Cluster Randomized Trials and the Stepped Wedge Design

Jim Hughes

Department of Biostatistics

University of Washington

Jan 26, 2005

Overview

- Cluster randomized trials
- Design alternatives
- Stepped Wedge Design
  - Power
    - time trends, # steps, delays
  - Analytic Approaches
  - Examples

Cluster Randomized Trials

- Randomization at group level (but unit of inference is the individual)
- Individual randomization not feasible, ethical or potential contamination
- Usually, less efficient than individually randomized trial
- Intervention effect on a community may be greater than the sum of the parts (e.g. herd immunity)

Cluster Randomized Trials

- Clusters may be large ... (cities, schools)
- ... or small (IDU networks, families)
- Often issues with blinding, self-selection, informed consent
- Key statistical challenge: multiple sources of variation

Independent Observations

$$Y_i = \mu + e_i \quad i = 1, \ldots, N$$

$$e_i \sim N(0, \sigma^2)$$

$$\bar{Y} \sim N(\mu, \frac{\sigma^2}{N})$$

Clustered Data

$$Y_{ij} = \mu + a_i + e_{ij} \quad i = 1, \ldots, I \;\; j = 1, \ldots, J$$

$$a_i \sim N(0, \tau^2), \quad e_{ij} \sim N(0, \omega^2), \quad N = IJ$$

$$Var(Y_{ij}) = \sigma^2 = \tau^2 + \omega^2$$

$$\bar{Y} \sim N(\mu, \frac{\sigma^2}{N}[1 + \rho(J-1)]) \text{ where } \rho = \frac{\tau^2}{\tau^2 + \omega^2}$$

Key Considerations

- What is the unit of randomization/inference?
- How is intervention delivered?
- How is the outcome measured?
- Examples
  - COMMIT
  - PREVEN
  - HPTN037
  - HPTN041

Designs

- Parallel Design
  - Most straightforward, common design
  - Half treatment, half control
  - Long followup possible
- Crossover Trial
  - Each group receives both treatments
  - Random order; Washout period
- Stepped Wedge Crossover Trial
  - Crossover in one direction only

Parallel Design

- One time point (no crossover)
- Half of the clusters receive intervention
- Analysis using t-test
- Power sensitive to "between-cluster" variation

		Time
		1
Cluster	1	1
	2	1
	3	0
	4	0

<back>

Crossover Design

- Two time points; half of the clusters receive intervention at each time point
- Washout period between treatements
- Analysis using paired t-test
- Power not sensitive to between-cluster variation

		Time	
		1	2
Cluster	1	1	0
	2	1	0
	3	0	1
	4	0	1

<back>

Stepped Wedge Design

- Multiple time points
- Time of crossover randomized; crossover is unidirectional
- Model-based analysis
- Reduced sensitivity to between-cluster variation

Cluster	Time 0	1	2	3	4
1	0	1	1	1	1
2	0	0	1	1	1
3	0	0	0	1	1
4	0	0	0	0	1

Stepped Wedge Design

- Pros
  - Useful when intervention cannot be introduced in all clusters at once
  - Intervention never “taken away”
  - Power less dependent on “between-cluster” effect
- Cons
  - Lengthy (multiple time intervals)
  - Ascertainment of outcome “immediate”
  - More complex analysis

Example 1 - Expedited Partner Treatment

- Expedited treatment of sex partners for gonorrhea or chlamydia (Golden et al., NEJM 2005)
  - Intervention - voucher for meds and condoms
  - "Control" - physician's referral
  - Effective at reducing re-infection in index case in individually randomized trial
- Desire to implement intervention in a cluster randomized trial over Washington state
  - 24 counties considered, 4 randomization steps (plus baseline), randomize 6 per step, 6 month intervals
  - Logistically difficult to start all clusters at same time point
  - Outcome (STD) measured in "sentinel sites"

Model

Define:

