# THE IMPORTANCE OF DONOR VARIABLES ON PEDIATRIC HEART TRANSPLANT SURVIVAL

---

A

Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Master of Science

by

Julia Sharff

May 2024

# APPROVAL SHEET

This

Thesis

is submitted in partial fulfillment of the requirements
for the degree of

Master of Science

Author: Julia Sharff

This Thesis has been read and approved by the examing committee:

Advisor: Michael D. Porter

Advisor:

Committee Member: William T. Scherer

Committee Member: Robert Riggs

Committee Member:

Committee Member:

Committee Member:

Committee Member:

Accepted for the School of Engineering and Applied Science:

Jennifer L. West, School of Engineering and Applied Science

May 2024

**Background:** Waitlist mortality for pediatric heart transplant candidates remains high, with nearly 10% dying prior to receiving a heart. Despite this, there are a large number of donor hearts available for transplant, but refused due to poor donor quality. The 2020 ISHLT consensus statement suggested that few donor characteristics have an impact on post-transplant mortality. To test the validity of this, we utilized two machine learning models to evaluate the importance of various donor characteristics on patient survival post-transplant.

**Methods:** Random Forest (RF) and Lasso Logistic Regression (LR) were used to predict 1-year, 3-year, and 5-year post transplant survival using OPTN/UNOS data (2010-2019) for pediatric heart recipients. Candidates listed for multiple organs, re-transplantation, or with donors over the age of 30, were excluded from this study leaving a total of 3882 patients. RF and LR models were fit using combinations of donor, candidate, donor-candidate compatibility, and transplant predictor variables. A comparison of the AUC values, Brier scores, and log loss from 10-fold cross validation was used to assess differences in model performance.

**Results:** The LR models had higher average AUC, and lower brier score and log loss score, model performance for 1-year survival, with the best overall model (AUC = .754) coming from the candidate and donor variables. The random forest candidate model performed best for 3-year survival with an average AUC of .69. The RF models achieved better performance over LR for all 5-year survival models. The comparison of the model metrics from the different variable groups show that there is no statistically significant model improvement from the addition of non-candidate related variables. Further exploration into ischemic times suggest longer times are associated with reduced survival probability, making it difficult to determine its impact in models.

**Discussion:** The use of additional variables outside of echo and ischemic time when determining post-transplant success may be unnecessary. While ischemic time and echocardiograph measures were statistically insignificant as well, it is important to note that these conclusions are drawn based on contemporary donor heart selection practice, which displays little variability in accepted donor characteristics. Very few data points fell outside what is generally considered acceptable. The findings suggest that further evaluation on ischemic time and echo abnormality should be done to determine their impact on post-transplant survival.

## Table of Contents

# Introduction

## *Motivation and Contributions*

The decision to accept or decline a heart is a time sensitive and potentially life-threatening action. There are many factors to consider in the moment the offer comes through, including numerous candidate and donor characteristics. Machine Learning models have the potential to ease the decision-making process through post-transplant survival predictions. The goal of the work summarized in this thesis is to provide a survival model for pediatric heart transplants. This survival model is a piece of a larger pediatric heart transplant project aiming to utilize machine learning to assess the entire transplant process from the decision to list to the post-transplant outcomes. The ultimate goal is to create a dashboard allowing doctors to visualize every component of the candidates transplant process and better make decisions to optimize the patients chance of survival.

In addition to the survival model, the work in this thesis explores the impact of donor related variables on post-transplant outcomes. The amount of information presented to a doctor at time of offer is abundant, making it difficult to fully assess. The focus of the study is based on the International Society for Heart and Lung Transplantation Consensus Report on donor characteristics impact on post-transplant outcomes in pediatric heart transplants (Kirk, 2020). The ability to focus primarily on the candidate characteristics would cut down on the amount of information to process at the time of offer and allow doctors to make a more confident decision.

The work summarized in this thesis presents a new way to assess the relevant characteristics on the outcome of pediatric heart transplant. Shifting from the traditional variable importance methods, variable selection is based on predictive model performance to determine what is necessary for donor heart offer assessment. Due to its design around the ISHLT consensus, the work provides a data driven assessment of the claims made on donor characteristics impact. Further, the final survival model produced can be used to predict post-transplant outcomes, a key component of the pediatric heart transplant system dashboard.

## Background

While there have been positive trends in the survival of pediatric patients post heart transplantation, mortality on the waitlist remains high (Kirk, 2020; Butts, 2021). Due to the transplant centers' reliance on high transplant success rates for center ratings, many hospitals and doctors remain strict when determining the quality and usefulness of a donor hearts (Khush, 2015). These narrow acceptance windows create longer wait times, likely contributing to higher pre-transplant mortality, with 10.9 deaths for every 100 patients on the wait list (Colvin, 2021). With over 30% of pediatric donor hearts going unused after refusal for poor donor quality, it is important to determine the impact that donor characteristics have on post-transplant outcome (Kirk, 2020).

### *Literature Review*

The importance of post-transplant success has resulted in the implementation of machine learning (ML) techniques to better understand the factors driving success. Studies regarding pediatric heart transplants attempt to better understand the potential for modeling in post-transplant outcomes and determine the relationship between various donor and patient characteristics on the outcome.

Miller et. al. (2019) studied pediatric heart transplants from 2006-2015, addressing 1, 3, and 5-year survival. The 1-year survival random forest model, containing donor and recipient characteristics was found to be the strongest model, with a training AUC of .74 and testing AUC of .72. Important features identified include recipient characteristics candidate diagnosis, ECMO at transplant, and mechanical ventilation at transplant, and donor characteristics gender and B1 antigen levels.

Killian et. al. (2023) studied pediatric heart transplants from 1987-2019, addressing 1, 3, and 5-year survival, using candidate and donor characteristics. The random forest model performed best at all levels, achieving a 1-year AUC of .697. The most important features coming from that model include candidate characteristics of days at status 1A, VAD device type, and malignancies since listing, as well as the post-transplant graft status. Additionally, the donor/recipient height ratio was found to be an important feature.

Miller et. al. (2022) studied both adult and pediatric heart transplants from 1994 – 2016 using both classification and survival modeling methods for 1-year and 90 day post-transplant survival. In the pediatric study, the 1-year survival random forest model achieved the highest AUC of .836. Finally, Ashfaq et. al. (2022) studied pediatric heart transplants from 2010 – 2020 to model 1-year post-transplant survival. The random forest model achieved the highest C-index at .68.

All of the above studies found the most success with random forest models, achieving a range of AUC values considered to be acceptable (.7-.8) and excellent (.8-.9) (Mandrekar, 2010). However, a lack of quality control within the data entry process across numerous institutions and the infrequency of post-transplant mortality make modeling challenging (Killian, 2023; Miller, 2019). Additionally, deciding on the most important features is difficult, with different features identified as related to post transplant outcome and little overlap between the studies.

Outside of modeling outcomes, ML has been used to aid in the identification of important features in the survival of a patient post-transplant, with a focus on the donor related variables impacting the decision to accept a heart or not. Currently, the window for donor heart quality is tight, with doctors hesitant to accept marginal donor hearts (Feingold, 2018). In an effort to capture these important candidate and donor heart metrics, risk scoring methods have been explored, assigning points based on how large the gap in these metrics are from what is considered optimal (Zafar, 2018; Fraser, 2019). While these systems have highlighted important relationships between the donor and candidate, others have questioned the necessity of such strict acceptance practices in pediatrics (Conway, 2020; Feingold, 2018). Conway et. al. (2020) addressed the concern on size matching in their literature review, stating that current guidelines are conservative and can be expanded upon. Further, a study performed to test the impact of donor heart refusal for organ quality on post-transplant outcomes found no correlation between number of refusals and patient survival (Rizwan, 2018).

### *ISHLT Consensus*

Due to the highly subjective nature of the donor acceptance decision process, attempts have been made to generalize the donor heart selection process. The International Society for Heart and Lung Transplantation released a consensus paper regarding acceptance practices of pediatric donor hearts (Kirk, 2020). Two important non-candidate characteristics were deemed to be

influential on post-transplant survival in pediatric patients. The report states the most important donor characteristic to be echocardiographic measurement of ejection fraction. When normal, the most favorable state for a donor heart, most donor characteristics become irrelevant and do not need to be considered. Additionally, ischemic time should be considered and times under 6 hours are advised.

