Preface

You and I have some important and interesting conversations coming up shortly. However, I propose that we postpone those for a moment while I share with you my motivations for writing this book.

The frustrations of doctors, nurses, judges, legislators, and administrators that arise as they interpret healthcare research efforts are the unfortunate and predictable products of their meager research backgrounds. It is only human for them to grab for whatever supporting grips are available; one such handhold is the ubiquitous *p*-value.

This reduction of a research effort to a single number is regrettable, but quite understandable. The complexity of a modern healthcare research endeavor requires a clear understanding of the circumstances in which one can generalize results from relatively small samples to large populations. Even though the concept of generalization is nonmathematical, many researchers are not its master. Recognizing their disadvantage, they latch onto the *p*-value, believing that it neatly binds these complicated features into one tidy package.

However, like continually substituting desserts for nutritious meals, the habitual replacement of *p*-values for clarity of vision is unfulfilling and dangerous. This book reaches out to these principle-starved people. Specifically I want to use the ubiquity of the *p*-value as an overture to the discussion of statistical reasoning in medicine.

Statistical reasoning in medicine is the process by which one determines whether sample-based results can be extended or generalized to the population at large. The concepts are straightforward, intuitive, and quite precise. However, their application requires thoughtful consideration.

For many years the tendency in the research community has been to replace this deliberation with a quick and simple assessment of the *p*-value's magnitude. The research community, in its quest for significant results, has created a polluted sea of *p*-values in which we all restlessly swim. Although *p*-values were designed to make a simple statement about sampling error, for many they have become the final arbiter of research efforts.

Investigators often gnash their teeth over this entity's value at their study's conclusion: is it less than 0.05 or ≥ 0.05? To these workers, *p*-values are the switching signal for the research train. If the *p*-value is less than 0.05, the research-train moves down the main track of manuscript publication, grant awards, regulatory approval, and academic promotion. However, if the *p*-value is greater than 0.05, the switch moves the other way, directing the research train off to the elephant's graveyard of discarded and useless studies. Replacing the careful consideration of a research effort's (1) methodology, (2) sample size, (3) magnitude of the effect of interest, and (4) variability of that effect size with a simple, hasty look at the *p*-value is a scientific thought-crime.

P-values continue to be the focus of research discussions in academic centers, remaining a staple of the medical community's research effort. The approval of a new medical intervention commonly includes consideration of the *p*-value, and arguments in courts of law for the scientific basis of an assertion frequently concentrate on the size of the *p*-value. Clearly, many researchers, journal editors, regulators, and judges cling doggedly to its use. It is therefore all the more curious that so few of these specialists understand either what the *p*-value is or precisely what information it is designed to convey. Although they understand the message that the *p*-value "had better be less than oh five," there is little understanding of either the source or justification of this ubiquitous mantra.

I don't think we statisticians have been as helpful as possible. A biostatistics professor at a school of public health once asked a statistics student sitting for his qualifying exam (that must be passed to enter the Ph.D. candidacy phase), "Explain what a *p*-value means." The professor never received a satisfactory response.[*] When biostatisticians do respond to this question, we often give the following response, "the *p*-value is the conditional probability of rejecting the null hypothesis in favor of the alternative hypothesis when the null hypothesis is true." I fear that to the non-statistical world, this answer smacks of Orwellian double-speak.

This text emphasizes an intuitive understanding of the role of the *p*-value in sample-based research, deemphasizing the underlying mathematics. This nonmathematical approach is available when the foundation principles of statistical reasoning in medicine are clearly articulated. Our purpose here is to clearly state and develop the principles that govern when and how one takes results from a small sample and applies them to a larger population in healthcare research. The enunciation of these principles brings the roles and limitations of *p*-values into sharp focus.

Lemuel A. Moyé
University of Texas
School of Public Health
April, 2006

[*] Related by Dr. Sharon Cooper, Chair of Epidemiology and Biostatistics, Texas A&M Rural School of Public Health.

Acknowledgments

Approximately 8,000 North Carolina Cherokee are descended from those who did not take the *Nunadautsun't* to the Oklahoma reservations; one who remained was my grandmother. An important lesson she taught her grandchildren was never to begin the day until you have given thanks to God for the people He has placed in your life to guide you.

I owe a special debt of thanks to the leaders and members of the Cardio-Renal Advisory Committee of the Food and Drug Administration, with whom I served as a statistician from 1995 to 1999. They are Ray Lipicki, JoAnn Lindenfeld, Alastair Wood, Alexander Shepherd, Barry Massie, Cynthia L. Raehl, Michael Weber, Cindy Grimes, Udho Thadani, Robert Califf, John DiMarco, Marvin Konstam, Milton Packer, and Dan Roden. The public sessions held three times a year are among the best training grounds for cardiovascular research epidemiology and biostatistics. I am also indebted to Lloyd Fisher, with whom I have often disagreed and from whom I have always learned.

I am indebted to my alma mater, The University of Texas School of Public Health, and to public school education in general, which has provided much of my education. To all of you, and especially Asha Kapadia, Palmer Beasley, Robert Hardy, Charles Ford, Barry Davis, Darwin Labarthe, Richard Shekelle, Fred Annegers, and C. Morton Hawkins, consider this book an early payment on a debt too large to ever repay. I would also like to thank Craig Pratt and John Mahmarian, who have been patiently supportive in the development of my ideas. Furthermore, I would like to express my gratitude to Marc Pfeffer, Frank Sacks, Eugene Braunwald – you have been more influential than you know.

To the fellow classmates and friends known as the "DELTS"— Derrick Taylor, MD, Ernest Vanterpool, and Tyrone Perkins. Ever since high school in Queens, New York, we have pushed and prodded each other to excellence. You are a part of all my worthwhile endeavors.

Additionally, I have been blessed with friends and colleagues who were always willing to share from their wellspring of good sense with me. Four of them are James Powell of Cincinnati, Ohio, John McKnight of Washington D.C., Joseph Mayfield of New York City; and Gilbert Ramirez, now in Iowa.

John Kimmel, the reviewing editor, and many chapter reviewers have provided good questions and additional references, improving the book's structure. In addition, Dr. Sarah Baraniuk, Dr. Yolanda Muñoz, and Dr. Claudia Pedroza each suggested important improvements in the last chapters of this book. Dr. Steve Kelder provided some especially helpful insights for the chapter on epidemiology.Lisa Inchani, a graduate student of mine, selflessly volunteered her time to patiently review and copyedit this text. Its readability is largely due to her steadfast efforts.

Finally, my dearest thanks go to Dixie, my wife, on whose personality, character, love, and common sense I have come to rely, and to my daughters Flora and Bella, whose continued emotional and spiritual growth reveals anew to me each day that, through God, all things are possible.

Lemuel A. Moyé
University of Texas
School of Public Health
April, 2006