Y_{ijk} to be a 0/1 response for individual k from cluster i at time j

$$i \in 1 \ldots I \quad \text{clusters}$$
$$j \in 1 \ldots T \quad \text{time points}$$
$$k \in 1 \ldots N \quad \text{individuals per cluster}$$

p_{ij} to be the true proportion of cases in cluster i at time point j

$$Y_{ijk} \sim B(1, p_{ij})$$

Model

The following random-effects model can be used for p_{ij}:

$$p_{ij} = p + \alpha_i + \beta_j + \theta X_{ij}$$

p overall prevalence of cases

α_i random effect of cluster i

$\alpha_i \sim N(0, \tau^2)$

β_j fixed time effect for time period $[j, j+1]$

θ fixed treatment effect

X_{ij} indicator for treatment in cluster i at time point j

Power Calculations

Interest is in determining the power to test

$$H_0 : \theta = 0 \text{ vs. } H_A : \theta = \theta_A$$

One can approximate the theoretical power by

$$power = \Phi\left(\sqrt{\frac{\theta_A^2}{Var(\hat{\theta})}} - Z_{1-\alpha/2}\right)$$

A closed form expression for $Var(\hat{\theta})$ can be derived using the random effects model

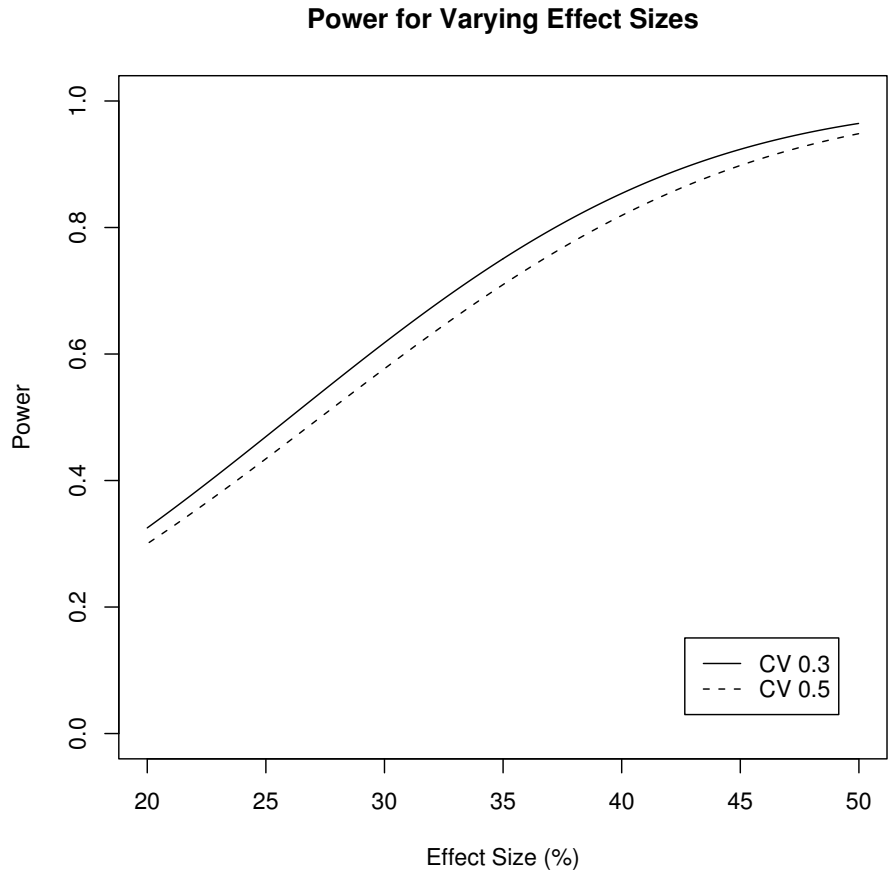
Power Calculations

- "Simple" formula for $Var(\hat{\theta})$ involves τ,σ, I, T, X_{ij}
- Variance formula applicable to the other designs (parallel, crossover)
- Assumes estimation of time effect parameters $\beta_1 \ldots \beta_{T-1}$

