BIOS 6623 Project 0 Written Report

Michaela Palumbo

September 13, 2017

Introduction

This project was done in collaboration with dental researchers who aimed to explore a new gel treatment for gum disease. The study recruited 130 subjects and data was measured on subjects at baseline and after one year. Subjects were divided evenly among 5 groups. One group was a control group that continued with their normal dental hygiene routine. The other 4 groups were treatment groups that received a gel to apply to their gums twice daily. There were 4 dose levels for the gel treatment: placebo (no active ingredient), low dose, medium dose, and high dose. The measurements for pocket depth and attachment loss were recorded at base line and one year for each subject. These recorded measurements were averages of multiple measurements taken at different sites throughout the subject's mouth. The number of sites in the mouth at which these two outcomes were measured was recorded in the dataset. The dataset also contained values for the following additional covariates: age, race, sex, and smoking status. The primary scientific question of interest for this study was whether treatment results in lower average pocket depth and attachment loss at one year. This scienfic question of interest can be rephrased with the following two statistical hypotheses. Is there a significant difference in the average change of pocket depth over 1 year between the controls and any of the treatment groups? Is there a significant difference in the average change of attachment loss over 1 year between the controls and any of the treatment groups? To address these statistical hypotheses and ultimately attempt to answer the scientific question of interest, the provided data was analyzed as described below.

Methods

The first step upon receiving the data was to analyze and clean the data. It was immediately noted that 27 subjects were lost to followup and did not have measurements for attachment loss or pocket depth at the one year visit. A graph was created to explore the trend of dropout rates across treatment groups. We thought it would be important to understand the pattern of missingness due to loss of followup for several reasons. Firstly, because it might impact the interpretation of the analyses. Also, understanding the pattern of dropout rates may shed light on the scientific question of interest or inform potential changes that may be

needed to improve the study design in future experiments. After examining the pattern of missingness, the subjects that had been lost to followup and did not have measurements at the one year visit were removed from the data set, leaving a new sample size of 103 subjects. All proceeding steps in the data analysis plan were performed on this new data set that excluded subjects without a one year measurement. This decision was made because we would not be including these subjects that had missing values at one year in our models, and we wanted our data exploration and descriptive statistics to match the sample of data that would be used in the final models. For clarity, when referring to the "dataset" or the "sample" throughout the rest of this report, it is referencing the new data set that was created to exclude subjects that dropped out.

Using this new data set, we next created histograms for each continuous variable and frequency tables for each categorical variable in the dataset to determine if there were any outliers or unrealistic values. It did not appear that there were any drastic outliers or unreasonable measures in the dataset. We also created 2-way tables and boxplots to compare the demographic information within each treatment group. We did this in order to determine whether each of the treatment groups were comprised of roughly similar samples. Next, we created two new variables. We created a variable for the change in pocket depth over one year and for the change in attachment loss over one year. This was done by subtracting the baseline outcome measurement from the outcome measurement at the one year visit. We created these difference variables for pocket depth and attachment loss because we planned to create models that used this difference as the outcome. We chose to use the difference between the measurements as the outcome, because we felt that comparing the changes in pocket depth and attachment loss between the 5 treatment groups would be an effective way to address the scientific question of interest and we would be able to interpret the results in a fairly simple and intuitive manner. Once these new change variables were created and we decided that we would be using them as our outcome we explored the relationships between the covariates and the outcome by creating boxplots and/or scatter plots. These plots were also examined to assess whether model assumptions such as normality and homoscedasticity appeared to be satisfied. The last phase of the descriptives stage was to calculate summary statistics including means, standard deviations, or proportions for the variables.

