Introduction to the NIH Rigor and Reproducibility Training Modules

Potential Discussion Points and Questions:

Starting Points:

- Concerns are mounting about the current system for ensuring reproducibility in biomedical research
- Science is often viewed as "self-correcting," which is valid over the long-term, but this is no longer true over the short-term
- Some factors contributing to lack of reproducibility and rigor in research[1]:
    - Poor training in experimental design
    - Increased emphasis on making provocative statements
    - Publications that are missing basic elements of experimental design
    - Pressures to publish, particularly in high-impact journals
    - Promotion and tenure incentives for publications
    - Dearth of publications on negative data or flaws in previously published work
- When considering reproducibility in research, several issues arise, including potential biases and rigorous experimental design
- Bias: prejudice in favor of or against one idea, thing, person, or group compared with another, usually in a way considered to be unfair
    - Bias in experiments can often be unconscious and unintentional. It can result from:
        - Unknown or unavoidable differences between comparison groups
        - Less than ideal experimental designs that introduce systematic erroneous differences between test groups
- Scientific rigor: the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation, and reporting of results.
- Rigor in scientific research includes consideration of[2]:
    - *Experimental Design*: Rationale for the selected models and endpoints (animal and/or cellular); Adequacy/comparability of the controls and other experimental conditions; Route and timing of intervention delivery/dosing; Justification of sample size, including power calculation; Statistical methods used in analysis and interpretation of results; Maintenance of a detailed laboratory record
    - *Minimizing Bias*: Methods of blinding (allocation concealment and blinded assessment of outcome); Strategies for randomization and/or stratification; Reporting of data missing due to attrition or exclusion; Criteria for identifying outliers; Reporting of all results (negative and positive)

[1] http://www.nature.com/news/policy-nih-plans-to-enhance-reproducibility-1.14586

[2] Based on concepts described in: Landis, et al. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature*. Oct 11, 2012; 490(7419): 187–191. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3511845/

    - *Results*: Independent validation/replication, if available; Robustness and reproducibility of the observed results; Dose-response results; Verification that interventional drug or biologic reached and engaged the target
    - *Interpretation of Results*: Consideration of alternative interpretations of the experimental data; Relevant literature in support of or in disagreement with the results; Discussion of effect size in relation to potential clinical impact; Potential conflicts of interest
- Issues surrounding reproducibility, or the lack thereof, do not necessarily imply research misconduct.
    - Research Misconduct: falsification, fabrication, or plagiarism in proposing, performing, or reviewing research, or in reporting research results; research misconduct does NOT include honest error or differences of opinion[3, 4]
- On the contrary, most issues surrounding reproducibility and rigor are often the result of unintentional behaviors, such as lack of transparency in reporting and improper statistical analysis, which would have been avoided, had the researcher been aware.
    - Example of scientific misconduct – a researcher who selectively excludes or manipulates data in favor of a biased outcome
    - Example of lack of reproducibility – a researcher does not provide all the detailed steps in the methods section of their paper due to length constraints, including information that may seem minor, but is essential to the experiment, which results in the inability of another lab to duplicate their results

Lead-in Questions:

- When you think of bias in an experiment, what comes to mind?
- What is an example of an unavoidable bias? What could be an example of an avoidable bias? An intentionally introduced bias?
- With regard to preparing samples for an experiment, what might constitute bias on the part of the experimenter? How might this type of bias manifest in the design of an experiment?
- How do you think bias might affect what is documented in a protocol, or detailed in a materials and methods section?
- Do you think there is a component of laboratory and research culture that must be considered when addressing reproducibility and rigor in research?
- Would people be willing to admit that their lab might not always conduct the most rigorously designed studies?

[3] http://grants.nih.gov/grants/research_integrity/research_misconduct.htm
[4] http://ori.hhs.gov/sites/default/files/42_cfr_parts_50_and_93_2005.pdf