

# STAT 590 Survival Analysis - Final Project

A clinical Trial in the Treatment of Carcinoma of the  
Oropharynx

Andres Gonzalez  
Yushan Zhao

May 13, 2023



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

# Table of contents

<b>1</b>	<b>Abstract</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>3</b>
<b>3</b>	<b>Exploratory Data Analysis</b>	<b>4</b>
3.1	Age . . . . .	4
3.2	Gender . . . . .	6
3.3	Time and Status . . . . .	7
<b>4</b>	<b>Methods</b>	<b>8</b>
<b>5</b>	<b>Results</b>	<b>9</b>
5.1	Model Selection . . . . .	9
5.2	PH Assumption Check . . . . .	11
5.3	Kaplan Meier Survival Curve For Treatment . . . . .	15
5.4	Research Questions . . . . .	16
<b>6</b>	<b>Conclusion</b>	<b>18</b>

# 1 Abstract

This is a report of 195 patients with squamous carcinoma of 15 sites in the mouth and throat, with 16 participating institutions, though only data on three sites in the oropharynx reported by the six largest institutions are considered here. Patients entering the study were randomly assigned to one of two treatment groups, radiation therapy alone or radiation therapy together with a chemotherapeutic agent. Stepwise regression was used to build up the model, and a cox regression proportional hazard model was adopted to analyze the significance of different variables. Dependent variable is survival time, and independent variables include T staging, N Staging, Sex, Age, Grade, General Condition, and Treatment. After the stepwise regression, only T staging, N Staging, Grade, General Condition, and Treatment are added to the final model. One goal of the study was to compare the two treatment policies with respect to patient survival. After controlling for T Staging, N Staging, Grade, and General Condition, results indicated that treatment does not play a significant role in patients' survival time.

## 2 Introduction

Pharynx cancer is a significant health challenge, impacting countless individuals globally. Survival analysis focuses on examining the time interval between a participant's entry into a study and the occurrence of a subsequent event. Treatment differences play a vital role in influencing survival rates and enhancing patients' quality of life. This study endeavors to perform a comprehensive survival analysis on pharynx cancer patients, contrasting the effectiveness of combined treatment approaches with traditional radiation therapy. Additionally, the interconnections between various patient and tumor attributes (covariates) and survival rates will be investigated to gain a deeper understanding of their influence on treatment outcomes.

Furthermore, 30% of the survival periods in this study are censored, primarily due to patients still being alive at the time of analysis. A limited number of patients were lost to follow-up as a result of moving or switching to institutions not involved in the study, although these cases were relatively uncommon. Even with the implementation of specific eligibility criteria to eliminate patients with severe disease progression, a substantial degree of heterogeneity persists among the study participants, leaving numerous factors unaccounted for.

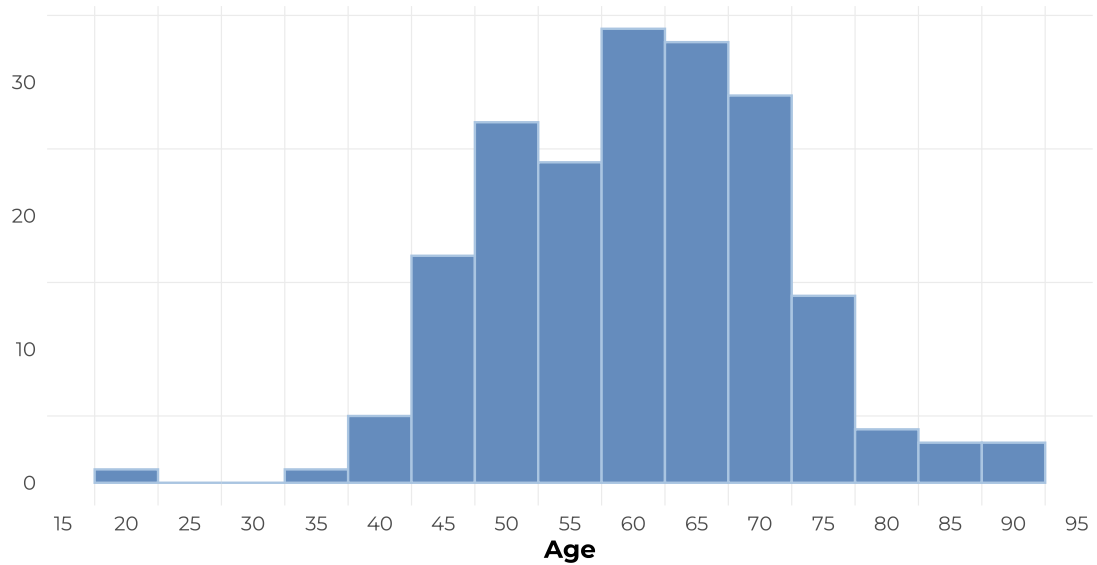
This study encompasses seven critical covariates believed to influence survival outcomes, comprising sex, T staging, N staging, age, general condition, grade and treatment. In the TX Treatment variable, 1 represents the standard treatment, while 2 indicates the test treatment. The T and N staging classification system quantifies the tumor's size at the primary site and involvement of regional lymph nodes, while the general condition variable estimates the patient's functional ability at diagnosis. T=1 denotes a small primary tumor with a diameter of 2 centimeters or less, while T=4 signifies a massive tumor extending into adjacent tissues. T=2 and T=3 represent intermediate cases. N=0 implies no clinical evidence of lymph node metastasis, and N=1, N=2, N=3 designate increasing degrees of lymph node involvement. Patients classified as T=1, N=0; T=1, N=1; T=2, N=0; or T=2, N=1, as well as those with distant metastases, were excluded from the study.

