Pragmatic and Group-Randomized Trials in Public Health and Medicine

Part 1: Introduction and Overview

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
Three Kinds of Randomized Trials

- Individually Randomized Clinical Trials (RCTs)
  - Individuals randomized to study conditions with no connection among participants after randomization.
  - Most surgical and drug trials, some behavioral trials

- Individually Randomized Group Treatment Trials (IRGTs)
  - Individuals randomized to study conditions with some connection among participants after randomization.
  - Many behavioral trials

- Group-Randomized Trials (GRTs)
  - Groups randomized to study conditions with some connection among participants before and after randomization.
  - Many trials conducted in communities, worksites, schools, etc.
Distinguishing Characteristics

- **Group-randomized trials**
  - The unit of assignment is an identifiable group.
  - Different groups are allocated to each condition.
  - The units of observation are members of the groups.
  - The number of groups allocated to each condition is usually limited.

- **Individually randomized group-treatment trials**
  - The unit of assignment is the individual participant.
  - Participants receive some of their treatment in physical or virtual groups or through a common change agent.
  - The number of groups or change agents is usually limited.
Alternative Labels

- Group-randomized trials are also called...
  - Cluster-randomized trials.
  - They are sometimes called community trials.
  - These labels are interchangeable.

- Individually randomized clinical trials are also called....
  - Randomized clinical trials,
  - Randomized controlled trials,
  - Controlled clinical trials.
  - These labels are interchangeable.
Pragmatic Trials

- GRTs are often used for pragmatic trials.
  - Pragmatic and explanatory trials were first described by Schwartz & Lellouch (1967).
    - Explanatory trials test causal research hypotheses.
    - Pragmatic trials help users choose between options for care.
  - Similar to efficacy and effectiveness trials (Cochrane, 1971).
    - Efficacy trials evaluate an intervention under controlled conditions.
    - Effectiveness trials evaluate an intervention under real-world conditions.

Examples

- Group-randomized trials: Health Care Systems Collaboratory
  - 9 pragmatic trials conducted in collaboration with health care systems, funded as UH2/UH3 trials by a variety of NIH ICs.
  - 8 are group-randomized trials.
    - Hospital acquired infections
    - CRC screening
    - Healthcare utilization in back pain care
    - Chronic pain management
    - Mortality in dialysis patients
    - Management of PTSD in trauma patients
    - Advanced care planning in nursing homes
    - Management of multiple chronic conditions
Examples

- **Group-randomized trials: Health Care Systems Collaboratory**
  - **Overview papers**
Examples

- Individually randomized group treatment trials: Childhood Obesity Prevention and Treatment Research (COPTR)
  - 4 trials funded by NHLBI as U01s
  - Two prevention studies targeting young children
  - Two treatment studies targeting youth
  - All involve substantial participant interaction post-randomization
- Overview paper
Notation

- Following Murray (1998)
  - Dependent variable (Y)
  - Condition, \( C_l \) (\( l=1\ldots c \)), will identify the study conditions
  - Time, \( T_j \) (\( j=1\ldots t \)), will identify the measurement occasion
  - Group, \( G_k \) (\( k=1\ldots g \)), will identify the unit of assignment
  - Member, \( M_i \) (\( i=1\ldots m \)), will identify the unit of observation
  - Covariate, \( X_o \) (\( o=1\ldots x \)), will identify covariates
  - Random effects will be **BOLD**, fixed effects will be **PLAIN**

Impact on the Design

- **Randomized clinical trials**
  - There is usually good opportunity for randomization to distribute potential confounders evenly, as most RCTs have $N>100$.
  - If well executed, confounding is not usually a concern.

- **Individually randomized group treatment trials**
  - There may be less opportunity for randomization to distribute potential confounders evenly, as most IRGTs have $N<100$.
  - Confounding can be more of a concern in IRGTs than in RCTs.

- **Group-randomized trials**
  - GRTs often involve a limited number of groups, often $<50$.
  - There may be limited opportunity for randomization to distribute potential confounders evenly.
  - Confounding is usually a concern in GRTs if $G$ is $<50$. 
Observations on randomized individuals who do not interact are independent and are analyzed with standard methods.

The members of the same group in a GRT will share some physical, geographic, social or other connection.

The members of groups created for an IRGT will develop similar connections.

Those connections will create a positive intraclass correlation that reflects extra variation attributable to the group.

\[ ICC_{m:g:c} = corr\left(y_{i:k:l}, y_{i':k:l}\right) \]

The positive ICC reduces the variation among the members of the same group so the within-group variance is:

\[ \sigma^2_c = \sigma^2_y \left(1 - ICC_{m:g:c}\right) \]
The between-group component is the one's complement:

$$\sigma^2_{g:c} = \sigma^2_y \left( \text{ICC}_{m:g:c} \right)$$

The total variance is the sum of the two components:

$$\sigma^2_y = \sigma^2_e + \sigma^2_{g:c}$$

The intraclass correlation (ICC) is the fraction of the total variation in the data attributable to the unit of assignment:

$$\text{ICC}_{m:g:c} = \frac{\sigma^2_{g:c}}{\sigma^2_e + \sigma^2_{g:c}}$$
Impact on the Analysis in a GRT

- Given $m$ members in each of $g$ groups...

