STEPPED WEDGE DESIGNS AND THE WASHINGTON STATE EPT TRIAL

Jim Hughes
Professor of Biostatistics
University of Washington
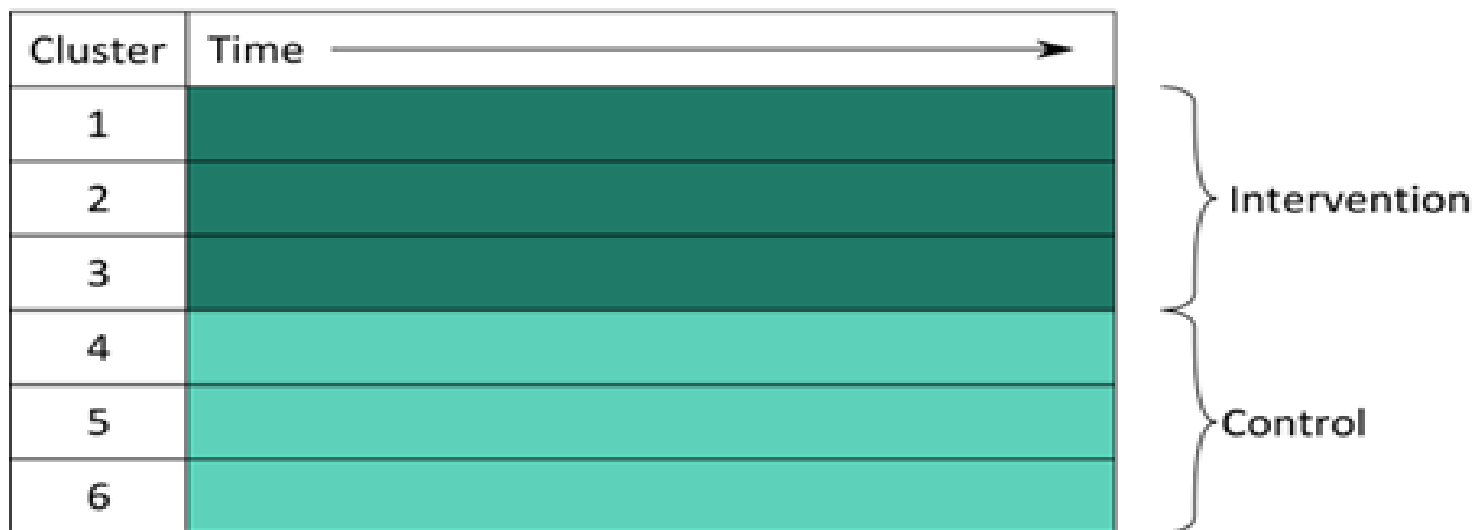
Cluster Randomized Trials

- Randomization at group level; outcome measured on individuals within the group

- Clusters may be large (cities, schools) … or small (IDU networks, families)

- Why? Individual randomization not feasible, potential contamination, or want to measure community effect