The objective of this study is to determine the validity of the statements made in the 2020 ISHLT consensus paper utilizing the modeling techniques carried out in the various pediatric post-transplant survival studies. We hypothesis that based on current donor practices, the only variables that will have an impact on post-transplant survival are echocardiograph measurements and ischemic time. No additional characteristics of the donor heart will have an impact on the patients outcome post-transplantation, and the focus when making the decision to transplant should be primarily based on the characteristics of the patient and not the donor.

## Methods

### *Data Collection*

The data used in this study comes from the UNOS registry data base. This data base contains information regarding each candidate registered for a transplant and their waitlist information. In the event that a transplant occurs, information on pre-transplant, post-transplant and the donor is stored. For the purpose of this study, pediatric recipients, aged <18, receiving a heart transplant between the years of 2010 and 2019 were used. Those undergoing re-transplantation or on the waitlist for multiple organ transplants were removed. Additionally, recipients with donors over the age of 30 were excluded. Finally, those missing donor match run information were removed resulting in a study population size of 3882 patients.

### *Censoring*

Several patients in the final data set were censored due to an out of date follow-up information. Patients lost to follow-up prior to the survival window were removed from the model. This removed 55 patients from the 1-year models, 632 from the 3-year models, and 1365 from the 5-year model.

### *Variable Selection*

Both candidate and donor related variables were used for this study. The process of variable selection included literature review, expert opinion and variable importance scoring with penalized lasso logistic regression and random forest. Variables were chosen based on the frequency in which they appeared across previous modeling attempts and their importance when modeling 1,3 and 5-year survival using penalized lasso regression and random forest (Miller, 2019; Killian, 2023; Ashfaq, 2022). Ischemic time and echocardiograph measurement variables were added due to their importance on post-transplant outcomes (Kirk, 2020). The expert opinion of the doctors on our team lead to the inclusion of additional donor and candidate related variables. The final set of predictors used across the models in this study consisted of 50 variables in total.

For the purpose of our study, four different variable groups were used. The first group is the candidate variables, including 20 variables pertaining to the health and demographics of the patient receiving the heart, as well as characteristics of the offer. The second and third group are ischemic time and the 11 echocardiograph measurement variables. Finally, the donor group contains 18 variables with health and demographic data pertaining to the donor and the donor heart and donor-recipient characteristics (Appendix B).

### *Experiment Design*

To test our hypotheses, the four variable groups outlined above were used in the formation of four data testing sets; 1. candidate, 2. candidate and ischemic time, 3. candidate and echocardiograph measurements, and 4. candidate, echocardiograph measurements and all other donor variables.

Two different models were used to determine the prediction power of the variables. The Random Forest (RF) model is a tree-based ensemble model that creates many decision trees through randomization and aggregates the results into a single output.[12]. Random Forest has the capacity to handle interactions between variables and non-linear relationships (Rigatti, 2017). Additionally, RF was used due to its prevalence in previous pediatric heart transplant survival studies as the best performing model. The lasso logistic regression (LR) model is a form of penalized regression that shrinks variable coefficients towards zero. By shrinking coefficients to zero, the LR model is performing variable selection, ultimately producing a model that is easy to interpret (James, 2021). 10-fold cross validation and a tuning grid of size 50 were used to select the optimal metrics, mtry in the random forest models and the penalty lambda in the lasso regression models. In addition to the tuned metrics, the random forest models were run with the number of trees set to 1000 and a minimum node size of 1. The values for AUC, brier score, and minimum log loss were recorded for the models with the optimal tuning parameters for each of the ten folds.

Both models were built with each of the four predictor variable sets to evaluate the primary outcome of 1-year survival post-transplant. Two additional secondary outcomes were tested, 3-year and 5-year post transplant survival.

To determine the impact of the different variable groups on model success, the model performance metrics were compared for each survival period. The 10 AUC scores, brier scores, and log loss scores from each fold for the different variable groups were plotted with confidence intervals to determine if there are any significant differences in performance.

## Results

A total of 3882 patients were used in the study. The entire set of patients were used to test and build the model, with performance and predictions coming from each of the 10-cross validated folds. Patients with unknown outcomes for each survival window were censored, with a number of patients being removed due to being lost to follow-up prior to the survival window. Of the remaining patients, the overall mortality rate is 7.5% for 1-year post transplant, 14.2% for 3-year post transplant, and 23.4% for 5-year post transplant.

**Table 1:** Cohort patient characteristics

| Recipient | Overall (N = 3882) | Alive or LTF (N = 3292) | Deceased (N = 590) |
|---|---|---|---|
| Age (years), mean (SD) | 6.24 (6.16) | 6.34 (6.14) | 5.63 (6.25) |
| Gender (female), n (%) | 1733 (44.64) | 1464 (44.47) | 269 (45.59) |
| Race, n (%) | | | |
| Asian | 153 (3.94) | 136 (4.13) | 17 (2.88) |
| Black | 751 (19.35) | 619 (18.80) | 132 (22.37) |
| Hispanic | 796 (20.50) | 686 (20.84) | 110 (18.64) |
| White | 2062 (53.12) | 1758 (53.40) | 304 (51.53) |
| Other | 120 (3.09) | 93 (2.83) | 27 (4.56) |
| BMI, mean (SD) | 17.98 (4.65) | 17.95 (4.62) | 18.17 (4.82) |
| Weight, mean (SD) | 26.41 (24.22) | 26.79 (24.18) | 24.30 (24.35) |
| Height, mean (SD) | 107.31 (42.37) | 108.43 (42.32) | 101.07 (42.13) |
| TBILI, mean (SD) | 1.06 (2.35) | .97 (1.63) | 1.58 (4.61) |
| Creatinine, mean (SD) | .50 (.74) | .50 (.68) | .52 (.98) |
| Status, n (%) | | | |
| Status 1A | 3321 (85.55) | 2870 (87.18) | 514 (87.12) |
| Status 1B | 436 (11.23) | 381 (11.57) | 55 (9.32) |
| Status 2 | 125 (3.22) | 104 (3.16) | 21 (3.56) |
| Days on waitlist, mean (SD) | 121.84 (232.36) | 122.07 (236.39) | 120.52 (208.65) |
| Primary Diagnosis, n (%) | | | |
| Cardiomyopathy | 1940 (49.97) | 1738 (52.79) | 202 (34.24) |
| Congenital Heart Disease | 1878 (48.38) | 1496 (45.44) | 382 (64.75) |
| Other | 64 (1.65) | 58 (1.76) | 6 (1.02) |

### *Ischemic Time*

Ischemic time and echocardiograph measures were said to be important in post-transplant outcomes according to the ISHLT consensus statement. A further assessment of these values within our data set can be seen in Table 2. The average ischemic time is roughly 3 hours, reaching as high as 11 hours.

**Table 2:** Distribution of ischemic time

| Ischemic Time (hours) | Minimum | Mean | Maximum |
|---|---|---|---|
| Overall (N=3882) | .18 | 3.60 | 11.35 |
| Alive or LTF (N = 3292) | .18 | 3.63 | 11.35 |
| Deceased (N=590) | .633 | 3.74 | 8.65 |

A univariate analysis between ischemic time and post-transplant survival in our population displays a potential trend in ischemic time and survival. Ischemic times were plotted against 1-, 3-, and 5-year survival using the generalized additive model smoothing technique to capture the impact of ischemic time on survival, shown in Figure 1. A consistent downward slope indicates that longer ischemic times may decrease chances of a post-transplant outcome, as stated in the ISHLT consensus. There appears to be a potential impact on 1-, 3- and 5-year survival, with 1-year survival having the most prominent relationship. Additionally, we can see that the majority of the ischemic times present in our data set are within 1 and 6 hours. There are few observations that fall outside of the recommended 6-hour window, making it difficult to assess what is happening around the tails, contributing to wide confidence intervals.

