Improving Pragmatic Clinical Trials: Lessons Learned from the NIH Collaboratory Biostatistics Core

Andrea J Cook, PhD

Associate Investigator
Biostatistics Unit, Group Health Research Institute
Affiliate Associate Professor
Dept. of Biostatistics, University of Washington

June 18, 2015
NIH Collaboratory

Acknowledgements

NIH Collaboratory Coordinating Center Biostatisticians

Elizabeth Delong, PhD, Andrea Cook, PhD, and Lingling Li, PhD

NIH Collaboratory Project Biostatisticians

Patrick Heagerty, PhD, Bryan Comstock, MS, Susan Shortreed, PhD, Ken Kleinman, PhD, and William Vollmer, PhD

NIH Methodologist

David Murray, PhD

Funding

This work was supported by the NIH Health Care Systems Research Collaboratory (U54 AT007748) from the NIH Common Fund.

Outline

- **NIH Collaboratory Pragmatic Trial Setting**
- **UH2 Phase: What did we do?**
  - Common themes across studies
  - How were the trials improved?
- **What are we doing now?**
  - UH3 Phase Issues
  - New UH2 Trials
  - Unanswered Questions?

Pragmatic vs. Explanatory Trials

GroupHealth

CMAJ

ANALYSIS

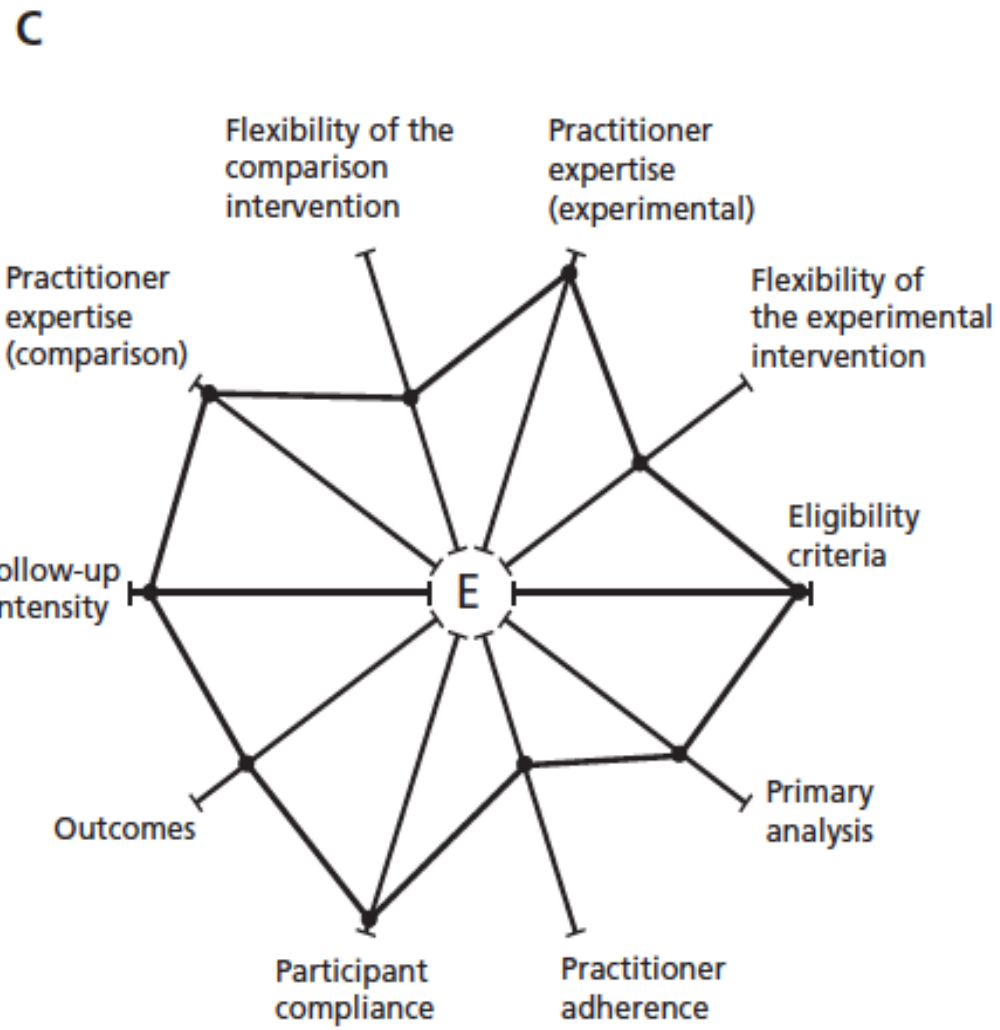
A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers

