**Name**_____**KEY**_____

**Group members**_____

**Part 1: Orienting yourself to the Example Project**

Read the Example Project description.

1) What is(are) the outcome(s) for this analysis?

   There are two outcomes. They could be represented in one of two ways for analysis.

   Whole mouth average pocket depth and whole mouth average attachment loss at one year is one possible set of outcomes.

   The change in whole mouth average pocket depth between year 1 and baseline and change in whole mouth average attachment loss between year 1 and baseline.

   Since this is a randomized trial, the groups are randomized at baseline and their average baseline values are likely similar. Thus, we can study year 1 outcomes adjusting for baseline values. This is often a more powerful analysis (but not always), as the baseline value likely acts as a precision variable.

2) What is(are) the primary variable(s) of interest for this analysis?

   The treatment variable.

   **A common question for this project is what to do with the two control arms.** Keep them separate. They fill two different roles. One is a control with no treatment (but they know they are a control) and the other is a more true control (not active treatment but the process is the same). There is concern that putting a wet substance on the gums actually does something. So, they also included a no-treatment control arm also.

   Another common question for this project is if there is an ordering effect for the treatment levels. This is often asked because who wants to analyze so many pairwise comparisons of treatment groups. Unfortunately, it is possible that higher doses are bad/irritating and thus the treatment group needs to be categorical in the analysis.

3) What is the primary hypothesis you think you will be answering in your analysis?

   Q1: Is average pocket depth at 1-year different between treatment groups? A subquestion might be: is the average pocket depth smaller in any treatment group compared to the control groups at 1-year?

   Q2: Is average attachment loss at 1-year different between treatment groups? A subquestion might be: is the average attachment loss smaller in any treatment group compared to the control groups at 1-year?

If you chose change as your outcome, you might use the following wording:

Q1: Is average change in pocket depth different between treatment groups? A subquestion might be: is the average change pocket depth greater in any treatment group compared to the control groups at 1-year?

Q2: Is average change in attachment loss at 1-year different between treatment groups? A subquestion might be: is the average attachment loss greater in any treatment group compared to the control groups at 1-year?

4) What are three questions that you have for the investigator?

Answers will vary.

**Part 2: Drafting an Introduction Section**

5) What is the primary hypothesis you think you will be answering in your analysis (i.e., did the answers to the question change anything)?

No new solution. This might change if you decided to change the form of the primary outcome or to change how you use the treatment variable after discussion with the investigator.

6) Draft an introduction section to your data analysis plan. It can be in bullet points or paragraph form.

We received data on 130 individuals who participated in a treatment trial to assess the effects of treatment(s) for gum disease. The goal of the treatment is to reduce attachment loss and pocket depth, two measure of gum health. These measures were repeated on various sites in the mouth and averaged for analysis. We received baseline data (pre-treatment) on pocket-depth and attachment loss and 1-year follow-up data on these two measures. Other demographic information was also provided (age, sex, race and smoking status). There were 5 treatment groups (2 placebo like groups (placebo and no-treatment), and low, medium and high dose treatment groups).

**Part 3: An initial analysis plan**

7) Develop a general analysis strategy for this project. It can be in bullet points or paragraph form.

The primary outcomes in this analysis will be follow-up [or change in] attachment loss and pocket depth. Each outcome will be modeled separately. Linear regression will be used to assess whether average attachment loss at 1-year differs between treatment and placebo groups after adjusting for baseline attachment loss. This analysis will be repeated for pocket

depth. In further analysis, the linear regression will be adjusted for age, sex and smoking status. There are too few non-white participants to make adjustment for race meaningful.

Note, if the outcomes at 1-year are fairly skewed, the primary outcomes for analysis will be log (base-e) transformed prior to analysis and the regression will be adjusted for log baseline outcome. A p-value < 0.05 will be considered statistically significant for the test of any treatment differences (the omnibus test). We will use a Bonferroni correction to adjust for multiple comparisons when comparing any two groups if there is evidence of any treatment differences using the omnibus test. R version XXX will be used for all analyses.

**Part 4: Table shells. The start of a results section.**

8) What types of descriptive statistics do you want to compute to assess the data and describe the data? How would you put this information into a Table in a report?

Means and standard deviations will be created for baseline, follow-up, and change from baseline to follow-up in pocket depth and attachment loss. Mean and standard deviation will also be calculated for age.

We will report the percentage of individuals who dropped out by group. Since there are only two time points in this study, only individuals who have complete data for both time points will be included in the analysis.

**Proposed Table 1:**

| | Control A | Control B | Low Dose | Medium Dose | High Dose |
|---|---|---|---|---|---|
| N | | | | | |
| Age (Mean, SD) | | | | | |
| Gender (N,%) | | | | | |
| Race (N,%) | | | | | |
| Smoking Status (N,%) | | | | | |
| Attachment Loss Baseline | | | | | |
| Pocket Depth Baseline | | | | | |
| Attachment Loss 1-Year | | | | | |
| Pocket Depth 1-Year | | | | | |
| N dropout (N,%) | | | | | |

9) What types of graphs do you want to create to assess and describe the data? Why, what are you looking for with these graphs?

Boxplots and histograms will be created by treatment group to assess skew in the data.

The frequency distributions of sex, race and smoking status will be created to assure there are enough data in each level of the group for analysis.

10) What statistics should you present to describe the results of your analyses and statistical models to address the investigator's hypotheses and research questions? What tables and/or graphs would be useful in communicating these results?

I would plan on presenting the p-values for the F test for any difference in PD or AL between treatment groups. I would want to present estimates, p-values, and 95% CI's for the pairwise comparisons between the treatment groups – since there are many comparisons, it might be easiest to put the information in a table. To illustrate the results, I might make a graph of the change by group or make a line graph of the baseline to 1-year change colored by group or make a bar chart with the pairwise comparisons between groups represented.

I might also consider including a supplementary table showing the estimates, 95% CI's and p-values for the adjustment variables included in the models for completeness.

See ExampleAnalysisPlanProject0.docx for a complete draft of a data analysis plan for this example project.