Clinical Trial on Treatment for Pharyngeal Cancer

Survival Analysis Final Project

Marian Lu
California State University Long Beach
May 8, 2024

Table of Contents

1 Abstract

This study aims to evaluate the effects of two different radiation therapies on a population of patients with cancer in the throat and mouth. Exploratory data analysis was carried out by developing graphs and plots to identify the presence of any patterns or interactions in the data. The model was built using survival analysis methods such as Proportional Hazard Regression and Likelihood Ratio Tests. The final model and results were then analyzed to draw a conclusion on the treatment effect.

2 Introduction

The patient data is from a clinical trial provided by the Radiation Therapy Oncology Group. Out of the sixteen institutions participating in the study, only the six largest institutions had their reported data taken into consideration. The study population consists of patients with squamous carcinoma of 15 sites in the throat and mouth, but only three of the sites are taken into consideration here: the faucial arch, tonsillar fossa, and pharyngeal tongue. The patients were randomly allocated to one of the treatment groups, which were radiation therapy alone and radiation therapy with a chemotherapeutic agent. The study aims to determine whether the combined treatment mode is preferable to the conventional radiation therapy.

Age

Age of the patient at the time of diagnosis.

Sex

The sex variable is a binary variable that classifies the patient as male or female.

Condition

Condition refers to the patient's able-bodiedness divided into 4 categories of increasing severity: no disability, restricted work, requires assistance with self care, and bed confined. There is also a category for missing data.

Site

Site refers to the location of the tumor. The data only contains three sites: which are the faucial arch, tonsillar fossa, and pharyngeal tongue. The faucial arches are located directly behind the oral cavity, the tonsillar fossa is a space within the lateral wall of the mouth, and the pharyngeal tongue is the part of the tongue closest to the throat.

Grade

Grade is a measure of how different the tumor is from the host cell and is recorded in three categories: well differentiated, moderately differentiated, and poorly differentiated.

T Stage

T staging classification describes the size of the tumor on an ordinal scale of increasing size, with 1 being the smallest size and 4 being the largest.

N Stage

N staging classification describes the magnitude of lymph node involvement on an ordinal scale of increasing severity, with 1 being no nodes detected and 4 being multiple nodes.

Time

Survival time in days since diagnosis

3 Results

3.1 Exploratory Data Analysis

Figure 3.1.1 Age Distribution by Sex

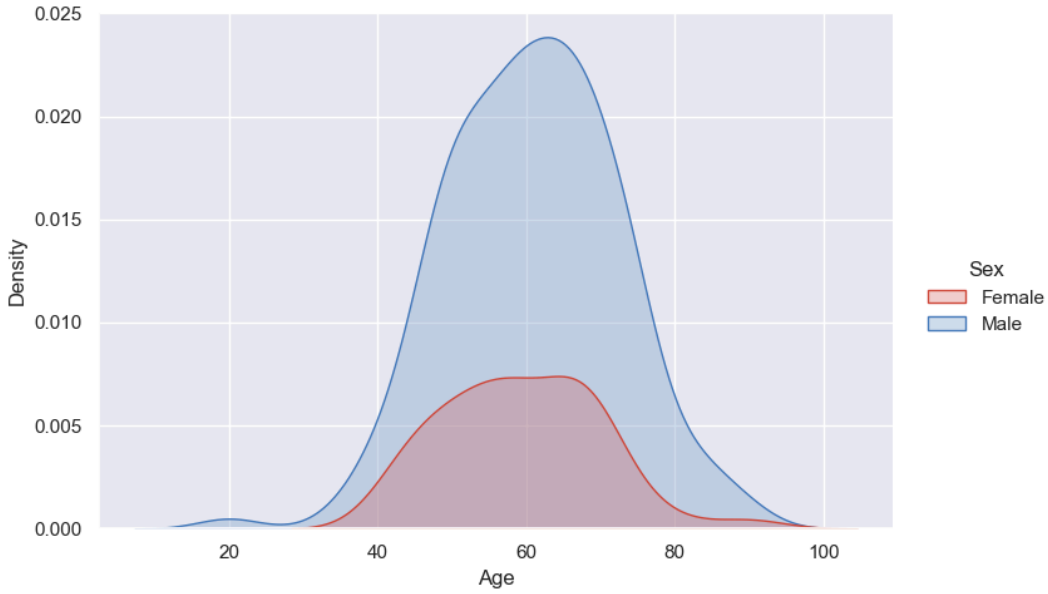
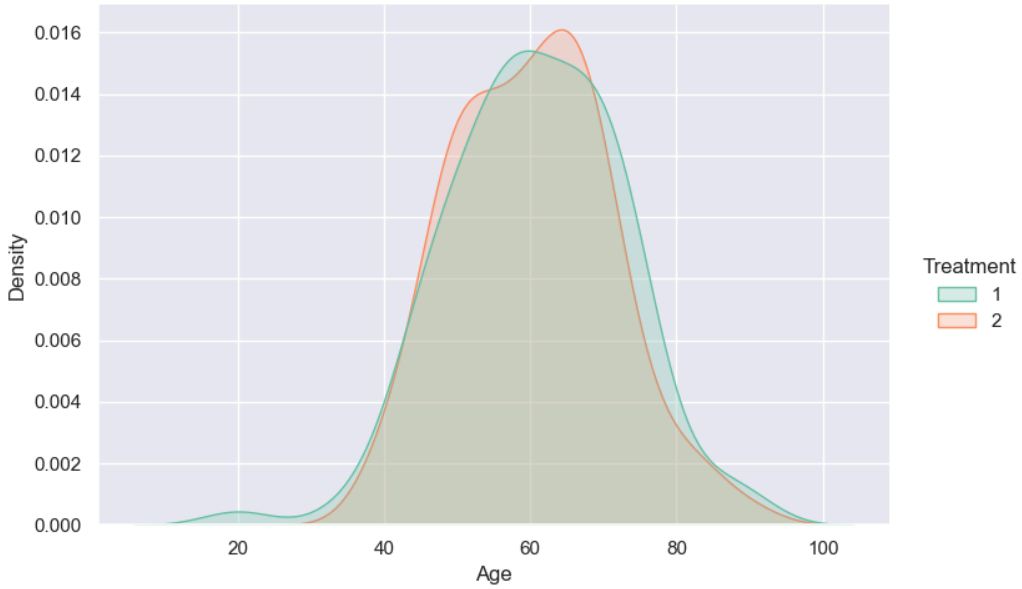


