

# **Exploring the Impact of Blood Transfusion and Lung Transplantation**

*BTC 1877 Team Project*  
*Dr. Nicholas Mitsakakis*  
*December 2, 2024*

Mausam Vadakkayil,  
Vian Tran,  
Ahmed Mokhtar

## 1. Introduction

Patients with advanced lung diseases who have exhausted other medical therapies and experience severe reductions in quality of life may consider lung transplantation as a treatment option to replace one or both of their damaged lungs with healthy donor lung(s). In Canada, there are currently five lung transplant centres located in Winnipeg, Edmonton, Vancouver, Toronto and Montreal; considering the limited transplantation centres there is a rigorous assessment process used to determine eligibility (Canadian Pulmonary Fibrosis Foundation, n.d.). In Ontario, an adult patient may be eligible to be waitlisted for lung transplants if they have the following advanced-stage lung diseases: Chronic Obstructive Pulmonary Disease (COPD), Interstitial Lung Diseases, Airway Diseases and Pulmonary Hypertension (Provincial Lung Working Group Transplant Steering Committee, 2024).

Despite being an established therapeutic procedure for end-stage lung disease, the success of lung transplantation can be impaired by perioperative complications such as primary graft dysfunction (PGD), infections and blood loss (Oechslin et al., 2018). Current literature particularly highlights the impact of perioperative blood transfusions on adverse outcomes for lung transplant patients, such as increased likelihood for PGD or transfusion-related acute lung injury (Siddiqui & Shakil, 2024; Toy et al., 2012). A literature review was performed to understand the relationship between blood transfusions and lung transplantation, with a focus on the characteristics and outcomes of lung transplant patients who received blood transfusions.

Oechslin et al. (2018) performed a narrative review to determine factors influencing blood loss in lung transplantation that lead to transfusions. The authors identified greater transfusion requirements for double-lung transplants, which can be attributed to the greater surgical complexity of double compared to single-lung transplants. Furthermore, patients with pulmonary hypertension and cystic fibrosis require perioperative blood products more frequently, which can also be attributed to these patients requiring a double lung transplant (Klapper et al., 2021; Oechslin et al., 2018). A recurrent transfusion factor in the literature was also extracorporeal membrane oxygenation (ECMO) and cardiopulmonary bypass (CPB) both of which support lung transplant patients who present with low cardiac output, refractory hypoxemia, and other cardiopulmonary instabilities (Ang et al., 2009; Klapper et al., 2021; Oechslin et al., 2018). Studies suggest patients who receive preoperative ECMO and CPB are more likely to receive large-volume transfusions as they require more daily red blood cells and fresh frozen plasma (FFP) (Ang et al., 2009).

For lung transplant patients who received transfusions perioperatively, the literature highlights adverse outcomes for this group in comparison to non-transfusion lung transplant patients. Siddiqui & Shakil (2024) found that PGD, intensive care unit (ICU) days, need for mechanical ventilation and ECMO support were significantly greater in patients who received blood transfusion perioperatively. Zalunardo et al. (2011) echoed this finding; the authors found that blood-product use was associated with greater mechanical ventilation and ICU stay. Multiple studies also indicated that intraoperative transfusion was a predictive factor for in-hospital death and mortality, compared to non-transfusion lung transplant groups (Klapper et al., 2021; Weber et al., 2013; Zalunardo et al., 2011).

The literature review provides an overview of the characteristics of lung transplant patients who received blood transfusion and associated outcomes post-transplantation, which helped guide our analysis. The aims of this study are to:

1. Identify predictors that influence the (1) need for transfusions and (2) amount of transfusion and compare it with relevant literature.
2. Investigate the impact of transfusions and confounding variables on patient outcomes such as mortality and ICU length of stay.

## 2. Methods

Data was retrieved from 192 lung transplant patients. The dataset includes variables related to patient demographics, underlying respiratory diagnosis, intraoperative descriptions, preoperative blood work, the usage of extracorporeal life support, blood product transfusion of patients, survival and ICU length of stay. The variables can be summarized in **Table 1**. Variables not following the categories mentioned above were removed. The variables were renamed for consistency before plotting and data analysis.

**Table 1:** Variables of interest from raw data

Variable Categories and Extracted Variables	
Variable Category	Extracted Variables
Patient Demographic Data	1. Gender 2. Height 3. Weight 4. Age 5. Body Mass Index (BMI)
Underlying Respiratory Diagnosis and Intraoperative Descriptions	1. Chronic obstructive pulmonary disease (COPD) 2. alpha1-Antitrypsin Deficiency 3. Cystic Fibrosis 4. Idiopathic Pulmonary Hypertension 5. Interstitial Lung Disease 6. Other Pulmonary Disease 7. First Lung Transplant 8. Redo Lung Transplant 9. ExVIVO Lung Perfusion 10. Preoperative Extracorporeal Life Support (ECLS) 11. Lung Allocation Score (LAS score) 12. Intraoperative ECLS
Life Support	1. ECLS Extracorporeal Membrane Oxygenation (ECMO) 2. ECLS Cardiopulmonary Bypass (CPB)
Blood Product Transfusion Data	1. Intra Fresh Frozen Plasma 2. Intra Packed Cells 3. Intra Platelets 4. Intra Cryoprecipitate 5. Red Blood Cell (RBC) 0-24hrs 6. RBC 24-48hrs 7. RBC 48-72hrs 8. RBC 72hr Total 9. Fresh Frozen Plasma (FFP) 0-24hrs 10. FFP 24-48hrs 11. FFP 48-72hrs 12. FFP 72hr Total 13. Platelet (Plt) 0-24hrs 14. Plt 24-48hrs 15. Plt 48-72hrs 16. Plt 72hr Total 17. Cryoprecipitate (Cryo) 0-24hrs 18. Cryo 24-48hrs 19. Cryo 48-72hrs 20. Cryo 72hr Total 21. Total 24hr RBC 22. Massive Transfusion
Survival and ICU Length of Stay	1. Duration of ICU Stay (days) 2. Death Date 3. Alive by 30 days 4. Alive by 90 days 5. Alive by 12 months 6. ICU Length of Stay (LOS) 7. Hospital LOS

A “Transfusion” variable was created to display patients who received blood transfusion(s). Those who had a total 24 hours red blood cell (RBC) greater than 0 were assigned TRUE and those with less than 0 total 24 hours RBC were assigned FALSE.

## 2.1. Exploratory Data Analysis

The descriptive statistics were calculated. The continuous variables were summarized to display the distribution, mean and standard deviations. The categorical variables were displayed as frequencies and percentages. Graphs were generated using the ggplot2 package to view the distribution and relationships.

### 2.1.1. *Transfusion Patients and Massive Transfusion Patients*

The data was further subsetted to explore individuals who received transfusions or massive transfusions. The descriptive statistics were calculated for this data set. The continuous variables were summarized to display the distribution, mean and standard deviations. The categorical variables were displayed as frequencies and percentages.

Graphs were generated using the ggplot2 package to view the distribution and relationships.

### **2.1.2. Collinearity of Variables - Correlation Plots**

The relationship between predictors was explored by creating a correlation plot. The predictors included variables in the Patient Demographic Data, Underlying Respiratory Diagnosis and Intraoperative Descriptions, and Pre-op Bloodwork categories, as shown in **Table 1**. The chosen categorical variables were converted to numeric values. A correlation matrix between the predictors was created using Pearson's correlation coefficient. The matrix was displayed as a heat map. The colour intensity indicated the strength of the correlation between the variables. Positive correlations were displayed in blue and negative correlations in red. Using a threshold for a high correlation of 0.7, variables paired with absolute correlation above the threshold were isolated.

### **2.1.3. Data Cleaning**

The data was assessed for the presence of blank spaces and missing values. The percentage of missing values across variables was quantified. Variables with >30% missingness were removed. Furthermore, repetitive variables that displayed the same information were removed such as duration of ICU stay days (vs. ICU LOS). For the preoperative blood work variables, any NAs were replaced with 0. Furthermore, the type of extracorporeal life support including ECMO, CPB or none was combined into one variable of “ECLS Type” and the individual variables were removed. Any categorical variables were factorized.

### **2.1.4. Handling of Missing Data**

The “LAS score” variable was imputed with predictive mean matching (PMM) using the “mice” package to minimize bias. PMM creates a regression model using the observed data to predict the missing values. It finds similar observed values closest to the predicted value. The final imputed value is randomly selected from the closest observed values. Predictors were chosen for each variable based on literature to ensure accurate imputation. Predictors of LAS scores include age, gender, BMI, diagnosis (including COPD, alpha1-antitrypsin deficiency, cystic fibrosis, idiopathic pulmonary hypertension, interstitial lung disease and other pulmonary diseases), and type of lung transplant. A single imputation was performed with the algorithm running for 5 iterations to refine the estimates. The diagnostic plots can be seen in **Supplementary Figure 1**.

### **2.1.5. Analysis - Objective One**

#### **2.1.5.1. Model Comparison**

The performance of predictive models, including Least Absolute Shrinkage and Selection Operator (LASSO) and Classification and Regression Tree (CART), was

compared to predict the need for transfusion. “Transfusion” (TRUE/FALSE) was used as the primary outcome variable. Predictors used for the models include the demographics, underlying diagnosis and intraoperative descriptions, pre-op bloodwork (INR) and life support types variables. A subset of the data was created with the “Transfusion” variable and the predictors for further analysis.

The best-performing model was selected based on the Area Under the Curve (AUC) score. The AUC score is a measure of discrimination.

#### **2.1.5.2. LASSO Classification Model Comparison**

The data was split 80/20 between training and testing, repeated for 5 iterations. LASSO classification applied an L1 penalty to shrink the coefficients to 0 for feature selection. The LASSO path was created by training the model on the training set using a sequence of lambda values which alters the penalty strength. A 10-fold cross-validation was performed to select the optimal lambda value that maximizes the AUC score. The optimal lambda was used to train the model and the features that remained when using the optimal lambda were noted. The model was tested on the testing data set to obtain the predicted probabilities. The predicted probabilities were evaluated using the Receiver Operating Characteristic (ROC) curves and the AUC scores were noted. The AUC score from each iteration was averaged.

#### **2.1.5.3. CART**

The data was split 80/20 between training and testing, repeated for 5 iterations. A full tree was created using the training data set. 10-fold cross-validation was used to prune the tree to minimize deviance while preventing overfitting the model. The variables that were used in the final model (full tree and pruned tree) were noted. The full tree and pruned tree models were tested on the testing data set to obtain the predicted probabilities. The predicted probabilities from the full and pruned tree were evaluated using the Receiver Operating Characteristic (ROC) curves and the AUC scores were noted for the full and pruned tree. The AUC score from each iteration was averaged.

The average AUC was calculated to determine the overall performance of each model to be compared. The LASSO classification model was chosen for further analysis.

#### **2.1.5.4. LASSO Classification**

The LASSO model was used to predict the need for transfusions with “Transfusion” (TRUE/FALSE) as the primary outcome variable. The entire dataset was used to train a LASSO classification model to identify predictors associated with transfusion needs. The LASSO path was created by training the entire data set using a sequence of lambda values which alters the penalty strength. A 10-fold cross-validation was

performed to select the optimal lambda value that maximizes the AUC score. The optimal lambda was used to train the model using the complete dataset and the features that remained when using the optimal lambda were noted.

#### **2.1.5.5. LASSO Regression**

The LASSO regression model was used to predict the amount of transfusion needed, with “Total 24-hour RBC” as the primary outcome variable. Predictors used for the models include demographics, underlying diagnosis and intraoperative descriptions, pre-op bloodwork (INR), and life support type variables. A subset of the data was created with the “Total 24 hours RBC” variable and the predictors for further analysis.

The entire dataset was used to train a LASSO regression model to identify predictions associated with the transfusion amount. The LASSO path was created by training the entire data set using a sequence of lambda values which alters the penalty strength. A 10-fold cross-validation was performed to select the optimal lambda value that minimizes the mean squared error (MSE). The optimal lambda was used to train the model using the complete dataset and the features that remained when using the optimal lambda were noted.

#### **2.1.6. Analysis - Objective Two**

The patient outcomes selected for analysis were mortality and ICU length of stay.

##### **2.1.6.1. Survival Analysis**

The survival data was collected longitudinally, with checkpoints at the 30-day, 90-day, and 12-month mark. The Kaplan-Meier curves were used to analyze survival while censoring for patients who did not experience the event or were lost to follow-up. The censoring time-point selected was the last recorded death date which was 2020-01-22. The survival time of the deceased patients is the number of days between their surgery date and their death date. For the censored patients which include patients who did not experience the event or lost to follow-up, the survival time is the number of days between their surgery and the last follow-up date. A new variable called “Status” was created to indicate whether the patient died (1) or was censored (0). A Kaplan-Meier survival curve was created with the survival time and status variables, with the model based on the transfusion variable to compare the survival probability of patients who received a transfusion against those who did not. A log-rank test was conducted to compare the differences between the two survival groups (transfusion and no transfusion). The result will show if there is a statistically significant difference in survival between patients with and without transfusion.

##### **2.1.6.2. Cox Proportional Hazards Model**

The Cox Proportional Hazard model was created to evaluate the effect of transfusion and other confounding variables on mortality. A list of predictors used for the model includes transfusion and the significant variables from the LASSO classification model. Censoring was also done with the same assumptions stated for the survival analysis (Kaplan-Meier curves). The proportional hazards assumptions of a Cox Regression were tested. The hazard ratio, 95% confidence intervals and p-value were noted.

#### **2.1.6.3. Wilcoxon Test for ICU Length of Stay**

A Wilcoxon Rank-Sum test was used to compare the ICU length of stay between patients who received a transfusion and those who did not. The distribution of ICU length of stay data was heavily right-skewed. Therefore, a non-parametric test was used to evaluate statistical significance.

### **3. Results**

#### **3.1. Exploratory Data Analysis**

The raw dataset contained 192 patient observations of 60 variables. The mean age of patients was 56 years (std = 14.80), with 54% (n = 104) of patients being male and 46% (n = 88) patients being female. Most patients received a bilateral or double lung transplant (n = 157 or 81%) and roughly 9% of patients received a single left (n = 18) or single right (n = 17) transplant. Most patients were also undergoing their first lung transplant (95% or n= 181) and only 10 patients were re-doing lung transplant (5%). The most common lung diseases in the patients were Interstitial lung disease (n= 95 or 48%) and COPD (n= 58 or 30%), followed by cystic fibrosis (n= 30 or 16%), alpha1-Antitrypsin Deficiency (n= 9, 5%), idiopathic pulmonary hypertension (n= 6, 3%) and 8% of patients had other pulmonary diseases (n = 15). More than half of all patients received transfusions (n= 114, 59%). **Supplementary Table 1 and Table 2** provide a comprehensive overview of all numerical and categorical variables included in the dataset.

