) The Influence of Biostatistics at the National Heart, Lung, and Blood Institute, The American Statistician, 69:2, 108-120, DOI: 10.1080/00031305.2015.1035962

To link to this article: http://dx.doi.org/10.1080/00031305.2015.1035962

The Influence of Biostatistics at the National Heart, Lung, and Blood Institute

David L. DeMets, Janet Turk Wittes, and Nancy L. Geller

Since the early 1950s, the National Heart, Lung, and Blood Institute (NHBLI) has conducted a long series of influential randomized clinical trials in heart, lung, and blood diseases. The biostatisticians at the Institute have been central to the design, conduct, monitoring, and final analyses of these trials. The uniquely favorable deck of cards the group of biostatisticians at the Institute has been dealt over the six and half decades of the group's life has led to contributions that have had a major impact on the fields of biostatistics and clinical trials. The leaders of the NHLBI and its several Divisions have valued the independence, creativity, and collaborative interactions of statisticians within the Institute. The medical problems the Institute faced impelled the statisticians to develop methodology that would address questions of great public importance. Perhaps most importantly, the individual members of the group had a collective vision passed from member to member over time that new methodology must fit the questions being asked. The group has always had the technical ability to develop new methods and the conviction that they were responsible for ensuring that they could explain their methods to the clinicians with whom they worked.

1. INTRODUCTION

This article discusses the contributions of the biostatisticians of the National Heart, Lung, and Blood Institute (NHLBI) with special emphasis on the development of methodology for clinical trials. We three authors represent the living subset of the six leaders of the biostatistics group. While the group had different names over the years (and is now known as The Office of Biostatistics Research), we shall simply refer to it here as the Branch.

In 1930, the Ransdell Act redesignated what was then the Laboratory of Hygiene, which had been created in 1887, as the National Institute of Health (NIH) (The Ransdell Act 1930). Our story of the Branch begins in the late 1940s after World War II when the NIH was still relatively new. An Office of Biometry, which was formed under the leadership of Harold Dorn (Figure 1), began to assemble a small group of statisticians. The first group Dorn recruited (Jerome Cornfield, Samuel Greenhouse, Marvin Schneiderman, Nathan Mantel, and Jacob Lieberman) were all housed in one large room in a temporary building at NIH, referred to as T-6 (Figure 2). All were new to the field of biomedical research but eager to learn and contribute. Not all were trained formally in statistics but had learned statistical principles while working for the Census Bureau or another Federal agency during World War II. Their tightly packed physical arrangement, which forced discussion among them, probably played a significant role in what was to follow.

At that time, most of the statistical problems emanated from laboratory research at the NIH. Discussions about the individual projects ensued as they were brought to these statisticians; the "discussions" often led to heated debates, but eventually some consensus arose among this group of statisticians as to how to address the problems statistically. Some questions related to study design, others to analysis. The group became well known for going to the cafeteria in Building I, which also housed the NIH Director's Office, to have lunch where their morning discussion continued, and continued, until they were often asked to leave to allow the cafeteria staff to go home. These interactions led to a progressive understanding of how statistics could influence and assist in the design and analysis of research being conducted at the NIH. The group became deeply engaged in the science; they believed that the problems they faced should drive the statistical methodology they needed to modify or develop; and they learned to convey statistical concepts and techniques to their laboratory colleagues and the NIH leadership. This early experience became the model for succeeding generations of biostatisticians at the NIH over the next six decades; it continues today (albeit in less cramped quarters).

In 1948, the National Heart Institute (NHI) was created along with a National Heart Advisory Council. The NHI started with an intramural program composed of nine sections that covered research, both basic and clinical, and cooperative research agreements with four major university medical centers. Around 1950, the Office of Biometry began to place their biostatistical team

David DeMets, Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI 53706 (E-mail: *demets@biostat.wisc.edu*). Janet Wittes is President, Statistics Collaborative, Inc., 1625 Massachusetts Ave., NW, Suite 600, Washington, DC 20036 (E-mail: *janet@statcollab.com*). Nancy Geller, Office of Biostatistics Research, National Heart, Lung and Blood Institute, Rockledge 2, Room 8210, Bethesda, MD 20892-7938 (E-mail: *ng@helix.nih.gov*). We thank Kent Bailey, Erica Brittain, Dean Follmann, Ed Lakatos, Gordon Lan, Joel Verter, and David Zucker for sharing with us their collective memory of the 1980s. We are grateful to Sunjana Supekar for her help with the document and, especially, the references.

Color versions of one or more of the figures in the article can be found online at *www.tandfonline.com/r/tas*.

Figure 1. Harold Dorn.

in several of the newly forming institutes, including the NHI. Felix Moore (Figure 3) became the first leader of the biostatistics group at NHI. Max Halperin, who had already become part of the initial T-6 group, joined the NHI; the others went to other institutes. While the group became separated administratively, those early individuals still joined for lunch regularly, continuing their heated discussions about the development of first principles. Over the next several years, the various NIH institutes hired additional statisticians to form what was then one of the largest (perhaps the largest) collections of biostatisticians in the country.

Leadership of the NHI Biostatistics Branch changed over time. Halperin (Figure 4) left the NHI in 1955 to become Branch Chief in the Division of Biologics at NIH. In 1960, Cornfield, who had previously been at the National Cancer Institute and that at Johns Hopkins, returned to the NIH to become the NHI Biostatistics Assistant Branch Chief. In 1963 he became Branch Chief, a position he held until 1967 (Figure 5). Halperin returned to the NHI in 1966 as Assistant Branch Chief and assumed the position of Branch Chief in 1968 following the retirement of Cornfield (Figure 2).

The NHI added lung diseases to its mission in 1969, becoming the National Heart and Lung Institute (NHLI). Although the National Sickle Cell Disease Program began at NHLI in 1971 and the NHLI included a Division of Blood Diseases and Resources in 1972, it was not until 1976 that the Institute became what it is today, the NHLBI (we shall use the moniker NHLBI henceforth).

Max Halperin was succeeded by David L. DeMets from 1977 to 1982. The next two Chiefs were Janet Wittes (1983–1989) and Nancy Geller (1990–present) (Figure 6). Joel Verter served

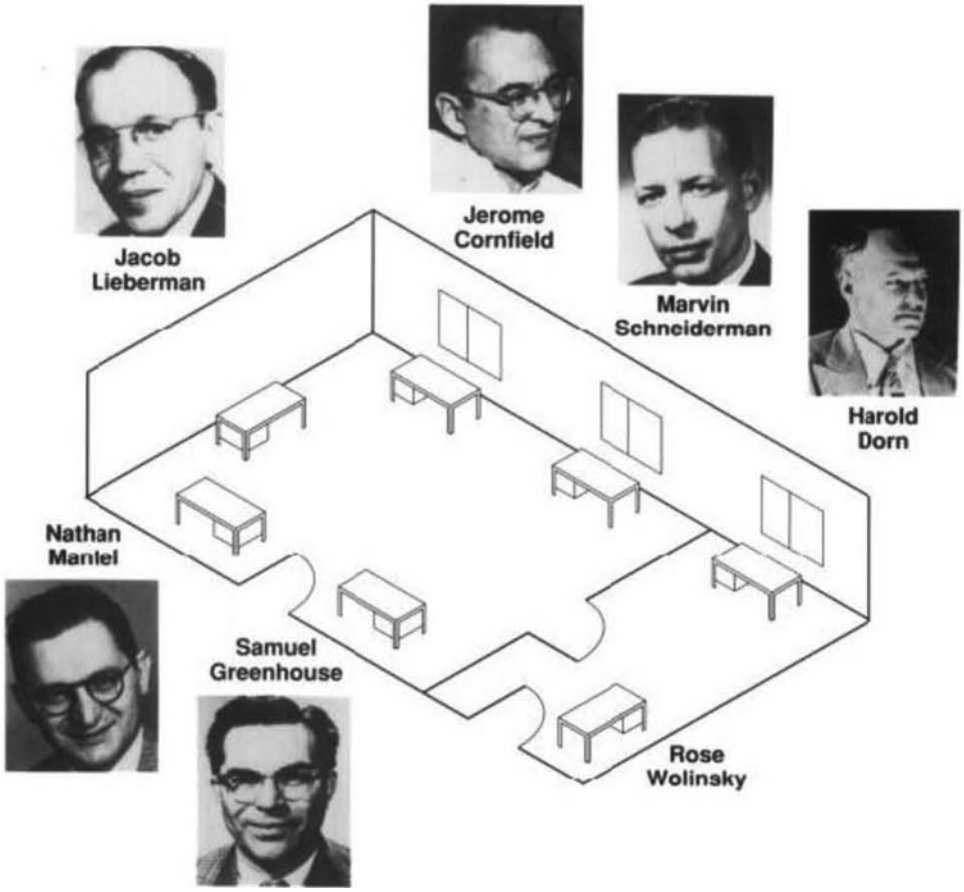

Figure 2. T-6: The original Office of Biometry at the National Institutes of Health.

