STAT 590 Survival Analysis - Final Project

A clinical Trial in the Treatment of Carcinoma of the Oropharynx

Andres Gonzalez
Yushan Zhao

May 8, 2023

CALIFORNIA STATE UNIVERSITY - LONG BEACH
DEPARTEMENT OF MATHEMATICS AND STATISTICS
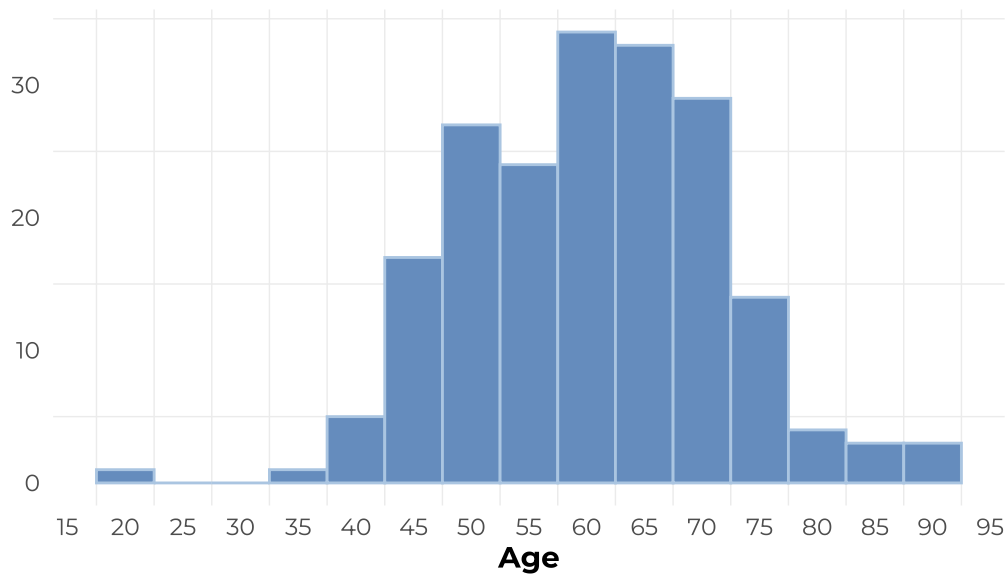1250 BELLFLOWER BOULEVARD
LONG BEACH, CALIFORNIA 90840

