Prediction of Mortality and Allograft Loss in Pediatric Heart Transplant Patients Using Machine Learning

# METHODS

## Modeling algorithms

We define the term ‘modeling algorithm’ as the collection of steps taken to develop a prediction model. We compared modeling algorithms based on the following characteristics: (1) number of predictors used in the model, (2) method to select the specified number of predictors, (3) method to develop a prediction model using the selected predictors. We considered seven options for the number of predictors, three options for variable selection, and four options for model development (Figure 1). All possible combinations of these options were assessed, leading to the evaluation of 84 unique modeling algorithms.

**Methods to select predictors:**

We assess contribution importance, permutation importance, and stepwise importance. Contribution importance represents fractional contribution of each predictor to the model based on the total gain of this feature’s splits (1). Higher percentage means a more important predictive feature. Permutation importance measures the importance of each predictor by measuring how much the model’s prediction error increases when the values of the predictor are randomly permuted (2). Stepwise importance follows the traditional stepwise technique that includes or excludes variables one at a time until a stopping criteria is met (3). Though it is not a valid procedure for statistical inference, stepwise modeling algorithms can be useful for prediction (4). There is no numeric importance value for stepwise importance but a subset of the most important variables is easily obtained by using a pre-specified number of steps.

**Methods to develop a risk prediction model:**

Proportional hazards regression is a semi-parametric modeling technique that estimates a baseline hazard function and assumes predictors have a multiplicative effect on the hazard for an event of interest. Gradient boosting (hereafter referred to as boosting) develops an ensemble of weak prediction models (1). Each learner in the ensemble attempts to correct errors from the previous learners, and the ensemble’s prediction is the aggregate of its individual learners’ predictions. We developed boosting models using decision trees as learners (5), a technique that has been recognized in numerous settings to be state of the art for statistical prediction (6). Random forests are also ensembles of decision trees, but in contrast to boosting, the random forest comprises de-correlated decision trees that are grown using bootstrapped replicates of the original data (2, 7). We applied the standard random forest algorithm for risk prediction, which grows decision tree nodes by splitting the data based on an individual variable (8). We also applied an extension of the random forest that splits data using linear combinations of variables (9). Random forests grown using linear combinations of input variables are referred to as ‘oblique’ random forests.

## Internal validation

We applied Monte-Carlo cross-validation to identify the most effective modeling algorithm for developing a final prediction model and to estimate the performance of modeling algorithms when they are applied to new data (10). Monte-Carlo cross-validation is an extension of split-sample testing, which splits available data into training and testing sets, then develops a model using the training set and validates the model in the testing set. Monte-Carlo cross-validation replicates this procedure, using a different split of the available data for each replicate, leading to a reduction in the variance of performance estimates for the modeling algorithms assessed (11). In the current analysis, we completed a total of 500 replications of split-sample testing.

## Measures of model performance

Model performance was evaluated based on discrimination and calibration, as recommended by published guidelines (12, 13). Discrimination and calibration were measured using a time-dependent concordance (C-) statistic and modified D’Agostino-Nam test, respectively (14, 15). Point estimates and 95% confidence intervals (CIs) for these metrics were computed empirically using the distribution of results from Monte-Carlo cross-validation. Point estimates were the median value of a performance metric, while the 2.5th and 97.5th percentiles defined lower and upper bounds for 95% CIs, respectively. We did not find previously published prediction models for graft loss among pediatric patients, so we did not analyze net reclassification improvement as it requires a baseline prediction model (16).

## Missing data

Missing values were imputed after data were split into training and testing sets during each replicate of Monte-Carlo cross-validation. All information from the testing data was withheld during this process (17). Specifically, missing values were imputed in both the training and testing data using the mean and mode of continuous and categorical variables, respectively, computed using the training data. While imputation to the mean is not appropriate for statistical inference, this technique has been shown to produce prediction models with Bayes consistency when missing values are non-informative (18).

## Bayesian analysis of model performance

We applied Bayesian hierarchical models to draw inferences regarding the expected performance of modeling algorithms, accounting for correlated performance within each replicate of Monte-Carlo cross-validation (19). Specifically, we estimated the posterior probability that a given technique to fit a risk prediction model had superior discrimination or calibration compared to other techniques. All comparisons were made holding the number of predictor variables and the method to select predictor variables fixed; i.e., both models used the same predictor variables.

## Statistical analysis

We conducted the current analysis following previously published guidelines on multivariable prediction models for individual prognosis and diagnosis (20). The 1 year incidence of graft loss or morality was computed overall and by transplant year. Characteristics of patients were calculated as mean with standard deviation or percent in the overall population and stratified by transplant year. Using the results from internal validation, we visually assessed the discrimination and calibration of candidate modeling algorithms to develop a final prediction model. We used posterior predicted probability to perform inference on pairwise comparisons between candidate modeling algorithms. To make results relevant for clinical settings, we restricted these pairwise comparisons to models that used 20 predictor variables. Based on the results, we selected a final prediction modeling algorithm and applied it to all the available data. We tabulated a summary of the variables included into the final prediction model - including the numeric importance value and rank along with the count and percent of how many missing values were present. We used partial dependence to estimate multivariable adjusted predicted risk as a function of each variable included in the final prediction model, separately (21).