Beyond the primary focus on the comparative effectiveness of combined treatment versus traditional radiation therapy, it is essential to pinpoint the covariates that have the most considerable impact or exhibit the most robust correlation with survival results. In this context, conducting an exploratory analysis proves valuable for gaining deeper insights into the variables and recognizing underlying patterns into the study which addresses in the following section.

## 3 Exploratory Data Analysis

### 3.1 Age

**Age Distribution**



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

The age histogram exhibits a normal distribution, with a potential outlier at age 20. The age range covers individuals from their twenties up to their nineties, and the mean age is 60.4.

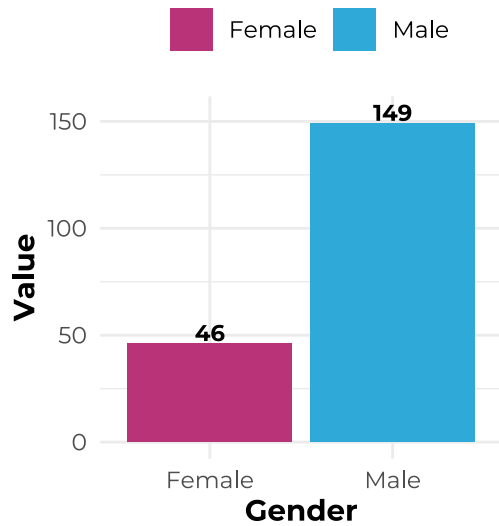


Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset

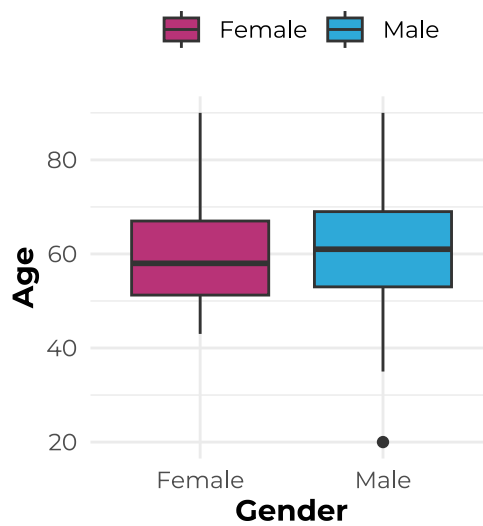
Upon examining the plots, distinct patterns in age distributions are observed. The grade differentiation plot reveals a higher prevalence of moderately differentiated individuals compared to other grades. The condition plot indicates that most individuals belong to the “no disability” category. The site distribution presents a balanced ratio among faucial arch, tonsillar fossa, and pharyngeal tongue. The T staging plot displays an even distribution for “more than 4cm” and “massive invasive tumor” classifications within the 40-80 age range. Lastly, the N staging plot exhibits an equitable distribution across all classifications.