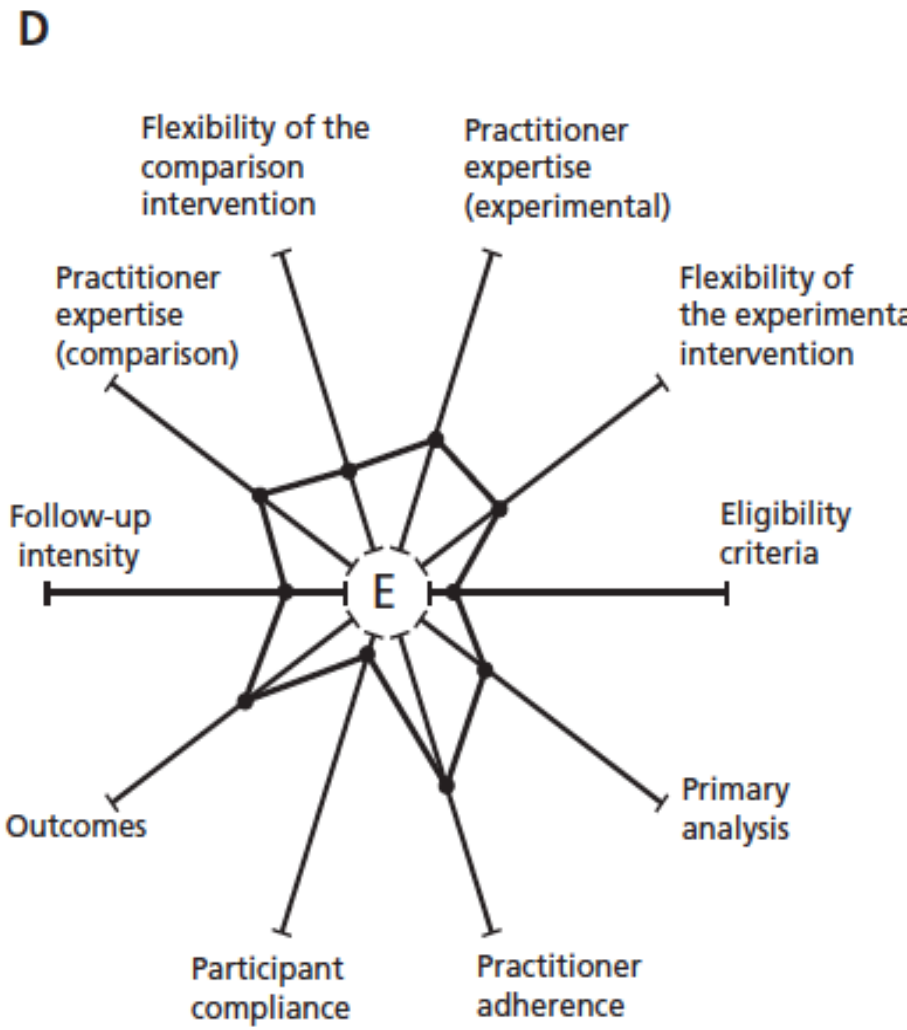
Kevin E. Thorpe MMath, Merrick Zwarenstein MD MSc, Andrew D. Oxman MD, Shaun Treweek BSc PhD, Curt D. Furberg MD PhD, Douglas G. Altman DSc, Sean Tunis MD MSc, Eduardo Bergel PhD, Ian Harvey MB PhD, David J. Magid MD MPH, Kalipso Chalkidou MD PhD

Published at www.cmaj.ca on Apr. 16, 2009. An abridged version of this article appeared in the May 12 issue of *CMAJ*. This article was published simultaneously in the May 2009 issue of the *Journal of Clinical Epidemiology* (www.jclinepi.com).

