Design and Analysis of Stepped Wedge Cluster Randomized Trials

Michael A. Hussey & James P. Hughes

Biost 572

Navneet R. Hakhu

May 7, 2013

Cluster Randomized Trial (CRT)

- Randomize (independent) clusters to intervention arm
  - Subjects within clusters are correlated

- Q: Why are CRTs useful?

Partner Notification

- Public health authorities contact sex partner
  - Of potential exposure to sexually transmitted infection (STI)
  - To seek treatment
  - Drawback: Implementation expensive

- Alternative: Patient Delivered Partner Therapy
  - Infected patient brings treatment to sex partner
    - Drugs or drug vouchers

Expedited Partner Therapy (EPT)

- Individually randomized trial [Golden et al., 2005]
  - 1998 to 2003 in King County, WA
  - Notification strategies (Intervention arms)
    - Patient delivered partner therapy, referred to as EPT
    - Standard partner notification (control)
  - Goal: To compare effectiveness of notification strategies for treating chlamydia and/or gonorrhea
    - Primary outcome: "presence of persistent or recurrent infection in the original index patient 3 – 19 weeks after treatment"
  - Study results
    - Significantly increased proportion of partners treated
    - Decreased risk of infection in patients

- Q: Successful trial, but are we done?

Limitation of EPT

- Q: What about all the other counties in WA state?
  - King county is not representative of every county in WA
- Goal for WA: To implement EPT in every county
  - Q: How?

Motivation for CRT

- Individually randomized trial completed
  - But only for one county (King)
- New trial
  - Counties represent clusters
  - Q: What kind of CRT should we use?

Possible CRT Designs

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

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

Stepped Wedge		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	0	1	1	1
	3	0	0	0	1	1
	4	0	0	0	0	1

- Q: Which design is best from a scientific perspective?
- Q: Which design is best from a statistical perspective?

Statistical Perspective

- Criteria for best design
  - Power
    - Probability of rejecting null when alternative is true
    - For stepped wedge: Consider different effect sizes (i.e., number of clusters randomized at each time point
  - Coefficient of Variation (CV)
    - Ratio of between-cluster standard deviation over mean prevalence
  - Sample sizes within clusters
    - Equal versus unequal

Statistical Model: Individual-level

$$Y_{ijk} = \mu_{ij} + e_{ijk}$$
$$= (\mu + \alpha_i + \beta_j + X_{ij}\theta) + e_{ijk}$$

- $i = 1, \ldots, I$ (number of clusters)
- $j = 1, \ldots, T$ (number of time intervals)
- $k = 1, \ldots, N$ (number of individuals within cluster i at time j)

- μ : mean prevalence
- $\alpha_i \sim N(0, \tau^2)$: random effect for cluster
- β_j : fixed effect for time; $\beta_T = 0$ (for identifiability)
- $X_{ij} = 1_{[cluster\ i\ receiving\ treatement\ at\ time\ j]}$
- θ : treatment effect
- $e_{ijk} \sim_{iid} N(0, \sigma_e^2)$

- $Var(Y_{ijk}) = \tau^2 + \sigma_e^2$

Statistical Model: Cluster-level

$$Y_{ij+} = \mu_{ij} + e_{ij+}$$
$$= (\mu + \alpha_i + \beta_j + X_{ij}\theta) + \left(\sum_{k=1}^{N} e_{ijk}\right)$$

- $e_{ij+} \sim_{iid} N(0,\ \sigma^2)$ where $\sigma^2 = \frac{\sigma_e^2}{N}$
- $Var(Y_{ij+}) = \tau^2 + \sigma^2 = \cdots = \frac{\tau^2 + \sigma_e^2}{N}\ [1 + (N-1)\,\rho]$
- $\rho \equiv Corr(Y_{ijk}, Y_{ij'k'}) = \frac{\tau^2}{\tau^2 + \sigma_e^2} \neq 0$ (intraclass correlation)
- $Cov(Y_{ijk}, Y_{ij'k'}) = \tau^2$

Predictor of Interest

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

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

Stepped Wedge		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	0	1	1	1
	3	0	0	0	1	1
	4	0	0	0	0	1

$$X_{ij} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 10 \\ 10 \\ 01 \\ 01 \end{bmatrix} \qquad \begin{bmatrix} 01111 \\ 00111 \\ 00011 \\ 00001 \end{bmatrix}$$

$$(I, T) = (4, 1) \qquad (4, 2) \qquad (4, 5)$$

(Approximate) Statistical Power

- Testing $H_0 : \theta = 0$ versus $H_1 : \theta = \theta_A$

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

  - Φ : Cumulative density function of $N(0,1)$
  - $z_{1-\alpha/2}$: $\left(1-\frac{\alpha}{2}\right)$-quantile of $N(0,1)$
  - $\mathrm{Var}(\hat{\theta})$: Estimated from weighted least squares (WLS)

Estimated Variance from WLS

- Z : Design matrix, $IT \times (T + 1)$
- $\eta = (\mu,\ \beta_1, \dots ,\ \beta_{T-1},\ \theta)$: parameter vector, $(T + 1) \times 1$
- $V = \mathrm{diag}(V_1, \dots ,\ V_I)$: block diagonal matrix, $IT \times IT$

$$V_i = \begin{bmatrix} \tau^2 + \sigma_e^2 & \tau^2 & \cdots & \tau^2 \\ \tau^2 & \ddots & \tau^2 & \vdots \\ \vdots & \tau^2 & \ddots & \tau^2 \\ \tau^2 & \cdots & \tau^2 & \tau^2 + \sigma_e^2 \end{bmatrix}_{T \times T}$$

Estimated Variance from WLS

- Z : Design matrix, $IT \times (T+1)$
- $\hat{\eta} = (Z^T V^{-1} Z)^{-1} Z^T V^{-1} Y$: WLS estimate
  - $\hat{\theta}$: $(T+1)$ entry of $\hat{\eta}$
- $\mathrm{Cov}(\hat{\eta}) = (Z^T V^{-1} Z)^{-1}$: $(T+1) \times (T+1)$ matrix
  - $\mathrm{Var}(\hat{\theta})$: $(T+1) \times (T+1)$ entry of $\mathrm{Cov}(\hat{\eta})$

Analysis of CRT

- Population-level approach
  - Generalized Estimating Equations (GEE)
- Individual-level approaches
  - Linear Mixed Models (LMM)
  - Generalized Linear Mixed Models (GLMM)
- Some considerations
  - Known versus unknown variance components
  - Normal versus non-normal data
- Goal: Compare power from GEE, LMM, GLMM

Simulation Setup

- Generate data from the model

$$Y_{ijk} = \mu + \alpha_i + \beta_j + X_{ij}\theta + e_{ijk}$$

  - $\mu = 0.05$
  - $\tau^2 = 0.000225$
  - $\sigma_e^2 = \mu(1-\mu)/N$
  - $\theta_A = 0.25$

- Problem occurs when trying to generate binomial data have negative values

- When not assuming binary response, estimated SE is very small → power of 1, which is not correct

Summary

- Motivated CRTs
  - Expedited Partner Therapy individually randomized trial
  - Three designs: parallel, crossover, stepped wedge
  - After scientific consideration, we want to consider statistical aspects of the three designs
    - Power
    - CV (prevalence estimated from cross-sectional sampling)

- Next steps:
  - Figure out problems with simulations
    - Focusing on Power calculations for different values of
      - Treatment effect
      - CV
  - Extension: Compare Power for parallel versus stepped wedge
    - More comparable sample sizes
    - Different time steps