## 3.2 Gender

**Distribution of Gender**



**Age Distribution by Gender**

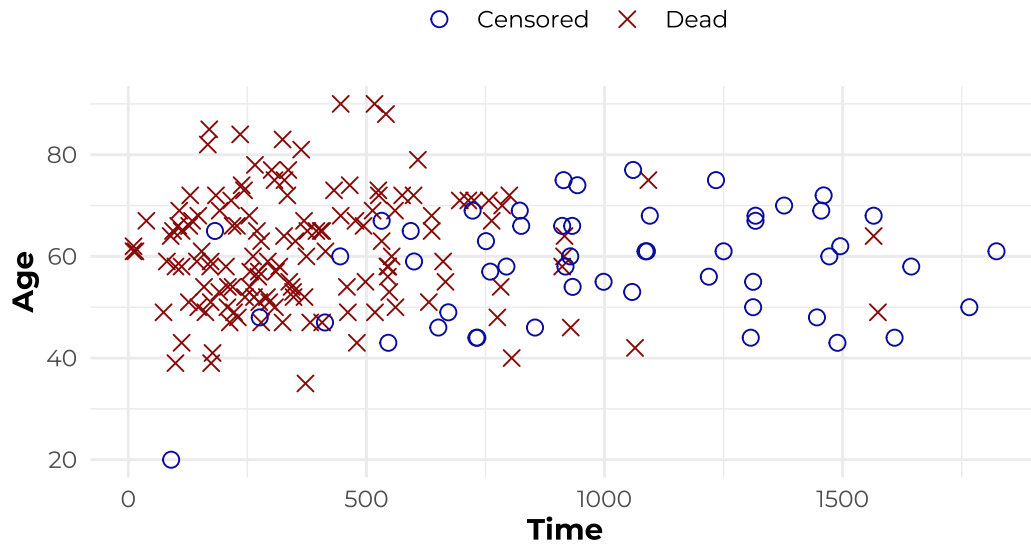


*Source: Clinical Trial in the Treatment of Carcinoma of the Oropharynx Dataset*

The distribution of gender reveals a higher count of males compared to females, with 149 males and 46 females. In the age distribution of gender, both genders display a relatively even distribution, with the exception of a single outlier in their 20s, as mentioned previously.

### 3.3 Time and Status

#### Scatter Plot of Time and Age by Status



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

The plot above suggests that, regardless of age, individuals who experience an event tend to have a shorter time to event, with only a few reaching over 1000 days. In contrast, those who are censored generally exhibit a longer duration before being censored.



## 4 Methods

Stepwise selection, utilizing the likelihood ratio test as a statistic, was employed to develop the model, while the Cox regression proportional hazard model was used for the analysis of this dataset.

## 5 Results

### 5.1 Model Selection

Table 5.1: Model Comparison Based on -2 Log-Likelihood

Variables	$-2\log\hat{L}$
None	1324.34
Sex	1323.64
T Staging	1316.46
N Staging	1321.20
Age	1323.89
General Condition	1314.33
Grade	1324.05
Treatment	1323.41
N Staging + Age + General Condition + Grade + Treatment	1307.94
T Staging + Age + General Condition + Grade + Treatment	1306.35
T Staging + N Staging + General Condition + Grade + Treatment	1300.99
T Staging + N Staging + Age + Grade + Treatment	1309.31
T Staging + N Staging + Age + General Condition + Treatment	1302.07
T Staging + N Staging + Age + General Condition + Grade	1301.92
T Staging + N Staging + Age + General Condition + Grade + Treatment	1301.95
T Staging + N Staging + General Condition + Grade + Treatment + Sex	1299.90

Stepwise regression is employed to develop the Cox regression proportional hazard model. This approach constructs a series of Cox regression models, incorporating or excluding specific variables based on their impact on the -2log likelihood rate. Decisions are made by determining if the variable significantly reduces the -2log likelihood rate.

The statistical method used for the hypothesis test for the reduced model and the full model is the Likelihood Ratio test (LRT). In the first step, each variable is added to the model one at a time, and then those models with separate variables are compared with models with null variables. From Table 1, it can be seen that, when compared with the null model, the model with T staging, N staging, Age, General Condition, Grade, and

Treatment (but not Sex) has a significant reduction in -2log likelihood rate. Thus, Sex is excluded from the model at step 1.

Afterward, a model with those significant variables (let's say  $n$  variables here) from the previous step is compared and contrasted with models with  $n-1$  variables. In this step, variables are included in the final model if they contribute the most to the reduction in -2log likelihood rate. For example, comparing the model with N Staging, General Condition, Grade, Age, and Treatment to the model with T Staging, N Staging, Age, General Condition, Grade, and Treatment, there is a significant decrease in -2log likelihood rate from 1307.94 to 1300.96, with a 6.96 drop. This indicates that the variable T Staging is important and adds a lot of information to the model, so it cannot be removed.

