



Home Visiting Programs

What Isn't There Matters: Attrition and Randomized Controlled Trials

August 2014

OPRE Report #2014-49

A randomized controlled trial (RCT) offers a highly credible way to evaluate the effect of a program. But a strong design can be offset by weaknesses in planning or execution. One common problem that weakens the conclusions of RCTs is **attrition**, or missing data. This brief describes what attrition is, why it matters, and how it factors into the study ratings in the Home Visiting Evidence of Effectiveness (HomVEE) review.

In the HomVEE review, researchers search, screen, review, and rate studies to identify home visiting programs with rigorous evidence of effectiveness. The results may help inform the decisions of state administrators, practitioners, and other stakeholders. Attrition is an essential consideration for HomVEE reviews of RCT studies and may determine whether the study earns a rating of high, moderate, or low.

This brief on research methods and standards was written by Sarah A. Avellar and Timothy Silman of Mathematica Policy Research. Since 2009, Mathematica has conducted the Home Visiting Evidence of Effectiveness (HomVEE) review under contract to the U.S. Department of Health and Human Services. The purpose of the review is to identify, assess, and rate the rigor of impact studies of home visiting programs for pregnant women and families with children from birth to age 5.

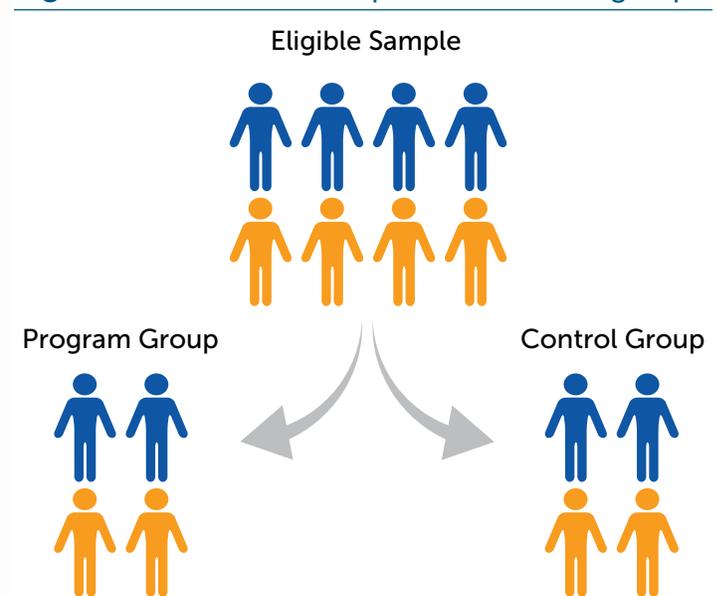
The HomVEE website:
<http://homvee.acf.hhs.gov/>

What Is an RCT, and Why Is It a Strong Research Design?

The distinguishing characteristic of an RCT is that it assigns study participants randomly—either to the **program group** (sometimes known as the “treatment group”) or to the **control group** (sometimes known as the “business-as-usual” group). Members of the program group can receive program services. Members of the control group do not receive program services (though they may receive other services offered in the community).

The main advantage of random assignment is that it creates two groups that are, on average, essentially the same (see Figure 1).¹ This equivalence holds, in theory, even for traits that would not or could not be measured. Thus, at **baseline** (before program services begin), the two groups differ from each other only in that one can receive program services, and the other cannot. As a result, any differences in the **outcomes** between groups at the end of an evaluation can be attributed to the **effect** of the program. In other words, an RCT captures the changes that are caused by the program, as opposed to changes caused by other factors. Thus, if program and control group children are equivalent at baseline, researchers would expect their development to be the same if neither group participated in the program.

Figure 1: Randomization produces similar groups



But because children in the program group received program services, any differences between their outcomes and the outcomes for the children in the control group represent the program’s effect.