- When group membership is established by random assignment,

\[ \sigma_{y_g}^2 = \frac{\sigma_y^2}{m} \]

- When group membership is not established by random assignment,

\[ \sigma_{y_g}^2 = \frac{\sigma_c^2}{m} + \sigma_g^2 \]

- Or equivalently,

\[ \sigma_{y_g}^2 = \frac{\sigma_y^2}{m} \left(1 + (m - 1) \text{ICC}\right) \]
Impact on the Analysis

- Nested factors must be random effects (Zucker, 1990).
- The variance of any group-level statistic will be larger.
- The df to estimate the group-level component of variance will be based on the number of groups, and so is often limited.
  - This is almost always true in a GRT, can be true in an IRGT.
- Any analysis that ignores the extra variation or the limited df will have a Type I error rate that is inflated, often badly.
  - Type I error rate may be 30-50% in a GRT, even with small ICC
  - Type I error rate may be 15-25% in an IRGT, even with small ICC
- Extra variation and limited df always reduce power.

Impact on the Analysis

Scott & Holt (1982) estimate the effect of the ICC as:

\[
\text{DEFF} = 1 + \left( m - 1 \right) \frac{\text{ICC}_y}{\text{ICC}_x}
\]

- DEFF is the ratio of the variance as observed to the variance under simple random sampling.
- \(\text{ICC}_y\) is the ICC for the dependent variable.
- \(\text{ICC}_x\) is the ICC for the independent variable.

Impact on the Analysis

- For most health related outcomes, ICC values are ...
  - 0.00-0.05 for large aggregates (e.g., schools, worksites),
  - 0.05-0.25 for small aggregates (e.g., classrooms, departments),
  - 0.25-0.75 for very small aggregates (e.g., families, spouse pairs).
- ICCs tend to be larger for knowledge and attitudes, smaller for behaviors, and smaller still for physiologic measures.
- If the groups are crossed with the levels of the exposure of interest (most observational studies), $\text{ICC}_x \approx \text{ICC}_y$.
- If the groups are nested within the levels of the exposure of interest (IRGTs, GRTs), $\text{ICC}_x = 1$, because all members of a group will have the same value for exposure.
Impact on the Analysis

- Given the ICC and m per group, DEFF is...

<table>
<thead>
<tr>
<th>Surveys</th>
<th>IRGTs</th>
<th>GRTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC&lt;sub&gt;y&lt;/sub&gt;=ICC&lt;sub&gt;x&lt;/sub&gt;</td>
<td>ICC&lt;sub&gt;x&lt;/sub&gt;=1</td>
</tr>
<tr>
<td>m</td>
<td>0.05</td>
<td>0.25</td>
</tr>
<tr>
<td>50</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>100</td>
<td>1.12</td>
<td>3.25</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.90</td>
</tr>
<tr>
<td>200</td>
<td>1.25</td>
<td>5.75</td>
</tr>
<tr>
<td>1.01</td>
<td>2.90</td>
<td>100</td>
</tr>
<tr>
<td>200</td>
<td>1.50</td>
<td>10.75</td>
</tr>
<tr>
<td>1.02</td>
<td>4.90</td>
<td>500</td>
</tr>
</tbody>
</table>

- The usual F-test, corrected for the ICC, is:

\[ F_{\text{corrected}} = \frac{F_{\text{uncorrected}}}{DEFF} \]
The Warning

Randomization by cluster accompanied by an analysis appropriate to randomization by individual is an exercise in self-deception, however, and should be discouraged.

Cornfield (1978)

- Though Cornfield's remarks were addressed only to GRTs, they also apply to IRGTs.

Summary

- A GRT remains the best comparative design available when the investigator wants to evaluate an intervention that...
  - operates at a group level
  - manipulates the social or physical environment
  - cannot be delivered to individuals without contamination

- An IRGT is the best comparative design when...
  - Individual randomization is possible without contamination
  - There are good reasons to deliver the intervention in groups

- The challenge is to create trials that are:
  - Rigorous enough to avoid threats to validity of the design,
  - Analyzed to avoid threats to statistical validity,
  - Powerful enough to provide an answer to the question,
  - And inexpensive enough to be practical.
Pragmatic and Group-Randomized Trials in Public Health and Medicine

Visit [https://prevention.nih.gov/grt](https://prevention.nih.gov/grt) to:

- Provide feedback on this series
- Download the slides, references, and suggested activities
- View this module again
- View the next module in this series:
  **Part 2: Designing the Trial**

Send questions to:

GRT@mail.nih.gov