At step 3, the variables indicating non-significance at step 1 are added back to the model. Thus, Sex is added to the model. The model with T Staging, N Staging, General Condition, Grade, and Treatment is compared with the model with T Staging, N Staging, General Condition, Grade, and Sex. The result shows that adding Sex to the model doesn't add information to explain variation in the model. Consequently, it is decided not to include Sex in the model. The process ends when the final model is identified. The final model includes: T Staging, N Staging, General Condition, Grade, and Treatment.

Model Comparison: AIC Values for Reduced, Multivariate and Reduced with Sex Models		
Model	df	AIC
<b>T Staging + N Staging + General Condition + Grade + Treatment</b>	<b>5</b>	<b>1310.992</b>
T Staging + N Staging + Age + General Condition + Grade + Treatment	6	1312.954
T Staging + N Staging + General Condition + Grade + Treatment + Sex	6	1311.900
<i>Pharynx Dataset</i>		

Figure 5.1: AIC Values for Selected Models

Also, the AIC comparison indicates that the model with five predictors (T Staging, N Staging, General Condition, Grade, and Treatment) is the best fit among the three, as it has the lowest AIC value. The small differences in AIC values suggest that the full model containing age or adding Sex to the reduced model does not significantly improve the model fit, making the first model the most parsimonious choice.

The final model is:

$$\lambda(t, x_1, x_2, x_3, x_4, x_5) = \lambda_0(t) \times e^{0.296x_1 + 0.170x_2 + 0.221x_3 - 0.122x_4 + 0.168x_5}$$

where

$x_1$  = T Staging (1 if 2 cm or less, 2 if 2 cm to 4 cm, 3 if more than 4 cm, and 4 if massive invasive tumor)

Patient Variables (N = 195)			
Variable	HR <sup>†</sup>	95% CI <sup>†</sup>	P-Value
T staging	1.34	1.07, 1.68	0.010
N staging	1.19	1.02, 1.37	0.022
General Condition	1.25	1.11, 1.40	<0.001
Grade	0.89	0.70, 1.12	0.3
Treatment	1.18	0.85, 1.65	0.3
<sup>†</sup> HR = Hazard Ratio, CI = Confidence Interval			
<i>Pharynx Dataset</i>			

Figure 5.2: Summary of the Cox Regression Proportional Hazard Model

$x_2$  = N Staging (0 if no clinical evidence of node metastases, 1 if single positive node 3 cm or less in diameter, not fixed, 2 if single positive node more than 3 cm in diameter, not fixed, and 3 if multiple positive nodes or fixed positive nodes)

$x_3$  = General Condition (1 if no disability, 2 if restricted work, 3 if requires assistance with self care, 4 if bed confined, and 9 if missing)

$x_4$  = Grade (1 if well differentiated, 2 if moderately differentiated, 3 if poorly differentiated, 9 if missing)

$x_5$  = Treatment (1 if standard, 2 if test)

## 5.2 PH Assumption Check

Before finalizing the model, it is highly necessary to check the PH assumption to ensure that it is appropriate to apply Cox regression PH model for this data set.

Observing the above plots, it is evident that the solid lines provide good fits to these plots. More specifically, the dashed lines indicate positive or negative 2 standard errors enclosed by the fit lines. These plots imply no obvious concerns and patterns as time passes by. This means that the PH assumption is satisfied, making it appropriate to use the Cox regression PH model for data analysis.

The log-log survival curve for each variable is also run, as shown below. These curves are overall parallel, which indicates that the PH assumption is satisfied as well.