Figure 3.1.2 Age Distribution By Treatment



Age appears to follow a normal distribution, with the average patient being 60.4 years old. The majority of the patients are male, with over three times as many males as females. The patients allocated to each treatment have similar age distributions as well as cohort sizes.

Observing Variable Frequencies with Barplots

Figure 3.1.3 Barplot by General Condition

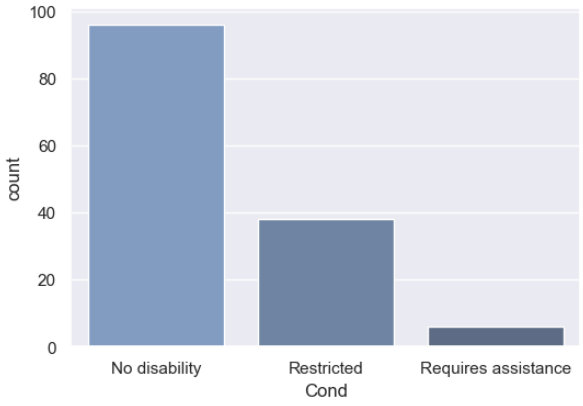


Figure 3.1.4 Barplot by Site

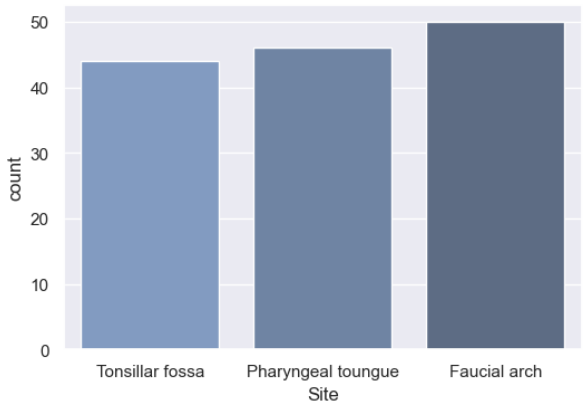


Figure 3.1.5 Barplot by T Stage

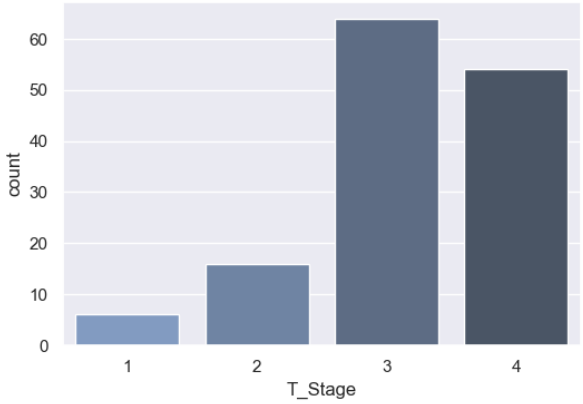


Figure 3.1.6 Barplot by N Stage

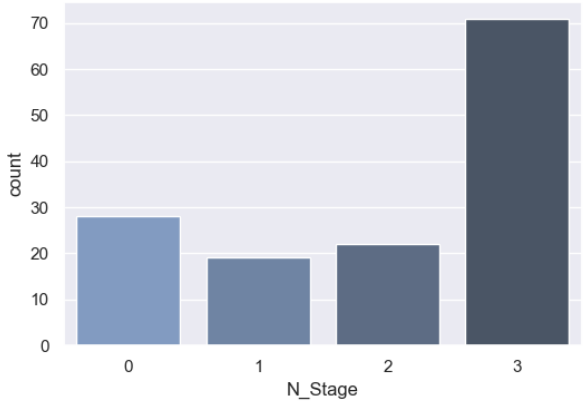


Figure 3.1.7 Barplot by Sex

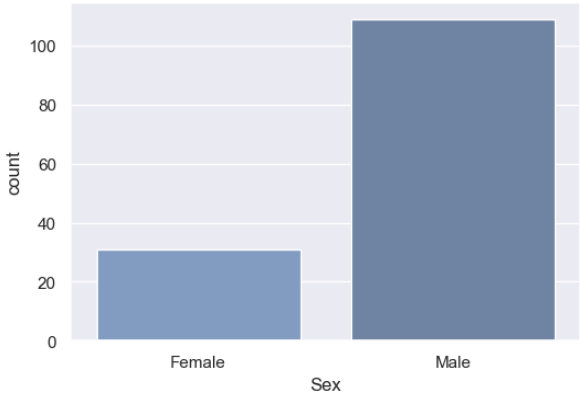
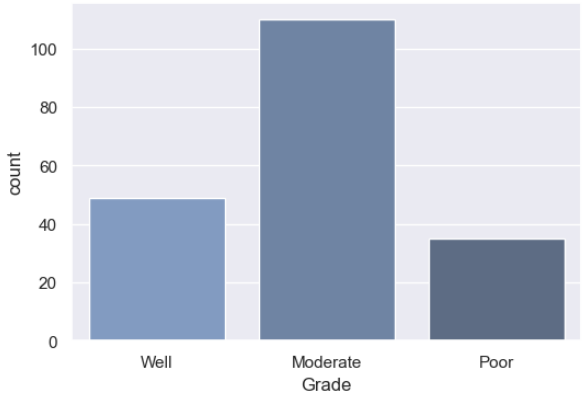


Figure 3.1.8 Barplot by Grade



The general condition frequency barplot shows that most patients entering the study have no disabilities. There is generally an even distribution of the tumor locations. However, the T Stage and N Stage categories show that the majority of patients are in the more severe categories. There is also an imbalance in the gender distribution, as mentioned before. Finally, the grade frequency bar plot shows that most of the patients' tumors moderately resemble the host cell.

