STEPPED WEDGE DESIGNS AND THE WASHINGTON STATE EPT TRIAL

Jim Hughes
Professor of Biostatistics
University of Washington

HSR&D cyberseminar
June 7, 2016

Poll Question #1

How familiar are you with Stepped Wedge designs?

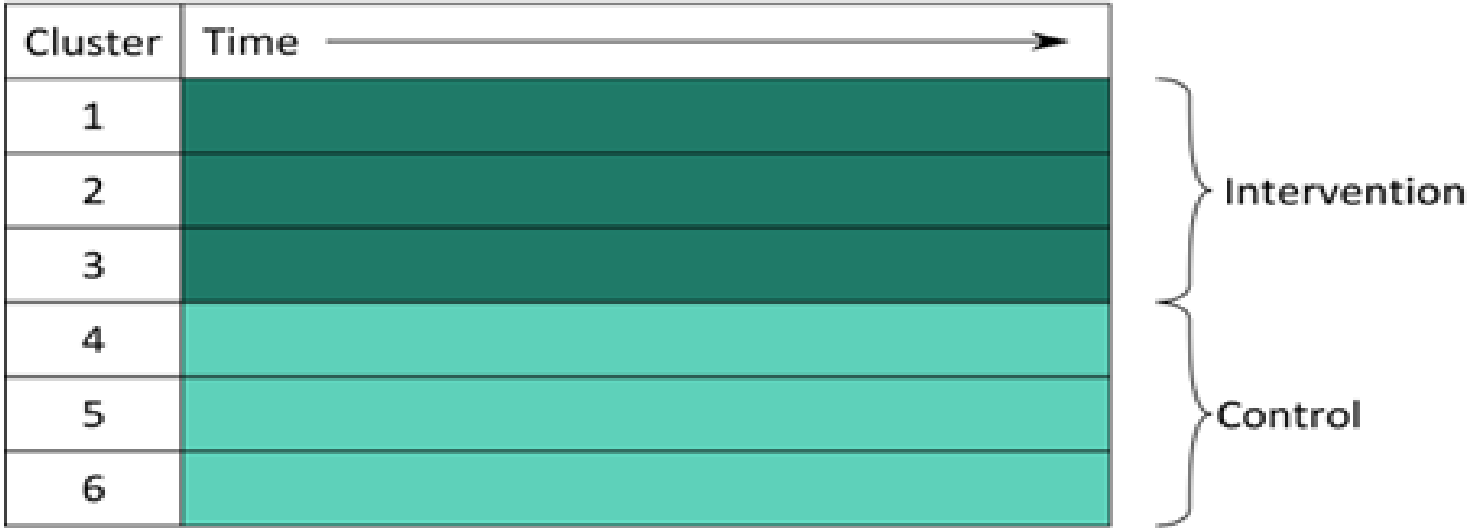
1. Never heard of them before this talk
2. I've heard the term but do not know the details
3. I'm familiar with the concept
4. I have used stepped wedge designs
5. I am an expert at stepped wedge designs

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
- Key statistical challenge: individuals not independent
- Usually, larger sample size than individually randomized trial