#### **3.1.1. Transfusion Patients and Massive Transfusion Patients**

The dataset was further subsetted to only include patients who received transfusions. The mean age of patients who received transfusions was 54 years (std = 14.90), with 44% (n = 50) of patients being male and 56% (n = 64) patients being female. Most transfusion patients also received a bilateral or double lung transplant (n = 105 or 92%) and roughly 4% of patients received a single left (n = 4) or single right (n = 5) transplant. Most transfusion patients were undergoing their first lung transplant (91% or n= 104) and all 10 patients who were re-doing lung transplant were receiving transfusions (9%). The most common lung diseases in the transfusion patients were Interstitial lung disease (n= 57 or 50%) followed by cystic fibrosis (n= 25 or 22%) and COPD (n= 23,

20%), with alpha1-Antitrypsin Deficiency (n= 5) and idiopathic pulmonary hypertension (n= 4) making up about 4% of patients. 11% of patients had other pulmonary diseases (n = 13). 9 patients also received transfusions of more than 10 RBC units, referred to as massive transfusions (8%). **Supplementary Table 3 and Table 4** provide a comprehensive overview of all numerical and categorical variables for patients who received transfusion.

### **3.1.2. Collinearity of Variables**

The correlation matrix was plotted to assess collinearity between the following groups of predictors: Patient Demographic Data, Underlying Respiratory Diagnosis and Intraoperative Descriptions, and Pre-op Bloodwork categories (**Supplementary Figure 2**.) An absolute threshold of 0.7 was used to determine which predictors exhibited high correlation with other predictors. **Table 2** outlines the correlated variables, and the associated strength of correlation.

**Table 2:** Variables with High Collinearity (correlation > |0.7|)

Variable1	Variable2	Correlation
BMI	Weight	0.8330637
HOSPITAL_LOS	ICU_LOS	0.7799568
Pre_Hct	Pre_Hb	0.9581300
Pre_Fibrinogen	Pre_Platelets	-0.9283573
Pre_INR	Pre_PT	0.9975031
Pre_Fibrinogen	Pre_INR	0.8003747
Pre_Fibrinogen	Pre_PTT	0.9163166
Pre_Creatinine	Pre_Fibrinogen	-0.9086222

Based on the identified variables in **Table 2**, the following variables were removed from downstream analysis due to high collinearity: Weight, hospital LOS, and pre-op bloodwork (Hct, Platelets, Fibrinogen and Creatinine) . BMI, ICU LOS and Preoperative international normalized ratio (Pre\_INR) were kept for downstream analysis.

## **3.2. Analysis**

### **3.2.1. Model Comparison**

The performance of predictive models, including Least Absolute Shrinkage and Selection Operator (LASSO) and Classification and Regression Tree (CART), was compared to predict the need for transfusion. “Transfusion” (TRUE/FALSE) was used as the primary outcome variable. For each iteration, the random nature of sampling to create the training and testing data sets resulted in different features present in the final model used for prediction after each iteration.

### **3.2.2. LASSO Classification**

A summary of the LASSO path, cross-validation plots for the LASSO classifier and AUC-ROC curve for each iteration can be seen in **Supplementary Figure 3**. A list of the predictors extracted from the LASSO classification model for each interaction is provided in **Supplementary Table 5**. The predictors that appeared in 4 or more models (iterations) were: BMI, underlying respiratory diagnosis (including COPD, cystic fibrosis, idiopathic pulmonary hypertension, interstitial lung disease, alpha1 antitrypsin deficiency and other pulmonary disease), height, LAS score, preoperative ECLS, redo lung transplant, the type of lung transplant (single left or right lung), age, ex-vivo lung perfusion, and male gender. To evaluate the predictive performance of the LASSO classification model, the AUC curve was calculated from the ROC curves for each iteration and summarized in **Supplementary Table 6**. The average AUC across all 5 iterations was 0.801, indicating the model has a good ability to distinguish between the two classes of transfusion.

### **3.2.3. CART**

A summary of the CART full tree, CART pruned tree and its respective AUC-ROC curves for each iteration can be seen in **Supplementary Figure 4**. A list of the predictors extracted from the CART full tree model for each interaction is provided in **Supplementary Table 7**. The predictors that appeared in 4 or more models (iterations) of the CART full tree model were: type of lung transplant, height, age, BMI, LAS score, and pre-op bloodwork (INR). To evaluate the predictive performance of the CART full tree model, the AUC curve was calculated from the ROC curves for each iteration and summarized in **Supplementary Table 8**. The average AUC across all 5 iterations was 0.667, indicating the model has a poor to moderate ability to distinguish between the two classes of transfusion. A list of the predictors extracted from the CART pruned tree model for each interaction is provided in **Supplementary Table 9**. The pruned trees contained one predictor for each iteration. The predictor varied for each interaction, ranging from height, type of lung transplant, LAS score and BMI. The height variable appeared in 2 iterations. To evaluate the predictive performance of the CART pruned tree model, the AUC curve was calculated from the ROC curves for each iteration and summarized in **Supplementary Table 10**. The average AUC across all 5 iterations was

0.602, indicating the model has a poor ability to distinguish between the two classes of transfusion.

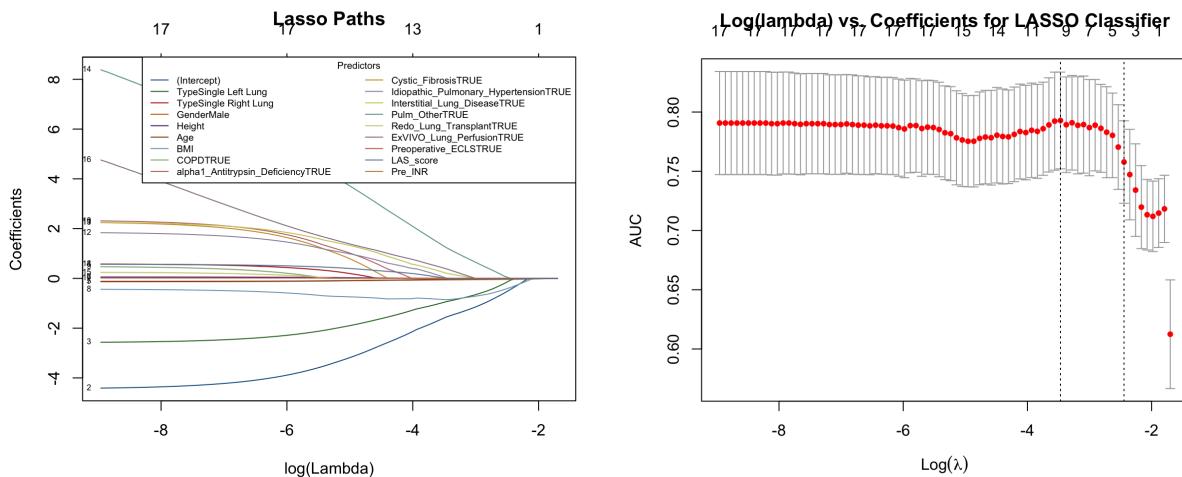
The comparison of the three models with the average AUC can be summarized in **Table 3**. The LASSO classification models outperformed the CART full and pruned tree models with an overall higher AUC score. Given the LASSO Classification model had the highest score, it was used for further analysis.

**Table 3:** Comparison of Average AUC Across the Classification Models

Comparison of Average AUC Across Models	
Model	Average AUC
LASSO Classification	0.801
CART Tree	0.667
CART Pruned Tree	0.602

### 3.2.4. LASSO Classification

The entire dataset was used to train a LASSO classification model to identify predictors associated with transfusion needs. The LASSO path displaying how Lasso regularization affects the coefficients of predictors in a model as  $\log(\lambda)$  increases can be seen in **Figure 1**.



**Figure 1:** Summary of the LASSO path, cross-validation plots for the LASSO Classifier

By around  $\log(\lambda)$  equals -3, most of the features' coefficients have shrunk to zero, which means the features no longer contribute to the model's predictions. The remaining predictors have non-zero coefficients even at higher levels of regularization. To obtain the specific predictors, 5-fold cross-validation is performed to obtain the optimal lambda value that maximizes the AUC score. The optimal lambda was 0.031 and the corresponding AUC was 0.793. The chosen predictors with non-zero coefficients at the optimal value of lambda can be summarized in **Table 4**.

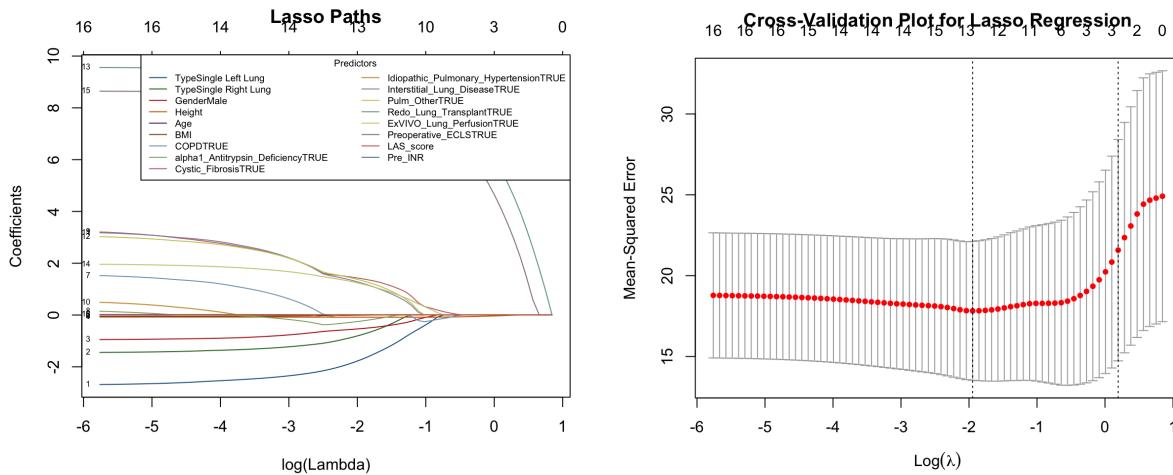
**Table 4:** LASSO Classifiers - Final Selected Predictors

Lasso Classifiers - Final Selected Predictors with Coefficients	
Predictors	Coefficient
TypeSingle Left Lung	-1.5514
TypeSingle Right Lung	-0.9365
Height	-0.0602
BMI	-0.0305
COPDTRUE	-0.8578
Interstitial_Lung_DiseaseTRUE	0.0145
Pulm_OtherTRUE	0.2077
Redo_Lung_TransplantTRUE	1.2383
Preoperative_ECLSTRUE	0.3870
LAS_score	0.0088

These predictors including the type of lung transplant (single left or right lung), height, BMI, underlying respiratory diagnosis (including COPD, interstitial lung disease, and other pulmonary diseases), redo lung transplant, preoperative ECLS and LAS score are identified as the most important for predicting the need for transfusion. These predictors consistently appeared in 4 or more iterations when we trained and tested the LASSO classification model five times for comparison against the CART model.

### 3.2.5. LASSO Regression

The entire dataset was used to train a LASSO regression model to identify predictors associated with the amount of transfusion (total 24 hours RBC). The LASSO path displaying how Lasso regularization affects the coefficients of predictors in a model as  $\log(\lambda)$  increases can be seen in **Figure 2**.



**Figure 2:** Summary of the LASSO path, cross-validation plots for the LASSO Regression

By around  $\log(\lambda)$  equals -1, the coefficient of most of the features has shrunk to zero, which means the features are no longer contributing to the model's predictions. The remaining predictors have non-zero coefficients even at higher levels of regularization. To obtain the specific predictors, 10-fold cross-validation is performed to obtain the optimal lambda value that minimizes the MSE value. The optimal lambda was 0.143. The chosen predictors with non-zero coefficients at the optimal value of lambda can be summarized in **Table 5**.

**Table 5:** LASSO Regression - Final Selected Predictors

Lasso Regression - Final Selected Predictors with Coefficients	
Predictors	Coefficient
TypeSingle Left Lung	-1.7238
TypeSingle Right Lung	-0.7672
GenderMale	-0.5250
Height	-0.0854
BMI	-0.0110
alpha1_Antitrypsin_DeficiencyTRUE	-0.2346
Cystic_FibrosisTRUE	1.4076
Interstitial_Lung_DiseaseTRUE	1.2428
Pulm_OtherTRUE	1.3401
Redo_Lung_TransplantTRUE	9.2354
ExVIVO_Lung_PerfusionTRUE	1.2102
Preoperative_ECLSTRUE	8.2592
LAS_score	0.0116

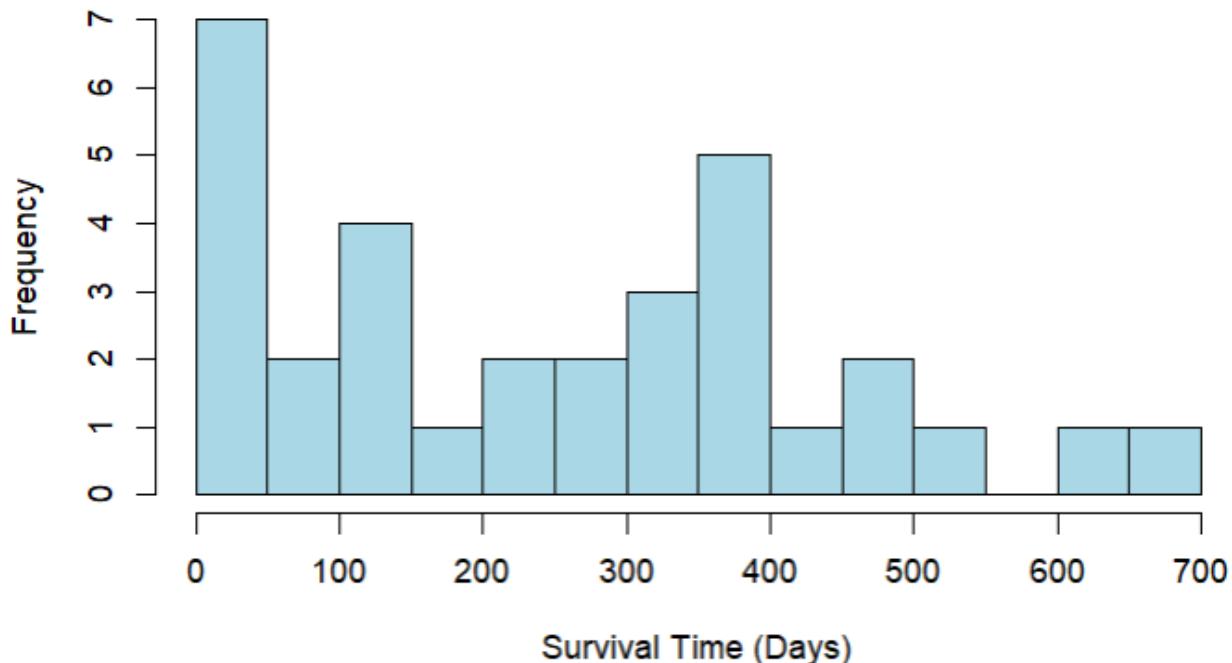
These predictors include type of lung transplant (single left or right lung), male gender, height, BMI, underlying respiratory diagnosis (including cystic fibrosis, interstitial lung

disease, alpha1 antitrypsin deficiency and other pulmonary disease), redo lung transplant, ex vivo lung perfusion, preoperative ECLS and LAS score.

