Lung Cancer Treatment Report

Submitted by: Hammad Muniem

# Objective

The objective of our study is **to guage the efficacy of the test treatment as compared to the standard treatment for advanced inoperable lung cancer and to understand how different charactersitics of the patient would affect the number of days the patient will survive**. In order to do this, we would be estimating the number of days the patient can survive depending on his choice to take the new test treatment or going for the standard treatment while also investigating other factors that might play a role.

# Method

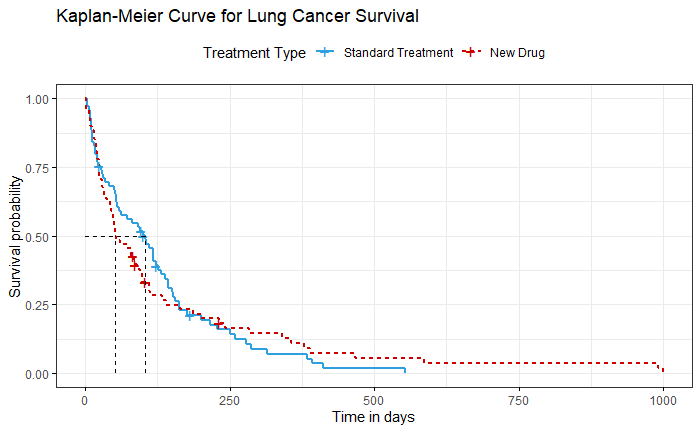
We would be using **data from a Veteran's Administration Lung Cancer Trial**, where 137 patients with advanced, inoperable lung cancer were treated with chemotherapy (standard treatment) vs chemotherapy combined with a new drug (test treatment). This data contains information on:

* Age of the patient
* Months from diagnosis
* Cell Type
* Karnofsky Score
* Prior Chemotherapy
* Number of days of survival

**We would be using various statistical techniques on this sample data to develop a data model that predict the number of days a specific patient will survive.** If the new drug is effective, then our data model will predict that if the patient opts out for the new treatment, he would survive for more days than he would undergoing Chemotherapy alone.

# Model 1 – Basic Model

We initially only considered the effect of the test treatment on days of survival while keeping all other factors constant.



The graph above shows our results for this initial model. It shows the number of days and the corresponding chance of survival for that specific number of days for both types of treatment. So, for example if we look at 250 days, the lines for the respective treatments show the chances that the patient will survive 250 days.

Some of the results for this model are shown below. **We are not considering any other factors such as age of patient etc at this point in time.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Chances that a patient has survived for a given number of days for each treatment** | | **Treatment Type** | |
| **Standard Treatment** | **New Drug** |
| **Days** | **365** | 6.7% | 9.8% |
| **183** | 19.9% | 23.1% |

The median survival i.e. the time at which half of the patients in the respective treatment groups are expected to be alive is also calculated. **The median survival time for standard treatment is an estimated 100 days while the median survival time for the new drug treatment is an estimated 52 days.**

Since there can be differences in the kind of patients that are opting for the new drug treatment, for example it is possible that only people with poor Karnofsky score are opting for the treatment. This would mean that the efficacy of the new drug would be underestimated in the previous model. This is why we need to dig deeper into the problem.

# Model 2 – Advanced Models

In this step, we attempted to include other characteristics of the patient to check how other factors (Age of the patient and months from diagnosis) play a part in the number of days the patient will survive.

We ran many models and have selected the top 3. The results of these models are shown below. The effects of each factor presented are independent meaning that one factor is changed at a time and the other factors are kept constant. So, for instance, if we are talking about the effect of age, then we would be talking about the changes caused by Age only assuming that treatment type and months from diagnosis are kept constant.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Models** | | |
| **Factors** | **Cox Model** | **Exponential Model** | **Weibull Model** |
| **New Treatment** | Patients who opt for the test treatment have a **0.3% more chance of surviving** till a specific day than the standard treatment. | Patients who opt for the test treatment have a **11% more chance of surviving** till a specific day than the standard treatment. | Patients who opt for the test treatment have a **7% more chance of surviving** till a specific day than the standard treatment. |
| **Age** | All models predict that **increasing age by 1 year would reduce the chances of survival by 1%** at a specific day than the person who is 1 year younger. | | |
| **Months from diagnosis** | All models predict **that increasing months from diagnosis by 1 month** **would reduce the chances of survival by 1%** at a specific day than the person who was diagnosed 1 month earlier. | | |

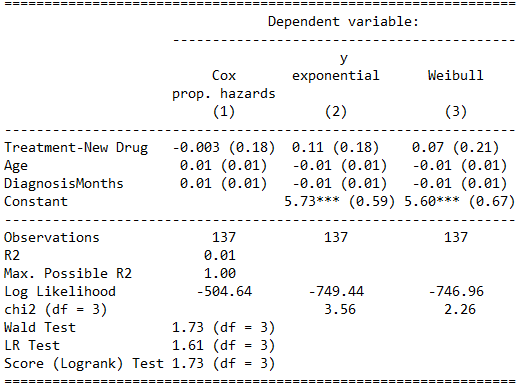
The effects of the 3 aforementioned factors are constant over time. Please note that we also tried to factor in Karnofsky Score and Cell type but their effect on the survival days is not constant over time so their effect would be needed to be investigated further.

# Conclusion

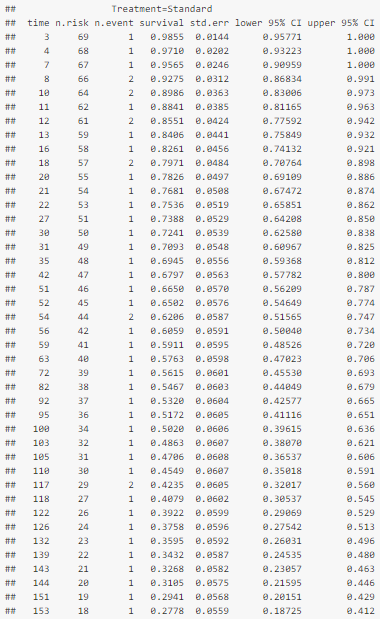
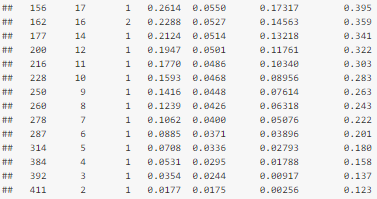
To conclude, all our models had the same predictions for Age and Months from diagnosis. Our models differ slightly when it comes to gauging efficacy of the test treatment. All models estimate that the test treatment would improve the chances of survival for the patient by varying amounts. The effect of the test treatment needs to be further investigated to make a better prediction but according to our models, the Test treatment does seem to work since it shows improvement in patients’ chances across 3 models.

# Appendix

## Details of the advanced models



## Probabilities for each day-Basic Model – Standard Treatment

## Probabilities for each day-Basic Model – Test Treatment

