# CBPR and Experimental Designs

Michele Heisler, MD, MPA
University of Michigan

# Overview

- Challenges to Conducting RCTs in CBPR: Lessons from REACH-Detroit

- Alternative Quasi-Experimental Designs

# Best design for this study question?

Is a community health worker-led diabetes self-management intervention developed and implemented using CBPR effective in improving participants' diabetes clinical outcomes?

# RCT as the gold standard for ‘efficacy’ of interventions

- Highly reliable form of scientific evidence because it accounts for spurious causality

- Treatments are allocated to participants at random; different treatment groups are statistically equivalent

# Experimental Designs

**Experimental design**

**Txt** **O---------X----------O----------O**

**Control** **O--------------------O----------O**

**Delayed control**

**Txt** **O---------X----------O----------O**

**Control** **O--------------------O--------X--------O----------O**

**Enhanced usual care design**

**Txt** **O--------X----------O----------O**

**EUC** **O---------Y----------O----------O**

# Challenges to RCTs in CBPR

- Mistrust from communities for medical research
- Withholding an intervention from individuals who are in need of services
- RCT alone does not provide insights into external validity (RE-AIM; process evaluation)

# Keys to overcoming challenges

- Community ownership and accountability
- Long standing relationship and collaboration
- Extensive planning phase that involves all major stakeholders
- Observed success of intervention during non-randomized, pilot stage
    - Recognition that to sustain need strong evidence of efficacy
- Demonstrated benefits to community with minimal risk

# Designs For When Everyone Needs to Get an Intervention

- Randomize who gets the intervention ***first***
- Randomize two interventions with different outcomes

# Case example: REACH Detroit

- Multi-level, CBPR intervention aimed at eliminating disparities among African Americans and Latinos with diabetes in Detroit.
- Family-level: Community Health Worker (CHW) intervention
- Health system: CMEs, residents program, workshops
- Community-level: Access to physical activity and healthy foods, support groups

# Family Health Advocates Led Six-month Program

- **Journey to Health/El Camino a la Salud:** 11 two-hour, culturally tailored group diabetes self management classes
- **One-on-One Support:** behavioral goal setting and follow-up ("action plans"), social support, linkage to resources
- **Clinic visits:** accompany clients to at least one provider visit, provide help navigating the health care system

# Aim of Intervention to Improve Diabetes Clinical Outcomes

- Improve diabetes self-management skills and behaviors
- Increase physical activity and healthy eating
- Enhance family-provider relationships
- Increase access to community resources
- Improve healthcare navigation skills

# Family Health Advocate Intervention – Design

- **Phase 1: Legacy of distrust in the community →**
  - non-randomized, 1 group with historical control, (n=180)

    *(Heisler et al, Diabetes Care 2006)(Two Feathers et al, AJPH 2006)*

- **Phase 2: Increased community trust →**
  - randomized intervention & delayed intervention control group (pilot with extensive evaluation of process)

# Initial Staff Training

- FHA staff and data collection coordinators provided with training on RCT designs:
  - What is random assignment--“chance”
  - Why it is important to randomize
  - Concept of ‘equipoise’
  - How we would be implementing the RCT
  - Reviewed orientation script and consent forms (6$^{th}$ grade reading level)
  - Role play explaining process and rationale of randomizing (“flip a coin”)

# Delayed intervention procedure

- Participants recruited from health systems and invited to orientation
- Participants informed of random assignment to treatment and delayed group and asked to sign consent form
- Participant given a sealed letter with group assignment
- Group assignment explained and baseline data collected

# Delayed Intervention procedures

- Delayed group called by phone once per month to update contact information
- Six months selected due to our ability to detect change at this time point in previous work.
- Delayed group over-sampled by 10% to account for possible attrition.
- Individuals in need of immediate medical attention were provided services and removed from the study (n=3)

# Recruitment and Retention

- 164 participants: n=77 in immediate group and n=87 in delayed group
- Six month retention 71%
- Drop out rates similar for both groups
- No significant difference in number of baseline interviews completed
- Women in the delayed group dropped out at a higher rate than men
- More Latinos in delayed group dropped out; more African Americans in the immediate group dropped out

