UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Thinking Outside the Box of RCT's: What Additional Rigorous Options are There?

David Henry

The Way Forward for ACF Research with American Indians and Alaska Natives

April 18, 2014

Acknowledgements

Center for Alaska Native Health Research (CANHR)

- Jerry Mohatt (d.)
- Jim Allen
- Carlotta Fok
- Stacy Rasmus
- Bill Charles
- People Awakening Team
- "Small Sample Conference"

Families and Communities Research Group

- Patrick Tolan
- Deborah Gorman-Smith
- Michael Schoeny

Chicago Center for Youth Violence Prevention

Rush University College of Nursing

Does banging pans cause elephants to flee?

- Charley has an effect, but no *counterfactual.*
- He does not know what would have happened if he had not banged the pans.
- The *counterfactual* is the same people, place and time but without Charley banging pans.
- Research methods exist to provide estimates of the *counterfactual*.
- The *counterfactual* is essential for determining whether our interventions have caused the effects we observe.

Cause and Effect: Other Views

- The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm. (Russell, 1913, p. 1).
- A “causal line,” as I wish to define the term, is a temporal series of events so related that, given some of them, something can be inferred about the others whatever may be happening elsewhere. (Russell, 1948, p. 459).

Cause and Effect

- Neyman-Rubin Causal Model
  - Let Y(0) be an outcome in the absence of treatment
  - Let Y(1) be the same outcome with treatment
  - Effect of the treatment = Y(1) – Y (0)
- The problem of estimating cause and effect is a problem of missing information.
  - Y(0) is missing for those who receive treatment
  - Y(1) is missing for those who do not receive treatment

Counterfactual

- The factual is a situation with certain actors at a certain time with intervention present.
- The counterfactual is exactly the same situation and actors at the same time absent the intervention.
- All research methods approximate the counterfactual
- All research methods have strengths and weaknesses for approximating the counterfactual.

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Threats to Internal Validity:

Why a study may fail to approximate the counterfactual

(Campbell & Stanley, 1967; Shadish, Cook and Campbell, 2002)

1. Ambiguous temporal precedence
2. Selection
3. History
4. Maturation
5. Regression
6. Attrition
7. Testing
8. Instrumentation
9. Interactive Effects

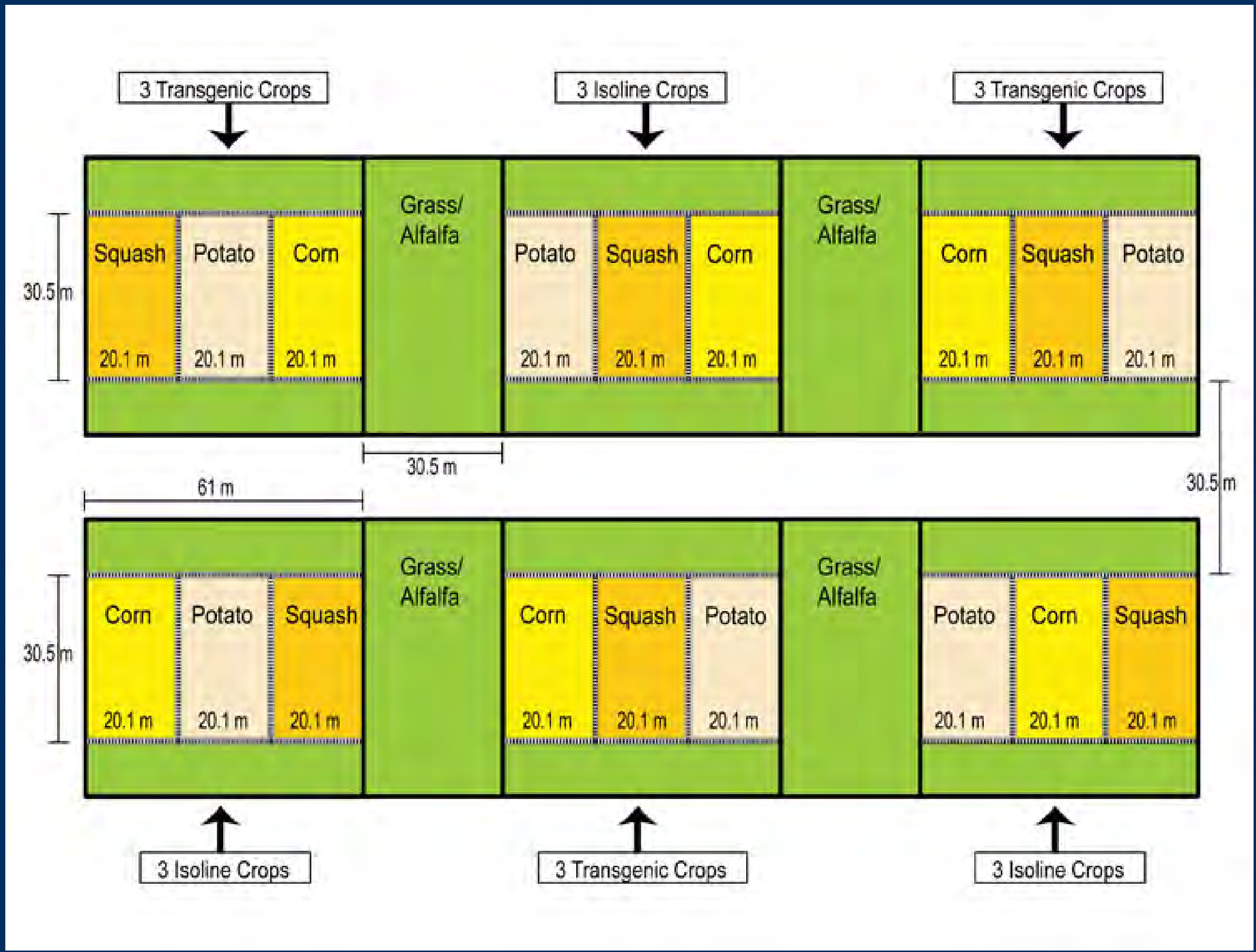
Examining the Box:

How RCT approximates the counterfactual

- Sampling representative of a defined population
- Random assignment
- Control conditions
- Blinding of participants and investigators.