Figure 3. Felix Moore.

as an Acting Chief in the periods between DeMets and Wittes, and then again between Wittes and Geller. The Branch has had a strong statistical staff over the course of its life.

2. THE DEVELOPMENT OF CLINICAL TRIALS: THE EARLY YEARS

In the early 1950s, cardiovascular disease was occurring in epidemic proportions; it led to more mortality and morbidity than any other disease. Physicians who treated wounded soldiers from World War II and the Korean War were noticing that atherosclerosis had already developed in young healthy males. Risk factors and causes for this finding were largely unknown. The challenge to the NHI was how to address this serious public health problem.

One major and, in retrospect, game-changing decision, was the establishment in 1948 of the Framingham Heart Study (FHS), an observational study that enrolled 5209 men and women from Framingham, Massachusetts (Dawber, Meadors, and Moore 1951). The FHS cohort continued in follow-up for over four decades, characterizing the cohort's long-term risk. The offspring of the original cohort formed another observational study. We now consider the risk factors of age, gender, weight, smoking, blood pressure, cholesterol, and diabetes common knowledge, but their relationship to cardiovascular disease was unknown in the 1950s when the FHS was first underway. Cornfield and his colleagues used logistic regression to establish these as potential risk factors associated with the cardiovascular epidemic; that set of associations was one of the most significant contributions to public health. In so doing, they introduced a new quantitative approach to epidemiology (Kannel et al. 1961; Truett, Cornfield, and Kannel 1967).

Even at the time of the formation of the FHS, the NHI recognized that identifying possible risk factors associated with coronary heart disease did not imply causation. The NHI understood that randomized clinical trials would be necessary to test hypotheses that modifying identified risk factors would result in a reduction in cardiovascular mortality and morbidity. How to perform such trials, however, was not clear. Therefore, the NHI commissioned the Heart Special Projects Committee, chaired by Dr. Bernard Greenberg, a professor of statistics at the University of North Carolina, to study the requirements and make recommendations. Those recommendations, issued in 1967, became known as the Greenberg Report which was published in 1988 (Heart Special Projects Committee 1988). The Committee proposed an organizational structure for NHI clinical trials. After some modification, the proposed structure became the model for future sponsored trials (Figure 7).

The Greenberg model introduced several important interlocking components of clinical trials. First, trials needed clinical centers to recruit patients. The lead investigators from each clinical center would form a Steering Committee. For trials with many clinical sites, a subcommittee of this Steering Committee might be formed to govern the trial. Central to the operations of the trial would be a data coordinating center responsible for data management and statistical analysis. One of the most important recommendations was the introduction of a policy advisory board that would be independent of the sponsor and the investigators. This board later became referred to as a Data and Safety Monitoring Board (DSMB) or Data Monitoring Committee (DMC). The DSMB was to be charged with periodic review of interim data to evaluate the emerging benefits and risks of the study intervention and to make recommendations to the NHI as to whether the trial should continue, be terminated, or be modified in some way. The report strongly emphasized that trials should not be terminated or modified without review of an external independent body. A number of articles and texts have subsequently described this process in detail (e.g., Ellenberg, Fleming, and DeMets 2002). Many industry-sponsored clinical trials have adapted, with some variations, this early model. The main modification industry trials introduced was the separation of the statistical analysis of the interim data from the data coordinating center. This separate statistical analysis center ensures a firewall between those who process the data and those who must do unblinded analysis in support of the DMC work (Ellenberg, Fleming, and DeMets 2002). The impact of the early recommendations of the Greenberg Report and the implementation by the Coronary Drug Project (CDP) described below has had tremendous impact on the conduct of clinical trials not only in the United States but worldwide as well.

Two major studies conducted or initiated in the 1960s were the Diet Heart Feasibility Study (The National Diet-Heart Study Research Group 1968) and the CDP (The Coronary Drug Project Research Group 1973). The CDP, the first major randomized trial sponsored by the NHI to benefit from the Greenberg Report, was a placebo-controlled trial of 8341 men with a recent heart attack. The trial was designed to evaluate five approaches

Figure 4. Max Halperin.

to lowering serum cholesterol levels. Total mortality was the primary outcome. The interventions were two different doses of estrogen, clofibrate, dextrothyroxine, and niacin. While only the niacin arm indicated a benefit in reducing mortality, its side effects made it an impractical intervention. (Later, a sustained release version greatly reduced the side effects and niacin became more frequently used.) The early results appeared to show benefit of the clofibrate arm, but the DSMB resisted terminating the arm, feeling that the data were premature. As the trial proceeded, the results ebbed and flowed, sometimes showing no difference, sometimes showing a marginal benefit, but ultimately settling down to indicate absolutely no effect on mortality. On the recommendation of the DSMB, all arms except niacin and placebo were terminated early for evidence of increased cardiovascular risk.

The CDP Research Group had to deal with the organizational structure as well as the operational details and statistical issues in design and analysis (The Coronary Drug Project Research Group 1981). While the results of the CDP did not identify successful interventions, except perhaps for niacin, the trial has become a useful didactic example for students learning about implementation of clinical trials and biostatistical methodology. The lessons of this trial became the basis of one of the first major texts on the fundamental principles of clinical trial design and conduct (Friedman, Furberg, and DeMets 2010).

In 1972, two members (Hal Kahn and Fred Ederer) of the Biostatistics Branch left the NHLBI to start a new biostatistics office in the National Eye Institute and, in that process, exported the clinical trial methodology lessons gained from the CDP.

Even before the CDP began to report its first results in 1973, the NHLBI began to design a series of trials in cardiology, cardiovascular surgery, pulmonary diseases, and blood diseases. Many of these trials have had a profound impact on the practice of cardiovascular and pulmonary medicine. Unlike the tight budgets of the current decade, in the relatively halcyon days of the early 1970's, the budget of NIH in general and the NHLBI in particular was large enough to allow the development and launch of these trials. For example, the Coronary Artery Surgery Study (CASS), one of the first randomized trials of cardiovascular surgery, was influential in comparing the benefits of coronary bypass graft surgery to best medical care (Tasaka et al. 1984). Results suggested that patients with multiple vessel disease benefited the most from surgery. The Hypertension Detection and Followup Project (HDFP), a trial of almost 11,000 men, demonstrated that treatment of mild hypertension (moderate high blood pressure) with a series of stepped drug interventions to lower blood pressure would reduce total five-year coronary mortality compared to usual or referred care (Hypertension Detection and Follow-up Program Cooperative Group 1979). Prior to this trial, physicians routinely treated only the most severely

Figure 5. Jerome Cornfield.

