Technical Appendix

Name: Braden Tam
Course: STA490

April 5th, 2019

Introduction

The following technical appendix contains the full code used for the paper "Having Sleep Conditions Can Give You Higher Levels of Stress". This document will outline the techniques and methodology used to conduct an exploratory data analysis and produce the results found in the paper. Statistical methods including linear mixed models, logistic regression models, and two-sample t-tests will be explained in detail as to why they were used for this analysis.

Libraries

```
library(readxl)
library(tidyverse)
library(lmerTest)
library(ggthemes)
#devtools::install_github('cttobin/ggthemr')
library(ggthemr)
library(table1)
library(reshape)
library(knitr)

ggthemr('flat')
#ggthemr_reset()
```

Reading in the data

This chunk simply reads in the data from an excel file.

```
SAQdataraw <- read_excel("cleandata.xlsx")
static = read_excel("cleandata.xlsx", sheet = "Static data")
t1 = read_excel("cleandata.xlsx", sheet = "Dynamic time 1")
t2 = read_excel("cleandata.xlsx", sheet = "Dynamic time 2")
t3 = read_excel("cleandata.xlsx", sheet = "Dynamic time 3 update")
```

Creating SAQ Functions

SAQ is given to us as both continuous and binary variables, however for the binary variable the 2 groups are very disproportionate in sample size. The following functions are for splitting up SAQ into 3 categories in order to have more even group sizes. The first cutoff (15) is based off the standard split while the second cutoff (25) is based off splitting the rest of the data equally.

```r
#0-15, 15-30, 31-50
splitSAQ3 <- function(saq){
  i <- 1
  while(i <= length(t1$SAQGlobal)){
    if(!is.na(saq[i])){
      if(saq[i] <= 15){
        saq[i] <- 1
      } else if (15 < saq[i] & saq[i] <= 24){
        saq[i] <- 2
      } else if (saq[i]>=25){
          saq[i] <- 3
      }
    }
        i <- i+1
    }
     return(saq)
}
```

Setting up the dataframe

This chunk is for creating a data frame for all of the relevent data that will be used.

```r
t1data = tibble(Time = 1,
                ID = t1$ID,
                SAQcont = t1$`SAQGlobal`,
                SAQ2 = t1$`SAQGlobal+`,
                SAQ3 = splitSAQ3(t1$`SAQGlobal`),
                Insomnia = t1$`SAQInsom+`,
                OSA = t1$`SAQOSA+`,
                NRS = t1$`SAQNRS+`,
                DSPD =  t1$`SAQSched+`,
                EDS = t1$`SAQEDS+`,
                RLS = t1$`SAQRestless+`,
                Distress = t1$`K6-Tot`,
                Anxiety = t1$`BAI-Tot`,
                Depression = t1$`CESD-Tot`,
                Attendance = t1$Attendance,
                Sex = static$Gender,
                `SAQ present` = !is.na(t1$SAQGlobal))

#T2 data with T1 SAQ data
t2data = tibble(Time = 2, ID = t1$ID,
                SAQcont = t1$`SAQGlobal`,SAQ2 = t1$`SAQGlobal+`,
                SAQ3 = splitSAQ3(t1$`SAQGlobal`), Insomnia = t1$`SAQInsom+`,
                OSA = t1$`SAQOSA+`, NRS = t1$`SAQNRS+`, DSPD =  t1$`SAQSched+`,
                EDS = t1$`SAQEDS+`, RLS = t1$`SAQRestless+`, Distress = t2$`K6-Tot`,
                Anxiety = t2$`BAI-Tot`, Depression = t2$`CESD-Tot`,
                Attendance = t1$Attendance, Sex = static$Gender,
                `SAQ present` = !is.na(t1$SAQGlobal))

#T2 data with T2 SAQ data
t2saqdata = tibble(Time = 2.1, ID = t1$ID,
```

```
                  SAQcont = t2$`SAQGlobal`, SAQ2 = t2$`SAQGlobal+`,
                  SAQ3 = splitSAQ3(t2$`SAQGlobal`), Insomnia = t2$`SAQInsom+`,
                  OSA = t2$`SAQOSA+`, NRS = t2$`SAQNRS+`, DSPD =  t2$`SAQSched+`,
                  EDS = t2$`SAQEDS+`, RLS = t2$`SAQRestless+`, Distress = t2$`K6-Tot`,
                  Anxiety = t2$`BAI-Tot`, Depression = t2$`CESD-Tot`,
                  Attendance = t1$Attendance, Sex = static$Gender,
                  `SAQ present` = !is.na(t1$SAQGlobal))

t12data = full_join(t1data, t2data) %>%
  arrange(ID)
fulldata = full_join(t12data, t2saqdata) %>% #combines the dataframes
  arrange(ID)
```

The purpose of this chunk is to extract the data of interest. Specifically, this will find only the patients that have both time 1 and 2 data. After extracting the patients that have full data (both time 1 and 2), we obtained their stress levels to be used in the interaction later on.

Why don't we just apply models to the full data? The models that will be used below will exclude missing values (NAs) which would lead to some patients having only 1 data point. Having one data point is completely trivial because we want to measure how their levels of stress improve from the program, thus requiring both time 1 and 2 data. With only 2 data points of interest, we don't want to impute any missing data because in the context of our problem, there is no reasonable way to impute. Missing data will be explored in a subsequent section in this technical appendix.

```
fullglobal = !is.na(t2$`K6-Tot`) & !is.na(t1$SAQGlobal)#full questionnaires and global
fullinsom = !is.na(t2$`K6-Tot`) & !is.na(t1$`SAQInsom+`) & t1$`SAQInsom+` == 1
fullosa = !is.na(t2$`K6-Tot`) & !is.na(t1$`SAQOSA+`) & t1$`SAQOSA+`== 1


#full questionnaire and others (at least 1 and not including insomnia/osa) (not including 0)
fullother = !is.na(t2$`K6-Tot`) & (!is.na(t1$`SAQNRS+`) & t1$`SAQNRS+` == 1 |
            !is.na(t1$`SAQSched+`) & t1$`SAQSched+` == 1 | !is.na(t1$`SAQEDS+`) &
            t1$`SAQEDS+` == 1 | !is.na(t1$`SAQRestless+`) & t1$`SAQRestless+` == 1) &
            !is.na(t1$`SAQInsom+`) & t1$`SAQInsom+` == 0 & !is.na(t1$`SAQOSA+`) &
            t1$`SAQOSA+` == 0

fullnone = !is.na(t1$`K6-Tot`) & (!is.na(t1$`SAQNRS+`) & t1$`SAQNRS+` == 0 &
            !is.na(t1$`SAQSched+`) & t1$`SAQSched+` == 0 & !is.na(t1$`SAQEDS+`) &
            t1$`SAQEDS+` == 0 & !is.na(t1$`SAQRestless+`) & t1$`SAQRestless+` == 0) &
            !is.na(t1$`SAQInsom+`) & t1$`SAQInsom+` == 0 & !is.na(t1$`SAQOSA+`) &
            t1$`SAQOSA+` == 0

ID = t1$ID[fullglobal]
time = c(rep(1, length(ID)), rep(2, length(ID)))
ID = rep(ID,2)
SAQcont = t1$SAQGlobal[fullglobal]
SAQcont = rep(SAQcont, 2)
SAQ2 = t1$`SAQGlobal+`[fullglobal]
SAQ2 = rep(SAQ2, 2)
SAQ3 = splitSAQ3(SAQcont)


distress1 = t1$`K6-Tot`[fullglobal]
distress2 = t2$`K6-Tot`[fullglobal]
distress = c(distress1, distress2)
```
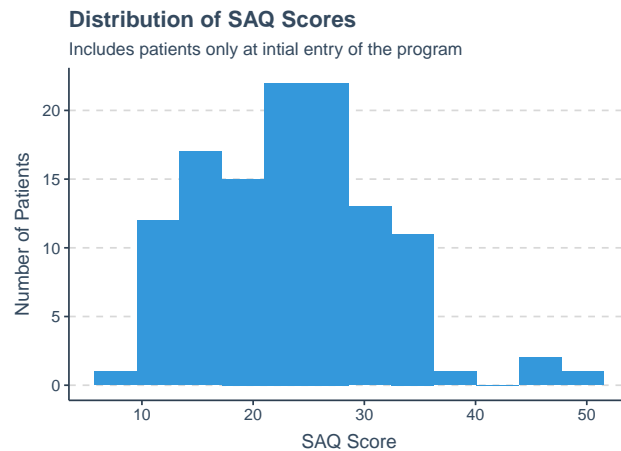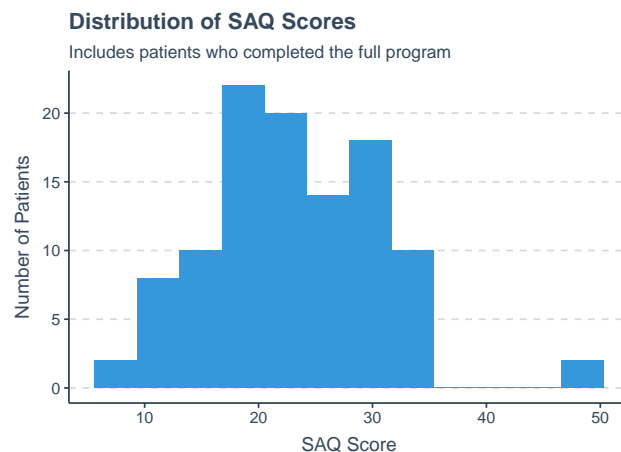
```
anxiety1 = t1$`BAI-Tot`[fullglobal]
anxiety2 = t2$`BAI-Tot`[fullglobal]
anxiety = c(anxiety1, anxiety2)
depression1 = t1$`CESD-Tot`[fullglobal]
depression2 = t2$`CESD-Tot`[fullglobal]
depression = c(depression1, depression2)
SAQdata = tibble(time, ID, SAQcont, SAQ2, SAQ3, distress, anxiety, depression)

distress1_osa = t1$`K6-Tot`[fullosa]
distress2_osa = t2$`K6-Tot`[fullosa]
distress1_insom = t1$`K6-Tot`[fullinsom]
distress2_insom = t2$`K6-Tot`[fullinsom]
distress1_other = t1$`K6-Tot`[fullother]
distress2_other = t2$`K6-Tot`[fullother]
distress1_none = t1$`K6-Tot`[fullnone]

anxiety1_osa = t1$`BAI-Tot`[fullosa]
anxiety2_osa = t2$`BAI-Tot`[fullosa]
anxiety1_insom = t1$`BAI-Tot`[fullinsom]
anxiety2_insom = t2$`BAI-Tot`[fullinsom]
anxiety1_other = t1$`BAI-Tot`[fullother]
anxiety2_other = t2$`BAI-Tot`[fullother]

depression1_osa = t1$`CESD-Tot`[fullosa]
depression2_osa = t2$`CESD-Tot`[fullosa]
depression1_insom = t1$`CESD-Tot`[fullinsom]
depression2_insom = t2$`CESD-Tot`[fullinsom]
depression1_other = t1$`CESD-Tot`[fullother]
depression2_other = t2$`CESD-Tot`[fullother]
```

