

Beyond the Classroom: RCT Design in the Real World



Mike Pergamit, Senior Fellow, Urban Institute

Session Agenda

- Why use an RCT design?
- Intent to treat vs. treatment on the treated
- Setting up the RCT: Referral and randomization
- Threats to randomization
- Questions and Discussion

Learning Objectives

- You will be able to explain why an RCT is the best evaluation design.
- You will be able to explain the difference between intent to treat and treatment on the treated.
- You will be able to see how the referral and randomization process relates to your study objectives.
- You will learn how to enhance your RCT design.
- You will be able to recognize and plan for threats to the RCT design.

Poll

- What experience have you had with RCTs?
 - A: Conducted multiple RCTs
 - B: Conducted one RCT
 - C: None

Poll: Have you designed an RCT, implemented an RCT, both, or neither?

- A: Designed
- B: Implemented
- C: Both
- D: Neither

We will focus on randomization at the individual level, but there are other RCT designs

- Staggered implementation
 - Randomized waitlist
 - Stepped wedge design
- Clustered RCT

Why Use an RCT?

RCTs Create Equivalent Groups

RCT



No
RCT



RCTs Are Considered the Gold Standard because They Are the Best Method to Determine if the Outcomes are the Result of the Intervention

Addresses selection bias

- Most quasi-experimental designs don't
- Provides internal validity
 - Internal validity means the two groups can be considered comparable
 - The assessment of effectiveness can be seen as valid for the groups studied



Agency Buy-In May Be a Challenge—Some People Think RCTs are Unethical because They Deny a Service to a Group

Here are reasons for agencies to accept an RCT:

- RCT is the best way to know that the agency is helping people and the funding is justified; findings from other designs may be misleading.
- A lottery is a fair way of providing a service with limited availability.
- It is not usually the case that the control group gets nothing—they continue to get usual service. Nothing is typically being taken away from clients.
- Funding may depend on proving with evidence that the intervention is effective.
- To help convince them, the evaluator can offer a small number of exemptions—that is where the agency can give the intervention to a few clients and not include them in the evaluation.

Discussion: What Are Your Experiences?

What challenges have you had in getting an agency or service provider to accept an RCT?

- What have you done or said to convince them?
- What successes have you had?

Intent to Treat (ITT) vs. Treatment on the Treated(TOT)

There Are Two Types of Analyses of RCTs: Intent to Treat (ITT) and Treatment on the Treated (TOT)

- In ITT analysis, all members of the study population are included and analyzed as part of the group they were assigned to.
- In TOT analysis, people who actually received the treatment are compared with those who did not receive the treatment.

Intent to Treat (ITT) Analysis Measures the Impact of the Intervention on the Target Population

- Allows policymakers to determine if the intervention has an impact for a population with particular needs.
- Reveals who wants the intervention by allowing anyone in the target population to be offered it.

In an Intent to Treat Design, Detecting Impacts may Be Difficult if Few People Take Up the Service

The expectation:



The reality:



Treatment on the Treated (TOT) Analysis Measures Outcomes for Those Who Actually Receive the Treatment

- Program managers are often interested in the impact on those who received the service, not those who were offered the service but turned it down.
- TOT analysis must account for selection bias.
 - Those who choose to take up the intervention may be different than those who do not take it up, leading to potential bias in impact estimates.
 - Angrist, Imbens, and Rubin (1995) devised a technique (using instrumental variables) to account for who actually received the treatment, which yields the Complier Average Causal Effect (CACE), also known as the Local Average Treatment Effect (LATE).
 - You cannot assume the TOT estimate is generalizable to the full target population.

Even If You Calculate the TOT, ALWAYS Report the ITT Estimates

- The ITT is the estimate that tells us what the overall effect of the intervention is.
- The TOT estimate provides information that can help you refine your target population.

Setting Up the RCT: Referral and Randomization

Referral Process

“Referral” is How a Person/Family Is Identified for the Evaluation

- Types of referral processes:
 - Point in time (e.g., all 17-year-olds in foster care with a case plan goal of emancipation)
 - Flow; predictable (e.g., all young people in foster care turning age 16 each month)
 - Flow; unpredictable (e.g., referred after assessment indicates substance misuse)

Ideally, Referrals Will All Come Through a Central Coordinator

- To control the randomization process, it is best that one person receives all referrals.
- If randomizing within offices, each office could have one person who receives all referrals.
- If feasible, create an electronic method to randomize—take out the potential for mistakes and having participants assigned to groups outside the randomization process

Randomization

When in the Referral Process You Randomize Depends on Several Things

- The earlier you can randomize in the referral process, the closer you get to a true Intent to Treat design.
 - If you can randomize at the time of referral, you are randomizing among those you intend to treat.
- If you need to get consent from participants, you can randomize after getting consent, but you may not need to.
 - You can randomize before gaining consent but not tell the person which group they are in.
 - This allows you to collect information from everyone to use to know if those who gave consent differ from those who did not give consent.