Pragmatic vs. Explanatory Trials

GroupHealth





Key features of most PCTs

Use of electronic health records (EHRs)

- EHRs allow efficient and cost-effective, recruitment, participant communication & monitoring, data collection, and follow up

Randomization at clinic or provider level

- Protocols can be tailored to local sites and can adapt to changes in a dynamic health care environment

Pragmatic Trials Concept

- **Size: Large simple trials→ precise estimates, evaluate heterogeneity**
- **Endpoints: patient oriented usually with minimal adjudication**
- **Setting: integrated into real world**
  - Non-academic centers
  - Leverage electronic data
  - Patients as partners

Outline

- NIH Collaboratory Pragmatic Trial Setting
- **UH2 Phase: What did we do?**
  - Common themes across studies
  - How were the trials improved?
- What are we doing now?
  - UH3 Phase Issues
  - New UH2 Trials
  - Unanswered Questions?

Round 1 Demonstration Projects

GroupHealth

Principal Investigator	Institution	Project
Gloria Coronado	Kaiser Foundation Research Institute	Strategies and Opportunities to Stop Colon Cancer in Priority Populations
Lynn DeBar	Kaiser Foundation Research Institute	Collaborative Care for Chronic Pain in Primary Care
Laura Dember	University of Pennsylvania	Pragmatic Trials in Maintenance Hemodialysis
Susan Huang	University of California--Irvine	Decreasing Bioburden to Reduce Healthcare-Associated Infections and Readmissions
Jeffrey Jarvik	University of Washington	A Pragmatic Trial of Lumbar Image Reporting with Epidemiology (LIRE)
Gary Rosenthal	University of Iowa	Nighttime Dosing of Anti-Hypertensive Medications: A Pragmatic Clinical Trial
Gregory Simon	Group Health Cooperative	Pragmatic trial of population-based programs to prevent suicide attempt

STUDY DESIGN

Study Design: Cluster RCT

GroupHealth

- **Mostly Cluster RCTs (except one)**
  - Randomization Unit:
    - Provider < Panel < Clinic < Region < Site
- **Average Size of Cluster**
  - Initial Proposals: Most large clinic level clusters
  - Goal: Smallest Unit without contamination
    - More clusters are better if possible
  - Smaller number of clusters increase sample size along with estimation issues (GEE)
  - Potential Solutions: Panel-level or physician-level

Study Design: Variable Cluster Size

GroupHealth

- **Variable Cluster Size**
  - Sample Size calculations need to take this into account
    - Design effects are different
    - Depends on the analysis choice
  - Analysis Implications: What are you making inference to?
    - Cluster vs Patient vs Something in-between
    - Marginal versus conditional estimates

DeLong, E, Cook, A, and NIH Biostatistics/Design Core (2014) Unequal Cluster Sizes in Cluster-Randomized Clinical Trials, *NIH Collaboratory Knowledge Repository*, https://www.nihcollaboratory.org/Products/Varying-cluster-sizes_V1.0.pdf

DeLong, E, Lokhnygina, Y and NIH Biostatistics/Design Core (2014) The Intraclass Correlation Coefficient (ICC), *NIH Collaboratory Knowledge Repository*, https://www.nihcollaboratory.org/Products/Intraclass-correlation-coefficient_V1.0.pdf

Study Design: Which Cluster Design?

GroupHealth

- **Cluster**
  - Randomize at cluster-level
  - Most common, but not necessarily the most powerful or feasible
  - Advantages:
    - Simple design
    - Easy to implement
  - Disadvantages:
    - Need a large number of clusters
    - Not all clusters get the interventions
    - Interpretation for binary and survival outcomes:
      - Mixed models within cluster interpretation problematic
      - GEE marginal estimates interpretation, but what if you are interested in within cluster changes?

Study Design: Which Cluster Design?

GroupHealth

- **Cluster with Cross-over**
  - Randomize at cluster but cross to other intervention assignment midway
  - Feasible if intervention can be turned off and on without “learning” happening
  - Alternative: baseline period without intervention and then have half of the clusters turn on

Study Design: Which Cluster Design?

GroupHealth

	Cluster	Period 1	Period 2
Simple Cluster	1	INT	
	2	UC	
	3	UC	
	4	INT	
Cluster With Crossover	1	INT	UC
	2	UC	INT
	3	UC	INT
	4	INT	UC
Cluster With Baseline	1	UC	INT
	2	UC	UC
	3	UC	UC
	4	UC	INT

Study Design: Which Cluster Design?