Table 1

This chunk creates a simple table of summary statistics of the population split by time and sex.

```
table1(~ `SAQ present` +  Anxiety + Depression + Distress | Time*Sex, data=t12data)
```

```
## [1] "<table class=\"Rtable1\">\n<thead>\n<tr>\n<th class=\"grouplabel\"></th>\n<th colspan=\"2\" cla
```

Distrubtion of SAQ

We want to see how the distribution of SAQ scores change over time so these histogram plots capture the distribution of SAQ scores.

```
#SAQ for patients with just time 1
ggplot(t1) +
  geom_histogram(aes(SAQGlobal),  bins = 12) +
  labs(title = "Distribution of SAQ Scores",
       x = "SAQ Score", y = "Number of Patients",
       subtitle = "Includes patients only at intial entry of the program") +
  theme_hc()
```

Distribution of SAQ Scores

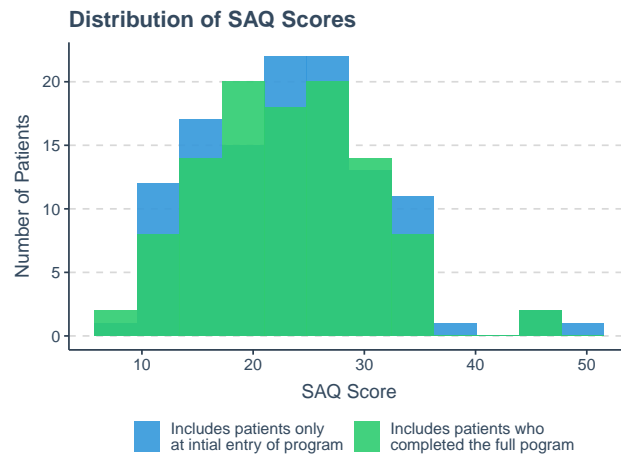Includes patients only at intial entry of the program


```
#SAQ for patients with both time 1 & 2
ggplot(SAQdata) +
  geom_histogram(aes(SAQcont), bins = 12) +
  labs(x = "SAQ Score", y = "Number of Patients",
       title = "Distribution of SAQ Scores",
       subtitle = "Includes patients who completed the full program") +
  theme_hc()
```
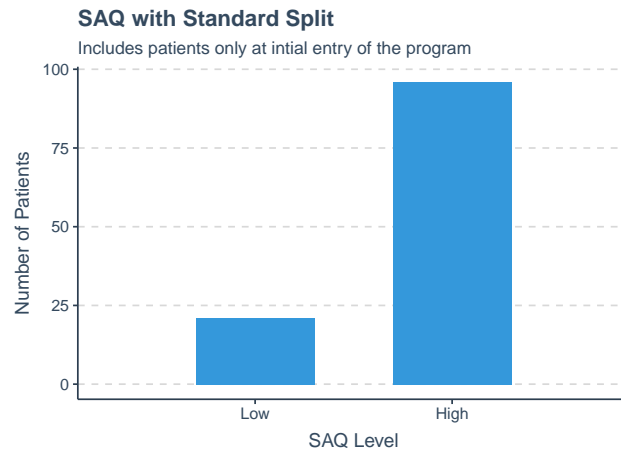
Distribution of SAQ Scores

Includes patients who completed the full program


```
SAQtime1 = tibble(SAQ = t1$SAQGlobal)
SAQboth = tibble(SAQ = append(SAQdata$SAQcont, rep(NA,
                length(t1$SAQGlobal) - length(SAQdata$SAQcont))))
SAQtime1$graph = "Includes patients only at intial entry of program"
SAQboth$graph = "Includes patients who completed the full pogram"

combhist = rbind(SAQtime1, SAQboth)
ggplot(combhist, aes(SAQ, fill = graph)) +
  geom_histogram(aes(SAQ, group = graph), bins = 12, position = "identity", alpha = .9) +
  labs(x = "SAQ Score", y = "Number of Patients", title = "Distribution of SAQ Scores") +
  theme_hc() + scale_fill_discrete(name = NULL,
                       labels = c("Includes patients only \nat intial entry of program",
                       "Includes patients who \ncompleted the full pogram"))
```

Distribution of SAQ Scores

These plots show how splitting up SAQ into 3 categories makes the groups a bit closer in size.

```
#SAQ for patients with just time 1
ggplot(t1) +
  geom_bar(aes(`SAQGlobal+`), width = 0.6) +
  labs(x = "SAQ Level", y = "Number of Patients", title = "SAQ with Standard Split",
       subtitle = "Includes patients only at intial entry of the program") +
  scale_x_discrete("SAQ Level",limits=c(0,1), labels = c("Low", "High"))+ theme_hc()
```
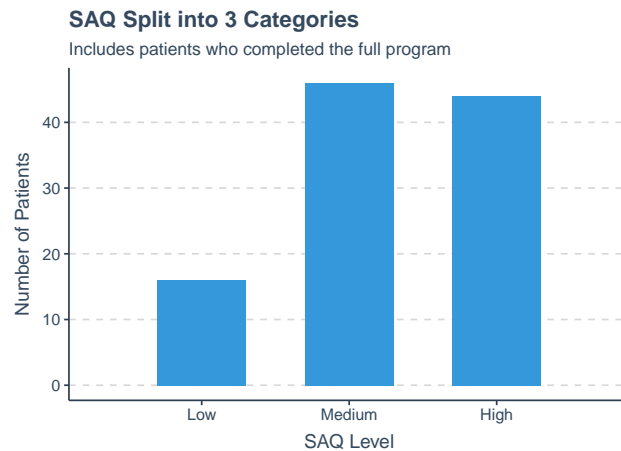
SAQ with Standard Split

Includes patients only at intial entry of the program


```
#SAQ for patients with both time 1 & 2
ggplot(SAQdata) +
  geom_bar(aes(SAQ2), width = 0.6) +
  labs(x = "SAQ Level", y = "Number of Patients", title = "SAQ with Standard Split",
       subtitle = "Includes patients who completed the full program") + theme_hc() +
  scale_x_discrete("SAQ Level",limits=c(0,1), labels = c("Low", "High"))
```