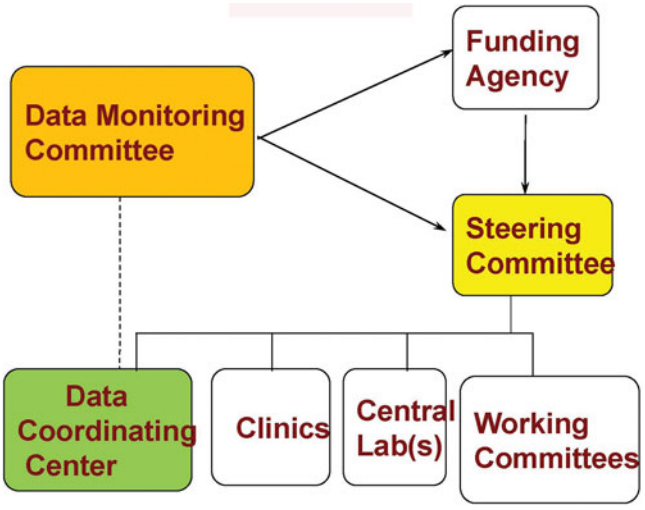

Figure 7. The NHLBI model for clinical trials.

hypertensive patients. The Lipid Research Clinic (LRC) Primary Prevention Trial, which randomized 3804 men with high serum cholesterol, showed that cholestyramine, which reduces serum cholesterol, led to a modest but significant reduction in coronary heart disease compared to placebo (7.0% vs 8.6% 7 year incidence) (Rifkind 1984). Two decades later, a series of trials using newer statin therapies proved even more beneficial in reducing mortality.

The Beta-blocker Heart Attack Trial (BHAT) was one of the first major randomized placebo-controlled trials to demonstrate that the administration of a beta-blocking drug rather than a matching placebo within a few days following a heart attack would substantially and significantly reduce mortality from a second heart attack (The Beta-Blocker Heart Attack Trial Research Group 1982). The trial, which randomized 3827 participants, showed a 25% reduction in mortality. Use of a beta-blocker is now standard of care.

Figure 6. Nancy Geller, Janet Wittes, and Dave DeMets (left to right).

The Multiple Risk Factor Intervention Trial tested the hypothesis that, for patients whose Framingham risk score indicated an elevated cardiovascular risk, an intervention focused on reducing blood pressure, increasing smoking cessation, and modifying diet would result in an overall reduction in cardiovascular disease (The Multiple Risk Factor Intervention Trial Research Group 1982). This trial randomized over 12,000 men aged 35–57 years and followed them for an average of 7 years. The men were randomized either into a special intervention group with a more intense approach to modifying risk factors or to a usual care group. Results were somewhat surprising in that the special intervention group showed only a small benefit in coronary heart disease mortality despite a reduction relative to usual care in the overall risk.

In pulmonary disease, three early trials of the 1970's are of special interest. The respiratory distress syndrome (RDS) trial evaluated the administration of a steroid, dextromethadone, to mothers at risk for premature delivery (The Collaborative Group on Antenatal Steroid Therapy 1981). The lungs are the last major organ system to develop causing babies born prematurely to struggle with respiratory distress and inadequate oxygen. This steroid intervention to the mother, which would pass the placental barrier, was believed to accelerate the process of lung development and thus reduce the distress. The RDS trial demonstrated a reduction in respiratory distress of the neonate; longer term follow-up indicated no deficit in infant development (The Collaborative Group on Antenatal Steroid Therapy 1984).

The Intermittent Positive Pressure Breathing Trial (IPPB), an early randomized trial of a medical device, was designed to evaluate whether the IPPB device was superior to a hand-held nebulizer in delivering a drug into the lungs for patients with obstructive pulmonary disease (The Intermittent Positive Pressure Breathing Trial Group 1983). The IPPB intervention used pressure to deliver the bronchodilator treatment, forcing drug into the lungs. The intervention is costly; it requires both a device and health care staff. The trial was well conducted but gave no evidence of any benefit for IPPB. The trial changed the practice of pulmonary care by reducing the use of a nonbeneficial intervention.

The Nocturnal Oxygen Therapy Trial (NOTT) was an early "noninferiority" trial, although the term was never used to describe NOTT (Nocturnal Oxygen Therapy Trial Group 1980). Patients with advanced chronic obstructive pulmonary disease often require oxygen supplementation to survive or to carry on some specific daily functions. Although the standard of care at the time was 24 hr of oxygen supplementation, little data were available to support its use. Oxygen supplied in a tank is expensive and Medicare was interested whether this cost was justified. Some clinicians considered it not ethical to compare no oxygen supplementation to 24 hr of supplementation. Others considered the trial ethical but difficult to recruit patients. Thus, the NOTT compared 24 hr of oxygen to 12 hr of oxygen administered during the evening or nocturnal hours. The results clearly showed that treatment with 24 hr of oxygen had a 50% lower mortality than the nocturnally treated group, despite the fact many of the other intermediate outcomes showed no difference between the groups. In this case, reliance on intermediate outcomes would have suggested that the less costly and perhaps more convenient nocturnal oxygen would have resulted in lower mortality. It was one of the many trials that cautioned clinical trialists about reliance on using an intermediate marker as a surrogate outcome.

The Cardiac Arrhythmia Suppression Trial (CAST), which was completed in 1989, had a profound effect on cardiology and clinical trials in general (Cardiac Arrhythmia Suppression Trial [CAST] Investigators 1989). At the time, it was well known that patients with cardiac arrhythmias are at increased risk for sudden death. Anti-arrhythmia drugs, which were developed and approved to suppress severe arrhythmias, became widely used generally in patients with less severe arrhythmias probably because the clinical community had accepted the proposition that reduction in arrhythmias could serve as a surrogate for a reduction in sudden death. The CAST trial was a randomized placebo-controlled trial testing whether anti-arrhythmics would in fact reduce the incidence of both sudden death and total mortality. Many people objected to CAST when it was designed on the grounds that they considered randomization to placebo unethical. To the surprise of many, the CAST trial was terminated when only approximately 10% of the total expected deaths had occurred, with convincing evidence of harm, not of benefit as most of the cardiology community had expected. This failure of an intermediate marker as a true surrogate for a clinical outcome was a landmark experience for many people; it alerted the clinical research community to the unreliability of intermediate markers as surrogates for clinical outcome, no matter how strong the correlation or the biological rationale.

These early trials formed a template for many trials performed in the subsequent years. Each trial addressed a question of importance to the public health; each followed the model described in the Greenberg report; and each involved active collaboration among the biostatisticians in the Branch, the other members of the Institute staff, and investigators from the academic community at large.

3. CARDIOVASCULAR CLINICAL TRIALS FROM THE 1980s ONWARD

The doubling of the NIH budget beginning in the late 1980s and extending into the 1990s resulted in a proliferation of cardiovascular clinical trials sponsored by NHLBI. The basic structure of Figure 7 continued to be used throughout that period (and continues to be used to the present day). Some trials used more complex designs, such as factorial and partially factorial layouts, to address more than one question within the same trial. To give a flavor of the types of trials conducted in this period, we mention briefly a few of them.

Because of the Institute's long-standing interest in hypertension, it conducted several trials of blood pressure reduction and hypertension prevention. For example, the Systolic Hypertension in the Elderly Program (SHEP) assessed whether treating elderly patients who had isolated systolic hypertension (ISH) would reduce the rate of stroke. After a mean of 4.5 years of follow-up, the incidence of fatal and nonfatal ischemic and hemorrhagic strokes was reduced in the active intervention arm, changing the standard of care for ISH (SHEP Cooperative Research Group 1991). Trials studying the prevention of hyperten-

sion included the trials of hypertension prevention (e.g., Whelton et al. 1992; The Trials of Hypertension Prevention Collaborative Research Group 1997) and the dietary approach to stop hypertension (DASH; Appel et al. 1997). The DASH diet, rich in fruits, vegetables, and low-fat dairy products, has become the paradigm of a heart-healthy diet.

An early example of a 2 × 2 factorial design was the Post-Coronary Artery Bypass Graft study (Post-CABG), a trial of aggressive vs moderate lipid-lowering with lovastatin (plus cholestyramine, if needed) as well as warfarin vs sham anticoagulation in patients who had had previous coronary artery bypass surgery (The Post Coronary Artery Bypass Graft Trial Investigators 1997). The Post-CABG trial also resulted in articles describing the methodology used to analyze the primary endpoint of the trial (Zucker and Wittes 1992; Canner et al. 1997).

A partial factorial design consists of separate trials of interventions believed to be independent as in a factorial design, but not all subjects are enrolled in all the trials. An early example of a partial factorial trial was the double-blind Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) which was planned to randomize 40,000 high risk hypertensive patients to three newer anti-hypertensive agents which were each compared to a control thiazide-type diuretic (chlorthalidone). The primary outcome was the occurrence of fatal coronary heart disease or nonfatal myocardial infarction. The doxazosin (an alpha-adrenergic blocker) treatment arm was stopped early because of an increased risk of heart failure. There was no difference in outcome in the other three arms, leading to the conclusion that thiazide-type diuretics should be the first-step antihypertensive therapy (The ALLHAT Officers and Coordinators 2004).