### **3.2.6. Survival Analysis**

The patient outcomes selected for this study are mortality and ICU length of stay. The method of performing the analysis for each outcome differed due to data availability. Survival data was collected longitudinally, with checkpoints at the 30-day, 90-day, and 12-month marks. Thus, Kaplan-Meier curves were utilized in order to analyze survival, while censoring for patients who did not experience the event. The censoring time point selected was 2020-01-22, which is the last recorded death date. It was observed in the data that although some patients survived at the aforementioned study check points, there was a date of death that was recorded following that period. As such, a reasonable assumption was made that the study followed up with patients in the study up until the last recorded date of death, which is presumably the last follow up date. ICU and hospital length of stay data was recorded in days. As such, a Wilcoxon Rank-Sum test was conducted in order to analyze whether there was a statistically significant difference in lengths of stay among subgroups.

**Histogram of Survival Times (Days) for Deceased Patients**



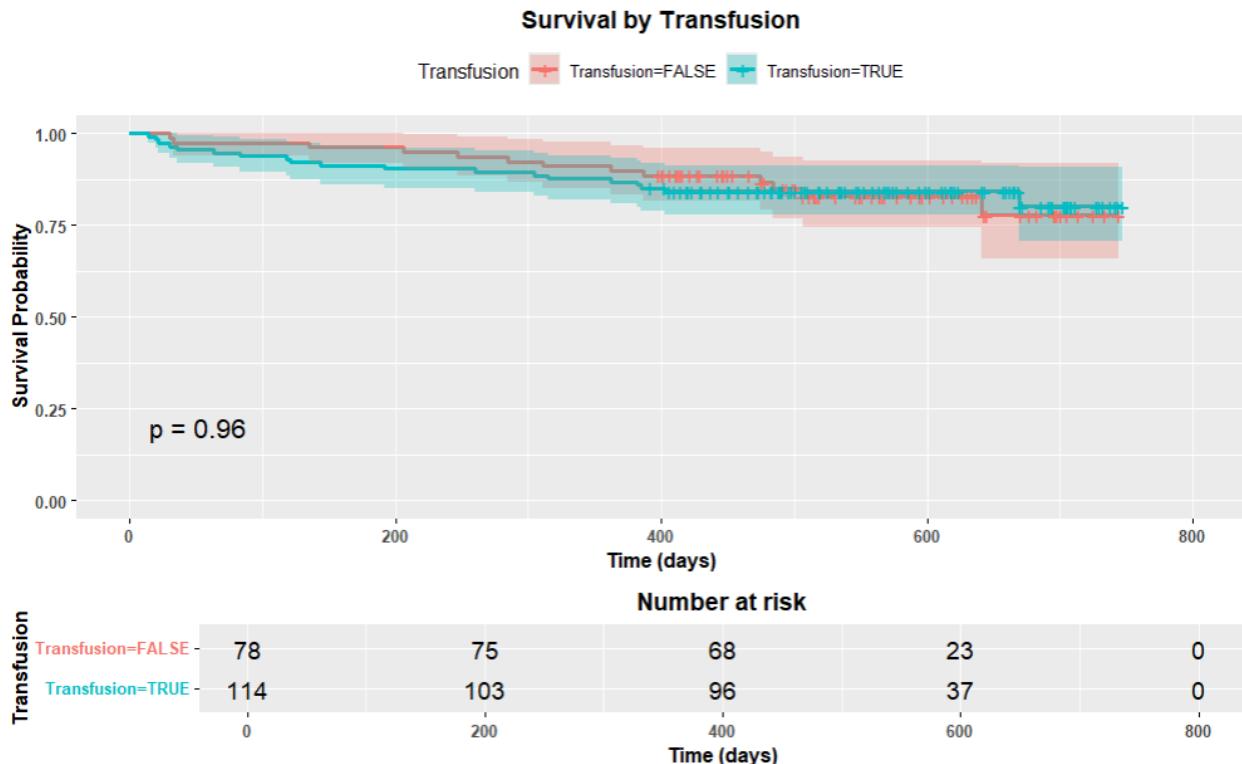
**Figure 3:** Histogram plot of the survival time in days for patients who experienced the event (death)

Overall, there were 32 patients who died in the duration of the study. **Figure 3** shows a histogram plot of the survival times of those patients. The median time of death was 253.5 days, suggesting that most patients died within the first year of their transplant. The data appears to be slightly right skewed, with a mean death date of 250.8 days. The last recorded death occurred on day 669.

### 3.2.6.1. Kaplan Meier Curve

Kaplan Meier curves were used to analyze survival probabilities at different time points while censoring for those who did not experience the event. The following were the subgroups on which the analysis was conducted:

- a. Transfusion: patients who received a transfusion vs patients who did not.



**Figure 4:** Kaplan Meier curve showing the survival probability at different time points among those who received a transfusion vs those who did not, in addition to a risk table.

The Kaplan Meier curve shown in **Figure 4** breaks down the survival probability of patients at different time points based on whether or not they had a transfusion. As expected, the survival probability starts at 100% for both groups at 0 days. This is reflected in the risk table, where all 78 patients who did not receive a transfusion, and 114 who did are at risk. As the number of days increases, the survival probability of both groups slowly decreases parallel to each other, with the 95% confidence intervals overlapping. This trend is also observed in the risk table, where the number of patients

at risk declines steadily between the 200-day interval. This decline in the number of patients still at risk of experiencing the event is due to some patients already dying. For example, 103 patients are at risk at the 200 day mark in the transfusion group, compared to 96 in the 400 day mark. This suggests that 7 patients died in that time or were lost to follow up.

The null hypothesis of this analysis is that there is no difference in survival distributions between those who received a transfusion and those who did not. Since a p-value of 0.96 was obtained, there is no statistical evidence to reject the null, thus we can conclude that there was no true difference in survival outcomes between the two groups. This is further supported by the proximity of the survival curves for the two groups, and the overlapping confidence intervals. Additionally, the effect size obtained for both groups was about 0.001 respectively, indicating a negligible difference in the observed and expected events, thus reinforcing the conclusion that there is no significant difference in survival between both groups.

### **3.2.6.2. Cox Proportional Hazard Model**

A Cox Proportional Hazard model was synthesized in order to determine the effect of transfusion and other confounding variables on mortality. The logical transfusion variable was used as a predictor, in addition to significant variables in the Lasso classification model. Censoring was used under the same assumptions stated for the Kaplan Meier curve.

**Table 6:** Results of the Cox Proportional Hazards Model

<b>Cox Proportional Hazards Model Results</b>				
Hazard Ratios, 95% CI, and P-Values				
Predictor	Hazard Ratio	95% CI	P-Value	
TypeSingle Left Lung	1.94	(0.55, 6.85)	0.302	
TypeSingle Right Lung	1.60	(0.49, 5.2)	0.431	
Height	1.01	(0.97, 1.05)	0.575	
BMI	1.11	(1, 1.23)	0.045	
Interstitial_Lung_DiseaseTRUE	0.74	(0.28, 1.98)	0.551	
COPDTRUE	0.86	(0.32, 2.31)	0.759	
Pulm_OtherTRUE	0.41	(0.05, 3.37)	0.408	
TransfusionTRUE	1.69	(0.7, 4.09)	0.242	
Preoperative_ECLSTRU	1.23	(0.13, 12.03)	0.859	
LAS_score	0.99	(0.94, 1.04)	0.626	
Pre_INR	1.25	(0.69, 2.26)	0.460	

**Table 6** presents the hazard ratios, 95% confidence intervals, and the p-values for the variables in the model. The significance threshold (alpha) was set at 0.05 a priori. The null hypothesis states that the hazard ratio of the predictor is equal to 1. The results of the model suggest that BMI is the only predictor significantly associated with mortality ( $p = 0.045$ ). This means that BMI has a hazard ratio that is statistically different from 1, suggesting that it has a significant effect on mortality. BMI has a hazard ratio of 1.11. This suggests that for every one unit increase in BMI, there is an 11% increase in hazard when all else is held constant (at zero) or at reference level.

Typically, a concordance score of 0.5 indicates that the model predicts risks accurately 50% of the time when compared to the observed data, thus only predicted risks by random chance. In this case, the model yielded a concordance of 0.639, thus indicating that it predicts risk accurately about 64% of the time, which is slightly better than random chance. The p-value obtained for this model is 0.7. This suggests that we fail to reject the null hypothesis that the model does not significantly explain the variability in survival.

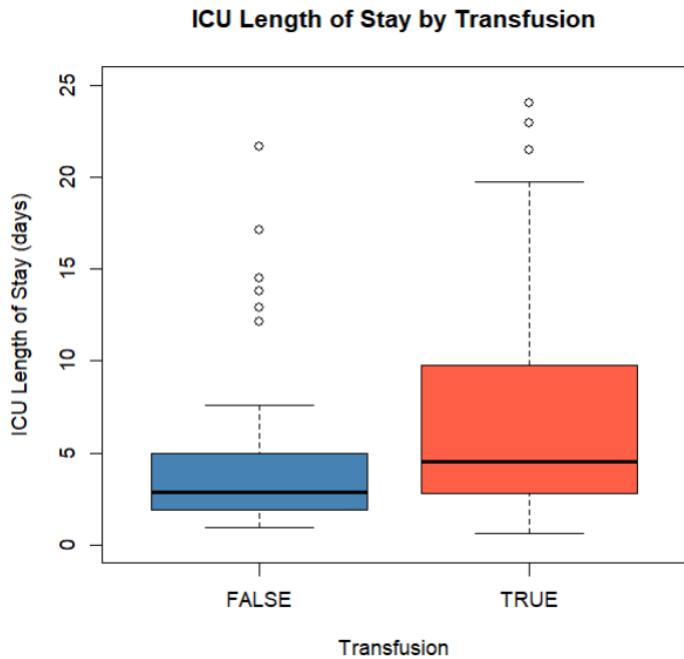
A per variable test was conducted to check if the proportional hazard assumptions were violated. A p-value less than 0.05 would indicate a violation of the proportional hazards assumption for a specific variable. Table 7 shows that none of the variables had a p-value greater than that threshold, thus the proportional hazard assumptions are upheld. Similarly, the Global test had a p-value greater than 0.05, indicating that the overall model did not violate proportional hazard assumptions.

**Table 7:** Results of the proportional hazards assumptions test.

Proportional Hazards Assumption Test Results			
Variable	ChiSquare	Degrees Of Freedom	PValue
Type	0.805	2	0.670
Height	0.768	1	0.380
BMI	0.457	1	0.500
Interstitial_Lung_Disease	2.629	1	0.100
COPD	2.866	1	0.090
Pulm_Other	0.449	1	0.500
Transfusion	2.569	1	0.110
Preoperative_ECLS	1.777	1	0.180
LAS_score	1.406	1	0.240
Pre_INR	1.135	1	0.290
GLOBAL	10.912	11	0.450

### 3.2.6.3. Wilcoxon Rank-Sum Test

A Wilcoxon Rank-Sum test was used to compare the ICU length of stay among those who received transfusions, compared to those who did not. Since the distribution of ICU length of stay data is heavily right skewed as shown in Figure 5, a non-parametric test was needed to evaluate statistical significance. The null hypothesis in this situation is that there is no difference in distribution of ICU length of stay between the two groups. Conversely, the alternative states that there is a difference in the distribution. The p-value obtained from the test is 0.0004, suggesting that there is strong evidence to reject the null. As such, there is statistically significant difference in the distribution of ICU length of stay between the two groups.



**Figure 5:** Box plot showing the distribution of ICU lengths of stay by transfusion status.

#### 4. Discussion

Our analysis findings were compared to the literature, to determine if our findings were consistent with existing research and/or any novel insights.

##### 4.1. Predictors for the need and amount of transfusion

We identified key factors predictive of the need for transfusion, including the type of lung transplant (single lung, left or right), height, BMI, underlying respiratory diagnosis (e.g., COPD, interstitial lung disease, and other pulmonary diseases), redo lung transplant, preoperative extracorporeal life support (ECLS), and LAS score. These findings align partially with the existing literature, which identified greater transfusion requirements for double-lung transplants due to their increased surgical complexity (Oechslin et al., 2018). While redo lung transplants were not explicitly identified as a predictor for transfusion need in the literature, the similarity in mechanics to double-lung transplants (e.g., increased complexity and surgical challenges) justifies their inclusion in our predictive model. The literature also supports our finding of ECLS as a predictor for transfusion as it was highlighted that patients requiring perioperative ECMO support are more likely to receive transfusions (Ang et al., 2009). Finally, although pulmonary hypertension and cystic fibrosis were prominent in the literature as predictors of transfusion need, these were not significant in our analysis potentially due to differences in population characteristics (Klapper et al., 2021; Oechslin et al., 2018).

We also identified factors predicting the amount of transfusion required. These included the type of lung transplant, male gender, height, BMI, underlying respiratory diagnosis (including cystic fibrosis, interstitial lung disease, alpha-1 antitrypsin deficiency, and other pulmonary diseases), redo lung transplant, ex vivo lung perfusion, preoperative ECLS, and LAS score. The literature echoes the association between double-lung transplants and higher transfusion volumes, but studies also had found preoperative ECMO and CPB in predicting large-volume transfusions, which is consistent with our analysis (Ang et al., 2009; Oechslin et al., 2018).

## **4.2. Impact of transfusion and confounding variables on patient outcomes**

### ***4.2.1. Transfusion vs. No Transfusion***

Our analysis revealed no significant difference in survival between patients who received transfusions, or massive transfusions, and those who did not. This is in contrast to the literature, which highlights intraoperative transfusion as a predictor of in-hospital mortality and poorer outcomes (Klapper et al., 2021; Weber et al., 2013; Zalunardo et al., 2011). The discrepancy may stem from differences in study design, population characteristics, or statistical methods.

### ***4.2.2. ICU length of stay for transfusion vs no transfusion***

We found a statistically significant difference in ICU length of stay between patients who received transfusion vs. patients who did not, such that those who received transfusion had longer ICU length of stay than their counterparts. This finding is consistent with the literature where ICU length of stay was significantly greater in patients who received blood transfusion perioperatively (Siddiqui & Shakil, 2024).

### ***4.2.3. BMI***

BMI was identified as a significant predictor of mortality in our analysis. The literature does not explicitly address BMI as a factor affecting post-transplant patient outcomes, and thus this is a factor unique to our patient population.

## **4.3. Limitations**

The main limitations from this study ultimately stem from the assumptions and methods used to analyze the data. The first limitation is the lack of statistical inference from the Lasso classification model. Lasso does not provide coefficients with p-values, preventing deductions about statistical significance of the predictors. Results from the

Lasso model also rely heavily on the penalization parameter. Although the statistical optimal lambda value is used, the results can vary depending on the degree of penalization selected, thus limiting the ability to draw concrete statistical inference. Furthermore, the Kaplan Meier curve is considered a univariate analysis, suggesting lack of control for confounding variables or interaction effects. Lastly, a major limitation is the censoring assumption of the study. The first assumption is that the last recorded death date is the last follow up date of the study. The second assumption is that the research team followed up with all patients up until the follow up date. These assumptions were made due to the lack of detail on how the data was collected. A deviation from these assumptions significantly impacts the interpretability of findings past the 365 day mark.