SAQ with Standard Split

Includes patients who completed the full program


```
#SAQ for patients with just time 1
ggplot(t1) +
  geom_bar(aes(splitSAQ3(t1$SAQGlobal)), width = 0.6) +
  labs(x = "SAQ Category", y = "Number of Patients", title = "SAQ Split into 3 Categories",
       subtitle = "Includes patients only at intial entry of the program") + theme_hc() +
  scale_x_discrete("SAQ Level",limits=c(1,2,3), labels = c("Low", "Medium", "High"))
```
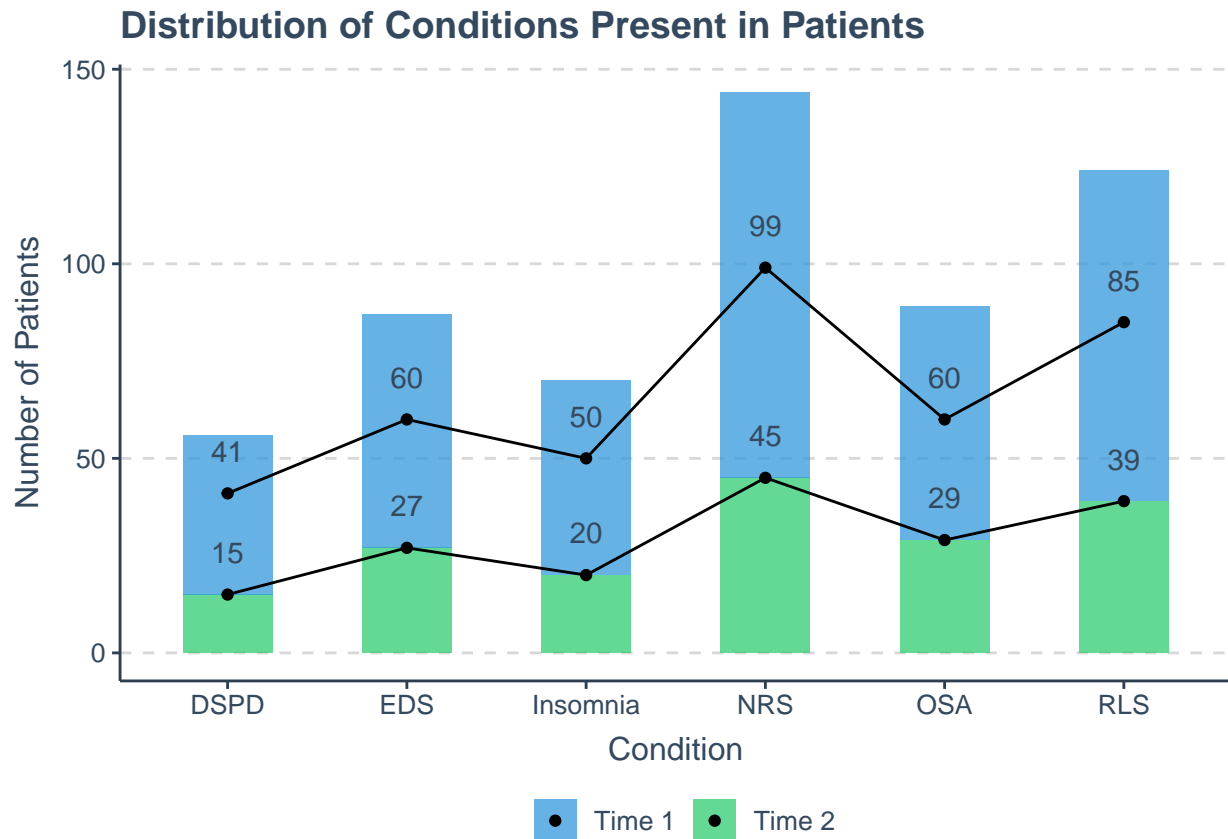
SAQ Split into 3 Categories

Includes patients only at intial entry of the program


```
#SAQ for patients with both time 1 & 2
ggplot(SAQdata) +
  geom_bar(aes(SAQ3), width = 0.6) +
  labs(x = "SAQ Category", y = "Number of Patients", title = "SAQ Split into 3 Categories",
       subtitle = "Includes patients who completed the full program") + theme_hc() +
  scale_x_discrete("SAQ Level",limits=c(1,2,3), labels = c("Low", "Medium","High"))
```

SAQ Split into 3 Categories
Includes patients who completed the full program

Distribution of Sleep Conditions

This chunk joins two dataframes in order to plot 2 bar graphs together.

```r
# Time 2 conditions using SAQ time 1 data
t1conditions = fulldata %>%
  filter(!is.na(Distress), Time == 1) %>%
  select(EDS, Insomnia, NRS, OSA, RLS, DSPD) %>%
  summarise(EDS = sum(na.omit(EDS)), Insomnia = sum(na.omit(Insomnia)),
            NRS = sum(na.omit(NRS)), OSA = sum(na.omit(OSA)),
            RLS = sum(na.omit(RLS)), DSPD = sum(na.omit(DSPD))) %>%
  gather(Condition, 'Time 1')

# Time 2 conditions using SAQ Time 2 data
t2conditions = fulldata %>%
  filter(!is.na(Distress), Time == 2) %>%
  select(Insomnia, OSA, NRS, DSPD, EDS, RLS) %>%
  summarise(EDS = sum(na.omit(EDS)), Insomnia = sum(na.omit(Insomnia)),
            NRS = sum(na.omit(NRS)), OSA = sum(na.omit(OSA)),
            RLS = sum(na.omit(RLS)), DSPD = sum(na.omit(DSPD))) %>%
  gather(Condition2, 'Time 2')

joint = cbind(t1conditions, t2conditions) %>%
  select(-Condition2)


jointplot = melt(joint)
ggplot(data = jointplot, aes(x=factor(Condition), y=value,
                             group=factor(variable), fill=factor(variable))) +
  geom_bar(position = "stack", stat = "identity", width = 0.5, alpha = .75)  +
  labs(x = "Condition", y = "Number of Patients",
       title = "Distribution of Conditions Present in Patients") +
  theme_hc() + geom_line(color = "black") + geom_point(color = "black") +
  scale_fill_discrete(name = NULL, labels = c("Time 1", "Time 2")) +
  geom_text(aes(label=value), vjust=-1.5)
```

Distribution of Conditions Present in Patients

```
joint = joint %>%
  mutate(Proportion = `Time 2`/`Time 1`) %>%
  format(digits=2)

kable(joint, align = c("l", "c", "c", "c"))
```

Condition	Time 1	Time 2	Proportion
EDS	60	27	0.45
Insomnia	50	20	0.40
NRS	99	45	0.45
OSA	60	29	0.48
RLS	85	39	0.46
DSPD	41	15	0.37

Observation:
We can note that the distribution of conditions remains relatively constant throughout the program indicating the program is not improving any specific condition.

Interaction Plots

These plots show the interaction of SAQ and sleep disorders with time for anxiety, depression, and distress which helps us give a visual indication of answers to our research question. We must first create a dataframe that groups the outcome variable by categories in order to examine the interaction.

```r
x = c(1,2,1,2,1,2)
group = c(1,1,2,2,3,3)

means = SAQdata %>%
  group_by(time, SAQ3) %>%
  summarise(Anxiety = mean(na.omit(anxiety)),
            Depression = mean(na.omit(depression)),
            Distress = mean(na.omit(distress)))

#anxiety by sleep disorder
values = c(mean(na.omit(anxiety1_insom)), mean(na.omit(anxiety2_insom)),
           mean(na.omit(anxiety1_osa)), mean(na.omit(anxiety2_osa)),
           mean(na.omit(anxiety1_other)), mean(na.omit(anxiety2_other)))
g = tibble(x, group, values)

ggplot() +
  geom_line(data = g, aes(x = x, y = values, colour = factor(group)), size = 0.75) +
  labs(x = "Time Period", y = "Anxiety Levels", title = "Anxiety over time split by Sleep Disorder",
       caption = "Sample size of people with 0 disorders = 3",
       subtitle = "*includes at least one of EDS, NRS, RLS, or DSPD") +
  scale_colour_manual("Sleep Disorder",values=c("blue","green","red"),breaks=c(1, 2, 3, 4),
                      labels = c("OSA", "Insomnia", "Other*", "No Sleep Disorder")) +
  scale_x_continuous("Time Period", limits=c(1,2), breaks=seq(1,2, by = 1)) + theme_hc()  +
  theme(legend.position = "right")+ annotate("text", x = 1.95, y =15.7, label = "n = 28") +
  annotate("text", x = 1.95, y =14.5, label = "n = 20") +
  annotate("text", x = 1.95, y =11.45, label = "n = 21") +
  geom_text(aes(label=), vjust=-1.5)
```
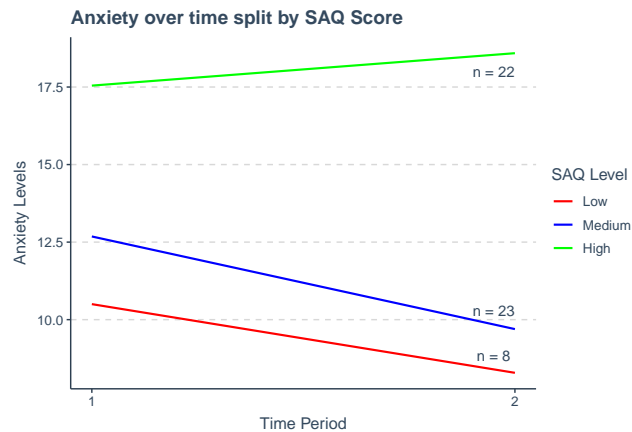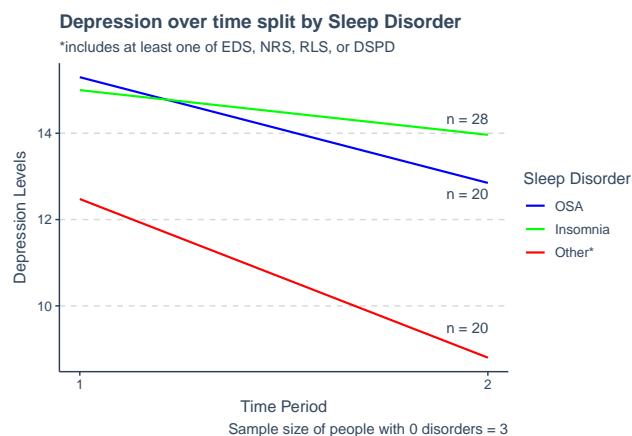

Anxiety over time split by Sleep Disorder
*includes at least one of EDS, NRS, RLS, or DSPD

Sample size of people with 0 disorders = 3

```r
#Anxiety by SAQ3
ggplot() +
  geom_line(data = means, aes(x = time, y = Anxiety, colour = factor(SAQ3)), size = 0.75) +
  labs(x = "Time Period", y = "Anxiety Levels", title = "Anxiety over time split by SAQ Score") +
  scale_colour_manual("SAQ Level", values=
        c("red", "blue", "green"), breaks=c(1,2,3), labels = c("Low", "Medium", "High")) +
  scale_x_continuous("Time Period",limits=c(1,2),breaks=seq(1,2, by = 1)) + theme_hc() +
  theme(legend.position = "right") + annotate("text", x = 1.95, y =18, label = "n = 22") +
  annotate("text", x = 1.95, y =10.3, label = "n = 23") +
  annotate("text", x = 1.95, y =8.85, label = "n = 8")
```