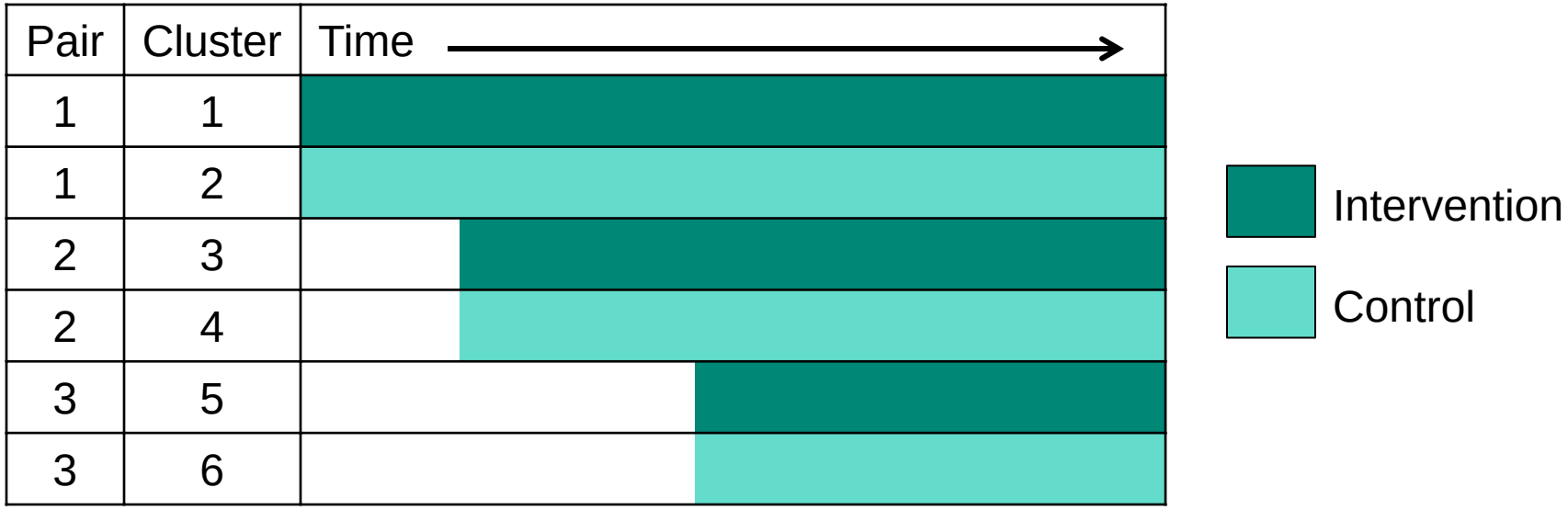
Common Trial Designs

(a) Parallel Cluster Study

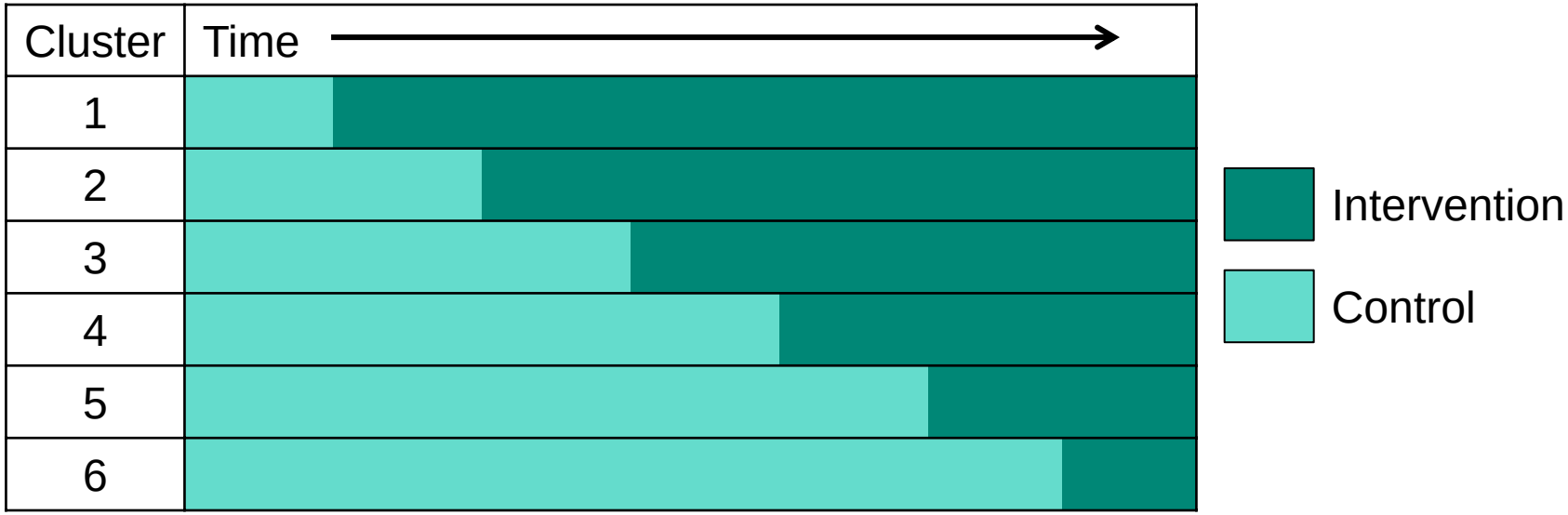


Common Trial Designs

(c) Matched Pair Parallel Cluster study



The stepped wedge design



- Phased implementation, all clusters receive intervention eventually
- Clusters are randomized as to <u>when</u> intervention is received

The stepped wedge design

- Time in NOT balanced between intervention and control periods
- Need to be able to measure outcome on each cluster, at each time step (to control for time trends)
- Cohort or cross-sectional 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
- Differential loss to followup
  - Potential for clusters scheduled for a later start to “jump the gun”
  - Equal loss to followup by cluster may result in unequal loss by intervention arm (!)
- Relatively complex analysis (intentional confounding of time and treatment must be resolved using e.g. regression analysis)

Example

- Introduction of HBV vaccination in infants in The Gambia (The Gambia Hepatitis Study Group, 1987)
- Cluster randomized (Health districts)
- 17 health districts, but program could not be implemented in all districts at the same time
- Immediate outcome: HBV antibody titre
- Longterm outcome: Hepatocellular cancer and other liver disease (results expected 2017!)

WA State EPT

- Expedited partner treatment for Gc and Ct in WA state
- EPT shown to be effective in reducing reinfection 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