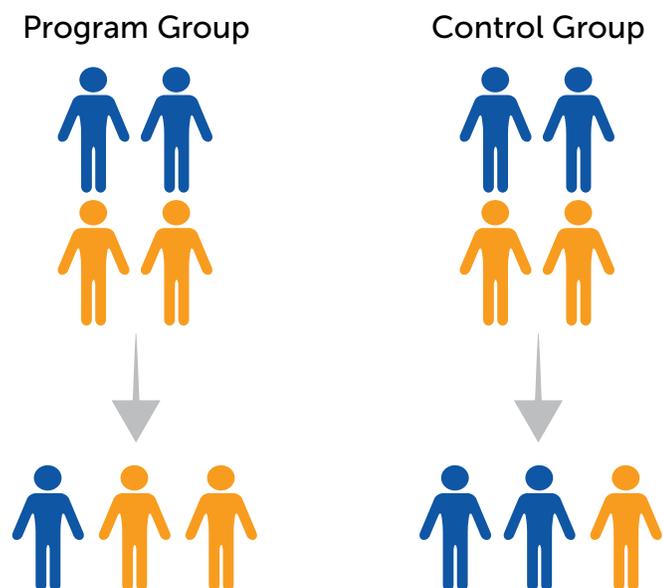
What Is Attrition, and How Does It Undermine the Benefits of RCTs?

Attrition happens when data are missing for some members of the program or control groups. Attrition may occur because study participants do not respond to surveys or are missing from administrative data sets, or it may occur for some other reason. For the purpose of this brief, we will use the example of nonresponse to follow-up surveys.

Attrition threatens most RCTs. **Overall attrition** is the combined loss of data for any sample member from either the program or the control group. For example, if 100 people are randomly assigned (50 to program and 50 to control), and 25 do not respond to a follow-up survey (regardless of whether those people are part of the program or control groups), overall attrition for the follow-up survey is 25 percent. **Differential attrition** refers to the difference in the rate of attrition between the program and control groups. Consider the example above, where overall attrition is 25 percent. If 15 of the 25 people who did not respond to the follow-up survey were in the program group, then program group attrition is 30 percent ($15/50 = .30$). That leaves 10 in the control group who did not respond, meaning that control group attrition is 20 percent ($10/50 = .20$). The differential attrition—30 percent program group attrition minus 20 percent control group attrition—is 10 percent.

Attrition undermines the assumption that the program and control groups have the same measured and unmeasured characteristics at baseline. It can create an imbalance between the two groups if the characteristics of those who have follow-up data differ from those who do not (Figure 2). For example, if less motivated people in the

Figure 2: After attrition, the program and control groups are no longer balanced



Key Terms

Attrition: Missing data for study participants

Baseline: Period of time before start of program services that are being evaluated

Bias: Erroneously shifting the results in one direction or another

Control group: Study participants who should not receive services from the program of interest but may receive other available services

Differential attrition: Difference between the proportion of the program group and the proportion of the control group who were randomly assigned and are missing outcome data

Effect: Changes in outcomes that can be attributed to the program

Equivalence: Similarity (on average) between two groups

Outcomes: Specific measures on which study participants are assessed

Overall attrition: Proportion of the sample who were randomly assigned and are missing outcome data

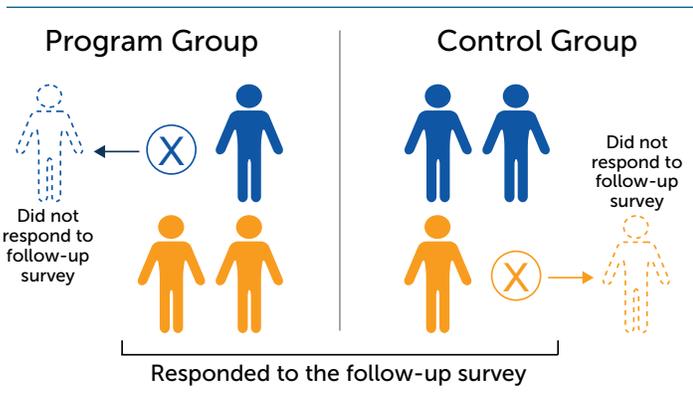
Program group: Study participants who can receive program services; sometimes known as a “treatment group”

Randomized controlled trial: Design in which sample members are assigned by chance to either the program or control group

program group are overwhelmed by the services and then fail to respond to the follow-up survey, the results for the program group will contain fewer unmotivated people. Similarly, if highly motivated people in the control group are frustrated with a lack of services and then do not respond to the follow-up survey, the results for the control group will contain fewer motivated people.