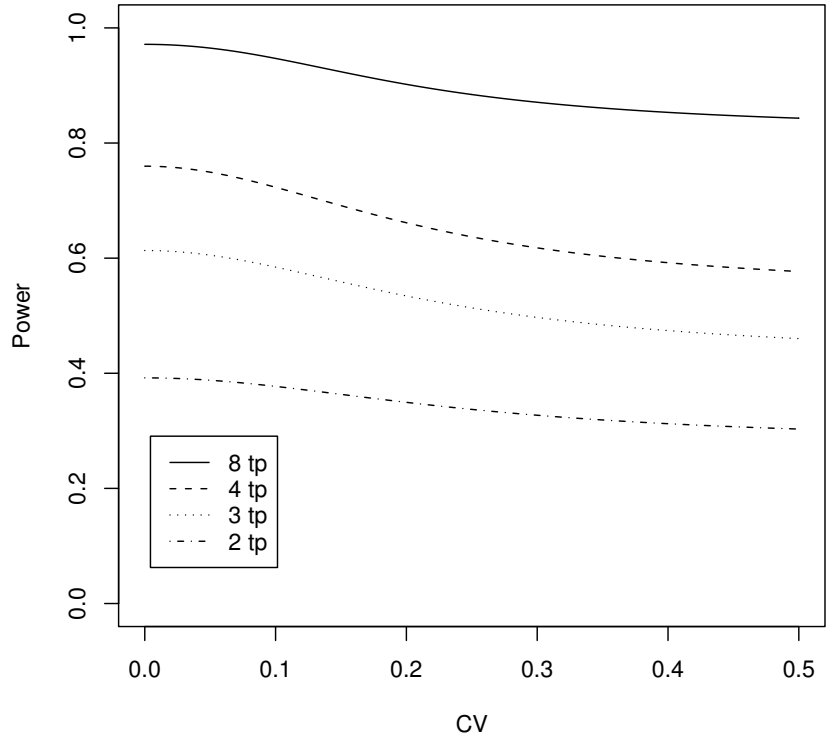
Time Effects

- $\beta_1 \ldots \beta_{T-1}$; how do these effect $\hat{\theta}$?
- Bias in $\hat{\theta}$ when time effects exist and are not estimated (e.g. paired t-test)
- If $\beta_1 \ldots \beta_{T-1}$ are small relative to the prevalence p, bias will be small
- Some power is lost if one estimates non-existent time effects

