CLUSTER RANDOMIZED TRIALS

L. Miriam Dickinson, PhD

ACCORDS
ADULT AND CHILD CONSORTIUM FOR HEALTH OUTCOMES RESEARCH AND DELIVERY SCIENCE
UNIVERSITY OF COLORADO | CHILDREN'S HOSPITAL COLORADO

What is a Cluster Randomized Trial?

- Key features of a CRT that distinguish it from a traditional randomized trial
  - Unit of randomization is a cluster, not the individual
    - A clusters can be a medical practice, community, county, hospital, school, etc.
  - Individuals are nested, or clustered within the larger unit of randomization
  - All individuals enrolled in the study from a particular cluster will be in the same study arm

Why Choose a Cluster Randomized Trial Design?

- Target of the intervention
  - Does the intervention focus primarily on the patient? Or does the intervention target a larger unit such as a clinic or community or environment?
- Is contamination a potential problem?
  - If there are patients from both study arms in the same setting can they exchange information or somehow influence each other (or the clinician delivering the care)?
- Other considerations
  - Setting: Where will the study take place: clinic, hospital, geographic unit (e.g. county, community)? Are there potential contextual effects of interest?
  - Is feasibility an issue? Is it possible/feasible to deliver all interventions in all settings (necessary for a patient randomized trial)
  - Cost: CRTs are sometimes more expensive but there could be tradeoffs

How Do I Conduct a CRT? Common Issues to Consider

- Clustering of patients within larger unit (e.g. patients within clinics)
  - Individuals within clusters are more similar to each other than members of other clusters
    - Violation of independence assumption
  - Power and sample size, statistical analysis are *all* affected by clustering
    - Reduced power for the same number of individuals
    - Possibly greater cost
    - More complex analyses
- Recruiting clusters from a larger pool can be challenging
  - Self-selection
- Blinding is often not possible
- Heterogeneity among clusters
- Generally, the number of units/clusters to be randomized is much smaller than trials in which individuals are randomized
  - Potential for covariate imbalance between study arms
  - Simple, or even stratified randomization of groups can result in study arms that are very different from each other

Power and Sample Size for CRTs: A Simple Approach

- Intraclass correlation coefficient: a measure of how similar patients within the same cluster are relative to patients in other clusters
- Steps in a power analysis:
  - Determine your primary outcome variables
  - Obtain an estimate of the ICC, either from the literature or based on actual data you may have
  - Calculate the variance inflation factor (VIF): (1 + (m − 1)ICC), where m is the number of patients per practice
    - Calculate the effective sample size: divide the proposed sample size (m x number of practices) by the VIF
- Do a traditional power analysis using effective sample size

Practices per arm	Patients per practice	ICC	VIF	Effective sample size	Effect size	power
6	50	5%	3.45	87	.43	>80%
6	50	10%	5.9	51	.56	80%
6	50	15%	8.35	36	.67	80%
6	100	10%	10.9	55	.55	>80%
10	50	10%	5.9	85	.44	>80%

Example 1: Connection to Health

- This is a very common use of a CRT in primary care practice settings
- Purpose: to test effectiveness of interactive behavior change technology (IBCT) with practice facilitation (PF) on improving self-management support (SMS) for patients with type 2 diabetes in primary care
  - Control arm: Education for clinicians and staff on patient self-management support (SMS)
  - Intervention arm: Education *plus* IBCT tool *with* practice facilitation to assist practices in implementing SMS using the IBCT tool in their practice
- Outcomes at the patient level evaluated in a random sample of patients from each practice by medical record review: 1) evidence of SMS and, 2) HbA1c over time
- Factors that influenced the choice of a CRT design
  - Intervention is focused on the practice as a whole (education, technology + facilitation) rather than directly on the patient
  - Contamination would be an issue if patients within the same practice were randomized to different approaches because care would be delivered by the same clinical team

Example 2: Population-based vs Practice-based Reminder Recall

- Purpose: Compare two approaches to increasing up-to-date immunization rates in 19-35 month old children in Colorado
  - *Population-based R/R*
    - Intervention delivered at the level of the population, in this case, the county
  - *Practice-based R/R*
    - Intervention targeted eligible practices (training for R/R) and delivered to patients by practices
- Setting: counties in Colorado, stratified by rural/urban location