# Interviews with 15 Participants

- Delayed group expressed disappointment
  - Viewed immediate group as “winning”
  - Prefer to be in immediate group
- Several in immediate group would have preferred to be in the delayed group
  - See what intervention is like and decide to participate later
- Lack of trust and clarity of RCT process for both groups
  - Questioned whether process was truly random
  - Questioned fairness of process (not first come first served)
  - Those who withdrew did not do so because of the process, but rather due to work schedules and other commitments

# Lack of understanding and clarity of RCT

*"I was disappointed, because I was hoping to get into the (immediate) group… I feel that because I wasn't on insulin--I was on pills--that's the reason why I probably wasn't picked for the group. I don't know."*

*[respondent who withdrew from delayed intervention group]*

# Lack of understanding and clarity of RCT

*"It seems like it [group selection by randomization] was like pulling a rabbit out of a hat…. And I remember thinking 'Oh my gosh'…"*
*[Intervention group]*

*"Well, I think they already had their plans… they already knew what they were doing…. They looked at people and knew what group they were going to assign them to because I mean, to do it by chance – this is not a betting game."*
*[Intervention group]*

# Fairness

*"I explained to them that I would need to be in the first group because of my work schedule… and then I ended up in the second group… I had told two co-workers about it [REACH], and they had ended up joining the first group, and that was really a slap in the face. --They were losing weight and enjoying themselves, but I wasn't able to participate like I wanted to." [withdrew from delayed group]*

*"I got in…and the second class they had to wait six months to get in. The class really wasn't that big, why couldn't we all be in the same class?" [intervention group]*

# Interviews with 4 FHAs and 2 other staff

- Staff still had misgivings about RCT design
  - Not completely convinced that randomization was necessary to show efficacy of program
  - Dissatisfaction with inability to provide help and services to participants in delayed group

*"If you know they need help, you can begin not liking the process of the randomization when you know a client needs help."*

*"Regarding randomization, the immediate group should be the people who most need it"*

# RESULTS--Staff

- Did they ever provide “special” or “secret” services for delayed group?
  - Two staff said yes
  - *“They might’ve needed some extra information or something and I provided it”*
- FHAs wanted to conduct survey interviews (they had done so in the first cohort)
  - Felt like we were restricting their job and imposing on their autonomy
  - *“I’d like to be free to do things with my clients…”*

# Lessons learned

- Successful collaboration to achieve RCT
- RCT successful from research perspective
  - Able to randomize and execute intervention
  - Able to detect differences
  - No increased attrition from delayed group
- Even with detailed training and procedures, maintaining integrity of RCT is a challenge
  - Community mistrust difficult to overcome
  - Retention of understanding is a major issue

# Lessons learned

- When participants bring little or no prior understanding about research methods and design, they may be overwhelmed or confused when explained the process of randomization (Kerr, Robinson, Stevens, et al., 2004)
  - Providing examples not enough, e.g., flipping a coin versus pulling rabbit from a hat
  - Many trials have found participants do not believe assignments are by chance
  - Staff must be brought on board at the very beginning, with periodic assessment and dialogue-- little chance for success otherwise

# Next Step: Enhanced Usual Care Control Group Design with RE-AIM

- Not all participants receive the full intervention
- Only treatment group receives the FHA services (home and clinic visits)
- All participants invited to community-level intervention
  - Activities and events readily available to the general public and free of charge
  - Usual care typically included regular medical services
  - Isolates the unique services of the FHA

# Study Question

What are the costs and benefits of implementing a comprehensive community-level diabetes prevention program?

