OVERSIGHT OF PRAGMATIC RANDOMIZED CLINICAL TRIALS

Susan S. Ellenberg, Ph.D.
Perelman School of Medicine
University of Pennsylvania

National Cancer Policy Forum Workshop
Contemporary Issues in Human Subjects Protection
February 25, 2014

Penn Medicine

WHAT IS A "PRAGMATIC" TRIAL?

- Concept first introduced by Schwartz and Lellouch, 1967*
- Made distinction between "explanatory" and "pragmatic" trials
- Explanatory trials
  - Purpose is to answer a scientific question
  - Implication: conduct trial controlling heterogeneity as much as possible so as to isolate treatment effect
- Pragmatic trials
  - Purpose is to answer a practical question: which treatment to use
  - Implication: conduct trial under "real world" conditions
  - Results intended to be widely generalizable

*Schwartz and Lellouch, *J Chron Dis*, 1967

TERMINOLOGY

- Starting in the 1980s, lots of discussion about "large, simple trials"*
  - Answering important health care questions
  - Identifying small but worthwhile improvements in major outcomes in common diseases
  - Confirming potentially important conclusions from meta-analyses of smaller trials
- LSTs are basically large pragmatic trials
  - Eligibility based on "uncertainty principle"
  - Minimal restrictions on care other than assigned treatment
  - Rigorous attention to control of bias by focusing on control of systematic errors rather than random errors
  - Reliable assessment of effect of treatment in "real world" setting

*Yusuf et al, *Stat Med*, 1984

THE ISIS TRIALS

- ISIS: International Studies of Infarct Survival
- ISIS-1*
  - RCT of IV atenolol vs placebo following MI
  - 16,000 subjects
  - 15% mortality reduction at day 7
- ISIS-4**
  - 2x2x2 factorial design, testing oral captopril, oral mononitrate and IV magnesium sulfate in immediate post-MI period
  - 58,050 subjects from 1086 hospitals
  - Captopril, but not others, was found to decrease mortality

*ISIS-1 Collaborative Group, *Lancet*, 1986

**ISIS-4 Collaborative Group, *Lancet*, 1995

ISIS TRIALS: SIMPLICITY

- Limited data collection
  - Baseline data collected by phone at randomization
  - Single page form collected at hospital discharge
- No eligibility criteria other than
  - Lack of contraindications
  - Uncertainty of physician and patient whether the treatment has positive benefit/risk ratio for that patient ("uncertainty principle")

INCREASING INTEREST IN PRAGMATIC TRIALS

- Comparative effectiveness studies
  - More information about effects of commonly used treatment approaches
- Quality improvement studies
  - Using randomized trials instead of arbitrary judgment to make decisions about optimal management of care

OVERSIGHT OF PRAGMATIC TRIALS

- Types of oversight to consider
  - IRB/institutional human research protection programs
  - Ongoing monitoring of data quality
  - Data monitoring committee
- Increasing discussion about what types of oversight mechanisms are needed for certain types of pragmatic trials

IRB OVERSIGHT

- Much ongoing debate
  - If a hospital doesn't need to get IRB review when deciding to switch brands of antibacterial soap, why should they need IRB review if they want to do a study comparing 2 brands of antibacterial soap, or 2 types of dispensers?
  - What if instead of soap the issue is whether to adopt a new diagnostic assay as a routine procedure?
  - What if instead of new assay the issue is the duration of a dialysis interval?
- Does the endpoint of the study make a difference, if both regimens are considered standard care?
- Does it make a difference if it's an individually randomized or cluster randomized trial?

MONITORING CONDUCT OF STUDY

- Quality of study conduct always important
- Most important monitoring can be done centrally, not requiring on-site checking of data accuracy
  - Timeliness of data
  - Range/consistency checks
  - Dropout rate
  - Ineligibility rate
- Errors in data entry should be minimal and random—will not create systematic bias—and many can be identified via central review
- Indications of quality problems from central review can lead to site visits as necessary

MONITORING DATA ON SAFETY AND EFFICACY

- Data Monitoring Committees (DMCs)/Data and Safety Monitoring Boards (DSMBs) are needed to monitor accumulating data from certain types of RCTs
  - Trials with serious outcomes
  - Trials with anticipated potential safety issues
- Pragmatic trials, as other RCTs, may or may not require a DMC

DATA MONITORING COMMITTEES

- Group of experts without other involvement in the trial, and with no relevant conflicts of interest regarding trial outcomes
- Will review emerging data on a regular basis and make recommendations to sponsor and study leadership re need for modifications
  - To monitor for any emerging safety issues
  - To recommend action if safety concerns are identified
  - To protect the integrity of trial results
  - To recommend whether the trial should continue as designed, be modified, or terminate

DMCs FOR PRAGMATIC TRIALS

- Principles and practices for DMCs for pragmatic trials mostly same as for any trial requiring a DMC
  - Pre-specification of statistical criteria for early termination
  - Regular review of data on safety and quality of study conduct
  - Include relevant disciplines
  - Avoid major conflicts of interest
  - Maintain confidentiality of interim results

IMPORTANT ISSUES FOR DMCs FOR PRAGMATIC TRIALS

- Comparative effectiveness
- Early stopping guidelines

COMPARATIVE EFFECTIVENESS

- Most pragmatic trials will be comparing one treatment approach to another
- Must be cautious in interpreting a finding of "no difference"
  - Could mean truly no difference
  - Could mean outcomes were too variable to permit detection of effect
  - Could mean trial quality was poor, diluting observable effect
- This is an issue in any active control trial, but may be enhanced in pragmatic trials where there will likely be substantial heterogeneity

CRITERIA FOR EARLY STOPPING

- Comparison of "standard" treatment approaches in pragmatic trials may require very conservative criteria for early termination
  - Practitioners may be very attached to their long standard practices and will be reluctant to adopt a new approach
  - How many zeros before the first nonzero digit in the p-value will be needed to foster change?
  - "Futility" probably not relevant for these trials for same reason—each approach will have its adherents who truly believe their approach is better

CLUSTER RANDOMIZED TRIALS (CRTs)

- Design being used more widely for pragmatic questions
- Randomization of units instead of individual subjects/patients/participants
- Examples
  - Randomization of emergency vehicles to different approaches of patient support during transport
  - Randomization of hospitals/clinics/hospital units to different approaches to infection control
  - Randomization of communities to different approaches to prevention of HIV infection
- Key design consideration: statistical power depends more on number of clusters than on number of individuals per cluster

CRTs PRESENT SPECIAL CHALLENGES FOR OVERSIGHT

- Relatively few researchers have been directly involved with CRTs; hard to assess what you don't understand
- In a CRT, where the unit of randomization is the cluster, there are typically far fewer units randomized than in a conventional RCT
  - Greater chance of important imbalance
- Because CRTs are less efficient than traditional RCTs, many more trial participants are required
  - More participants exposed to potentially inferior intervention
- Difficult to keep investigators blinded to interim results when everyone at site receives same treatment
  - May raise concern about interim changes in study conduct

BIG ISSUE: INTRA-SITE CORRELATION OF OUTCOMES

- In a CRT we have to account for the extent to which outcomes at a given site will be similar
- The higher this correlation, the more sites are needed—adding participants to sites doesn't help much if correlation is high
  - Extreme case: if everyone at a given site is expected to have exactly the same outcome, you wouldn't need more than 1 participant per site
- Problem: often difficult to estimate this correlation in planning study
- DMC must monitor this factor as data emerge to assess whether study has adequate power