## 5. Conclusion

Overall, our analysis identified key factors predictive of the need and amount for transfusion, and the impact of transfusion on patient outcomes. Many of our findings overlap with literature in areas such as the importance of transplant type and preoperative ECLS, however they diverge in others, particularly regarding survival outcomes post-transfusion highlighting the need for more research in this area.

## References

- Ang, A. L., Teo, D., Lim, C. H., Leou, K. K., Tien, S. L., & Koh, M. B. C. (2009). Blood transfusion requirements and independent predictors of increased transfusion requirements among adult patients on extracorporeal membrane oxygenation – a single centre experience. *Vox Sanguinis*, 96(1), 34–43.

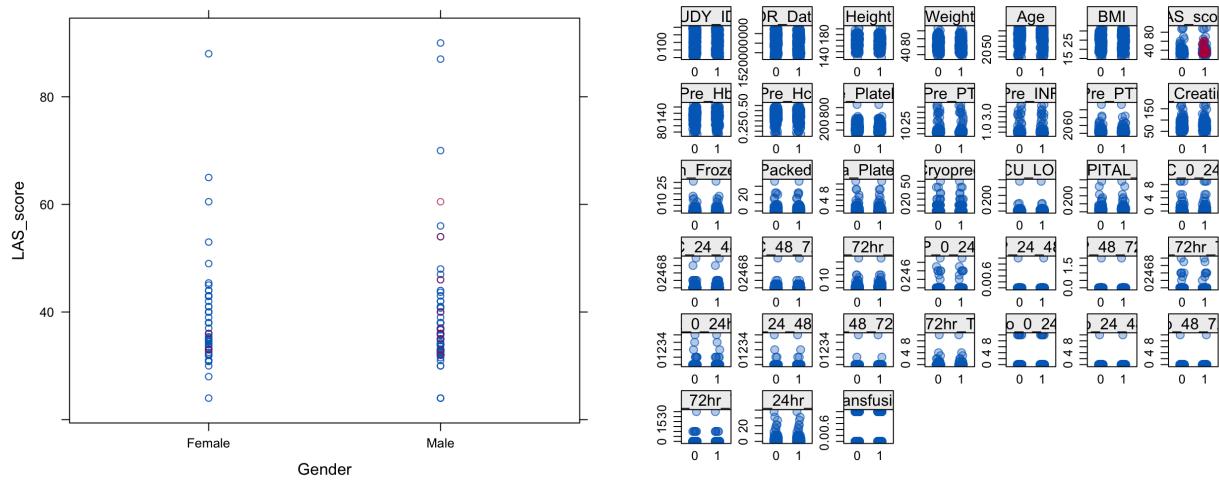
<https://doi.org/10.1111/j.1423-0410.2008.01110.x>

- Canadian Pulmonary Fibrosis Foundation. (n.d.). *Lung Transplantation*. Canadian Pulmonary Fibrosis Foundation. Retrieved November 30, 2024, from <https://cpff.ca/understanding-pf/treatment-and-care/lung-transplantation/>
- Klapper, J. A., Hicks, A. C., Ledbetter, L., Poisson, J., Hartwig, M. G., Hashmi, N., Welsby, I., & Bottiger, B. A. (2021). Blood product transfusion and lung transplant outcomes: A systematic review. *Clinical Transplantation*, 35(10), e14404. <https://doi.org/10.1111/ctr.14404>
- Oechslin, P., Zalunardo, M. P., Inci, I., Schlaepfer, M., & Grande, B. (2018). Established and potential predictors of blood loss during lung transplant surgery. *Journal of Thoracic Disease*, 10(6). <https://doi.org/10.21037/jtd.2018.05.165>
- Provincial Lung Working Group Transplant Steering Committee. (2024, June). *Ontario's Adult Listing Criteria for Lung Transplantation*. [https://www.giftolife.on.ca/resources/pdf/transplant/ON\\_Adult\\_Listing\\_Criteria\\_Lung\\_Transplantation\\_v4.pdf](https://www.giftolife.on.ca/resources/pdf/transplant/ON_Adult_Listing_Criteria_Lung_Transplantation_v4.pdf)
- Ramos, K. J., & Lease, E. D. (2016). Recipient Selection for Lung Transplantation. *American Journal of Respiratory and Critical Care Medicine*, 193(9), P15-16. <https://doi.org/10.1164/rccm.1939P15>
- Said, S. A., Okamoto, T., Nowacki, A. S., Niikawa, H., Ayyat, K. S., Sakanoue, I., Yun, J., & McCurry, K. R. (2021). The Effect of Blood Transfusion in Lung Donors on Recipient Survival. *The Annals of Thoracic Surgery*, 112(4), 1109–1117. <https://doi.org/10.1016/j.athoracsur.2020.10.027>

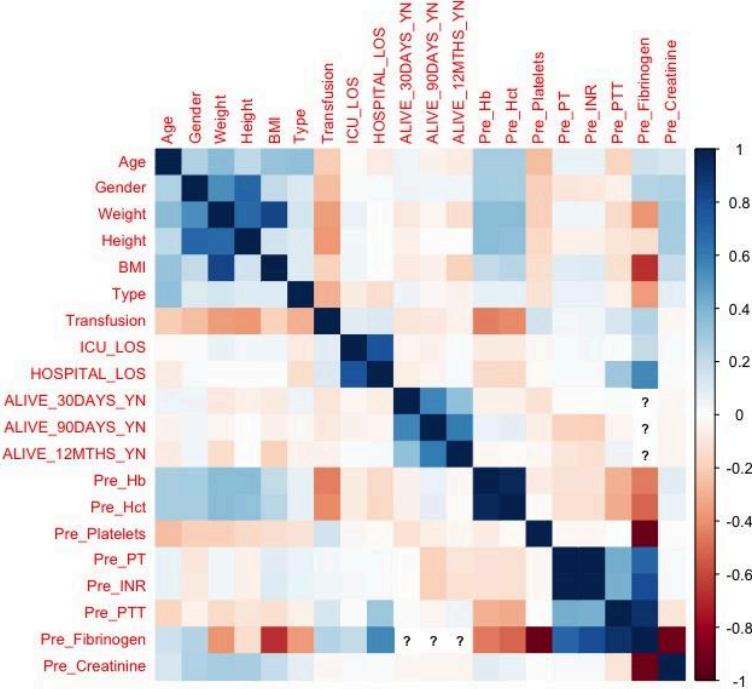
- Siddiqui, A. S., & Shakil, J. (2024). Impact of Blood Products Transfusion on Patients in the Immediate Post-Lung Transplant Period: A Cohort Study. *Annals of Transplantation*, 29, e943652-1-e943652-11. <https://doi.org/10.12659/AOT.943652>
- Toy, P., Gajic, O., Bacchetti, P., Looney, M. R., Gropper, M. A., Hubmayr, R., Lowell, C. A., Norris, P. J., Murphy, E. L., Weiskopf, R. B., Wilson, G., Koenigsberg, M., Lee, D., Schuller, R., Wu, P., Grimes, B., Gandhi, M. J., Winters, J. L., Mair, D., ... Matthay, M. A. (2012). Transfusion-related acute lung injury: incidence and risk factors. *Blood*, 119(7), 1757–1767. <https://doi.org/10.1182/blood-2011-08-370932>
- Weber, D., Cottini, S. R., Locher, P., Wenger, U., Stehberger, P. A., Fasshauer, M., Schuepbach, R. A., & Béchir, M. (2013). Association of intraoperative transfusion of blood products with mortality in lung transplant recipients. *Perioperative Medicine*, 2(1), 20. <https://doi.org/10.1186/2047-0525-2-20>
- Zalunardo, M. P., Thalmann, C., Seifert, B., D'Cunja, J., Weder, W., Boehler, A., & Spahn, D. R. (2011). Impact of preoperative right-ventricular function and platelet transfusion on outcome after lung transplantation☆. *European Journal of Cardio-Thoracic Surgery*, 39(4), 538–542. <https://doi.org/10.1016/j.ejcts.2010.07.040>

## Appendix A. Supplementary Figures

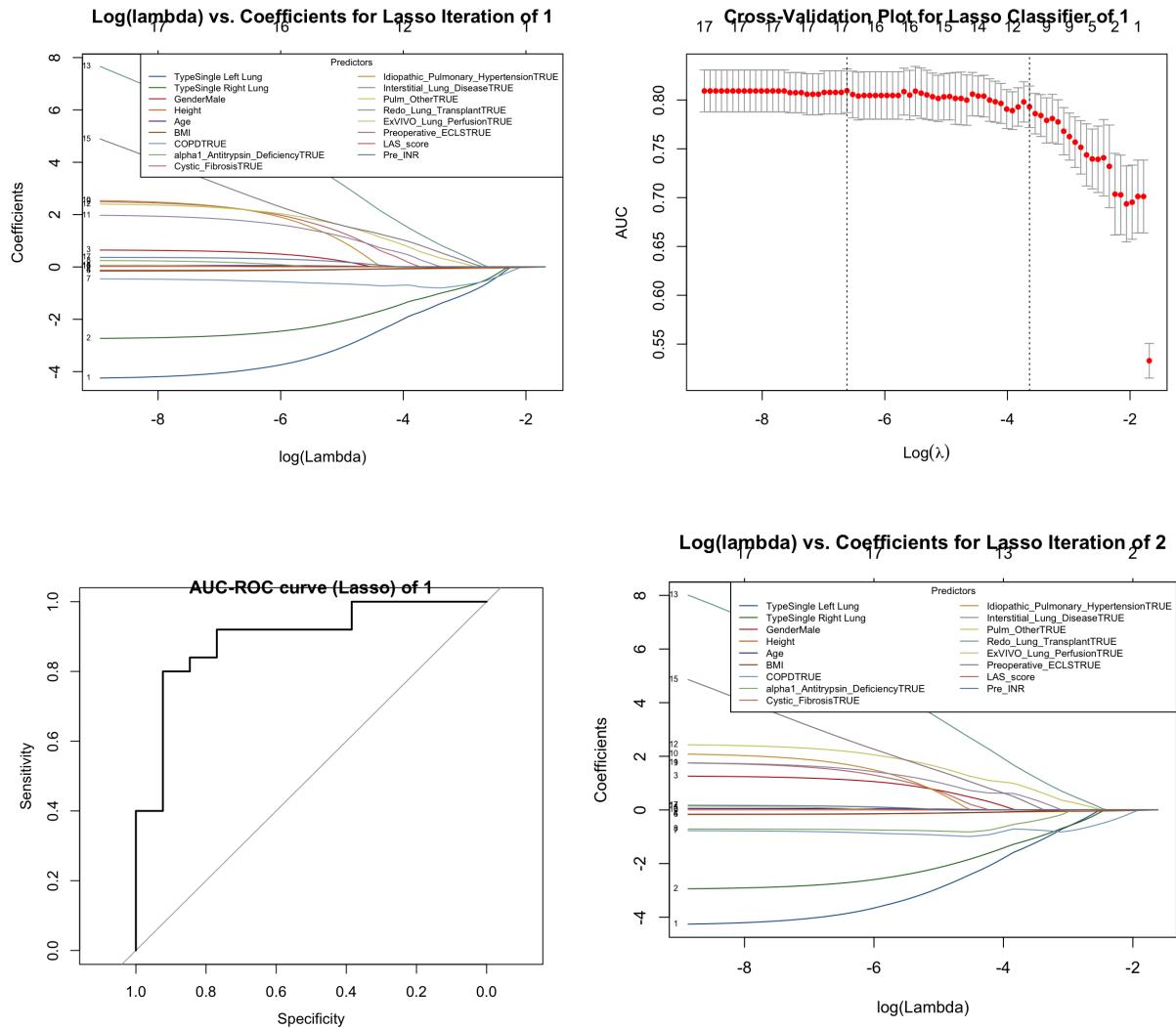
**Figure 1: Diagnostic plots for imputation**