# Why not an RCT?

- Randomizing participants in activities does not answer the question and is probably not feasible
- Randomizing enough communities is prohibitively expensive

# An RCT to Examine Community Diabetes Prevention Program

Pros:

- Gold standard for efficacy evaluation

Cons:

- Would need to either:
  - Randomize > 20-40 sites, or
  - Randomize within sites

# Aren't Quasi-Experiments (QEs) a Weak Methodology?

- Most published QE's are terrible, but
- Can produce strong evidence

# Inherent Shortcoming of QE's

Can never eliminate the possibility of confounding

(Threat to Internal Validity)

# What is a Quasi-Experiment?

- An intervention is studied but

1. Receipt of the intervention is non-random,

or

2. Randomized units are too few to be analyzed as random effects (<10-20 per intervention group)

# Common Quasi-Experimental Designs

- Pre-Post comparison
- Non-random (non-equivalent) comparison groups
- Time-series designs
- Preference allocation

# Simple Pre-Post Design

**Intervention Group:**

# Non- randomized (nonequivalent) control groups

**Intervention Group**

**O → X ⟶ O**

**O → X ⟶ O**

**Comparison Group**

**O ⟶ O**

**O ⟶ O**

# Most Significant Rival Hypotheses

- What else might explain improvement in the treatment group more than control group?

# Selection-Maturation Effect

- Selecting one of two groups in way that its participants develop faster
- To eliminate, many researchers try to match participants across the groups—but have to do this very carefully!
- Regression to the mean can increase differences between two groups on the matched factor at post-test

# Local History Effect

- When some event affects either the treatment or the control group but not both

# Regression to the mean

- Predicts increase if treatment group selected because of unusually poor scores

- Provide more than one assessment to evaluate this possible artifact

# When Will Regression towards the Mean Occur

Whenever:

1) Baseline measures are imperfectly correlated with follow-up measures (in the absence of intervention),

AND

2) You preferentially sample subjects who are above or below the mean.

# How to Prevent Regression towards the Mean Biasing Results

- Do not sample based upon baseline values (you do not need to intervene on the entire sample, however)
- Use a control group, or
- Use time-series analyses (increase baseline replicates)

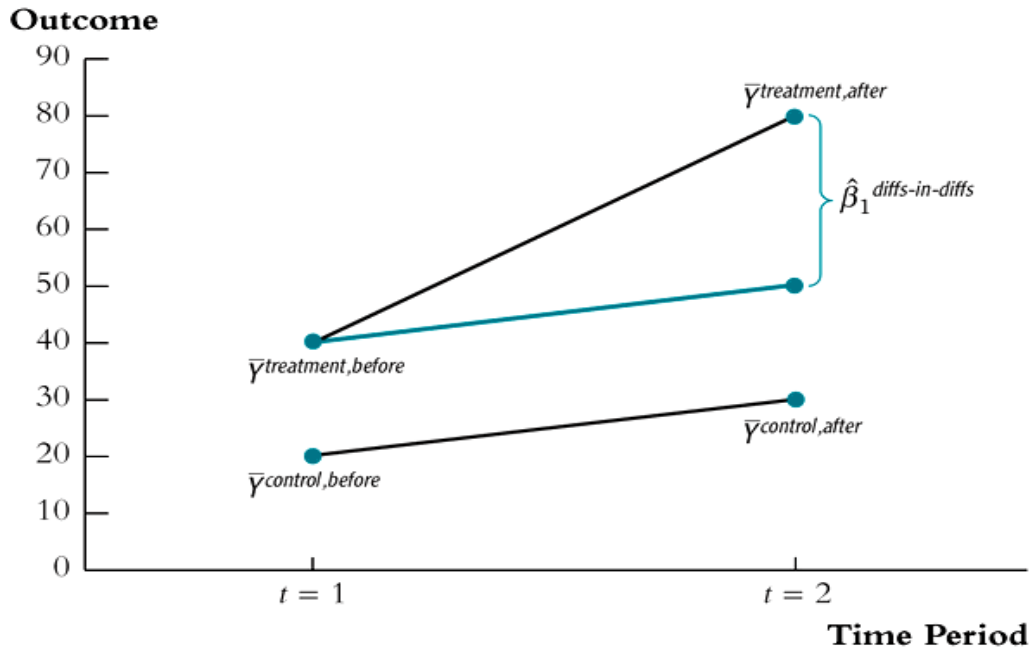
# Precision of Baseline Measurement is Critical

- Replicate measures to estimate reliability
- Measurements over time to estimate temporal stability or trends