Anxiety over time split by SAQ Score


```
#depression by sleep disorder
values = c(mean(na.omit(depression1_insom)), mean(na.omit(depression2_insom)),
          mean(na.omit(depression1_osa)), mean(na.omit(depression2_osa)),
          mean(na.omit(depression1_other)), mean(na.omit(depression2_other)))
g = tibble(x, group, values)

ggplot() +
  geom_line(data = g, aes(x = x, y = values, colour = factor(group)), size = 0.75) +
  labs(x = "Time Period", y = "Depression Levels", title = "Depression over time split by Sleep Disorder",
       caption = "Sample size of people with 0 disorders = 3",
       subtitle = "*includes at least one of EDS, NRS, RLS, or DSPD") +
  scale_colour_manual("Sleep Disorder", values= c("blue","green", "red"),
      breaks=c(1, 2, 3, 4), labels = c("OSA", "Insomnia", "Other*", "No Sleep Disorder")) +
  scale_x_continuous("Time Period", limits=c(1,2), breaks=seq(1,2, by = 1)) + theme_hc()  +
  theme(legend.position = "right") + annotate("text", x = 1.95, y =14.35, label = "n = 28") +
  annotate("text", x = 1.95, y =12.6, label = "n = 20") +
  annotate("text", x = 1.95, y =9.5, label = "n = 20")
```

Depression over time split by Sleep Disorder
*includes at least one of EDS, NRS, RLS, or DSPD


```
#depression by SAQ3
ggplot() +
  geom_line(data = means, aes(x = time, y = Depression, colour = factor(SAQ3)), size = 0.75) +
  labs(x = "Time Period", y = "Depression Levels", title = "Depression over time split by SAQ Score") +
  scale_colour_manual("SAQ Level", values=c("red", "blue", "green"),
        breaks=c(1,2,3), labels = c("Low", "Medium", "High")) + scale_x_continuous("Time Period",
        limits=c(1,2), breaks=seq(1,2, by = 1))+ theme_hc() + theme(legend.position = "right") +
  annotate("text", x = 1.95, y =18.2, label = "n = 22") +
```

11

```
annotate("text", x = 1.95, y =8.25, label = "n = 23") +
annotate("text", x = 1.95, y =7, label = "n = 8")
```

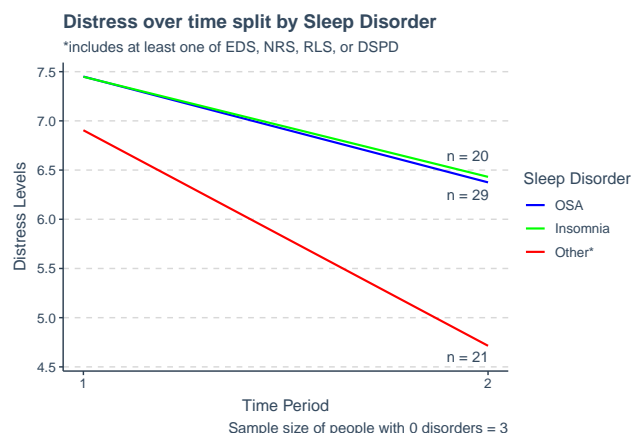**Depression over time split by SAQ Score**


```
#distress by sleep disorder
values = c(mean(na.omit(distress1_insom)), mean(na.omit(distress2_insom)),
          mean(na.omit(distress1_osa)), mean(na.omit(distress2_osa)),
          mean(na.omit(distress1_other)), mean(na.omit(distress2_other)))
g = tibble(x, group, values)

ggplot() +
  geom_line(data = g, aes(x = x, y = values, colour = factor(group)), size = 0.75) +
  labs(x = "Time Period", y = "Distress Levels", title = "Distress over time split by Sleep Disorder",
      caption = "Sample size of people with 0 disorders = 3",
      subtitle = "*includes at least one of EDS, NRS, RLS, or DSPD") + scale_colour_manual(
      "Sleep Disorder", values= c("blue","green", "red"),breaks=c(1, 2, 3, 4), labels =
       c("OSA", "Insomnia", "Other*", "No Sleep Disorder")) + scale_x_continuous("Time Period",
       limits=c(1,2), breaks=seq(1,2, by = 1)) + geom_label() + theme_hc()  +
  theme(legend.position = "right")+ annotate("text", x = 1.95, y =6.25, label = "n = 29") +
  annotate("text", x = 1.95, y =6.65, label = "n = 20") +
  annotate("text", x = 1.95, y =4.6, label = "n = 21")
```

Distress over time split by Sleep Disorder
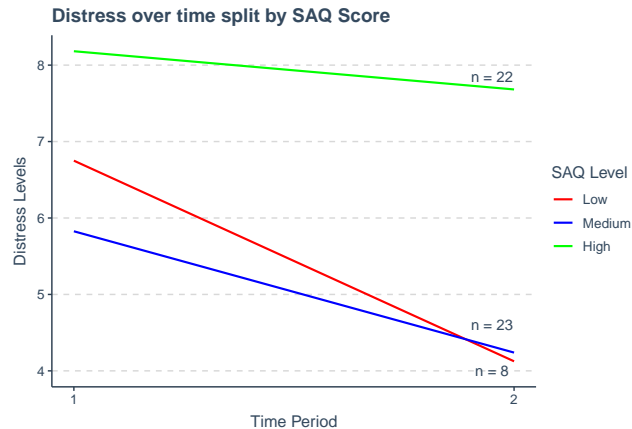*includes at least one of EDS, NRS, RLS, or DSPD


```
#Distress by SAQ3
ggplot() +
  geom_line(data = means, aes(x = time, y = Distress, colour = factor(SAQ3)), size = 0.75) +
  labs(x = "Time Period", y = "Distress Levels", title = "Distress over time split by SAQ Score") +
  scale_colour_manual("SAQ Level", values = c("red", "blue", "green"),
```

```
     breaks=c(1,2,3), labels = c("Low", "Medium", "High")) + scale_x_continuous("Time Period",
     limits=c(1,2), breaks=seq(1,2, by = 1)) + theme_hc() + theme(legend.position = "right") +
  annotate("text", x = 1.95, y =7.85, label = "n = 22") +
  annotate("text", x = 1.95, y =4.6, label = "n = 23") +
  annotate("text", x = 1.95, y =4, label = "n = 8")
```


Distress over time split by SAQ Score

Observations:

Looking at the interaction plots, we see what we expect which is that patients with high SAQ scores will not benefit as much as people with low SAQ scores. A shallower (flatter) slope indicates less improvement, thus high SAQ scores have a shallower slope. For stress levels split by sleep disorders, OSA and insomnia seem to hinder patients' abilities to improve their stress levels compared to other disorders.

