A PLACE FOR RIGOR

Since 1949, UNC's researchers set the pace for the field of biostatistics.

BY ANGELA SPIVEY

Today it's a given that biostatisticians work hand in hand with geneticists, oncologists and drug manufacturers to design studies and crunch the numbers. But in 1949, when the UNC Department of Biostatistics was first established, as far as medical research was concerned, biostatisticians didn't really exist.

"In the first days we were looked on as interlopers," says Dr. Jim Grizzle, professor emeritus of biostatistics at Carolina. In 1957, when Grizzle started working as a graduate research assistant to Dr. Bernard Greenberg, the first chair of UNC's Department of Biostatistics, many investigators were skeptical of biostatisticians and their ideas about study design. "We had to earn the researchers' respect. We were speaking a different language: things like 'randomization' and 'uniform ascertainment of outcomes,'" Grizzle says. "Suppose you're studying a treatment that's supposed to prevent heart attacks? Well, what's a heart attack? For your outcomes to be uniform, all the centers conducting the trial have to use the same diagnostic criteria."

At that time, such principles of statistics had been used in agricultural studies, particularly in the United Kingdom, but not so much in medicine. Biostatisticians gradually began to carve out the basic tenets for collecting and analyzing data in biomedical research, and Greenberg was right in the middle of it. He and others in the UNC department laid out many of the ideas that today still govern the conduct of medical research, especially clinical trials.

The state of North Carolina's first department of statistics was established at what would become North Carolina State University, and its founder, Gertrude Cox, recruited Greenberg to teach biostatistics at UNC's schools of medicine and public health while he was still a graduate student at N.C. State. Soon, in 1949, he was the first faculty member appointed to a new department of biostatistics at UNC-Chapel Hill. "The two departments

were sort of viewed as one by Miss Cox, who founded them," Grizzle says. Students would often take classes at both institutions.

Soon after Greenberg was appointed chair, he began to grow the UNC department and the field by obtaining funding for training students and faculty. In 1953 the National Institutes of Health's National Heart Institute awarded one of NIH's first training grants to UNC to train students, postdoctoral fellows, and faculty and to bring in visiting faculty. "These training grants came about in part because Greenberg was asked by the NIH to help them evaluate research proposals from scientists," Grizzle says. "NIH became aware that the proposals needed statistical help in the worst way."

As Greenberg's research assistant, Grizzle worked on one of Greenberg's early consulting projects—a contract that the National Cancer Institute awarded UNC in 1955 to serve as the statistical coordinating center for the Southeastern Cooperative Cancer Chemotherapy Study Group. This group of 10 medical schools worked together to conduct clinical trials and test chemotherapy agents for treating leukemia and lymphoma. "Greenberg and I were among the first cancer chemotherapy statisticians," Grizzle says. The department also began to work on other projects, including a contract to provide statistical services to Veterans Administration hospitals evaluating four operations in the treatment of duodenal ulcers.