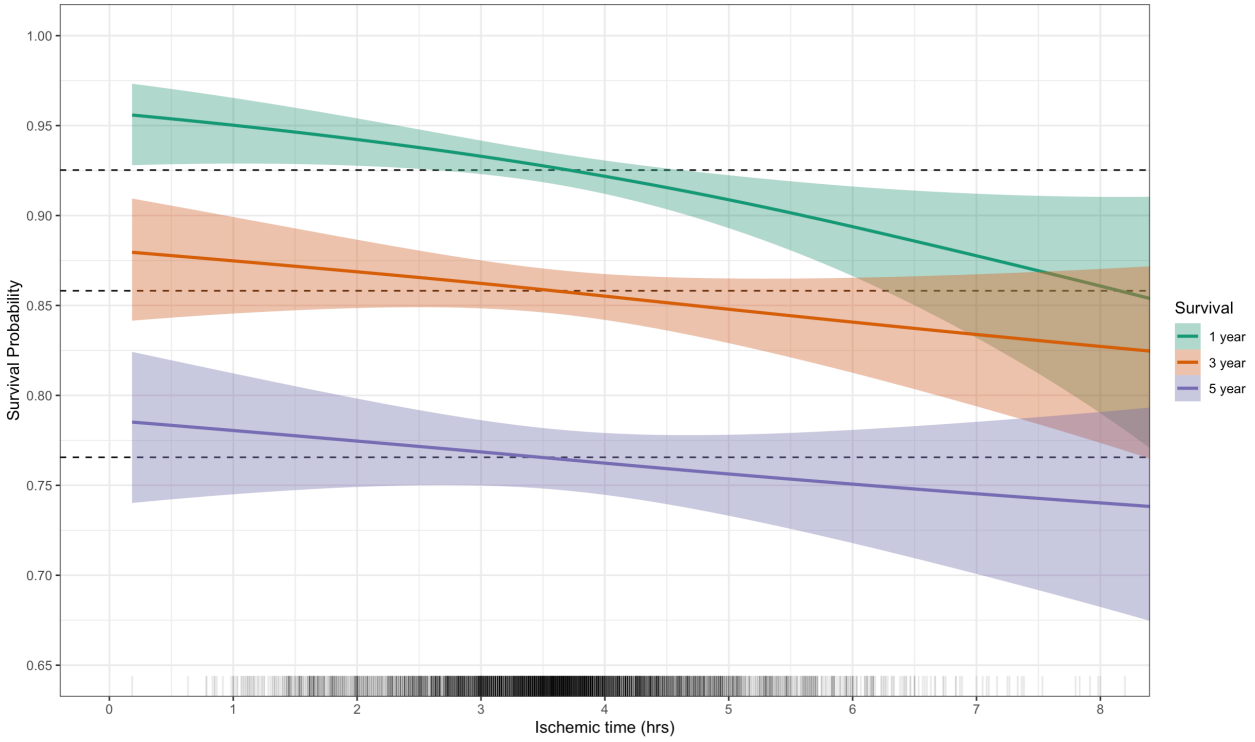


**Figure 1**: Ischemic time vs survival probability for 1-, 3-, and 5-year with shaded CI

### *Echocardiograph measurements*

There were 11 echocardiograph measures included in the echo variable set, with their distributions and completeness shown in Table 3. The primary variable of concern is ABNL_CUM, a cumulative measure that indicates if the echo measurement is abnormal. The declaration of abnormality comes from the assessment of the additional echo measurements. Within our data set, 357 sets of echo readings were declared abnormal in at least one measurement, roughly 9% of the observations.

A univariate analysis was performed on the cumulative echo assessment variable against the survival for 1-, 3-, and 5-years post-transplant, shown in Figure 2. In the case of 1-year survival, there is a slight decrease in survival as you move from the normal to abnormal echo groups. There is, however, uncertainty around the survival probability for abnormal echos, due to the limited number in our data set, which increases with each time frame.

**Table 3**: Distribution of echocardiograph measurements

| Echo Measure | Overall (N = 3882) | Alive or LTF (N = 3292) | Deceased (N = 590) |
|---|---|---|---|
| ABNL_CUM (1), n (%) | 357 (9.1) | 305 (9.3) | 52 (8.8) |
| *% missing* | *4.1* | *3.9* | *4.7* |
| LVSWMA (1), n (%) | 88 (2.3) | 75 (2.3) | 13 (2.2) |
| *% missing* | *90.8* | *90.5* | *92.2* |
| OBJ (1), n (%) | 229 (5.9) | 194 (5.9) | 35 (5.9) |
| *% missing* | *19.7* | *19.5* | *20.5* |
| OBJ_SCALE, mean | .0834 | .0834 | .0832 |
| *% missing* | *19.7* | *19.5* | *20.5* |
| VALVE_FXN (1), n (%) | 36 (.9) | 28 (.9) | 8 (1.4) |
| *% missing* | *4.3* | *4.2* | *4.9* |
| QUAL (1), n (%) | 10 (.3) | 10 (.3) | 0 (0) |
| *% missing* | *4.7* | *4.6* | *5.1* |
| GLOBAL (1), n (%) | 102 (2.6) | 89 (2.7) | 13 (2.2) |
| *% missing* | *4.6* | *4.5* | *5.1* |
| Biplane ef, mean | 62.83 | 62.84 | 62.82 |
| *% missing* | *87.4* | *87.1* | *89* |
| Qual ef, mean | 61.83 | 61.84 | 61.73 |
| *%missing* | *82.2* | *82.1* | *82.7* |
| Four chamber ef, mean | 63.14 | 63.25 | 62.5 |
| % missing | *71.4* | *71* | *73.9* |
| LV_EJECT_DON, mean | 64.25 | 64.54 | 64.43 |
| % missing | 2.6 | 2.5 | 3.2 |

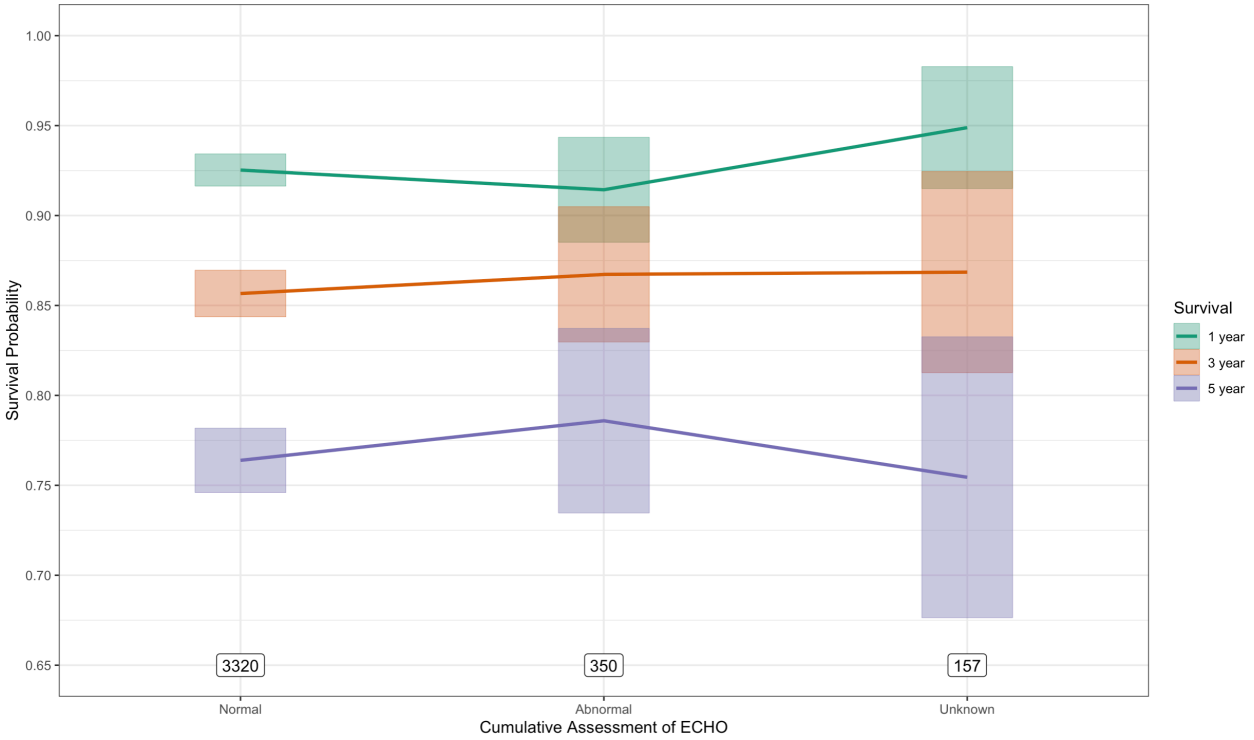


**Figure 2**: Cumulative assessment of echos vs survival probability for 1-, 3-, and 5-year. Counts displayed above each echo type and shading represents the CI around the survival probabilities at each echo type.

### *Predictive performance*

The results of the random forest and lasso logistic regression models for 1-, 3-, and 5-years are shown in Table 4 and 5. The Lasso regression models based on average AUC from the 10-fold cross validation performed best for 1-year survival, with the highest performing model coming from the candidate and donor variable model. The random forest models yielded higher average AUC values for 3- and 5-year survival with the candidate only model performing best for both years. In addition to the four variable group models, donor only models were made using the donor related variables to compare the performance capability of only donor variables with the performance capability of only candidate variables. Across both random forest and lasso regression for all years, the worst performing models were those with only donor characteristics. The brier scores and log loss scores confirm these results.

