STATISTICIANS AND THE EVOLUTION OF THE RANDOMIZED CLINICAL TRIAL

Susan S. Ellenberg
Perelman School of Medicine
University of Pennsylvania

Virginia Chapter of the ASA
Annual Meeting
October 25, 2019

WHAT IS A CLINICAL TRIAL?

- A clinical trial is an experiment to evaluate the effect of an intervention on a human being
- We have learned a lot over the past decades about optimal approaches to clinical trials, including avoidance of pitfalls in the design, conduct, analysis and reporting of trials that undermine reliability of results
- Statisticians have been highly influential in this process

MY JOURNEY IN STATISTICS

- Like most of us—I did not grow up planning to be a statistician
- I got a masters degree in secondary mathematics education, and taught high school math for 3 years, until our first child came along
- Janet Wittes was working as a research associate for Jerry Cornfield, who had recently retired as the Chief of the Biometrics Research Branch at the (then) National Heart Institute
- They needed a programmer
- This was 1971

MY INTRODUCTION TO CLINICAL TRIALS

- Cornfield was a leading statistician of his time
- He was in great demand for service on many clinical trial executive committees and steering committees
- He brought Janet and me to meetings of these committees
- I started taking statistics courses at GWU
- After a few years I was assigned a clinical trial of my own
- I was hooked! And I have been involved in the world of clinical trials ever since.

PRELIMINARY ALERT!

- The evaluation of clinical trial data has relied primarily on frequentist significance testing
- Much of what follows relates to appropriate and inappropriate application of significance testing to data from clinical trials
- I will comment on the current debate about the value of significance testing at the end of this talk

THE CONCEPT OF RANDOMIZED CLINICAL TRIALS GOES WAY BACK

"Let us take out of the Hospitals, out of the camps, or from elsewhere, two hundred, or five hundred poor people, that have fevers, pleurisies, etc. Let us divide them into halfes, let us cast lots, that one halfe of them may fall to my share, and the other to yours; I will cure them without bloodletting and sensible evacuation; but do you do, as ye know...We shall see how many funerals both of us shall have."

Van Helmont, *Oriatrike*, 1662

CLINICAL TRIALS BEFORE THE MID-20TH CENTURY

- Observation of results of new treatment approaches, compared to historical data or impressions
- Before-and-after comparisons
- Early 20th century: more systematic methods came into use
  - alternating treatment assignments
  - Assignments by first letter of last name
  - Assignment by day of the month on which patient presented
- Randomization: "paradigm shift"

THE CONCEPT OF RANDOMIZATION

- Fisher introduced randomization into agricultural experiments in 1920's
- Fisher recognized that testing soil treatments, etc., on systematically selected plots left it to the investigator's judgment to "balance" for all other factors that might affect outcome
- Randomization permitted assumption that there were no differences between treatment groups--except for the treatment itself

RANDOMIZATION IN MEDICAL RESEARCH

- Austin Bradford Hill was professor of medical statistics at the London School of Hygiene and Tropical Medicine
- Recognized, and was frustrated by, the inevitable confounding of treatment effects with other factors in observational studies of medical treatments
- Study of streptomycin for treatment of tuberculosis, late 1940s, provided opportunity for use of randomized allocation
  - Inadequate supply; randomization seemed fairest way of allocating treatment
- Widely viewed as beginning of modern era of RCTs

IN THE BEGINNING…

- The National Institutes of Health and other federal agencies began to fund randomized clinical trials
- One of the first NIH randomized clinical trials was a study of oxygen use in premature infants (1953)
  - Many premature infants required supplementary oxygen to survive
  - Some astute physicians noticed a rise in retrolental fibroplasia (now called retinopathy of prematurity) concurrent with the increased use of oxygen in the 1940s—condition often led to blindness
  - Observational evidence was not consistent
- Trial showed that lower oxygen levels greatly reduced eye problems without adversely affecting survival

SALK POLIO VACCINE (1954)

- Vaccine was to be studied in multiple states, under auspices of departments of public health
- Approach to testing the vaccine was controversial
  - Many states rejected the notion of randomization and use of a placebo control, and chose to vaccinate only second graders, using first and third graders as "observed controls"
  - Other states were persuaded that a randomized trial was the best way to get a reliable answer
  - Over 400,000 children randomized in the 11 states that elected to randomize
  - One of Paul Meier's first projects