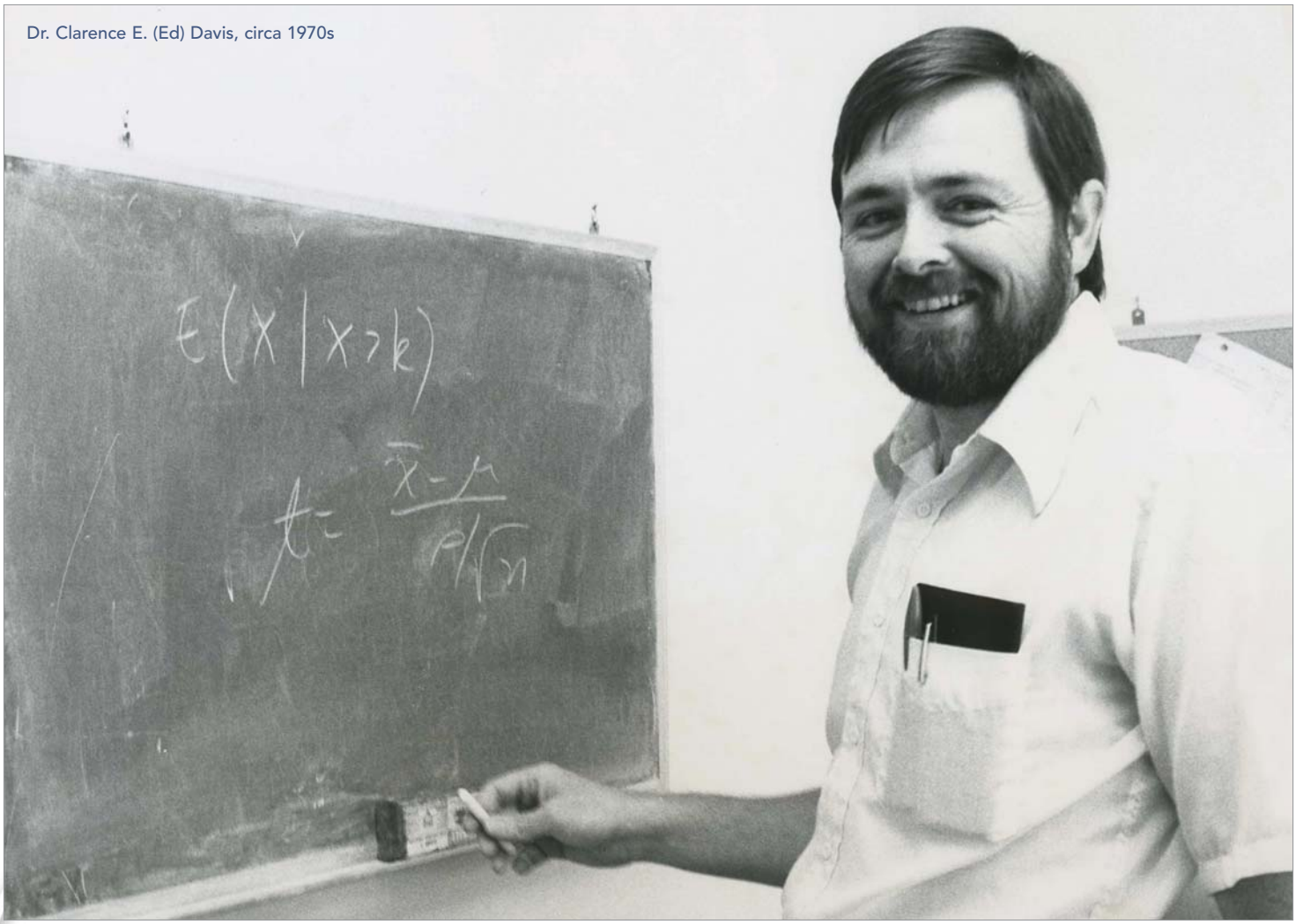
Laying the framework for clinical trials

As Greenberg did this early consulting work, he began to publish his ideas for putting rigor into designing studies and analyzing data in biomedical research. In 1959, he published in *The American Statistician* the first article describing the design and conduct of cooperative field and clinical trials. At the same time he was nurturing the people in the fledgling UNC department. "Greenberg was like a second father. He had an enormous impact on my life," Grizzle says. "I was essentially a country boy trying to get ahead. He kept throwing opportunity in my path, and he did the same for others on the fledgling faculty."

Dr. Bernard Greenberg, the first chair of UNC's Department of Biostatistics, chaired an NIH committee in 1967 directed to develop a process for conducting large, multi-center clinical studies. The landmark paper developed by the committee came to be known as the "Greenberg Report." The paper, commissioned by the NIH's National Heart Institute (now called the National Heart, Lung, and Blood Institute) outlined the way clinical trials should be run.

A bit later, in 1967, Greenberg chaired a committee appointed by NIH's National Advisory Heart Council, to come up with a process for conducting large, multi-center clinical studies. The paper that committee wrote came to be known as the "Greenberg Report." (See page 22 and *www.sph.unc.edu/cph/weblinks.*) The report was commissioned by the NIH's National Heart Institute (now called the National Heart, Lung, and Blood Institute).

"It was a landmark paper that basically outlined the way clinical trials should be run by the NIH," says Dr. Clarence "Ed" Davis,



Dr. Clarence E. (Ed) Davis, circa 1970s

Dr. Clarence E. (Ed) Davis, UNC research professor of biostatistics, was the primary biostatistician for the 1984 Lipid Research Clinics clinical trial — one of the first trials to show that lowering blood cholesterol reduces the risk of heart disease. Davis served as interim chair and chair of the Biostatistics Department from 1997 through 2005.

professor of biostatistics at the School who served as interim chair and chair of the department from 1997 to 2005.

Davis explains that up until the '50s and '60s, many studies were small enough that a student recording data in a lab notebook was enough. But things got a lot more complicated when researchers began studying the best way to treat cancer or prevent a heart attack.

"Those questions require studying hundreds or up into the thousands of people to compare therapies and see which ones are better," Davis says. "When you have more than one school or university collecting data for a study, you have to have some centralized way of doing that. That's the concept that Greenberg set out. It has allowed the NIH and other research institutes around the world to ask questions that can only be answered by very large studies."

The Greenberg report is also cited as introducing the idea that a formal independent committee should monitor data in clinical trials as it accumulates and should review interim analyses of the data.

In 1972, NIH began setting up coordinating centers to administrate large, multicenter trials. One of the first few centers was awarded to UNC, with Grizzle as principal investigator. Since then UNC's Collaborative Studies Coordinating Center (see page 22) has evolved to take on other issues such as design and data management for these large studies. Among other landmarks, the center coordinated one of the earliest studies to show that lowering cholesterol reduces the risk for heart disease, the Lipid Research Clinics study.