Key issue in a CRT: $Corr(Y_{ijk}, Y_{ij'k'}) \neq 0$

Express the correlation 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

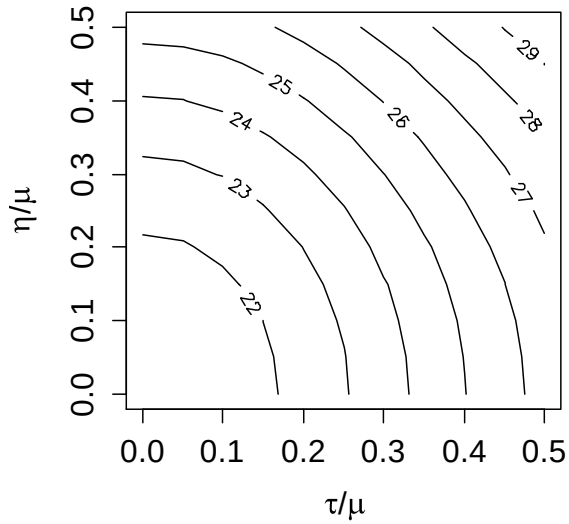
$\mu = 0.05$

I = 24 LHJ's

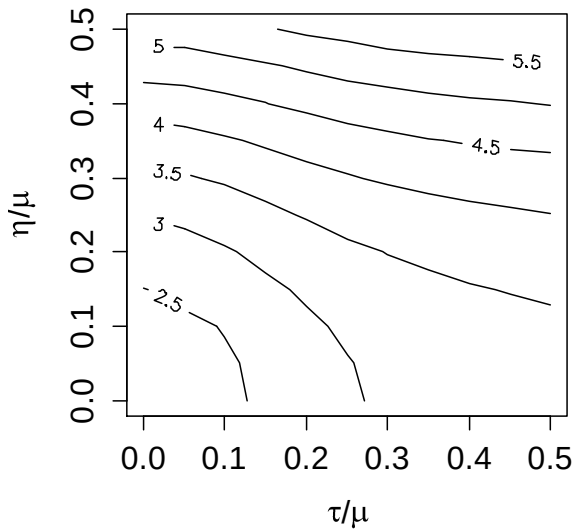
J = 4 waves

$\theta = -0.015$