Randomness and the Counterfactual

- Randomness in selection and assignment assumes interchangeable units.
  - Exclusion criteria
- Randomness in selection and assignment assumes stable units.
  - ITT
- Randomness in selection and assignment assumes consistent environments
  - Split plot, HLM



- **Figure 1.** Split-plot experimental design consisting of 3 replicates of two genotypes: transgenic and isoline. Main plots were split among 3 crops consisting of sweet corn, potatoes, and winter squash. Transgenic cultivars targeted Lepidoptera in corn, Coleoptera in potato and aphid-transmitted viruses in squash.

Blinding and the Counterfactual

- If research participants do not know what treatment they are receiving, there is less likelihood of biased reports and ratings.
- If assessors do not know what condition they are assessing, there is lower likelihood of "halo" effects.

RCT is a good design in some circumstances.

- Best when
  - Condition can be hidden from subjects
  - Condition can be hidden from assessors
  - Condition can be hidden from investigators
  - Individuals can be randomly selected from a pure population
  - Assignment procedures are truly random
- Has problems if
  - Conditions cannot be hidden
  - People move
  - People may not cooperate with treatment
  - Naïve hypotheses can affect measures
  - Focal issue does not occur in isolation

Randomized Controlled Trial

Threat to internal validity	Performance
Ambiguous temporal precedence	+ Intervention precedes outcome change
Selection	? Recruits may opt out when assigned
History	+ affects I and C equally
Maturation	+ affects I and C equally
Regression	+ affects I and C equally
Attrition	? Differential attrition of I and C
Testing	+ affects I and C equally
Instrumentation	+ does not change between I and C
Interactive Effects	? cooperation with treatment

Thinking Outside the Box: Three Alternative Designs

1. Roll-out Designs (DWLD and SWD)
2. Interrupted Time Series Design (ITSD)
3. Regression Point Displacement Design (RPDD)

1. Roll-out Designs

- Logistics may prohibit starting all communities, schools, settings at the same time.
- Power can be increased by randomizing start times as well as settings (Brown et al., 2006, 2007).

Time frame		Standard wait-listed design		Full dynamic wait-listed design	
Period	Time block	Trained	Not trained	Trained	Not trained
1	1	16	16	4	28
	2			8	24
	3			12	20
	4			16	16
2	1	32	0	20	12
	2			24	8
	3			28	4
	4			32	0

Dynamic Waitlisted Design

- Works well for time-limited interventions where maintenance of effects is a concern.
- Power is increased because every community serves as its own control group.
- Analysis: Mixed effects regression model with random effect for community and fixed effects for time, period (observation, intervention, followup) and interactions.

Dynamic Waitlisted Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O					
	2	O					
	3	O					
	4	O					
	5	O					

X=intervention, O=observation

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Dynamic Waitlisted Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	O				
	2	O	X				
	3	O	O				
	4	O	O				
	5	O	O				

X=intervention + observation, O=observation only

Dynamic Waitlisted Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	O	O			
	2	O	X	O			
	3	O	O	O			
	4	O	O	X			
	5	O	O	O			

X=intervention + observation, O=observation only

Dynamic Waitlisted Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	O	O	X		
	2	O	X	O	O		
	3	O	O	O	O		
	4	O	O	X	O		
	5	O	O	O	O		

X=intervention + observation, O=observation only

Dynamic Waitlisted Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	O	O	X	O	
	2	O	X	O	O	O	
	3	O	O	O	O	O	
	4	O	O	X	O	O	
	5	O	O	O	O	X	

X=intervention + observation, O=observation only

Dynamic Waitlisted Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	O	O	X	O	O
	2	O	X	O	O	O	O
	3	O	O	O	O	O	X
	4	O	O	X	O	O	O
	5	O	O	O	O	X	O

X=intervention + observation, O=observation only

Dynamic Waitlisted Design

Example: QPR Study

- Brown et al., 2006; Wyman et al., 2008
- Gatekeeper training
- 32 north Georgia schools selected
- Standard waitlist design used in the first year
- Training inefficiency led to switch to DWLD
- Poisson mixed effects regression used to detect change in rate of referral for suicidality

Stepped Wedge Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O					
	2	O					
	3	O					
	4	O					
	5	O					

X=intervention + observation, O=observation

Stepped Wedge Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	X				
	2	O	O				
	3	O	O				
	4	O	O				
	5	O	O				

X=intervention + observation, O=observation

Stepped Wedge Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	X	X			
	2	O	O	X			
	3	O	O	O			
	4	O	O	O			
	5	O	O	O			

X=intervention + observation, O=observation

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Stepped Wedge Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	X	X	X		
	2	O	O	X	X		
	3	O	O	O	X		
	4	O	O	O	O		
	5	O	O	O	O		

X=intervention + observation, O=observation

Stepped Wedge Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	X	X	X	X	
	2	O	O	X	X	X	
	3	O	O	O	X	X	
	4	O	O	O	O	X	
	5	O	O	O	O	O	

X=intervention + observation, O=observation

Stepped Wedge Design

		Time					
	Community	0	1	2	3	4	5
Randomize	1	O	X	X	X	X	X
	2	O	O	X	X	X	X
	3	O	O	O	X	X	X
	4	O	O	O	O	X	X
	5	O	O	O	O	O	X

X=intervention + observation, O=observation

Stepped Wedge Design

- Works well for interventions with anticipated dose effects.
- Like the dynamic waitlisted design, power is increased because every community serves as its own control group.
- Analysis: Mixed effects regression model with random effect for community and fixed effects for time, intervention condition and the interaction.