# Difference-in-Differences Estimation

**FIGURE 11.1 The Differences-in-Differences Estimator**

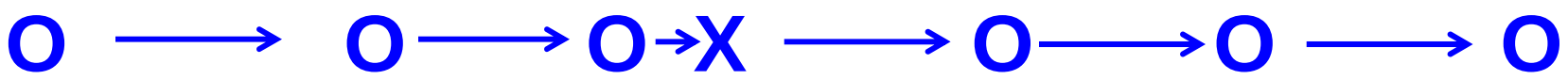
The post-treatment difference between the treatment and control groups is 80 – 30 = 50, but this overstates the treatment effect because before the treatment $Y$ was higher for the treatment than the control group by 40 – 20 = 20. The differences-in-differences estimator is the difference between the final and initial gaps, so that $\hat{\beta}_1^{\text{diffs-in-diffs}} = (80 - 30) - (40 - 20) = 50 - 20 = 30$. Equivalently, the differences-in-differences estimator is the average change for the treatment group minus the average change for the control group, that is, $\hat{\beta}_1^{\text{diffs-in-diffs}} = \Delta\bar{Y}^{\text{treatment}} - \Delta\bar{Y}^{\text{control}} = (80 - 40) - (30 - 20) = 30$.



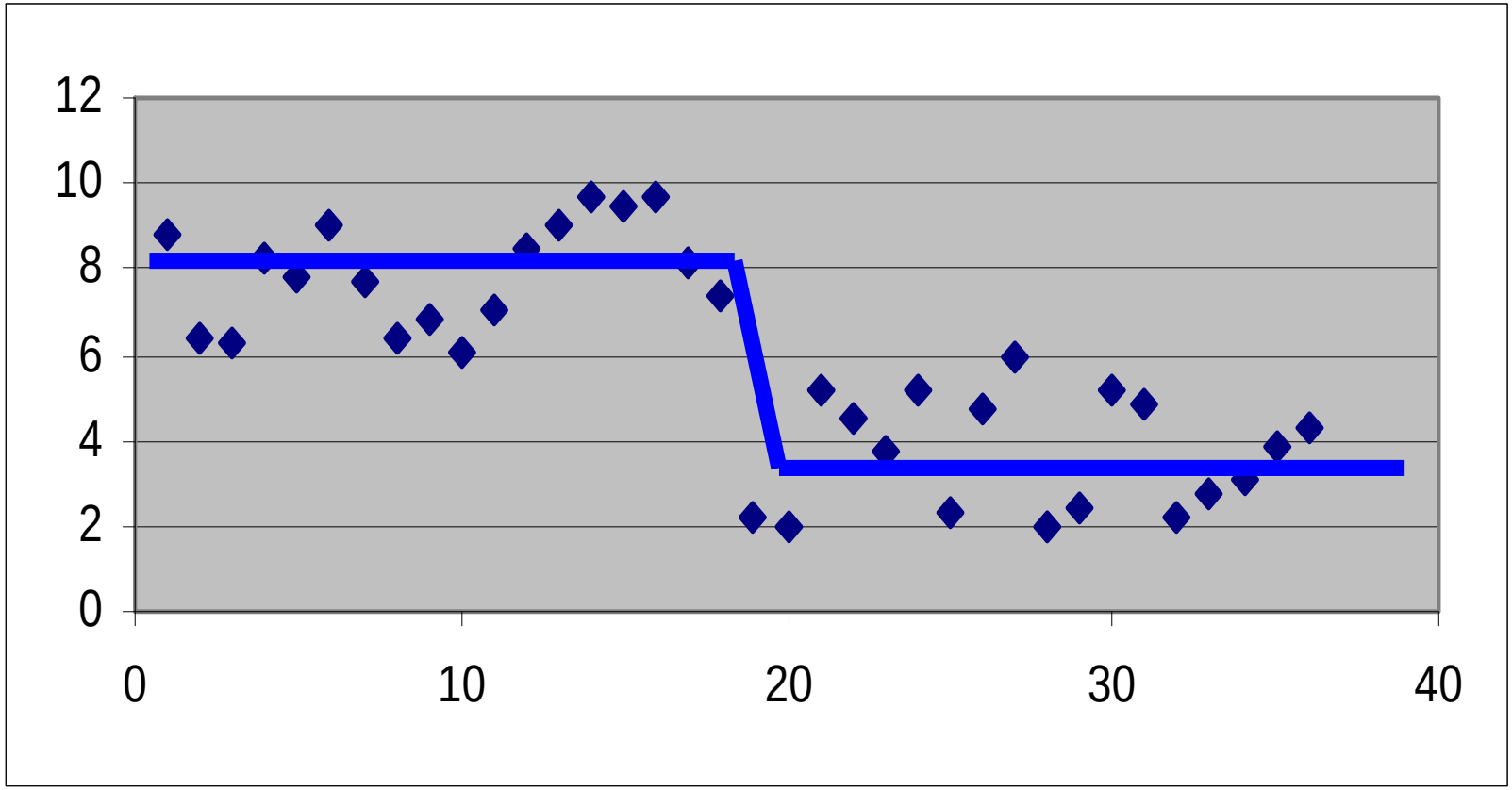

JH Stock & MW Watson. Introduction to Econometrics

# Interrupted Time Series Analysis

**Intervention Group:**



# Time-series trial



# Transfer & Impulse-Response Functions
## (Box-Jenkins, dynamic regression, etc)

- Examines for 2 things
  - Abrupt change
  - Change in slope of temporal changes

# Key Weakness of Single Group Interrupted Time Series Design

- Design fails to control for effects of history
- If an extraneous event occurs at about the same time as treatment, could change behavior
- You must consider all other events taking place at same time as treatment and whether could be rival hypotheses

# Multiple Time-Series Design

**Intervention Group:**

**O O O X O O O**

**Control Group:**

**O O O O O O**

# A Couple of Other QE Designs

- **Cross-Lagged Designs**

  (Tries to infer causality by examining temporal correlations between two variables)

- **Regression-Discontinuity Design**

  (Allows statistically robust design even if you preferentially offer the intervention to outliers)