Analyses were conducted using SAS version 9.4, R version 4.0.4, and a number of open-source R packages (22–26). All R code for the current analysis is publicly available at <https://github.com/bcjaeger/graft-loss>. Data for the current analysis are available by request from **FILL IN**.

# RESULTS

Overall, the 1-year incidence of graft loss or mortality after transplant was 0.08 (0.07, 0.09). The incidence was highest during 2014 and lowest during 2016, with incidence rates (95% CI) of 0.10 (0.07, 0.14) and 0.06 (0.04, 0.09), respectively. The overall sample was 55.0% male, 65.1% white, and had mean (standard deviation) age of years at the time of transplant (**Table 1**).

## Internal validation

Internally validated estimates of the C-statistic showed a progressive increase in model discrimination as more predictors were included, except when proportional hazards regression was applied to develop the risk prediction model (**Figure 2**). The oblique random survival forest obtained the highest C-statistic in 14 of 21 comparisons, and obtained the highest overall C-statistic when 15 or more predictor variables were selected, regardless of the variable selection method.

With the exception of boosting, all modeling algorithms obtained adequate calibration (i.e., p-value for mis-calibration ≥ 0.05; **Figure 3**). The oblique random survival forest obtained the highest p-value for mis-calibration in 17 of 21 comparisons. When predictors were selected using permutation importance, the oblique random survival forest obtained the first and second highest median p-values for mis-calibration using 35 and 20 predictor variables, respectively.

**Bayesian analysis of model performance**: As discrimination and calibration were generally higher when permutation importance was applied, we focused our pairwise comparisons to models developed using 20 predictors selected by permutation importance (**Figure 4**). In this setting, oblique random survival forests obtained the highest C-statistic (0.764) and p-value for mis-calibration (0.422). The posterior probability that oblique random survival forests obtained superior discrimination versus other modeling algorithms ranged from 0.69 (versus boosting) to 0.94 (versus proportion hazards). The posterior probability that oblique random survival forests obtained superior calibration versus other modeling algorithms ranged from 0.6 (versus standard random forests) to 0.85 (versus boosting).

**Selection of the final modeling algorithm**: To maintain clinical relevance, we limited the number of predictors in our final model to 20. Under this setting, using permutation importance to select predictors and oblique random survival forests to fit a risk prediction provided the best discrimination (C = 0.76 and calibration (P-value for mis-calibration = 0.39). Therefore, we fit our final prediction model using the oblique random survival forest after selecting 20 predictors using permutation importance.

## Model summary

The three variables with greatest permutation importance were cardiopulmonary bypass time, primary etiology (cardiomyopathy, congenital heart disease, or other), and ECMO at transplant (**Table 2**). Among the variables selected, the percentage of values missing ranged from 0 to 22.7, and the overall percentage of values missing in the training data was 5.5

**Table 1:** Characteristics of patients included in the current analysis.

| **Variable** | **Overall** | **Transplant year** | | |
| --- | --- | --- | --- | --- |
| **Overall (N = 3,787)** | **Before 2014 (N = 1,400)** | **2014 through 2016 (N = 1,286)** | **After 2017 (N = 1,101)** |
| Recipient Age (years) at Transplant1 | 4.88 (0.74, 13.1) | 4.38 (0.67, 12.9) | 4.85 (0.76, 13.0) | 5.49 (0.88, 13.3) |
| F0 Male recipient | 55.0% | 52.9% | 56.1% | 56.6% |
| Race | | | | |
| Black | 17.1% | 18.5% | 15.9% | 16.5% |
| Other | 17.9% | 14.9% | 18.0% | 21.5% |
| White | 65.1% | 66.6% | 66.1% | 61.9% |
| Recipient Hispanic or Latino | 20.1% | 15.4% | 22.0% | 24.2% |
| Cardiopulmonary bypass time (minutes) | 161 (123, 211) | 154 (118, 199) | 163 (126, 214) | 170 (127, 226) |
| Primary Etiology | | | | |
| Cardiomyopathy | 49.6% | 52.0% | 49.8% | 46.2% |
| Congenital heart disease | 48.0% | 44.4% | 48.7% | 51.7% |
| Other | 2.48% | 3.64% | 1.56% | 2.09% |
| ECMO at Transplant | 4.30% | 4.50% | 4.74% | 3.54% |
| Surgeries Prior to Listing | 46.9% | 44.4% | 47.6% | 49.4% |
| CHD: Single Ventricle | 35.1% | 26.9% | 37.7% | 43.2% |
| MSCD at Transplant | 30.3% | 27.9% | 28.5% | 35.5% |
| Transplant Lab: Serum Albumin g/dL | 3.70 (3.20, 4.20) | 3.70 (3.20, 4.20) | 3.70 (3.20, 4.20) | 3.70 (3.20, 4.20) |
| Medical History at Listing | 83.8% | 99.6% | 83.9% | 68.6% |
| Transplant Lab: BUN mg/dL | 16.0 (12.0, 22.0) | 16.0 (11.8, 22.0) | 16.0 (12.0, 22.0) | 17.0 (12.0, 23.0) |
| 1Table values are median (25th percentile, 75th percentile) and percent for continuous and categorical variables, respectively. | | | | |