- Usually, larger, more complex than individually randomized trial

Common Trial Designs

(a) Parallel Cluster Study

Cluster	Time
1	
2	
3	
4	
5	
6	

Intervention

Control

Common Trial Designs

(b) Matched Pair Parallel Cluster study

Pair	Cluster	Time ⟶		
1	1			
1	2			
2	3			
2	4			
3	5			
3	6			

Intervention
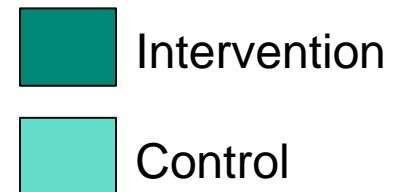
Control

Stepped wedge design

Cluster	Time →					

Intervention
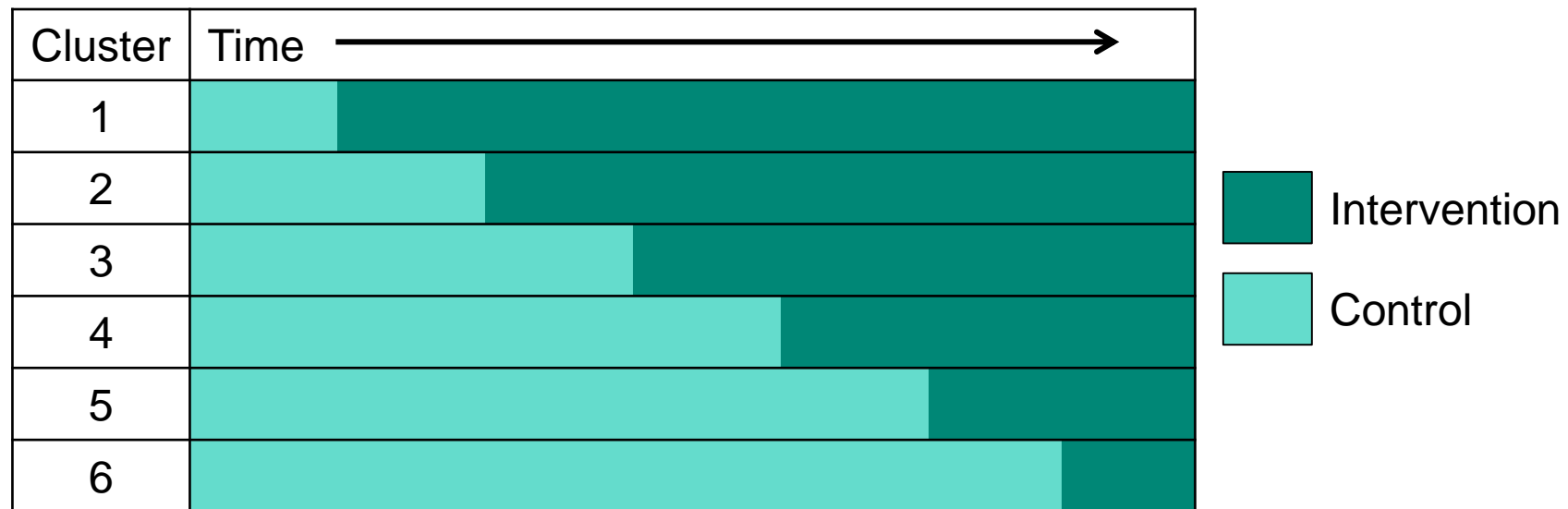
Control

- Clusters are randomized as to <u>when</u> intervention is received
- All clusters receive intervention eventually

Stepped wedge design

- Time in NOT balanced between intervention and control periods

- Need to be able to measure outcome <u>on each cluster</u>, <u>at each time step</u> (to control for time trends)

- Cross-sectional or cohort sampling possible

    - Repeated measurements on members of a cohort may result in significant participant burden

Advantages

- Acceptability (social, political, ethical) - all clusters receive the intervention

- Logistical or financial - cannot introduce the intervention in all units at once

- Efficiency: Units act as their own control, so (likely) fewer clusters needed

- Possible to study the effect of time on intervention effectiveness (i.e. seasonality, time since introduction)

Disadvantages

- Long time to completion
  - Increased potential for contamination
  - Increased potential for external events to influence study
  - Potential for clusters scheduled for a later start to "jump the gun"

- Relatively complex analysis
  - Intentional confounding of time and treatment must be resolved using e.g. regression analysis
  - Dependent on assumptions

WA State EPT

- Expedited partner treatment for Gc and Ct in WA state

- EPT shown to be effective in reducing <u>reinfection</u> in IRT (Golden et al., NEJM, 2005)

- EPT to be implemented throughout Washington state; logistically difficult to implement the program in all counties simultaneously

- Solution: use a SW design; (24) LHJs are the randomization units; randomize 6 per time period

- Outcome (STI) measured in sentinel sites

- Six month intervals – 3 to implement, 3 to assess outcome

WA State EPT

	TIME				
	Baseline	Step 1 10/7	Step 2 6/08	Step 3 1/09	Step 4 8/09
Wave 1 (6 LHJs)		Intervention			
Wave 2 (6 LHJs)			Intervention		
Wave 3 (6 LHJs)				Intervention	
Wave 4 (5 LHJs)					Intervention

Statistical Issues - Model

Model:

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

$a_i \sim N(0,\tau^2)$ – variation in mean between clusters

$c_i \sim N(0,\eta^2)$ – variation in tx effect between clusters

$e_{ijk} \sim N(0,\sigma^2)$ – random variation

Notes:

1) "Standard" SW model does not include treatment heterogeneity

2) Model shown above assumes same time effect in all clusters

3) Express the variation in terms of the *coefficient of variation* (CV) between clusters – τ/μ and η/μ

Statistical Issues - Power

- Power = Probability of detecting a treatment effect when the treatment really works

- Depends on …
  - strength of treatment effect
  - number of clusters
  - number of steps
  - number participants per cluster per step,
  - variance components: σ^2 (easy to know) , η^2 ,τ^2 (hard to know).

Power – Variance Components

WA State EPT
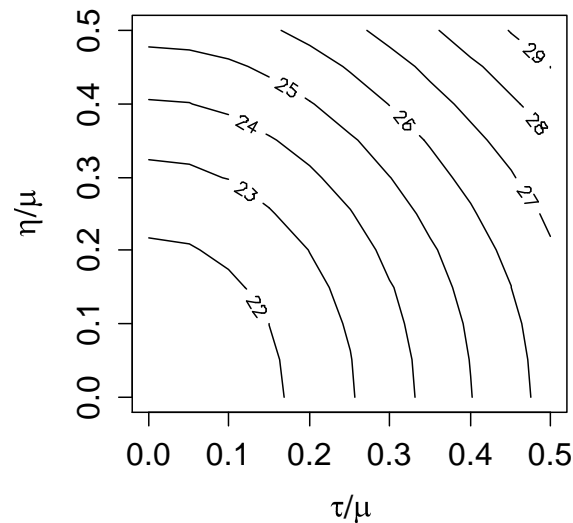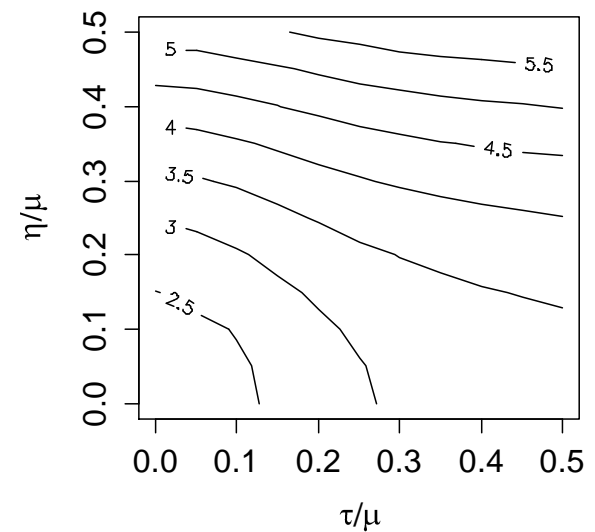μ = 0.05
I = 24 LHJ's
J = 4 waves
θ = -0.015