Planning: Study Design Challenges

- Early decisions involved unit of randomization
  - Individual level randomization not feasible and didn't fit the goals of the study
- County would be the cluster and unit of randomization
- Also interested in context: rural vs urban
- Baseline data could be obtained from CIIS database by county of residence
- All children in age range with at least 2 immunization records in CIIS, residing in selected counties, would be included in the trial if they needed 1 or more vaccines

Study Design Challenges

- Implications of using a county-based population
  - PB arm
    - All eligible practices in PB intervention counties would be invited to participate in training, thus eliminating potential selection bias
    - But practice participation was not a requirement
    - Individual affiliation with a practice was not a requirement for data to be included
  - Population-based arm
    - All eligible children, regardless of practice affiliation (or not) would be included in the trial
  - Analysis: population-based sample

Cluster Selection

- Pre-specified criteria for selecting counties
  - Minimum 70% in CIIS
  - Urban or rural (frontier counties with <10,000 excluded)
  - No ongoing existing county-wide reminder/recall efforts
  - Other county-specific exclusions (e.g. high refusal rates, smaller population relative to other urban)
- Setting:16 counties in Colorado, stratified by rural/urban location
  - Rural: Alamosa, Eagle, Fremont, Garfield, Grand, Logan, Otero, Rio Grand
  - Urban: Adams, Arapahoe, Douglas, El Paso, Jefferson, Larimer, Pueblo Weld

Study Design Challenges: Covariate Imbalance

- Relatively few units for randomization and heterogeneity among clusters
- Imbalance in clinical trials is not a new problem
- Stratification is not always sufficient to overcome this problem
  - Motivating factor to explore alternatives to simple (or stratified) randomization came from experience with a previous cluster randomized trial (type 2 diabetes) and imbalanced study arms
- Minimization methods for randomization of individuals were first described in the 1960's and 1970's
- Extended to CRTs in early 2000s

Methods for Randomization

- Raab and Butcher (2001) consider the effects of covariate imbalance on an optimal design criterion: difference between crude and adjusted treatment effect
  - Showed that differences between crude and adjusted treatment effect are minimized when differences in treatment group means on covariates to be included in the analysis are small
- Covariate constrained randomization methods described
  - Moulton LH. Covariate-based constrained randomization of group-randomized trials. Clinical Trials 2004
  - Glynn RJ, Brookhart A, Stedman M, Avorn J, Solomon DH. Design of cluster-randomized trials of quality improvement interventions aimed at medical care providers. Medical Care. 2007
- But relatively few CRTs had used these approaches at the time we planned this trial

Procedure for Covariate Constrained Randomization

- Baseline data on units of randomization must be available
- All possible randomizations of units into study groups are generated (for 2 arm trial)
- A balance criterion (B), defined as the sum of squared differences between study groups on relevant standardized variables, is calculated for each randomization
  - $B=(w_1(x_{11} - x_{21})^2 + w_2(x_{12} - x_{22})^2 + \ \dots)$
  - Where w is the weight for each selected variable, x_{11} is the mean for study arm 1, variable 1, x_{21} is the mean for arm 2, variable 1, etc.
- Establish a criterion for maximum allowable difference between study arms and define a set of “optimal randomizations” in which the differences between treatment groups on covariates are minimized
- A single randomization is then chosen from the set of “optimal randomizations”

Covariate Constrained Randomization for R/R trial

- All possible randomizations generated using SAS Proc IML
- Standardize randomization variables (z-scores)
- Generate a file containing data on each randomization and calculate group means on all randomization variables
- Variables weighted equally
- For each randomization
  - Balance criterion calculated (sum of total squared differences across all variables)

Covariate Constrained Randomization for R/R trial

- Stratification variable (urban/rural) can be included in the process by limiting possible randomizations to those that are balanced
- In this case, each study arm should include exactly 4 rural counties; all other combinations are eliminated
- This results in smaller set of possible randomizations that are already balanced on rural/urban location

Covariate Constrained Randomization for R/R trial

- Variables for balance criterion (county level)
  - Total number of children in age range
  - Up-to-date rates for early childhood immunizations
  - % African American in county
  - % Hispanic in county
  - Average income
  - Pediatric to family medicine ratio
  - # of community health clinics
- For each randomization balance criterion calculated (total squared difference)
  - $B = (nKIDSg1 - nKIDSg2)^2 + (UTDg1 - UTDg2)^2 + (\%blackG1 - \%blackG2)^2 + (\%HispG1 - \%HispG2)^2 + (incomeG1 - incomeG2)^2 + (pedsfmratioG1 - pedsfmratioG2)^2 + (nchcG1 - nchcG2)^2$