Observing Survival Rates with Kaplan Meier Curves

Figure 3.1.9 KM Survival Curve of General Condition

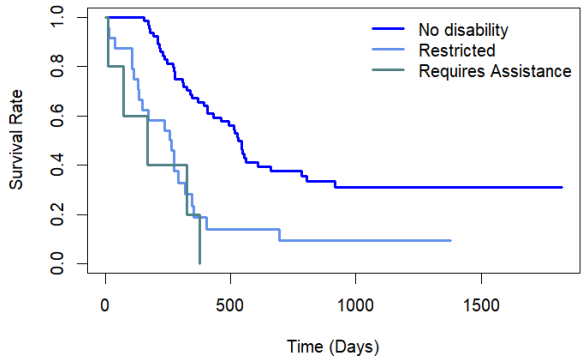


Figure 3.1.10 KM Plot of Site

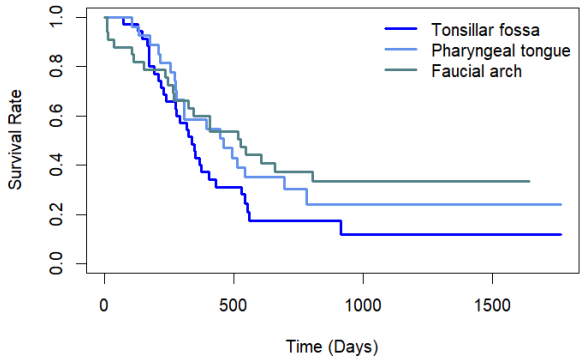


Figure 3.1.11 KM Survival Curve of T Stage

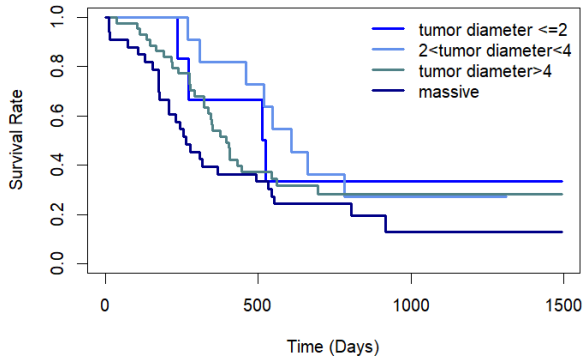


Figure 3.1.12 KM Survival Curve of N Stage

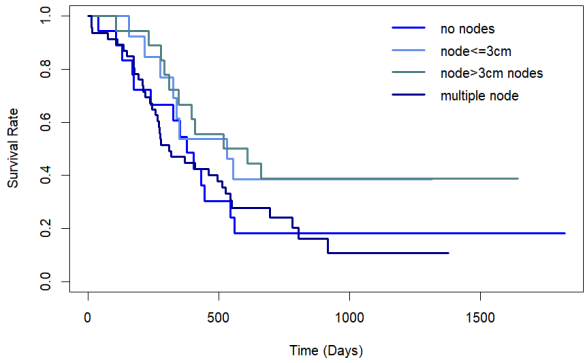


Figure 3.1.13 KM Survival Curve of Sex

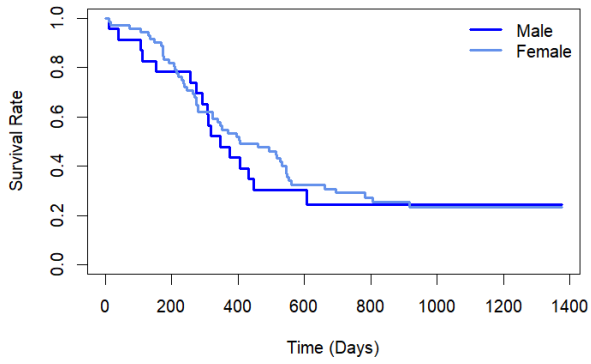
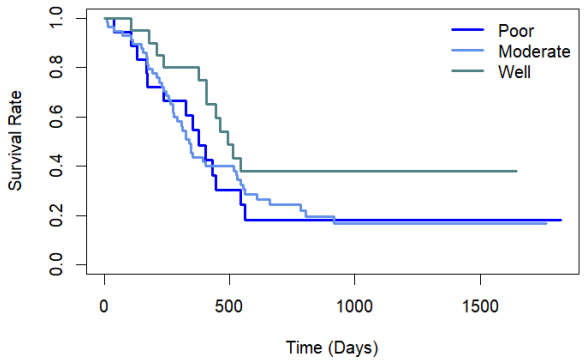


Figure 3.1.14 KM Survival Curve of Grade



Based on the Kaplan Meier Survival Curves, Condition, T Stage, and N Stage show noticeable differences in the survival curves. As expected, the survival curves for more mild magnitudes of the conditions show higher survival rate than those for more severe magnitudes in most cases.

Figure 3.1.15 Scatterplot of Patient Status by Age and Time

0 = Censored
1 = Dead

The majority deaths occurred relatively early in the study, mostly before 750 days. There was no censored data for patients above the age of 80, meaning that older patients have a high death rate, which is to be expected.

3.2 Methods

The methods used in model development are:

- Cox Proportional Hazard Regression Model
- Likelihood Ratio Test
- Testing for Proportional Hazard Assumption
- Stratified Cox Regression Model

The methods used in model evaluation are:

- Hazard Ratio Confidence Interval
- Hypothesis Testing
  - Log-Rank Test
  - Wilcoxon Test
  - Likelihood Ratio Test
- Kaplan-Meier Survival Curve Modeling

3.3 Model Development

A Cox Proportional Hazard Regression Model is a semi-parametric model that will be fitted to the survival data based on the covariates.

$$\lambda(t,\ X)\ =\ \lambda_0(t)e^{\beta X}$$

where X is the vector of covariates and β is a vector of the covariate parameter estimates.

The significant parameters will be selected using the Likelihood Ratio Tests between -2log(L) values.

$$-\ 2log(\frac{L_R}{L_F}) \sim \chi^2_{df=1}$$

The Likelihood Ratio between the full model and reduced model will be compared to a critical value of χ^2 =3.84. If it is greater than the critical value, then the full model is significant.

A backwards selection approach was taken to determine which parameters should be included in the model. A null model and full model containing all the relevant variables was created. Variables were removed one by one from the full model and the Likelihood Ratio Test was conducted with each new model.