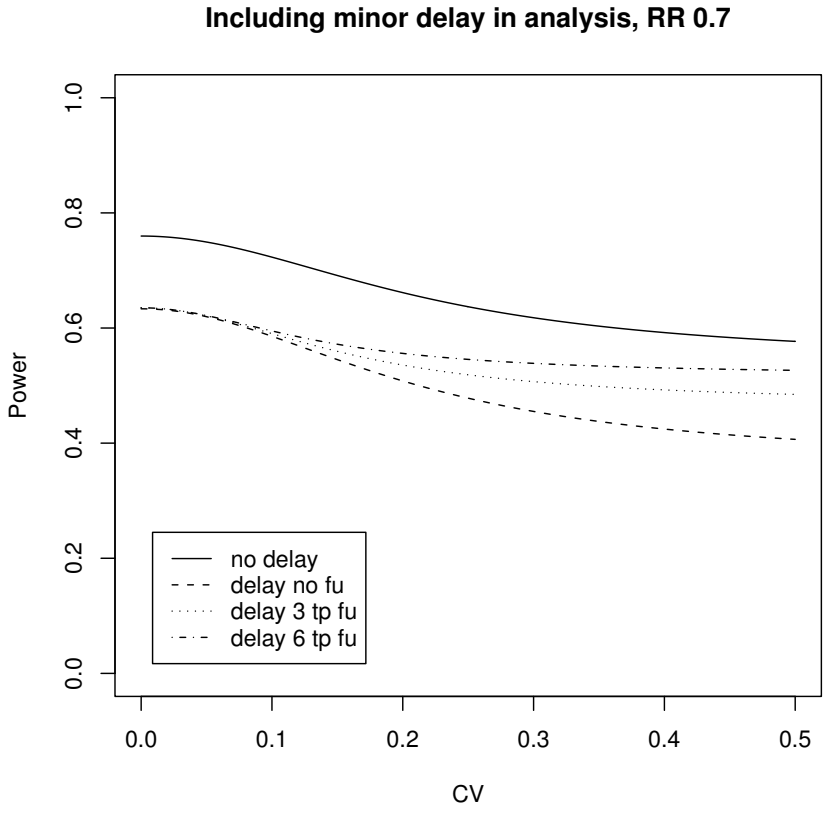
Example 1 - Expedited Partner Treatment

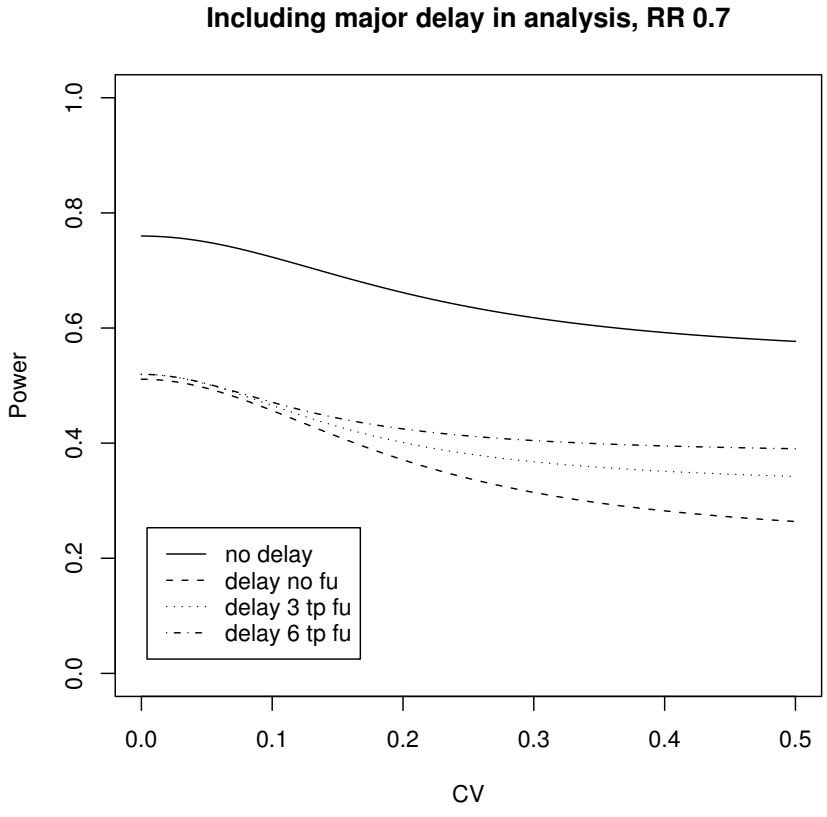


Number of steps



Delay in Treatment Effect





Data Analysis

- Linear Mixed Models
  - Assumes random effects model
  - Best if outcome measured at cluster level since ...
  - Transformations (e.g. logistic) difficult
- Generalized estimating equations
  - Robust variance structure
  - More “natural” for binary data
  - Outcome measured at cluster or individual level

Data Analysis

Table 1: Estimated power comparing clusters that have the same sample size ($N = 100$) and clusters with different sample sizes (24 clusters, 5 time points, $\tau^2 = 0.000225$, $\mu = 0.05$, 500 iterations)

	Same cluster sizes		Different cluster sizes	
Odds Ratio	LMM	GEE	LMM	GEE
1.0	0.054	0.056	0.040	0.044
0.7	0.688	0.706	0.298	0.694
0.6	0.912	0.914	0.510	0.896
0.5	0.976	0.976	0.704	0.984