When You Randomize Affects Who Is in Your Analysis Sample

- As discussed in the previous session, the need for consent produces a study population that might differ from the target population.
- If you want a design that reduces the effect of low take-up of the intervention on impact estimates, you could randomize after people express interest in the intervention.
 - This is like building TOT into the RCT design.
 - Sometimes we do only want to know how the intervention works for people who want the intervention. But be cautious—the impacts may not be those that would be experienced by all people who might benefit from the intervention.

Randomization Can Be as Simple as a Flip of a Coin but Doesn't Have to Be: Treatment and Control Groups Can Be Different Sizes

- For a given total sample size, the greatest power to detect impacts comes from equal-sized treatment and control groups. But it isn't necessary, and more power comes from bigger total samples.
- Sometimes the target population is more than double what can be served by the intervention and you can randomize at rates that yield a smaller treatment group than control group.
- Conversely, the number in the target population may be less than double the number that can be served. Then the treatment group can be larger than the control group.
- In a multisite study, you could have both situations.

Concerns about a String of Controls (Flipping “Tails” Many Times in a Row) Can Be Handled with Block Randomization

- Block randomization is where the numbers of treatment and control cases are determined by the size of a “block.” For example, one could have a block of size 8, where there will be four treatments and four controls. The block enforces the numbers and prevents a long string of controls.
- The smaller the block, the less likely a long string of controls; but the more likely someone can guess whether the next case will get treatment or control.

An Example: The First 3 Blocks of an RCT with Block Size 8

Block #1

1. Treatment
2. Treatment
3. Control
4. Treatment
5. Control
6. Control
7. Treatment
8. Control

Block #2

1. Control
2. Treatment
3. Control
4. Control
5. Treatment
6. Treatment
7. Control
8. Treatment

Block #3

1. Treatment
2. Control
3. Control
4. Control
5. Treatment
6. Control
7. Treatment
8. Treatment

A Special Case of Block Randomization Is “Yoking” Two Cases

- If caseload slots need to be filled, and you don't want the risk of a string of controls, then a block size of two (2), one treatment and one control, can be used.
- To avoid knowing what case will be assigned next, “yoke” two cases, that is, require two referrals to come together; the “coin flip” determines which gets treatment and which gets control.

If You Believe That Subgroups May Be Impacted Differently by the Intervention, Consider Stratifying Randomization

- Is there a reason to believe the intervention may have different impacts on different subpopulations?
- Stratification means randomizing within each subgroup. Doing so creates equivalent treatment and control groups for analysis (confirmatory analysis).
 - Examples: parents with children at home versus parents with children in out-of-home care; boys versus girls; youth entering care at young ages versus youth entering care at older ages; children new to the system versus children in care for a while; formal versus informal caregivers in kinship navigator programs.

Stratification Has Strengths and Weaknesses

- Strengths:
 - Creates equivalent subgroups
 - Improves efficiency of estimates
- Weaknesses:
 - Small sample sizes within subgroups may make it hard to have enough sample in each group
 - Stratifying by many subgroups can make randomizing and analysis more complicated

If You Don't Stratify, Any Subgroup Analysis Will Be Considered Exploratory

- Often we don't have a theory about whether an intervention might have different impacts across groups.
- To learn whether the impact differs by subgroups, we conduct analysis for subgroups after collecting data.
 - Without randomization, treatment and control groups are not necessarily equivalent.
- This analysis is considered exploratory, not confirmatory—that is, it isn't confirming a theory.
- If we learn from exploratory analysis that there appear to be differential impacts, we should build that into stratification in the next RCT.

Any questions about referral and randomization?

Threats to Randomization

It's Important to Prevent Control Cases from Getting the Treatment (Crossovers)

Treatment



Control



Some Controls got the treatment



It's Important to Prevent Control Cases from Getting the Treatment (Crossovers)

- If people assigned to the control group get the treatment, you lose the value of the randomization.
- You can't just put those control cases into the treatment group because then you create groups that are not equivalent.
 - This likely introduces bias because control cases that get the treatment may differ from those who don't get treatment.

There Are Several Ways Controls May Get the Treatment

- Staff don't buy into the evaluation:
 - When staff have direct access to the intervention, staff who don't believe the evaluation is fair may get a Control case into the intervention.
- Re-referrals:
 - Caseworkers who want their client to get an intervention may refer them a second time.
- Multiple doors into the randomization process:
 - If randomization occurs at one location but the intervention is provided at other locations, this may provide ways for the Control cases to access the intervention.

Design a Referral and Randomization Process That Keeps Control Cases from Getting the Treatment

- Enlist the agency in taking an active role in the evaluation. Have them understand the need to prevent crossovers.
 - They are your best means of prevention.
 - Include staff, not just program managers, in “selling” the evaluation.
- Set up the system to monitor who is getting the treatment and checks their assignment.
 - You can check for previous referrals, but this may not solve the multiple office problem.

Questions and Discussion

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: Urban Institute et al. (2021). Slide Deck Session 6: Beyond the Classroom: RCT Design in the Real World - Child Welfare Evidence-Building Academy. OPRE Report 2021-112, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.