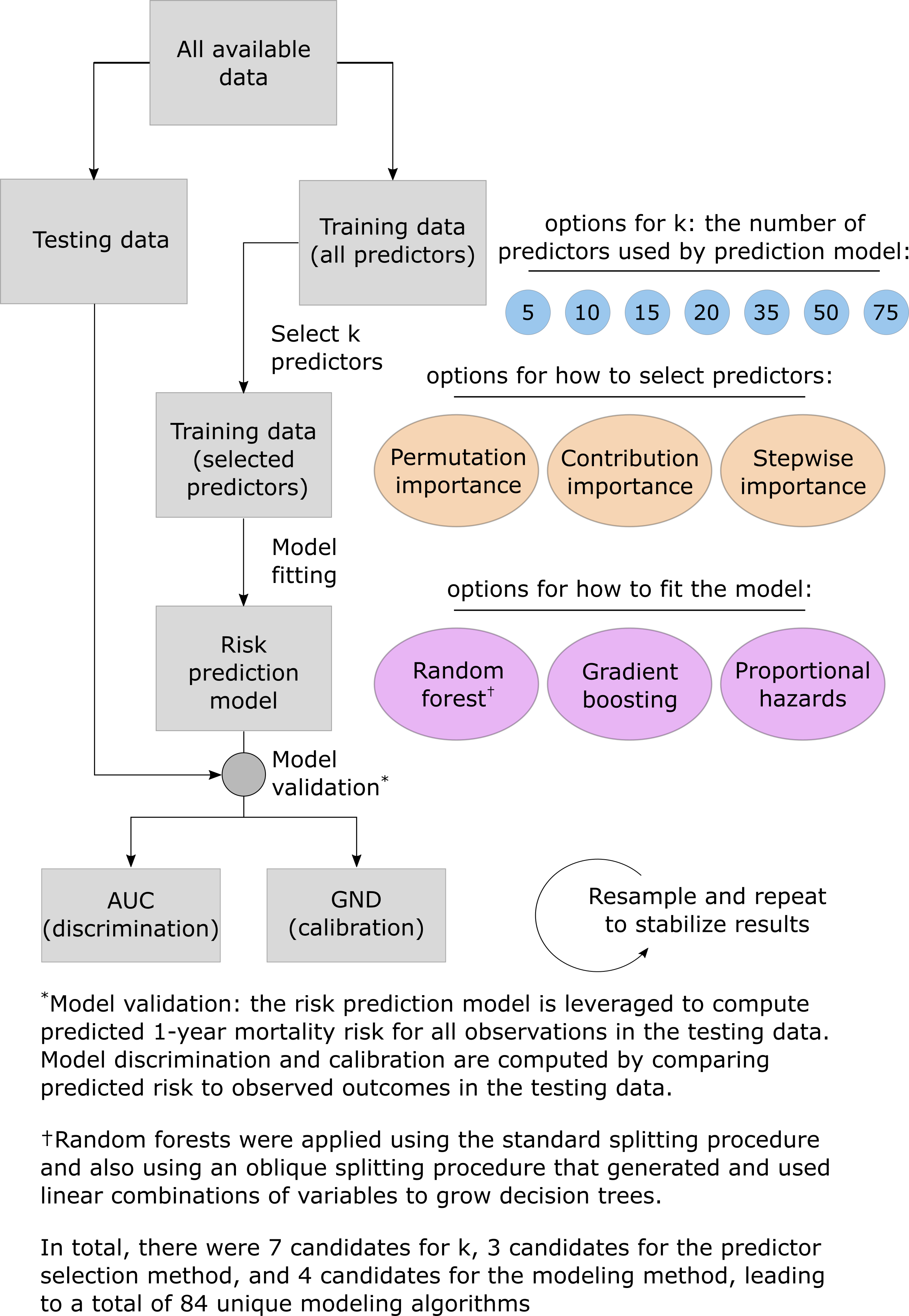
**Table 2:** Variable missingness and importance for predictors used in the final prediction model.