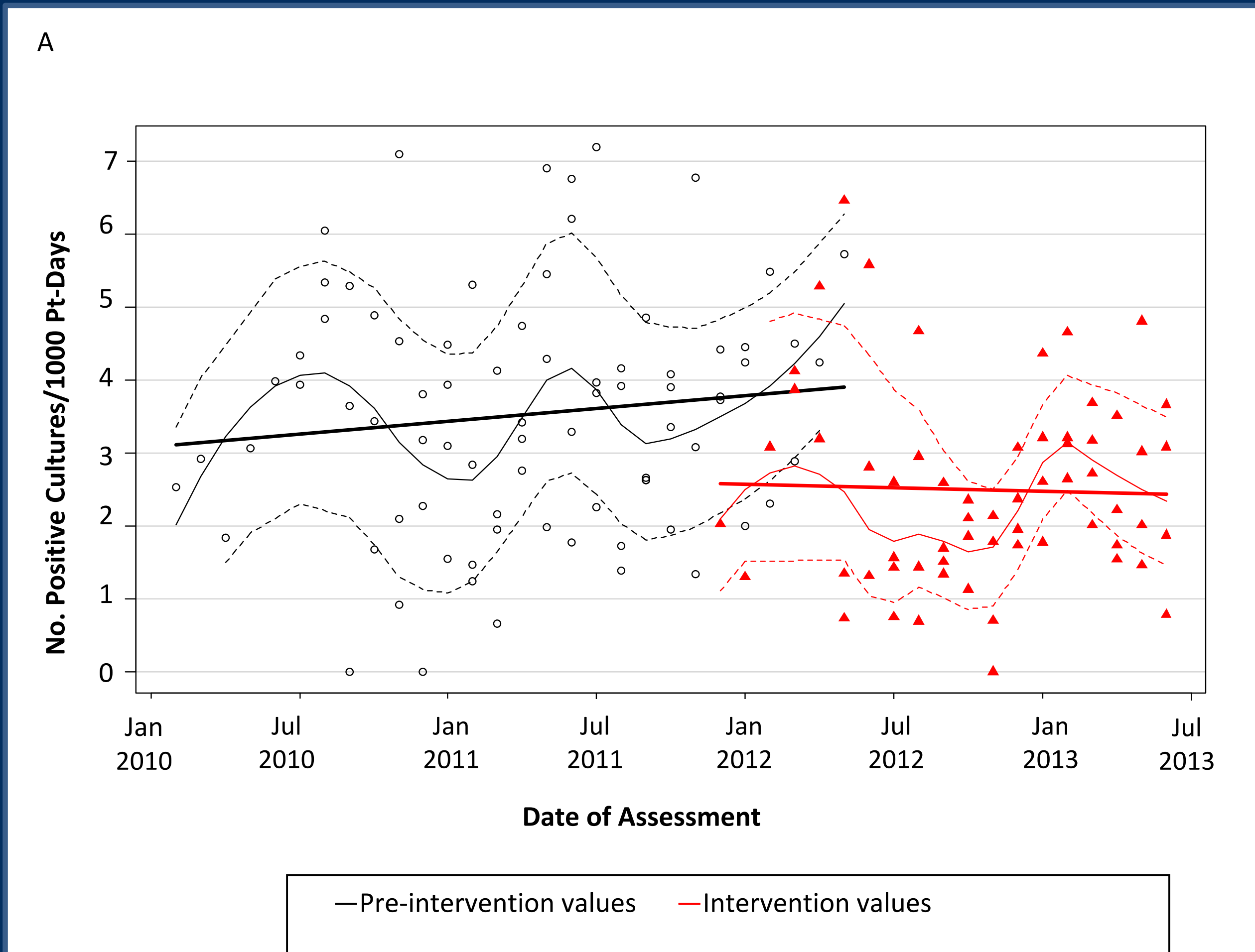
Stepped Wedge Design

Example:

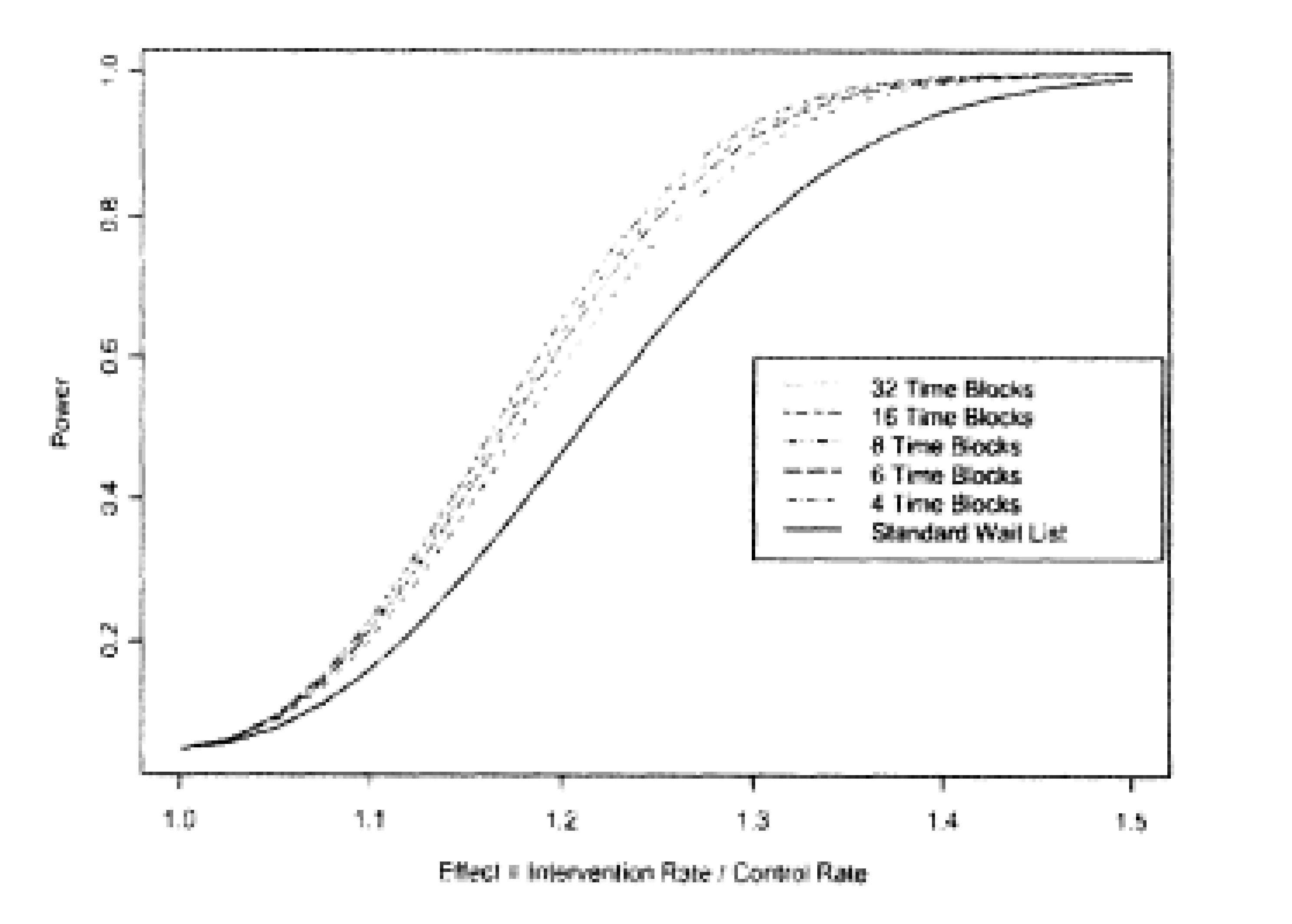
- Intervention to address outbreak of antibiotic resistant bacteria in long-term care (LTC)
- Four LTC hospitals randomly assigned to starting times over a 6 month rollout period
- 42 months of assessment per facility
- 18-20 months of baseline before intervention*