Creating the dataframe for our model

The following 3 chunks creates the dataframes for anxiety, depression, and distress. Only patients who have both time 1 and 2 data for the levels of stress are included. In order to use our model, we must create 1 dataframe per outcome variable that contains data for both times. These chunks also create new variables including a variable for the number of disorders a patient has and a variable for the specific disorder (but aggregating NRS, DSPD, EDS, RLS into an "other" category because insomnia and OSA are theorized to have a larger impact than the rest of the sleep conditions). Treating sleep disorders in multiple ways allows us to have different interpretations of our results.

```
fullanxiety = !is.na(t1$`BAI-Tot`) & !is.na(t2$`BAI-Tot`)
ID = t1$ID[fullanxiety]
Time = c(rep(1, length(ID)), rep(2, length(ID)))
ID = rep(ID,2)
PTSD = rep(t1$`PTSD+`[fullanxiety],2)
Stress = rep(t1$`INTChronicStress+`[fullanxiety],2)
SAQcont = rep(t1$SAQGlobal[fullanxiety],2)
SAQ2 = rep(t1$`SAQGlobal+`[fullanxiety],2)
SAQ3 = splitSAQ3(SAQcont)
OSA = rep(t1$`SAQOSA+`[fullanxiety],2)
Insomnia = rep(t1$`SAQInsom+`[fullanxiety],2)
NRS = rep(t1$`SAQNRS+`[fullanxiety],2)
DSPD = rep(t1$`SAQSched+`[fullanxiety],2)
EDS = rep(t1$`SAQEDS+`[fullanxiety],2)
RLS = rep(t1$`SAQRestless+`[fullanxiety],2)
Disorder = 0
Numdisorder = 0
```

```r
anxiety1 = t1$`BAI-Tot`[fullanxiety];anxiety2 = t2$`BAI-Tot`[fullanxiety]
Anxiety = c(anxiety1, anxiety2)
dataanxiety = tibble(ID, PTSD, Stress, Time, SAQcont, SAQ2, SAQ3, OSA, Insomnia,
                     NRS, DSPD, EDS, RLS, Disorder, Numdisorder, Anxiety)

#for Disorder 1= OSA, 2= Insomnia, 3= Other
i = 0
for(i in 1:nrow(dataanxiety)){
  if(dataanxiety$OSA[i] == 1 & !is.na(dataanxiety$OSA[i])){
    dataanxiety$Disorder[i] = 1
  }
  else if(dataanxiety$Insomnia[i] == 1 & !is.na(dataanxiety$Insomnia[i])){
    dataanxiety$Disorder[i] = 2
  }
  else if(is.na(dataanxiety$OSA[i]) & is.na(dataanxiety$Insomnia[i]) &
          is.na(dataanxiety$NRS[i]) & is.na(dataanxiety$DSPD[i]) &
          is.na(dataanxiety$EDS[i]) & is.na(dataanxiety$RLS[i])){
    dataanxiety$Disorder[i] = NA
  }
  else(dataanxiety$Disorder[i] = 3)
}


#for Numdisorder
i = 0
k = 0 #num of Disorders
for(i in 1:nrow(dataanxiety)){
  l = 0 #num of Disorders added up
  for(j in 8:13){
    if(is.na(dataanxiety[i,j])){
      j = j + 1
      }
  else{
    l = l + dataanxiety[i,j]
     j = j + 1
     }
   }
  if(l >= 3){
    dataanxiety$Numdisorder[i] = 3
  }
  else if(is.na(dataanxiety$OSA[i]) & is.na(dataanxiety$Insomnia[i]) &
          is.na(dataanxiety$NRS[i]) & is.na(dataanxiety$DSPD[i]) &
          is.na(dataanxiety$EDS[i]) & is.na(dataanxiety$RLS[i])){
    dataanxiety$Numdisorder[i] = NA
  }
  else{
    dataanxiety$Numdisorder[i] = as.numeric(l)
  }
}

fulldepression = !is.na(t1$`CESD-Tot`) & !is.na(t2$`CESD-Tot`)
ID = t1$ID[fulldepression]
Time = c(rep(1, length(ID)), rep(2, length(ID)))
ID = rep(ID,2)
PTSD = rep(t1$`PTSD+`[fulldepression],2)
```

```r
  Stress = rep(t1$`INTChronicStress+`[fulldepression],2)
  SAQcont = rep(t1$SAQGlobal[fulldepression],2)
  SAQ2 = rep(t1$`SAQGlobal+`[fulldepression],2)
  SAQ3 = splitSAQ3(SAQcont)
  OSA = rep(t1$`SAQOSA+`[fulldepression],2)
  Insomnia = rep(t1$`SAQInsom+`[fulldepression],2)
  NRS = rep(t1$`SAQNRS+`[fulldepression],2)
  DSPD = rep(t1$`SAQSched+`[fulldepression],2)
  EDS = rep(t1$`SAQEDS+`[fulldepression],2)
  RLS = rep(t1$`SAQRestless+`[fulldepression],2)
  Disorder = 0
  Numdisorder = 0
  depression1 = t1$`CESD-Tot`[fulldepression];depression2 = t2$`CESD-Tot`[fulldepression]
  Depression = c(depression1, depression2)
  datadepression = tibble(ID, PTSD, Stress, Time, SAQcont, SAQ2, SAQ3, OSA, Insomnia,
                          NRS, DSPD, EDS, RLS, Disorder, Numdisorder, Depression)


  #for Disorder 1= OSA, 2= Insomnia, 3= LOW SAQ, 4= Other
  i = 0
  for(i in 1:nrow(datadepression)){
    if(datadepression$OSA[i] == 1 & !is.na(datadepression$OSA[i])){
      datadepression$Disorder[i] = 1
    }
    else if(datadepression$Insomnia[i] == 1 & !is.na(datadepression$Insomnia[i])){
      datadepression$Disorder[i] = 2
    }
    else if(is.na(datadepression$OSA[i]) & is.na(datadepression$Insomnia[i]) &
            is.na(datadepression$NRS[i]) & is.na(datadepression$DSPD[i]) &
            is.na(datadepression$EDS[i]) & is.na(datadepression$RLS[i])){
      datadepression$Disorder[i] = NA
    }
    else(datadepression$Disorder[i] = 3)
  }


  #for Numdisorder
  i = 0
  k = 0 #num of Disorders
  for(i in 1:nrow(datadepression)){
    l = 0 #num of Disorders added up
    for(j in 8:13){
      if(is.na(datadepression[i,j])){
        j = j + 1
          }
    else{
      l = l + datadepression[i,j]
      j = j + 1
      }
    }
    if(l >= 3){
      datadepression$Numdisorder[i] = 3
    }
    else if(is.na(datadepression$OSA[i]) & is.na(datadepression$Insomnia[i]) &
            is.na(datadepression$NRS[i]) & is.na(datadepression$DSPD[i]) &
```

```r
           is.na(datadepression$EDS[i]) & is.na(datadepression$RLS[i])){
    datadepression$Numdisorder[i] = NA
  }
  else{
    datadepression$Numdisorder[i] = as.numeric(l)
  }
}
```

```r
fulldistress = !is.na(t1$`K6-Tot`) & !is.na(t2$`K6-Tot`)
ID = t1$ID[fulldistress]
Time = c(rep(1, length(ID)), rep(2, length(ID)))
ID = rep(ID,2)
PTSD = rep(t1$`PTSD+`[fulldistress],2)
Stress = rep(t1$`INTChronicStress+`[fulldistress],2)
SAQcont = rep(t1$SAQGlobal[fulldistress],2)
SAQ2 = rep(t1$`SAQGlobal+`[fulldistress],2)
SAQ3 = splitSAQ3(SAQcont)
OSA = rep(t1$`SAQOSA+`[fulldistress],2)
Insomnia = rep(t1$`SAQInsom+`[fulldistress],2)
NRS = rep(t1$`SAQNRS+`[fulldistress],2)
DSPD = rep(t1$`SAQSched+`[fulldistress],2)
EDS = rep(t1$`SAQEDS+`[fulldistress],2)
RLS = rep(t1$`SAQRestless+`[fulldistress],2)
Disorder = 0
Numdisorder = 0
distress1 = t1$`K6-Tot`[fulldistress];distress2 = t2$`K6-Tot`[fulldistress]
Distress = c(distress1, distress2)
datadistress = tibble(ID, Time, PTSD, Stress, SAQcont, SAQ2, SAQ3, OSA, Insomnia,
                      NRS, DSPD, EDS, RLS, Disorder, Numdisorder, Distress)


#for Disorder 1 = OSA, 2 = Insomnia, 3 = Other
i = 0
for(i in 1:nrow(datadistress)){
  if(datadistress$OSA[i] == 1 & !is.na(datadistress$OSA[i])){
    datadistress$Disorder[i] = 1
  }
  else if(datadistress$Insomnia[i] == 1 & !is.na(datadistress$Insomnia[i])){
    datadistress$Disorder[i] = 2
  }
  else if(is.na(datadistress$OSA[i]) & is.na(datadistress$Insomnia[i]) &
          is.na(datadistress$NRS[i]) & is.na(datadistress$DSPD[i]) &
          is.na(datadistress$EDS[i]) & is.na(datadistress$RLS[i])){
    datadistress$Disorder[i] = NA
  }
  else(datadistress$Disorder[i] = 3)
}

#for Numdisorder
i = 0
k = 0 #num of Disorders
for(i in 1:nrow(datadistress)){
  l = 0 #num of Disorders added up
  for(j in 8:13){
```

```
    if(is.na(datadistress[i,j])){
      j = j + 1
        }
  else{
    l = l + datadistress[i,j]
    j = j + 1
   }
  }
  if(l >= 3){
    datadistress$Numdisorder[i] = 3
  }
  else if(is.na(datadistress$OSA[i]) & is.na(datadistress$Insomnia[i]) &
          is.na(datadistress$NRS[i]) & is.na(datadistress$DSPD[i]) &
          is.na(datadistress$EDS[i]) & is.na(datadistress$RLS[i])){
    datadistress$Numdisorder[i] = NA
  }
  else{
    datadistress$Numdisorder[i] = as.numeric(l)
  }
}
```