| **Predictor variable** | **Missing values** | | | **Permutation Importance** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Number** | **Percent** | **Imputed to** | **Change in C-statistic** | **Scaled** | **Rank** |
| Cardiopulmonary bypass time (minutes) | 158 | 4.17 | 161 | 0.0097168775 | 1.00000000 | 1 |
| Primary Etiology | 0 | 0.00 | cardiomyopathy | 0.0094045540 | 0.96785763 | 2 |
| ECMO at Transplant | 0 | 0.00 | no | 0.0067196087 | 0.69153992 | 3 |
| Surgeries Prior to Listing | 8 | 0.21 | no | 0.0057533531 | 0.59209897 | 4 |
| CHD: Single Ventricle | 156 | 4.12 | no | 0.0043769877 | 0.45045209 | 5 |
| MSCD at Transplant | 0 | 0.00 | no | 0.0034450069 | 0.35453847 | 6 |
| Transplant Lab: Serum Albumin g/dL | 422 | 11.1 | 3.70 | 0.0028812727 | 0.29652249 | 7 |
| Medical History at Listing | 533 | 14.1 | yes | 0.0022693445 | 0.23354669 | 8 |
| Recipient Age (years) at Transplant | 5 | 0.13 | 4.88 | 0.0019788482 | 0.20365063 | 9 |
| Transplant Lab: BUN mg/dL | 102 | 2.69 | 16.0 | 0.0018293602 | 0.18826626 | 10 |
| PRA Max List | 326 | 8.61 | 0.00 | 0.0016294804 | 0.16769589 | 11 |
| Transplant Lab: Total Protein g/dL | 844 | 22.3 | 6.40 | 0.0015449353 | 0.15899503 | 12 |
| CHD: Hypoplastic Left Heart | 860 | 22.7 | no | 0.0015393821 | 0.15842354 | 13 |
| Surgeries Prior to Listing: Norwood Stage I: BT Shunt | 0 | 0.00 | no | 0.0015229812 | 0.15673566 | 14 |
| Recipient Height (inches) at Transplant | 64 | 1.69 | 40.2 | 0.0014523551 | 0.14946726 | 15 |
| Recipient Age (years) at Listing | 0 | 0.00 | 4.31 | 0.0013204170 | 0.13588902 | 16 |
| BSA (m2) at Transplant | 65 | 1.72 | 0.68 | 0.0012916666 | 0.13293021 | 17 |
| Transplant Lab: eGFR | 123 | 3.25 | 97.1 | 0.0009546629 | 0.09824791 | 18 |
| Inotropes, Pressors, or Thyroid Hormones: T3 (Tri-iodothyronine) | 532 | 14.0 | no | 0.0009470231 | 0.09746167 | 19 |
| Surgeries Prior to Listing: PA Banding | 0 | 0.00 | no | 0.0009172054 | 0.09439302 | 20 |

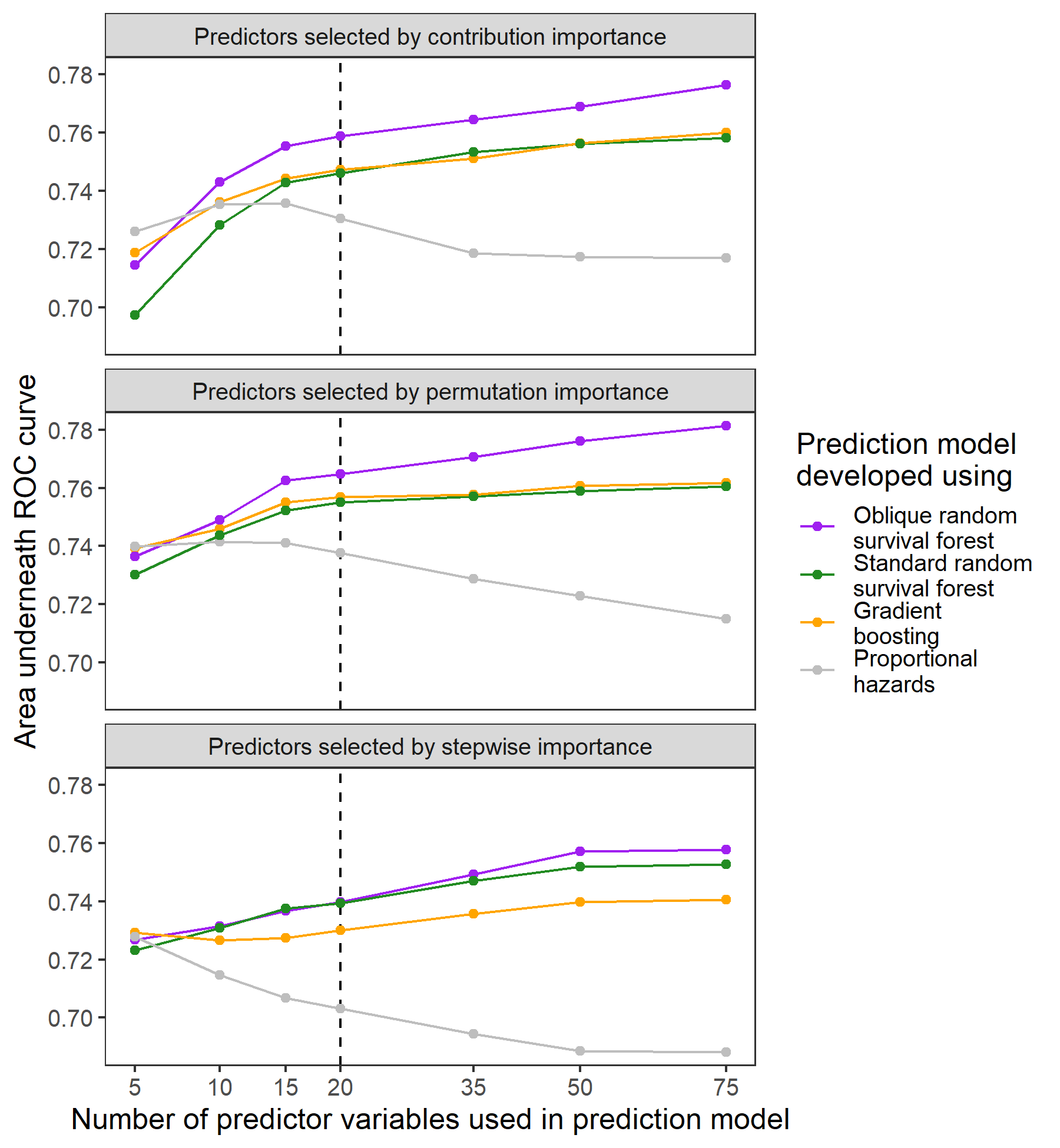
**Table 3:** Predicted risk for graft loss or mortality at 1 year following transplant according to the 10 most important variables selected for inclusion in the final prediction model.