# Regression-Discontinuity Design

- Allows more statistically accurate assessment of studies that target interventions preferentially to those who are most likely to need (and benefit) from the intervention

- Key issue = Measure all subjects pre and post, but preferentially intervene on those with greater need for the intervention

# Two Ways to Enhance Internal Validity

- Qualitative Evaluations for:
  - Evaluations of confounders
  - Examination of pathways

- Quantitative evaluation of pathways (intervening variables)

# Other Key Partnership Issues

- Your community partners will usually be evaluated on tangible results
- Unhappy constituents can get them in trouble
- They are often overwhelmed with other tasks

# Basic Design Issues

- Unit of intervention allocation (individual, block, neighborhood?)
- Pre-Post, Time-series, Control sites?
- Baseline data?
- Intervening variable?
- Qualitative Evaluations?
- Partnership Issues?

# Steps in conducting QEs

- Carefully evaluate your design
- Carefully evaluate the data you will collect
- Carefully evaluate your data after you collect (and patterns of results)
- Be aware of potential sources of threats to internal validity
- Determine how you will assess potential rival hypotheses to observed pattern
- Be careful but creative with your design!

# Key Points

- QE's are often the optimal design for answering policy-relevant effectiveness questions

- QE's can be very strong designs, especially if:
  - Combine time-series & NR comparison group designs
  - Qualitative examination of micro-environment
  - Examine pathways (intervening factors)

For more information contact:


***Michele Heisler***
***mheisler@umich.edu***

***Gloria Palmisano***
***Project Manager***

**gpalmisano@chasscenter.org**

Or visit our website:

**www.reachdetroit.org**

Funded by:
Centers for Disease Control and Prevention
Cooperative Agreement No. U50/CCU417409, NIDDK, and
The Michigan Diabetes Research and Training Center