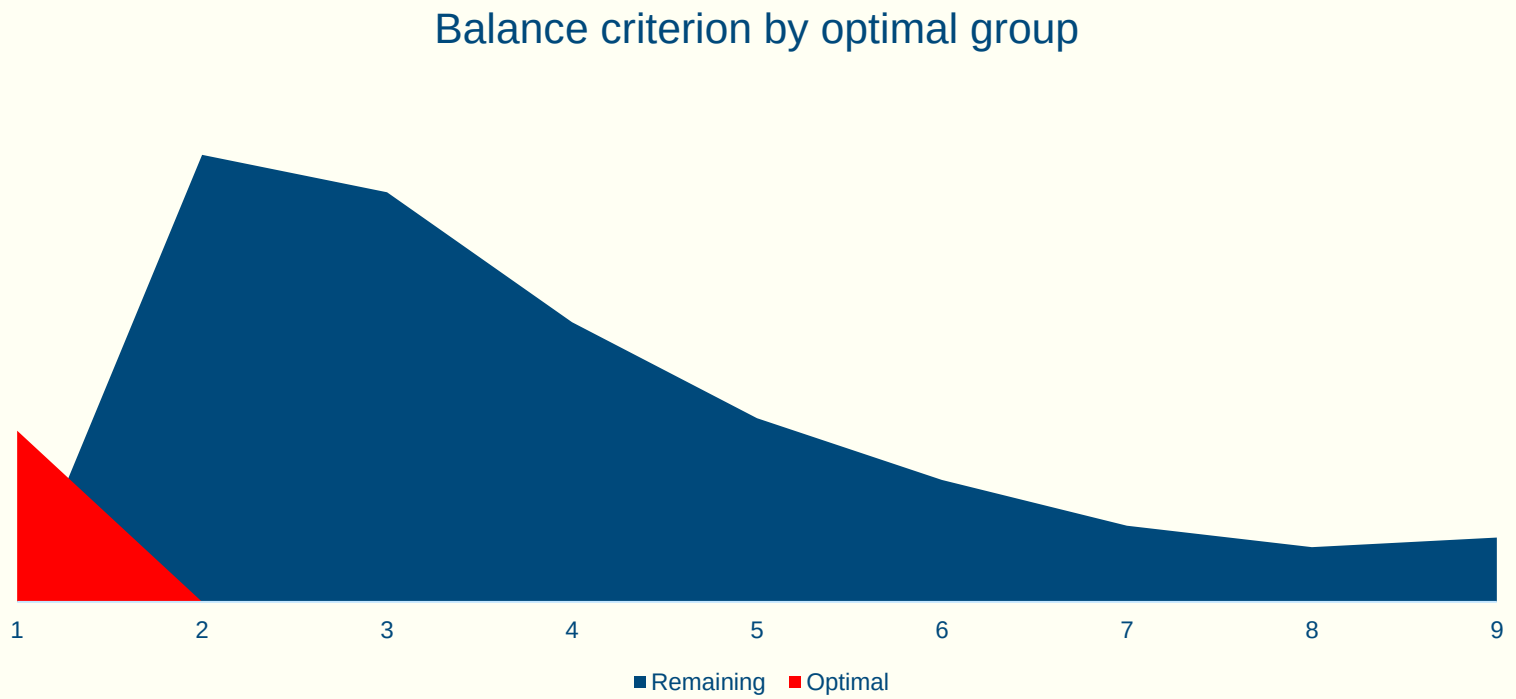
Covariate Constrained Randomization for R/R trial

- Examined the distribution of the balance criterion and set a value for defining a candidate set
  - Early work, including this study, used the best 10% to define the candidate set
  - That could be unnecessarily restrictive and a larger candidate set will work just as well
- Optional: compare differences in means on raw variables for “optimal set” vs others
- Randomly selected a final randomization from the optimal set and assigned counties to study arms