**Table 4**: Average AUC value random forest model for 1-, 3-, and 5-year survival

| Year | Candidate | Candidate and Ischemic | Candidate and Echo | Candidate and donor | Donor Only |
|---|---|---|---|---|---|
| 1 year | .728 | .724 | .724 | .720 | .636 |
| 3 year | .691 | .689 | .675 | .674 | .605 |
| 5 year | .743 | .738 | .736 | .726 | .597 |

**Table 5**: Average AUC value lasso logistic regression model for 1, 3, and 5-year survival

| Year | Candidate | Candidate and Ischemic | Candidate and Echo | Candidate and donor | Donor Only |
|---|---|---|---|---|---|
| 1 year | .748 | .746 | .748 | .754 | .630 |
| 3 year | .684 | .685 | .681 | .685 | .595 |
| 5 year | .707 | .708 | .703 | .704 | .591 |

### *Hypothesis test results*

The values for AUC, brier score, and log loss across the 10-folds for each model type were compared. The results from random forest are shown in figures 3-5 and lasso regression in figures 6-8, with T0 representing the candidate variable model, T1 representing the candidate and ischemic time variable model, T2 representing the candidate and echocardiograph measurements, and T3 representing the candidate and donor variables model. The T0 model represents the baseline, with an average AUC over the 10-folds of (year 1: [.73, .75] , year 3: [.69, .68], year 5: [.74, .71]), an average brier score of ([.07, .07], [.11, .12], [.14, .15]), and an average log loss of ([.24, .24], [.38, .38], [.45, .48]) in the random forest models and lasso regression models.