Table 3.3.1 -2logL Values for Cox Regression Model Selection

Variable	$-\ 2log\hat{L}$
None	1324.53
Sex + Tx + Grade + Age + Cond + Site + T_Stage + N_Stage	1298.404
Tx + Cond + Site + T_Stage+ N_Stage	1299.527
Cond + Site + T_Stage + N_Stage	1300.909
Cond + T_Stage + N_Stage	1301.856
Cond + T_Stage + N_Stage + T_Stage*N_Stage	1298.413
Age + Cond + T_Stage + N_Stage + T_Stage*N_Stage	1298.272
Age + Cond + T_Stage + N_Stage + Age*Cond	1286.122
Age + Cond + T_Stage + N_Stage + T_Stage*N_Stage + Age*Cond	1281.700
Tx + Age + Cond + T_Stage + N_Stage + T_Stage*N_Stage + Age*Cond	1280.820

The only significant parameters in the full model were Condition, T stage and N stage based on p-value. Thus, the rest of the variables were removed. Interactions between the significant terms were checked, yielding a significant interaction between T stage and N stage. The previously removed variables were added back into the model and tested for interactions as well. This procedure revealed a significant interaction between Age and Condition in the model. Although Tx, the treatment variable, was removed from the model due to its high p-value and small contribution from its Likelihood Ratio Test, it was added back into the model because it is the variable of interest in this study.

Checking PH Assumption

The proportional hazard assumption states that the hazard ratio is constant throughout time. To ensure that the variables satisfy this assumption, log-log survival curves were plotted for the categorical variables and their slopes were observed.

Figure 3.3.1 Log-log Survival Curve for Treatment

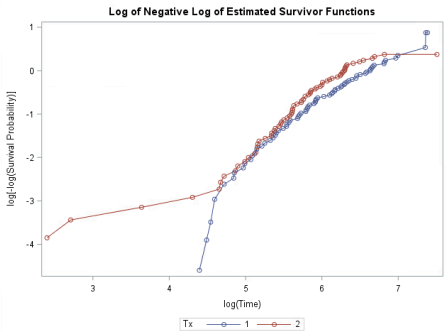


Figure 3.3.2 Log-log Survival Curves for Condition

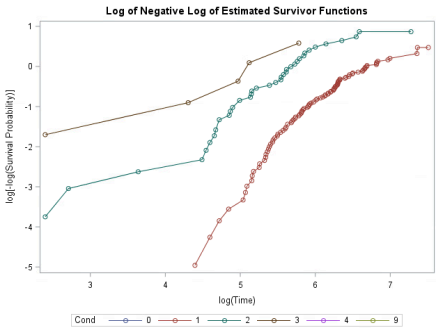


Figure 3.3.3 Log-log Survival Curves for T_Stage

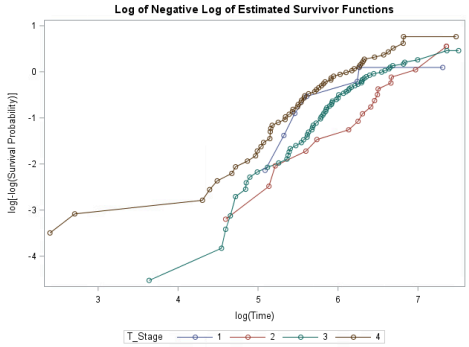
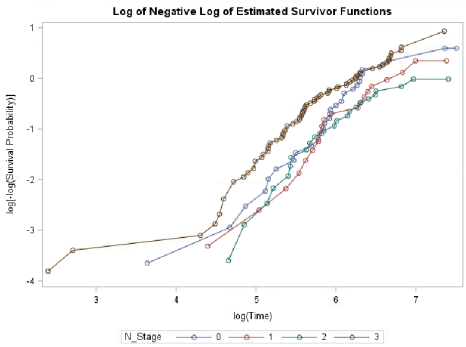


Figure 3.3.4 Log-log Survival Curves for N_Stage



The slopes of the log-log survival curves look similar, except for the presence of a few outliers. Disregarding the outliers, the variables Condition, T Stage, and N Stage satisfy the Proportional Hazard Assumption.

Stratified Cox Regression Model

Stratified Cox Models estimate baseline hazards for each level of the covariate, which accounts for non-proportional covariates. The log-log survival curves for treatment and condition satisfy the PH assumption, while the curves for T Stage and N Stage are borderline cases. An imbalance in classes may be affecting the shape of the line, and it is hard to judge whether or not the overlaps are significant. However, if stratification took place, both T Stage and N Stage would be stratified but the resulting model would be too complex.

Therefore, we accept the final Cox Regression Model:

$$\lambda(t) = \lambda_0(t)e^{0.159Tx + 0.0526Age + 3.36Cond + 0.913Tstage + 1.0Nstage - 0.266Tstage*Nstage - 0.040Age*Cond}$$

3.4 Model Evaluation

Table 3.4.1 Analysis of Maximum Likelihood Estimates for Final Model

Parameter	Parameter Estimate	$Pr > \chi^2$	Hazard Ratio	95% Hazard Ratio Confidence Interval
Treatment	0.159	0.348	1.172	(0.841, 1.634)
Age	0.0525	0.0004	1.054	(1.024, 1.085)
Condition	3.364	<.0001	28.893	(7.098, 117.640)
T Stage	0.912	0.0026	2.491	(1.375, 4.514)
N Stage	1.009	0.0158	2.743	(1.209, 6.225)
T Stage*N Stage	-0.266	0.0290	0.766	(0.603, 0.973)
Age*Condition	-0.0406	<.0001	0.960	(0.942, 0.978)

In this table, the p-values for the majority of the variables are very small, showing their significant contribution to the model. Their hazard ratio confidence intervals also show that they are significant, since most of them do not contain 1. The only exception is the Treatment variable, which remains in the model because it is relevant to the primary question of whether the test treatment is significant.