|  | **1-year predicted risk for graft loss or mortality** | | |
| --- | --- | --- | --- |
| **Characteristic** | **Risk (95% CI)** | **Ratio (95% CI)** | **Difference (95% CI)** |
| Cardiopulmonary bypass time (minutes) | | | |
| 25th percentile: 125 | 3.96 (3.80, 4.11) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 161 | 4.98 (4.70, 5.21) | 1.26 (1.22, 1.29) | 1.01 (0.87, 1.18) |
| 75th percentile: 209 | 5.65 (5.34, 5.95) | 1.42 (1.40, 1.46) | 1.68 (1.56, 1.82) |
| Primary Etiology | | | |
| Cardiomyopathy | 4.47 (4.31, 4.83) | 1 (Reference) | 0 (Reference) |
| Congenital heart disease | 6.68 (6.46, 6.83) | 1.49 (1.43, 1.57) | 2.20 (2.01, 2.39) |
| Other | 8.28 (7.69, 8.47) | 1.85 (1.80, 1.89) | 3.81 (3.55, 3.91) |
| ECMO at Transplant | | | |
| No | 4.52 (4.35, 4.81) | 1 (Reference) | 0 (Reference) |
| Yes | 19.3 (19.0, 19.9) | 4.26 (4.11, 4.35) | 14.8 (14.5, 15.2) |
| Surgeries Prior to Listing | | | |
| No | 4.76 (4.52, 5.16) | 1 (Reference) | 0 (Reference) |
| Yes | 6.15 (5.97, 6.27) | 1.29 (1.22, 1.33) | 1.38 (1.10, 1.54) |
| CHD: Single Ventricle | | | |
| No | 4.67 (4.37, 4.95) | 1 (Reference) | 0 (Reference) |
| Yes | 6.58 (6.32, 6.77) | 1.41 (1.35, 1.48) | 1.91 (1.68, 2.11) |
| MSCD at Transplant | | | |
| No | 4.15 (4.00, 4.34) | 1 (Reference) | 0 (Reference) |
| Yes | 7.24 (6.66, 7.71) | 1.74 (1.68, 1.80) | 3.09 (2.71, 3.38) |
| Transplant Lab: Serum Albumin g/dL | | | |
| 25th percentile: 3.30 | 4.93 (4.83, 5.13) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 3.70 | 4.63 (4.50, 4.86) | 0.94 (0.92, 0.95) | -0.30 (-0.39, -0.23) |
| 75th percentile: 4.10 | 4.60 (4.45, 4.80) | 0.93 (0.91, 0.95) | -0.32 (-0.44, -0.25) |
| Medical History at Listing | | | |
| No | 3.70 (3.56, 3.90) | 1 (Reference) | 0 (Reference) |
| Yes | 4.97 (4.71, 5.17) | 1.34 (1.31, 1.37) | 1.27 (1.13, 1.38) |
| Recipient Age (years) at Transplant | | | |
| 25th percentile: 0.75 | 6.04 (5.72, 6.22) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 4.88 | 5.01 (4.77, 5.24) | 0.83 (0.81, 0.85) | -1.04 (-1.14, -0.92) |
| 75th percentile: 13.0 | 4.75 (4.58, 4.91) | 0.79 (0.77, 0.80) | -1.29 (-1.43, -1.17) |
| Transplant Lab: BUN mg/dL | | | |
| 25th percentile: 12.0 | 4.25 (4.13, 4.56) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 16.0 | 4.68 (4.44, 4.94) | 1.10 (1.08, 1.12) | 0.43 (0.35, 0.53) |
| 75th percentile: 22.0 | 5.23 (4.96, 5.52) | 1.23 (1.21, 1.27) | 0.98 (0.90, 1.14) |