THE GREENBERG REPORT

- Bernard Greenberg was the founding chair (1949) of the Department of Biostatistics at the University of North Carolina School of Public Health and an early advocate for randomized clinical trials
- As the NIH increased its support for clinical trials, Greenberg was asked by the then-National Heart Institute to chair a committee to provide recommendations on the conduct of (multicenter) randomized clinical trials
- This report, issued in 1967, focused on the important roles of a leadership group (steering/executive committee) and the data coordinating center
  - Published *Controlled Clinical Trials*, 1988

REGULATORY STANDARDS

- Until passage of the Federal Food, Drug and Cosmetic Act in 1938, drug manufacturers could market their products without prior approval by the FDA
- Up through mid-20th century, manufacturers only had to demonstrate safety of the product
- Many drugs approved with little to no evidence that they had any benefit
- Companies began performing randomized trials in the 1950s but they didn't have to, and often didn't
- Requirement to demonstrate efficacy with "adequate and well-controlled trials" instituted in 1962
  - Kefauver-Harris Amendment to the Food, Drug and Cosmetic Act

EARLY CHALLENGES

- Pushback from researchers
  - Assigning treatments by coin flips is unethical

EARLY CHALLENGES

- Pushback from researchers
  - Assigning treatments by coin flips is unethical
  - The new treatment is likely to be much better

EARLY CHALLENGES

- Pushback from researchers
  - Assigning treatments by coin flips is unethical
  - The new treatment is likely to be much better
  - Randomized studies take too much time and are too expensive

EARLY CHALLENGES

- Pushback from researchers
  - Assigning treatments by coin flips is unethical
  - The new treatment is likely to be much better
  - Randomized studies take too much time and are too expensive
  - "We didn't need randomized trials to know that penicillin worked!"

EARLY CHALLENGES

- Pushback from researchers
  - Assigning treatments by coin flips is unethical
  - The new treatment is likely to be much better
  - Randomized studies take too much time and are too expensive
  - "We didn't need randomized trials to know that penicillin worked!"

PUSHBACK: ETHICS

- Concerns about the ethics of randomization patients
  - In the course of treatment, the physician is obligated to the patient and to no one else. He is not the agent of society, nor of the interests of medical science, the patient's family, the patient's co-sufferers, or future sufferers from the same disease. The patient alone counts when he is under the physician's care. Jonas, *Daedalus*, 1969
- Particular resistance from cancer researchers
  - Results from the study of last year's potential wonder drug were not as great as had been hoped; much better for everyone to have the opportunity to be treated with a potentially better regimen
  - Researchers proposed carefully done historically controlled studies; Gehan and Freireich, *New England Journal of Medicine*, 1974

PUSHBACK: THERE MUST BE OTHER WAYS

- Some argued that alternative approaches could be used to avoid bias and thereby sidestep the discomfiting ethical concerns
- Weinstein, *NEJM*, 1974
  - : …to control for variables that can be identified…as interfering factors, matching, blocking or adjusting may be far more efficient…than purely randomizing."
- Hellman and Hellman, *NEJM*, 1991
  - "It is fallacious to suggest that only the randomized clinical trial can provide valid information or that all information acquired by this technique is valid."

PUSHBACK: EFFICIENCY

- Improved computing capabilities in the 1980s led to proposals that large databases of patient experiences could often substitute for randomized trials
  - Horwitz RI and Feinstein AN: "Improved observational method for studying therapeutic efficacy," *JAMA*, 1981
  - Kunitz et al: "The Pilot Stroke Data Bank," *Stroke*, 1984
  - Starmer et al: "On the complexity of investigating chronic illness," *Biometrics*, 1980

DEFENSE OF THE RCT

- Many responses to such pushback
  - Byar et al: "Randomized clinical trials: perspectives on some recent ideas," *New England Journal of Medicine*, 1976
  - Ellenberg and Dambrosia: "Statistical considerations for a medical database," *Biometrics*, 1980
  - Byar and Green: "Using observational data from registries to compare treatments: The fallacy of omnimetrics," *Statistics in Medicine*, 1984
- Advocates for RCTs greatly outnumbered the dissenters
- RCTs became entrenched as the "gold standard" for evaluating clinical interventions
- But…experience in early trials revealed pitfalls that needed to be addressed

EARLY CHALLENGES

- Learning all the ways that validity of the randomized comparison could be compromised

SOME THINGS WE LEARNED PRETTY QUICKLY

- How do we deal with mid-trial results that are much stronger than—or in the opposite direction from—what we anticipated?

