Maximizing Statistical Interactions Part III:
Analysis and Write-up
Provided by: The Biostatistics Collaboration Center (BCC) at Northwestern University

If a statistician has been involved with your project from the beginning, the analysis and write-up should be fairly straightforward. Otherwise, you will have to let him or her in on all the issues that have occurred during the study. Please also be sure to refer to part I and II of this series!

1. What is your deadline?
   a. Is this analysis for an abstract submission deadline, poster, or presentation?
   b. If you don't have a deadline, per se, but are anxious about getting a manuscript out – please provide the current draft so that we can answer the questions as they are written in the manuscript.
2. What was the design?
   a. What was the sampling: simple random, cluster, systematic, stratified?
   b. Was the "intervention" independent in terms of other factors, or have you matched on age and gender; was it a cross-over study?
   c. Was there anything "particular" about the study in the data collection or sampling that might be helpful to know? For instance, if you collected study participants from two different sites or clinics, it might be informative to examine any potential differences in site populations.
3. What is the outcome of interest?
   a. Are you interested in if an event happened, or when it happened?
   b. In analyses that you have seen regarding this outcome, how is it usually treated? Are logarithms commonly used? (A statistician will look at the data to determine if it could be considered "normally distributed" but some sense of "if it usually is" can be garnered from the existing literature).
   c. If you have a categorical outcome, does it make sense to look at "all categories" or might it make more sense to consider a dichotomized version of the outcome as well (e.g., stage of disease or any disease)?
   d. Note: you don't have to be specific in terms of guessing what analysis you want. A collaborative environment is usually more productive. Of issue is that your specific data may not "fit" assumptions required for popular analyses in the literature, additionally those analyses may not be able to address the question you truly want answered. If you are more upfront with the question you want answered rather than suggesting specific analyses, you will generally be happier with your experience with the BCC.
4. What is your specific question (or hypothesis)?
   a. Are you interested in a difference between groups?
   b. Are you interested in whether two (or more) groups differ after a long intervention? Or are you more interested in whether the pattern of the change in the groups differed over time?
   c. Are you interested in subsets of the data or the whole dataset?

d. Are you interested in any post-hoc tests? If you have three "new" treatments and a control group, are you just interested in how each fared relative to the control, or are you interested in all possible comparisons?

5. Are there any potential confounders or effect modifiers that should be examined?
   a. Occasionally, there might be a "best" statistical model, which may or may not make biological sense. We are great at finding the best statistical model.
   b. Again, through your expertise and literature review, you may be aware that certain variables need to also be considered in multivariable analyses.

6. Are there any secondary outcomes that should be examined? (any confounders/modifiers for that?) - It is generally best to mention these from the start.
   a. From a statistically appropriate point of view, we want to be sure that we are not "fishing" for results or "data dredging." If we present overall non-significant findings, it would be inappropriate (statistically) to start examining every possible hypothesis in order to find something that is significant.
   b. There are methods available for finding the "best" cut point if you wish to dichotomize continuous variables, or "best" multivariable predictive models from a statistical point of view. These models may not be optimal in answering your question, or adjusting for covariates that clinically may be relevant. Collaboration brings the best of both "worlds" (clinical application and statistical theory) together so that your data can tell an accurate, informative story.

7. As with any paper, authorship is generally expected whenever substantive input on the design or analysis is provided.
   a. As for the BCC's approaches to authorship, we endorse the criteria recommended by the International Committee of Medical Journal Editors (http://www.icmje.org/#author). These criteria state that authorship credit should be based on:
      i. Substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data;
      ii. Drafting the article or revising it critically for important intellectual content; and
      iii. Final approval of the version to be published.
   b. Authors should meet all conditions.
   b. Since the statistician often contributes to each of these three levels of intellectual and scientific involvement in manuscript preparation, we generally feel that authorship is appropriate for this kind of collaborative effort.
   c. It should also be noted that there is no Feinberg policy allowing investigators to provide payment as a replacement for authorship. When the BCC work is substantive, both payment for services and authorship on papers is justified.