**Figure 1:** Model development algorithm considered in the current analysis.



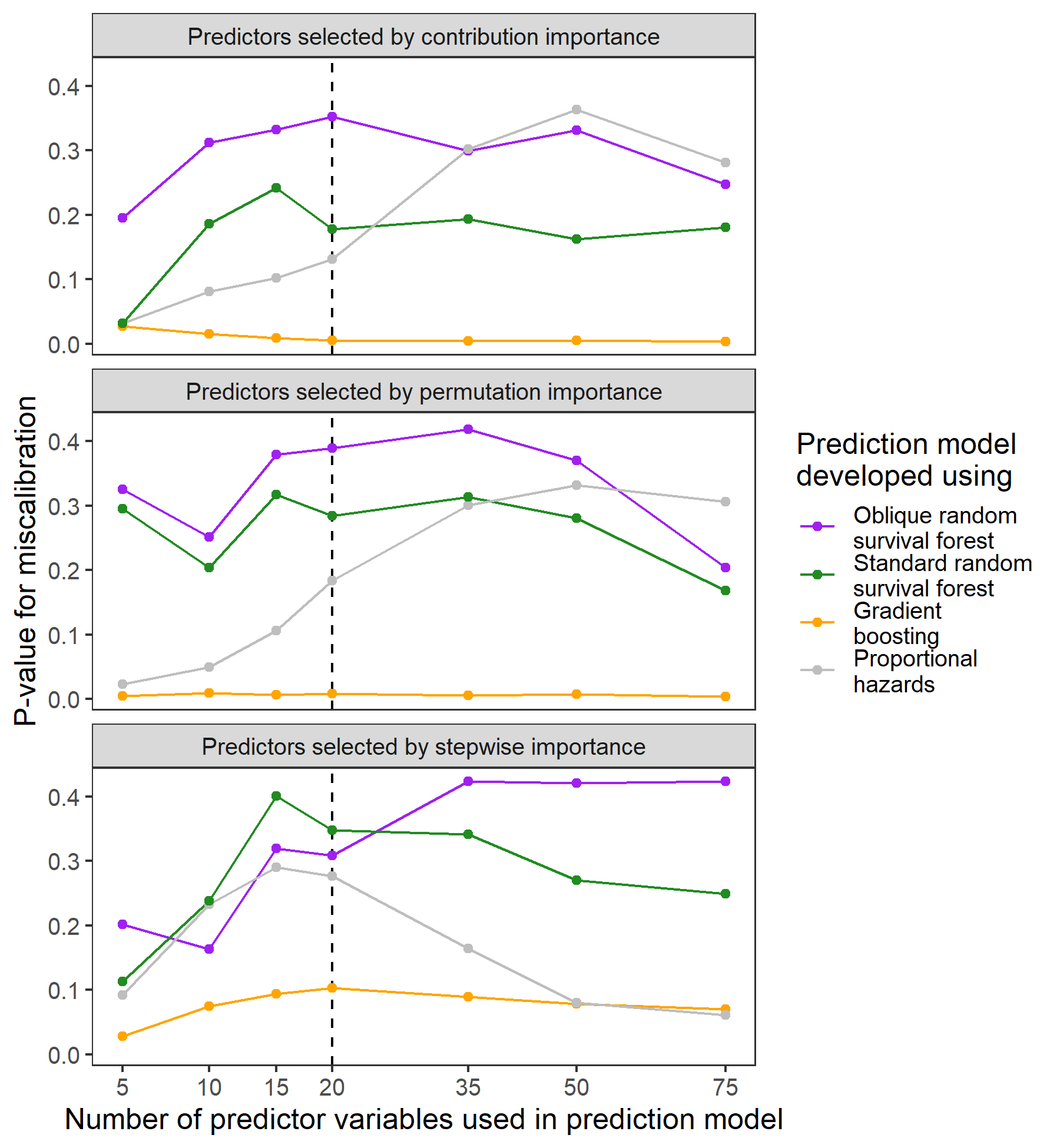
**Figure 2:** Internally validated estimates of model discrimination as a function of the number of predictor variables included in the prediction model.



Each point in the figure is the median value of model performance aggregated across 500 replicates of Monte-Carlo cross validation.

Discrimination was assessed at 1 year after transplant using **FILL IN**.

**Figure 3:** Internally validated estimates of model calibration as a function of the number of predictor variables included in the prediction model.