Table 3.4.2 Median Survival Time for Treatment Groups

Treatment	Median Survival Time $\widehat{S}(t) = 0.5$
Standard (Tx=1)	517
Test (Tx=2)	376

There is a considerable difference in median survival times between the two treatments, with standard treatment having a median survival time 141 days longer than the test treatment. However, there is no proof that this difference is statistically significant as of now, so further analysis of treatment effect must take place.

Hazard Ratio Confidence Interval

The Analysis of Maximum Likelihood Estimates table provides 95% confidence intervals for hazard ratios of each variable. Consider the excerpt from the Analysis of Maximum Likelihood table output from the final model:

Excerpt from Table 3.4.1

	Parameter Estimate	$Pr > \chi^2$	Hazard Ratio	95% Hazard Ratio Confidence Interval
Treatment	0.159	0.347	1.172	(0.841, 1.634)

The hazard ratio indicates that the test treatment would result in a 17% increase in hazard rate when compared to the standard treatment. However, since the 95% Hazard Ratio Confidence Interval includes 1, it is uncertain if the true value of the hazard ratio is below or over 1. Therefore, there is no significant difference between the hazard rates of the two treatments based on hazard ratio confidence interval.

Equality Tests

H_0: The two treatments have no impact on patient survival time

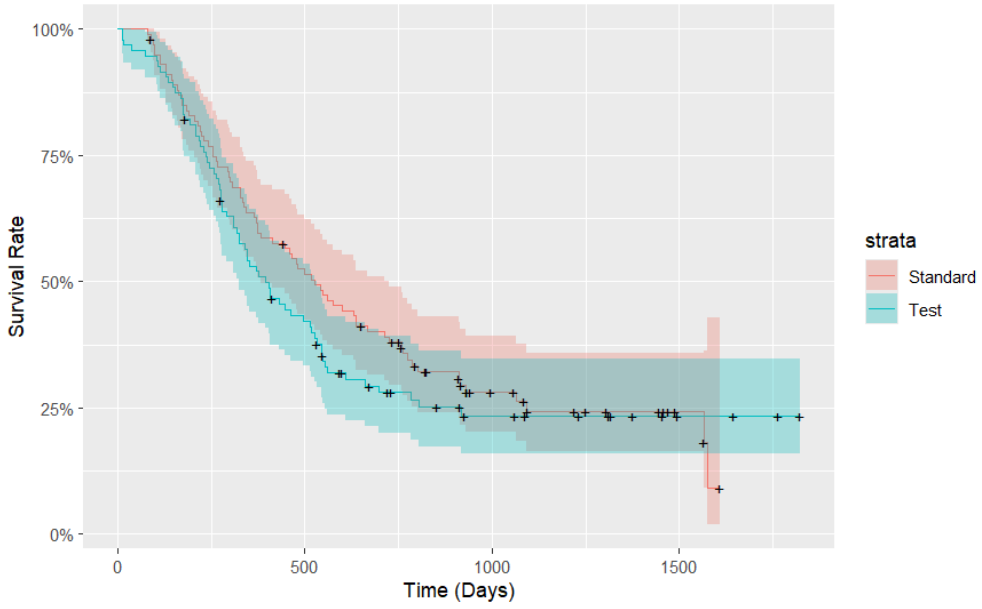
H_1: The two treatments have a significant impact on patient survival time

Table 3.4.3 Log-Rank, Wilcoxon, and Likelihood Ratio Tests for Treatment

Test	p-value
Log-Rank	0.336
Wilcoxon	0.179
-2log(LR)	0.385

The Log Rank test, Wilcoxon test, and Likelihood Ratio test for equality of treatment effect was conducted using the final model. These are different hypothesis tests that compare the survival distributions of two samples. Based on the table above, none of the p-values are significant when tested against the standard significant level of $\alpha = 0.1$. The p-value from the Wilcoxon test is noticeably lower than the other tests, and this difference can be explained by a graphical approach.

Figure 3.4.5 Kaplan-Meier Plot of Survival Estimates for Treatment



At a surface level glance, the curves seem to follow the same general trend with no obvious differences in survival rate. However, at around Time=300, a gap appears between the survival curves of the two treatments. This accounts for the lower p-value from the Wilcoxon test because the Wilcoxon test places greater weights to deaths at earlier time points.