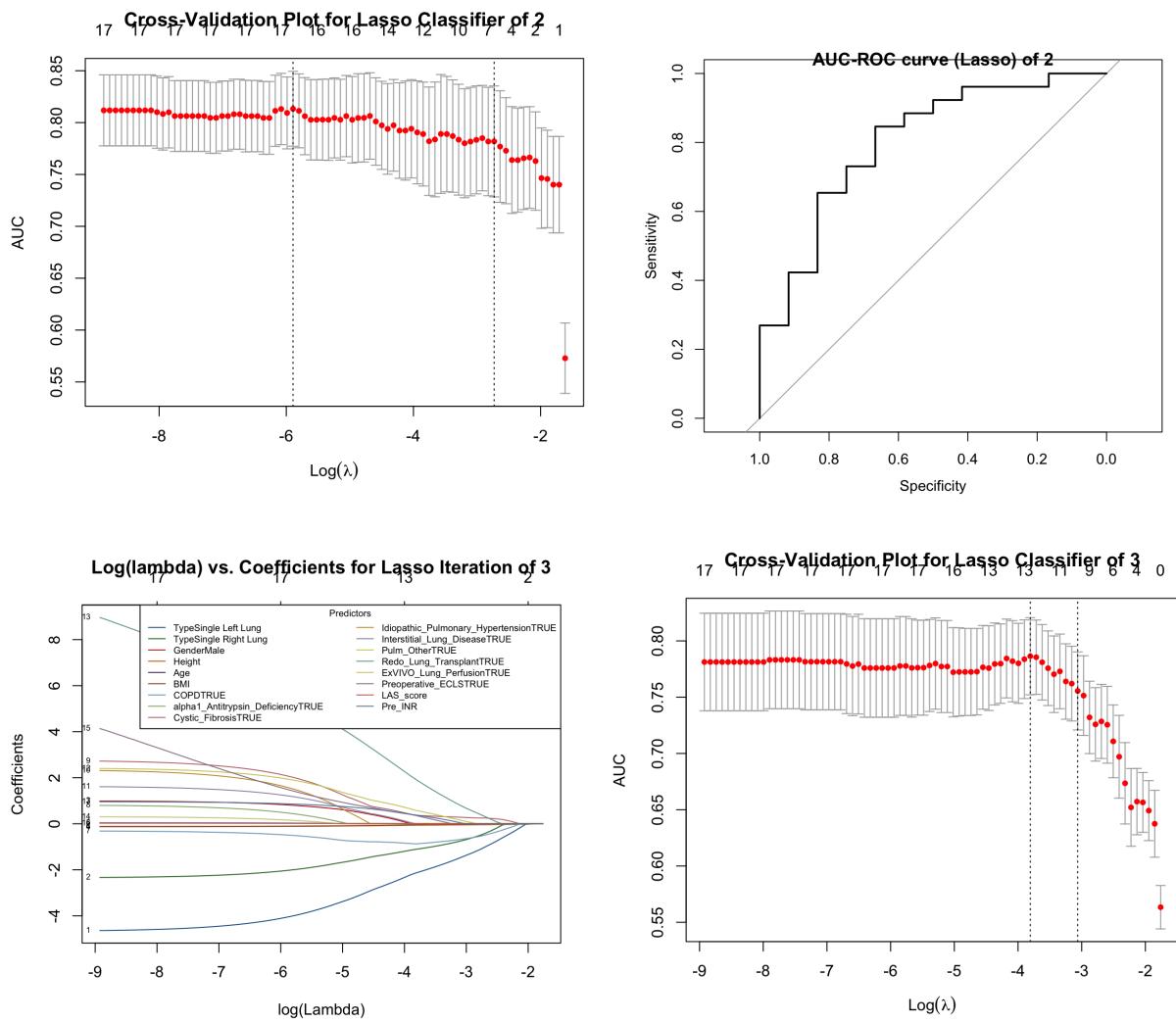


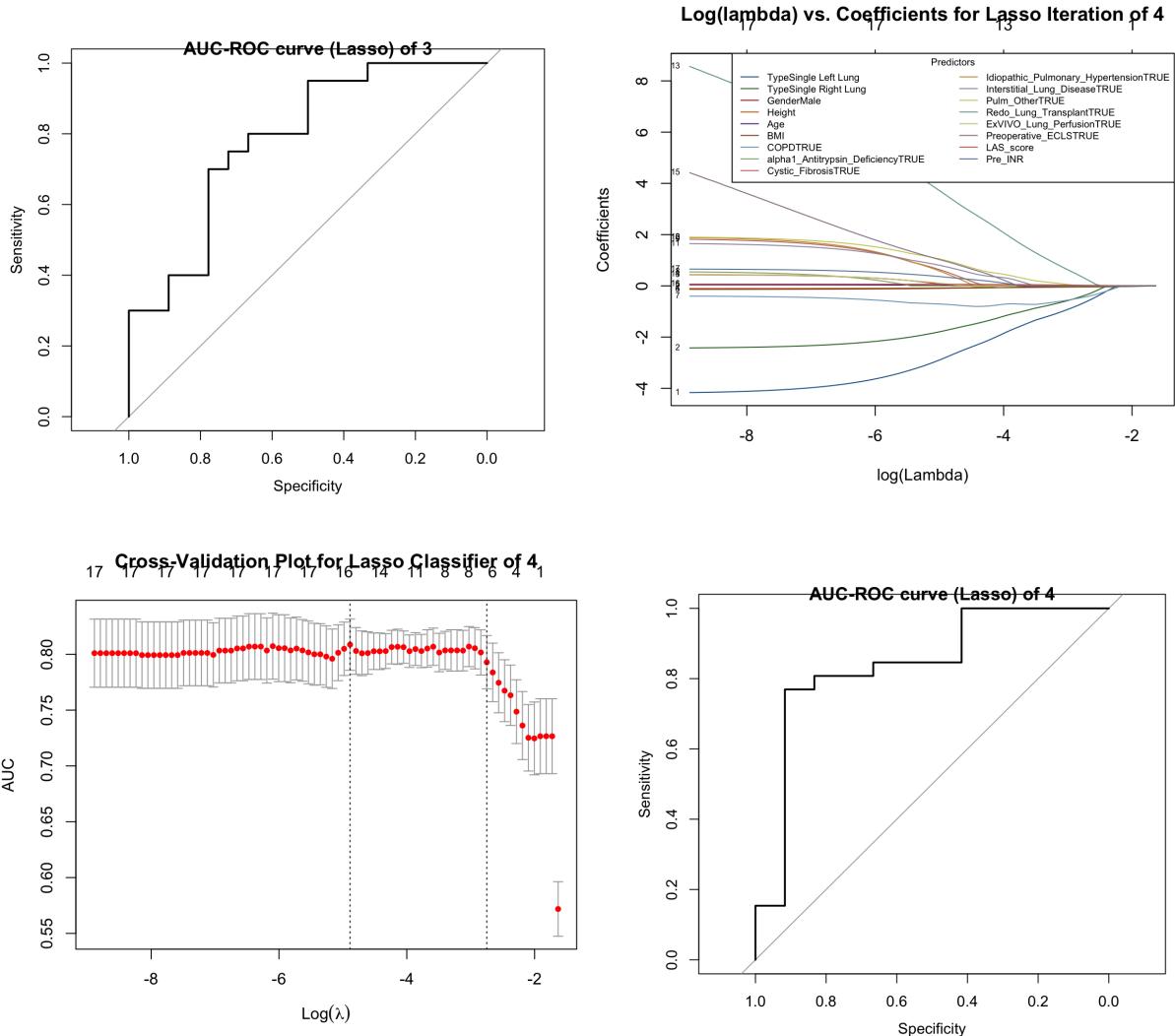
**Figure 2: Correlation Matrix of Predictors**

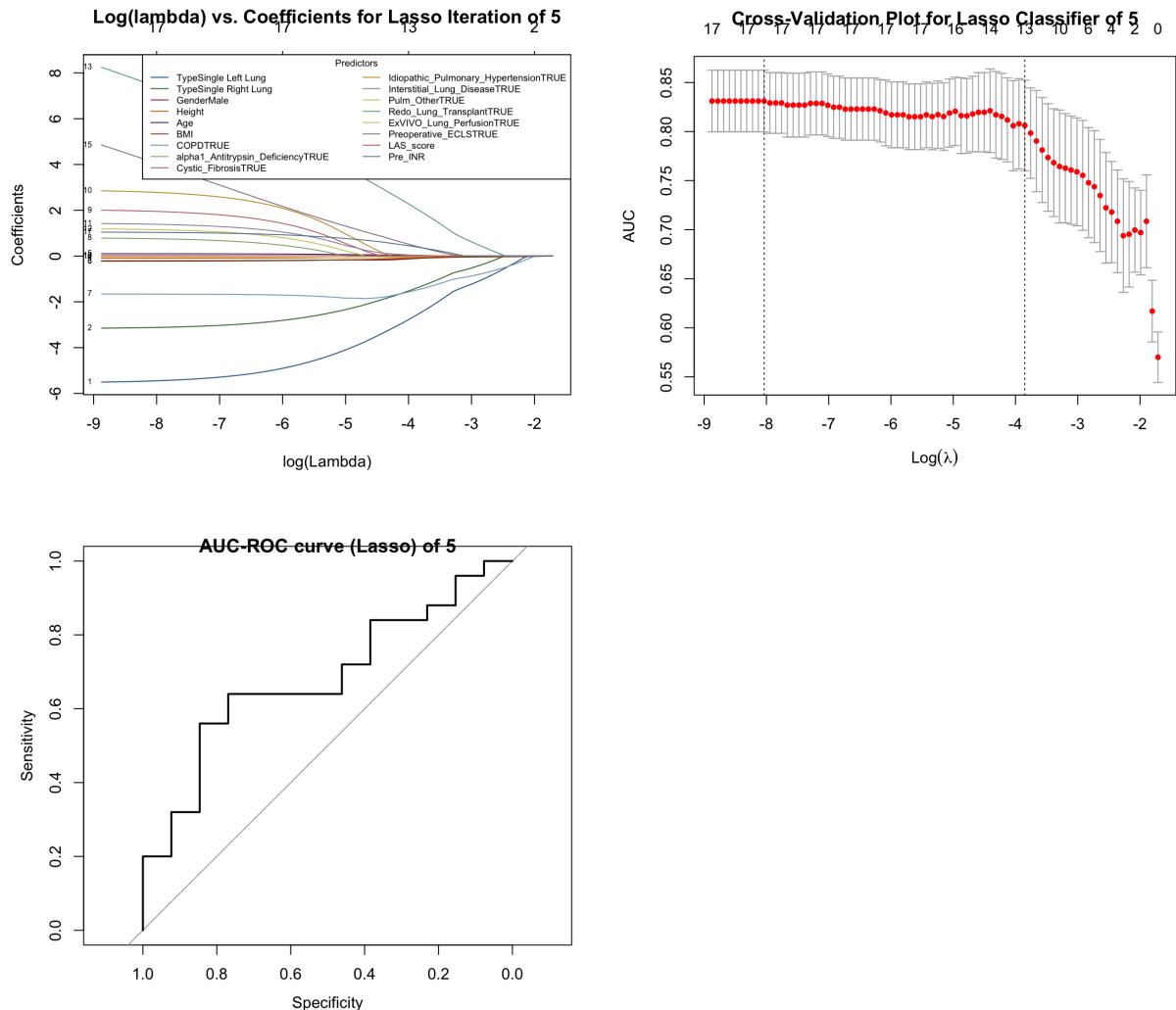


**Figure 3: Summary of the LASSO path, cross-validation plots for the LASSO Classifier and the AUC-ROC curves for each iteration. A total of 5 iterations.**

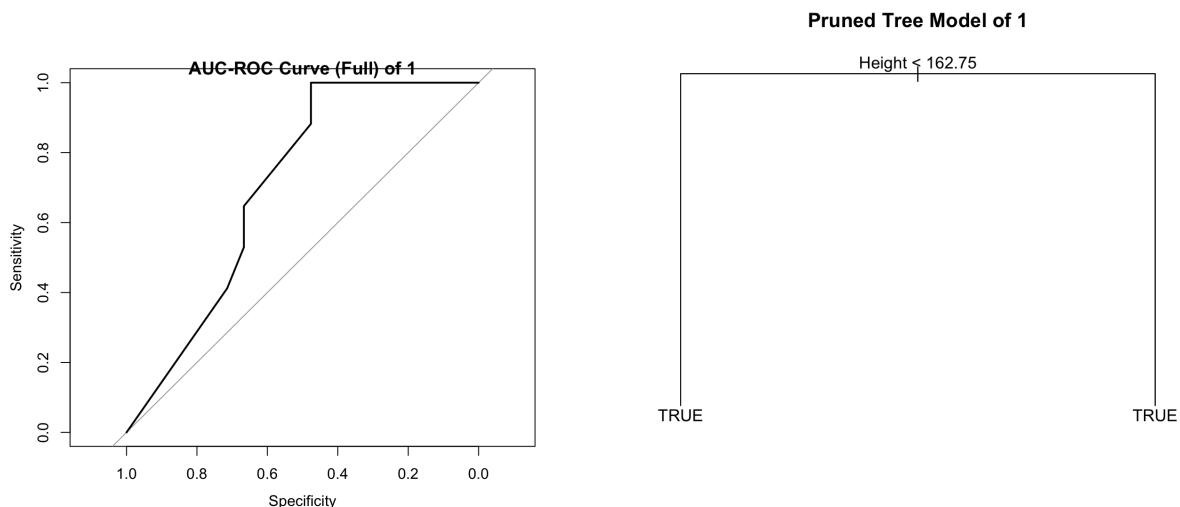
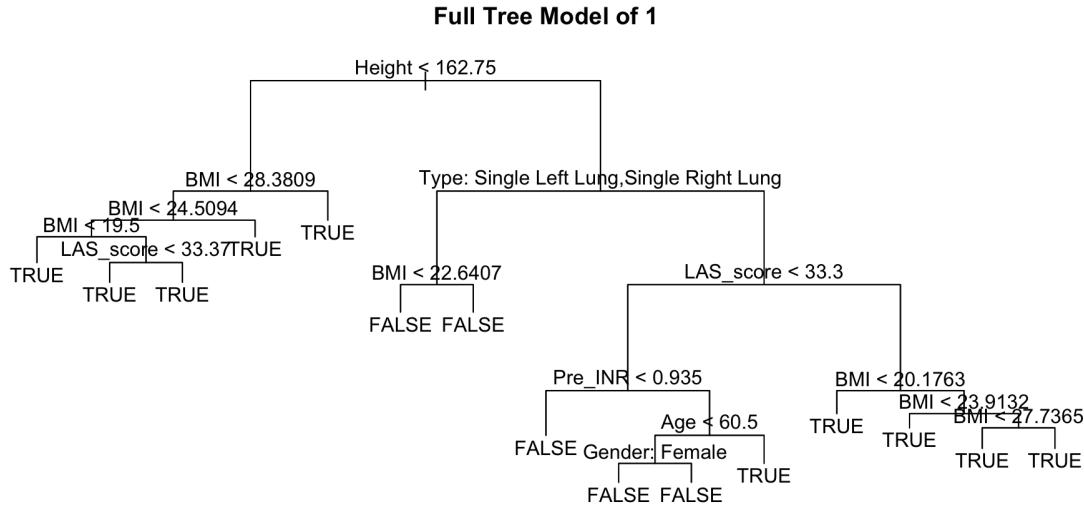