At the same time as the NHBLI was studying prevention of hypertension, it was also studying treatment of heart disease. For example, in 1992, the double-blind Studies of Left Ventricular Dysfunction (SOLVD) treatment trial showed that patients with heart failure had reduced mortality compared to placebo (The SOLVD Investigators 1992).

The Digitalis Investigation Group (1997) conducted a large simple trial comparing digitalis to placebo in 6800 patients with heart failure (left ventricular ejection fractions of 0.45 or less), all of whom received background medication. The simplicity of the data form was controversial at the time (one page for eligibility, one page for background medication, one page for each follow-up visit), but it represented one of the Institute's initiatives in making trials more efficient.

The Thrombolysis in Myocardial Infarction (TIMI) trials tested whether thrombolytic agents administered in the early hours after an acute myocardial infarction would reduce mortality (e.g., The TIMI Study Group 1994).

Other randomized trials studied coronary artery bypass grafting (CABG) and percutaneous transluminal coronary balloon angioplasty (PTCA). In 2004, the Bypass Angioplasty Revascularization Investigation (BARI) compared CABG and PTCA with respect to 5-year survival in patients with multivessel disease who were suitable candidates for either revascularization procedure. The initial strategy of PTCA did not compromise overall survival, but those in the PTCA group more frequently required at least one repeated revascularization (The BARI Investigators 1996). Subgroup analysis (at the 0.01 level) suggested that those with a history of diabetes were more likely to benefit from CABG. Although this was not an initially designed subgroup, the observed difference led to BARI 2D, a trial that enrolled 2368 patients with both Type 2 diabetes and CAD eligible for both PCI and CABG. Patients were randomized in a factorial design to either prompt revascularization or medical therapy as well as to insulin-sensitization therapy or insulin-provision therapy to achieve hemoglobin A1c (HbA1c) below 7%. No differences in death rates or major cardiovascular event rates were found for either intervention (The BARI 2D Study Group 2009). This was just one of the many trials that emphasize the need for confirmatory trials when a subgroup unexpectedly shows benefit.

The early NHBLI trials of heart disease included only men. In 1992, the NHLBI initiated the Women's Health Initiative (WHI), which included a large observational study and a large prevention trial in postmenopausal women. At that time, hormone replacement therapy (HRT) was in widespread use in the United States. On the basis of observational studies, many people believed that HRT would protect women from heart disease. The WHI involved four clinical trials in partial factorial designs comparing HRT to placebo to study heart disease (and other outcomes), a low-fat eating pattern compared to usual dietary advice to study breast cancer (and other outcomes), and calcium plus vitamin D compared to placebo to study hip fractures (and other outcomes). The WHI randomized approximately 68,000 women in one or more of the trials; the planned duration was 8.5 years. Because of the trial's complexity and visibility, the DSMB monitored many endpoints of interest. The statistical stopping rules for the HRT trials were designed to monitor not only the primary and secondary endpoints, but also some prespecified adverse events. Because multiple endpoints were monitored, it was necessary to assure that the trial would not stop too early. Thus, monitoring included an overall measure of risk and benefit as part of the stopping guidelines (Anderson et al. 2007). The first WHI results were published in 2002, earlier than initially expected, because of an increased risk of breast cancer in the HRT trial in women with an intact uterus. Contradicting the findings of many epidemiological studies, the data also showed an increased risk of cardiovascular disease (Writing Group for the Women's Health Initiative Investigators 2002). Results of the other trials showed a neutral effect in women without a uterus and in the dietary and calcium trials. The WHI has had a major effect on medical practice with far fewer women taking HRT owing to the adverse effects seen in the trials.

The Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial enrolled 10,251 subjects with Type 2 diabetes who were at high risk for cardiovascular disease and randomized them to intensive or standard glucose lowering. The primary outcome was a composite of fatal or nonfatal myocardial infarction or fatal or nonfatal stroke. Any of a long list of drugs could be used to attain the HbA1c goal. Approximately half of the participants were also enrolled in a double-blind lipid trial (randomized to fenofibrate or placebo) and the other half were enrolled in a blood pressure trial that targeted SBP of <120 mm

Hg or <140 mm Hg. All three trials had the same primary outcome (The ACCORD Study Group 2007). Notable was that the glycemia trial was stopped early for an excess of deaths on the intensive glycemia arm, although the primary endpoint favored the intensive lowering arm (The ACCORD Study Group 2011). The other trials were neutral and the ACCORD cohort continues to be followed to assess the long-term effects of randomized treatment assignment.

4. OTHER CLINICAL TRIAL INITIATIVES FROM NHLBI

In order to facilitate the administration of multiple trials in specific disease areas, the NHLBI has established clinical trials networks; each has one coordinating center responsible for multiple trials over a period of several years. The first networks sponsored by NHLBI were in lung diseases, the Asthma Clinical Trials Network and the Childhood Asthma Research and Education Network. These undertook multiple trials and led to the establishment of asthma treatment guidelines (Denlinger et al. 2007). Networks in cardiovascular diseases came later, with the Heart Failure Clinical Research Network (*https://www.hfnetwork.org/*), the Cardiovascular Cell Therapy Network (https://ccct.sph.uth.tmc.edu/cctrn/), the Cardiothoracic Surgery Network (*http://www.ctsurgerynet.org/*) and the Pediatric Heart Network (*http://www.pediatricheartnetwork.org/ResourcesPublications/AboutthePHN.aspx*). They continue to conduct multiple clinical trials in their respective areas. By having a single administrative and statistical coordinating structure in place and a group of investigators working together, these successful networks are able to design and conduct trials efficiently.

5. THE FUTURE OF TRIALS AT NHBLI

The Institute has long been interested in developing methods for increasing the efficiency of trials (Wittes et al. 1990; Waclawiw et al. 2012). One approach has been designs of trials that collect only necessary data (Verter 1990). The current need for improving efficiency results from the high cost of large trials in an era of tight budgets. To ensure continuation of its tradition of high quality randomized trials, the NHLBI is exploring ways to decrease costs per subject. One way of enhancing efficiency is through the use of available high quality registries to assess eligibility of participants. The NHLBI is currently investigating use of registries to facilitate the continuation of its long heritage of performing influential clinical trials (Lauer and D'Agostino 2013).

6. STATISTICAL METHODOLOGY

Only a few large randomized trials had been conducted before 1970 when the huge buildup in clinical trials began at NHLBI. New statistical methods were needed to address challenges in design and analysis. The entire NIH biostatistics community, as well as the biostatistics community nationally and internationally, was very active in evaluating, modifying, and innovating new statistical methods. In this article, we focus only on the methods developed by the NHLBI statistical teams. The most prominent methodology of the NHLBI Biostatistics Branch might be thought of in three major categories: (i) issues related to design, (ii) data monitoring methods, and (iii) analysis for longitudinal studies. Other methodologies in which the Branch has made important contributions included errors-in-variables and informative censoring. Halperin, DeMets, and Ware (1990) described some of these methods in detail.

The articles to which we refer are only a small, nonrandom selection of articles written by Branch members over the years when they were at the Branch. Our goal is to provide a once-over-lightly view of the type of research in which the Branch took leadership. We hope this gives the reader a sense of the variety of problems the Branch faced and the approaches members took to address them.

