

5.3: Replication, bias, and nuisance

Introduction

This page describes additional elements of experimental design — replication and bias — which can impact the validity of a study.

Replication in science refers to repeating an experiment under identical conditions multiple times. To the extent the same conclusions are achieved, we can say the outcome is **reproducible** (Simmons 2014) and, therefore, likely to be **objectively true** (Earp 2011). Replication of experimental and sampling units is an essential component of good experimental design. Confusing technical replicates for biological replicates is called **pseudoreplication**.

Bias refers to any of a number of sociological and cognitive errors which may influence conclusions in science (Ionnidis 2005). In a meta-analysis, Fanelli et al (2017) reported that while the size of effect from identifiable biases were small, some disciplines were more prone to these sources of error. That researchers testing similar hypotheses reach different conclusions may result from bias. One such bias is positive outcome bias, where reports of null hypotheses are more often rejected than they should be based on sample size and other concerns of statistical power (see [Chapter 11 – Power Analysis](#)).

The page concludes with a discussion of **nuisance variables**, variables that may be causally associated with the outcome of interest, but themselves are either of no interest in the study, or may not be known in advance (Meehl 1970; see also discussion of **spurious correlations** in [Chapter 16.3 – Data aggregation and correlation](#)).

Technical vs. biological replicates

A laboratory technician retrieves a vial of an immortalized cell line from cryostorage and initiate propagation of the cells for the week's work. After a couple of days of subculturing, the technician has grown millions of cells and is ready to set up an assay. After washing the cells in PBS (phosphate buffered saline), the technician adjusts the concentration of cells to 10,000 per mL media, places the tube on ice, and proceeds to set up a 48-well microplate. To each well, the technician adds media, one or more agents, plus cells, so the total volume in each well is one mL, three replicates per treatment.

In our example there were four treatments plus a control (media only)

- DMSO only (D)
- DMSO + Cytokine (DC)
- DMSO + Small molecule Inhibitor (DS)
- DMSO + Cytokine + Small molecule Inhibitor (DCS)
- Media only (M)

The plate table might look something like Table 5.3.1.

Table 5.3.1. Schematics of a set up for a hypothetical 48-well microplate.

	1	2	3	4	5	6	7	8
A	D	D	D	○	○	○	○	○
B	DC	DC	DC	○	○	○	○	○
C	DS	DS	DS	○	○	○	○	○
D	DCS	DCS	DCS	○	○	○	○	○
E	M	M	M	○	○	○	○	○
F	○	○	○	○	○	○	○	○

The technician has done a lot of work, but from a statistician's point of view, effectively the technician has only one sampling unit to show for his efforts; there is one biological sampling unit because as far as we can tell, the cells were all related. **Biological replication** refers to repeat measures (e.g., same conditions) on biologically distinct entities (Blainey et al 2014). **Technical replication**, in contrast, refers to repeat measures on the same entity, e.g., the same individual (Blainey et al 2014).

However, there were more than one technical samples in the experimental design: twelve in all, because there were twelve wells. Many of you will recognize this as the distinction between technical and biological replicates. In thinking about experimental

design, replicates pertain to the number of sampling units that are treated the same way. In our example cells may be “treated the same way” in at least two distinct ways. Biologically, the cells may be the same: derived from the same clonal cell, all at the same passage (generation) number, each cell genetically, morphologically, physiologically, etc. identical to the next cell. Technical replication on the other hand pertains to the levels of the treatments.

Within each of these treatments we had 3 wells, and that’s at the level of the technical replication.

A final note on this experiment. We as molecular biologists call each of these treatment groups, but statisticians thinking of this design would recognize only one treatment (called in statistics a factor), with four levels of the treatment. We will discuss this beginning in chapter 12, when we introduce analysis of variance.

Types of bias

When we make decisions, we like to think we are rational; that we make decisions based on an evaluation of evidence. And yet, an increasing body of literature suggests that our decisions often made in a manner that falls short of rational processing. We outlined some sources of bias in [Chapter 2.6](#). Deviance from rational decision making is due to any number of **cognitive biases** we may have. Researchers and medical doctors make many decisions and, unfortunately, are just as susceptible to cognitive biases as the rest of us. One kind of bias that is bedeviling to research is **confirmatory bias**. Confirmatory bias refers to a researcher in effect seeking evidence in a way that confirms a prior conclusion (Kaptchuck 2003). Confirmatory bias would be exhibited if we report positive effects of aspirin to relieve migraines in subjects, but exclude cases where the subjects report no improvement. Medical doctors may be risk-averse and therefore tend to over-prescribe, or they may be risk-takers (e.g., rise and fall of high-dose chemotherapy, cf. discussion in Howard et al. 2011).

The concept of bias was first introduced in Chapter 3 and will be returned to in subsequent chapters. In general, bias refers to the objectivity of measurement and inferences about such measurements. Bias implies that a series of measurements consistently fail to return the true value of the variable. Bias is systematic error and may be associated with a poorly calibrated instrument or even the use of improper sets of rules for measurement given the nature of the characteristic. Bias can be captured by the concept of **accuracy**.