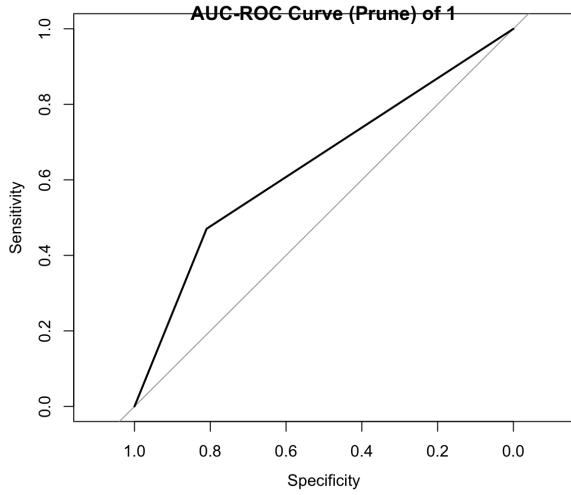




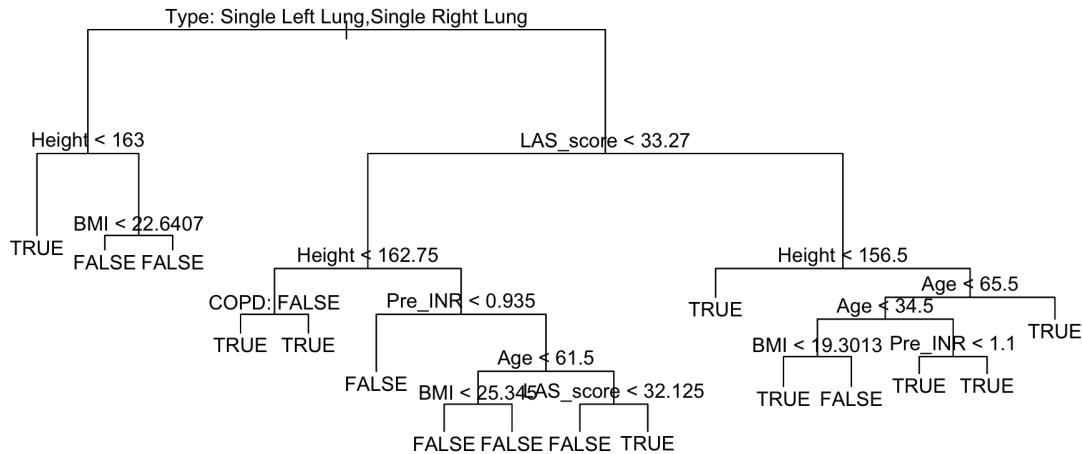


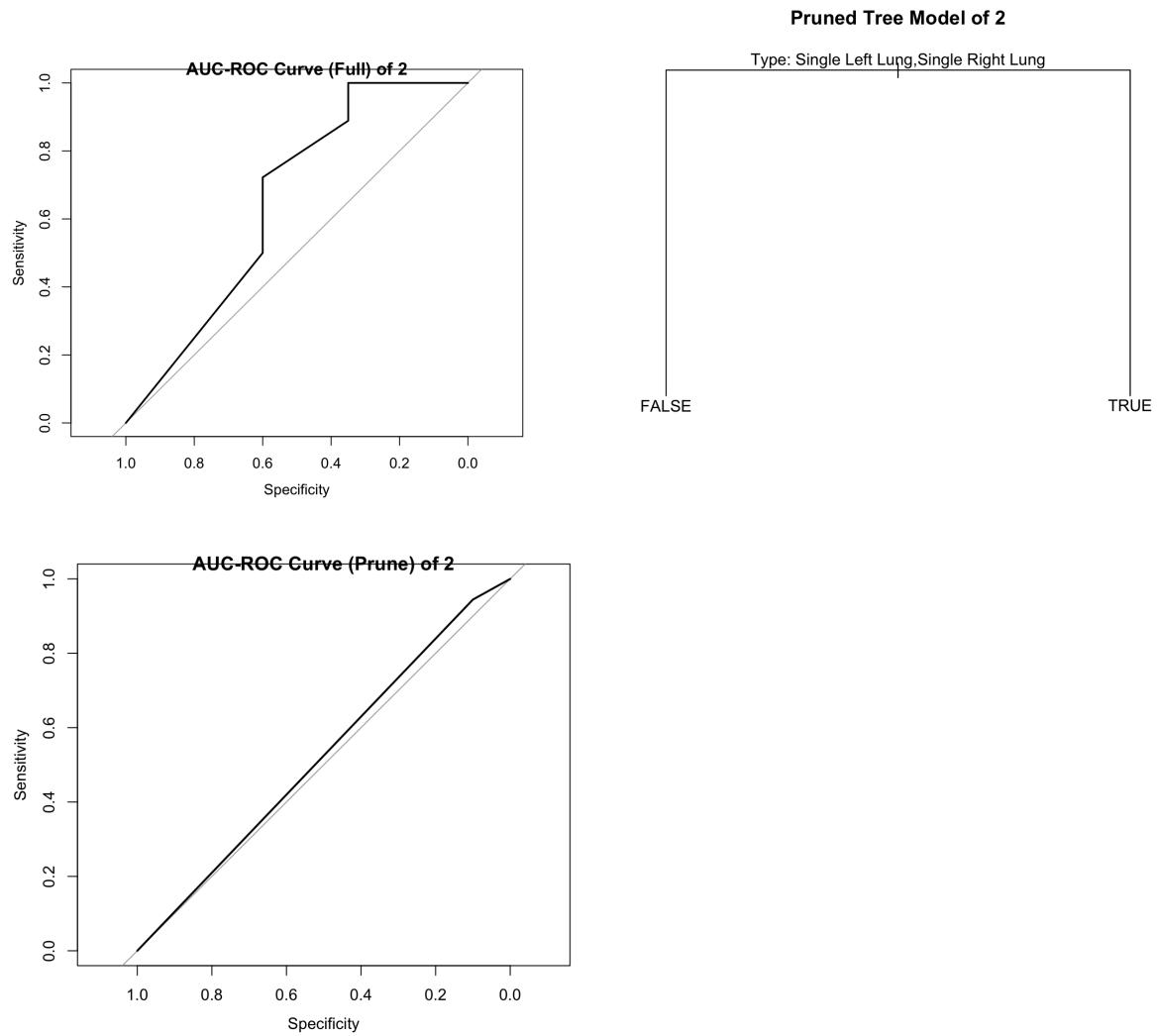
**Figure 4: Summary of the full tree, its respective AUC-ROC curve, pruned tree, and its respective AUC-ROC curve for each iteration. A total of 5 iterations.**