THE UNIVERSITY GROUP DIABETES PROJECT

- This was an early multicenter placebo-controlled trial to see whether certain antidiabetic agents would reduce cardiovascular complications of diabetes
- Study began entering patients in 1961
- In 1969, the investigators noticed an apparent excess mortality on one of the 3 active agent arms
- They had developed a monitoring plan based on simulations using historical data on mortality in diabetics
- They stopped the trial but mortality data did not meet conventional levels of statistical significance
- There was a huge uproar from the drug manufacturer, and from diabetologists, who loved this drug

EARLY MONITORING DESIGNS

- Recognition of need for statistically based monitoring procedures emerged in 1950's
- Two schools of thought
  - Frequentist: base clinical trial monitoring plans and decision-making on control of Type 1 and Type 2 errors
  - Bayesian: base decision-making on the "likelihood principle"

ARMITAGE AND CORNFIELD

- Peter Armitage (UK)
  - Began publishing on sequential trials in 1950's
  - First book on topic, *Sequential Medical Trials,* appeared in 1960
- Jerome Cornfield (US)
  - Investigated sequential methods from Bayesian perspective
  - Key publication: "Sequential trials, sequential analysis and the likelihood principle," *American Statistician*, 1966

INTERESTING READING

- Lengthy review of Armitage's book published in 1963 in *Journal of the American Statistical Association* by FW Anscombe; gave Bayesian critique
- Response by Armitage in same issue
- Despite advocacy for Bayesian methods by eminent statisticians such as Anscombe, Savage and Cornfield, frequentist methods became the standard in clinical research
  - Computationally simpler
  - Concerns about prior beliefs influencing the formal evaluation of data

STATISTICIANS WERE DEVELOPING MORE TRANSPARENT MONITORING APPROACHES

- In 1969, Peter Armitage published a design based on pairs of subjects randomly assigned to two study arms
- Marvin Zelen published his "play the winner" paper the same year
- Neither approach was really practical for most trials
- In the meantime, common practice was to simply keep looking at the data and stopping the trial as soon as the magic "$p < 0.05$" was hit

GETTING CLINICIANS' ATTENTION

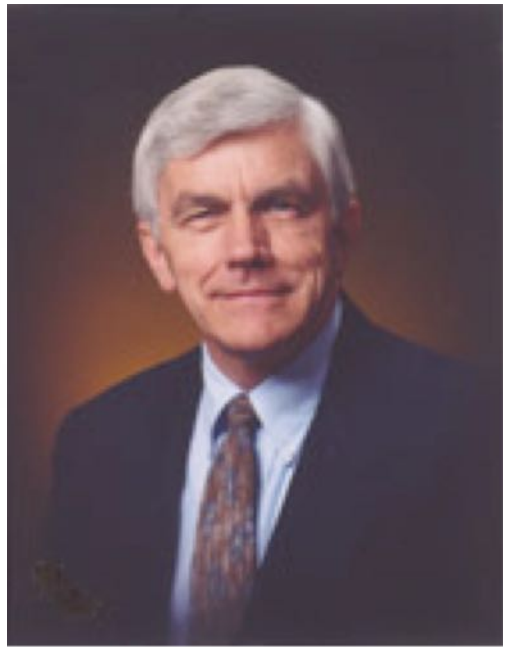
- Armitage and others had pointed out the inferential problems associated with regular testing at the nominal level
- Most statisticians understood the problem, but many clinicians worried about the ethics of continuing a trial once the results appeared nominally significant
- When Armitage's student Klim McPherson published an eminently readable article in the *New England Journal of Medicine* in 1974, showing how much the false positive rate could be inflated, clinical investigators finally took notice