A personal interest

In 1972, the Collaborative Studies Coordinating Center (CSCC) was established, Greenberg become dean of the School of Public Health, and Grizzle took over as chair of the Department of Biostatistics. Davis joined the faculty that year as well, though he had first met Greenberg almost a decade earlier during a summer research program.

"Bernie (Greenberg) was one of the few people in the school who seemed to value every other part of the school," Davis says. "We all get involved in our particular areas, but Bernie was always very interested in all aspects of public health, not just the quantitative and biostatistics parts. We were all in awe of him because he seemed to be able to do everything well. I call him Bernie now, but

believe me when I met him walking in the hall, it was always 'Hello Dr. Greenberg' or 'Hello Dean Greenberg.'"

Greenberg took a personal interest in each faculty member. If someone had something happen, such as the death of a parent, Greenberg would send a handwritten note. "You wouldn't even know how he would know, but you would get a note," Davis says.

Greenberg nurtured the research life of faculty as well, encouraging them not only to consult with biomedical investigators, but to also advance the field through methodological research. Dr. Dana Quade, UNC professor emeritus of biostatistics, remembers this mentoring. He was an associate professor when, in 1968, he received a Research Career Development Award from the National Institute of General Medical Sciences. Quade says he was focused on teaching and never would have thought of applying for the grant if it weren't for Greenberg.

"Prior to that time I had almost no publications," Quade says. He had been a coauthor on a couple of publications with some investigators in epidemiology — studies in which Quade had helped analyze the data. But his only methodological publication was his PhD dissertation. "I had some ideas but no time to develop them. But by the time I was done with that award, I had several good papers in major journals of the field," Quade says. Those included a publication in the *Journal of the American Statistical Association.*

Today: data overload

More than 40 years after the Greenberg report, UNC biostatisticians still push the field forward, developing new tools to help investigators explore research questions that Greenberg probably couldn't have imagined. These questions in genomics and other areas require huge and complex sets of data — and the need for new statistical methodology to go with them, says Dr. Michael Kosorok, professor and chair of the department.

Dr. Barry Margolin served as chair of the Department of Biostatistics from 1987 through 1999 and director of the biostatistics facility at UNC's Lineberger Comprehensive Cancer Center from 1989 through 1999. His research specialty was the statistical design and analysis of experiments, particularly studies in genetic toxicology.

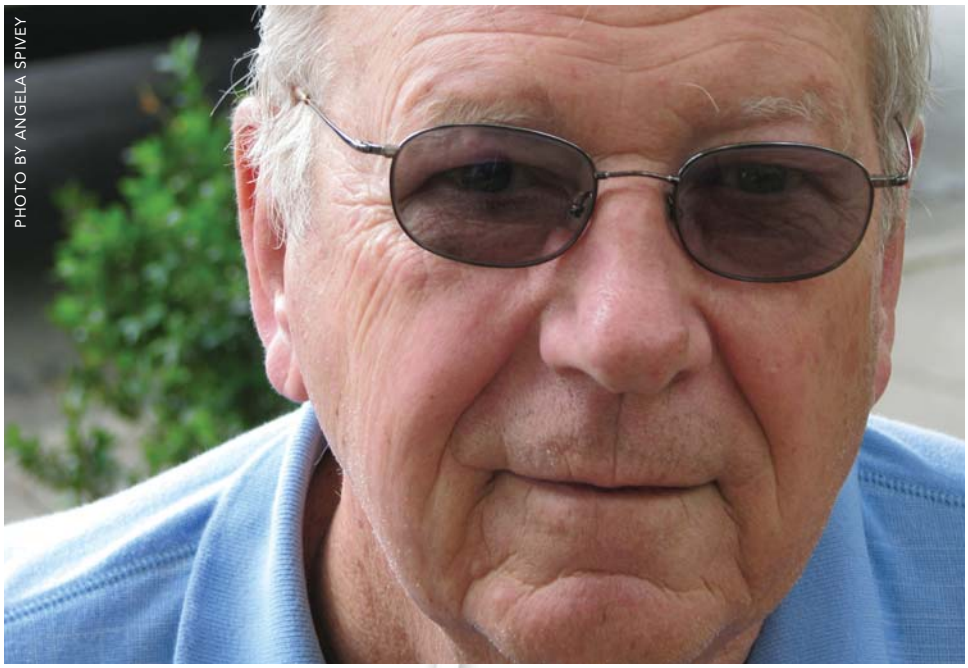


PHOTO BY ANGELA SPIVEY

Dr. Jim Grizzle, professor emeritus of biostatistics at Carolina, served as chair of the Biostatistics Department from 1973 through 1987. In the early years of the department, biostatisticians were looked on as "interlopers," Grizzle says. "We had to earn the researchers' respect. We were speaking a different language: things like 'randomization' and 'uniform ascertainment of outcomes.'"