Model for Primary Research Question

Anxiety, depression, and distress were all modeled separately because the lmer function does not allow for multiple response variables within one model. Subject `ID` is treated as a random effect to account for the idiosyncratic variation amount patients while all the other covariates are treated as fixed effects. Each outcome variable consists of four different models which uses the four variables that measure sleep disorders and all of the models include time, PTSD, and stress. Including `SAQ3`, `Disorder`, `Numdisorder`, and the individual sleep disorders (`OSA`, `EDS`, `Insomnia`, `NRS`, `DSPD`, `RLS`) in a single model would not make sense because these variables are different ways of capturing the same informatino. The function `factor()` was used on the variables `SAQ3`, `Disorder`, and `Numdisorder` in order to examine the categories of the variables. `factor` is not applied to the variables of individual sleep disorders because they are binary variables. `REML = TRUE` removes all fixed effects and evaluates remaining variance that belongs to random effects (It ignores fixed effects). In order to assess fixed effects, we need to set `REML = FALSE` to use maximum likelihood (ML).

```
anxietymodel1 = lmer(Anxiety ~ Time * (factor(SAQ3) + PTSD + Stress) + (1|ID),
                     data = dataanxiety, REML = FALSE)
anxietymodel2 = lmer(Anxiety ~ Time * (OSA + EDS + Insomnia + NRS + DSPD +
                                        RLS + PTSD + Stress) + (1|ID),
                     data = dataanxiety, REML = FALSE)
anxietymodel3 = lmer(Anxiety ~ Time * (factor(Disorder) + PTSD + Stress) + (1|ID),
                     data = dataanxiety, REML = FALSE)
anxietymodel4 = lmer(Anxiety ~ Time * (factor(Numdisorder) + PTSD + Stress) + (1|ID),
                     data = dataanxiety, REML = FALSE)

depressionmodel1 = lmer(Depression ~ Time * (factor(SAQ3) + PTSD + Stress) + (1|ID),
                        data = datadepression, REML = FALSE)
depressionmodel2 = lmer(Depression ~ Time * (OSA + EDS + Insomnia + NRS + DSPD +
                                              RLS + PTSD + Stress) + (1|ID),
                        data = datadepression, REML = FALSE)
depressionmodel3 = lmer(Depression ~ Time * (factor(Disorder) + PTSD + Stress) + (1|ID),
                        data = datadepression, REML = FALSE)
```

```
depressionmodel4 = lmer(Depression ~ Time * (factor(Numdisorder) + PTSD + Stress) + (1|ID),
                        data = datadepression, REML = FALSE)

distressmodel1 = lmer(Distress ~ Time * (factor(SAQ3) + PTSD + Stress) + (1|ID),
                      data = datadistress, REML = FALSE)
distressmodel2 = lmer(Distress ~ Time * (OSA + EDS + Insomnia + NRS + DSPD +
                                           RLS + PTSD + Stress) + (1|ID),
                      data = datadistress, REML = FALSE)
distressmodel3 = lmer(Distress ~ Time * (factor(Disorder) + PTSD + Stress) + (1|ID),
                      data = datadistress, REML = FALSE)
distressmodel4 = lmer(Distress ~ Time * (factor(Numdisorder) + PTSD + Stress) + (1|ID),
                      data = datadistress, REML = FALSE)


summary(distressmodel1)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: Distress ~ Time * (factor(SAQ3) + PTSD + Stress) + (1 | ID)
##    Data: datadistress
##
##      AIC      BIC   logLik deviance df.resid
##    539.5    571.5   -257.8    515.5       94
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -1.7665 -0.5085 -0.0550  0.4541  2.2311
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 7.231    2.689
##  Residual             3.246    1.802
## Number of obs: 106, groups:  ID, 53
##
## Fixed effects:
##                  Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)        6.3408     1.9538 99.9884   3.245  0.00160 **
## Time              -1.8293     1.0277 53.0000  -1.780  0.08083 .
## factor(SAQ3)2     -0.7769     2.0126 99.9884  -0.386  0.70031
## factor(SAQ3)3     -1.7957     2.0236 99.9884  -0.887  0.37700
## PTSD               4.5300     1.7175 99.9884   2.638  0.00968 **
## Stress             3.0427     1.5102 99.9884   2.015  0.04661 *
## Time:factor(SAQ3)2  0.7303    1.0587 53.0000   0.690  0.49334
## Time:factor(SAQ3)3  2.4131    1.0645 53.0000   2.267  0.02750 *
## Time:PTSD         -1.1591     0.9035 53.0000  -1.283  0.20508
## Time:Stress       -0.8095     0.7944 53.0000  -1.019  0.31283
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) Time   f(SAQ3)2 f(SAQ3)3 PTSD   Stress T:(SAQ3)2
## Time       -0.789
## fctr(SAQ3)2 -0.808  0.638
## fctr(SAQ3)3 -0.669  0.528  0.698
## PTSD       -0.106  0.084  0.113   -0.063
```

```
## Stress        -0.431  0.340  0.077   -0.122   -0.235
## Tm:f(SAQ3)2  0.638 -0.808 -0.789   -0.551   -0.089 -0.061
## Tm:f(SAQ3)3  0.528 -0.669 -0.551   -0.789    0.049  0.096  0.698
## Time:PTSD     0.084 -0.106 -0.089    0.049   -0.789  0.185  0.113
## Time:Stress  0.340 -0.431 -0.061    0.096    0.185 -0.789  0.077
##               T:(SAQ3)3 T:PTSD
## Time
## fctr(SAQ3)2
## fctr(SAQ3)3
## PTSD
## Stress
## Tm:f(SAQ3)2
## Tm:f(SAQ3)3
## Time:PTSD    -0.063
## Time:Stress -0.122      -0.235
```

Explanation for distress w/ SAQ3 and confounders: - There is weak evidence that on average people improve their levels of distress over time (t-value = -1.8, p-value = 0.08) - There is evidence that people within SAQ category 3 improve less than category 1 and 2 (t-value = 2.3, p-value = 0.03) - There is no evidence that SAQ category 1 and 2 are significantly different (t-value = 0.69, p-value = 0.5) - Standard deviations are very high relative to the estimate indicating that there is not enough data or the data has too much noise

Drop-out and Participation

I modeled dropout by first creating a dataframe with every predictor and a column of whether or not they dropped out and then used the logistic regression. Logistic regression is a method to model a binary outcome variable and is appropriate for modelling drop-out (dropped out/did not drop out). The function `glm()` allows us to perform logistic regression and specifying `family = "binomial"` indicates using standard logistic regression. For participation, I used the same method but with `family = "quasibinomial"` which accounts for overdispersion because the participation observations are not independent which can lead to overdispersion.

```
dropout_attendance_all = tibble(SAQcont = t1data$SAQcont,
                    SAQmiss = 0,
                    SAQavg = t1data$SAQcont,
                    Insomnia = t1data$Insomnia,
                    OSA = t1data$OSA,
                    NRS = t1data$NRS,
                    DSPD = t1data$DSPD,
                    EDS = t1data$EDS,
                    RLS = t1data$RLS,
                    Distress = t1data$Distress,
                    Anxiety = t1data$Anxiety,
                    Depression = t1data$Depression,
                    Dropped = is.na(t2$`CESD-Tot`),
                    Attendance = t1data$Attendance)

dropout_attendance_all$SAQmiss = is.na(dropout_attendance_all$SAQcont)
dropout_attendance_all$SAQmiss = as.numeric(dropout_attendance_all$SAQmiss)
dropout_attendance_all$SAQavg[is.na(dropout_attendance_all$SAQavg)] =
  mean(na.omit(dropout_attendance_all$SAQcont))


summary(glm(Dropped~ SAQcont +  Distress + Anxiety + Depression +
```
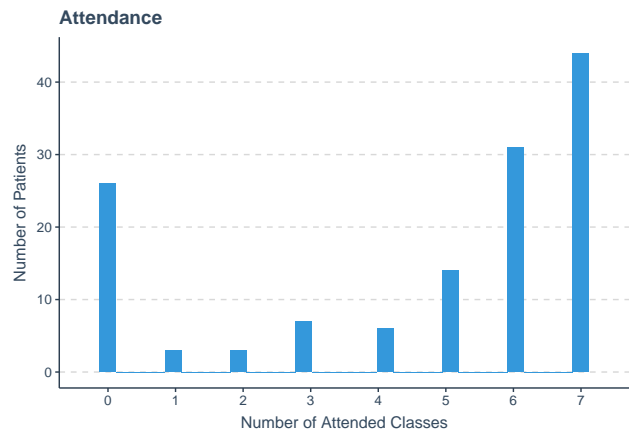
```
              Insomnia + OSA + NRS + DSPD + EDS + RLS,
           data = dropout_attendance_all, family = "binomial"))
```

```
##
## Call:
## glm(formula = Dropped ~ SAQcont + Distress + Anxiety + Depression +
##     Insomnia + OSA + NRS + DSPD + EDS + RLS, family = "binomial",
##     data = dropout_attendance_all)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0355  -1.0985   0.5486   1.0033   1.6408
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.009228   0.766593   1.317   0.1880
## SAQcont     -0.108888   0.060821  -1.790   0.0734 .
## Distress    -0.061824   0.082980  -0.745   0.4562
## Anxiety      0.007365   0.029395   0.251   0.8022
## Depression   0.070904   0.035385   2.004   0.0451 *
## Insomnia     1.002138   0.565506   1.772   0.0764 .
## OSA          0.358745   0.506726   0.708   0.4790
## NRS         -0.199867   0.575039  -0.348   0.7282
## DSPD         1.135108   0.521031   2.179   0.0294 *
## EDS         -0.033426   0.477796  -0.070   0.9442
## RLS          0.761928   0.507586   1.501   0.1333
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 157.46  on 114  degrees of freedom
## Residual deviance: 142.17  on 104  degrees of freedom
##   (36 observations deleted due to missingness)
## AIC: 164.17
##
## Number of Fisher Scoring iterations: 4
```

```
#number of patients attending up to 7 classes
ggplot(data = dropout_attendance_all) +
  geom_histogram(aes(Attendance)) +
  labs(title = "Attendance", y = "Number of Patients", x = "Number of Attended Classes") +
  scale_x_discrete(limits=0:7) + theme_hc()
```

```
summary(glm(cbind(Attendance, 7 - Attendance) ~ SAQcont + Distress +
            Anxiety + Depression + Insomnia + OSA + NRS + DSPD + EDS + RLS,
        data = dropout_attendance_all, family = "quasibinomial"))
```

```
##
## Call:
## glm(formula = cbind(Attendance, 7 - Attendance) ~ SAQcont + Distress +
##     Anxiety + Depression + Insomnia + OSA + NRS + DSPD + EDS +
##     RLS, family = "quasibinomial", data = dropout_attendance_all)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.4388  -1.8646   0.7536   2.0342   3.4590
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.316507   0.643020  -0.492    0.624
## SAQcont      0.053777   0.050359   1.068    0.288
## Distress    -0.004189   0.071151  -0.059    0.953
## Anxiety      0.015263   0.025461   0.599    0.550
## Depression  -0.025731   0.027404  -0.939    0.350
## Insomnia    -0.234927   0.458949  -0.512    0.610
## OSA         -0.126734   0.433480  -0.292    0.771
## NRS          0.460771   0.485055   0.950    0.345
## DSPD        -0.431148   0.417182  -1.033    0.304
## EDS         -0.455035   0.418229  -1.088    0.279
## RLS         -0.042994   0.437715  -0.098    0.922
##
## (Dispersion parameter for quasibinomial family taken to be 4.705802)
##
##     Null deviance: 576.79  on 103  degrees of freedom
## Residual deviance: 541.09  on  93  degrees of freedom
##   (47 observations deleted due to missingness)
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

Model for Secondary Research Question

Although the interaction plots were indicative of significant results, the low sample size resulted in results that were not very strong. Therefore, we are unable to make any conclusions on he question of how the SMT program improves the levels of stress of patients with sleep disorders. It may be a good idea to answer a slightly different question: Whether or not people entering the program with high SAQ having initially higher levels of stress and if this difference is larger than the improvement due to the program.

This plot contains side by side boxplots of each response variables for each specific sleep disorder. This gives us a good visual for the difference in stress levels between having / not having each sleep disorder.