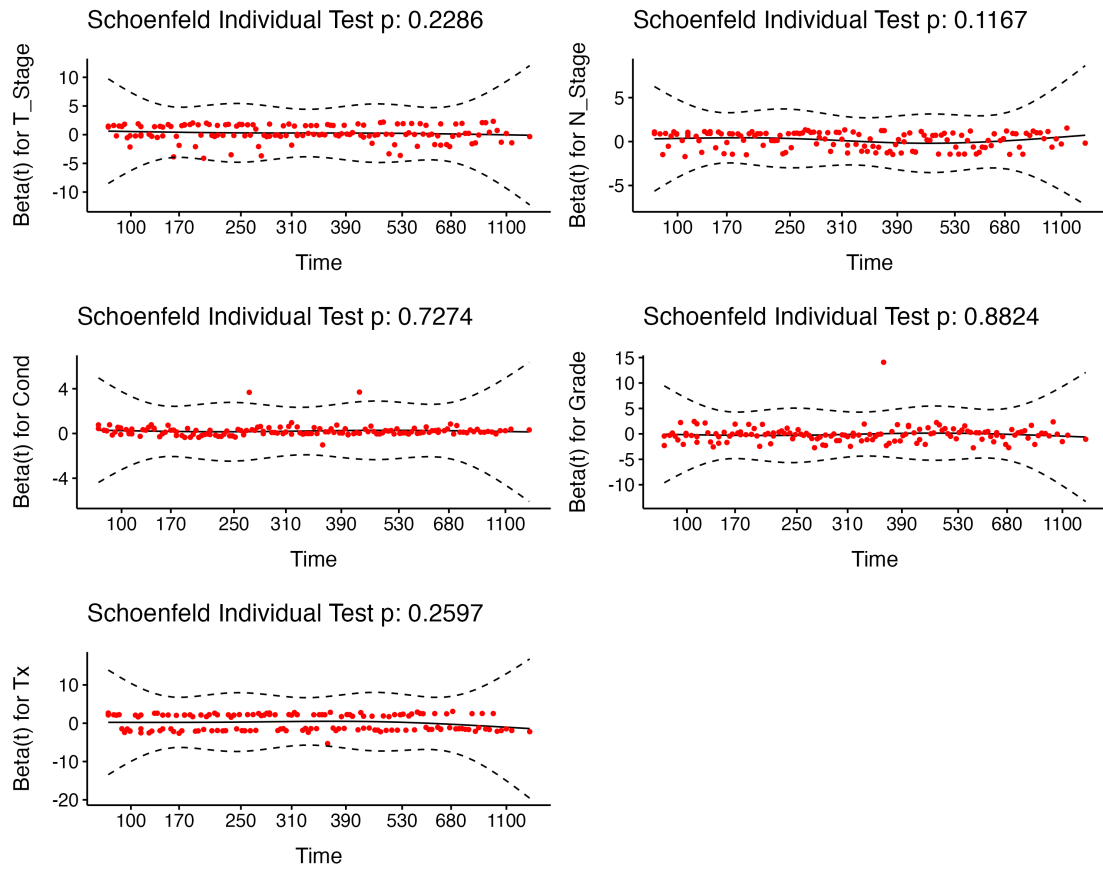
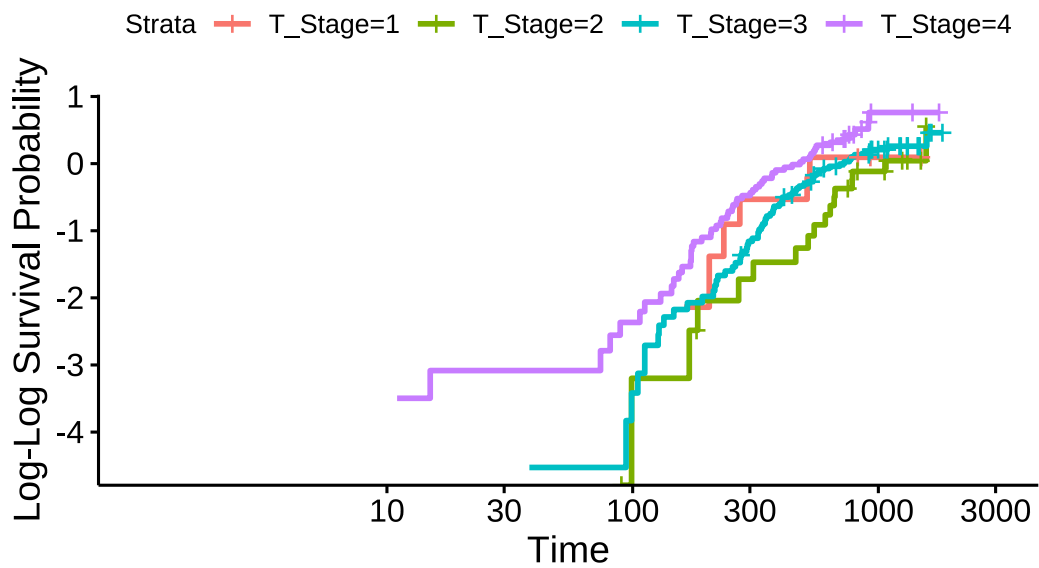
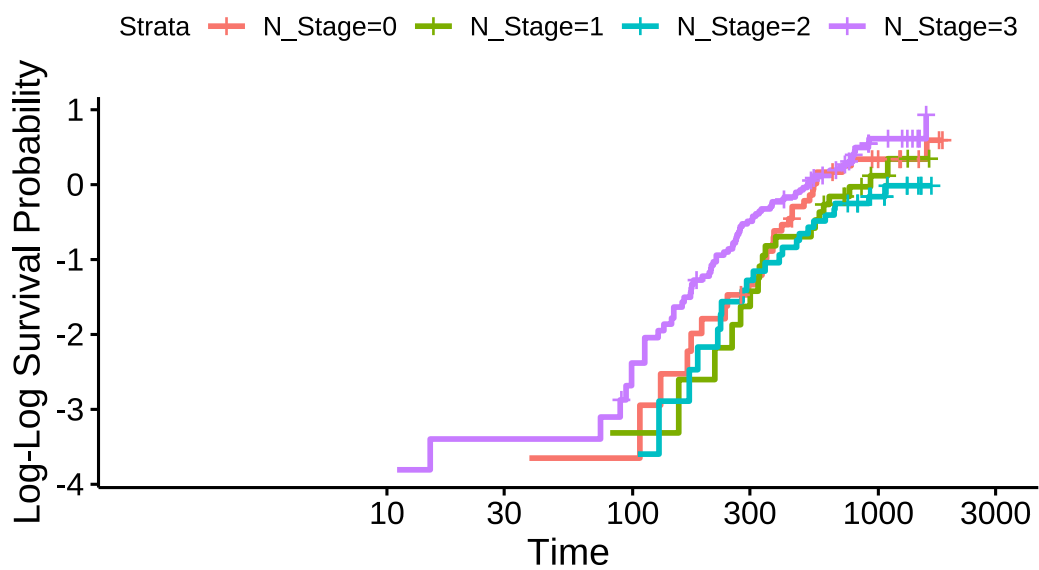


