Pragmatic and Group-Randomized Trials in Public Health and Medicine

Part 7: Alternative Designs

David M. Murray, Ph.D.
Associate Director for Prevention
Director, Office of Disease Prevention
National Institutes of Health

A free, 7-part, self-paced, online course from NIH with instructional slide sets, readings, and guided activities
Target Audience

- Faculty, post-doctoral fellows, and graduate students interested in learning more about the design and analysis of group-randomized trials.
- Program directors, program officers, and scientific review officers at the NIH interested in learning more about the design and analysis of group-randomized trials.
- Participants should be familiar with the design and analysis of individually randomized trials (RCTs).
  - Participants should be familiar with the concepts of internal and statistical validity, their threats, and their defenses.
  - Participants should be familiar with linear regression, analysis of variance and covariance, and logistic regression.
Learning Objectives

And the end of the course, participants will be able to...

- Discuss the distinguishing features of group-randomized trials (GRTs), individually randomized group-treatment trials (IRGTs), and individually randomized trials (RCTs).
- Discuss their appropriate uses in public health and medicine.
- For GRTs and IRGTs...
  - Discuss the major threats to internal validity and their defenses.
  - Discuss the major threats to statistical validity and their defenses.
  - Discuss the strengths and weaknesses of design alternatives.
  - Discuss the strengths and weaknesses of analytic alternatives.
  - Perform sample size calculations for a simple GRT.
- Discuss the advantages and disadvantages of alternatives to GRTs for the evaluation of multi-level interventions.
Organization of the Course

- Part 1: Introduction and Overview
- Part 2: Designing the Trial
- Part 3: Analysis Approaches
- Part 4: Power and Sample Size
- Part 5: Examples
- Part 6: Review of Recent Practices
- Part 7: Alternative Designs and References
What About Alternative Designs?

- Many alternatives to GRTs have been proposed.
  - Multiple baseline designs
  - Time series designs
  - Quasi-experimental designs
  - Dynamic wait-list or stepped-wedge designs
  - Regression discontinuity designs

- Murray et al. (2010) compared these alternatives to GRTs for power and cost in terms of sample size and time.


Multiple Baseline Designs

- Intervention introduced into groups one by one on a staggered schedule
  - Measurement in all groups with each new entry.
  - Often used with just a few groups, e.g., 3-4 groups.
  - Data examined for changes associated with the intervention.
Multiple Baseline Designs

Figure 1. Hypothetical example of a multiple baseline design used to assess behavior change following an intervention in four communities.
Multiple Baseline Designs

- Evaluation relies on logic rather than statistical evidence.
  - Replication of the pattern in each group, coupled with the absence of such changes otherwise, is taken as evidence of an intervention effect.
  - With just a few groups, there is little power for a valid analysis.
- Good choice if effects are expected to be large and rapid.
- Poor choice if effects are expected to be small or gradual.
- Very poor choice if the intervention effect is expected to be inconsistent across groups.

Time Series Designs

- Often used to evaluate a policy change in a single group.
- Require repeated and reliable measurements.
  - Standard methods require ~50 observations before and again after the intervention.
- Rely on a combination of logic and statistical evidence.
  - Standard methods provide evidence for change in a single group.
  - One-group designs provide no statistical evidence for between-group comparisons.
- Best used in with an archival data collection system.
  - Could be a strong approach with archival data on many groups.
- May require several cycles of data.
Quasi-Experimental Designs

- QEs have all the features of experiments except randomization.
  - Causal inference requires elimination of plausible alternatives.
- If groups are assigned and members are observed, analysis and power issues are the same as in GRTs.
- Useful when randomization is not possible.
  - Can provide experience with recruitment, measurement, intervention.
  - Can provide evidence of treatment effects if executed properly.
- Well-designed and analyzed QEs are usually more difficult and more expensive than well-designed and analyzed GRTs.
- cf. Shadish et al. (2000).
Stepped-Wedge Designs

- Sometimes called Dynamic Wait-List Designs
- Combine the features of multiple baseline designs and GRTs.
  - Measurement is frequent and on the same schedule in all groups.
  - Time is divided into intervals.
  - Groups selected at random for the intervention in each interval.
  - By the end of the study, all the groups have the intervention.
- Both *Trials* (2015) and the *Journal of Clinical Epidemiology* (2013) recently published issues focused on the design and analysis of stepped wedge designs.

- See also Hughes JP, Granston TS, Heagerty PJ. Current issues in the design and analysis of stepped wedge trials. *Contemporary Clinical Trials*. 2015;45(Pt A):55-60. PMC4639463.
Stepped Wedge Design

- **Control Condition**
- **Intervention Condition**
Stepped Wedge Design

- The analysis estimates a weighted average intervention effect across the intervals.
  - Assumes that the intervention effect is rapid and lasting.
  - Not very sensitive to intervention effects that develop gradually or fade over time.
- These designs can be more efficient but usually take longer to complete and cost more than the standard GRT (Rhoda, 2011).

Regression Discontinuity Designs

- Individuals are assigned to conditions based on a score, often reflecting the need for the intervention (Shadish et al., 2002).
- The analysis models the relationship between the assignment variable and the outcome.
  - The difference in intercepts at the cutoff is the intervention effect.
- Several recent papers have focused on regression discontinuity designs in public health and medicine (Moscoe et al., 2015; Bor et al., 2015).

Regression Discontinuity Design

Figure 1. Hypothetical regression discontinuity experiments: (a) ineffective treatment and (b) effective treatment.
Regression Discontinuity Design

- Because assignment is fully explained by the assignment variable, proper modeling supports causal inference (Rubin, 1977).
- RDs avoid randomization, but are as valid as a RCT or GRT.
- RDs are less efficient than the standard RCT or GRT, often requiring twice as many participants.
- RDs can be used in the context of GRTs (Pennell, et al., 2011).

A GRT remains the best comparative design available whenever the investigator wants to evaluate an intervention that...
- operates at a group level
- manipulates the social or physical environment
- cannot be delivered to individuals

GRTs provide better or equal quality evidence and are either more efficient or take less time than the alternatives.

Even so, GRTs are more challenging than the usual RCT.
- IRGTs present many of the same issues found in GRTs.
- Investigators new to GRTs and IRGTs should collaborate with more experienced colleagues, especially experienced biostatisticians.
Summary

- Many alternatives to GRTs have been proposed.
  - Multiple baseline designs
  - Time series designs
  - Quasi-experimental designs
  - Dynamic wait-list or stepped-wedge designs
  - Regression discontinuity designs
- Under the right conditions, these alternatives can provide good evidence for causal inference.
  - Some rely on logic more than statistical evidence.
    - Multiple baseline designs, time-series designs
  - Others require studies as large or larger than GRTs and may take longer to complete
    - Quasi-experimental designs, stepped wedge, regression discontinuity
Pragmatic and Group-Randomized Trials in Public Health and Medicine

Visit https://prevention.nih.gov/grt to:

- Provide feedback on this series
- Download the slides, references, and suggested activities
- View the any of the modules in this series

Send questions to:
GRT@mail.nih.gov