```
conditionsplot = t1data %>%
  select(Insomnia, OSA, NRS, DSPD, EDS, RLS) %>%
  gather(Condition, Status)

distresslevels1 = t1data %>%
  select(Distress = Distress)
distresslevels2 = t1data %>%
  select(Distress = Distress)
distresslevels3 = t1data %>%
  select(Distress = Distress)
distresslevels4 = t1data %>%
  select(Distress = Distress)
distresslevels5 = t1data %>%
  select(Distress = Distress)
distresslevels6 = t1data %>%
  select(Distress = Distress)
distresslevels =
  bind_rows(distresslevels1, distresslevels2, distresslevels3,
            distresslevels4, distresslevels5, distresslevels6) %>%
  gather("Stress Type", "Stress Level")

anxietylevels1 = t1data %>%
  select(Anxiety = Anxiety)
anxietylevels2 = t1data %>%
  select(Anxiety = Anxiety)
anxietylevels3 = t1data %>%
  select(Anxiety = Anxiety)
anxietylevels4 = t1data %>%
  select(Anxiety = Anxiety)
anxietylevels5 = t1data %>%
  select(Anxiety = Anxiety)
anxietylevels6 = t1data %>%
  select(Anxiety = Anxiety)
anxietylevels =
  bind_rows(anxietylevels1, anxietylevels2, anxietylevels3,
            anxietylevels4, anxietylevels5, anxietylevels6) %>%
  gather("Stress Type", "Stress Level")

depressionlevels1 = t1data %>%
  select(Depression = Depression)
depressionlevels2 = t1data %>%
  select(Depression = Depression)
depressionlevels3 = t1data %>%
```
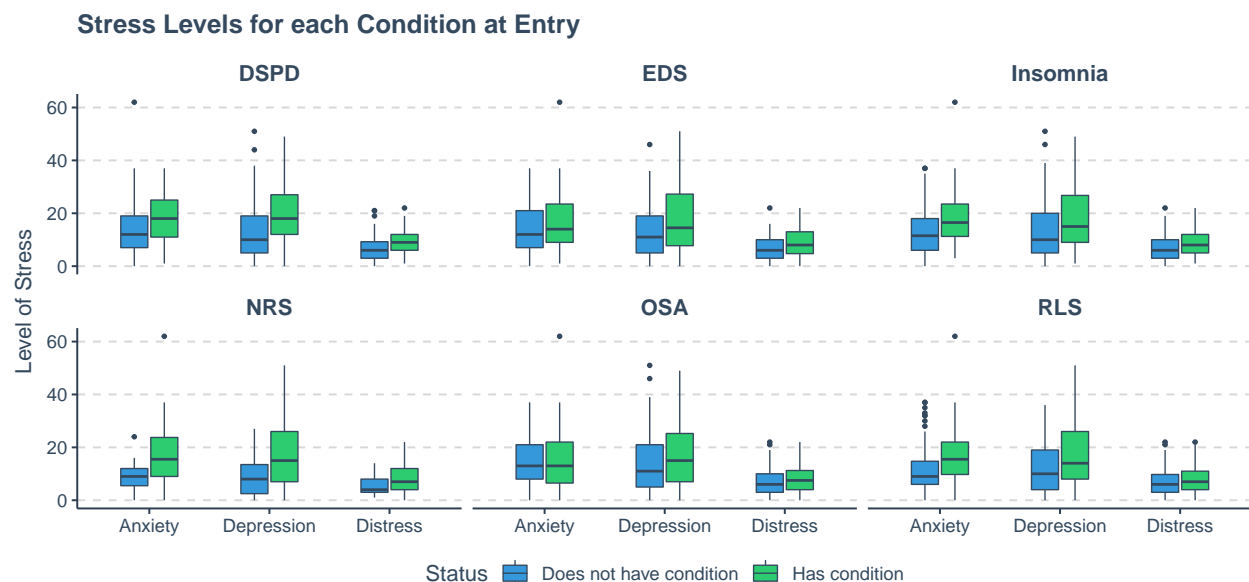
```
    select(Depression = Depression)
depressionlevels4 = t1data %>%
    select(Depression = Depression)
depressionlevels5 = t1data %>%
    select(Depression = Depression)
depressionlevels6 = t1data %>%
    select(Depression = Depression)
depressionlevels =
    bind_rows(depressionlevels1, depressionlevels2, depressionlevels3,
              depressionlevels4, depressionlevels5, depressionlevels6) %>%
    gather("Stress Type", "Stress Level")


conditionsplotdistress = bind_cols(conditionsplot, distresslevels)
conditionsplotanxiety = bind_cols(conditionsplot, anxietylevels)
conditionsplotdepression = bind_cols(conditionsplot, depressionlevels)

allplots = full_join(conditionsplotdistress, conditionsplotanxiety,
                     by = c("Condition", "Status", "Stress Type", "Stress Level"))
allplots = full_join(allplots, conditionsplotdepression,
                     by = c("Condition", "Status", "Stress Type", "Stress Level"))


ggplot(na.omit(allplots)) +
    geom_boxplot(aes(x=factor(`Stress Type`), y= `Stress Level` ,fill=factor(Status)),
                 width = 0.5, size = 0.35, outlier.size = 0.75) +
    labs(title="Stress Levels for each Condition at Entry",
         x = NULL, y = "Level of Stress")  + scale_fill_discrete(name = "Status",
           labels = c("Does not have condition", "Has condition")) +
    facet_wrap(~Condition) + theme(legend.position = "bottom") + theme_hc() +
    theme(strip.text = element_text(size = 12, face = "bold"))
```


Stress Levels for each Condition at Entry

The above plot shows us some evidence for difference in means. Conducting two-sample t-tests will test whether or not these differences are significant. The function `t.test()` was used with `var.equal = FALSE` because the variances differed.

```r
#Sleep conditions
pvaldistress = tibble(EDS = t.test(Distress ~ EDS,
                                   var.equal = FALSE, data = t1data)$p.value,
                      Insomnia = t.test(Distress ~ Insomnia,
                                        var.equal = FALSE, data = t1data)$p.value,
                      NRS = t.test(Distress ~ NRS,
                                   var.equal = FALSE, data = t1data)$p.value,
                      OSA = t.test(Distress ~ OSA,
                                   var.equal = FALSE, data = t1data)$p.value,
                      RLS = t.test(Distress ~ RLS,
                                   var.equal = FALSE, data = t1data)$p.value,
                      DSPD = t.test(Distress ~ DSPD,
                                    var.equal = FALSE, data = t1data)$p.value)
pvaldistress = pvaldistress %>%
  gather(Condition, "p-value") %>%
  mutate(Stress = "Distress")

estdistress0 = tibble(EDS = t.test(Distress ~ EDS,
                                   var.equal = FALSE, data = t1data)$estimate[1],
                      Insomnia = t.test(Distress ~ Insomnia,
                                        var.equal = FALSE, data = t1data)$estimate[1],
                      NRS = t.test(Distress ~ NRS,
                                   var.equal = FALSE, data = t1data)$estimate[1],
                      OSA = t.test(Distress ~ OSA,
                                   var.equal = FALSE, data = t1data)$estimate[1],
                      RLS = t.test(Distress ~ RLS,
                                   var.equal = FALSE, data = t1data)$estimate[1],
                      DSPD = t.test(Distress ~ DSPD,
                                    var.equal = FALSE, data = t1data)$estimate[1])
estdistress0 = estdistress0 %>%
  gather(Condition, "Mean for group 0") %>%
  mutate(Stress = "Distress")
estdistress1 = tibble(EDS = t.test(Distress ~ EDS,
                                   var.equal = FALSE, data = t1data)$estimate[2],
                      Insomnia = t.test(Distress ~ Insomnia,
                                        var.equal = FALSE, data = t1data)$estimate[2],
                      NRS = t.test(Distress ~ NRS,
                                   var.equal = FALSE, data = t1data)$estimate[2],
                      OSA = t.test(Distress ~ OSA,
                                   var.equal = FALSE, data = t1data)$estimate[2],
                      RLS = t.test(Distress ~ RLS,
                                   var.equal = FALSE, data = t1data)$estimate[2],
                      DSPD = t.test(Distress ~ DSPD,
                                    var.equal = FALSE, data = t1data)$estimate[2])
estdistress1 = estdistress1 %>%
  gather(Condition, "Mean for group 1") %>%
  mutate(Stress = "Distress")

estdistress = full_join(estdistress0, estdistress1)
testdistress = full_join(pvaldistress, estdistress)

pvalanxiety = tibble(EDS = t.test(Anxiety ~ EDS,
                                  var.equal = FALSE, data = t1data)$p.value,
```