N=12

N=120

Contours of $Var(\hat{\theta})$ ($\times 10^5$) as a function of τ and η

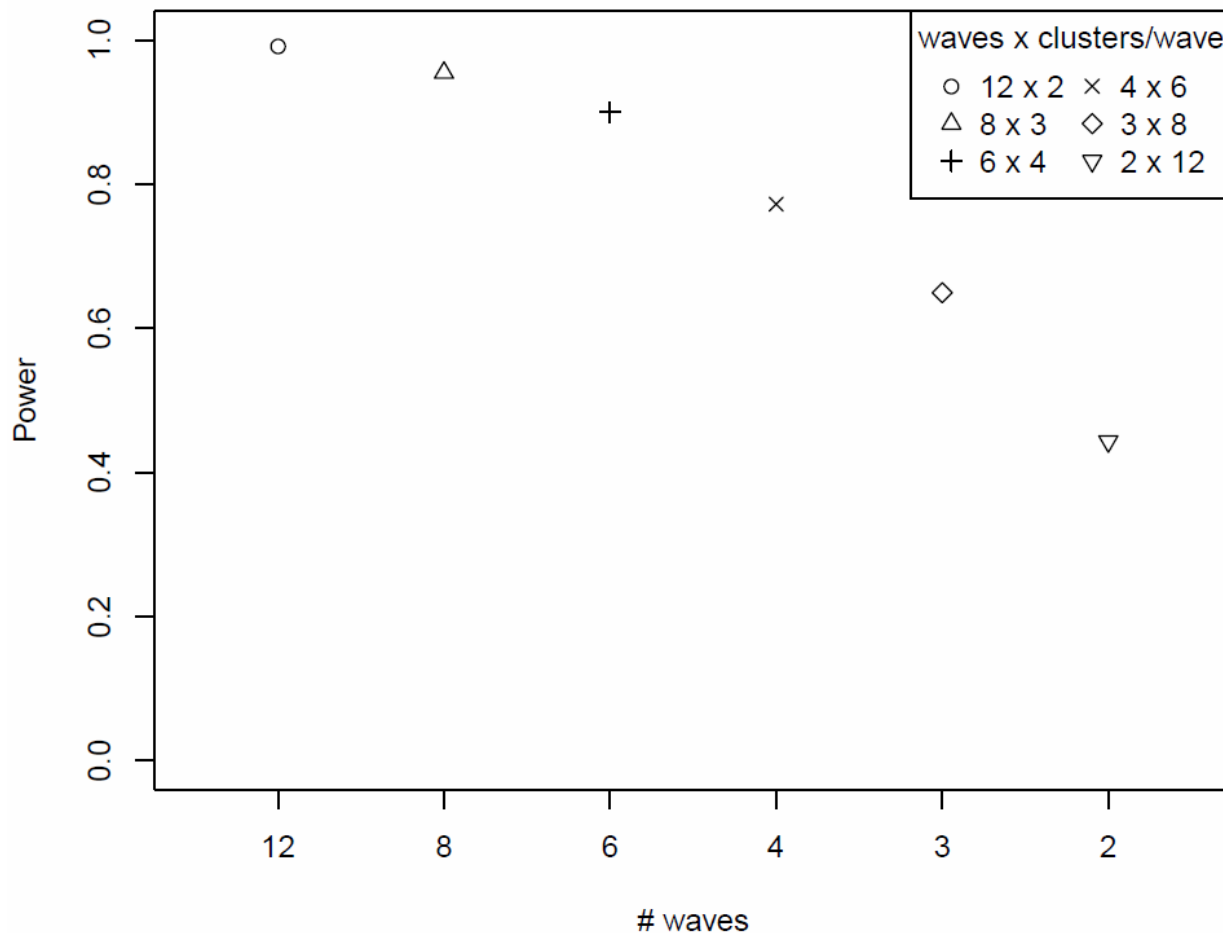# Power vs # waves

WA State EPT

$\mu = 0.05$
$I = 24$ LHJ's
$N = 120$
$\theta = -0.015$
$\tau / \mu = 0.20$
$\eta/\theta = 0.30$