N=12



N=120

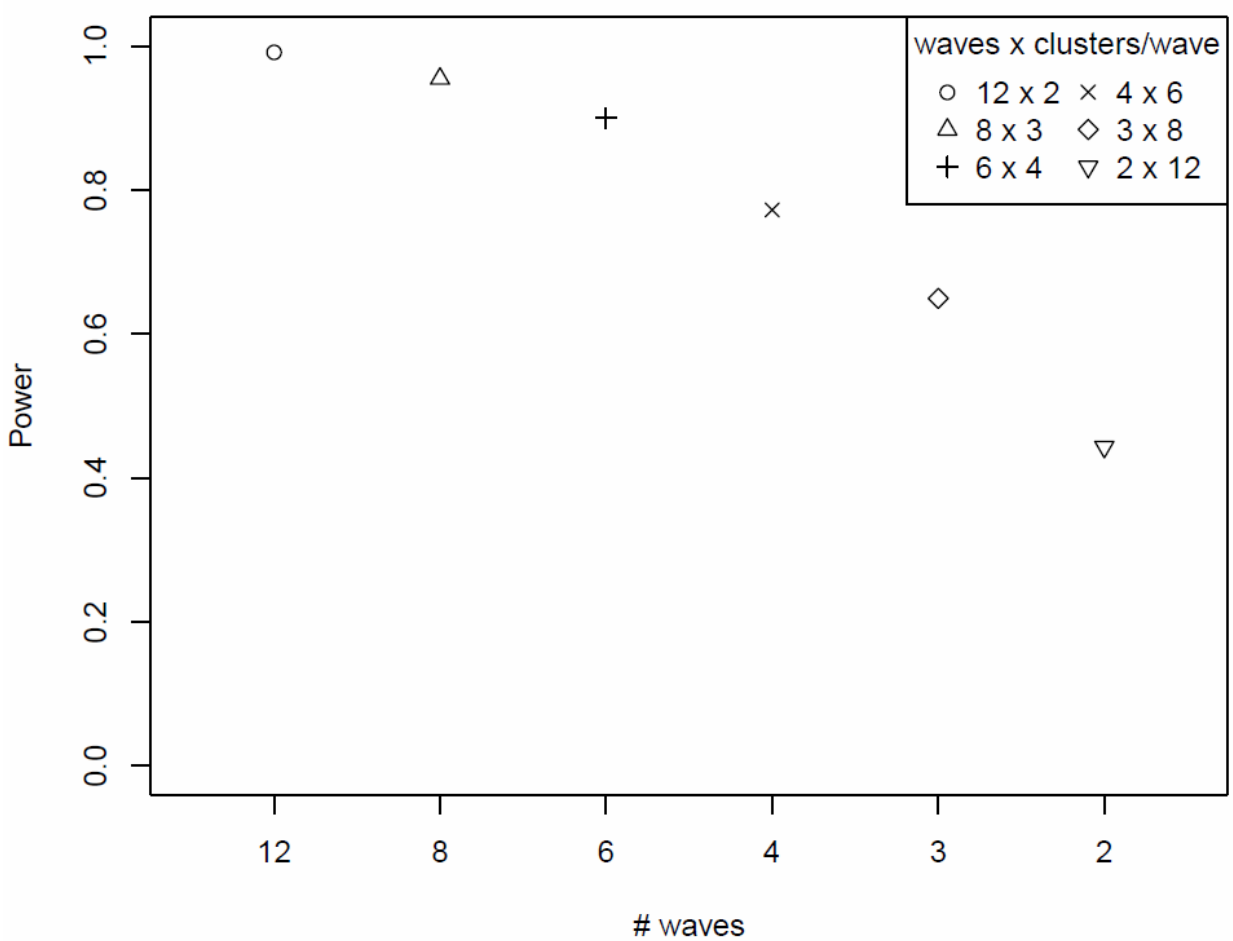


Contours of Var(θ) (× 10^5) as a function of τ and η

Power vs # waves

WA State EPT

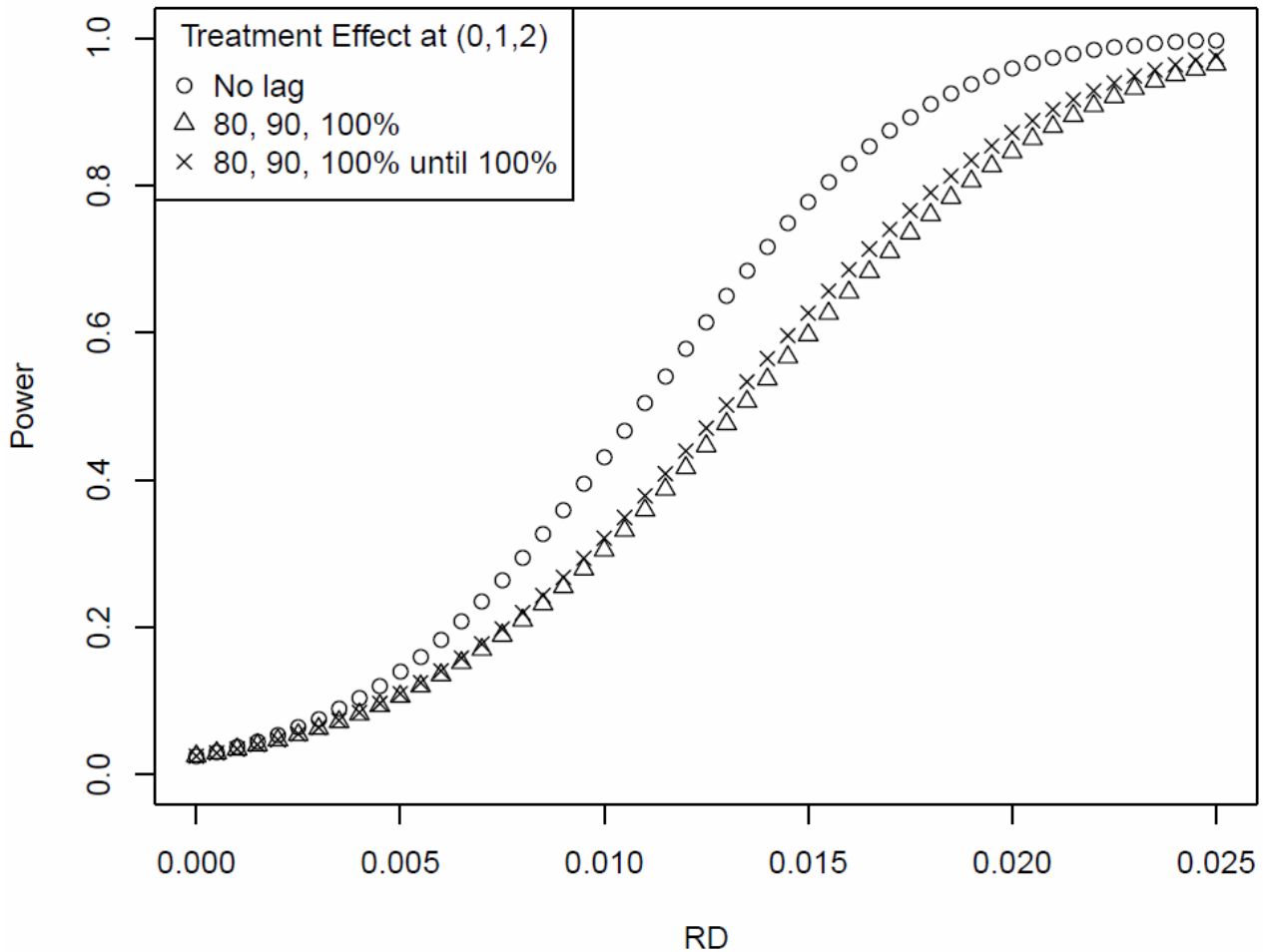
$\mu = 0.05$
I = 24 LHJ's
N = 120
$\theta = -0.015$
$\tau / \mu = 0.20$
$\eta / \theta = 0.30$