4 Conclusion

The objective of the study was to determine if the two treatments, conventional radiation therapy and conventional radiation therapy with a chemotherapeutic agent, had an effect on survival time in patients with throat cancer. Based on the research and analysis in this report, there is statistical evidence that the two treatments have no significant impact on patient survival time. However, there may be underlying factors still undiscovered that may improve the model's performance, or provide a more in-depth explanation for the results. Stratifying one of the covariates could have also affected the outcome. Overall, the current conclusion is that the impact of the test treatment has no statistically significant impact on patient survival time.

5 Appendix

Citations

"A Clinical Trial in the Treatment of Carcinoma of the Oropharynx." *Radiation Therapy Oncology Group*.

Views, R. "Survival Analysis with R." · *R Views*, 25 Sept. 2017, rviews.rstudio.com/2017/09/25/survival-analysis-with-r/.

Code

Figure 3.1.1 Age Distribution by Sex

```
sns.displot(dat2, x="Age", hue="Sex",
            kind="kde", fill=True,
palette=bicol, aspect=3/2)
```

Figure 3.1.2 Age Distribution by Treatment

```
sns.displot(dat2, x="Age",
hue="Treatment",
            element="bars",
multiple="stack", fill=True,
palette=mycol, aspect=3/2)
```

Figure 3.1.3 Frequency Barplot by General Condition

```
sns.countplot(dat4, x="Cond",
palette=seqcol)
```

Figure 3.1.4 Frequency Barplot by Site

```
sns.countplot(dat4, x="Site",
palette=seqcol)
```

Figure 3.1.5 Frequency Barplot by T Stage

```
sns.countplot(dat4, x="T_Stage",
palette=seqcol)
```

Figure 3.1.6 Frequency Barplot by N Stage

```
sns.countplot(dat4, x="N_Stage",
palette=seqcol)
```

Figure 3.1.7 Frequency Barplot by Sex

```
sns.countplot(dat4, x="Sex",
palette=seqcol)
```

Figure 3.1.8 Frequency Barplot by Grade

```
sns.countplot(dat5, x="Grade",
palette=seqcol, legend=False)
```

Figure 3.1.9 KM Survival Curve by General Condition

```
km_nstage <- survfit(Surv(Time,
Status) ~ Cond, data=trt2)
par(mar = c(4,4,1,1))
plot(km_nstage[2], col="blue",
conf.int=F, xlab="Time (Days)",
ylab="Survival Rate", lw=3)
lines(km_nstage[3],
col="cornflowerblue", conf.int=F,
lw=3)
lines(km_nstage[4], col="cadetblue4",
conf.int=F, lw=3)
lines(km_nstage[5], col="darkblue",
conf.int=F)
legend("topright", legend=c("No
disability", "Restricted", "Requires
Assistance"), levels(trt2$Cond),
col=c("blue", "cornflowerblue",
"cadetblue4"), lty=1, lw=3, bty="n")
```

Figure 3.1.10 KM Survival Curve by Site

```
km_site <- survfit(Surv(Time, Status)
~ Site, data=trt2)
par(mar = c(4,4,1,1))
plot(km_site[1], col="blue",
conf.int=F, xlab="Time (Days)",
ylab="Survival Rate", lw=3)
lines(km_site[2],
col="cornflowerblue", conf.int=F,
lw=3)
lines(km_site[3], col="cadetblue4",
conf.int=F, lw=3)
legend("topright",
legend=c("Tonsillar fossa",
"Pharyngeal tongue", "Faucial arch"),
levels(trt2$Site), col=c("blue",
"cornflowerblue", "cadetblue4"),
lty=1, bty="n", lw=3)
```

Figure 3.1.11 KM Survival Curve by T Stage

```
km_tstage <- survfit(Surv(Time,
Status) ~ T_Stage, data=trt2)
par(mar = c(4,4,1,1))
plot(km_tstage[1], col="blue",
conf.int=F, xlab="Time (Days)",
ylab="Survival Rate", lw=3)
lines(km_tstage[2],
col="cornflowerblue", conf.int=F,
lw=3)
lines(km_tstage[3], col="cadetblue4",
conf.int=F, lw=3)
lines(km_tstage[4], col="darkblue",
conf.int=F, lw=3)
legend("topright", legend=c("tumor
diameter <=2", "2<tumor diameter<4",
"tumor diameter>4", "massive"),
levels(trt2$T_Stage), col=c("blue",
"cornflowerblue", "cadetblue4",
"darkblue"), lty=1, bty="n", lw=3)
```