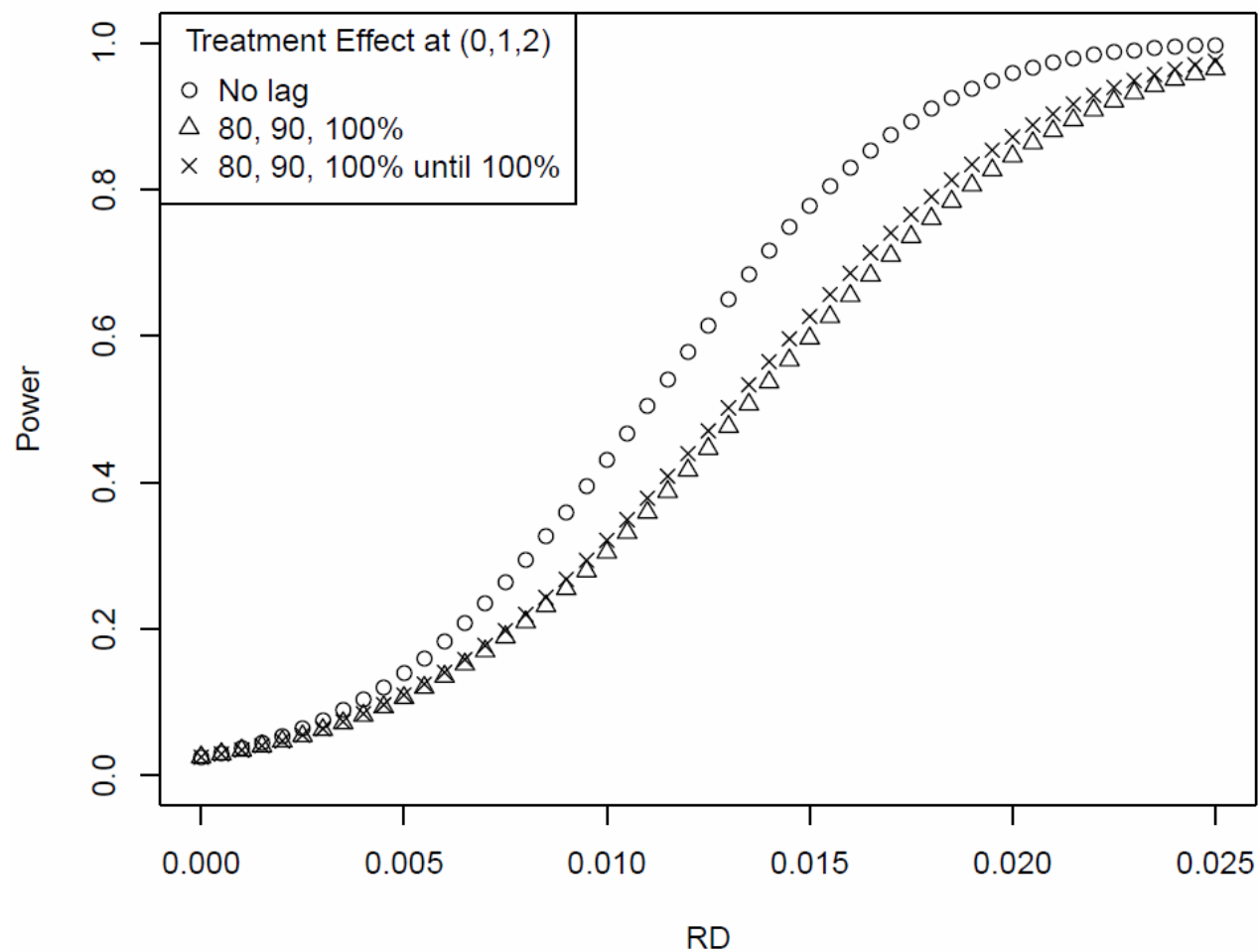# Power – Delayed treatment effect

WA State EPT

$\mu = 0.05$
$I = 24$ LHJ's
$J = 4$ waves
$N = 120$
$\tau/\mu = 0.20$
$\eta/\theta = 0.30$

Statistical Issues - Analysis

- Use regression based analysis (GEE, GLMM)
  - Controls for time trends and correlated data
  - Uses both within and between cluster info
  - Analyze cluster level means (if equal cluster sizes) or individual level data
  - Dependent on modelling assumptions
- "Vertical" analyses
  - Compare intervention and SOC at each time point and combine
  - Valid but less efficient
  - More robust?

Is the SW design the right design?

- Consider logistical and ethical issues, social and political acceptability

- SW useful for rollout/implementation studies
  - For intervention A vs intervention B, parallel cluster RCT (perhaps matched) may be better

- SW is sensitive to cluster variation in intervention effect

- SW confounds time trends with the intervention effect
  - → ALWAYS need to control for time trends (possibly within strata)

- Lag (time delay) in intervention effect reduces power
  - → Design step length > time lag

- Consider potential for changes in policy, other external factors not under investigator control

Resources

Recent Reference

- Hughes JP, Granston TS, Heagerty PJ. On the design and analysis of stepped wedge trials. Contemporary Clinical Trials. 45(Pt A):55-60, 2015.

Software: http://faculty.washington.edu/jphughes/pubs.html

- Excel spreadsheet for power calculations (does NOT include cluster to cluster variation in treatment effect)

- R package for power calculation (including cluster to cluster variation in treatment effect), data tabulation, plotting