REPEATED TESTING INFLATES TYPE I ERROR

significance level (%)	No. of repeated tests						
	1	2	3	4	5	10	200
1	1	1.8	2.4	2.9	3.3	4.7	12.6
5	5	8.3	10.7	12.6	14.2	19.3	42.4

McPherson K, *New England Journal of Medicine*; 290:501-2, 1974



NEW APPROACHES WERE DEVELOPED

- Haybittle 1971 (popularized by Peto 1976)
  - Look at interim data as often as you want but don't stop early unless interim difference is very extreme
- Pocock 1977; O'Brien and Fleming, 1979
  - Group sequential designs: look at interim results a fixed number of times; use p-values calculated to control overall type 1 error
- Lan and DeMets 1983
  - Provided basis for flexibility in number of interim looks
- All these methods could be applied to most trials requiring regular interim monitoring with the possibility of early termination for definitive results
- All these methods were readily understandable by clinical trials investigators and were simple to implement

SOME THINGS WE LEARNED PRETTY QUICKLY

- How do we deal with mid-trial results that are much stronger—or in the opposite direction—than we anticipated?
- What do we do about people who don't adhere to their randomized assignment or who are later found to have been ineligible?

THE PROBLEM OF EXCLUSIONS

- In the early days of clinical trials, not uncommon for investigators to exclude some information from analysis
  - The trial subject stopped (or never started) taking the assigned treatment
  - The trial subject added another treatment
  - A basis for ineligibility had been overlooked
  - Reported event did not meet all criteria
- Investigators recognized that including this information would dilute the observable treatment effect
- They did not appreciate that excluding the information could bias the results

EXAMPLE (circa 1980)

- Randomized trial of cancer therapy following surgery to remove tumor
- Control group: no further therapy
- Not blinded—side effects of chemotherapy would reveal treatment
- Protocol called for treatment to commence no later than 6 weeks post-surgery
  - Rationale for therapy is that it will kill any remaining cancer not removed at surgery
  - If therapy not started shortly after surgery, won't work—including such patients will dilute treatment effect

EXAMPLE (cont.)

- Study chair wanted to exclude all subjects randomized to receive therapy who did not start until more than 6 weeks post-surgery
- It took me a long time to persuade him that this was a bad idea

EXAMPLE (cont.)

- Study chair wanted to exclude all subjects randomized to receive therapy who did not start until more than 6 weeks post-surgery
- It took me a long time to persuade him that this was a bad idea
  - Only those assigned to post-surgical treatment are at risk of being excluded

EXAMPLE (cont.)

- Study chair wanted to exclude all subjects randomized to receive therapy who did not start until more than 6 weeks post-surgery
- It took me a long time to persuade him that this was a bad idea
  - Only those assigned to post-surgical treatment are at risk of being excluded
  - What if those who start therapy late are those who had the most extensive surgery and thus required longer recovery period?

EXAMPLE (cont.)

- Study chair wanted to exclude all subjects randomized to receive therapy who did not start until more than 6 weeks post-surgery
- It took me a long time to persuade him that this was a bad idea
  - Only those assigned to post-surgical treatment are at risk of being excluded
  - What if those who start therapy late are those who had the most extensive surgery and thus required longer recovery period?
  - What if those with most extensive surgery are most likely to have remaining unseen cancer?

MANY EXAMPLES

- Common for investigators reviewing data to toss out "unevaluable" patients
  - Especially pernicious in single-arm trials of new anticancer drugs when such exclusions bumped up success rate ("How to Succeed in Clinical Trials Without Really Trying")
    - Moertel, *Statistics in Medicine*, 1984: " …[if] you do not really get a breakthrough type response rate, you can always use the 'unevaluable patient gambit'. That works exceptionally well in Phase II trials. The technique is to take those patients who quickly get too sick to continue treatment, or who die shortly after treatment is initiated, and discard them from your analysis as unevaluable. That manoeuvre very neatly contracts your denominator and greatly inflates your response rate."