Figure 5.3: Schoenfeld Individual Test for Each Variable

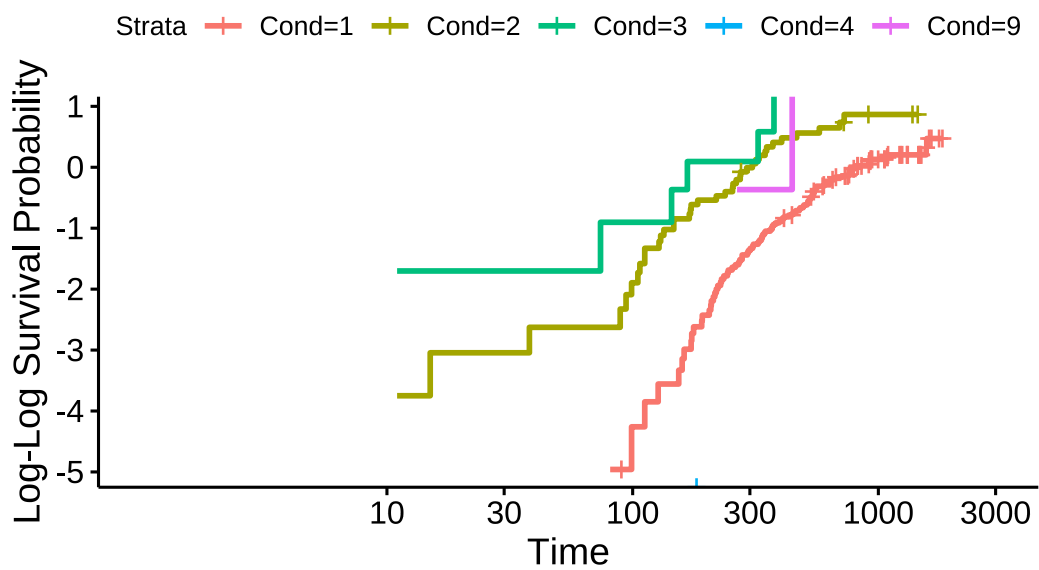
## Log-Log Kaplan-Meier Survival Estimates for T Sta