Power – Delayed treatment effect

WA State EPT

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

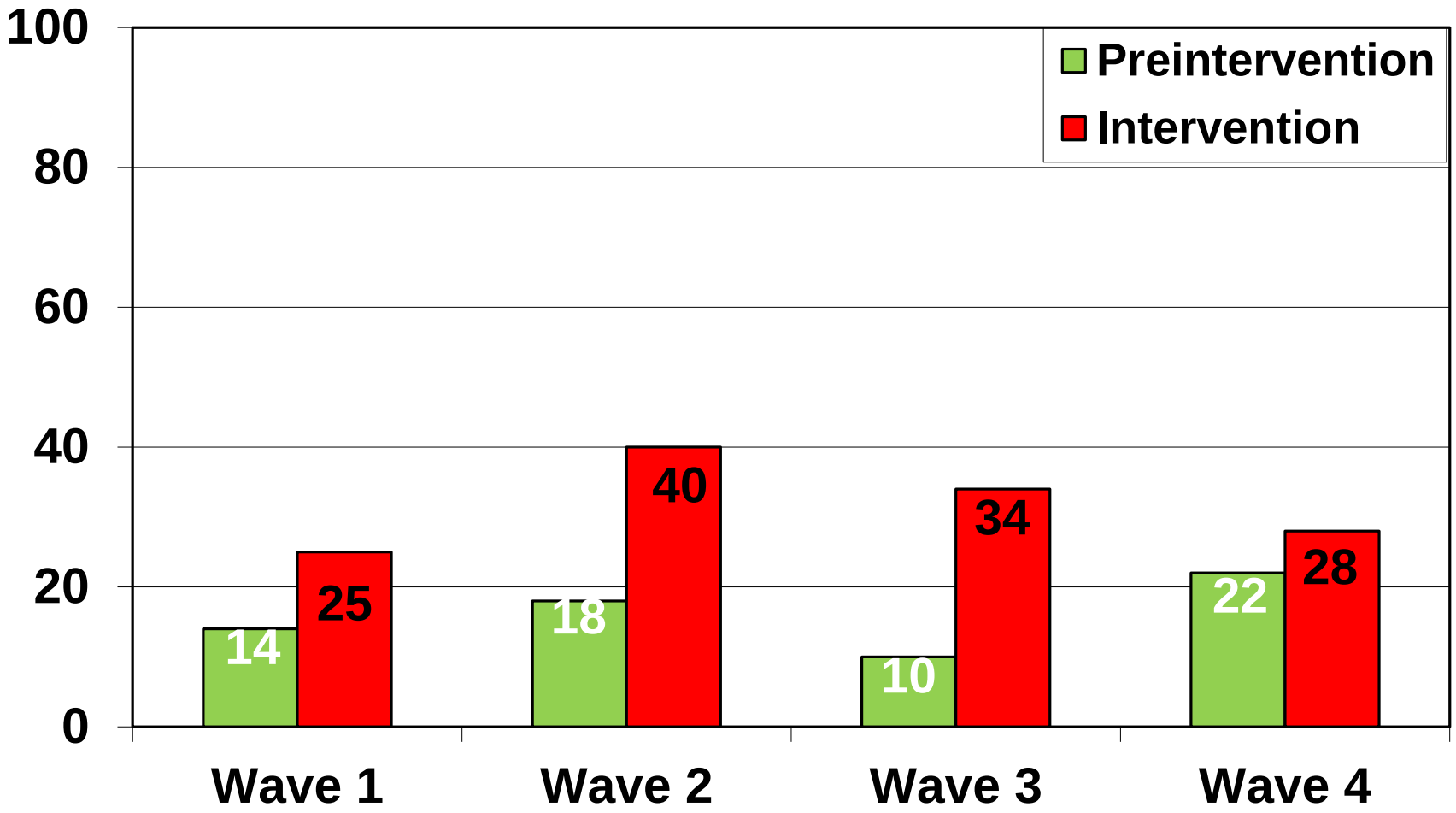


Statistical Issues - Analysis

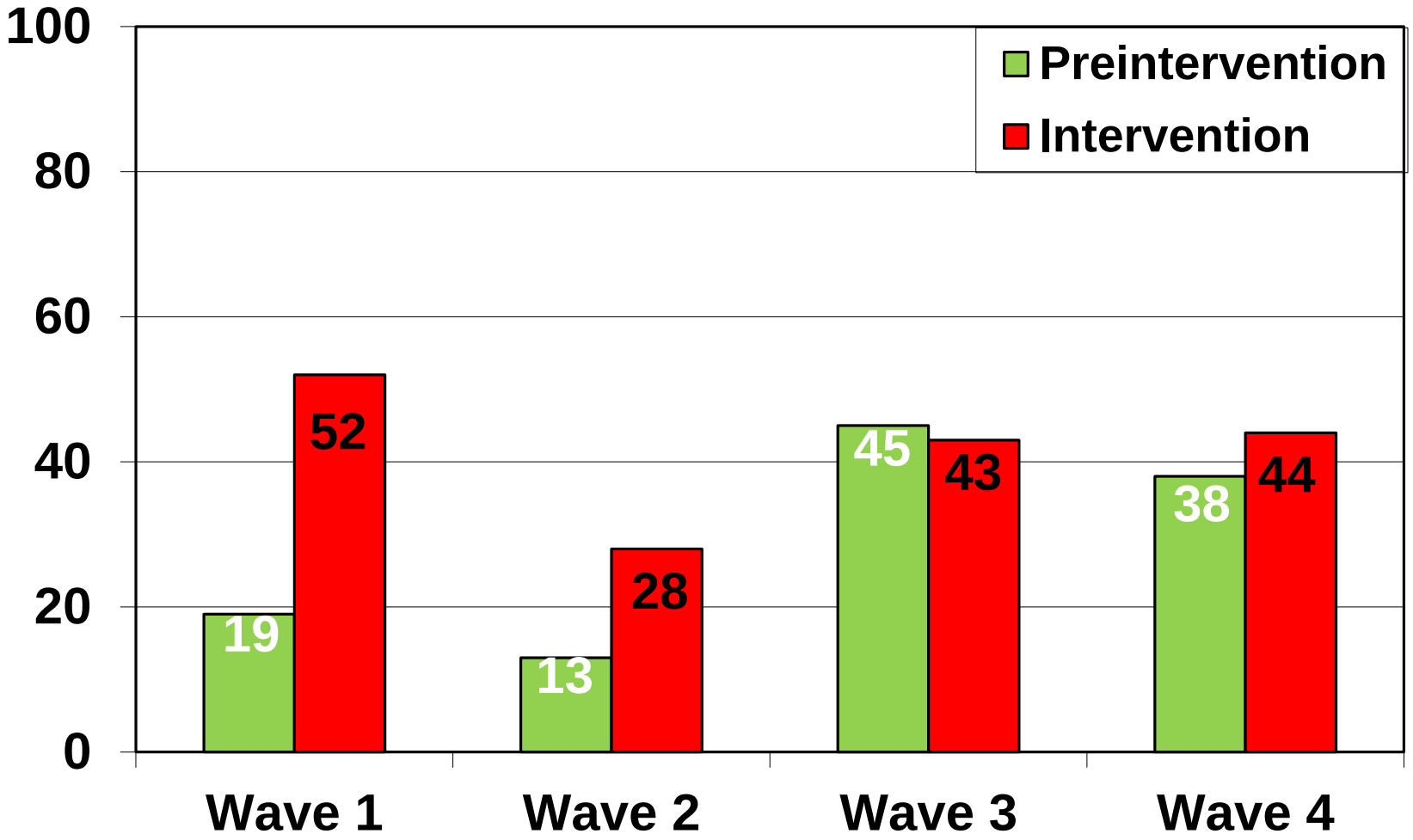
- Paired t-test
  - Analyze cluster means, before vs after
  - Not valid! Likely biased if there are time trends
- Repeated cross-sectional (in time) comparisons
  - Controls for time, so valid (unbiased)
  - Loses strength of within-unit comparisons
- GEE, GLMM
  - Use both within and between cluster info
  - Equal cluster sizes – analyze cluster level means
  - Unequal cluster sizes – analyze individual level data
  - Small number of clusters – use small sample adjustment to Wald statistic (Scott, 2008)