County Level Characteristics

	County-Level Variables for Randomization	
Variable	Rural and Urban Counties Mean (SD)	Min, max
Number of children age 19-35 months	4197 (4432)	234, 12354
% Up-to-date at baseline	40.8% (8.3)	27.0%, 54.0%
% Hispanic	22.3% (12.9)	6.0%, 44.0%
% African American	2.9% (2.7)	0%, 10.0%
Average Income ($)	$53481 (15793)	$29738, $93819
Pediatric to Family Medicine ratio	0.28 (0.25)	0, 1.0
# CHCs	4.4 (3.5)	0, 11

Distribution of Balance Criterion



Dickinson LM, Beaty B, Fox C, Pace W, Dickinson WP, Emsermann C, Kempe A.. Pragmatic cluster randomized trials using covariate constrained randomization: A method for practice-based research networks (PBRNs). J Am Board Fam Med. 2015 Sep-Oct;28(5)

Magnitude of Differences in Means on Raw Variables

Differences Between Study Groups on Raw Variables		
Variable	Optimal Mean (Max)	Remaining Randomizations Mean (Max)
Number of children age 19-35 months	223 (613)	1264 (6325)
% Up-to-date at baseline	2.1% (5.0)	4.9% (15.0)
% Hispanic	5.6% (11.3)	7.9% (23.3)
% African American	<1% (1.0)	1.4% (4.5)
Average Income ($)	$3659 (9702)	$9731 (27131)
Pediatric to Family Medicine ratio	0.20 (0.40)	0.15 (0.40)
# CHCs	1.3 (2.8)	1.6 (4.8)

[+]absolute value of differences taken for each randomization

Worst Randomization from Optimal Set

Variable	Arm 1 Means of County-Level Variables (SD)	Arm 2 Means of County-Level Variables (SD)
Number of children age 19-35 months	4275 (4628)	4118 (4546)
% Up-to-date at baseline	40.1% (8.8)	41.5% (8.3)
% Hispanic	23.8% (14.8)	20.9% (11.6)
% African American	2.5% (2.4)	3.3% (3.1)
Average Income $	$56264 (18004)	$50699 (13877)
Pediatric to Family Medicine ratio	0.33 (0.33)	0.23 (0.15)
# CHCs	4.8 (4.5)	4.0 (2.4)

Selected Randomization by Location

Variable	Rural		Urban	
	Arm 1 Mean (SD)	Arm 2 Mean (SD)	Arm 1 Mean(SD)	Arm 2 Mean(SD)
Number of children age 19-35 months	682 (695)	618 (465)	7467 (3915)	8049 (3855)
% Up-to-date at baseline	39.0 (7.5)	36.3 (6.5)	44.8 (9.1)	43.3 (10.1)
% Hispanic	26.5 (17.6)	22.3 (12.1)	18.3 (14.5)	22.3 (11.1)
% black	1.3 (.5)	2.3 (2.2)	4.3 (3.9)	3.8 (3.1)
Average Income $	47115 (16755)	49493 (15475)	61298 (23090)	56019 (5326)
Pediatric to Family Medicine ratio	.43 (.38)	.10 (.16)	37.8 (18.8)	21.3 (10.9)
# CHCs	2.5 (2.6)	1.8 (1.5)	5.3 (2.9)	8.0 (3.6)

Data and Analytic Challenges

- Establishing a cohort
  - Baseline cohort: data obtained from CIIS database in June 2010
  - Follow-up CIIS database obtained December 2010
  - Final analytic database involved matching baseline and follow-up records: 98.3% match

Data and Analytic Challenges

- Generalized linear mixed effects models
  - Study arm, county baseline up-to-date rates and rural/urban location included as fixed effects
- Clustering
  - Clustering within practice was important so we used site of last service used as random effect (most children assigned to a cluster this way)
  - For children with no practice affiliation or very small clusters we aggregated and created an “unaffiliated” cluster for each county
    - Convergence problems with numerous singletons and very small clusters
- Secondary analysis within practice-based arm
- We were also interested in rural vs urban differences

Conclusions and Acknowledgements

- Cluster randomized pragmatic trials present unique challenges but, in most situations, reasonable solutions to study design, data and analytic challenges can be found

- I would like to acknowledge Brenda Beaty for her collaboration on this project

Questions? Thoughts?

CONSORT statement: see extension for CRTs
http://www.consort-statement.org/extensions/overview/cluster-trials