7. METHODS RELATED TO THE DESIGN OF TRIALS

Since many of the interventions that the NHLBI was considering in the early 1970s involved long-term follow-up, the biostatisticians working on the design of these trials encountered challenges. The CDP and the HDFP had demonstrated that many trial participants would not adhere religiously to the intervention the protocol specified. Some participants randomized to the intervention arm would fail to take all of their medication or would not follow their assigned intervention completely; some might even "drop out" of the intervention altogether. In circumstances where the intervention was readily available outside the trial, participants in the control arm might start taking some of the intervention, or "drop in." Lessons from the CDP had taught clinical trialists, including biostatisticians, of the serious inferential bias that can arise by removing non-adhering participants from the analysis (The Coronary Drug Project Research Group 1980). This understanding cemented the importance of the principle of "intent to treat" (or "analyze as randomized"). Halperin et al. (1982) developed a model to adjust the sample size for certain patterns of nonadherence. The greater the degree of nonadherence, the more the sample size must be increased to compensate. Their sample size model assumed a fixed rate of non-adherence and a fixed hazard rate. Several other articles from the Branch in the 1980s addressed more complicated patterns of nonadherence or censoring by competing risks (Wu, Fisher, and DeMets 1980, Wu, Ware, and Feinlieb 1980; Wu 1988). Lakatos (1986, 1988) used Markov models to allow sample size calculation under nonproportional hazards. Shih (1995) provided software that has become widely used for the design of complex clinical trials. For example, the hazard for a patient following a heart attack changes over the first year, with the risk greatest during the first 30 days; risk then decreases over the next 3 months and becomes close to normal risk after one year. Nonadherence that begins immediately after the first heart attack and at the beginning of the new intervention has much greater impact on statistical power than non-adherence that occurs later in follow-up. Members of the Branch applied patterns of adherence rates from previous trials to the design of future trials. This type of modeling proved useful as the NHLBI designed many long-term follow-up trials with longer term interventions.

Zucker and Lakatos (1990) and Zucker (1992) proposed using weighted log-rank tests to deal with nonproportional hazards.

Blackwelder discussed the challenge of designing trials to demonstrate "equivalence" of interventions, perhaps one being the standard and the other a new less invasive, less toxic or less expensive alternative to the standard (Blackwelder 1982). He proposed using a margin of indifference for designing the trial and determining the sample size determination. This general paradigm, which has evolved over time, is now referred to as noninferiority; such trials are widely discussed in FDA guidelines and in the statistical literature.

In recent years, the Branch has developed methodology for the design of "biological assignment" trials (Logan et al. 2008) and genetics-based trials (Joo et al. 2010).

8. DATA MONITORING

As the Greenberg Report discussed, clinical trials needed to be monitored to assess progress and to look for early convincing evidence of benefit or important suggestions of harm. Over the life of the Branch, the NHLBI biostatisticians have made substantial contributions to the methodologies of interim monitoring of emerging clinical trial data.

The CDP provided an early opportunity to explore various statistical methods for monitoring interim results. Broadly speaking, these methods can be categorized as Bayesian approaches, curtailment procedures which lead to the concept of conditional power, and sequential testing. Cornfield described how the likelihood principle and Bayesian approaches could address sequential methods for interim analyses. His 1966 articles described this approach for acute response variables (Cornfield 1966a, 1966b). Cornfield, DeMets, and Ware (1969) introduced a Bayesian adaptive approach for using interim results to allocate trial participants to the leading treatment. Cornfield (1969) proposed a method of assessing interim results, rejecting the null hypothesis when the posterior odds became relatively small; he referred to these measures as relative betting odds (RBOs). The CDP investigators explored this method for their trial, but RBOs did not gain widespread use at that time. Three decades later, with substantially increased computational power and further methodological development, Bayesian methods are now being applied to the design and monitoring of clinical trials, especially phase 1 and 2 trials. Cornfield's early work certainly influenced the continued interest in this Bayesian approach.

The CDP also stimulated another approach to data monitoring. Initially, discussions between Halperin and Alling (at NIH but not at NHLBI) dealt with the question "what if" the future results not yet observed occurred in some proportion in each arm of the study. The idea was that a trial could be terminated when future results could no longer substantially alter the final test statistic. This thinking led to a formal approach dubbed curtailed sampling (Alling 1966). Halperin and Ware pursued the idea of curtailed sampling for the Wilcoxon test used in censored survival analysis (1974). Later, DeMets and Halperin considered an analogous strategy for the two sample t-test (1982).

The curtailed sampling approach is quite conservative because it demands that the final results cannot be changed, no matter what happens in the future. Statisticians at the NHLBI began to pursue the idea of what became known as stochastic curtailment or conditional power. Here, the probability of rejecting the null hypothesis at the scheduled end of the trial is computed, given that some of the results are already known, and speculating on the effect that will be observed during the remainder of the trial. If that probability is sufficiently low, then the trial might be terminated for futility. This method requires postulating the effect of the intervention in the remainder of the trial, perhaps using the current trend, perhaps the effect hypothesized in the design, perhaps a null effect, or perhaps other effect sizes in between. This idea can also be applied to a trial where the interim results are extremely favorable and the question is whether the probability of future results might diminish these favorable trends to make the final test statistic not significant. After some experience with this approach, NHLBI statisticians published two articles in 1982 describing the method, as well as a more rigorous presentation of this stochastic curtailment methodology (Halperin et al. 1982; Lan, Simon, and Halperin 1982). Later, Lan and Wittes (1988) proposed a simple way to compute these probabilities; this article was helpful in making this methodology accessible to clinical trialists (Lan and Wittes 1988).

Canner of the CDP coordinating center along with Cornfield and Halperin of the Branch experimented with a solution for the repeated testing problem, using a more conservative critical value for the final test statistic to compensate for early looks (Canner 1977).

DeMets and Ware explored asymmetric group sequential boundaries where trends in the wrong direction suggested that benefit was not likely, but falling short of evidence indicating harm (1980, 1982). This concept was stimulated by the previously described RDS trial where there was increased sensitivity about allowing a new intervention to continue to the point of demonstrating harm. While the concept of asymmetric boundaries was an important contribution to the field of clinical trials, these boundaries were replaced in practice with boundaries for futility generated by conditional power.

The BHAT trial, completed in 1982, was one of the first large multicenter trials to apply the O'Brien and Fleming sequential approach (a method not developed at the Branch) (O'Brien and Fleming 1979). In fact, BHAT was terminated early for convincing benefit using in part these sequential boundaries (The Beta-Blocker Heart Attack Trial Research Group 1982). In applying the O'Brien–Fleming sequential procedure to the BHAT trial, Lan and DeMets recognized that fixing the number of interim analyses in advance and demanding equally spaced times of analyses limited the practical utility of the method. In 1983, they introduced a more flexible sequential procedure that relaxed both these restrictions (Lan and DeMets 1983). Their method, which took advantage of basic Brownian motion theory and methods, came to be known as the alpha-spending function method for sequential analysis. They showed that either a Pocock or an O'Brien–Fleming type alpha-spending function could be defined to induce a sequential boundary similar to the more standard O'Brien–Fleming or Pocock (1977) boundaries but without their constraints. Other more general classes of alpha-spending functions were proposed to offer a rich set of

such sequential boundaries with different early stopping properties. In a subsequent article, they discussed information time in more detail (Lan and DeMets 1989a). Of special note is that the frequency of interim analyses can be changed during the trial using the alpha-spending function, even be influenced by the strength of the emerging trends (Lan and DeMets 1989b). The alpha-spending function approach is now widely applied in monitoring clinical trials.

Several Branch members explored various other methods of interim analysis. Brittain and Bailey studied optimal interim times to look at accruing data (1993). A recent topic of interest is adaptive trials. Wittes and Brittain (1990) and Proschan and Hunsberger (1995) developed various approaches to using interim data to change sample size of a study without compromising the Type I error rate. Follman, Proschan, and Geller (1994) and Proschan, Follman, and Geller (1994) addressed the stopping rules for more complex trials.

9. ANALYSIS OF LONGITUDINAL STUDIES

Many clinical trials and observational studies measure continuous outcomes over time. The analysis of those types of studies can be challenging because of missing data, censoring, unequally spaced measurements, and the usual challenges inherent in modeling. Many of the studies sponsored by the NHLBI, including the FHS, used continuous outcomes and, therefore, faced those challenges.