ADHERENCE TO MEDICATION

- Similar issue: how to handle subjects who discontinue assigned treatment
- Very common approach, even today—just analyze those who stayed on treatment, or who received at least some fraction of assigned treatment
- Big problem: those who adhere to treatment are not a random sample of the randomized study population
- Very famous example: NIH study of the 1970s

CORONARY DRUG PROJECT

FIVE-YEAR MORTALITY BY ADHERENCE TO CLOFIBRATE AND PLACEBO

	Clofibrate		Placebo	
Adherence	N	% mortality	N	% mortality
<80%	357	24.6	882	28.2
≥80%	708	15.0	1813	15.1

Coronary Drug Project Research Group, *NEJM*, 1980

ADHERENCE TO ASSIGNED TREATMENT

- Huge issue in clinical trials
- We want to know how people will do on treatment A, as compared to control
- People assigned to treatment A who refuse or stop treatment A create problems
  - They will have no (or a reduced) treatment effect, so won't give a true picture of how people do on the treatment
  - But if we remove them, we may bias our comparisons because their prognosis may be different from those who accept assigned treatment

INTENTION-TO-TREAT

- "Intention-to-treat" approach established early in modern era of clinical trials to prevent bias
  - Randomization produces groups that should be prognostically comparable
  - Strict analysis of groups as randomized preserves validity of comparison
- Well-established today (especially in regulatory setting) as the primary approach to comparison of outcomes by treatment

INTENTION-TO-TREAT

- First use I found in Pub Med: 1978 article in the *British Heart* Journal (Peter et al) describing results of a randomized trial of treatment for myocardial infarction
  - 3 mentions 1975-80 (all in cardiology)
  - 17 mentions 1981-85
  - 94 mentions 1986-90
  - 326 mentions 1991-95
  -
  - 6663 mentions 2011-15 (2044% increase vs 107% increase in Pub Med citations)

LOOKING FOR BETTER WAYS OF ACCOUNTING FOR ADHERENCE

- Efron and Feldman (1991) provided an early approach applied to data from a placebo-controlled trial of a cholesterol-lowering drug
  - Used a causal model as described by Rubin in the social science literature, with adherence as an outcome along with outcome of clinical interest
  - Problematic assumption: non-adherence is an inherent subject characteristic unaffected by treatment
- While underlying assumption was largely felt to be too stringent, this was a very innovative step and led to much further work using the causal model approach

MODELING ADHERENCE PATTERNS

- Methods for incorporating time-varying covariates were not available at the time the Coronary Drug Project data were analyzed
- Murray and Hernan (2016, 2018) reanalyzed these data, showing that using such methods, along with inverse probability weighting, virtually eliminates the mortality difference between adherers and non-adherers
- These models are very useful in helping to understand the potential impact of poor adherence—but results still depend on underlying assumptions of model

REGULATORY PERSPECTIVE

- The FDA and other regulatory authorities have long insisted on analyses using the "intention-to-treat" approach
- Recently issued draft guidance on "estimands" indicates recognition that in many situations a pure intention-to-treat approach doesn't really provide the result of interest
- More relevant (but still valid) results can be obtained by carefully matching estimation approach to the desired estimand, and performing appropriate sensitivity analyses

SOME THINGS WE LEARNED PRETTY QUICKLY

- How do we deal with mid-trial results that are much stronger—or in the opposite direction—than we anticipated?
- What do we do about people who don't adhere to their randomized assignment?
- What about people who get randomized but drop out so we don't know what happened to them?

THE HEARTBREAK OF MISSING DATA

- Missing data is an unfortunate fact of life in clinical trials
- Trying to ensure that you can still draw valid conclusions in the presence of missing data is also an unfortunate fact of life in clinical trials
- For many years, the problem was mostly ignored; data analysis was based on the data at hand
- Some simple approaches acknowledged the problem
  - Comparing baseline characteristics for those with missing and nonmissing outcomes
  - For studies assessing outcomes on a regular basis and with the primary analysis to be conducted at a fixed point in time, using the last value measured as the final value