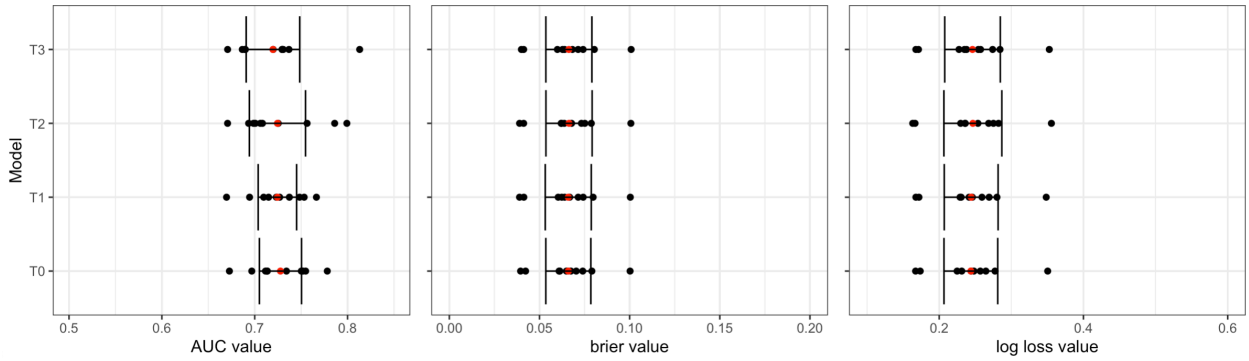


**Figure 3**: Year 1 survival random forest model performance from each fold

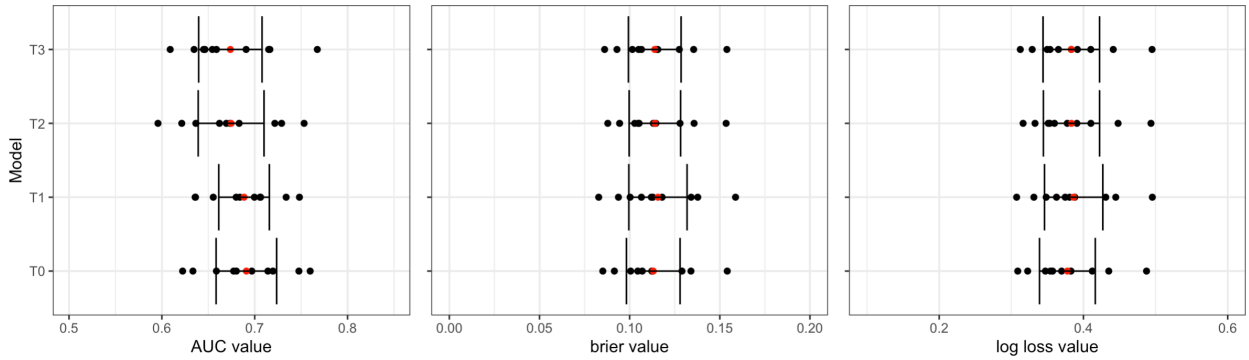


**Figure 4**: Year 3 survival random forest model performance from each fold

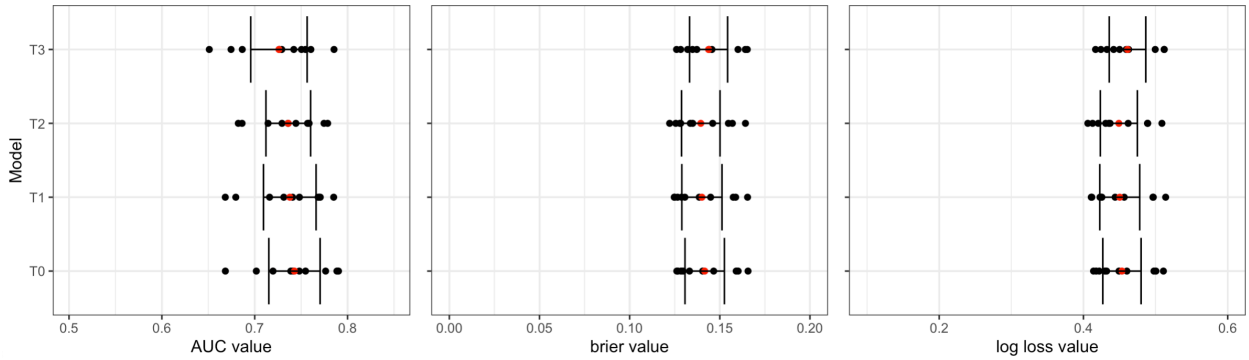


**Figure 5**: Year 5 survival random forest model performance from each fold

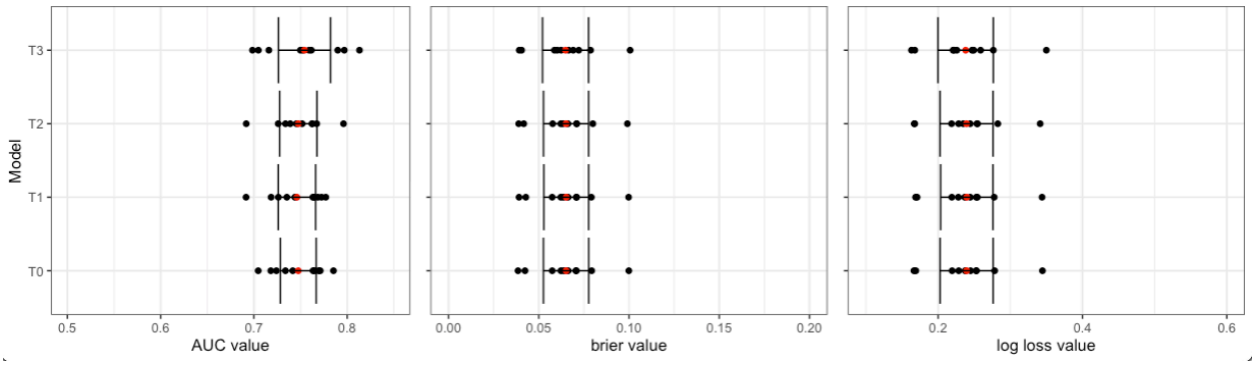


**Figure 6**: Year 1 survival lasso logistic regression model performance from each fold

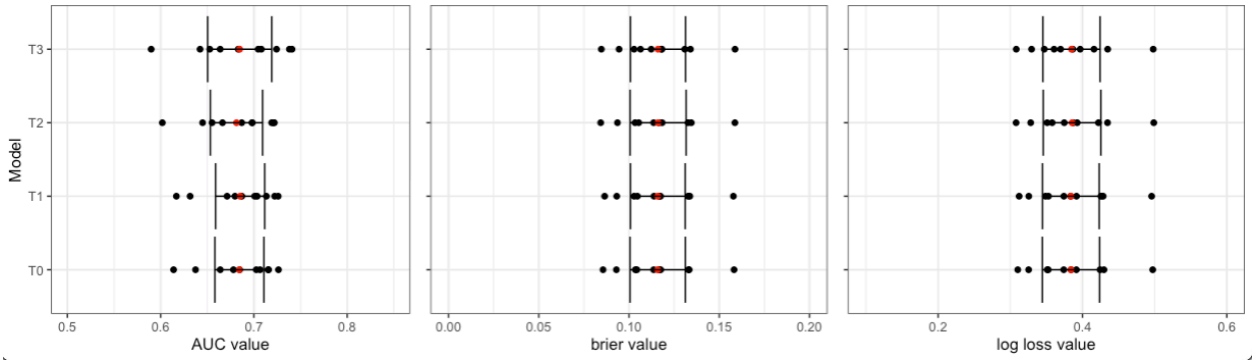


**Figure 7**: Year 3 survival lasso logistic regression model performance from each fold

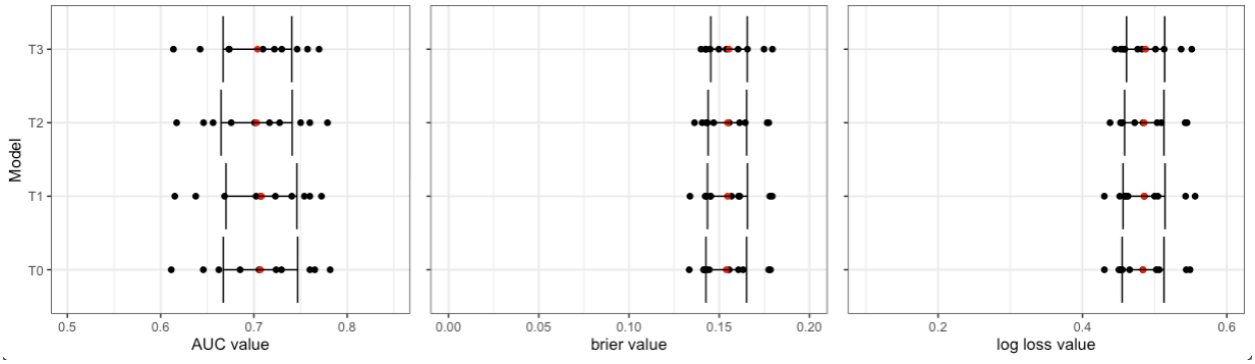


**Figure 8**: Year 5 survival lasso logistic regression model performance from each fold

Both the random forest and lasso regression tables for years 1, 3, and 5 display a lack of effect in the additional variables over the candidate only model. There is no significant difference in the model performance across model types and the corresponding confidence intervals all overlap.

### *Further assessment of ischemic time*

Despite the trends displayed in figure 1, ischemic time did not lead to an improvement in the model performance. Ischemic time is often considered an important feature in post-transplant survival (Ford, 2011; Magdo, 2017; Zafar, 2017; Fraser, 2019; Singh, 2019; Kirk, 2020). It's prevalence across studies indicating its impact lead to further analysis into its lack of impact on model performance. Additional tests were performed on the ischemic time variable used in the models to determine the potential explanation for the lack of performance change. To assess the model's ability to capture survival without the presence of ischemic time, the residuals from the candidate only model predictions were plotted against ischemic time for 1-, 3-, and 5-year survival. Figure 9 lacks any trend in ischemic time for year 3 and 5 random forest and lasso

regression models indicating that the model is not lacking due to the absence of ischemic time. There is a slight downward trend in 1-year survival, however, the confidence interval contains zero.

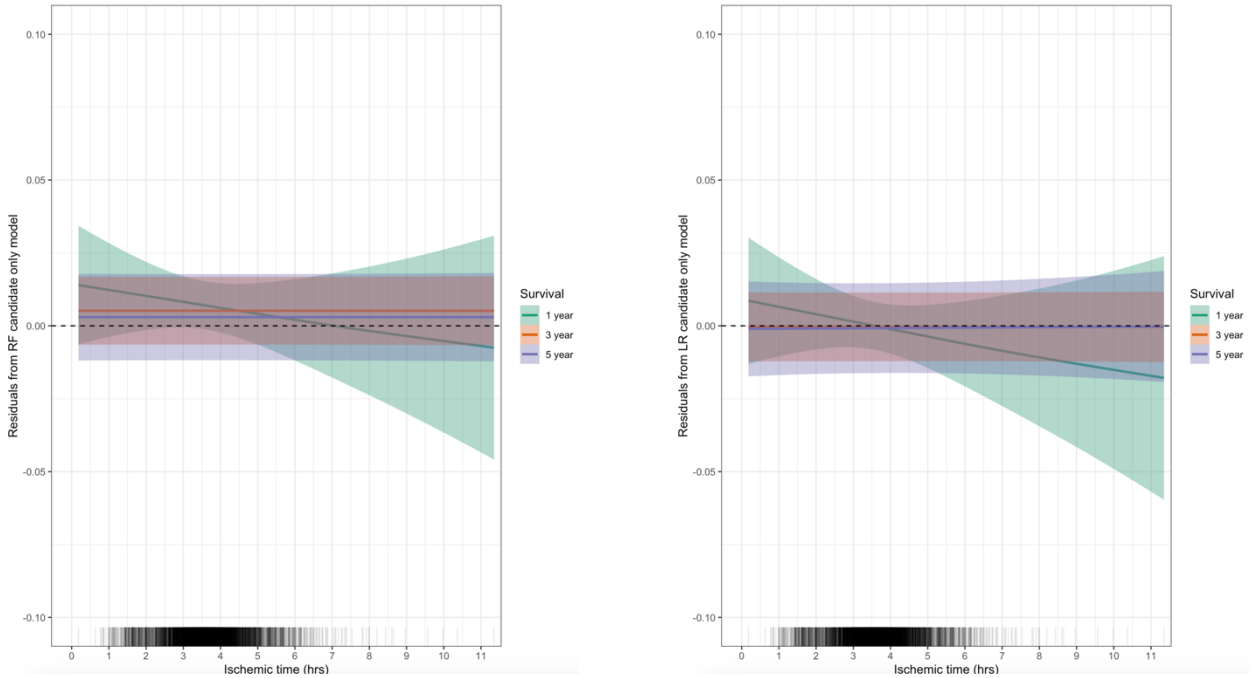


**Figure 9**: Ischemic time vs residuals from candidate only model predictions coming from random forest (left) and lasso logistic regression (right)

Ischemic time may be having little impact due to the candidate variables ability to accurately represent ischemic times impact on survival. Univariate analysis was performed on all candidate variables against ischemic time. Several numeric and categorical variables displayed potential relationships with ischemic time. The most prominent are the size related candidate metrics, with potential non-linear relationships between ischemic time and candidate BMI, height, and weight (figure 10). Additionally, there may be an association between the candidate diagnosis and ischemic time, with congenital heart disease having longer times (figure 11).