We then moved on to fitting models and the model selection process. Separate univariate linear models were fit for each of the outcomes of interest: change in pocket depth and change in attachment loss. The primary predictor of interet for these models was treatment group. For both outcomes, the following models were fit: a simple model (including only the outcome and treatment group), 6 crude models that included the outcome, treatment group and the addition of only one of the additional covariates (baseline measure, sites, age, race, sex, smoking status), a full model (including the outcome, treatment group, baseline measure, sites, age, race, sex, and smoking status), and 6 reduced models that were equal to the full model but exclued only

one of the covariates (baseline measure, sites, age, race, sex, or smoking status). Before the models were fit, we reordered the treatment group variable so that the control group would be used as the reference group. We chose to use the control group as a reference group because the goal of this study was to determine if treatment, at any dose level, impacted the change in pocket depth or attachment loss over one year. This study was not interested in comparing the different treatment dose levels with eachother. For this reason we made controls our reference group so that we could easily compare each of the treatment group levels with the control group. During the model selection process when all of these models were fit for each outcome, we did not look at parameter estimates or p-values. During this stage, we only used AIC and residual plots to compare models. The same model was chosen for each outcome. This was primarily based on the fact that similar models fit the data the best for each outcome. Models that fit the data best had a low AIC and nice residuals plots demonstrating that model assumptions were met. Also, we wanted the models to be the same for each outcome if possible, that way the models would address the questions and hypotheses of interest in a uniform manner for each of the outcomes. Once final models were selected for each of the two outcomes, we were able to look at parameter estimates and p-values in order begin to interpret the results of our models.

Results

When examining the pattern of dropout from this study, we saw a trend across treatment groups (appendix, figure 1). It appears that there was higher dropout as treatment dose increased. The high dose treatment group had 10 subjects lost to followup. This was followed by the medium dose group then the low dose group that had 6 and 5 subjects lost to followup, respectively. The control and placebo groups each only had 3 subjects lost to followup. As a result, the sample size for the high dose group was quite smaller than the sample sizes in the control and placebo groups. When examining the summary statistics across groups (appendix, table 1), it is notable that all groups had a higher percentage of females than males. However, group 5 in particular had the highest percentage of females (81.2%) compared to the rest of the treatment groups whose percentage of females ranged from 56.5% to 69.9%. In regards to other demographic variables such as age, race, and smoking status, these characteristics were fairly even across the treatment groups.

When looking at the boxplots of change in attachment across treatment group and change in pocket depth across treatment group (appendix, figures 2 and 3) it doesn't appear that there are differences in changes in attachment or pocket depth across treatment group.

For our two final models we chose a model that included the outcome (either change in pocket depth or change in attachment loss), treatment group, and baseline measure as a covariate (pocket depth or attachment

loss at baseline visit). We chose to include the baseline measure as a covariate because it improved the AIC compared to the model with only the outcome and treatment group. The results of the linear univariate model with change in pocket depth as the outcome are as follows. There were no significant differences in average change in pocket depth over one year between any of the treatment groups compared to the control group. The results of the linear univariate model with change in attachment loss as the outcome are as follows. There was a significant difference ($p = 0.0232$) in average change in attachment loss between the medium dose treatment group and the control group. The average change in attachment loss over 1 year was 0.176 units (95% CI: 0.025 to 0.327 units) higher for subjects in the medium dose treatment group compared to subjects in the control group.