Figure 3.1.12 KM Survival Curve by N Stage

```
km_nstage <- survfit(Surv(Time,
Status) ~ N_Stage, data=trt2)
par(mar = c(4,4,1,1))
plot(km_nstage[1], col="blue",
conf.int=F, xlab="Time (Days)",
ylab="Survival Rate", lw=3)
lines(km_nstage[2],
col="cornflowerblue", conf.int=F,
lw=3)
lines(km_nstage[3], col="cadetblue4",
conf.int=F, lw=3)
lines(km_nstage[4], col="darkblue",
conf.int=F, lw=3)
legend("topright", legend=c("no
nodes", "node<=3cm", "node>3cm
nodes", "multiple node"),
levels(trt2$N_Stage), col=c("blue",
"cornflowerblue", "cadetblue4",
"darkblue"), lty=1, bty="n", lw=3)
```

Figure 3.1.13 KM Survival Curve by Sex

```
km_sex <- survfit(Surv(Time, Status)
~ Sex, data=trt2)
par(mar = c(4,4,1,1))
plot(km_sex[1], col="blue",
conf.int=F, xlab="Time (Days)",
ylab="Survival Rate", lw=3)
lines(km_sex[2],
col="cornflowerblue", conf.int=F,
lw=3)
legend("topright", legend=c("Male",
"Female"), levels(trt2$T_Stage),
col=c("blue", "cornflowerblue"),
lty=1, lw=3, bty="n")
```

Figure 3.1.14 KM Survival Curve by Grade

```
km_gr <- survfit(Surv(Time, Status) ~
Grade, data=trt2)
par(mar = c(4,4,1,1))
plot(km_nstage[1], col="blue",
conf.int=F, xlab="Time (Days)",
ylab="Survival Rate", lw=3)
lines(km_gr[2], col="cornflowerblue",
conf.int=F, lw=3)
lines(km_gr[3], col="cadetblue4",
conf.int=F, lw=3)
legend("topright", legend=c("Poor",
"Moderate", "Well"),
levels(trt2$Grade), col=c("blue",
"cornflowerblue", "cadetblue4"),
lty=1, lw=3, bty="n")
```

Figure 3.1.15 Scatterplot of Status by Age and Time

```
sns.scatterplot(data=dat, x="Age",
y="Time", hue="Status",
style="Status",
palette=sns.diverging_palette(250,
15, n=2))
```

Table 3.3.1 -2logL Table for Model Selection

```
*null model;
proc phreg data=pharynx;
    model time*status(0) = /RL;
run;

proc phreg data=pharynx;
    model time*status(0) = tx/RL;
run;

*full model;
proc phreg data=pharynx;
    model time*status(0) = sex tx
grade age inst cond site t_stage
n_stage/RL;
run;

*reduced model 1;
proc phreg data=pharynx;
    model time*status(0) =  tx inst
cond site t_stage n_stage/RL;
run;

*reduced model 2;
proc phreg data=pharynx;
    model time*status(0) = tx cond
site t_stage n_stage/RL;
run;

*reduced model 3;
proc phreg data=pharynx;
    model time*status(0) = cond
t_stage n_stage/RL;
run;

*adding interaction term;
data phar2;
    set pharynx;
    int = t_stage*n_stage;
    int2 = age*cond;
run;
proc phreg data=phar2;
    model time*status(0) = tx age
cond t_stage n_stage int int2/RL;
run;
```

Figure 3.3.1 Log-log Survival Curves for Treatment

```
proc lifetest data=phar2 plots=(lls);
    time time*status(0);
    strata tx;
```

```
run;
```

Figure 3.3.2 Log-log Survival Curves for Condition

```
proc lifetest data=phar2 plots=(lls);
    time time*status(0);
    strata cond;
run;
```

Figure 3.3.3 Log-log Survival Curves for T Stage

```
proc lifetest data=phar2 plots=(lls);
    time time*status(0);
    strata t_stage;
run;
```

Figure 3.3.4 Log-log Survival Curves for N Stage

```
proc lifetest data=phar2 plots=(lls);
    time time*status(0);
    strata n_stage;
run;
```

Table 3.4.1 Analysis of Maximum Likelihood Estimates for Final Model

```
proc phreg data=phar2;
    model time*status(0) = tx age
cond t_stage n_stage int int2/RL;
run;
```

Table 3.4.2 Median Survival Times for Treatment Groups

```
proc lifetest data=phar2 plots=(lls);
    time time*status(0);
    strata tx;
run;
```

*Median was obtained by finding survival rate=0.5 from the Product-Limit Survival Estimates table in the output, and recording the corresponding time

Table 3.4.3 Table for Equality Tests

```
proc lifetest data=phar2 plots=(lls);
    time time*status(0);
    strata tx;
run;
```

Figure 3.4.5 Kaplan Meier Survival Curve for Treatment Groups

```
km_trt_fit <- survfit(Surv(Time,
Status) ~ Tx, data=dat2)
autoplot(km_trt_fit, xlab="Time
(Days)", ylab="Survival Rate")
```