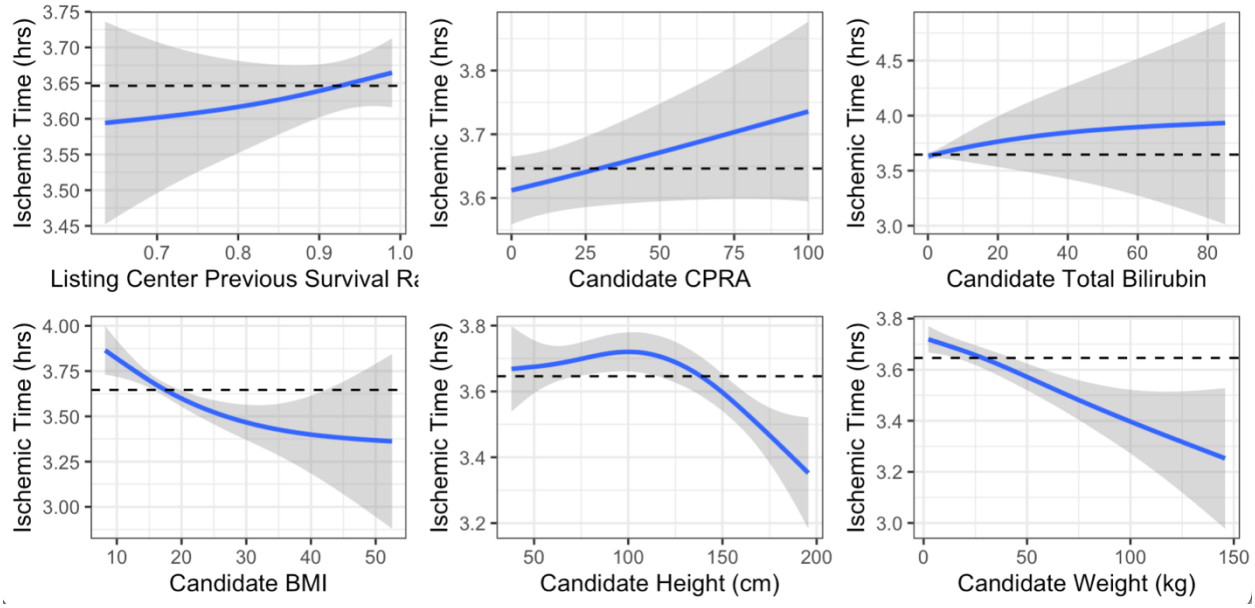


**Figure 10**: Univariate analysis of continuous candidate variables vs ischemic time

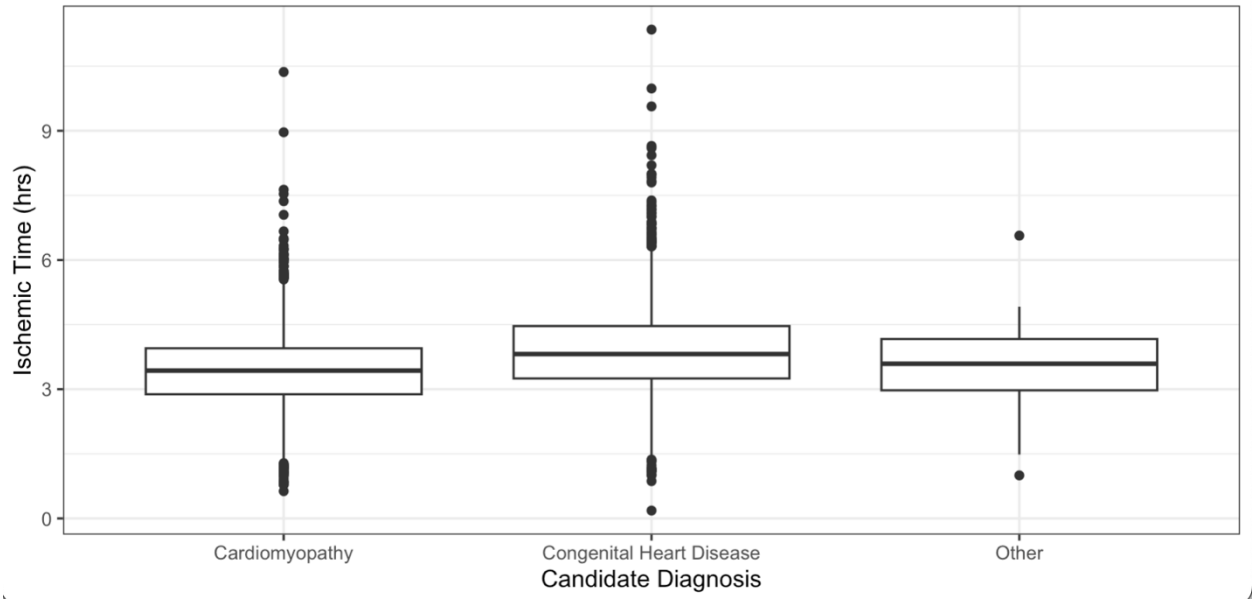


**Figure 11**: Candidate diagnosis vs ischemic time

Two models, a linear regression model and a random forest model, were constructed to determine the predictive strength of candidate variables for ischemic time. Fitting a simple linear model with all candidate variables as predictors and ischemic time as the response yields a low adjusted R squared of .05. These results indicate a poor ability to accurately predict ischemic time. The variables contributing most to the modeling of ischemic time are shown in figure 12.

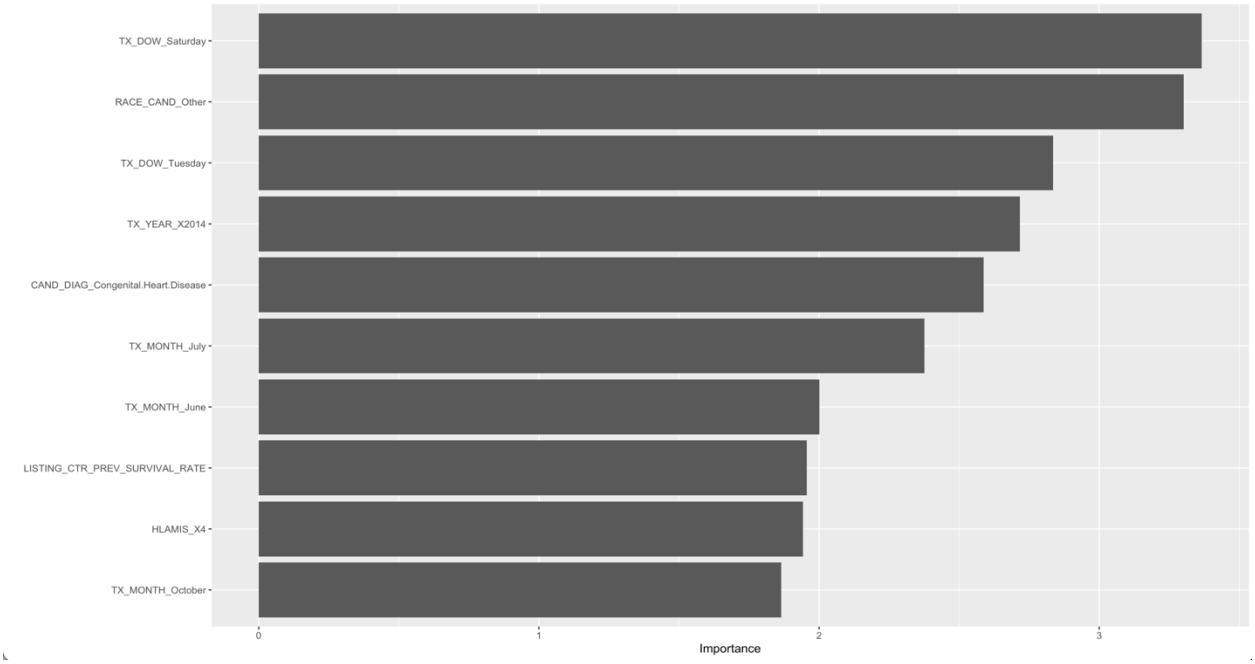


**Figure 12**: Variable importance from linear regression model of ischemic time

Due to the non-linear relationships seen in univariate analysis a 10-fold cross validated random forest model was implemented. Model performance remained poor with an average RMSE around 1 and MAE of .78 for the 10 folds. The variable importance from the random forest model indicated that several of the variables found through univariate analysis to be important, with weight, BMI, and height as the top 3 predictors (figure 13). While model performance is weak, the relationships highlighted through univariate analysis and variable importance may explain the lack of model improvement with the addition of ischemic time.