Conclusions

Recall that the scientific question was whether or not the new gel treatment would result in lower pocket depth and attachment loss at one year. Based on our analysis for change in pocket depth, we can conlude that gel gum treatment at any of the dose levels does not significantly result in a smaller change in pocket depth over one year compared to subjects receiving no gel treatment. Based on our analysis for change in attachment loss, we can conclude that only the medium treatment dose results in a significantly different change in attachment loss over one year compared to controls who had no gel treatment. However, the change for the medium treatment dose group is in a harmful or negative direction compared to controls. Over one year, the average change in attachment loss for subjects in the medium dose treatment group was 0.176 units larger than that for control subjects. This means that that on average subjects in the medium dose treatment group had a significantly greater increase in attachment loss over the year compared with controls. This indicates that the medium dose treatment level may have harmful or negative affects on gum health, specifically attachment loss. However, we saw no significant effect of the medium dose treatment (or any of the treatment levels) on change in pocket depth over the year. The limitations of the results presented here are that there was no way of verifying whether subjects actually used the treatment twice daily as it was prescribed. Additionally, the trend we saw with increasing dropout for increasing treatment dose level may indicate that some aspect of the higher treatment dose levels (potentially harmful results or negative side effects) led subjects in the medium and high dose groups to dropout of the study. If this is the case, it is also possible that subjects in these higher dose groups that remained in the study did not use the treatment as directed throughout the year or discontinued use completely for similar reasons to subjects that dropped out. Also in regards to the missing data, after removing subjects who did not return for followup the sample sizes across groups were no longer equal, with the high dose group being especially small. If possible, it

would be helpful to conduct some sort of survey to explore the reasons for dropout and to explore how well participants adhered to the prescribed treatment plan. It would also be useful to collect data on any reported side effects that subjects experienced in different treatment groups. This would allow the investigators to better understand dropout rates or lack of adherence. This information could also help inform the results we saw in this data and would be useful for designing future studies of a similar nature.

Reproducible Research Information

The code for this report can be found in the "Code" folder within the "Project0" folder in my github directory. My github directory is as follows: BIOS6623-UCD/bios6623-micpalumbo. A description of each of the files in my "Project0" folder including this report can be found in the readme.md file within the "Project0" folder. The data was stored in the following location: "C:/Users/Michaela Palumbo/Documents/CU AMC Fall 2017/BIOS6623/Project0_dental_data.csv" and the code I used to read in the data is given below and can be found in the file "Proj0descriptives.R" located in the code folder described above.

```
dat <- read.csv("~/Documents/CU AMC Fall 2017/BIOS6623/Project0_dental_data.csv")
```

Appendix

Table 1: Summary of descriptive statistics (excluding subjects lost to followup)

	control	placebo	low	med	high
n	23	23	21	20	16
age (mean (sd))	51.42 (10.16)	47.18 (8.86)	51.47 (10.26)	48.63 (9.87)	52.61 (10.71)
sex = female (%)	16 (69.6)	13 (56.5)	12 (57.1)	13 (65.0)	13 (81.2)
race (%)					
native american	1 (4.3)	0 (0.0)	1 (4.8)	0 (0.0)	1 (6.2)
african american	1 (4.3)	1 (4.3)	3 (14.3)	0 (0.0)	1 (6.2)
asian	1 (4.3)	1 (4.3)	0 (0.0)	1 (5.0)	0 (0.0)
white	20 (87.0)	21 (91.3)	17 (81.0)	19 (95.0)	14 (87.5)
smoker = yes (%)	7 (30.4)	10 (43.5)	6 (28.6)	8 (42.1)	5 (31.2)
sites (mean (sd))	153.91 (11.24)	159.65 (10.66)	161.71 (8.16)	153.50 (17.12)	158.25 (8.73)
attachbase (mean (sd))	2.55 (0.65)	1.83 (0.66)	2.10 (1.09)	2.24 (0.67)	2.31 (1.01)

	control	placebo	low	med	high
atdiff (mean (sd))	-0.22 (0.28)	-0.09 (0.24)	-0.02 (0.27)	-0.01 (0.23)	-0.16 (0.33)
pdbase (mean (sd))	3.29 (0.49)	3.10 (0.39)	3.22 (0.64)	3.05 (0.42)	3.18 (0.28)
pddiff (mean (sd))	-0.34 (0.23)	-0.35 (0.28)	-0.21 (0.28)	-0.20 (0.27)	-0.38 (0.24)