MAJOR STEPS TOWARD BETTER APPROACHES

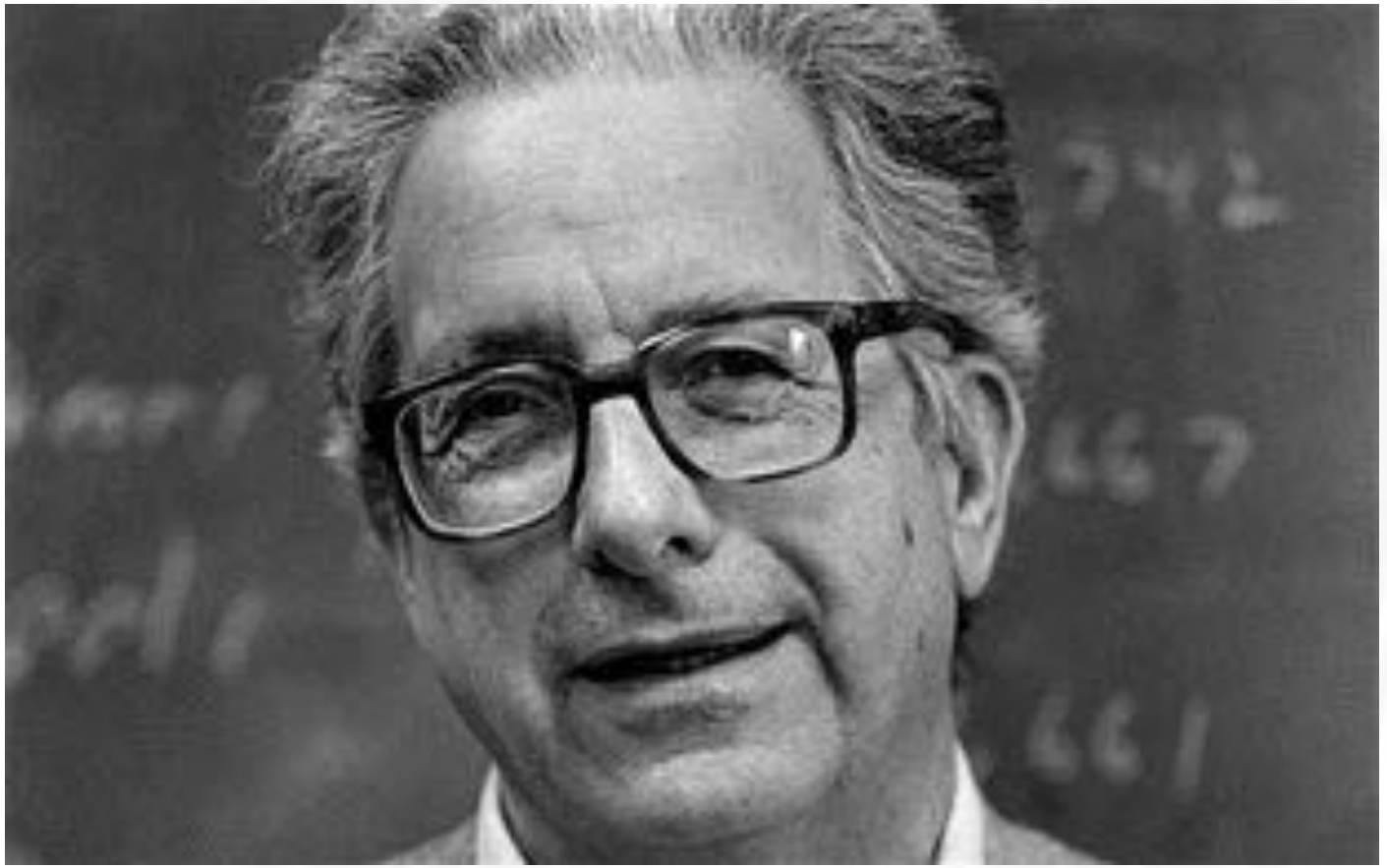
- Donald Rubin introduced the concept of data "missing at random" in a 1976 paper
- Offered a modeling approach that under certain assumptions would provide valid estimates of the missing outcomes
- These assumptions more plausible than those for analysis that ignored missing data, or used "last observation carried forward"
- Rubin's method of multiple imputation handled concern that simple imputation using baseline or last values, or control group averages, etc., did not only inadequately deal with bias, but underestimated variance
- Methods of causal inference have provided ways to help understand the potential influence of missing data on estimates of treatment effects