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Stepped Wedge Design: Results



Statistical Power for Roll-out Designs



1. Roll-out Designs

Ambiguous temporal precedence	+ Intervention precedes outcome change
Selection	+ Random assignment of start times
History	+ Affects I and C equally
Maturation	+ Affects I and C equally
Regression	+ Affects I and C equally
Attrition	? I or C may not want to wait
Testing	+ Affects I and C equally
Instrumentation	+ Does not change between I and C
Interactive Effects	+ Should affect I and C equally

2. Interrupted Time Series Design

- Sequence of observations interrupted by intervention.
- Observations before intervention establish baseline levels.
- Observations during the intervention measure immediate effects.
- Observations after the intervention ends measure maintenance of effects.

2. Interrupted Time Series Design

0 0 0 0 0 0 0 0 0 0 0

X → → X

Best for interventions with a defined term where maintenance of effects is a concern.

Also works well with no control and only one intervention unit.

Historical events threaten internal validity.

2. Interrupted Time Series Design

Example 1: Impact of Chicago Center for Youth Violence Prevention (CCYVP) on violence in one targeted Chicago community

Outcome: Number of violent crimes per month per police beat

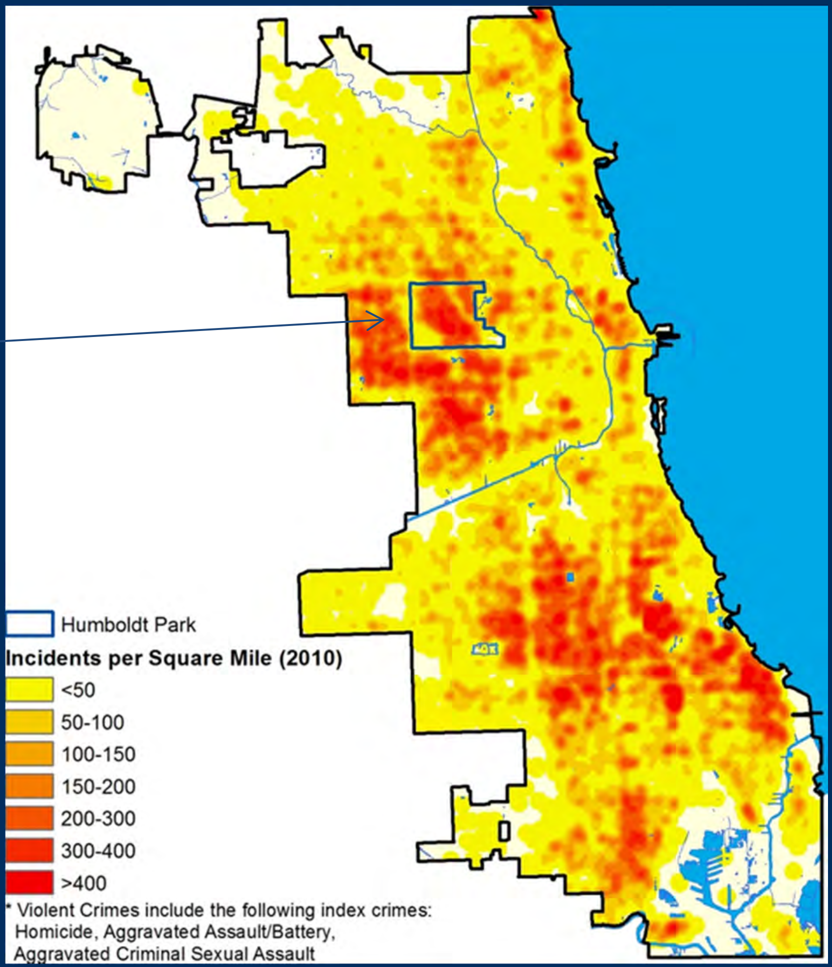
Predictors: Time, period, time*period interaction

Analysis: Poisson linear mixed model with comparison to other communities

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Violent Crime in Chicago 2010