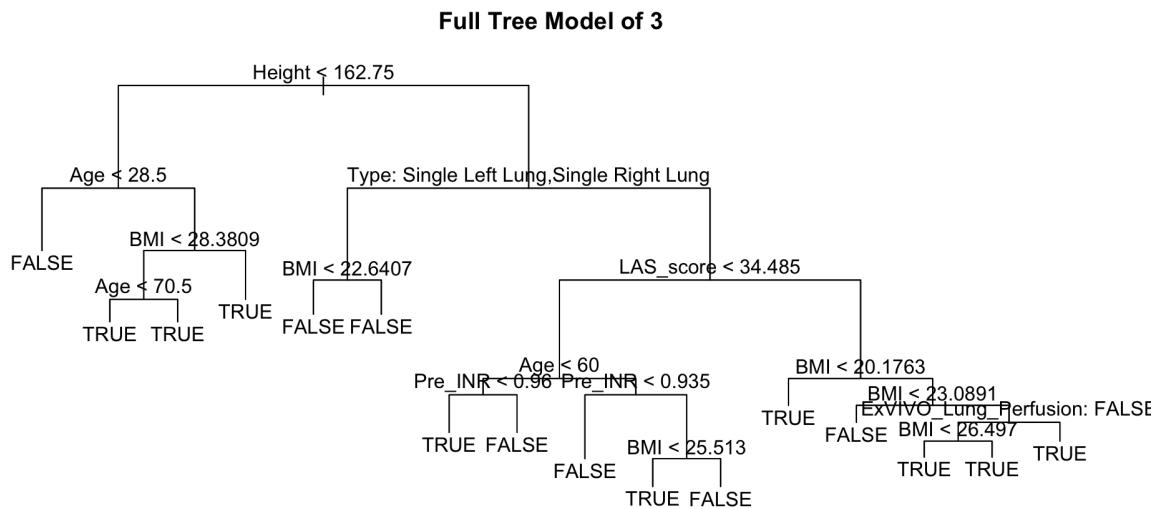


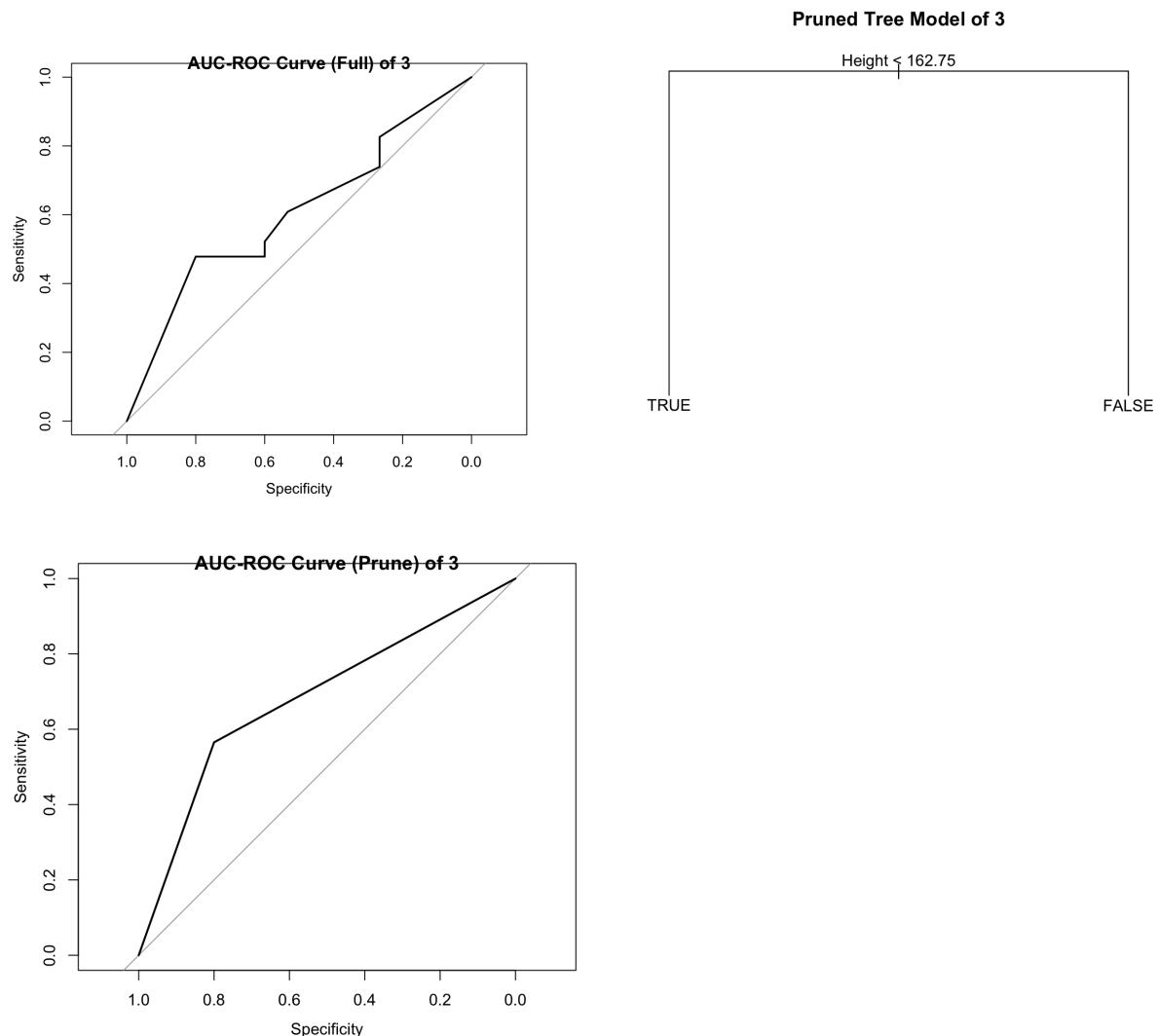


### Full Tree Model of 2

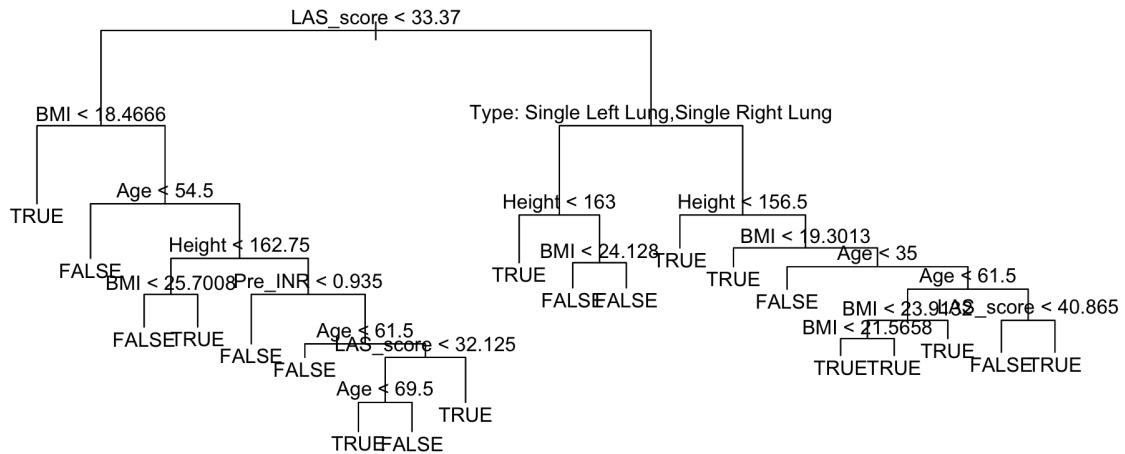




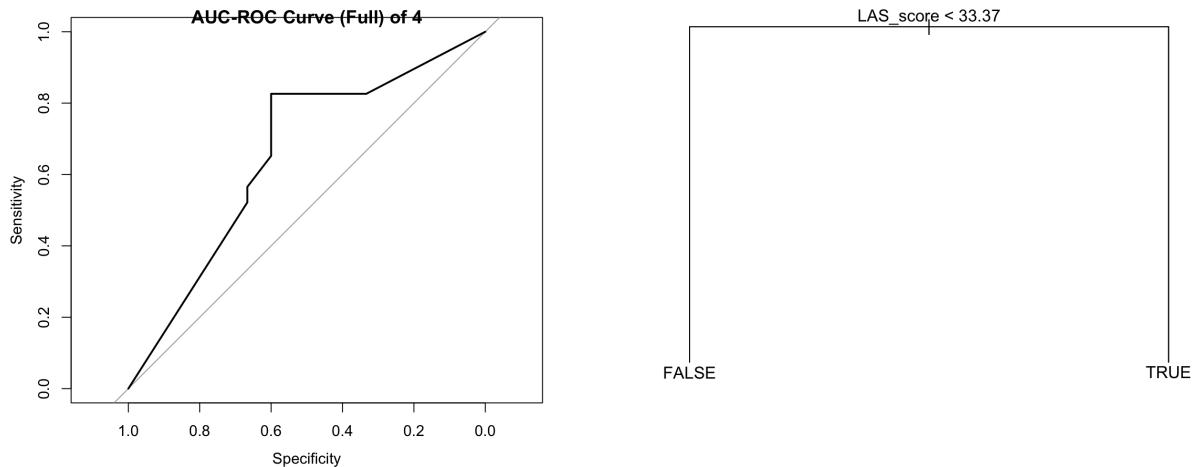


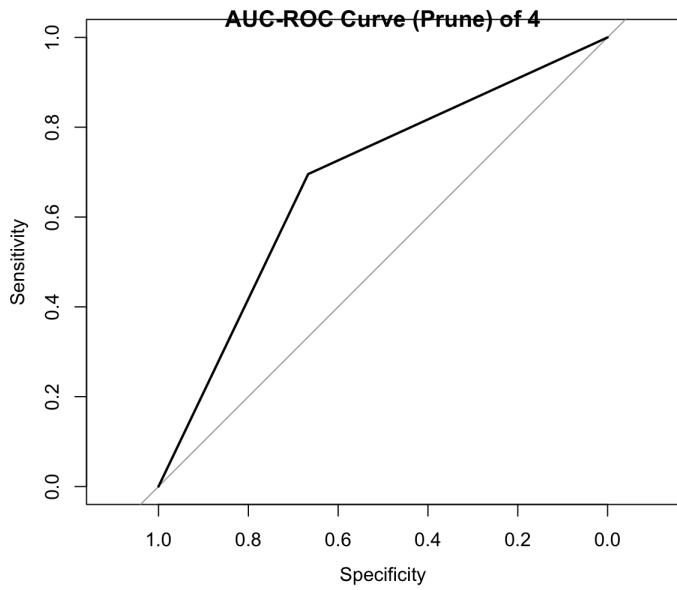


## Full Tree Model of 4

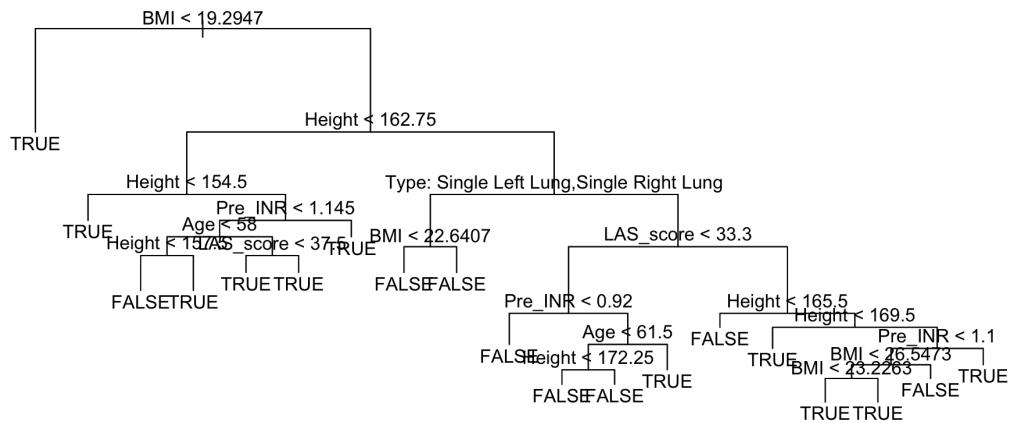


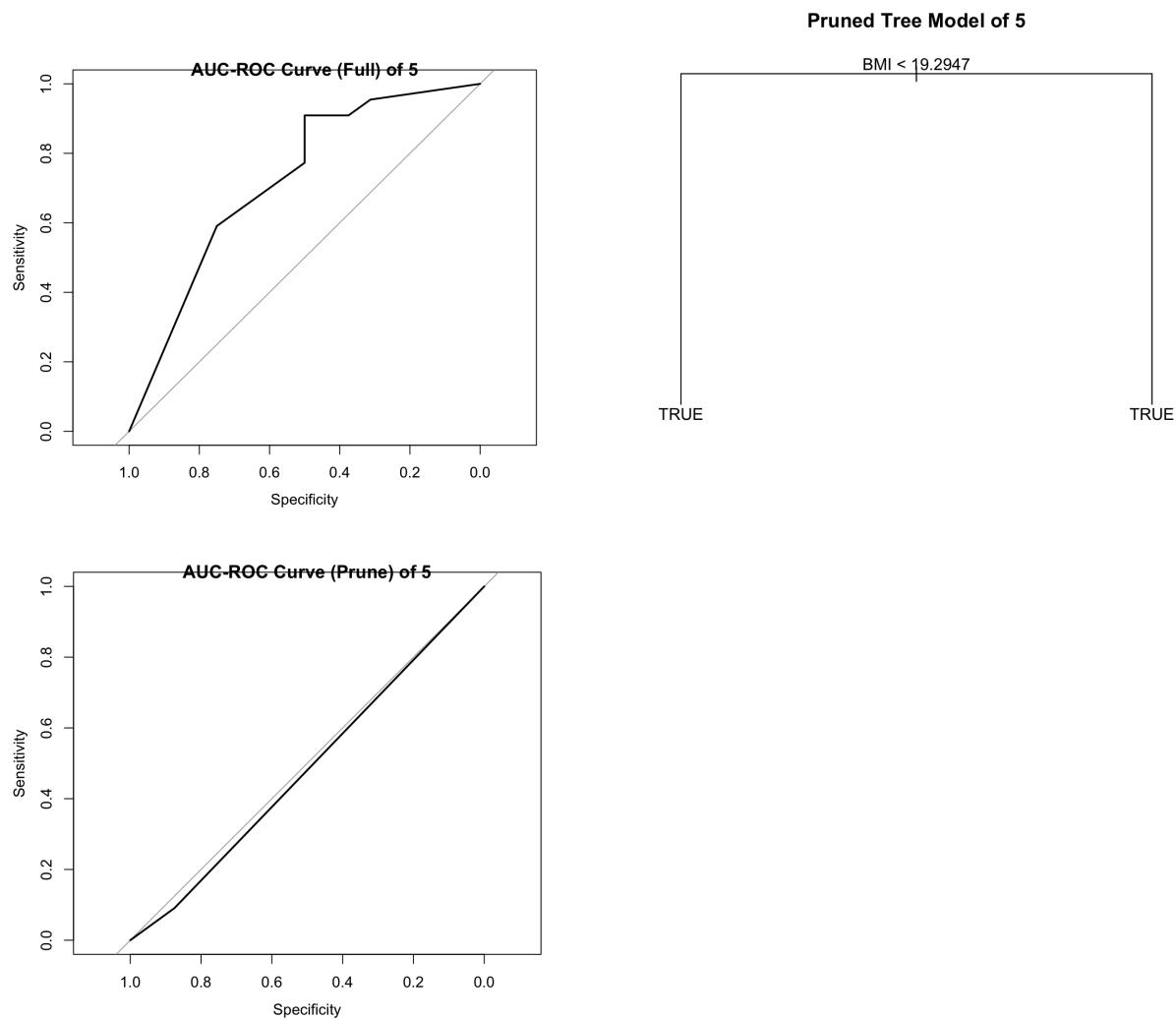
## Pruned Tree Model of 4



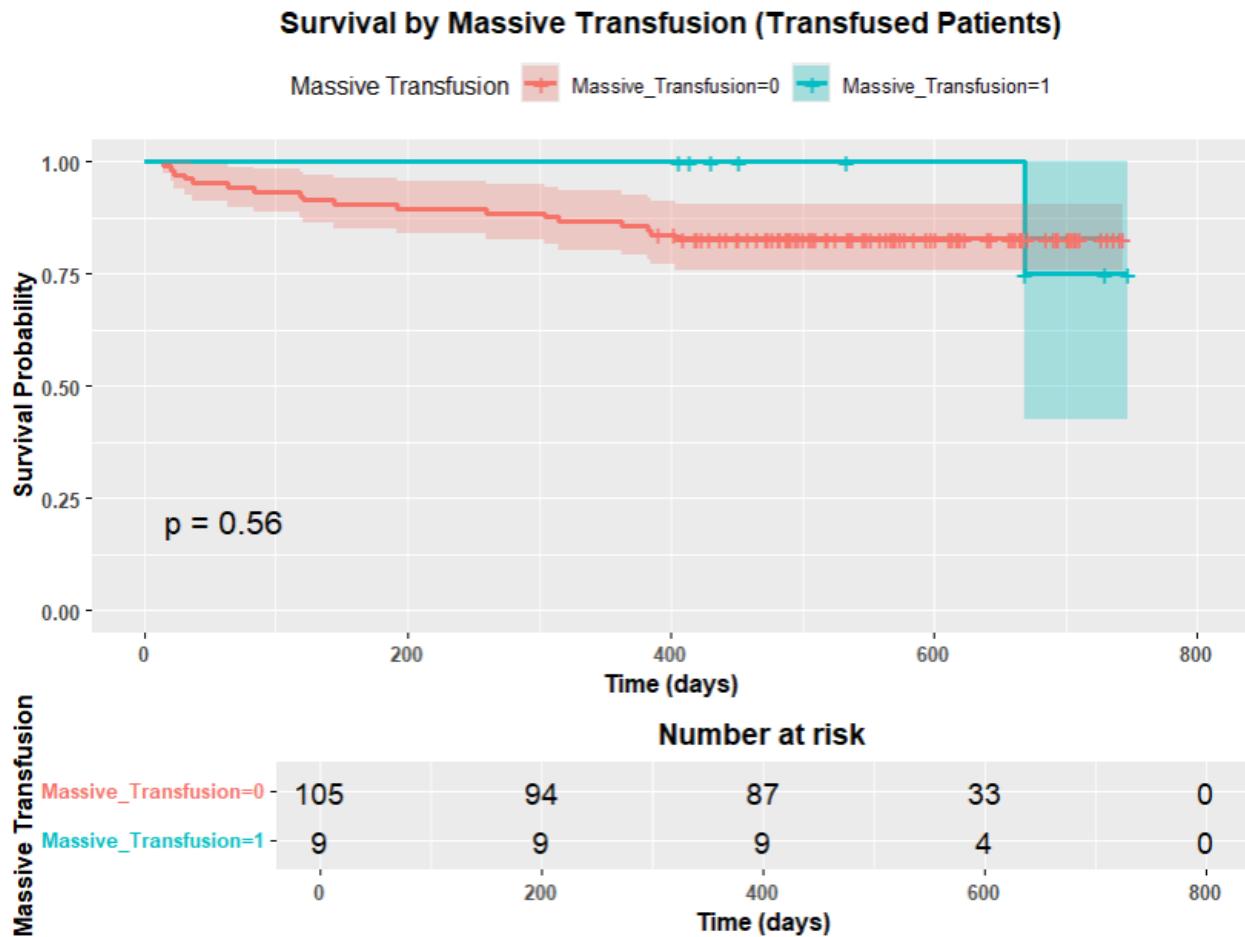


## Full Tree Model of 5

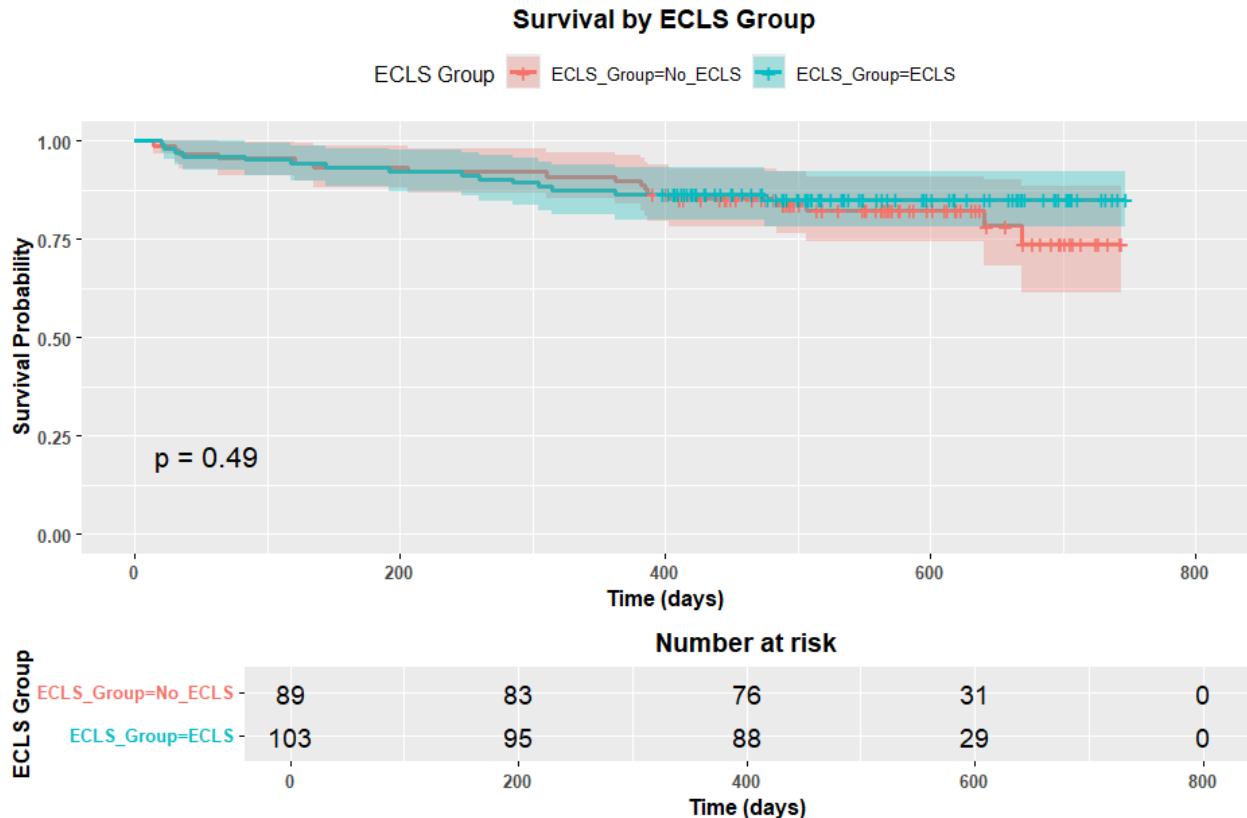




**Figure XX:** Kaplan Meier curve showing the survival probability at different time points among those who received a massive transfusion vs those who did not, in addition to a risk table.



**Figure XX:** Kaplan Meier curve showing the survival probability at different time points among those who received life support vs those who did not, in addition to a risk table.



## Appendix B. Supplementary Tables

**Table 1: Summary of Continuous Variables**

Variables	min	max	median	mean	var	std.dev	coef.var
Height	140.00000	192.00000	167.00000	166.69010415	102.11555701	10.10522424	0.06062282
Weight	33.10000	107.80000	67.10000	67.44479184	213.22144363	14.60210408	0.21650455
Age	19.00000	76.00000	61.00000	56.25520833	218.95547011	14.79714399	0.26303598
BMI	15.54134	31.78146	24.37445	24.10988047	15.84228289	3.98023654	0.16508736
LAS_score	24.00000	90.00000	34.00000	36.90772222	83.63285456	9.14510003	0.24778283
Pre_Hb	73.00000	168.00000	131.00000	127.84375000	383.59849476	19.58567065	0.15320006
Pre_Hct	0.22000	0.51000	0.40000	0.39384375	0.00299074	0.05468766	0.13885622
Pre_Platelets	89.00000	882.00000	243.50000	264.02604167	10,671.69041776	103.30387417	0.39126396
Pre_PT	9.60000	37.40000	11.60000	12.95156249	22.06984027	4.69785486	0.36272495
Pre_INR	0.80000	3.70000	1.00000	1.11317708	0.21946891	0.46847509	0.42084507
Pre_PTT	17.70000	93.50000	24.50000	26.06806283	63.78923732	7.98681647	0.30638320
Pre_Fibrinogen	1.30000	6.22000	3.20000	3.95399994	4.49737979	2.12070266	0.53634363
Pre_Creatinine	37.00000	169.00000	69.00000	74.55208333	426.80355585	20.65922447	0.27711130
Intra_Fresh_Frozen_Plasma	0.00000	26.00000	0.00000	0.61458333	6.43706370	2.53713691	4.12822277
Intra_Packed_Cells	0.00000	36.00000	1.00000	2.26562500	17.88195353	4.22870589	1.86646329
Intra_Platelets	0.00000	11.00000	0.00000	0.43229167	1.93256435	1.39016702	3.21580806
Intra_Cryoprecipitate	0.00000	50.00000	0.00000	2.70833333	54.40663176	7.37608512	2.72347758
Duration_of_ICU_Stay__days_	1.00000	99.00000	4.00000	7.29319372	131.11358501	11.45048405	1.57002330
ICU_LOS	0.63000	380.86000	3.62000	9.24203125	855.83771156	29.25470409	3.16539766
HOSPITAL_LOS	9.66000	381.52000	22.84000	33.60000000	1,794.23689319	42.35843355	1.26066767
RBC_0-24hrs	0.00000	9.00000	1.00000	1.30000000	4.51864407	2.12571025	1.63516173
RBC_24-48hrs	0.00000	8.00000	0.00000	0.80769231	2.11915535	1.45573190	1.80233473
RBC_48-72hrs	0.00000	8.00000	1.00000	0.70833333	1.48758865	1.21966744	1.72188344
RBC_72hr_Total	0.00000	23.00000	0.00000	0.80208333	4.95538831	2.22607015	2.77536019
FFP_0-24hrs	0.00000	7.00000	0.00000	0.86842105	3.30654339	1.81839033	2.09390401
FFP_24-48hrs	0.00000	1.00000	0.00000	0.03030303	0.03030303	0.17407766	5.74456265
FFP_48-72hrs	0.00000	2.00000	0.00000	0.06060606	0.12121212	0.34815531	5.74456265
FFP_72hr_Total	0.00000	8.00000	0.00000	0.18750000	0.95942408	0.97950196	5.22401044
Plt_0-24hrs	0.00000	4.00000	0.00000	0.36842105	0.77951636	0.88290224	2.39644893
Plt_24-48hrs	0.00000	4.00000	0.00000	0.22222222	0.57777778	0.76011695	3.42052628
Plt_48-72hrs	0.00000	4.00000	0.00000	0.14705882	0.49286988	0.70204692	4.77391904
Plt_72hr_Total	0.00000	10.00000	0.00000	0.14062500	0.70786976	0.84134997	5.98293309
Cryo_0-24hrs	0.00000	10.00000	0.00000	1.89189189	15.76576577	3.97061277	2.09875246
Cryo_24-48hrs	0.00000	10.00000	0.00000	0.30303030	3.03030303	1.74077656	5.74456265
Cryo_48-72hrs	0.00000	10.00000	0.00000	0.30303030	3.03030303	1.74077656	5.74456265
Cryo_72hr_Total	0.00000	30.00000	0.00000	0.46875000	7.63252618	2.76270269	5.89376575
Total_24hr_RBC	0.00000	38.00000	1.00000	2.67187500	24.77658704	4.97760857	1.86296461