THE PREVENTION AND TREATMENT OF MISSING DATA IN CLINICAL TRIALS

- National Academy of Sciences report (2010)
- Meticulous and clearly written summary of approaches to missing data in clinical trials
- Led by Rod Little; other statisticians contributing to the report were Michael Cohen, Ralph D'Agostino, Kay Dickersin, Scott Emerson, Constantine Frangakis, Joe Hogan, Geert Molenberghs, Susan Murphy, Jim Neaton, Andrea Rotnitzky, Dan Scharfstein, Joe Shih, Hal Stern

IMPORTANT ADVANCES

- We have developed increasingly useful and powerful ways of studying health outcomes
  - Survival analysis
  - Repeated measures analysis
  - Meta-analysis
  - Cluster-randomized studies
  - Adaptive designs



SURVIVAL ANALYSIS

- Prior to the 1960s, long-term survival data was analyzed using landmark approach—eg, 5 year survival rates
- No good way to handle those who had dropped out, or were lost-to-follow-up, or hadn't yet reached the landmark
- Development of life-table methods was a major advance
  - Cutler and Ederer, *Journal of Chronic Diseases*, 1958
- The product-limit estimate (Kaplan and Meier, *JASA* 1958) allowed much more precise estimates of survival probability at any time point (and remains the primary graphical depiction of survival data to the present day)

SURVIVAL ANALYSIS: TWO MAJOR ADVANCES

- 1966: extension of the Mantel-Haenszel test to survival data (Mantel, *Cancer Chemotherapy Reports*, 1966)
  - Became known as the logrank test (per Peto and Peto, 1972)
  - Simple case of soon-to-be-seen Cox model
- 1972: the proportional hazards model (Cox, *JRSS B*, 1972)
  - Huge impact on analysis of long-term studies with dichotomous outcomes
- These remain the primary tools for those conducting and analyzing studies with time-to-event endpoints




REPEATED MEASURES

- The objective of many clinical trials is to assess improvement of a chronic condition over time
  - Reduction in seizure frequency
  - Reduction in frequency of asthma exacerbations
  - Improvement on quality of life measures
- This requires multiple outcome measures over the course of the trial
- Must account for correlation of within-person measures
- Mixed model analysis has become standard in clinical trials looking at longitudinal data

META-ANALYSIS

- Approaches to integrating data from multiple studies were developed by Cochran, following Fisher, in the 1930s
- These techniques were first popularized in the social sciences by Gene Glass, who coined the term "meta-analysis," in the 1970s
- In the 1980s, Richard Peto and colleagues conducted in-depth meta-analyses (or "overviews") of treatments for myocardial infarction, leading to large randomized trials that provided definitive evidence of benefit of these treatments
- At the same time, the statistical methodology for performing meta-analyses was being worked out
  - Hedges and Olkin, *Statistical Methods for Meta-Analysis*, 1985
- Meta-analysis of clinical trials now widely accepted as a valuable approach to assessing a series of similar studies

CLUSTER-RANDOMIZED TRIALS

- Early publication: Cornfield, "Randomization by group: a formal analysis," *American Journal of Epidemiology*, 1978
  - Recognized need to study interventions that could only be delivered to groups: classroom techniques, hospital infection control protocols
- Donner has been the research pioneer in this area
- With recent emphasis on "pragmatic" trials, cluster-randomized trials are of increasing interest

ADAPTATIVE DESIGNS

- Clinical trials have always been, and will always have to be adaptive
- Early adaptive approaches
  - Multi-stage designs (Gehan, *J Chron Dis*, 1961)
  - Play the winner (Zelen, *JASA*, 1969)
  - Sequential designs
- Newer adaptive approaches
  - Bayesian adaptive designs (Berry)
  - Multi-arm, multi-stage designs (Parmar)
  - Sequential multi-arm trials (SMART) (Murphy)

ADAPTIVE DESIGNS

- Less controversial adaptations
  - Early trial termination for benefit, harm or futility
  - Modification of sample size based on overall event rate
- More controversial: changing the randomization ratio
  - This concept considered early in the modern clinical trials era: Cornfield, Halperin, Greenhouse: "An adaptive procedure for sequential clinical trials," *JASA*, 1969
  - Studied and advocated by Rosenberger (1990s)
  - Essential component of current generation of Bayesian adaptive designs
- More controversial: enlarging the sample size based on interim comparisons