Humboldt Park

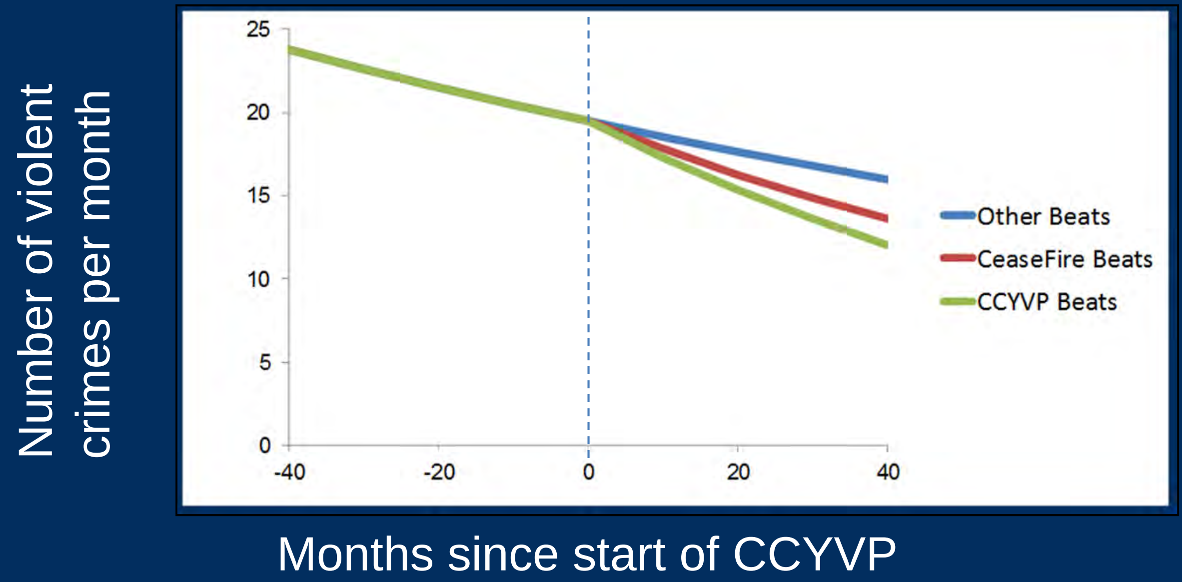


2. Interrupted Time Series Design

3 Years of Intervention

Significant reductions in violent crime

Compared with all other police beats in Chicago, and other police beats where CeaseFire is present



Chicago Police Department
CCYVP: Chicago Center for Youth Violence Prevention

2. Interrupted Time Series Design

Example 2: Impact of a culturally-focused CBPR intervention on 54 Yup'ik Youth from a single community

Outcomes: Reasons for Life and Reflective Processes

Predictors: Time, period, time*period interaction

Analysis: 3 ways of thinking about time

- Intervention dose at each wave
- Chronological time centered when individual began intervention
- Cohort with which the individual began

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

2. Interrupted Time Series Design

Example 2: Results

	Estimate	*SE*	*df*	*t*	95% CI Lower	95% CI Upper	*p*	Effect size (*d*)
Reflective Processes								
Dose	0.015	0.008	119	1.96	0.0005	0.030	.05	.35
Dose X Protective	0.001	0.009	119	0.11	-0.017	0.019	.91	.02
Reasons for Life								
Dose	0.012	0.007	120	1.76	-0.001	0.026	.08	.32
Dose X Protective	0.011	0.009	120	1.25	-0.006	0.028	.21	.23

Interrupted Time Series Design

Threat to internal validity	Performance
Ambiguous temporal precedence	+ Baseline precedes intervention
Selection	+ Units are their own controls
History	? Problem if no comparison
Maturation	? Problem if no comparison
Regression	+ Affects baseline and I phases equally
Attrition	+ Units are their own controls
Testing	+ Affects baseline and I phases equally
Instrumentation	+ Does not change between baseline and I
Interactive Effects	? Interaction of testing and intervention

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

3. Regression Point Displacement Design

- How to evaluate a CBPR intervention conducted in a single community?

Figure 2. The Yukon-Kuskokwim (Y-K) Delta with YKHC medical facilities marked.

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Use population data to estimate expected values.