**Table 2: Summary of Categorical Variables**

Variable	Count	Proportion
Type - Bilateral	157	81.77
Type - Single Left Lung	18	9.38
Type - Single Right Lung	17	8.85
Gender - Female	88	45.83
Gender - Male	104	54.17
COPD - FALSE	134	69.79
COPD - TRUE	58	30.21
alpha1-Antitrypsin_Deficiency - FALSE	183	95.31
alpha1-Antitrypsin_Deficiency - TRUE	9	4.69
Cystic_Fibrosis - FALSE	162	84.38
Cystic_Fibrosis - TRUE	30	15.62
Idiopathic_Pulmonary_Hypertension - FALSE	186	96.88
Idiopathic_Pulmonary_Hypertension - TRUE	6	3.12
Interstitial_Lung_Disease - FALSE	100	52.08
Interstitial_Lung_Disease - TRUE	92	47.92
Pulm_Other - FALSE	177	92.19
Pulm_Other - TRUE	15	7.81
First_Lung_Transplant - FALSE	11	5.73
First_Lung_Transplant - TRUE	181	94.27
Redo_Lung_Transplant - FALSE	182	94.79
Redo_Lung_Transplant - TRUE	10	5.21
ExVIVO_Lung_Perfusion - FALSE	115	59.90
ExVIVO_Lung_Perfusion - TRUE	77	40.10
Preoperative_ECLS - FALSE	185	96.35
Preoperative_ECLS - TRUE	7	3.65
Intraoperative_ECLS - FALSE	89	46.35
Intraoperative_ECLS - TRUE	103	53.65
ECLS_ECMO - FALSE	91	47.40
ECLS_ECMO - TRUE	101	52.60
ECLS_CPB - FALSE	190	98.96
ECLS_CPB - TRUE	2	1.04
ALIVE_30DAYS_YN - N	3	1.56
ALIVE_30DAYS_YN - Y	189	98.44
ALIVE_90DAYS_YN - N	9	4.69
ALIVE_90DAYS_YN - Y	183	95.31
ALIVE_12MTHS_YN - N	23	11.98
ALIVE_12MTHS_YN - Y	169	88.02
Massive_Transfusion - 0	183	95.31
Massive_Transfusion - 1	9	4.69
Transfusion - FALSE	78	40.62
Transfusion - TRUE	114	59.38

**Table 3: Summary of Continuous Variables for Patients who received Transfusion**

Variables	min	max	median	mean	var	std.dev	coef.var
Height	140.00000	183.00000	163.00000	163.58947366	95.653340051	9.78025256	0.05978534
Weight	33.10000	95.00000	63.35000	63.22280707	194.270804814	13.93810621	0.22046010
Age	21.00000	76.00000	58.00000	53.72807018	221.898851110	14.89626970	0.27725302
BMI	15.54134	31.78146	23.84518	23.49959271	17.327710583	4.16265667	0.17713740
LAS_score	24.00000	90.00000	35.00000	38.39439252	96.806221090	9.83901525	0.25626178
Pre_Hb	73.00000	164.00000	122.50000	120.60526316	388.152538426	19.70158721	0.16335595
Pre_Hct	0.22000	0.50000	0.38000	0.37492983	0.003134951	0.05599063	0.14933629
Pre_Platelets	89.00000	882.00000	257.00000	276.89473684	13,043.121564974	114.20648653	0.41245452
Pre_PT	9.60000	37.40000	11.60000	13.09210526	23.788698178	4.87736591	0.37254252
Pre_INR	0.80000	3.70000	1.00000	1.12991228	0.243641582	0.49360063	0.43684863
Pre_PTT	17.70000	93.50000	24.80000	26.97017537	91.416800671	9.56121335	0.35451061
Pre_Fibrinogen	1.30000	6.22000	4.62000	4.18999993	5.625199758	2.37175036	0.56605021
Pre_Creatinine	37.00000	169.00000	67.00000	74.04385965	558.307793821	23.62853770	0.31911542
Intra_Fresh_Frozen_Plasma	0.00000	26.00000	0.00000	1.03508772	10.441235833	3.23129012	3.12175486
Intra_Packed_Cells	0.00000	36.00000	3.00000	3.81578947	24.257801584	4.92522097	1.29074756
Intra_Platelets	0.00000	11.00000	0.00000	0.70175439	3.060704859	1.74948703	2.49301901
Intra_Cryoprecipitate	0.00000	50.00000	0.00000	4.38596491	83.247942866	9.12403106	2.08027908
Duration_of_ICU_Stay__days_	1.00000	99.00000	4.00000	8.22123894	128.870259166	11.35210373	1.38082640
ICU_LOS	0.63000	380.86000	4.55500	11.47298246	1,345.055668902	36.67500060	3.19664052
HOSPITAL_LOS	11.08000	381.52000	25.54000	37.50307018	2,091.696207305	45.73506540	1.21950190
RBC_0-24hrs	0.00000	9.00000	1.00000	1.77272727	5.342494715	2.31138372	1.30385748
RBC_24-48hrs	0.00000	6.00000	1.00000	0.91176471	1.355614973	1.16430880	1.27698384
RBC_48-72hrs	0.00000	8.00000	1.00000	0.93548387	1.995698925	1.41269208	1.51011912
RBC_72hr_Total	0.00000	23.00000	0.00000	1.21052632	7.300419190	2.70192879	2.23202813
FFP_0-24hrs	0.00000	7.00000	0.00000	1.50000000	4.833333333	2.19848433	1.46565622
FFP_24-48hrs	0.00000	1.00000	0.00000	0.05882353	0.058823529	0.24253563	4.12310563
FFP_48-72hrs	0.00000	2.00000	0.00000	0.11764706	0.235294118	0.48507125	4.12310563
FFP_72hr_Total	0.00000	8.00000	0.00000	0.31578947	1.580810433	1.25730284	3.98145900
Plt_0-24hrs	0.00000	4.00000	0.00000	0.63636364	1.194805195	1.09307145	1.71768371
Plt_24-48hrs	0.00000	2.00000	0.00000	0.16666667	0.264705882	0.51449576	3.08697453
Plt_48-72hrs	0.00000	4.00000	0.00000	0.27777778	0.918300654	0.95828005	3.44980818
Plt_72hr_Total	0.00000	10.00000	0.00000	0.19298246	1.042074212	1.02082036	5.28970551
Cryo_0-24hrs	0.00000	10.00000	0.00000	2.85714286	21.428571429	4.62910050	1.62018517
Cryo_24-48hrs	0.00000	10.00000	0.00000	0.58823529	5.882352941	2.42535625	4.12310563
Cryo_48-72hrs	0.00000	10.00000	0.00000	0.58823529	5.882352941	2.42535625	4.12310563
Cryo_72hr_Total	0.00000	30.00000	0.00000	0.70175439	11.892563267	3.44855959	4.91419742
Total_24hr_RBC	1.00000	38.00000	3.00000	4.50000000	33.579646018	5.79479473	1.28773216

**Table 4: Summary of Categorical Variables for Patients who received Transfusion**

Variables	Count	Proportion
Type - Bilateral	105	92.11
Type - Single Left Lung	4	3.51
Type - Single Right Lung	5	4.39
Gender - Female	64	56.14
Gender - Male	50	43.86
COPD - FALSE	91	79.82
COPD - TRUE	23	20.18
alpha1-Antitrypsin_Deficiency - FALSE	109	95.61
alpha1-Antitrypsin_Deficiency - TRUE	5	4.39
Cystic_Fibrosis - FALSE	89	78.07
Cystic_Fibrosis - TRUE	25	21.93
Idiopathic_Pulmonary_Hypertension - FALSE	110	96.49
Idiopathic_Pulmonary_Hypertension - TRUE	4	3.51
Interstitial_Lung_Disease - FALSE	57	50.00
Interstitial_Lung_Disease - TRUE	57	50.00
Pulm_Other - FALSE	101	88.60
Pulm_Other - TRUE	13	11.40
First_Lung_Transplant - FALSE	10	8.77
First_Lung_Transplant - TRUE	104	91.23
Redo_Lung_Transplant - FALSE	104	91.23
Redo_Lung_Transplant - TRUE	10	8.77
ExVIVO_Lung_Perfusion - FALSE	75	65.79
ExVIVO_Lung_Perfusion - TRUE	39	34.21
Preoperative_ECLS - FALSE	107	93.86
Preoperative_ECLS - TRUE	7	6.14
Intraoperative_ECLS - FALSE	35	30.70
Intraoperative_ECLS - TRUE	79	69.30
ECLS(ECMO) - FALSE	37	32.46
ECLS(ECMO) - TRUE	77	67.54
ECLS(CPB) - FALSE	112	98.25
ECLS(CPB) - TRUE	2	1.75
ALIVE_30DAYS_YN - N	3	2.63
ALIVE_30DAYS_YN - Y	111	97.37
ALIVE_90DAYS_YN - N	7	6.14
ALIVE_90DAYS_YN - Y	107	93.86
ALIVE_12MTHS_YN - N	15	13.16
ALIVE_12MTHS_YN - Y	99	86.84
Massive_Transfusion - 0	105	92.11
Massive_Transfusion - 1	9	7.89

**Table 5: LASSO Classifiers - Selected Predictors for Each Iteration**

Lasso Classifiers - Selected Predictors for Each Iteration	
Iteration	Predictors
1	TypeSingle Left Lung, TypeSingle Right Lung, GenderMale, Height, Age, BMI, COPDTRUE, alpha1_Antitrypsin_DeficiencyTRUE, Cystic_FibrosisTRUE, Idiopathic_Pulmonary_HypertensionTRUE, Interstitial_Lung_DiseaseTRUE, Pulm_OtherTRUE, Redo_Lung_TransplantTRUE, ExVIVO_Lung_PerfusionTRUE, Preoperative_ECLSTRAUE, LAS_score, Pre_INR
2	TypeSingle Left Lung, TypeSingle Right Lung, GenderMale, Height, Age, BMI, COPDTRUE, alpha1_Antitrypsin_DeficiencyTRUE, Cystic_FibrosisTRUE, Idiopathic_Pulmonary_HypertensionTRUE, Interstitial_Lung_DiseaseTRUE, Pulm_OtherTRUE, Redo_Lung_TransplantTRUE, ExVIVO_Lung_PerfusionTRUE, Preoperative_ECLSTRAUE, LAS_score, Pre_INR
3	TypeSingle Left Lung, TypeSingle Right Lung, Height, BMI, COPDTRUE, Cystic_FibrosisTRUE, Pulm_OtherTRUE, Redo_Lung_TransplantTRUE, Preoperative_ECLSTRAUE, LAS_score, Pre_INR
4	TypeSingle Left Lung, TypeSingle Right Lung, GenderMale, Height, Age, BMI, COPDTRUE, Cystic_FibrosisTRUE, Idiopathic_Pulmonary_HypertensionTRUE, Interstitial_Lung_DiseaseTRUE, Pulm_OtherTRUE, Redo_Lung_TransplantTRUE, ExVIVO_Lung_PerfusionTRUE, Preoperative_ECLSTRAUE, LAS_score, Pre_INR
5	TypeSingle Left Lung, TypeSingle Right Lung, GenderMale, Height, Age, BMI, COPDTRUE, alpha1_Antitrypsin_DeficiencyTRUE, Cystic_FibrosisTRUE, Idiopathic_Pulmonary_HypertensionTRUE, Interstitial_Lung_DiseaseTRUE, Pulm_OtherTRUE, Redo_Lung_TransplantTRUE, ExVIVO_Lung_PerfusionTRUE, Preoperative_ECLSTRAUE, LAS_score, Pre_INR

**Table 6: LASSO Classifiers - AUC for Each Iteration**

Lasso Classifiers - AUC for Each Iteration	
Iteration	AUC
1	0.90
2	0.80
3	0.78
4	0.84
5	0.68

**Table 7: CART Tree - Selected Predictors for Each Iteration**

CART Tree - Selected Predictors for Each Iteration	
Iteration	Predictors
1	Height, BMI, LAS_score, Type, Pre_INR, Age, Gender
2	Type, Height, BMI, LAS_score, COPD, Pre_INR, Age
3	Height, Age, BMI, Type, LAS_score, Pre_INR, ExVIVO_Lung_Perfusion
4	LAS_score, BMI, Age, Height, Pre_INR, Type
5	BMI, Height, Pre_INR, Age, LAS_score, Type

**Table 8: CART Tree - AUC for Each Iteration**

CART Tree - AUC for Each Iteration	
Iteration	AUC
1	0.70
2	0.65
3	0.60
4	0.65
5	0.72

**Table 9: CART Pruned Tree - Selected Predictors for Each Iteration**

CART Pruned Tree - Selected Predictors for Each Iteration	
Iteration	Predictors
1	Height
2	Type
3	Height
4	LAS_score
5	BMI

**Table 10: CART Pruned Tree - AUC for Each Iteration**

CART Pruned Tree - AUC for Each Iteration	
Iteration	AUC
1	0.64
2	0.52
3	0.68
4	0.68
5	0.48