GroupHealth

- **Cluster with Cross-over**
  - Advantages:
    - Can make within cluster interpretation
    - Potential to gain power by using within cluster information
  - Disadvantages:
    - Contamination can yield biased estimates especially for the standard cross-over design
    - May not be feasible to switch assignments or turn off intervention
    - Not all clusters have the intervention at the end of the study

Study Design: Which Cluster Design?

GroupHealth

- **Stepped Wedge Design**
  - Randomize timing of when the cluster is turned **on** to intervention
  - Staggered cluster with crossover design
  - Temporarily spaces the intervention and therefore can control for system changes over time

Study Design: Which Cluster Design?

GroupHealth

	Cluster	Baseline	Period 1	Period 2	Period 3	Period 4
Stepped Wedge	3	UC	INT	INT	INT	INT
	2	UC	UC	INT	INT	INT
	1	UC	UC	UC	INT	INT
	4	UC	UC	UC	UC	INT

Study Design: Which Cluster Design?

GroupHealth

- **Stepped Wedge Design**
  - Advantages:
    - All clusters get the intervention
    - Controls for external temporal trends
    - Make within cluster interpretation if desired
  - Disadvantages:
    - Contamination can yield biased estimates
    - Heterogeneity of Intervention effects across clusters can be difficult to handle analytically
    - Special care of how you handle random effects in the model
    - Relatively new and available power calculation software is relatively limited

GroupHealth

RANDOMIZATION

Randomization

- **Crude randomization not preferable with smaller number of clusters or need balance for subgroup analyses**
- **How to balance between cluster differences?**
  - Paired
    - How to choose the pairs best to control for important predictors?
    - Implications for analyses and interpretation
  - Stratification
    - Stratify analysis on a small set of predictors
    - Can ignore in analyses stage if desired
  - Other Alternatives

DeLong, E, Li, L, Cook, A, and NIH Biostatistics/Design Core (2014) Pair-Matching vs stratification in Cluster-Randomized Trials, *NIH Collaboratory Knowledge Repository*, https://www.nihcollaboratory.org/Products/Pairing-vs-stratification_V1.0.pdf

Randomization: Constrained Randomization

GroupHealth

- **Balances a large number of characteristics**
- **Concept**
  1. Simulate a large number of cluster randomization assignments (A or B but not actual treatment)
  2. Remove duplicates
  3. Across these simulated randomizations assignments assess characteristic balance
  4. Restrict to those assignments with balance
  5. Randomly choose from the restricted pool a randomization scheme.
  6. Randomly assign treatments to A or B

Randomization: Constrained Randomization

GroupHealth

- **Is Constrained randomization better then unconstrained randomization**
- **How many valid randomization schemes do you need to be able to conduct valid inference?**
- **Do you need to take into account randomization scheme in analysis?**
  - Ignore Randomization
  - Adjust for variables in regression
  - Permutation inference

=> Conduct a simulation study to assess these properties

Randomization: Constrained Randomization Simulation Design

GroupHealth

- **Outcome Type:** Normal
- **Randomization Type:** Simple versus Constrained
- **Inference Type:** Exact (Permutation) versus Model-Based (F-Test)
- **Adjustment Type:** Unadjusted versus Adjusted
- **Clusters:** Balanced designs, but varied size and number
- **Correlation:** Varied ICC from 0.01 to 0.05
- **Potential Confounders:** Varied from 1 to 10

Li, F., Lokhnygina, Y., Murray, D, Heagerty, P., Vollmer, W., Kleinman, K., and Delong, E. (2015) A comparison of the model-based F-test and the permutation test under simple versus constrained randomization for the analysis of data from group-randomized trials (In Submission).

Randomization: Constrained Randomization Simulation Results

GroupHealth

- **Adjusted F-test and the permutation test perform similar and slightly better for constrained versus simple randomization.**
- **Under Constrained Randomization:**
  - Unadjusted F-test is conservative
  - Unadjusted Permutation holds type I error (unless candidate set size is not too small)
  - Unadjusted Permutation more powerful then Unadjusted F-Test
- **Recommendation: Constrained randomization with enough potential schemes (>100), but still adjust for potential confounders**