SIGNIFICANCE TESTING

- Classical frequentist analysis has been central to the world of clinical trials
- There are unquestionably many issues about how significance testing is and has been used and misused
- Overall, however, the discipline of designing experiments based on limiting the probabilities of error in drawing conclusions about the efficacy and safety of medical treatments has served the public health well
- Any approach to statistical inference will be insufficiently understood by many and misused by some
- Significance testing is just a tool to support decision-making; it has been used successfully to identify many effective treatments, and to prevent widespread use of ineffective and even harmful treatments
- Cornfield: "We all do significance tests, and then go on to do more interesting things."

A FEW REMAINING COMMENTS

SOCIETY FOR CLINICAL TRIALS

- Established in 1978 "To promote the development and exchange of information for design and conduct of clinical trials…"
- Board of Directors
  - First Board: 12 men
  - Current Board: 6 men, 8 women
- Presidents
  - First woman president Genell Knatterud, 1991
  - Of first 20 presidents: 3 women
  - Of most recent 20 presidents: 7 women

SOME PERCEPTIONS OF THE ROLE STATISTICIANS PLAY IN CLINICAL TRIALS

DONALD FREDERICKSON, 1968*

- "The anarchy of guess and intuition [in the design of clinical trials] has given way to a benevolent tyranny of statisticians."

*Frederickson DS, The field trial: some thoughts on the indispensable ordeal." *Bulletin of the New York Academy of Medicine*, 1968

BERNIE FISHER, 1980s

- "Statisticians are the terrorists of clinical trials!"

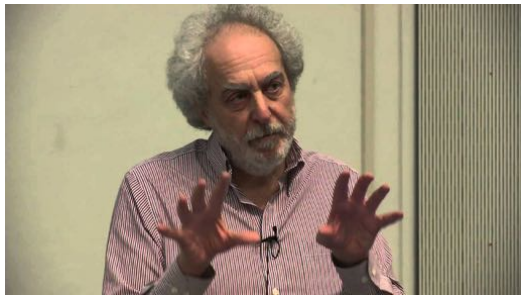
ALVAN FEINSTEIN, 1983

- Clinician perspective on clinical trials: **pragmatic**
  - Incorporate heterogeneity of ordinary clinical practice
- Statistician perspective on clinical trials: **fastidious**
  - Insistence on homogeneity of trial population and study conduct

ON THE OTHER HAND

- *The Progress of Experiment*, Harry Marks, Cambridge U Press, 1997
  - Part II: Of Methods and Institutions, or the Triumph of Statistics
  - Marks give statisticians enormous credit for the move to randomized trials in the mid-20th century, and the consequent increased reliability of the results of clinical investigations

STATISTICIANS NEED TO BE TEACHERS



STATISTICIANS AS TEACHERS

- Statisticians were the first scientists to welcome AIDS activists into the world of medical research. They taught us how to look at data, to discern the real from the apparent, to weigh the strengths and weaknesses of analyses, how statistics can help us escape our own biases, how data can be manipulated with the best of intentions.
- They gave us a set of tremendous tools with which to investigate and probe data on new drugs being developed by pharmaceutical companies, gave us the ability to sit on FDA panels to provide a rigorous community-based assessment of these drugs as they came up for approval, gave us the power to go toe-to-toe with researchers in the design and analysis of clinical trials.

Gregg Gonsalves, former ACT UP leader, now a professor at Yale and a recent McArthur winner, on the death of Paul Meier

THANKS TO ALL MY WONDERFUL MENTORS

- Jerome Cornfield
- Janet Wittes
- Sam Greenhouse
- Max Halperin
- Nathan Mantel
- Larry Shaw
- Richard Simon
- Judy O'Fallon
- Judy Goldberg
- Tony Lachenbruch
- Dave DeMets
- Ed Gehan
- Genell Knatterud
- Dave Byar
- Steve Lagakos
- Paul Meier
- Lynne Billard
- Barry Margolin
- Curt Meinert
- Bob O'Neill
- Dick Landis
- And of course…JONAS!