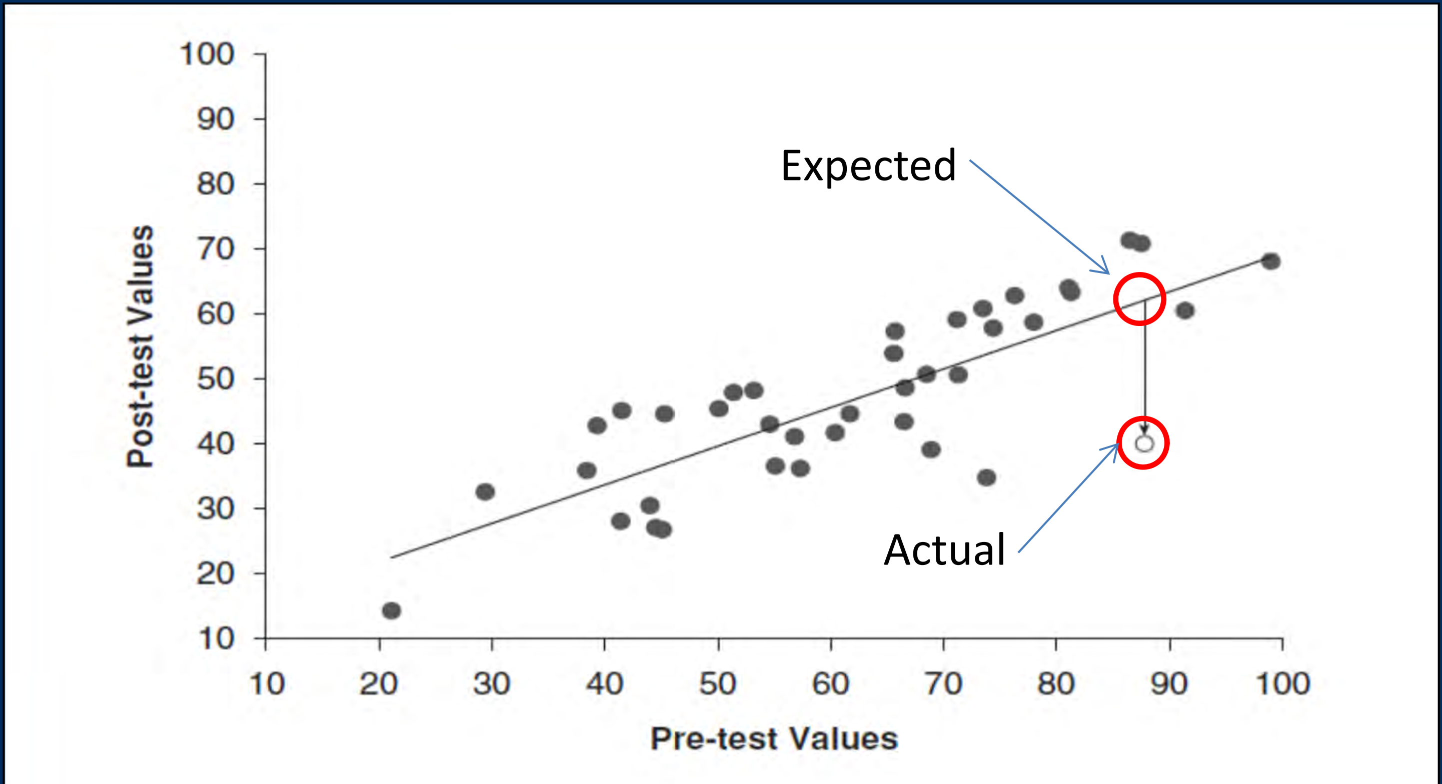


Figure 3. Illustration of the regression point displacement research design (Linden et al., 2006).

3. Regression Point Displacement Design (RPDD)

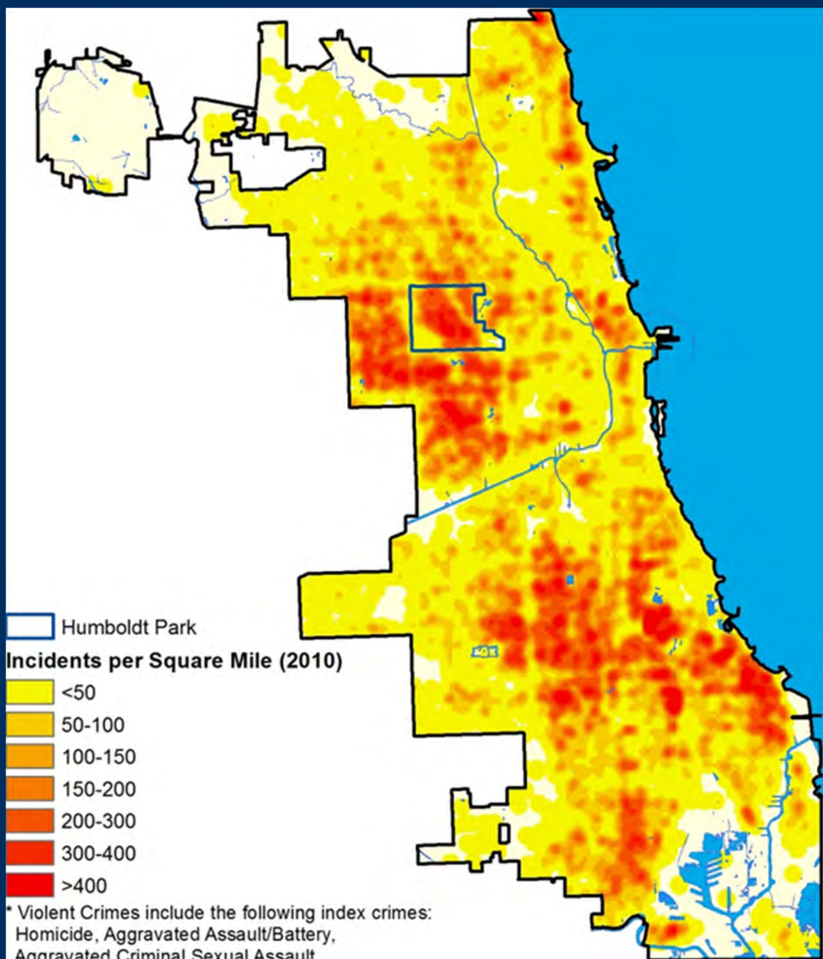
- Variant of the Regression Discontinuity Design
- Helpful when there is only one community or a few communities receiving intervention.
- Uses archival data on the population from which the intervention communities was drawn.
- Helpful if intervention communities were randomly drawn, but not necessary.
- Propensity scores can be used as covariates adjust for non-randomness of selection.

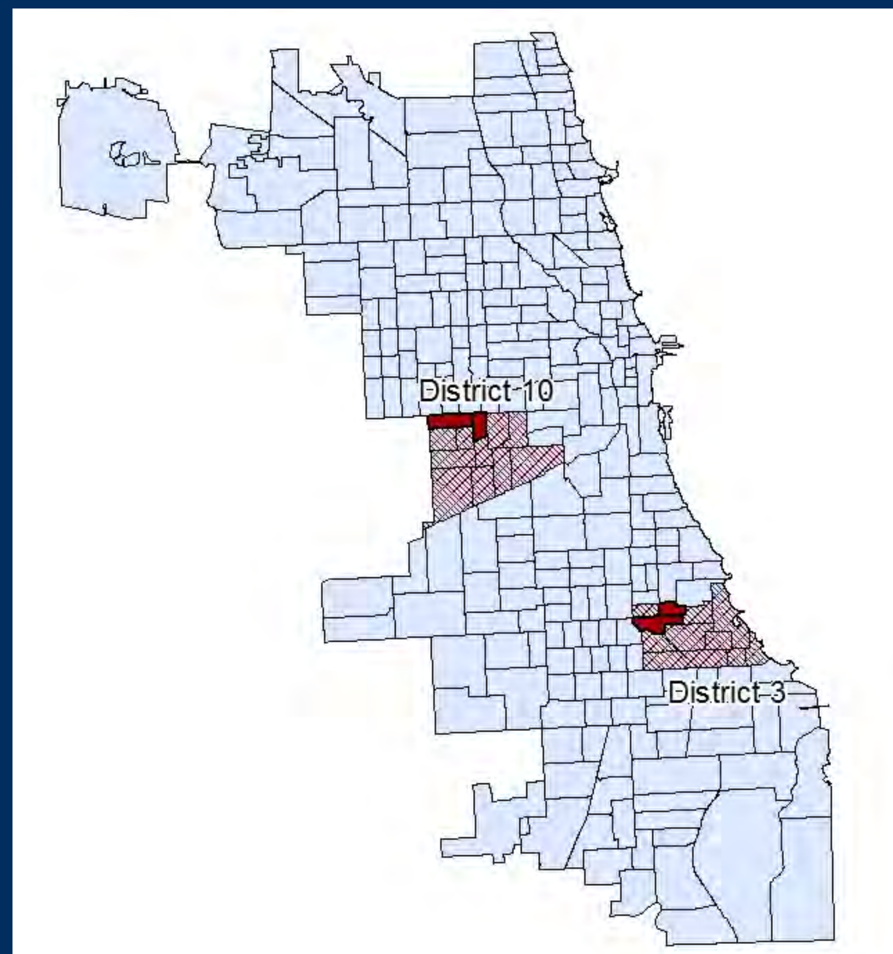
3. Regression Point Displacement Design (RPDD)