WA State EPT Trial - Results

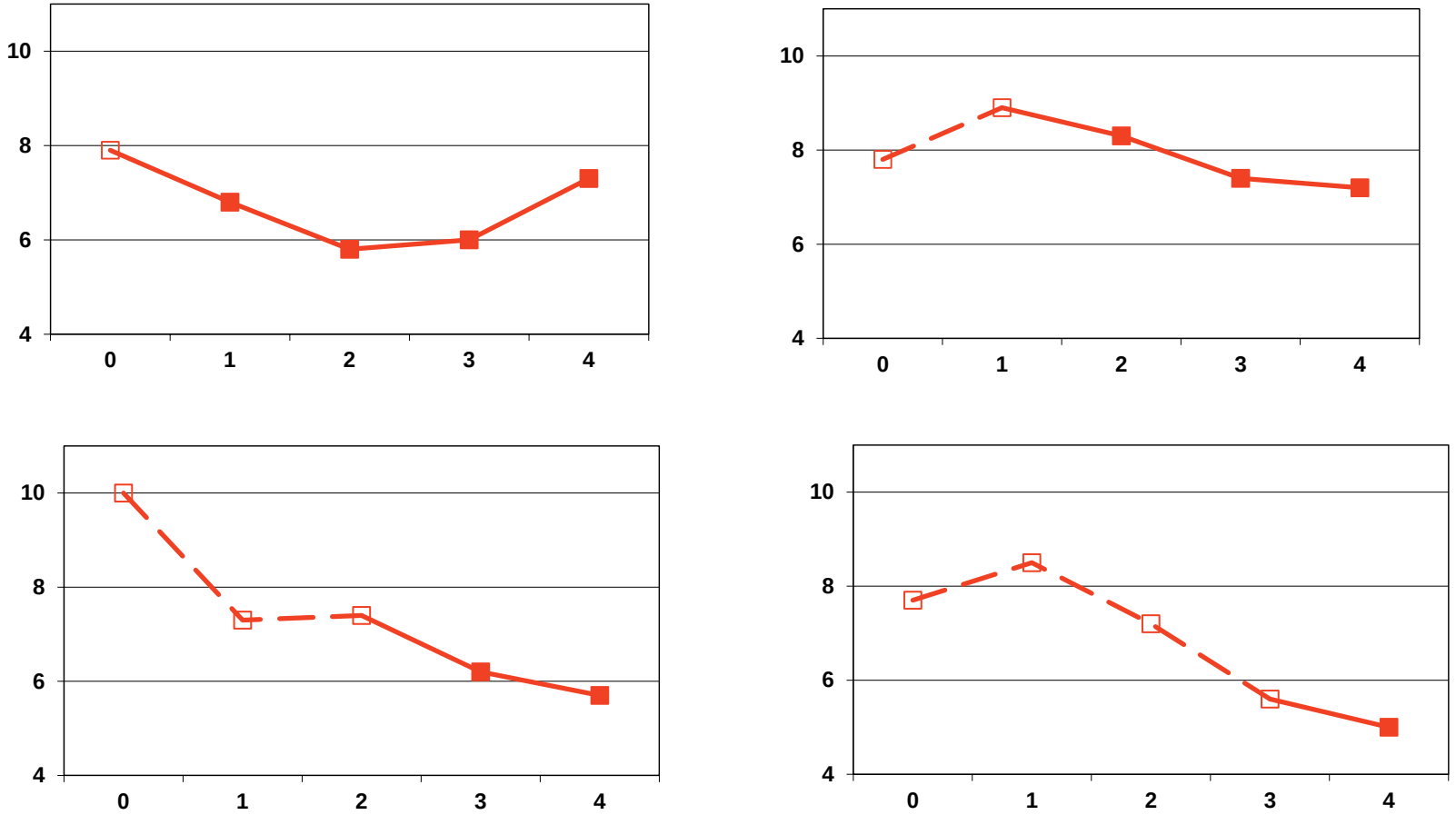
Percentage of Heterosexuals with Gonorrhea or Chlamydia Receiving PDPT from Medical Providers, Before and After Institution of the Study Intervention