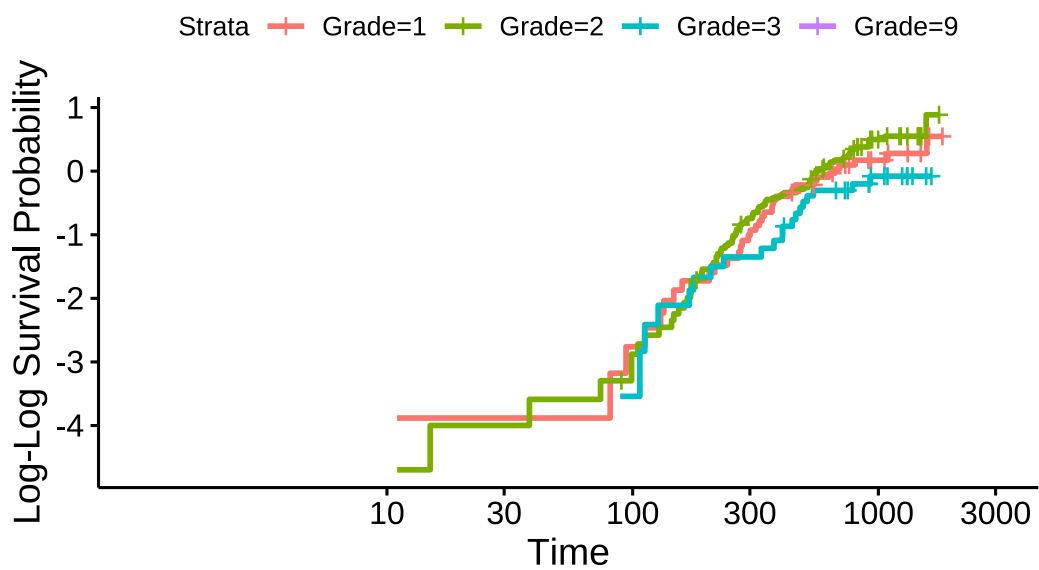
## Log-Log Kaplan-Meier Survival Estimates for N Sta



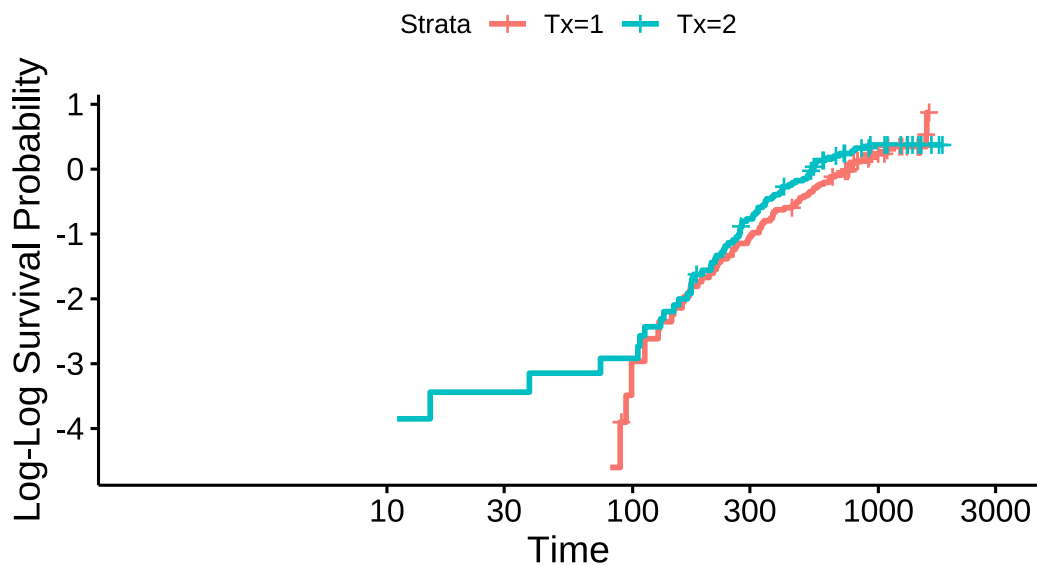
## Log-Log Kaplan-Meier Survival Estimates for Gene