Figure 1

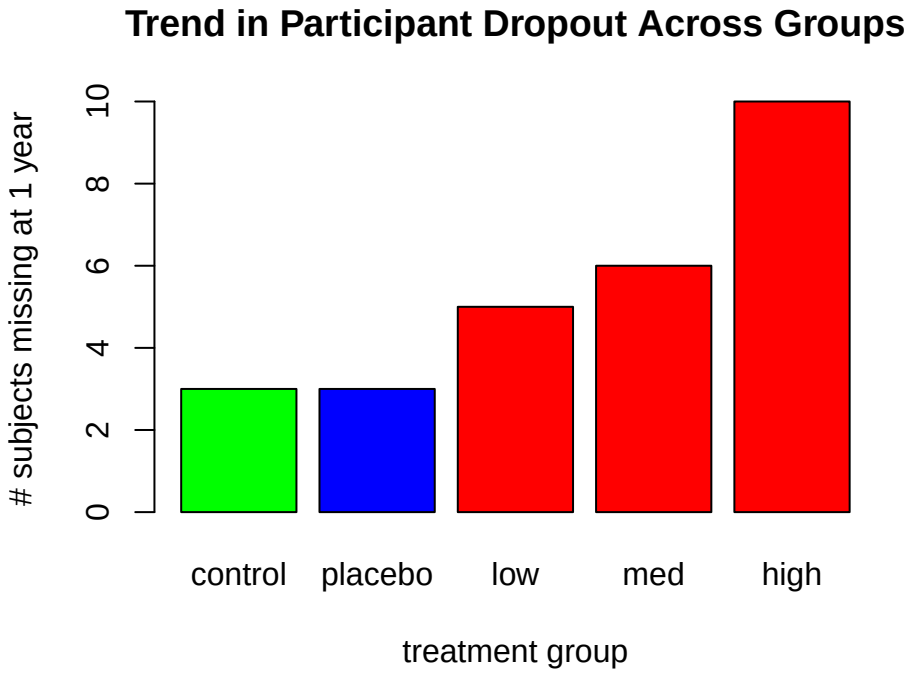

Figure 2

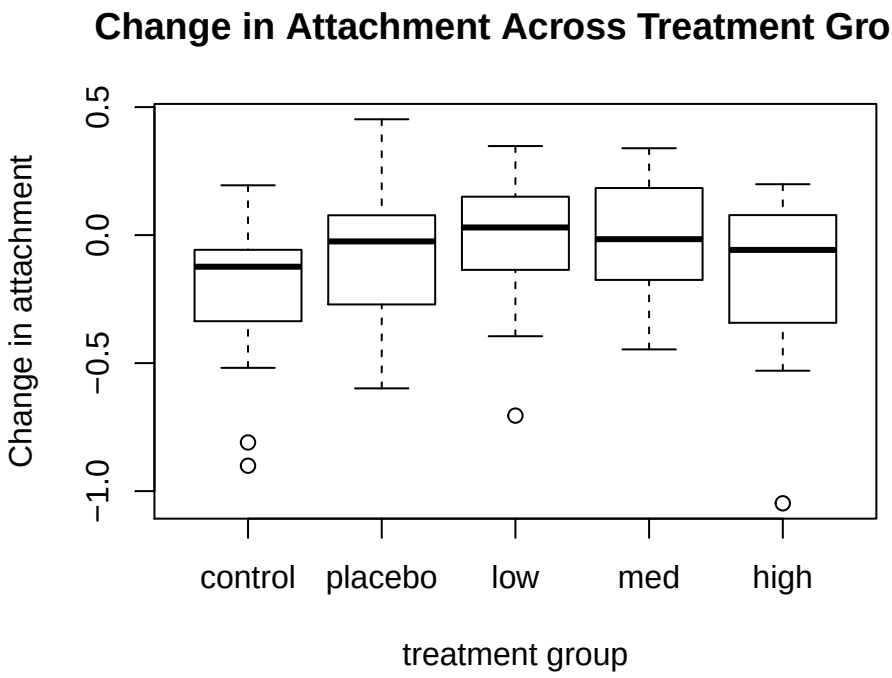

Figure 3