**When Should I See a Statistician?**

There are two basic reasons for seeking statistical help. The first is when you are planning your study and want to know how large a sample you will need, or when outlining planned analyses. The second is when you are giving him or her a copy of the data to analyse.

When designing your study and are seeking assistance on sample size determination and sampling, make certain that you provide her or him with the following information:

- What is the overall goal of your study (e.g., did the intervention improve medication compliance)?
- Are there any other questions that you would like to ask of your data (e.g., does ethnicity and age have an overall effect on medication compliance)?
- What is the overall design of the study?
- Will the data be stratified (e.g., will you have subgroups of patients that you need to recruit)?
- Do you have any information from previous studies on expected dropout/refusal rates?
- How many variables/measures will there be?
- Are any measures repeated on the same individual, how often?
- Will the data be time based (e.g., time to death or discharge)?

   Depending on your study design, there may be other issues that you will have to provide as well. You should note that sample size is dealt with differently in feasibility studies and studies exploring reliability/validity issues. In feasibility studies, the goal is primarily to gather information about methodology: scales/questionnaire being used, recruitment strategy etc. As effect size is normally not examined in such studies, sample size is more a question of will I have enough to analyze the pilot data to explore these issues for a much larger study. With respect to reliability/validity, typically referred to as a psychometric study, one needs to determine if the sample is large enough to conduct the analysis in question. For example, some reliability analyses require a subject/item ratio of 10:1. In such a scenario, a scale of 10 items would require up 100 people. Other forms of reliability analysis will require less. If uncertain which form of psychometric analysis is best, discuss it with the statistician.

When approaching the statistician for assistance with analysis, have the above (you will have to remind her or him about your study!) and following information readied or done.

1. Make certain you provide a spreadsheet file that contains the raw data and not data already summarized. Charts and figures are fine to look at but difficult to use to answer questions regarding significant findings.
2. Your data file should list subjects (one subject per row) by all the variables (column). Do not use variable names that exceed 12 characters, have spaces, or use characters found in mathematical manipulations (e.g., do not use "*", "+", "/" characters etc.). Also, make certain the variable names are unique (e.g., CS1, CS2, CS3 etc.).
3. Before handing over the data file to the statistician, check each variable for data entry errors or outliers by generating the frequency distribution for each variable using your spreadsheet

program. Outliers are values that exceed what you would normally expect. Either the outlier is a rare but valid number and needs to be included, or it needs to be removed or modified. If the later, you can discuss with the statistician the options for dealing with it. Remember, though, that unless the study falls within the expertise of the statistician, your data values will only be a "number" to him or her. The final decision regarding how an outlier is dealt with should be yours.

4. Give her or him sufficient time to fit the analysis into their own work schedule. Remember, the statistician will have their own work to attend to as well as that of other colleagues and residents. The sooner you arrange your meeting with the statistician and plan for the analysis, the better the experience will be for you and the statistician.  Always remember, your sense of urgency will not be theirs!