

## **STAT 8801 – Case Study #3**

**Group 3: Jenny, Joel, Rachel, Qi, Xiaoyi, and Lindsey**

**February 10, 2012**

### **Background**

A client wants to run an experiment to compare a new dressing for minor operations to a standard dressing. Subjects will be recruited and will have two simulated wound sites, one with the new dressing applied and the other with the standard dressing. Each site will be rated on a scale of 1 to 5 for factors such as pain and infection; these measurements will be repeated at 7, 14, and 42 days.

### **Questions for the Client**

This will likely be an expensive study, since we probably need to pay subjects for their time and discomfort, along with medical professionals to carry out the experiment. We will definitely want to discuss budget constraints with the client and talk about sample size expectations, raising the possibility that we might have some attrition. We would also like more details about how the wounds will be simulated and who will be doing the dressing and the recording. Will all ratings be self reported by patients, or will a medical professional be taking some of the more objective ratings on infection and healing? How many total responses are being measured? Additionally, we'll want to know if certain comparisons are more important than others to the client. For instance, should we give more attention to the pain scores or healing time? Are the day 7 results deemed more significant than day 42 results? It's also important that we understand how the client interprets the scale of 1 to 5 and how the scale is explained to the subjects.

### **Recommendations**

First, we recommend that this be a blind trial. If possible, it is best if the patient does not know which site has the new dressing and which has the standard dressing. If someone else is doing the rating afterward, it would be even better if that person also didn't know which dressing was used on each site. Next, we would want the wound sites to be comparable on each patient; for instance, we'd suggest putting wounds on the right knee and left knee instead of a right wrist and left shoulder. We'd want to make sure that the same person dresses both wounds on each subject. Also, record who dresses and records measurements for each wound. This way we can test for differences due to the medical team if necessary. Generally we'd suggest having as large a sample size as possible within the budget constraints. In this case we will likely have a relatively small variance; after running some preliminary power calculations, we'd prefer a sample size of at least 15, but hopefully closer to 40. Using over 50 subjects would probably not be necessary.

### **Ideas for Analysis**

Once all data is collected, we suggest computing the difference for all corresponding data points (standard score– new dressing score for each subject, response, and date). Then we could test the null hypothesis that these differences are equal to zero using paired t-tests. We would of course do an overall test to check if the mean scores of new and standard dressings are equal, but we could also slice and dice the data in lots of ways, depending on the interests of the client. For example, we could compare only the 14 day scores, or only the scores for pain. This would give the client plenty of detail around where the new dressing scored better or worse, in case not all results are conclusive. Additionally, we could do both one-sided and two-sided tests to see if the new dressing is in fact better, or just significantly different from the standard dressing. This manner of analyzing the data has many advantages. The results would be easy for the client to understand, and method offers a great level of granularity with many different points of comparison possible. Further, the analysis wouldn't be too difficult to carry out, and would be easily repeated if the client wanted to do a similar study later on.