Geneticists, for instance, use specially-made chips called microarrays to look for patterns in genes that could be related to disease. But one microarray can contain as many as 100,000 genes. Biostatisticians at UNC such as Kosorok and Dr. Fred Wright are working to develop statistically valid methods for analyzing this "high dimensional" data.

UNC faculty also work to quantify images used in cancer studies or neurological studies. Somehow those pictures have to be translated into a meaningful set of numbers. "There are very few biostatistics departments around the country that are investing a lot into brain imaging, but many of our faculty are working on questions such as how to use images to look at the relationship between risk factors for Alzheimer's and changes in brain structure," Kosorok says. Biostatistics faculty collaborate with scientists in computer science and operations research to develop methodologies to analyze such questions.

Designing and analyzing clinical trials still poses many challenges, and UNC is addressing many of them through the School's Center for Innovative Clinical trials (see page 29). For example, the field needs better methods for handling missing data, for conducting

post-marketing safety assessment of drugs, as well as methods for determining when to stop a trial early, either because of toxicity, futility or exceptional efficacy. "The main goal of the center is to form collaborations with industry, the FDA, and other academic institutions to work together to develop new methods to solve some of these big research problems," says Dr. Joseph Ibrahim, Alumni Distinguished Professor of biostatistics and director of the center.

Ibrahim studies Bayesian methods, which are gaining popularity in clinical trials because they provide sound mathematical tools for analyzing data about a problem while taking into account information from previous studies. In most Phase III clinical trials (the final step before drug approval), data is analyzed almost in a vacuum, as if previous studies don't exist. Ibrahim explains it this way: "Suppose you're driving a car and you come to the same stop sign nine days in a row, and each time the same car doesn't stop and crosses the line. The tenth time you come to that intersection, you're thinking 'I've seen this guy the last nine days fly through the stop sign, I better slow down. That's the essence of the Bayesian paradigm; it says I'm going to use this previous information to make my current decisions. But the classical paradigm essentially ignores that previous information and treats today's visit to the stop sign as if it's your first."

Most of the faculty such as Dr. Danyu Lin, Dennis Gillings Distinguished Professor of Biostatistics, develop new methodology—and computer programs that apply that methodology—to take on the ever more complex problems of genomics. "A new statistical method, no matter how wonderful it is, will not be used in practice unless there is software for it," Lin says. "So in recent years I have devoted more of my time and effort to develop software implementing my new methods."

Lin's methods and computer programs can be applied to genetic association studies, which try to relate diseases or conditions to genetic markers. For example, his SPREG program can be used to do secondary analyses in case-control association studies, in which researchers genotype large numbers of people for maybe 500,000 different genetic markers, then compare genetic markers in people with a disease such as diabetes to those without. But researchers often want to use the same data to analyze how these genetic markers affect other traits such as obesity or height. "If you do that with standard statistical methods, your analysis can be very biased," Lin says. Lin developed a method and the SPREG program to do those secondary analyses in a valid and efficient way.

> The work of UNC's biostatistics faculty is the right mix of theory, methodology, and applied research to push the field forward in these areas.

Ibrahim says UNC's department is strong because the faculty's consulting work feeds into their work developing new statistical methodology. For instance, a biostatistician will collaborate with oncologists on a melanoma trial and coauthor their paper in an oncology journal, then author a paper in a biostatistics journal about the methodology developed for that study. Methodology and practice are intertwined.

Kosorok says that UNC's faculty work is the right mix of theory, methodology, and applied research to push the field forward in these areas. Just as physicists have quantum theory, statisticians have their own theory—a body of tools and principles they use to construct new methods to be used in applications. While some biostatistics departments have no faculty at all working on theory, Kosorok says that proving current methods with theory is crucially important as research questions become more complex.

"Things are changing so rapidly that you really can't think about the theory all the time, you just have to do something that seems to work," Kosorok says. "But if we don't pay enough attention to the theory, we may end up coming up with numbers that don't mean as much as we think they do, that don't answer the scientific questions that we want to answer. So we need to also be thinking theoretically to find the proper way to handle these new, complicated situations. This is going to help us in the coming years to set the pace for the field of biostatistics." ■


PHOTO BY ANGELA SPIVEY

Dr. Michael R. Kosorok, professor and chair of the Department of Biostatistics at UNC's School of Public Health since 2006, has developed methods for using surrogate and multiple outcomes to increase cost-effectiveness of clinical trials. His work in cystic fibrosis has included being senior statistician on a large randomized trial which led to a change in national policy favoring nationwide newborn screening for cystic fibrosis.