- Requirements
  - Archival measures on population prior to and after intervention period
  - Outcome measure that is stable over time: Correlation > .9 preferred.
  - Outcome measure need not be identical at pretest and posttest
  - Random selection of intervention unit, or
  - Information to predict selection for intervention

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Example: Evaluation of CeaseFire



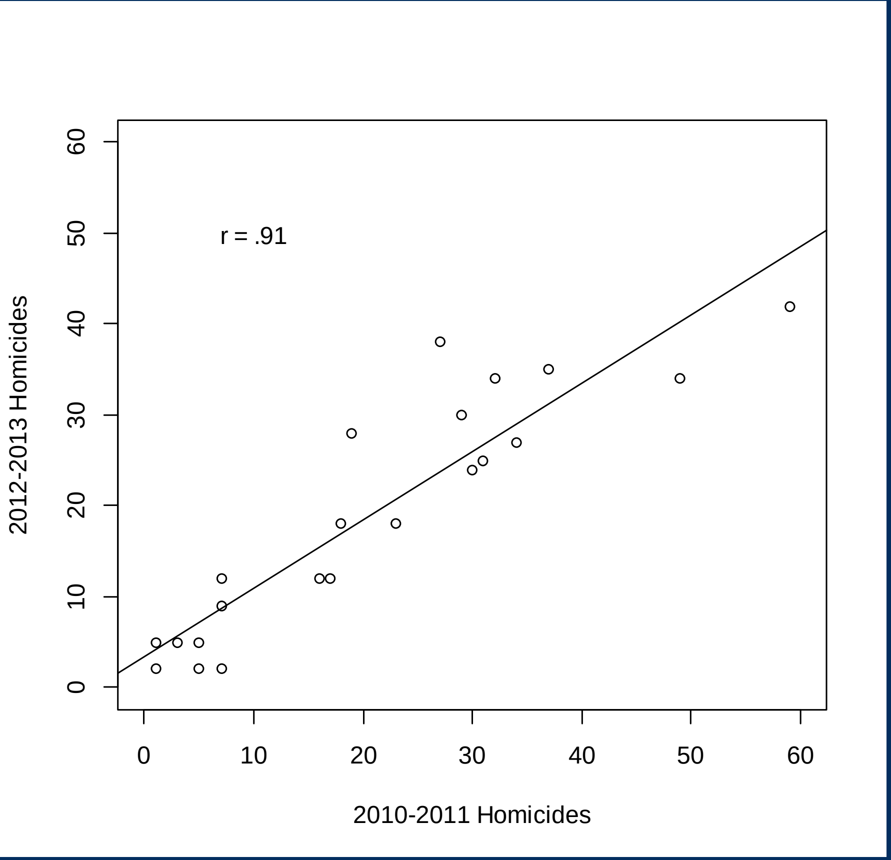


Example: Evaluation of CeaseFire

- Outreach workers and violence interrupters
- Organized by police districts and beats
- Evaluation of 1 year of CeaseFire in two districts
- 2 of 25 police districts received intervention
- Pre-post correlation for homicide = .91 ($p<.01$)
- Predictors of selection: Crime, poverty, political influence

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Homicides in Chicago Police Districts