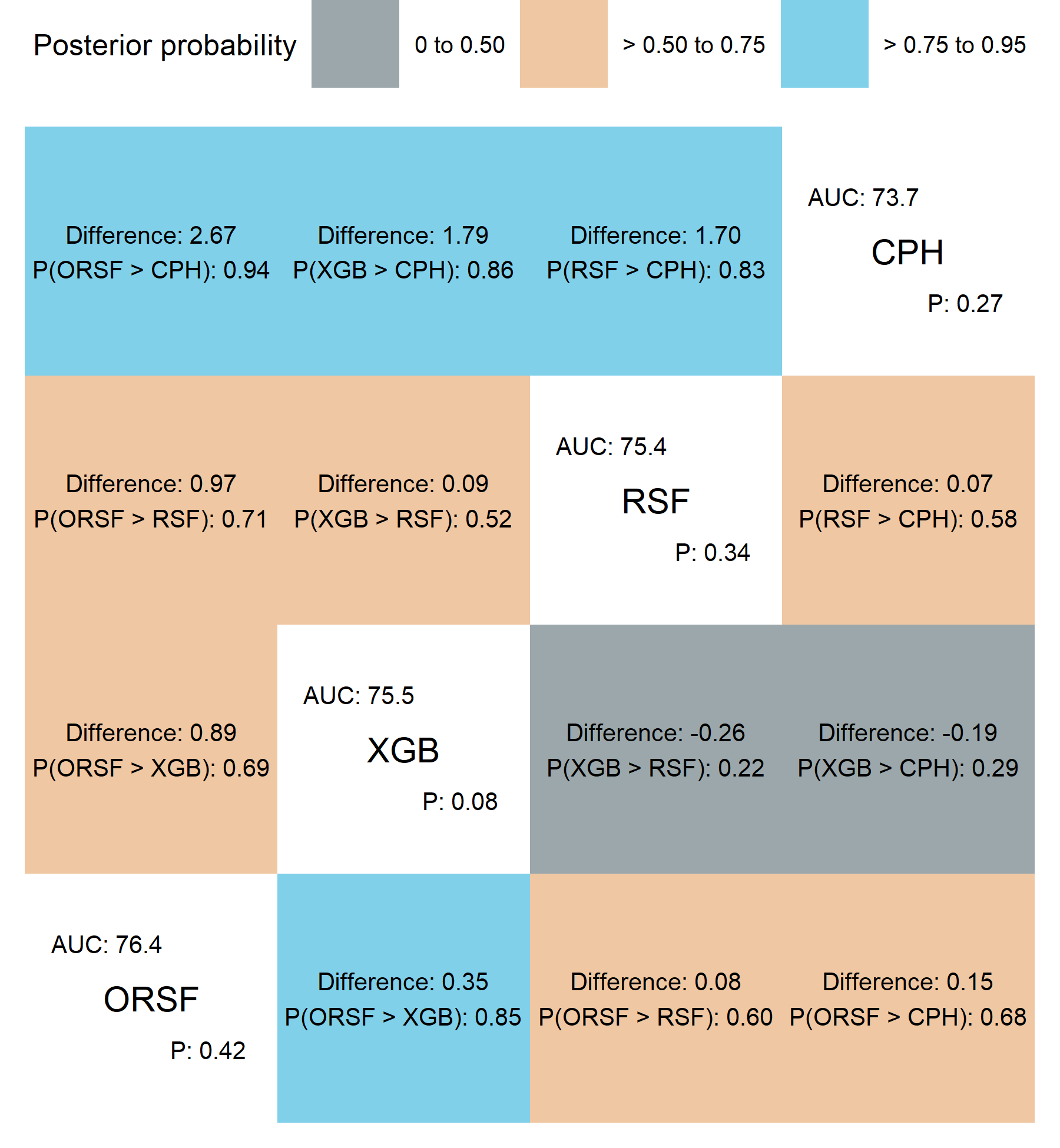


Model performance was defined using the p-value from a statistical test for miscalibration (REF).

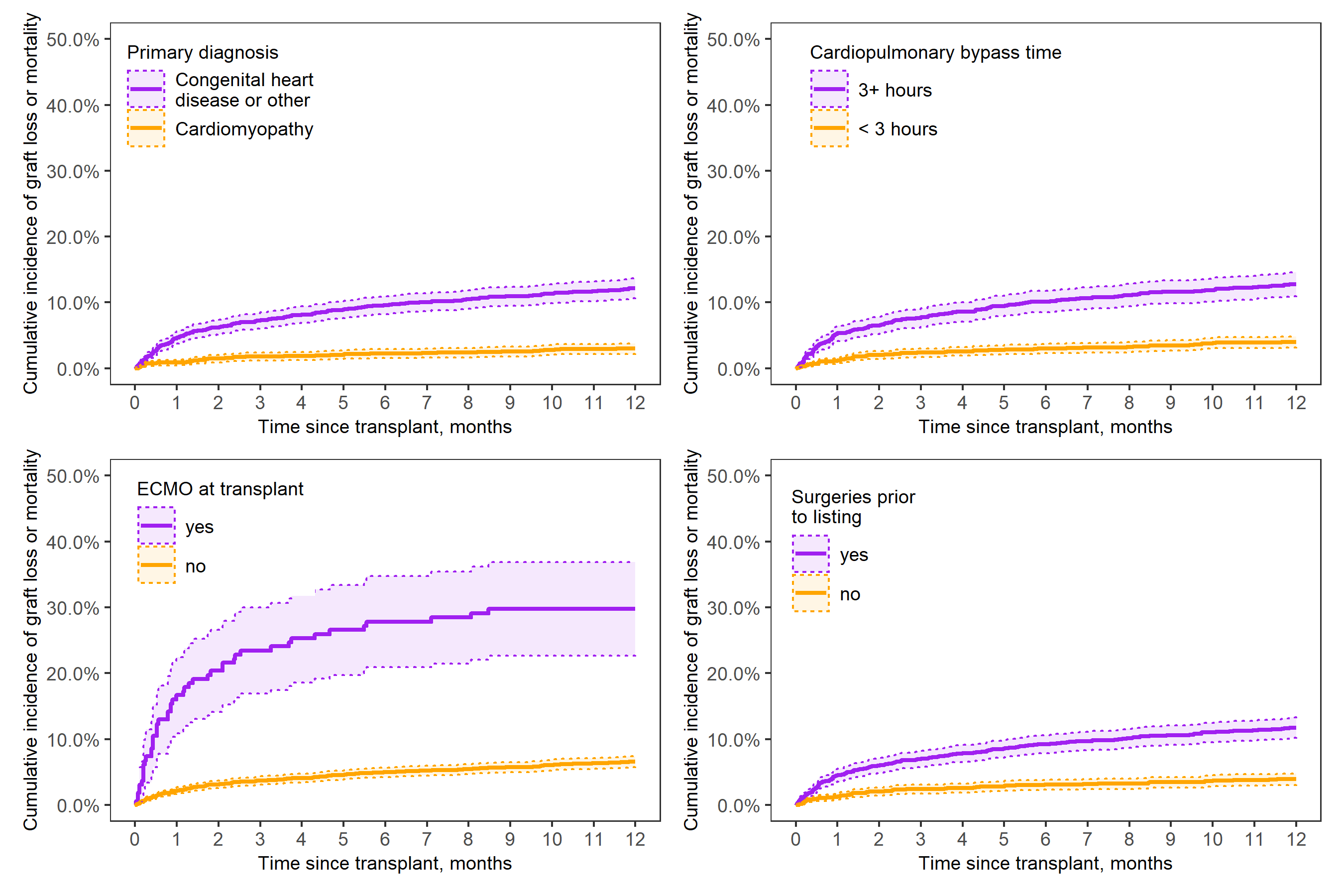
Each point in the figure is the median value of model performance aggregated across 500 replicates of Monte-Carlo cross validation.