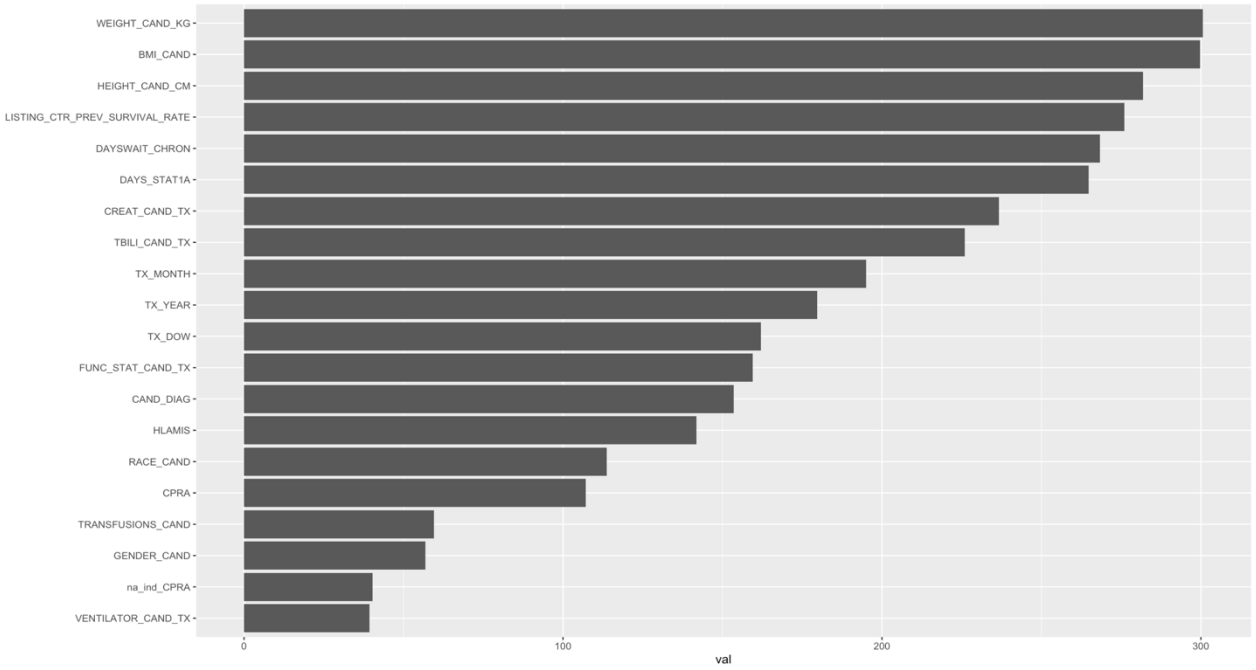


**Figure 13**: Variable importance from random forest model of ischemic time

## Discussion

The performance achieved in the modeling efforts of post-transplant survival is consistent with existing work. Similar to Miller et. al (2019) and Killian et. al. (2023) model performance based on AUC varied from over .6, in the donor only models to roughly .75, in the LR candidate and donor model. With the highest performing models considered acceptable (.7-.8), not excellent (.8-.9) or outstanding (>.9) according to Mandrekar (2010), there is room for improvement. Aside from model performance, the hypothesis findings are useful. Our study attempted to address the necessity of the various donor variables in the donor heart selection process. Through the implementation of two widely used modeling techniques, over the standard survival windows, we were able to assess the impact of non-candidate related variables on model performance. Our results show that the candidate model could not be improved upon with the addition of other variables.

While hypothesized to increases the performance, the addition of ischemic time does not add statistical improvement to the models, it is difficult to state that ischemic times are not important for outcomes. The initial univariate analysis displayed negative trends in longer ischemic times and survival. Within our data set, there was a relatively small window for ischemic time, with less than 3% over the 6-hour mark. The lack of representation of longer ischemic times could make it difficult for our model to accurately capture the effect on survival. As the majority of times within our population were between 1 and 6 hours, within the acceptable window, there may not have been enough variability for it to impact the ability to predict survival post-transplant. Despite statistical insignificance of the addition of ischemic time, it may not be advisable to deviate from the current acceptance window without further assessment into the relationship between increased times and survival.

Similarly, echocardiograph measurements, hypothesized to be impactful, did not appear to be important. Beginning with univariate analysis, there was a slight negative trend in the probability of 1-year survival in the case of negative echos. However, in modeling, there was little impact across all years when adding echo related variables to the candidate model. While this could be an indicator of their impact on post-transplant survival, it is worth noting the limited variability present across echo variables. Many were highly missing, requiring imputation. The cumulative

echo measurement that determines whether or not an echo is normal or abnormal, the primary variable of interest, was well documented with roughly 96% completion. Despite this, the variability is still low with a small portion of abnormal echo's at only 9.1%, making it difficult to fully understand the impact. The cumulative echo variable used in the study came from an expert assessment based on the echo variables to determine the overall normality of the heart (McCulloch, 2024). Ultimately, this determination of abnormality comes from the doctor prior to acceptance of the heart and can differ based on institution. Due to the limited impact shown throughout the models, further research into the best echocardiograph measurements and classifications should be done to better understand what readings impact post-transplant survival.

The lack of variability found in the echocardiograph measures and ischemic time provides a data set of primarily normal echos and acceptable ischemic times. We can confirm the validity of the ISHLT consensus statement that when ischemic time is within an acceptable range, and echos are normal, no other donor characteristics will have an impact. Across all years and metric evaluations, the addition of donor characteristics to the candidate model did not lead to a statistically significant improvement in the model. All other donor characteristics do not have an impact on post-transplant survival when ischemic time is less than 6 hours and echos suggest normal heart function.

It should be noted that these findings come from a small and incomplete data set. Information regarding various donor aspects that is missing could limit the model's ability to accurately capture factors contributing to post-transplant success. Additionally, the small population size used, especially when assessing 3- and 5-year survival, potentially lead to the inability to capture all relationships among the variables. The current practices in donor heart acceptance for pediatric transplant candidates leave little variability in the characteristics of the donor hearts. Therefore, the statement that all other donor heart characteristics do not impact the ability to predict post-transplant outcomes is based on the assumption that these acceptance practices do not change drastically.

## References


1. Khush, K. K., Zaroff, J. G., Nguyen, J., Menza, R., & Goldstein, B. A. (2015). National Decline in Donor Heart Utilization With Regional Variability: 1995-2010. *American Journal of Transplantation*, *15*(3), 642–649. https://doi.org/10.1111/ajt.13055
2. Colvin, M. M., Smith, J. M., Ahn, Y. S., Messick, E., Lindblad, K., Israni, A. K., Snyder, J. J., & Kasiske, B. L. (2021). OPTN/SRTR 2021 Annual Data Report: Heart. *Scientific Registry of Transplant Recipients*. https://srtr.transplant.hrsa.gov/annual_reports/2021/Heart.aspx#tab:HRactivity-ped-waiting
3. Kirk, R., Dipchand, A. I., Davies, R. R., Miera, O., Chapman, G., Conway, J., Denfield, S., Gossett, J. G., Johnson, J., McCulloch, M., Schweiger, M., Zimpfer, D., Ablonczy, L., Adachi, I., Albert, D., Alexander, P., Amdani, S., Amodeo, A., Azeka, E., Ballweg, J., … Zuckerman, W. A. (2020). ISHLT consensus statement on donor organ acceptability and management in pediatric heart transplantation. *The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation*, *39*(4), 331–341. https://doi.org/10.1016/j.healun.2020.01.1345
4. Miller, R., Tumin, D., Cooper, J., Hayes, D., Jr, & Tobias, J. D. (2019). Prediction of mortality following pediatric heart transplant using machine learning algorithms. *Pediatric transplantation*, *23*(3), e13360. https://doi.org/10.1111/petr.13360
5. Killian, M. O., Tian, S., Xing, A., Hughes, D., Gupta, D., Wang, X., & He, Z. (2023). Prediction of Outcomes After Heart Transplantation in Pediatric Patients Using National Registry Data: Evaluation of Machine Learning Approaches. *JMIR cardio*, *7*, e45352. https://doi.org/10.2196/45352
6. Conway, J., Ballweg, J. A., Fenton, M., Kindel, S., Chrisant, M., Weintraub, R. G., Danziger-Isakov, L., Kirk, R., Miera, O., Davies, R. R., & Dipchand, A. I. (2020). Review of the impact of donor characteristics on pediatric heart transplant outcomes. *Pediatric Transplantation*, *24*(3), e13680. https://doi.org/10.1111/petr.13680
7. Zafar, F., Jaquiss, R. D., Almond, C. S., Lorts, A., Chin, C., Rizwan, R., Bryant, R., Tweddell, J. S., & Morales, D. L. S. (2018). Pediatric Heart Donor Assessment Tool (PHDAT): A novel donor risk scoring system to predict 1-year mortality in pediatric heart

transplantation. *The Journal of Heart and Lung Transplantation*, *37*(3), 332–339. https://doi.org/10.1016/j.healun.2017.03.002