Table of contents

1 Exploratory Data Analysis

1.1 Age

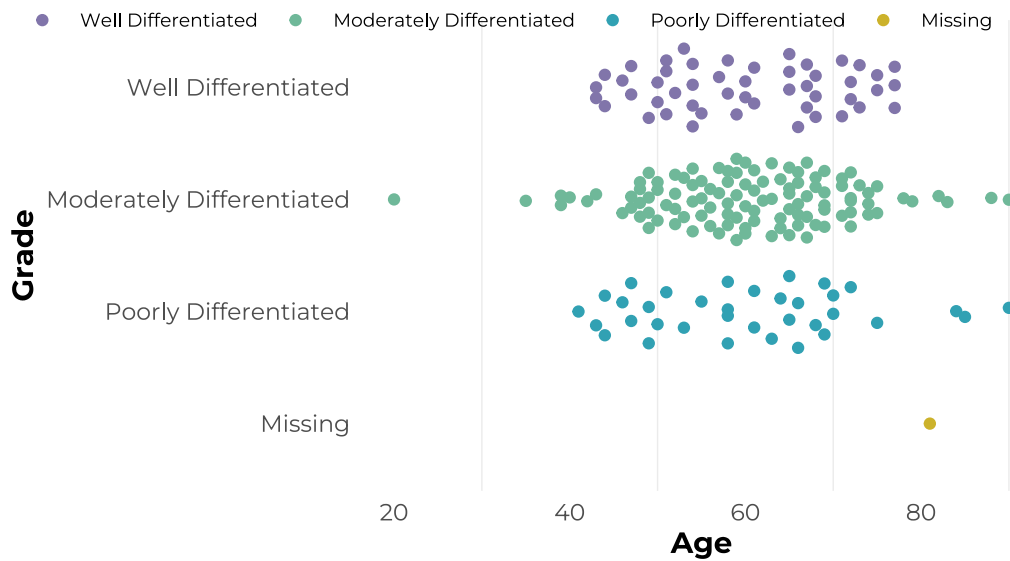**Age Distribution**

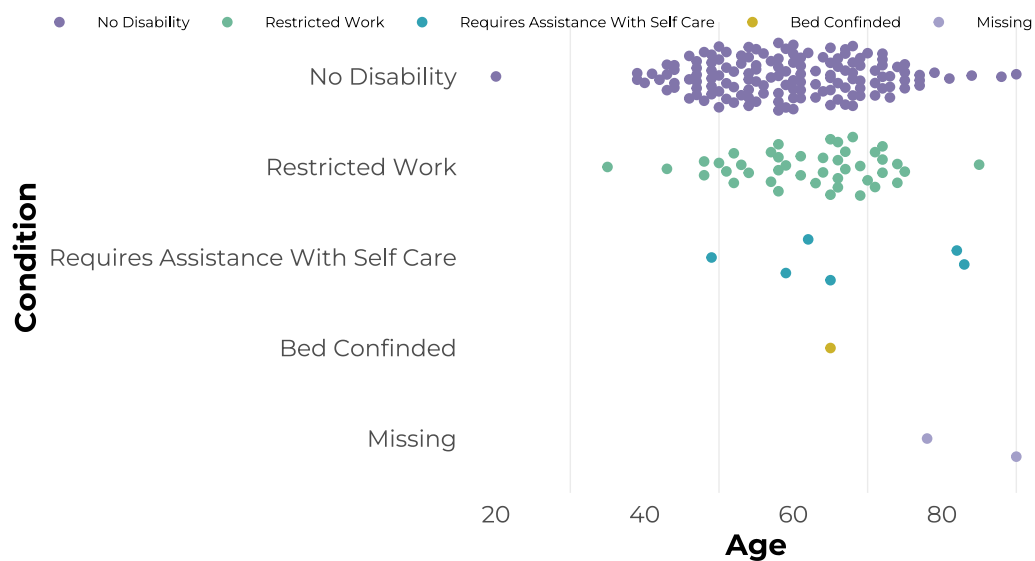*Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset*

Age Distribution by Grade Differentiation

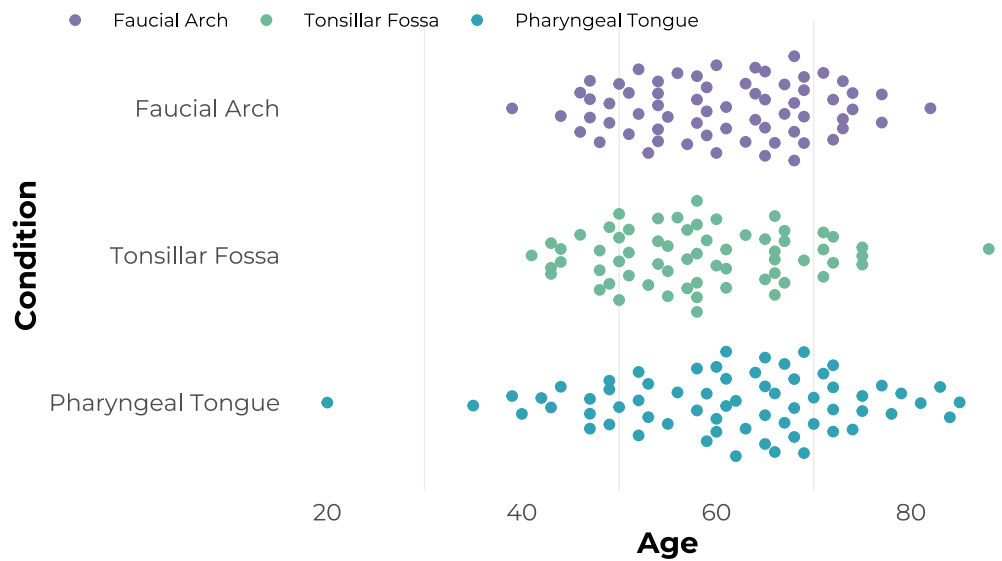*Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset*