Bias is challenging to eliminate; the best way is to design the experiment so that the observer, the researcher, is unaware of the specific questions to be tested. The research is conducted “blind.” Clinical research provides the easiest examples. A trial can be blind in a couple of ways:

- The subjects do not know which treatment they receive, but the researcher does know
- The researcher does not know which treatment subjects receive, but the subjects know
- Neither the researcher nor the subjects know which treatment they receive.
- The other possibility is that both the researcher and the subjects know treatments received — this would be a poor design.

Bias can enter a research project at multiple levels. A partial list of sources of bias in research includes (modified after Pannucci and Wikins 2010):

Bias that occurs before the experiment:

- Inadequate planning of the experiment, flawed design
- Bias in selection of subjects from the reference population
- Surveillance bias, where one group is studied more closely than another group.

Performance bias refers to conditions of an experiment that introduce unintended differences between groups. For example, subjects enrolled in a weight-loss study randomly assigned to the control group may react poorly when they realize they are not receiving the experimental intervention. This leads to potential for a systematic bias — participants in the control group may behave differently, counteracting the point of randomization (McCambridge et al 2014).

Another well-known performance bias is associated with eliciting maximal performance from animals, such as running stamina or maximum sprint running speed (Losos et al 2002). For example, sprint running speed of lizards may be measured by placing lizards onto a high-speed treadmill, then increasing belt speed to match the individual’s burst performance (e.g., Dohm et al 1998). Clearly, not all individuals will perform to maximum physiological capabilities under these conditions.

Bias after the experiment:

- Citation bias
- Confounding

Another way to view sources of bias is that, at least from the perspective of the researcher, the bias is likely to be **unconscious bias**. After all, if we knew about the sources of bias we would work our experiments in such a way to minimize known error-producing sources (Holman et al 2015). Selecting subjects at random and keeping record keeping blind with respect to treatments are among the best tools we have to avoid this kind of bias.

Nuisance variables

A key quality of experimental design is the opportunity afforded the researcher to apply control. In a classic view of the experiment, one allows only one aspect to vary, the treatment applied, all else is controlled. Thus, any difference in outcome may be attributed to the treatment. However, it is likely impossible to control for all possible variables that may effect our outcome; **confounding** clear interpretation of results (see Chapter 16.2 and 16.3). Consider, for example, the effects of age on running performance (Fig. 5.3.2).

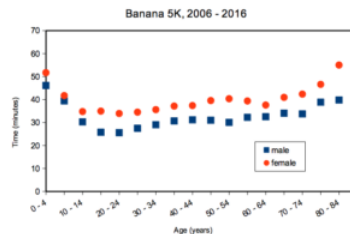


Figure 5.3.2: Mean 5K running times (minutes) by age and gender (2006 – 2016, Jamba Juice Banana 5K race, Honolulu, HI).

If we are interested in gender differences, we must also deal with the change in performance with age: older runners regardless of gender tend to run slower than younger runners. Thus, in the comparison of genders, age is a *nuisance variable*. In the literature, you may see “lurking variable” or other such names used. The concept is that these variables may be causal, but are either of no interest in the study, or may not be known in advance. Therefore, age must be accounted for before we can address gender differences, if any. Age, in this case, is considered a nuisance variable because we are not primarily interested in age effects, but age clearly covaries (is associated) with running performance. How best to handle nuisance variables? One approach is to match by age (**blocking effect**, see Chapter 14.4); another approach would be to randomly select with respect to age when compiling female and male times.

Here is another example of confounding. Results from ecological experiments intended to test for presence of competition for resources by plant species by removing plants may alter herbivory levels; thus, survivorship may not be the result of reduced plant-plant competition, but the result of changes in herbivore behavior (Reader 1992).

Questions

1. I cite a number of articles in Mike’s Biostatistics Book. If you’re paying attention, you’ll have noticed that all of my citations, at least for articles, link to a source that you can either download or read online. What kind of error am I committing by following this approach?
2. A doctor has been seeing patients with upper respiratory tract infections (URTI) all week. Although most URTI are caused by viral infections (Thomas and Bomar 2018), the doctor has prescribed each patient a ten-day dose of antibiotics. As you should know, antibiotics have no effect on the course of viral infections; antibiotics work against bacterial infections. What kind of error has the doctor committed by following this routine prescription of antibiotics?
3. Figure 5.3.2 is a plot of mean running times for men and women across different ages. What statistic is missing from the graph so that we can’t conclude from the graph anything more than there is a trend that men are faster than women?
4. In large part because of the tendency for over use of antibiotics, doctors are less likely today to prescribe antibiotics for patients with upper respiratory tract infections than in the past (Zoorob et al 2012). URTI are mostly caused by rhinovirus infections and, therefore, the course of infection in a patient should be unaffected by addition of antibiotic. However, although rare compared to URTI, life-threatening illnesses like acute epiglottitis caused by bacteria infection can sometimes develop or accompany URTI. Thus, not prescribing antibiotics for a diagnosis of URTI may risk a worsening condition for the patient. This sets the doctor up for a potential cognitive bias – to prescribe or not to prescribe antibiotics prophylactically? Discuss this decision in the context of potential sources of cognitive biases and possible outcomes of the decision.

This page titled 5.3: Replication, bias, and nuisance is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Michael R Dohm via source content that was edited to the style and standards of the LibreTexts platform.