Figure 3 illustrates this dilemma. In this example, researchers could assume that, initially, there were equal proportions of motivated and unmotivated people in the program and control groups. But after attrition, more-motivated people were disproportionately represented in the program group, whereas less-motivated people were disproportionately represented in the control group. Even if the rate of attrition is the same in each group, the reason for attrition is different.² The differences in underlying motivation could cause **bias** (erroneously shifting the results in one direction

Figure 3: Remaining groups differ in underlying motivation



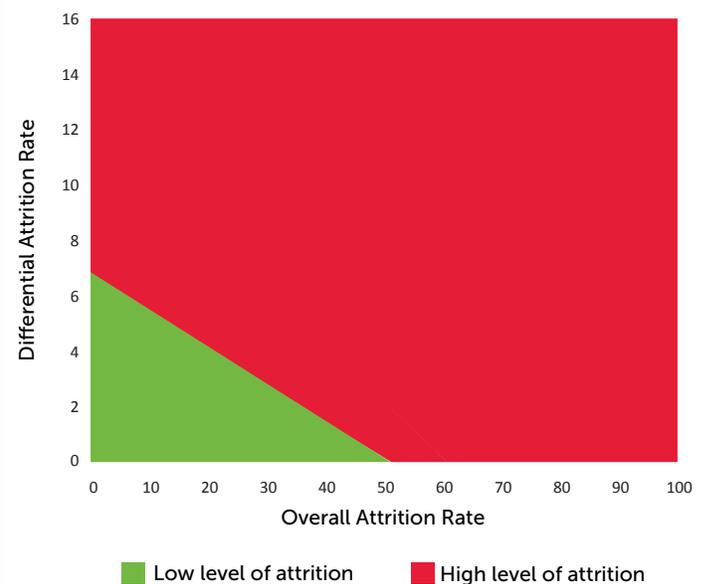
or another)—in this example, leading the researchers to overstate the effect of the program. In practice, however, researchers cannot know all the ways in which the remaining groups (after attrition) may have differed at baseline. Consequently, researchers do not know whether bias overestimates or underestimates program effects.

What Are the HomVEE Standards for Attrition?

In HomVEE, RCTs are eligible to receive a high rating because of their potential to produce highly credible findings. Attrition can introduce bias, however, so “moderate” is the highest rating that an RCT with high attrition can receive. Indeed, an RCT with high attrition can be rated as “moderate” only if the program and control groups included in the analysis (that is, the sample remaining after attrition) are shown to be equivalent at baseline for selected characteristics (for more details, see <http://hom-vee.acf.hhs.gov/document.aspx?rid=4&sid=19&mid=5>).

The HomVEE attrition standards match those developed for the What Works Clearinghouse (WWC), established by the U.S. Department of Education’s Institute of Education Sciences (see <http://ies.ed.gov/ncee/wwc/>). According to the WWC attrition standard, evaluations are classified as having either “high” or “low” attrition based on a combination of overall and differential attrition (Figure 4). The WWC standards also account for an important trade-off between overall and differential attrition—namely, that a study can have a higher overall rate of attrition if it has a low rate of differential attrition. This trade-off is illustrated in Figure 4.

Figure 4: Cutoffs for WWC attrition standards take into account overall and differential attrition



How Can Researchers Minimize Attrition?

To generate valid results and meet the HomVEE attrition standard, researchers should minimize both overall and differential attrition. A few methods to minimize attrition are discussed here. The attrition rate is based on the number of people randomly assigned, so researchers are encouraged to collect data (for example, follow-up surveys or administrative data) for everyone who has been randomly assigned. It does not matter if they have dropped out of the program or did not receive services.

Overall attrition can be minimized by using multiple methods to contact sample members at follow-up; contact can be made in person, by phone, via email, or online. Researchers can also continually track response rates to monitor the

success of data collection efforts, adjusting their strategies and increasing their effort if response rates are low.

Minimizing differential attrition is more important than minimizing overall attrition. Researchers should collect data at equal rates from everyone in the sample (both program and control group members). They also should recognize that more effort may be required to reach people in one group or the other. For example, it may be easier to contact program group members because of their connection with the program.

Attrition is a challenge in most studies. But it must be minimized if an RCT is to produce the most credible findings possible and thereby the most comprehensive understanding of program effects.

Endnotes

¹ Random assignment ensures that there are no *systematic* differences between the program and control groups, but there can be *chance* differences. Researchers use tests of statistical significance to distinguish between measured differences that are probably systematic (significant) and those that may have arisen by chance (nonsignificant).

² In this example, differential attrition may be zero, but it still represents a problem in that it leads to groups with different average motivation. In practice, we expect there to be more overlap in the reasons for attrition from the program and control groups (for example, less-motivated people in both groups may be less likely to respond to the follow-up). Thus, we are more concerned about differences between groups when differential attrition is high.