

## **NIH Rigor and Reproducibility Training Module 4: Sample Size, Outliers, and Exclusion Criteria**

### **Potential Discussion Points and Questions:**

#### Starting Points:

- Sample: here, a sample is defined as a single value or observation from the larger set of values
- Sample size: the optimal number of samples that should be used to reach sufficient statistical power; also referred to as 'n'
- Outliers: an observation that lies an abnormal distance, typically +/- 3 standard deviations, from other values in a random sample from a group of results<sup>1</sup>
- Exclusion criteria: standards set out before a study or review to determine whether a sample should be included or excluded from the study or analysis<sup>2</sup>
- Characterization of "normal" for a specific experiment is an important component to identifying outliers and determining exclusion criteria

#### Lead-in Questions:

- Do you have a standard approach to determining the appropriate sample size and setting criteria for outliers – how you determine the numbers that go into your power analysis?
- How do you know what "normal" is if you don't know the result? Can you do this initially? Will determination of the best statistical method and approach be useful in defining normal?

#### Follow-up Questions:

- Do you think it is common to report data from a single experiment (technical replicates) to generate an "exciting" finding? How often is this type of practice viewed as a way to expedite the research process?
- Since this is a grant application with preliminary results, is it acceptable to include results in such a manner?
- Is it appropriate for the applicant to purposely leave information about the type of replicates out and plot the data in such a way as to suggest significance over multiple experiments? Can it be considered falsification and therefore possible misconduct? If so, what are the potential consequences? What if it was simply an oversight?
- If this was your grant application, how would you have portrayed the data? Would you clearly state the "n" in the figure legend and/or describe this in the body of the grant? Would you have indicated the exclusion of data?
- Do you think papers or grant applications should delineate the use of biological vs. technical replicates in the figure legends (or elsewhere in the document)?
- The reviewer provides an analogy of "taking a thousand cells from one animal" and getting "just one point" from the resulting data. Is this always the case?<sup>3</sup>
- Do you think the review of the project will be affected?
- Do you think a typical review session discussing this issue would be as collegial?
- The reviewers appeared to be convinced easily that the figure was misleading. Do you think this transition in thought would have been so quick and painless if it were a real review session?

---

<sup>1</sup> <http://grants.nih.gov/grants/guide/pa-files/PAR-13-383.html>

<sup>2</sup> <http://www.nature.com/nmeth/journal/v11/n9/full/nmeth.3091.html>

<sup>3</sup> Arts, E et al. A solution to dependency: using multilevel analysis to accommodate nested data. *Nature Neuroscience*. 2014 April; 17(4): 491-496. <http://www.nature.com/neuro/journal/v17/n4/full/nn.3648.html>