# How Big Can Regression-to-the-Mean Be?

$$P_{rm} = 100\ (1\text{-}r)$$

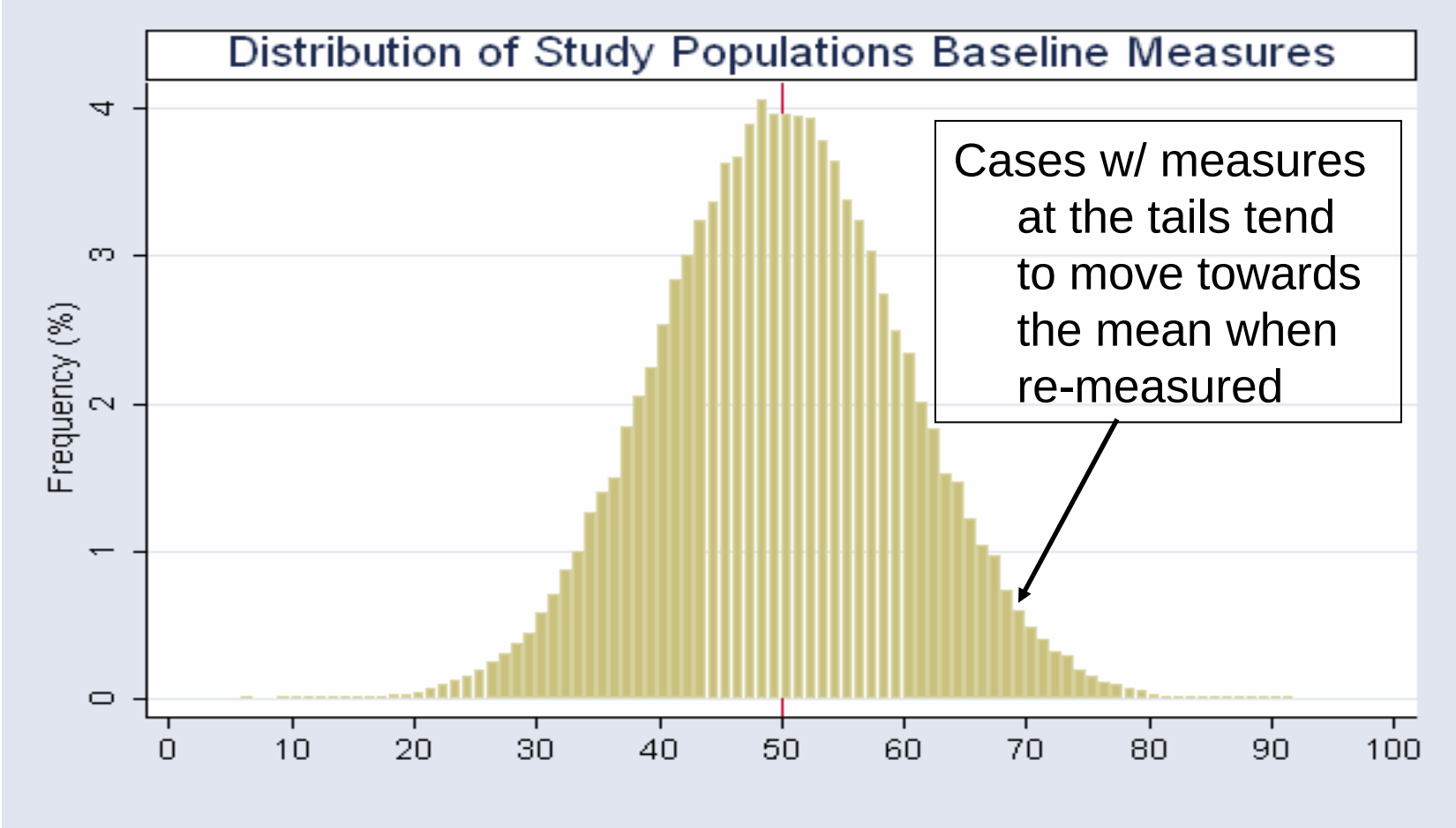
$P_{rm}$ = % regression towards mean

r = correlation of pre-post measures

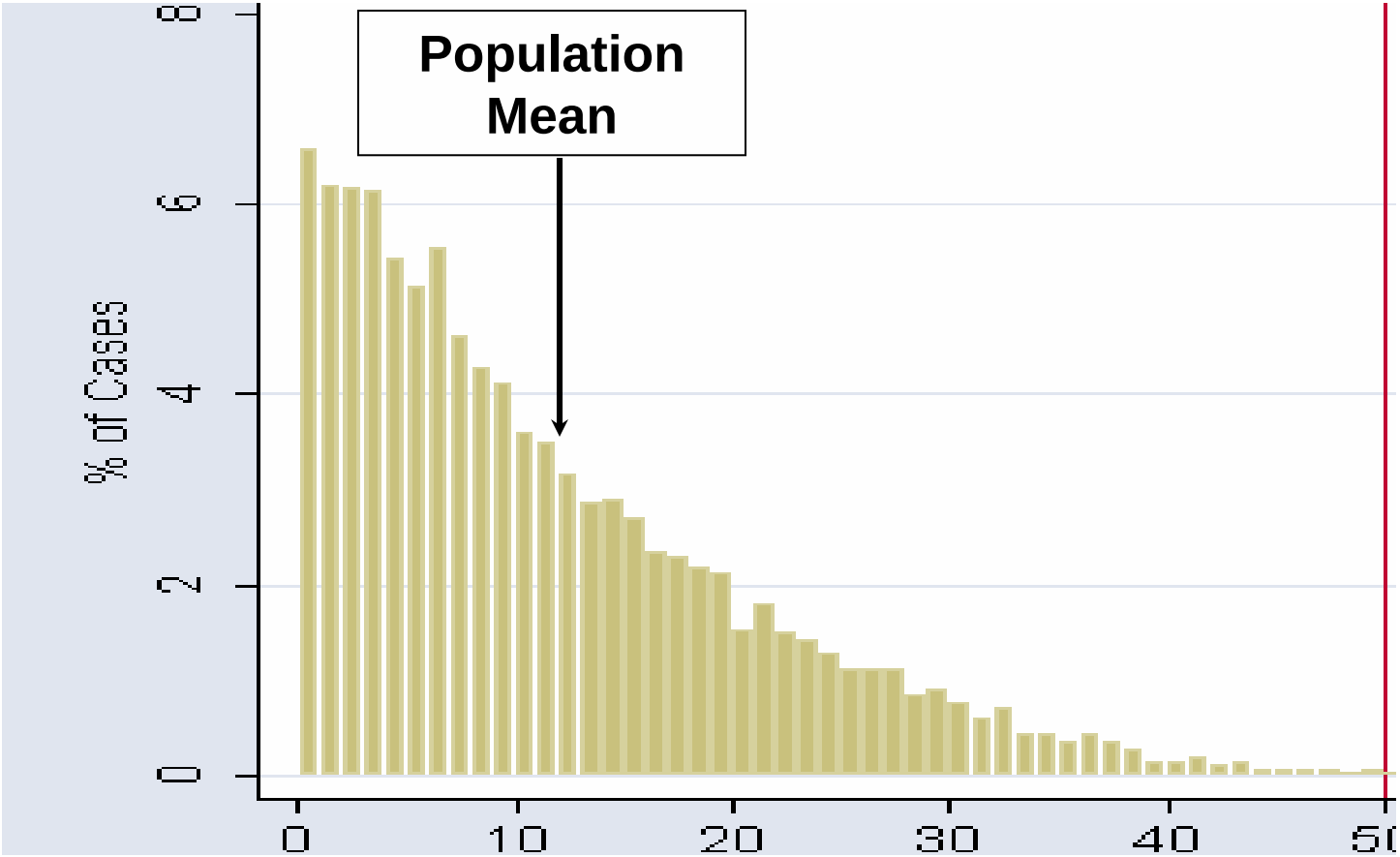
# An Evaluation of An Intervention (Pre-Post)

- Select 1000 people with A1c > 9%
- At 6 months, A1c decreased by 1.5 points ($p < 0.001$)

# Regression Towards the Mean



# Regression Towards the Mean



# Why Does Regression towards the Mean Occur

- Imprecise measurement leads to both misclassification errors and over-estimation of population variance

or

- Averaging over time decreases population variance

# How Big Can Regression to the Mean Be?

$$P_{rm} = 100\ (1\text{-}r)$$

$P_{rm}$ = % regression towards mean

r = correlation of pre-post measures

If r = 1, then no regression

If r =0.5, then 50% regression towards mean

If r = 0, then complete regression