Age Distribution by Condition

Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset
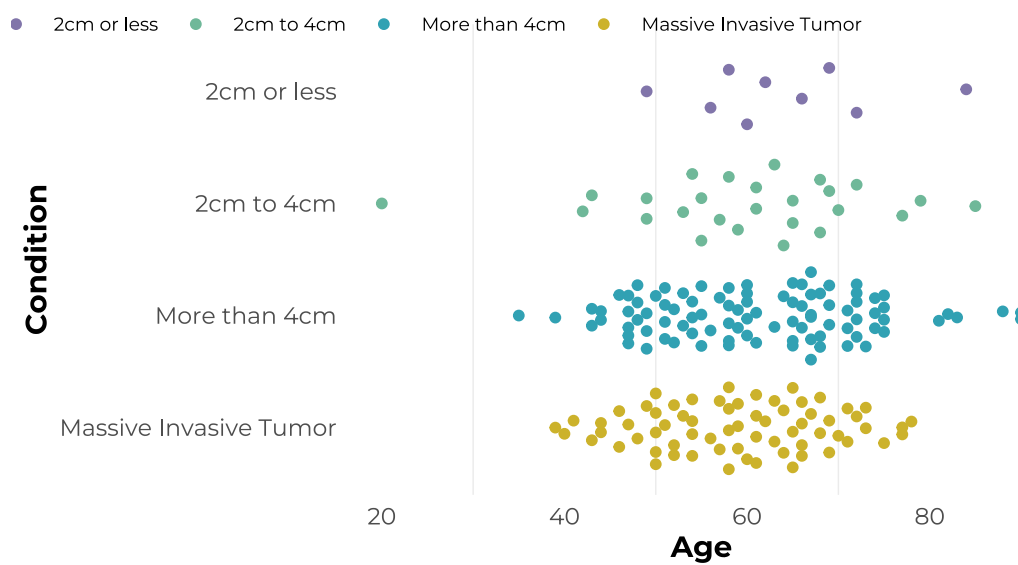
Age Distribution by Site

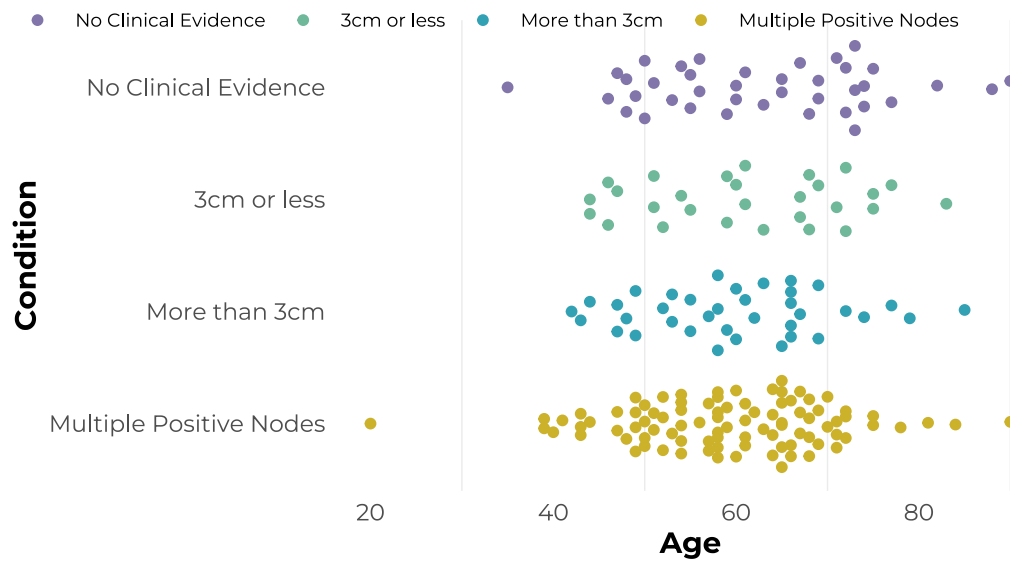*Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset*