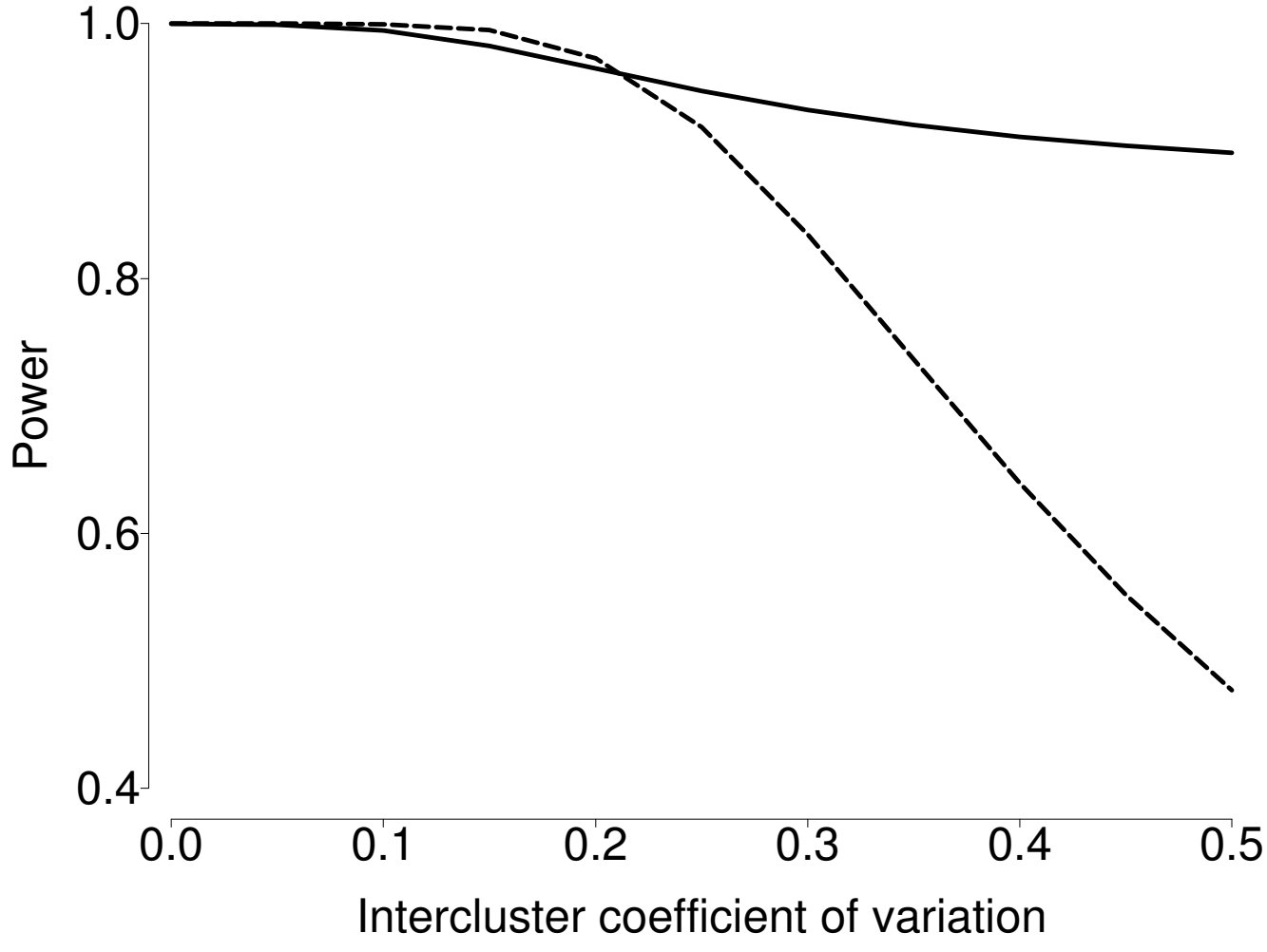
Example 2 - HIV research

- Mother to Child transmission of HIV can be drastically reduced using single dose Nevirapine
- Compare Targeted access vs Combined access
- Clinic randomized, outcome is % of HIV+ with NVP in cord blood
- Limited # clinics, unidirectional randomization needed

Example 2 - HIV research

		Time	
		1	2
Zambia	1	0	0
	2	0	1
	3	0	1
	4	1	1
Uganda	1	0	0
	2	0	1
	3	0	1
	4	1	1

Example 2 - HIV research



Summary - Stepped Wedge

- Stepped wedge useful for evaluation of public health interventions and “phase IV” trials
- Power relatively insenstive to between-cluster variation
- Maximize number of time steps
- Delayed treatment effect hurts power
- GEE most convienent for analysis

Acknowledgements

- Mike Hussey
- Dr. Matt Golden
- Dr. Jeff Stringer