NIH Collaboratory

Health Care Systems Research Collaboratory

Biostatistics and Study Design Core

Lessons Learned

Elizabeth DeLong, PhD
Duke University

Who We Are

- **Three members from Central Coordinating Center**
  - Elizabeth DeLong, Duke University
  - Andrea Cook, Group Health and University of Washington
  - Rui Wang, Harvard T. H. Chan School of Public Health
  - Overall direction and consistency of approach
  - Communication among projects, best practices, and dissemination
- **David Murray (NIH)** – advisor on all biostatistics issues
- **Biostatisticians from each Demonstration Project**
  - Design and analysis of Demonstration Projects
  - Communication and adoption of common practices across projects

Objectives

- Work with Demonstration Projects to address gaps and limitations in statistical plans and study designs during UG3 planning phase
- Gather information on key methodological issues and make it accessible to the research community
- Identify areas in need of methods development and address challenges
- Generate new knowledge by studying applications of statistical techniques in pragmatic trial designs

Objectives

- **Work with Demonstration Projects to address gaps and limitations in statistical plans and study designs during UG3 planning phase**
- Gather information on key methodological issues and make it accessible to the research community
- Identify areas in need of methods development and address challenges
- Generate new knowledge by studying applications of statistical techniques in pragmatic trial designs

Addressing Gaps and Limitations

- **Example 1**: Trade-off between risk of contamination (ICC) and sample size
  - Effective sample size is influenced by ICC and number of clusters
  - ICD-Pieces changed from randomizing providers to randomizing clinics because of overlapping staff and clinic procedures
  - PPACT did the opposite after preliminary assessment of potential contamination of outcomes

Addressing Gaps and Limitations

- **Example 2**: Accounting for variable cluster sizes in sample size calculations
  - ICD-Pieces had to recruit more sites to accommodate variability
- **Example 3**: Incorporating changes over time; feasibility of rolling out intervention and ability to recruit sites
  - TSOS changed from parallel to 6-period stepped wedge design
  - All sites receive intervention, but preparation to go live is staggered

Addressing Gaps and Limitations

- **Example 4**: Discovering issues after data are collected
  - STOP CRC planned clinic-level analysis, but ICC was higher than expected and number of patients in age–sex–race/ethnicity subgroups varied across clinics
  - Race/ethnicity was missing for many patients, so could not be used in analyses

BPMedTime Study: Value of UH2 Planning Period

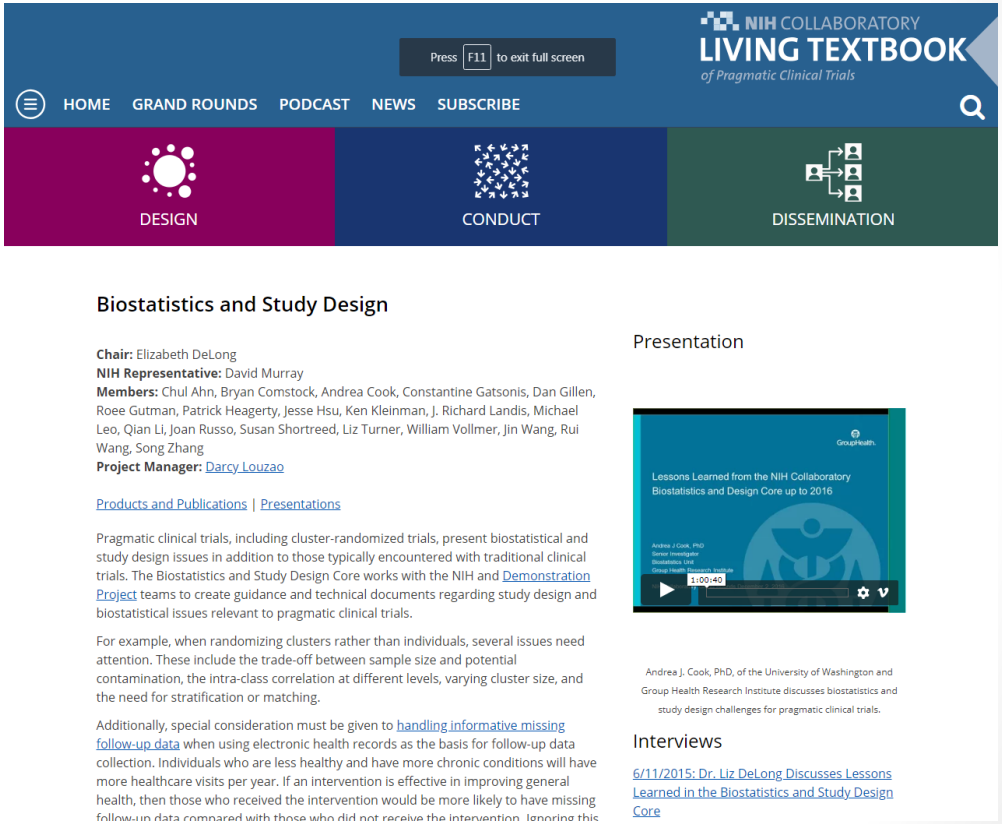
- Randomized pragmatic trial evaluating risk of adverse cardiovascular events in patients taking antihypertensive medication at bedtime vs morning or afternoon
- Sample size required to detect low effect rate grew from 1000 to 5000 patients
  - Difficult to budget within Collaboratory
  - Alternative design and analysis plans not deemed acceptable
  - Concern that potency of intervention not significant enough to reintroduce behavior change
- Potentially better suited as larger trial for network like PCORnet
- PI received positive feedback for Coordinating Center, Core/Working Groups, and Collaboratory concept