The IPPB was one of the many trials that stimulated the methodological interests of the Branch. The IPPB, which measured pulmonary function tests over several years, had censoring due to mortality and data that were missing probably not at random (The Intermittent Positive Pressure Breathing Trial Group 1983). At that time, when confronted with longitudinal data, the NHLBI statisticians often fit linear models for each participant, averaged the estimated slopes, and then compared the mean slopes between intervention arms to assess intervention effects. The Scottish statistician Wishart had suggested this simple method in 1938. While the Wishart approach was attractive, the Branch began to feel it failed to exploit the richness of longitudinal data and it ignored informative censoring. These problems led to research on other methods to deal with these data. Ware and colleagues published an article on repeated measures for studying circadian rhythm (Ware and Bowden 1977) and one on blood pressure changes and normative growth (Wu, Fisher, and DeMets 1980; Wu, Ware, and Feinlieb 1980). Wu and Ware teamed up to publish an influential article on how to use serial patient measurements to predict future values (1980). A article by Laird and Ware on random effects models was probably the most influential article from the efforts of this era on the analysis of longitudinal data (1982). Wu and colleagues extended these methods for informative censoring, with specific reference to declines in lung function (Wu 1988; Wu and Carroll 1988; Wu and Bailey 1988, 1989). Modeling of longitudinal data continues to be of interest to the Branch, both in epidemiological studies (Wu and Tian 2013) and clinical trials (Jeffries and Geller 2015).

10. OTHER STATISTICAL TOPICS

While general approaches to design of trials, interim monitoring of clinical trials, and longitudinal analyses were probably the consistent statistical themes that have engaged the Branch from its inception, the group addressed other important topics as well, for instance, errors in variables (Carroll et al. 1984), subgroups (Yusuf et al. 1991), community randomized trials (Zucker et al. 1995), meta-analysis (Yusuf, Zucker, and Peduzzi 1994), survival analysis (Shih 1995; Yang and Prentice 2010), and statistical genetics (e.g., Jeffries 2009; Troendle and Mills 2011; Zheng et al. 2012). All of these problems arose when various members of the Branch found themselves faced with a practical problem that had no obvious solution. An article by Bailey on detecting fabrication in data has helped many statisticians think about what types of data are so "good" that they probably represent falsifications (1991).

11. SUMMARY

We aimed in this article to provide the readers some insight into the contributions of the Biostatistics Branch of the NHLBI to medical research, especially to the design, monitoring, and analysis of clinical trials. Of special interest to us was to explore why the Branch at NHLBI had such a profound influence on clinical trials. One set of reasons relates to the qualities of the members of the Branch. Many were talented, they worked hard, and they were lucky to be at the right place at the right time. The Branch was fortunate to have supportive Institute and Division directors who understood the importance of statistics to the scientific success of their studies and who were eager to have us involved early in planning. We forged partnerships with each other and with the physicians and scientists with whom we worked. Many of these relationships have lasted long after members left the Branch either for other Institutes within NIH or for different organizations entirely. We may not have argued as vociferously and passionately as the T-6+Halperin group did, but the culture they established has characterized the Branch throughout its life. We learned from them how enjoyable statistics can be.

We three believe that the most important lesson for future generations of biostatisticians to carry with them is the importance of engaging in the science of the projects they will encounter. Throughout its history, Branch members have listened carefully to the problems faced by the clinicians with whom we have worked. We looked for ways we could use our statistical training and insight to contribute to solutions. We did not like to separate "statistical" from "medical" issues; we regarded them all—to a greater or lesser extent—as part of a whole. It was important for us to understand the biology of the studies with which we were involved; it was important also for us to make sure that the clinicians with whom we worked understood our problems and the reasons for the directions we were taking.

The future will likely bring Big Data, data sharing, computational biology, translational medicine, personalized medicine, and other challenges that we cannot predict. We hope that future generations of statisticians will take note of what made this

group so successful for so many years and continue the tradition of collaboration and fun.

REFERENCES

Alling, D. W. (1966), "Closed Sequential Tests for Binomial Probabilities," *Biometrika*, 53, 73–84. [116]

Anderson, G. L., Kooperberg, C., Geller, N., Rossouw, J. E., Pettinger, M., and Prentice, R. L. (2007), "Monitoring and Reporting of the Women's Health Initiative Randomized Hormone Therapy Trials," *Clinical Trials*, 4, 207–217. [114]

Appel, L. J., Moore, T. J., Obarzanek, E., Vollmer, W. M., Svetkey, L. P., Sacks, F. M., Bray, G. A., Vogt, T. M., Cutler, J. A., Windhauser, M. M., et al. (1997). "A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure," *New England Journal of Medicine*, 336, 1117–1124. [114]

Bailey, K. R. (1991), "Detecting Fabrication of Data in a Multicenter Collaborative Animal Study," *Controlled Clinical Trials*, 12, 741–752. [117]

Blackwelder, W. C. (1982), "Proving the Null Hypothesis in Clinical Trials," *Controlled Clinical Trials*, 3, 345–353. [116]

Brittain, E. H., and Bailey, K. R. (1993), "Optimization of Multistage Testing Times and Critical Values in Clinical Trials," *Biometrics*, 49, 763–772. [117]

Canner, P. L. (1977), "Monitoring Treatment Differences in Long-Term Clinical Trials," *Biometrics*, 33, 603–615. [116]

Canner, P. L., Thompson, B., Knatterud, G. L., Geller, N., Campeau, L., and Zucker, D. (1997), "An Application of the Zucker-Wittes Modified Ratio Estimate Statistic in the Post-Coronary Artery Bypass Graft (CABG) Clinical Trial," *Controlled Clinical Trials*, 18, 318–327. [114]

Cardiac Arrhythmia Suppression Trial (CAST) Investigators (1989), "Preliminary Report: Effect of Encainide and Flecainide on Mortality in a Randomized Trial of Arrhythmia Suppression After Myocardial Infarction," *New England Journal of Medicine*, 321, 406–412. [113]

Carroll, R. J., Spiegelman, C. H., Lan, G. K. K., Bailey, K. R., and Abbott, R. D. (1984), "Errors-in-Variables for Binary Regression Models," *Biometrika*, 71, 19–25. [117]

Cornfield, J. (1966a), "Sequential Trials, Sequential Analysis and the Likelihood Principle," *The American Statistician*, 20, 18–23. [116]

——— (1966b), "A Bayesian Test of Some Classical Hypotheses—With Applications to Sequential Clinical Trials," *Journal of the American Statistical Association*, 61, 577–594. [116]

——— (1969), "Bayesian Outlook and Its Application," *Biometrics*, 25, 617–657. [116]

Cornfield, J., Halperin, M., and Greenhouse, S. (1969), "An Adaptive Procedure for Sequential Clinical Trials," *Journal of the American Statistical Association*, 64, 759–770. [116]

Dawber, T., Meadors, G. F., and Moore, F. E. (1951), "Epidemiological Approaches to Heart Disease: The Framingham Study," *American Journal of Public Health*, 41, 279–286. [110]

Denlinger, L. C., Sorkness, C. A., Chinchilli, V. M., and Lemanske, R. F., Jr. (2007), "Guideline-Defining Asthma Clinical Trials of the NHLBI ACRN and CARE Networks," *Journal of Allergy and Clinical Immunology*, 119, 3–13. [115]

DeMets, D. L., and Halperin, M. (1982), "Early Stopping in the Two-Sample Problem for Bounded Random Variables," *Controlled Clinical Trials*, 3, 1–11. [116]

DeMets, D. L., and Ware, J. H. (1980), "Group Sequential Methods for Clinical Trials With a One-Sided Hypothesis," *Biometrika*, 67, 651–660. [116]

DeMets, D. L., and Ware, J. H. (1982), "Asymmetric Group Sequential Boundaries for Monitoring Clinical Trials," *Biometrika*, 69, 661–663. [116]

Ellenberg, S., Fleming, T., and DeMets, D. (2002), *Data Monitoring Committees in Clinical Trials: A Practical Perspective*, West Sussex: Wiley. [110]

Follmann, D. A., Proschan, M. A., and Geller, N. L. (1994), "Monitoring Pairwise Comparisons in Multiarmed Clinical Trials," *Biometrics*, 50, 325–336. [117]

Friedman, L., Furberg, C., and DeMets, D. (2010), *Fundamentals of Clinical Trials* (4th ed.), New York: Springer Science + Business Media, LLC. [111]