8. Fraser, C. D., 3rd, Grimm, J. C., Zhou, X., Lui, C., Giuliano, K., Suarez-Pierre, A., Crawford, T. C., Magruder, J. T., Hibino, N., & Vricella, L. A. (2019). Children's Heart Assessment Tool for Transplantation (CHAT) Score: A Novel Risk Score Predicts Survival After Pediatric Heart Transplantation. *World journal for pediatric & congenital heart surgery*, *10*(3), 296–303. https://doi.org/10.1177/2150135119830089
9. Feingold, B. (2018). The challenges of donor-derived risk, donor shortage and waitlist mortality in children: Time for a new measuring stick? *The Journal of Heart and Lung Transplantation*, *37*(3), 317–318. https://doi.org/10.1016/j.healun.2017.05.002
10. Rizwan, R., Zafar, F., Bryant, R., Tweddell, J. S., Lorts, A., Chin, C., & Morales, D. L. (2018). The Number of Refusals for Donor Organ Quality Does Not Impact Heart Transplant Outcomes in Children. *The Annals of Thoracic Surgery*, *105*(4), 1223–1230. https://doi.org/10.1016/j.athoracsur.2017.10.006
11. Singh, N., Raees, M. A., & Zafar, F. (2019). Donor considerations in pediatric heart transplantation. *Translational Pediatrics, 8(4)*, 284–289. https://doi.org/10.21037/tp.2019.08.02
12. Rigatti, S. J. (2017). Random Forest. *Journal of Insurance Medicine*, *47*(1), 31–39. https://doi.org/10.17849/insm-47-01-31-39.1
13. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning: With Applications in R. *Springer US*. https://doi.org/10.1007/978-1-0716-1418-1
14. Butts, R. J., Baez Hernandez, N., Kirk, R., Bano, M., & Davies, R. (2021). Center Donor Refusal Rate Is Associated With Worse Outcomes After Listing in Pediatric Heart Transplantation. *Transplantation, 105(9)*, 2080–2085. https://doi.org/10.1097/TP.0000000000003514
15. Denfield, S. W., Azeka, E., Das, B., Garcia-Guereta, L., Irving, C., Kemna, M., Reinhardt, Z., Thul, J., Dipchand, A. I., Kirk, R., Davies, R. R., & Miera, O. (2020). Pediatric cardiac waitlist mortality—Still too high. *Pediatric Transplantation, 24(3), e13671*. https://doi.org/10.1111/petr.13671
16. Miller, R. J. H., Sabovčik, F., Cauwenberghs, N., Vens, C., Khush, K. K., Heidenreich, P. A., Haddad, F., & Kuznetsova, T. (2022). Temporal shift and predictive performance of machine

learning for heart transplant outcomes. *The Journal of Heart and Lung Transplantation*, *41*(7), 928–936. https://doi.org/10.1016/j.healun.2022.03.019

17. Ashfaq, A., Gray, G. M., Carapellucci, J., Amankwah, E. K., Ahumada, L. M., Rehman, M., Quintessenza, J. A., & Asante-Korang, A. (2022). Predicting One Year Mortality Using Machine Learning After Pediatric Heart Transplantation: Analysis of the United Network of Organ Sharing (UNOS) Database. *The Journal of Heart and Lung Transplantation*, *41*(4, Supplement), S152. https://doi.org/10.1016/j.healun.2022.01.360
18. Mandrekar, J. N. (2010). Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *Journal of Thoracic Oncology*, *5*(9), 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d
19. Ford, M. A., Almond, C. S., Gauvreau, K., Piercey, G., Blume, E. D., Smoot, L. B., Fynn-Thompson, F., & Singh, T. P. (2011). Association of graft ischemic time with survival after heart transplant among children in the United States. *The Journal of Heart and Lung Transplantation*, *30*(11), 1244–1249. https://doi.org/10.1016/j.healun.2011.05.001
20. Magdo, H. S., Friedland-Little, J. M., Yu, S., Gajarski, R. J., & Schumacher, K. R. (2017). The impact of ischemic time on early rejection after pediatric heart transplant. *Pediatric Transplantation*, *21*(8), e13034. https://doi.org/10.1111/petr.13034
21. McCulloch, M. A., Alonzi, L. P., White, S. C., Haregu, F., & Porter, M. D. (2024). Pediatric donor heart acceptance practices in the United States: What is really being considered?. *Pediatric transplantation*, *28*(1), e14649. https://doi.org/10.1111/petr.14649

# Appendix A: Models from literature review

| Author and Year | Models | Best Metric | Top 5 features |
|---|---|---|---|
| Ashfaq (2022) | CPH, C-EN, GB, SVM, RF, CGB, ST<br><br>1-year survival | .68 (C-index; RF) | Serum bilirubin, distance, BMI, D terminal SGPT/ALT, D PCO2 |
| Miller (2019) | ANN, CART, RF<br><br>1, 3, 5-year survival | .74 (AUC; RF) | Diagnosis, ecmo AT mech vent AT, D Gender, D B1 antigen levels |
| Killian (2023) | XGB, LR, SVM, RF, SGD, MLP, AdaBoost, NN<br><br>1, 3, 5-year survival and rejection | .697 (AUC; RF) | Grf stat, malig, vad device type, days status 1, height ratio |
| Miller (2022) | RF, XGB, L2-LR, L2-Cox, SGB, RSF<br><br>90-day, 1-year survival | .836 (AUC; RF w/ shuffled CV) | N/A |

* CPH, Cox proportional hazard; C-EN, Cox with elastic net; GB, gradient boosting; SVM, support vector machines; RF, random forest; CGB, component gradient boosting; ST, survival trees; ANN, artificial neural network; CART, classification and regression tree; XGB, extreme gradient boosting; LR, logistic regression; SGD, stochastic gradient decent; MLP, multi-layer perceptron; AdaBoost, adaptive boosting; NN, neural network; L2-LR, L2 regularized logistic regression; L2-Cox, L2 regularized Cox regression; SGB, survival gradient boost; RSF, random survival forest

## Appendix B: Variable Groups

| Candidate | Donor | Echo |
|---|---|---|
| GENDER_CAND | WEIGHT_DON_KG | ECHO_ABNL_CUM |
| RACE_CAND | HEIGHT_DON_CM | ECHO_LVSWMA |
| WEIGHT_CAND_KG | CARDARREST_POSTNEURO_DON | ECHO_OBJECTIVE |
| HEIGHT_CAND_CM | DEATH_CIRCUM_DON | ECHO_OBJECTIVE_SCALE |
| BMI_CAND | RISK_HEP_DON | ECHO_VALVE_FXN |
| TX_YEAR | HRS_DECEASED_AT_CLAMP | ECHO_QUAL |
| TX_MONTH | SEPTAL_WALL | ECHO_GLOBAL_VENT_DYSF |
| TX_DOW | ABO_MATCH | ECHO_biplane_eject_frac |
| LISTING_CTR_PREV_SURVIVAL_RATE | GENDER_STR | ECHO_qual_eject_frac |
| ECMO_CAND_TX | BMI_DIFF | ECHO_four_chamber_eject_frac |
| VENTILATOR_CAND_TX | HEIGHT_RATIO | LV_EJECT_DON |
| FUNC_STAT_CAND_TX | CREAT_DON | |
| TRANSFUSIONS_CAND | DA1_DON | |
| CAND_DIAG | DB1_DON | |
| CPRA | PCO2_DON | |
| HLAMIS | TRANSFUS_TERM_DON | |
| DAYS_STAT1A | WEIGHT_RATIO | |
| DAYSWAIT_CHRON | AGE_DIFF | |
| TBILI_CAND_TX | **Ischemic** | |
| CREAT_CAND_TX | ISCHTIME | |

## Appendix C

**Table**: Average brier score random forest model for 1-, 3-, and 5-year survival

| Year | Candidate | Candidate and Ischemic | Candidate and Echo | Candidate and donor | Donor Only |
|---|---|---|---|---|---|
| 1 year | .066 | .066 | .066 | .066 | .069 |
| 3 year | .113 | .116 | .114 | .114 | .120 |
| 5 year | .142 | .140 | .134 | .144 | .176 |

**Table**: Average brier score lasso logistic regression model for 1-, 3-, and 5-year survival

| Year | Candidate | Candidate and Ischemic | Candidate and Echo | Candidate and donor | Donor Only |
|---|---|---|---|---|---|
| 1 year | .065 | .065 | .065 | .065 | .068 |
| 3 year | .116 | .116 | .116 | .116 | .120 |
| 5 year | .154 | .155 | .155 | .155 | .177 |

**Table**: Average log loss random forest model for 1-, 3-, and 5-year survival

| Year | Candidate | Candidate and Ischemic | Candidate and Echo | Candidate and donor | Donor Only |
|---|---|---|---|---|---|
| 1 year | .244 | .244 | .247 | .246 | .261 |
| 3 year | .378 | .386 | .383 | .383 | .401 |
| 5 year | .453 | .450 | .449 | .461 | .534 |

**Table**: Average log loss lasso logistic regression model for 1-, 3-, and 5-year survival

| Year | Candidate | Candidate and Ischemic | Candidate and Echo | Candidate and donor | Donor Only |
|---|---|---|---|---|---|
| 1 year | .239 | .240 | .239 | .238 | .260 |
| 3 year | .384 | .386 | .386 | .385 | .402 |
| 5 year | .484 | .486 | .486 | .487 | .537 |