Randomization: Constrained Randomization Next Steps

GroupHealth

- **What about Binary and Survival Outcomes??**
- **Hypothesized Results (Mine not NIH Collaboratories):**
  - Constrained Randomization probably still wins
  - Binary Outcomes: Likely less of a preference for adjusted versus unadjusted analyses (mean and variance relationship (minimal precision gains))
  - Survival Outcomes: Depends on scenario and model choice (frailty versus robust errors)

GroupHealth

OUTCOME ASCERTAINMENT

Outcome Ascertainment

- **Most trials use Electronic Healthcare Records (EHR) to obtain Outcomes**
  - Data **NOT** collected for research purposes
- **If someone stays enrolled in healthcare system - assume that if you don't observe the outcome it didn't happen**
  - In closed system this is likely ok
  - Depends upon cost of treatment (likely to get a bill the more the treatment costs)

Outcome Ascertainment (Cont)

- **Do you need to validate the outcomes you do observe?**
  - Depends on the Outcome (PPV, sensitivity)
  - Depends on the cost (two-stage design?)
- **How do you handle Missing Outcome Data?**
  - Leave healthcare system
    - Type of Missing Data: Administrative missingness (MCAR), MAR or non-ignorable?
    - Amount of Missing Data: how stable is your population being studied?
  - Depends on the condition and population being studied.

DeLong, E, Li, L, Cook, A, and NIH Biostatistics/Design Core (2014) Key Issues in Extracting Usable Data from Electronic Health Records for Pragmatic Clinical Trials, *NIH Collaboratory Knowledge Repository*, https://www.nihcollaboratory.org/Products/Extracting-EHR-data_V1.0.pdf

Outline

- **NIH Collaboratory Pragmatic Trial Setting**
- **UH2 Phase: What did we do?**
  - Common themes across studies
  - How were the trials improved?
- **What are we doing now?**
  - Current UH3 Phase Issues
  - New UH2 Trials
  - Unanswered Questions?

UH3 Phase

GroupHealth

- **Submitted new UH3 proposals last summer**
  - New design choices submitted
  - Improved sample size calcs using pilot data collected in UH2 phase and modifications
  - Improved and finalized analysis plans with feedback from all Collaboratory participants
- **Those funded moved to UH3 phase this Fall or Spring**
- **Very early in the UH3 phase**
  - Most studies are already randomizing participants
  - Some new issues have come up…

UH3 Phase: DSMB

GroupHealth

- **Are pragmatic clinical trials different?**
  - Depends on the study
  - Main difference: how we collect, and timeliness of the collection, of adverse events and outcomes
  - Formal Primary Outcome Monitoring
    - How do you handle the fact that you likely don't have the validated outcome available in a timely manner?
    - IRB has restricted the population that the DSMB can monitor to those that receive the intervention in the intervention arm only (e.g. internet intervention if they passively refuse by not going to the website we can't get their outcome data until the end of the study)

Data Safety Monitoring

UH3 Phase: DSMB

GroupHealth

- **Are pragmatic clinical trials different?**
  - Depends on the study
  - Main difference: how we collect, and timeliness of the collection, of adverse events and outcomes
  - Formal Primary Outcome Monitoring
    - How do you handle the fact that you likely don't have the validated outcome available in a timely manner?
    - IRB has restricted the population that the DSMB can monitor to those that receive the intervention in the intervention arm only (e.g. internet intervention if they passively refuse by not going to the website we can't get their outcome data until the end of the study)

New UH2's

GroupHealth

Principal Investigator	Institution	Project
Mor, Vincent; Volandes, Angelo; Mitchell, Susan	Brown University School of Medicine	Pragmatic Trial of Video Education in Nursing Homes
Vazquez, Miguel	UT Southwestern Medical Center	Improving Chronic Disease Management with Pieces (ICD-Pieces)
Zatzick, Douglas	University of Washington	A Policy-Relevant U.S. Trauma Care System Pragmatic Trial for PTSD and Comorbidity (Trauma Survivors Outcomes and Support [TSOS])

Conclusions

- **Pragmatic Trials are important to be able to move research quickly into practice**
- **Pragmatic Trials add Complication**
  - First Question: Can this study be answered using a pragmatic trial approach??
  - Study Design is essential and needs to be flexible
  - Using EHR data is valuable, but understanding the performance of all measures is important
  - Appropriate analysis taking into account design, randomization, and outcome ascertainment is key
- **Lot's of open statistical questions still to be addressed**