Halperin, M., DeMets, D. L., and Ware, J. (1990) "Early Methodological Developments in Clinical Trials at the National Heart, Lung and Blood Institute," *Statistics in Medicine*, 9, 881–892. [115]

Halperin, M., Lan, G. K. K., Ware, J. H., Johnson, N. J., and DeMets, D. L. (1982), "An Aid to Data Monitoring in Long-Term Clinical Trials," *Controlled Clinical Trials*, 3, 311–323. [115,116]

Halperin, M., and Ware, J. (1974), "Early Decision in a Censored Wilcoxon Two-Sample Test for Accumulating Survival Data," *Journal of the American Statistical Association*, 69, 414–422. [116]

Heart Special Projects Committee (1988), "Organization, Review and Administration of Cooperative Studies (Greenberg Report), A Report From the Heart Special Projects Committee to the National Advisory Heart Council, May 1967," *Controlled Clinical Trials*, 9, 137–148. [110]

Hypertension Detection and Follow-up Program Cooperative Group (1979), "Five Year Findings of the Hypertension Detection and Follow-Up Program, 1. Reduction in Mortality of Persons With High Blood Pressure Including Mild Hypertension," *Journal of the American Medical Association*, 213, 1143–1152. [111]

Jeffries, N. O. (2009), "Ranking Bias in Association Studies," *Human Heredity*, 67, 267–275. [117]

Jeffries, N. O., and Geller, N. L. (2015), "Longitudinal Clinical Trials With Adaptive Choice of Follow-up Time," *Biometrics* (in press), DOI: 10.1111/biom.12287. [117]

Joo, J., Geller, N. L., Kimmel, S., Rosenberg, Y., French, B., and Ellenberg, J. (2010), "Prospective Alpha Allocation in the Clarification of Optimal Anticoagulation Through Genetics (COAG) Trial," *Clinical Trials*, 7, 597–604. [116]

Kannel, W. B., Dawber, T. R., Kagan, A., Nevotskie, N., and Stokes, J. (1961), "Factors of Risk in the Development of Coronary Heart Disease—Six Year Followup Experience: The Framingham Study," *Annals of Internal Medicine*, 55, 33–50. [110]

Laird, N., and Ware, J. E. (1982), "Random Effects Models for Longitudinal Data," *Biometrics*, 38, 963–974. [117]

Lakatos, E. (1986), "Sample Size Determination in Clinical Trials With Time-Dependent Rates of Losses and Noncompliance," *Controlled Clinical Trials*, 7, 189–199. [115]

Lakatos, E. (1988), "Sample Sizes Based on the Log-Rank Statistic in Complex Clinical Trials," *Biometrics*, 44, 229–241. [115]

Lan, K. K. G., and DeMets, D. L. (1983), "Discrete Sequential Boundaries for Clinical-Trials," *Biometrika*, 70, 659–663. [116]

——— (1989a), "Group Sequential Procedures: Calendar Versus Information Time," *Statistics in Medicine*, 8, 1191–1198. [117]

——— (1989b), "Changing Frequency of Interim Analysis in Sequential Monitoring," *Biometrics*, 45, 1017–1020. [117]

Lan, K. K. G., Simon, R., and Halperin, M. (1982), "Stochastically Curtailed Tests in Long Term Clinical Trials," *Sequential Analysis*, 1, 207–219. [116]

Lan, K. K. G., and Wittes, J. (1988), "The B-Value: A Tool for Monitoring Data," *Biometrics*, 44, 579–585. [116]

Lauer, M. S., and D'Agostino, R. G., Sr. (2013), "The Randomized Registry Trial—The Next Disruptive Technology in Clinical Research?" *The New England Journal of Medicine*, 369, 1579–1581. [115]

Logan, B., Leifer, E., Bredeson, C., Horowitz, M., Ewell, M., Carter, S., and Geller, N. (2008), "Use of Biological Assignment in Hematopoietic Stem Cell Transplantation Clinical Trials," *Clinical Trials*, 5, 607–616. [116]

O'Brien, P. C., and Fleming, T. R. (1979), "A Multiple Testing Procedure for Clinical Trials," *Biometrics*, 35, 549–556. [116]

Pocock, S. J. (1977), "Group Sequential Methods in the Design and Analysis of Clinical Trials," *Biometrika*, 64, 191–199. [116]

Post Coronary Artery Bypass Graft Trial Investigators (1997), "The Effect of Aggressive Lowering of Low-Density Lipoprotein Cholesterol Levels and Low-Dose Anticoagulation on Obstructive Changes in Saphenous-Vein Coronary-Artery Bypass Grafts," *New England Journal of Medicine*, 336, 153–162. [114]

Proschan, M. A., Follmann, D. A., and Geller, N. L. (1994), "Monitoring Multiarmed Trials," *Statistics in Medicine*, 13, 1441–1452. [117]

Proschan, M. A., and Hunsberger, S. A. (1995), "Designed Extension of Studies Based on Conditional Power," *Biometrics*, 51, 1315–1324. [117]

Rifkind, B. M. (1984), "The Lipid Research Clinics Primary Prevention Trial Results. 1. Reduction in Incidence of Coronary Heart Disease," *Journal of the American Medical Association*, 251, 351–364. [112]

SHEP Cooperative Research Group (1991), "Prevention of Stroke by Antihypertensive Drug Treatment in Older Persons With Isolated Systolic Hypertension. Final Results of the Systolic Hypertension in the Elderly Program (SHEP)," *Journal of the American Medical Association*, 265, 3255–3264. [113]

Shih, J. O. (1995), "Sample Size Calculations for Complex Clinical Trials With Survival Endpoints," *Controlled Clinical Trials*, 16, 395–407. [115,117]

Tasaka, Y., Sekine, M., Wakatsuki, M., Ohgawara, H., and Shizume, K. (1984), "Myocardial Infarction and Mortality in the Coronary Artery Surgery Study (CASS) Randomized Trial," *New England Journal of Medicine*, 310, 750–758. [111]

The ACCORD Study Group (2007), "Action to Control Cardiovascular Risk in Diabetes (ACCORD) Trial: Design and Methods," *American Journal of Cardiology*, 99, 21i–31i. [115]

The ACCORD Study Group (2011), "Long-Term Effects of Intensive Glucose Lowering on Cardiovascular Outcomes," *The New England Journal of Medicine*, 364, 818–828. [115]

The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group (2004), "Major Outcomes in High-Risk Hypertensive Patients Randomized to Angiotensin-Converting Enzyme Inhibitor or Calcium Channel Blocker vs Diuretic," *Journal of the American Medical Association*, 288, 2181–2196. [114]

The BARI 2D Study Group (2009), "A Randomized Trial of Therapies for Type 2 Diabetes and Coronary Artery Disease," *The New England Journal of Medicine*, 360, 2503–2515. [114]

The Beta-Blocker Heart Attack Trial Research Group (1982), "A Randomized Trial of Propranolol in Patients With Acute Myocardial Infarction. I. Mortality Results," *Journal of the American Medical Association*, 247, 1707–1714. [112,116]

The Bypass Angioplasty Revascularization Investigation (BARI) Investigators (1996), "Comparison of Coronary Bypass Surgery With Angioplasty in Patients With Multivessel Disease," *The New England Journal of Medicine*, 335, 217–225. [114]

The Collaborative Group on Antenatal Steroid Therapy (1981), "Effect of Antenatal Dexamethasone Administration on the Prevention of Respiratory Distress Syndrome," *American Journal of Obstetrics and Gynecology*, 141, 276–287. [113]

The Collaborative Group on Antenatal Steroid Therapy (1984), "Effects of Antenatal Dexamethasone Administration in the Infant: Long-Term Follow-Up," *Journal of Pediatrics*, 104, 259–267. [113]

The Coronary Drug Project Research Group (1973), "The Coronary Drug Project, Design, Methods and Baseline Results," *Circulation*, 47, 1–179. [110]

The Coronary Drug Project Research Group (1980), "Influence of Adherence to Treatment and Response of Cholesterol on Mortality in the Coronary Drug Project," *The New England Journal of Medicine*, 303, 1038–1041. [115]

The Coronary Drug Project Research Group (1981), "Practical Aspects of Decision Making in Clinical Trials: The Coronary Drug Project as a Case Study," *Controlled Clinical Trials*, 1, 363–376. [111]