## Log-Log Kaplan-Meier Survival Estimates for Grad

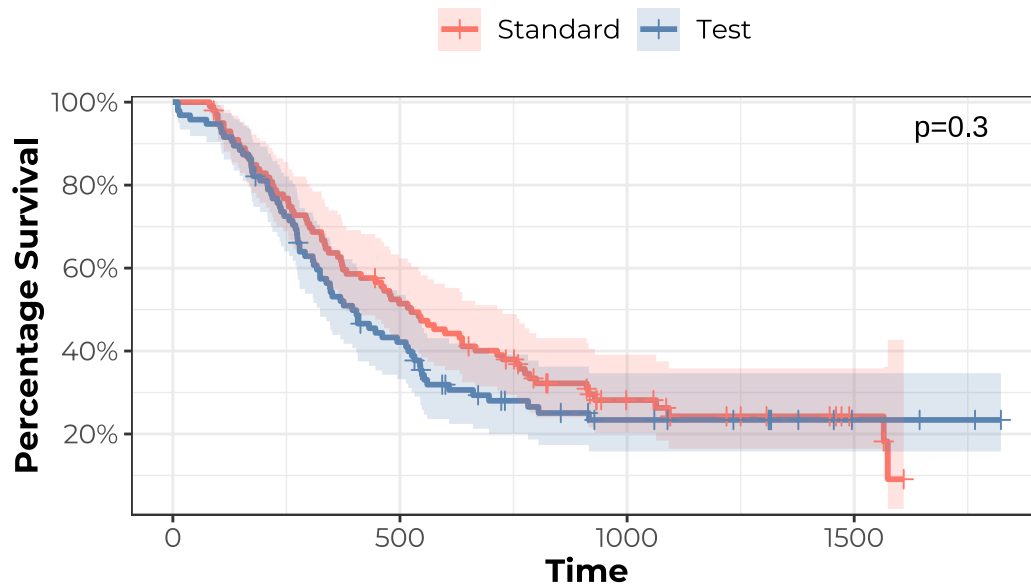


## Log-Log Kaplan-Meier Survival Estimates for Treat



### 5.3 Kaplan Meier Survival Curve For Treatment

#### Kaplan-Meier Survival Curve by Treatment



From the Kaplan-Meier curve plot, it becomes evident that there is no significant difference in treatment outcomes, as both lines overlap and remain quite similar. Furthermore, after 1000 days, the survival rate plateaus at a relatively low level of 20%.



## 5.4 Research Questions

Find the hazard ratio for an individual on a particular treatment relative to another individual on a different treatment with the same T staging, same N staging, same General Condition, and same Grade. Also construct a 95% confidence interval.

The difference in the log-hazard function for the two groups can be written as:

$$\log \lambda(t, x_1, x_2 = 1, x_3 = 1, x_4 = 1, x_5 = 2) - \log \lambda(t, x_1, x_2 = 1, x_3 = 1, x_4 = 1, x_5 = 1)$$

we get

$$\begin{aligned} & (\log \lambda_0(t) + \beta_1 + \beta_2 + \beta_3 + \beta_4 + 2\beta_5) \\ & - (\log \lambda_0(t) + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5) \\ & = \beta_5 \\ & = 0.16746 \end{aligned}$$

The point estimate for the hazard ratio (HR) is:

$$\hat{\text{HR}} = e^{0.16746} = 1.18$$

This suggests that the hazard rate for patients who received the testing treatment is 16.7% of the patients who received the standard treatment.

The confidence interval for the hazard ratio is:

$$e^{0.16746 \pm 1.96(0.17119)} = (e^{-0.1681}, e^{0.50299}) = (0.8453, 1.6537)$$

The confidence interval suggests that the hazard ratio is as low as 0.8453 and as high as 1.6537 at  $\alpha = 0.05$ . The confidence interval of the hazard ratio includes 1, which means that the treatment effect is not significant.

The median survival times on figure 5.4 vary across independent variables. Patients with smaller T staging tumors and a single positive node have longer survival times. Those without disabilities and with poorly differentiated tumors also fare better. Patients receiving radiation therapy alone tend to survive longer than those receiving combined treatments.

Median Survival for Key Variables	
Variable	Median Survival (95% CI)
T staging	
1	513 (235, —)
2	782 (608, —)
3	465 (376, 631)
4	317 (254, 480)
N staging	
0	446 (374, 561)
1	587 (347, —)
2	661 (465, —)
3	327 (270, 494)
General Condition	
1	553 (517, 763)
2	255 (170, 317)
3	156 (74, —)
4	— (—, —)
9	356 (266, —)
Grade	
1	432 (336, 696)
2	404 (324, 545)
3	544 (446, —)
9	363 (—, —)
Treatment	
1	525 (382, 714)
2	395 (324, 532)
Pharynx Dataset	

Figure 5.4: Median Survival at 95% Confidence Interval

## 6 Conclusion

The objective of this report is to examine and compare the treatment effects between two treatments: radiation therapy alone and radiation therapy combined with a chemotherapeutic agent. The results suggest that these treatments do not significantly impact the survival rate of patients with pharynx cancer. Other factors, such as family history of pharynx cancer or eating habits, might play a more significant role in determining the survival rate of these patients. Nevertheless, acknowledging the model's limitations and the data used in the analysis is essential. Pursuing further research and validation with larger datasets can help improve the model's accuracy and applicability in a wider context.