```r
                       Insomnia = t.test(Anxiety ~ Insomnia,
                                         var.equal = FALSE, data = t1data)$p.value,
                  NRS = t.test(Anxiety ~ NRS,
                                 var.equal = FALSE, data = t1data)$p.value,
                  OSA = t.test(Anxiety ~ OSA,
                                 var.equal = FALSE, data = t1data)$p.value,
                  RLS = t.test(Anxiety ~ RLS,
                                 var.equal = FALSE, data = t1data)$p.value,
                  DSPD = t.test(Anxiety ~ DSPD,
                                  var.equal = FALSE, data = t1data)$p.value)
pvalanxiety = pvalanxiety %>%
  gather(Condition, "p-value") %>%
  mutate(Stress = "Anxiety")

estanxiety0 = tibble(EDS = t.test(Anxiety ~ EDS,
                                    var.equal = FALSE, data = t1data)$estimate[1],
                  Insomnia = t.test(Anxiety ~ Insomnia,
                                      var.equal = FALSE, data = t1data)$estimate[1],
                  NRS = t.test(Anxiety ~ NRS,
                                 var.equal = FALSE, data = t1data)$estimate[1],
                  OSA = t.test(Anxiety ~ OSA,
                                 var.equal = FALSE, data = t1data)$estimate[1],
                  RLS = t.test(Anxiety ~ RLS,
                                 var.equal = FALSE, data = t1data)$estimate[1],
                  DSPD = t.test(Anxiety ~ DSPD,
                                  var.equal = FALSE, data = t1data)$estimate[1])
estanxiety0 = estanxiety0 %>%
  gather(Condition, "Mean for group 0") %>%
  mutate(Stress = "Anxiety")
estanxiety1 = tibble(EDS = t.test(Anxiety ~ EDS,
                                    var.equal = FALSE, data = t1data)$estimate[2],
                  Insomnia = t.test(Anxiety ~ Insomnia,
                                      var.equal = FALSE, data = t1data)$estimate[2],
                  NRS = t.test(Anxiety ~ NRS,
                                 var.equal = FALSE, data = t1data)$estimate[2],
                  OSA = t.test(Anxiety ~ OSA,
                                 var.equal = FALSE, data = t1data)$estimate[2],
                  RLS = t.test(Anxiety ~ RLS,
                                 var.equal = FALSE, data = t1data)$estimate[2],
                  DSPD = t.test(Anxiety ~ DSPD,
                                  var.equal = FALSE, data = t1data)$estimate[2])
estanxiety1 = estanxiety1 %>%
  gather(Condition, "Mean for group 1") %>%
  mutate(Stress = "Anxiety")

estanxiety = full_join(estanxiety0, estanxiety1)
testanxiety = full_join(pvalanxiety, estanxiety)


pvaldepression = tibble(EDS = t.test(Depression ~ EDS, var.equal =
                                        FALSE, data = t1data)$p.value,
                  Insomnia = t.test(Depression ~ Insomnia,
                                      var.equal = FALSE, data = t1data)$p.value,
```

```r
                      NRS = t.test(Depression ~ NRS,
                                   var.equal = FALSE, data = t1data)$p.value,
                      OSA = t.test(Depression ~ OSA,
                                   var.equal = FALSE, data = t1data)$p.value,
                      RLS = t.test(Depression ~ RLS,
                                   var.equal = FALSE, data = t1data)$p.value,
                      DSPD = t.test(Depression ~ DSPD,
                                    var.equal = FALSE, data = t1data)$p.value)
pvaldepression = pvaldepression %>%
  gather(Condition, "p-value") %>%
  mutate(Stress = "Depression")

estdepression0 = tibble(EDS = t.test(Depression ~ EDS,
                                     var.equal = FALSE, data = t1data)$estimate[1],
                        Insomnia = t.test(Depression ~ Insomnia,
                                          var.equal = FALSE, data = t1data)$estimate[1],
                        NRS = t.test(Depression ~ NRS,
                                     var.equal = FALSE, data = t1data)$estimate[1],
                        OSA = t.test(Depression ~ OSA,
                                     var.equal = FALSE, data = t1data)$estimate[1],
                        RLS = t.test(Depression ~ RLS,
                                     var.equal = FALSE, data = t1data)$estimate[1],
                        DSPD = t.test(Depression ~ DSPD,
                                      var.equal = FALSE, data = t1data)$estimate[1])
estdepression0 = estdepression0 %>%
  gather(Condition, "Mean for group 0") %>%
  mutate(Stress = "Depression")
estdepression1 = tibble(EDS = t.test(Depression ~ EDS,
                                     var.equal = FALSE, data = t1data)$estimate[2],
                        Insomnia = t.test(Depression ~ Insomnia,
                                          var.equal = FALSE, data = t1data)$estimate[2],
                        NRS = t.test(Depression ~ NRS,
                                     var.equal = FALSE, data = t1data)$estimate[2],
                        OSA = t.test(Depression ~ OSA,
                                     var.equal = FALSE, data = t1data)$estimate[2],
                        RLS = t.test(Depression ~ RLS,
                                     var.equal = FALSE, data = t1data)$estimate[2],
                        DSPD = t.test(Depression ~ DSPD,
                                      var.equal = FALSE, data = t1data)$estimate[2])
estdepression1 = estdepression1 %>%
  gather(Condition, "Mean for group 1") %>%
  mutate(Stress = "Depression")

estdepression = full_join(estdepression0, estdepression1)
testdepression = full_join(pvaldepression, estdepression)

options(scipen=999)

testresults = full_join(testdistress, testanxiety)
testresults = full_join(testresults, testdepression)
testresults$`p-value` = signif(testresults$`p-value`, digits = 2)
testresults$`Mean for group 0` = signif(testresults$`Mean for group 0`, 2)
testresults$`Mean for group 1` = signif(testresults$`Mean for group 1`, 2)
```

```r
testresults$Difference = testresults$`Mean for group 1` - testresults$`Mean for group 0`
```

```r
kable(testresults)
```

Condition	p-value	Stress	Mean for group 0	Mean for group 1	Difference
EDS	0.007600	Distress	6.6	9.0	2.4
Insomnia	0.025000	Distress	6.9	9.0	2.1
NRS	0.000130	Distress	5.4	8.4	3.0
OSA	0.260000	Distress	7.2	8.2	1.0
RLS	0.190000	Distress	7.0	8.0	1.0
DSPD	0.007400	Distress	6.8	9.4	2.6
EDS	0.061000	Anxiety	14.0	17.0	3.0
Insomnia	0.007000	Anxiety	13.0	18.0	5.0
NRS	0.000000	Anxiety	8.6	17.0	8.4
OSA	0.470000	Anxiety	15.0	16.0	1.0
RLS	0.012000	Anxiety	13.0	17.0	4.0
DSPD	0.067000	Anxiety	14.0	18.0	4.0
EDS	0.014000	Depression	13.0	18.0	5.0
Insomnia	0.017000	Depression	13.0	19.0	6.0
NRS	0.000021	Depression	9.6	18.0	8.4
OSA	0.220000	Depression	15.0	17.0	2.0
RLS	0.007800	Depression	12.0	17.0	5.0
DSPD	0.009500	Depression	13.0	19.0	6.0

Mean differences in stress levels per disorder:

Condition	# Cases	# Non-cases	Anxiety	Depression	Distress
DSPD	41	93	**4.0** (0.07)	**6.0** (0.01)	**2.6** (0.007)
EDS	60	85	3.0 (0.06)	**5.0** (0.01)	**2.4** (0.008)
Insomnia	50	87	**5.0** (0.01)	**6.0** (0.02)	2.1 (0.3)
NRS	99	40	**8.4** ($\ll 0.01$)	**8.4** ($\ll 0.01$)	**3.0** (< 0.01)
OSA	60	82	1.0 (0.5)	2.0 (0.2)	1.0 (0.3)
RLS	85	62	**4.0** (0.01)	**5.0** (0.008)	1.0 (0.2)

Similar to the model used to answer the primary research question, we used the same method but without all the covariates. We are only interested in how anxiety, depression, and distress all vary with time.

```r
#improvement of anxiety for all
anxietymodel = lmerTest::lmer(Anxiety ~ Time +(1|ID), data = t12data, REML = FALSE)

#improvement of depression for all
depressionmodel = lmerTest::lmer(Depression ~ Time +(1|ID), data = t12data, REML = FALSE)

#improvement of distress for all
distressmodel = lmerTest::lmer(Distress ~ Time +(1|ID), data = t12data, REML = FALSE)

summary(distressmodel)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
```

```
## Formula: Distress ~ Time + (1 | ID)
##    Data: t12data
##
##      AIC      BIC   logLik deviance df.resid
##   1244.9   1258.5   -618.5   1236.9      216
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.07887 -0.41923 -0.03236  0.41315  2.45967
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID       (Intercept) 18.246   4.271
##  Residual             4.378    2.092
## Number of obs: 220, groups:  ID, 150
##
## Fixed effects:
##             Estimate Std. Error       df t value        Pr(>|t|)
## (Intercept)   9.0366     0.5738 173.9308  15.749 < 0.0000000000000002 ***
## Time         -1.5037     0.3440  71.8932  -4.372          0.0000408 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## Time -0.751
```

Improvement of stress levels in all patients:

Anxiety	Depression	Distress
1.8 (0.05)	2.6 (< 0.01)	1.5 ($\ll 0.01$)