The Digitalis Investigation Group (1997), "The Effect of Digoxin on Mortality and Morbidity in Patients With Heart Failure," *The New England Journal of Medicine*, 336, 525–533. [114]

The Intermittent Positive Pressure Breathing Trial Group (1983), "Intermittent Positive Pressure Breathing Therapy of Chronic Obstructive Pulmonary Disease—A Clinical Trial. *Annals of Internal Medicine*, 99, 612–620. [113,117]

The Multiple Risk Factor Intervention Trial Research Group (1982), "Multiple Risk Factor Intervention Trial," *Journal of the American Medical Association*, 248, 1465–1477. [113]

The National Diet-Heart Study Research Group (1968), "The National Diet-Heart Study Final Report," *Circulation*, 37, 1–428. [110]

The Nocturnal Oxygen Therapy Trial Group (1980), "Continuous or Nocturnal Oxygen Therapy in Hypoxemic Chronic Obstructive Lung Disease—A Clinical Trial," *Annals of Internal Medicine*, 93, 391–398. [113]

The Ransdell Act (1930), "An Act to Establish And Operate a National Institute of Health, to Create a System of Fellowships in Said Institute, and to Authorize the Government to Accept Donations for Use in Ascertaining the Cause, Prevention, and Cure of Disease Affecting Human Beings, and For Other Purposes," (ch. 251, Pub.L. 71–251, 46 Stat. 379, enacted May 26, 1930, codified as amended at 42 U.S.C. § 21, 42 U.S.C. § 22, 42 U.S.C. §§ 23a–23g) [108]

The SOLVD Investigators (1992), "Effect of Enalapril on Survival in Patients With Reduced Left Ventricular Ejection Fractions and Congestive Heart Failure," *The New England Journal of Medicine*, 327, 685–691. [114]

The TIMI Study Group (1994), "Effects of Tissue Plasminogen Activator and a Comparison of Early Invasive and Conservative Strategies in Unstable Angina and Non-Q-Wave Myocardial Infarction. Results of the TIMI IIIB Trial," *Circulation*, 89, 1545–1556. [114]

The Trials of Hypertension Prevention Collaborative Research Group (1997), "Effects of Weight Loss and Sodium Reduction Intervention on Blood Pressure and Hypertension Incidence in Overweight People With High-Normal Blood Pressure. The trials of Hypertension Prevention, Phase II," *Archives of Internal Medicine*, 157, 657–667. [114]

Troendle, J. F., and Mills, J. L. (2011), "Correction for Multiplicity in Genetic Association Studies of Triads: The Permutational TDT," *Annals of Human Genetics*, 75, 284–291. [117]

Truett, J., Cornfield, J., and Kannel, W. B. (1967), "A Multivariate Analysis of Risk Factors of Coronary Heart Disease in Framingham," *Journal of Chronic Diseases*, 50, 511–524. [110]

Verter, J. (1990), "How Much Data Should We Collect in a Randomized Clinical Trial?" *Statistics in Medicine*, 9, 103–113. [115]

Waclawiw, M., Wu, C. O., Yang, S., and DeMets, D. (2012), Guest Editors of "Papers from 2010 NHLBI Workshop on Clinical Trials: Past, Present and Future," *Statistics in Medicine*, 31, 2937–3072. [115]

Ware, J. E., and Bowden, R. (1977), "Circadian Rhythms of Responses Collected at Intervals," *Biometrics*, 64, 156–160. [117]

Wishart, J. (1938), "Growth Rate Determination in Nutrition Studies With the Bacon Pig and Their Analyses," *Biometrics*, 30, 16–28. [117]

Whelton, P. K., Appel, L., Charleston, J., Dalcin, A. T., Ewart, C., Fried, L., Kaidy, D., Klag, M. J., Kumanyika, S., Steffen, L., et al. (1992), "The Effects of Nonpharmacologic Interventions on Blood Pressure of Persons With High Normal Levels Results of the Trials of Hypertension Prevention, Phase I," *Journal of the American Medical Association*, 267, 1213–1220. [114]

Wittes, J., and Brittain, E. (1990), "The Role of Internal Pilot Studies in Increasing Efficiency of Clinical Trials," *Statistics in Medicine*, 9, 65–72. [117]

Wittes, J., Duggan, J., Held, P., and Yusuf, S. (eds.), (1990), "Proceedings of the NHLBI Workshop on Cost and Efficiency in Clinical Trials," *Statistics in Medicine*, 9, 1–199. [115]

Writing Group for the Women's Health Initiative Investigators (2002), "Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative Randomized Con-

trolled Trial," *Journal of the American Medical Association*, 288, 321–333. [114]

Wu, C. O., and Tian, X. (2013), "Nonparametric Estimation of Conditional Distributions and Rank-Tracking Probabilities With Time-Varying Transformation Models in Longitudinal Studies," *Journal of the American Statistical Association*, 108, 971–982. [117]

Wu, M. (1988), "Sample Size for Comparison of Changes in the Presence of Right Censoring Caused by Death, Withdrawal and Staggered Entry," *Controlled Clinical Trials*, 9, 32–46. [115,117]

Wu, M., and Carroll, R. J. (1988), "Estimation and Comparison of Changes in the Presence of Informative Right Censoring, Modeling the Censoring Process," *Biometrics*, 44, 175–188. [117]

Wu, M. C., and Bailey, K. R. (1988), "Analyzing Changes in the Presence of Informative Right Censoring Caused by Death and Withdrawal," *Statistics in Medicine*, 7, 337–346. [117]

——— (1989), "Estimation and Comparison of Changes in the Presence of Right Censoring, a Conditional Linear Model," *Biometrics*, 45, 939–959. [117]

Wu, M. C., Fisher, M., and DeMets, D. (1980), "Sample Sizes for Long-Term Medical Trial With Time-Dependent Dropout and Event Rates," *Controlled Clinical Trials*, 1, 111–124. [115,117]

Wu, M., Ware, J. E., and Feinlieb, M. (1980), "On the Relation Between Blood Pressure Change and Initial Values," *Journal of Chronic Diseases*, 37, 637–644. [115,117]

Yang, S., and Prentice, R. L. (2010), "Improved Logrank-Type Tests for Survival Data Using Adaptive Weights," *Biometrics*, 66, 30–38. [117]

Yusuf, S., Wittes, J., Probstfield, J., and Tyroler, H. A. (1991), "Analysis and Interpretation of Treatment Effects in Subgroups of Patients in Randomized Clinical Trials," *Journal of the American Medical Association*, 266, 93–98. [117]

Yusuf, S., Zucker, D., and Peduzzi, P. (1994), "Effect of Coronary Artery Bypass Graft Surgery on Survival: Overview of 10-Year Results From Randomised Trials by the Coronary Artery Bypass Graft Surgery Trialists Collaboration," *Lancet*, 344, 563–570. [117]

Zheng, G., Wu, C. O., Kwak, M., Jiang, W., Joo, J., and Lima, J. A. C. (2012), "Joint Analysis of Binary and Quantitative Traits with Data Sharing and Outcome-Dependent Sampling," *Genetic Epidemiology*, 36, 263–273. [117]

Zucker, D. M. (1992), "The Efficiency of a Weighted Log-Rank Test under a Percent Error Misspecification Model for the Log Hazard Ratio," *Biometrics*, 48, 893–899. [116]

Zucker, D. M., and Lakatos, E. (1990), "Weighted Log Rank Type Statistics for Comparing Survival Curves When There is a Time Lag in the Effectiveness of Treatment," *Biometrika*, 77, 853–864. [116]

Zucker, D. M., Lakatos, E., Webber, L. S., Murray, D. M., McKinlay, S. M., Feldman, H. A., Kelder, S. H., and Nader, P. R. (1995), "Statistical Design of the Child and Adolescent Trial for Cardiovascular Health (CATCH): Implications of Cluster Randomization," *Controlled Clinical Trials*, 16, 96–118. [117]

Zucker, D., and Wittes, J. (1992), "Testing the Effect of Treatment in Experiments with Correlated Binary Outcomes," *Biometrics*, 48, 695–709. [114]