Objectives

- Work with Demonstration Projects to address gaps and limitations in statistical plans and study designs during UG3 planning phase
- **Gather information on key methodological issues and make it accessible to the research community**
- Identify areas in need of methods development and address methodological challenges
- Generate new knowledge by studying applications of statistical techniques in pragmatic trial designs

Contributions to Website

- Produced guidance on statistical considerations for pragmatic trials (available in Knowledge Repository and Living Textbook)
- Provided content for the Health Care Systems Interactions Core's Introductory Toolkit
- Learned there is a need for easily accessible information on pragmatic trials



Objectives

- Work with Demonstration Projects to address gaps and limitations in statistical plans and study designs during UG3 planning phase
- Gather information on key methodological issues and make it accessible to the research community
- **Identify areas in need of methods development and address methodological challenges**
- **Generate new knowledge by studying applications of statistical techniques in pragmatic trial designs**

Constrained Randomization

- Crude randomization risks imbalances with smaller numbers of clusters
- How to balance between-cluster differences?
  - Paired: How to choose pairs to control for important predictors?
  - Stratification: Stratify analysis on a small set of predictors, and ignore in analysis stage after stratification?
  - Constrained randomization
    - Achieve balance among known potential confounders by “constraining” possible randomization schemes to a set for which each scheme is suitably balanced, then randomly selecting one scheme
    - Effective method of controlling confounding?
    - What analyses work best in terms of type 1 error and power?

Constrained Randomization

- Balances a large number of characteristics
- Concept
  - Generate large number of cluster randomization assignments
  - Remove duplicates
  - Assess balance according to a prespecified metric
  - Restrict to assignments with sufficient balance
  - Randomly choose randomization scheme from the restricted pool
  - Assign clusters according to selected scheme

Constrained vs Simple Randomization

Constrained randomization

- Unadjusted *F* test too conservative
- Permutation test maintains type 1 error rate
- Permutation test must be referenced to appropriate distribution
- Adjusted *F* test yields highest power, but adjusted permutation test is close
- Adjusted *F* and permutation tests more powerful than unadjusted counterparts

Simple randomization

- *F* and permutation tests maintain type 1 error rate
- Little difference in performance between *F* and permutation tests
- Power of adjusted *F* test competitive with adjusted tests under constrained randomization

Constrained Randomization: Lessons Learned

- Adjusted *F* and permutation tests perform similarly and are slightly better with constrained randomization in terms of power
- Power under constrained randomization improves with decreasing candidate set size, as long as set is not too small
- Unadjusted permutation tests can be improved with additional analysis-based adjustment, even with constrained randomization
- If investigators want to control for more group-level characteristics than available groups will support for model-based analysis, permutation tests are more practical than mixed-model methods
- Constrained randomization by itself can offer design-based control of group-level potential confounders with unadjusted permutation tests

An Additional Lesson

- The UH2(UG3)/UH3 process worked well
  - Pilot studies couldn't have been carried out without initial funding
  - UH2 pilot phase provided evidence that the study could be implemented
  - Simultaneous Work Group discussions provided additional input and guidance
  - Avoided funding studies that were unlikely to recruit necessary sites/patients or to implement the intervention

Balancing Act With Trade-Offs

Getting Everyone on the Same Page

Tolerance and Patience: Lots of Bumps in the Road

Involves Cooperation and Teamwork

Thank You