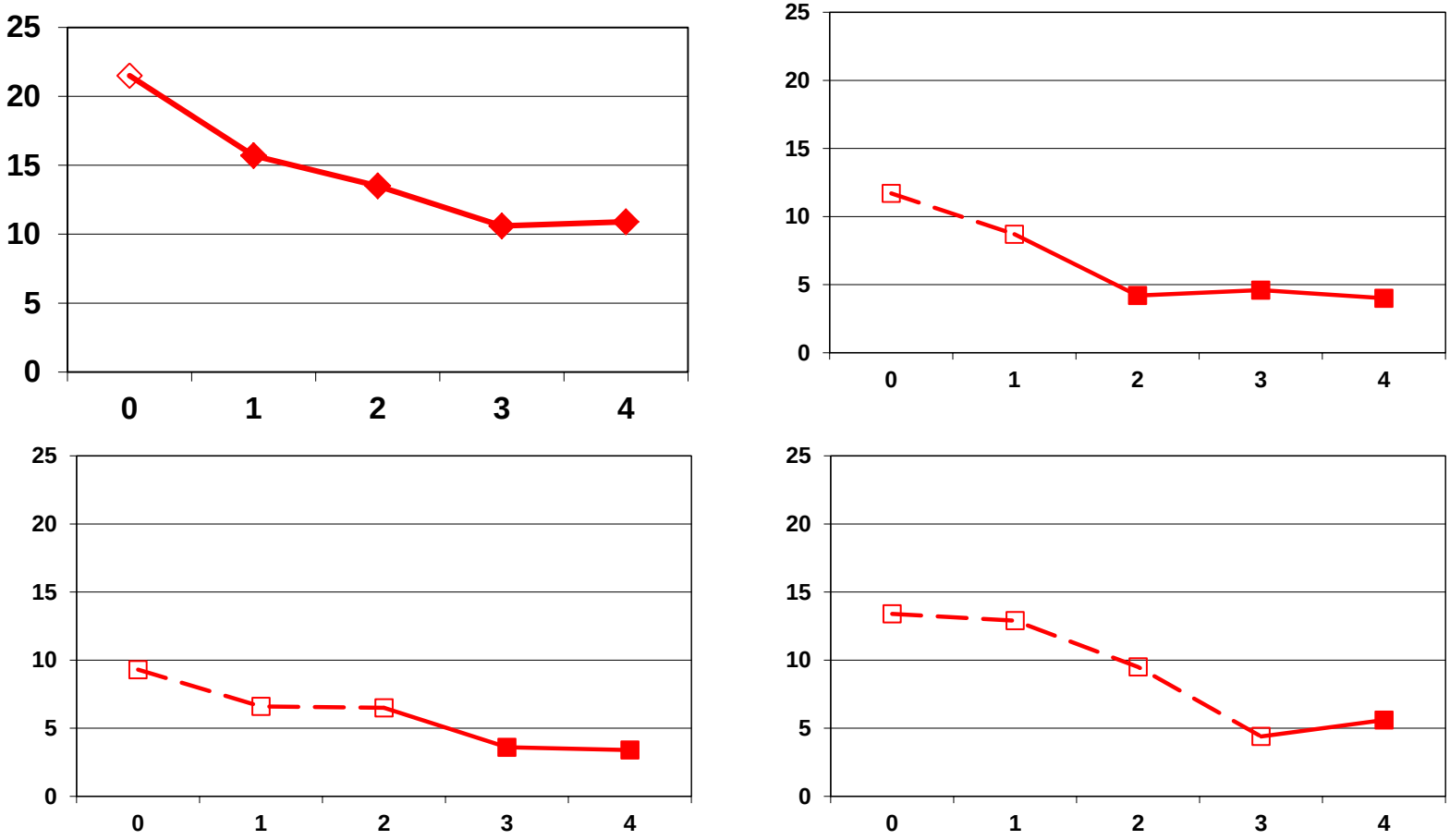
Percentage of Heterosexuals with Gonorrhea or Chlamydia Receiving Public Health Partner Services Before & After Institution of the Study Intervention



WA State EPT - Chlamydia Positivity Among Women Age 15-25 in Sentinel Populations of Women, by Study Wave



WA State EPT - Gonorrhea Incidence in Women, by Study Wave



WA State EPT Trial - Results

- Primary analysis based on last 3 months of each step
- Chlamydia positivity among women age 15-25 (ave. n = 300/LHJ/step)
  - Mixed effects generalized linear model with log link
  - Random effects for LHJ & testing clinic to adjust for variable number of tests in IPP clinics during different periods
- Gonorrhea Incidence in women
  - Based on state surveillance data
  - Marginal Poisson regression with robust variance, pop. size offset