Age Distribution by T_Stage

Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset
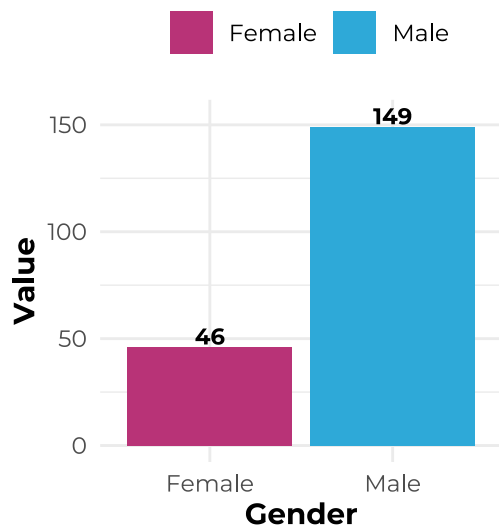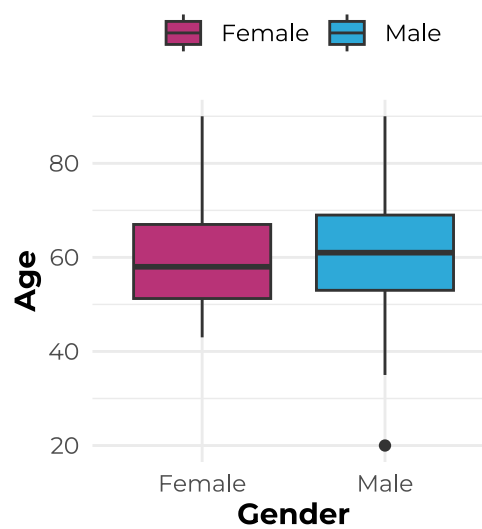
Age Distribution by T_Stage

Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset

1.2 Gender

Distribution of Gender

Age Distribution by Gender

Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset

Condition

195 Participants

Bed Confined · Missing · No Disability · Requires Assistance with Self Care · Restricted Work
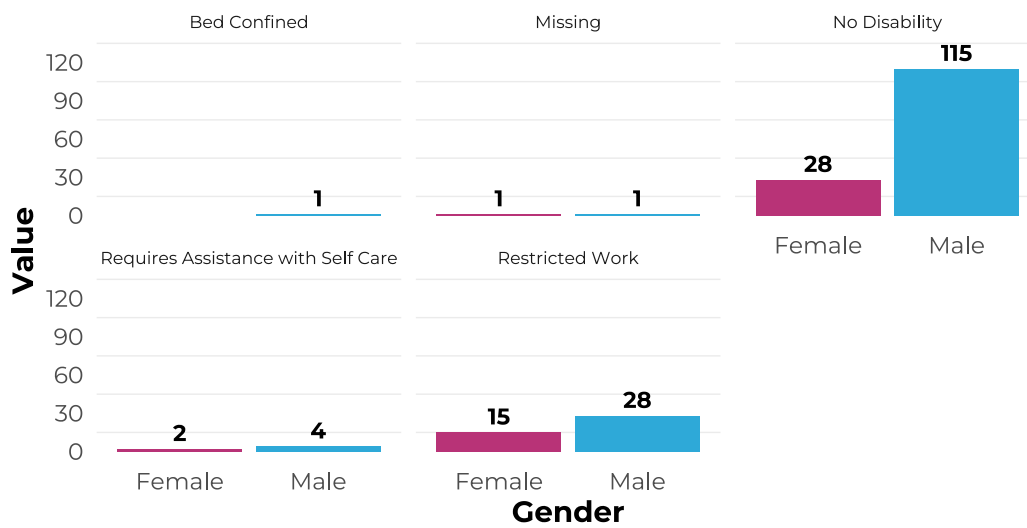
Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset
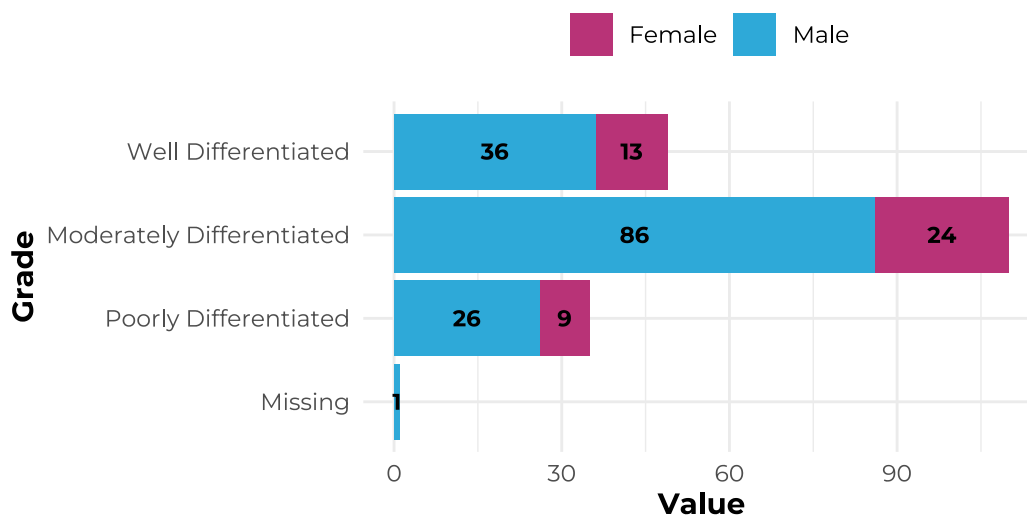
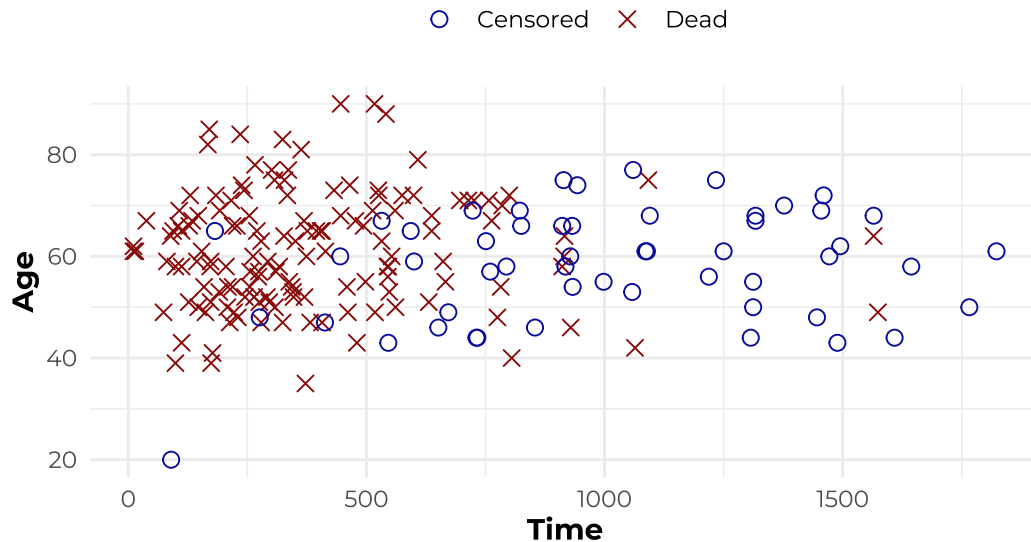## Stacked Distribution of Grades by Gender