Example: Evaluation of CeaseFire

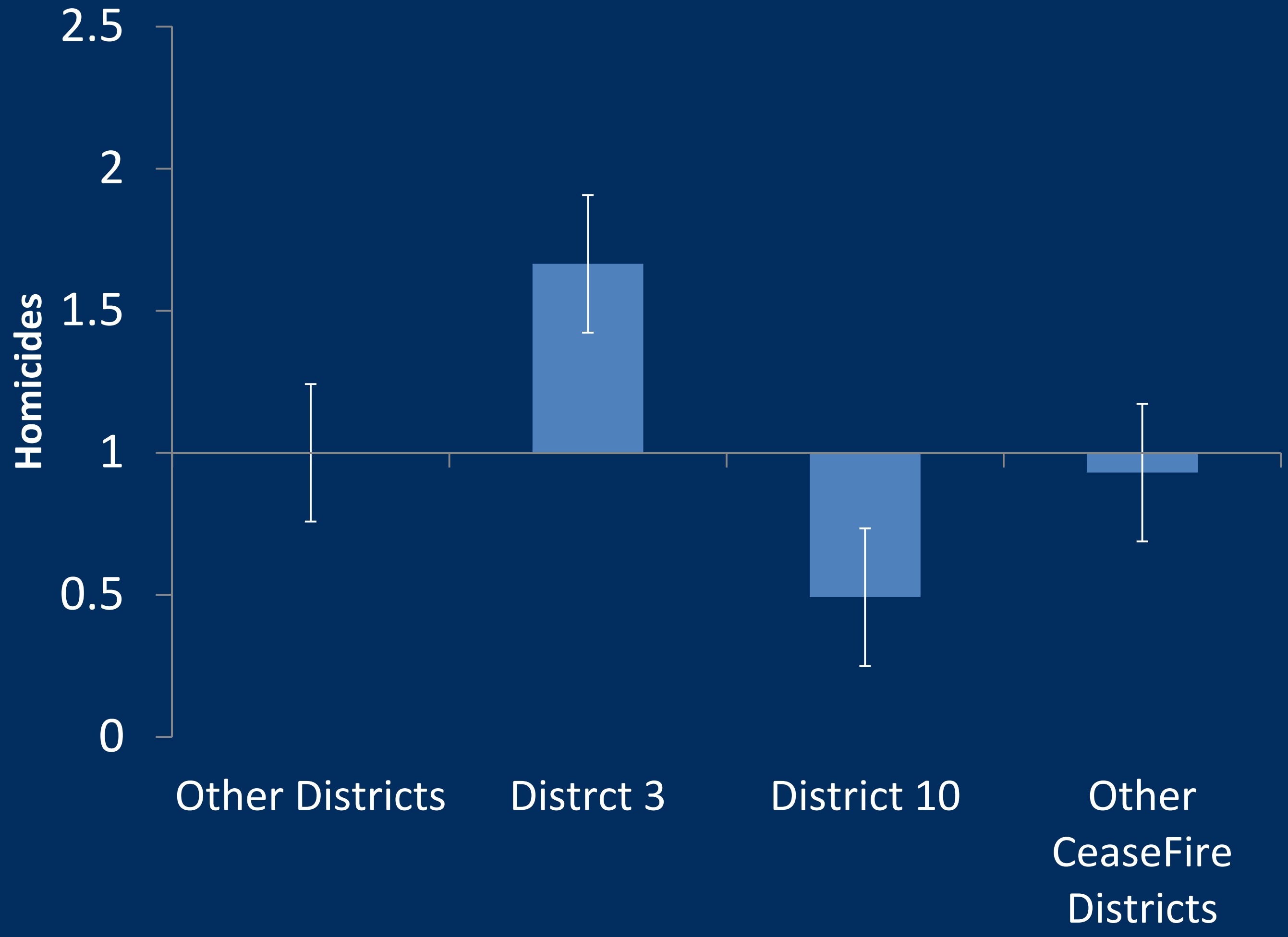
- Analysis
  1. Logistic regression to predict selection of any police beat for CeaseFire
  2. Regression
     - Outcome: 2012-2013 homicides
     - Predictors
       – 2010-2011 homicides
       – Codes for Districts that received intervention
       – Propensity score for selection from Step 1
       – Overall number of police reports*

Example: Evaluation of CeaseFire

- Results:
  - One city-funded CeaseFire district had significantly higher than expected homicides ($B = .54$, $p < .01$) and the other had significantly lower than expected homicides ($B = .54$, $p < .01$) .
  - There was a difference in CeaseFire tactics between the two districts.

UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

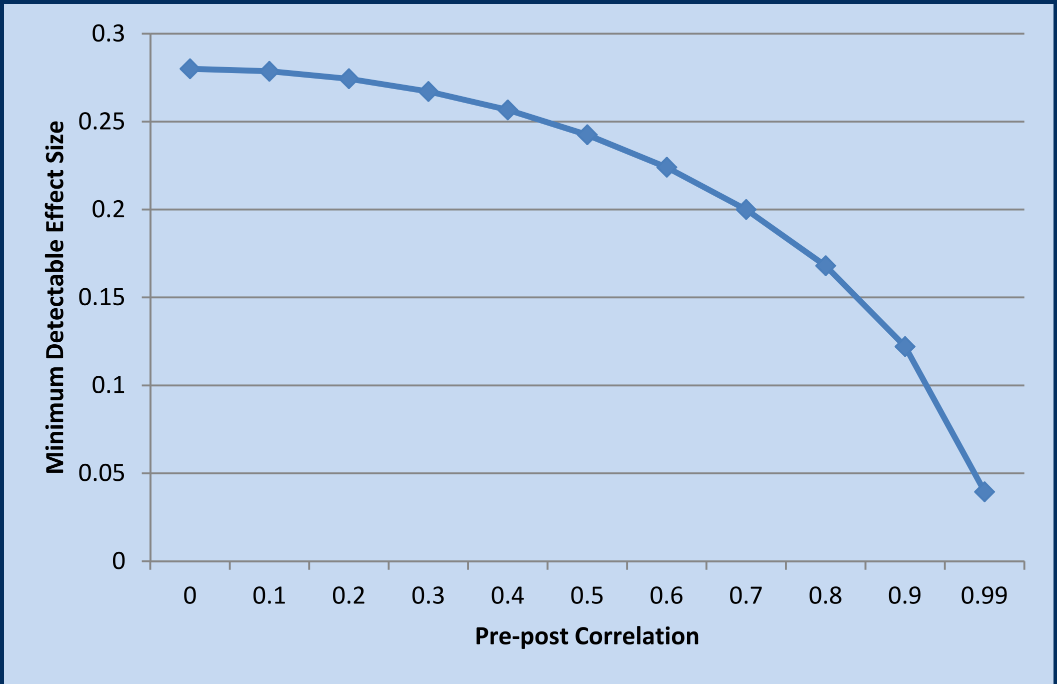
City-funded Districts (1-year of intervention)



UIC INSTITUTE FOR HEALTH RESEARCH AND POLICY

Power for the RPDD

As the pre-post correlation in the population increases, the minimum detectable effect size decreases.



3. Regression Point Displacement Design

Threat to internal validity	Performance
Ambiguous temporal precedence	? Depends on pre-post correlation
Selection	? Selection usually not random
History	+ Applies to entire population
Maturation	+ Applies to entire population
Regression	+ Expected in design
Attrition	? Attrition of Intervention units a problem
Testing	+ Archival records used
Instrumentation	+ Archival records used
Interactive Effects	+ Any that affect archival records

Summary: RCTs and Alternatives

	Ambiguous temporal precedence	Selection	History	Maturation	Regression	Attrition	Testing	Instrumentation	Interactive Effects
RCT	+	?	+	+	+	?	+	+	?
DWLD & SWD	+	+	+	+	+	?	+	+	+
ITSD	+	+	?	?	+	+	+	+	?
RPDD	?	?	+	+	+	?	+	+	+