	Risk Ratio	**95% CI**
Chlamydial positivity in women	.89	.77-1.04
Gonorrhea incidence in women	.86	.69-1.08

WA State EPT - Issues

- Competing intervention outside of investigator control
- Potential for contamination
  - of outcome
  - of intervention
- Lagged effects?
  - Delay in implementation
  - Lag in effect

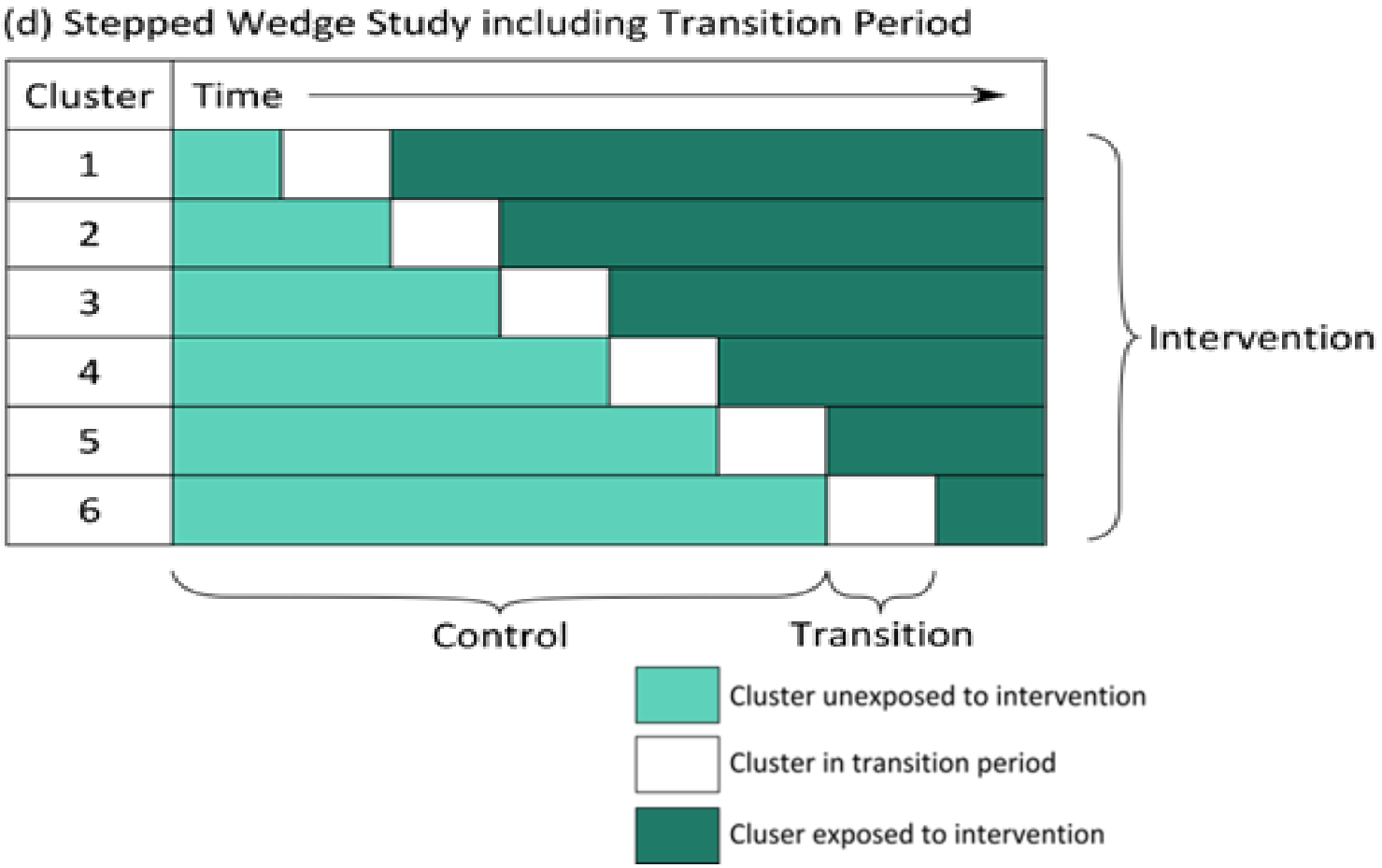
Is the SW design the right design?

- SW useful for rollout/implementation studies, esp intervention vs status quo
  - For intervention A vs intervention B, parallel cluster RCT (perhaps matched) may be better
- SW requires fewer clusters
- Less sensitive to cluster variation in mean, still 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, or add transition period

Stepped wedge with transition period



Is the SW design the right design ?

- Consider logistical and ethical issues, social and political acceptability
- Phased implementation may be necessary, BUT …
  - Must stick to implementation schedule
  - Must measure outcomes at same calendar times in all clusters (to control for time trends)
  - Maintain engagement with control clusters to avoid dropout
- Consider potential for changes in policy, other external factors not under investigator control

Thanks

Mike Hussey, MS

Navneet Hakhu, MS

Tanya Granston, PhD

Matt Golden, MD

Jeff Stringer, MD

Andrew Forbes, PhD

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