195 Participants

Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset

1.3 Time and Status

Scatter Plot of Time and Age by Status

Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset

1.4 Kaplan Meier

```r
pharynx <- pharynx %>%
  mutate(
    Status = as.numeric(Status),
    TX = as_factor(Tx),
    SEX = as_factor(Sex),
    T_STAGE = as_factor(T_Stage)
  )


pharynx_surv <- Surv(pharynx$Time, pharynx$Status)

# Fit a Kaplan-Meier curve, stratified by treatment
pharynx_km_tx <- survfit2(pharynx_surv ~ TX, data = pharynx)

# Plot the Kaplan-Meier curve, stratified by treatment
ggsurvfit(pharynx_km_tx, linewidth = 1) +
  add_confidence_interval() +
  add_censor_mark() +
```

```
add_pvalue(location = "annotation") +
labs(y = "Percentage Survival",
     title = "Kaplan-Meier Survival Curve by Treatment") +
scale_y_continuous(label = scales::percent,
                   breaks = seq(0,1, by = 0.2),
                   expand = c(0.015,0)) +
scale_color_manual(values = c("#FC766AFF", "#5b84B1FF"),
                   labels = c("Standard","Test")) +
scale_fill_manual(values = c("#FC766AFF", "#5b84B1FF"),
                  labels = c("Standard","Test")) +
theme(
  text = element_text(family = "montse"),
  title = element_text(face = "bold"),
  plot.title.position = "plot",
  legend.position = "top"
)
```

Kaplan-Meier Survival Curve by Treatment

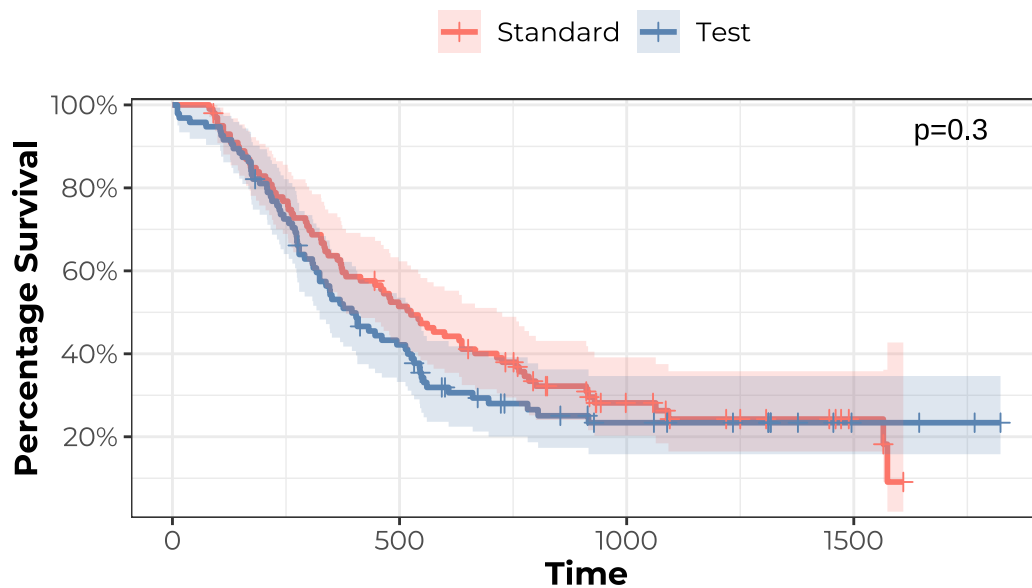

```
pharynx_surv <- Surv(pharynx$Time, pharynx$Status)

# Fit a Kaplan-Meier curve, stratified by treatment
pharynx_km_tx <- survfit2(pharynx_surv ~ Sex, data = pharynx)

# Plot the Kaplan-Meier curve, stratified by treatment
```

```
ggsurvfit(pharynx_km_tx, linewidth = 1) +
  add_confidence_interval() +
  add_censor_mark() +
  add_pvalue(location = "annotation") +
  labs(y = "Percentage Survival",
       title = "Kaplan-Meier Survival Curve by Treatment") +
  scale_y_continuous(label = scales::percent,
                     breaks = seq(0,1, by = 0.2),
                     expand = c(0.015,0)) +
  scale_color_manual(values = c("#B83377", "#2EA9D8")) +
  scale_fill_manual(values = c("#B83377", "#2EA9D8")) +
  theme(
    text = element_text(family = "montse"),
    title = element_text(face = "bold"),
    plot.title.position = "plot",
    legend.position = "top"
  )
```

Kaplan-Meier Survival Curve by Treatment