Calibration was assessed at 1 year after transplant using **FILL IN**.

**Figure 4:** Bayesian estimates of differences in model performance when 20 variables were selected using permutation importance.



**Figure 5:** Incidence of graft loss or mortality, stratified by the four variables with greatest permuation importance in the final prediction model.



# SUPPLEMENT

**Table S1**: Summary of all continuous candidate predictor variables.

| **Categorical predictor variable** | **Mean (standard deviation)** | **Median (25th, 75th percentile)** |
| --- | --- | --- |
| Listing Year | 2,014 (2.73) | 2,014 (2,012, 2,016) |
| Recipient Age (years) at Listing | 6.50 (6.20) | 4.31 (0.50, 12.8) |
| BSA (m2) at Listing | 0.83 (0.56) | 0.64 (0.33, 1.31) |
| BMI at Listing | 17.2 (4.64) | 16.1 (14.2, 18.8) |
| Listing Lab: Total Bilirubin mg/dL | 1.11 (1.75) | 0.60 (0.30, 1.10) |
| Listing Lab: AST U/L | 63.1 (117) | 36.0 (27.0, 53.0) |
| Listing Lab: ALT U/L | 58.0 (125) | 29.0 (19.0, 45.0) |
| Listing Lab: Creatinine mg/dL | 0.52 (0.35) | 0.44 (0.30, 0.64) |
| Listing Lab: Total Protein g/dL | 5.93 (1.58) | 6.10 (5.30, 6.90) |
| Listing Lab: Serum Albumin g/dL | 3.57 (0.73) | 3.60 (3.10, 4.10) |
| PRA Max List | 13.0 (24.8) | 0.00 (0.00, 13.0) |
| Listing eGFR | 95.4 (37.3) | 92.2 (71.1, 117) |
| Transplant Year | 2,015 (2.70) | 2,015 (2,012, 2,017) |
| Recipient Age (years) at Transplant | 6.85 (6.24) | 4.88 (0.74, 13.1) |
| Recipient Height (inches) at Transplant | 42.7 (16.7) | 40.2 (26.6, 59.1) |
| Recipient Weight (pounds) at Transplant | 58.7 (53.2) | 35.3 (16.6, 94.2) |
| BSA (m2) at Transplant | 0.86 (0.55) | 0.68 (0.37, 1.34) |
| BMI at Transplant | 17.9 (4.51) | 16.8 (14.9, 19.5) |
| Patient in ICU at Transplant | 1.16 (24.5) | 1.00 (0.00, 1.00) |
| Transplant Lab: Total Bilirubin mg/dL | 0.93 (1.26) | 0.50 (0.30, 1.00) |
| Transplant Lab: AST U/L | 65.3 (97.0) | 39.0 (28.0, 65.0) |
| Transplant Lab: ALT U/L | 48.9 (96.0) | 29.0 (20.0, 45.0) |
| Transplant Lab: Creatinine mg/dL | 0.51 (0.39) | 0.41 (0.30, 0.62) |
| Transplant Lab: BUN mg/dL | 18.9 (12.4) | 16.0 (12.0, 22.0) |
| Transplant Lab: Total Protein g/dL | 6.34 (1.19) | 6.40 (5.60, 7.10) |
| Transplant Lab: Serum Albumin g/dL | 3.67 (0.72) | 3.70 (3.20, 4.20) |
| Transplant Lab: eGFR | 102 (38.6) | 97.1 (75.6, 124) |
| Cardiopulmonary bypass time (minutes) | 175 (71.6) | 161 (123, 211) |
| Donor ischemic time (minutes) | 219 (68.5) | 218 (181, 255) |
| PRA Max Txpl | 12.7 (24.4) | 0.00 (0.00, 11.0) |

**Table S2**: Summary of all categorical candidate predictor variables.

| **Categorical predictor variable** | **Count (percent)** |
| --- | --- |
| Recipient Sex | - Female: 1,703 (45.0%) - Male: 2,084 (55.0%) |
| Race | - Black: 646 (17.1%) - Other: 677 (17.9%) - White: 2,464 (65.1%) |
| Recipient Hispanic or Latino | - No: 2,944 (79.9%) - Yes: 741 (20.1%) |
| Recipient: Blood Type | - A: 1,377 (36.4%) - Ab: 147 (3.88%) - B: 529 (14.0%) - O: 1,733 (45.8%) |
| Recipient: Rh | - Negative: 452 (12.0%) - Positive: 3,306 (88.0%) |
| First listing for a heart transplant | - No: 14 (0.37%) - Yes: 3,764 (99.6%) |
| Primary Etiology | - Cardiomyopathy: 1,877 (49.6%) - Congenital hd: 1,816 (48.0%) - Other: 94 (2.48%) |
| CHD: Hypoplastic Left Heart | - No: 2,070 (70.7%) - Yes: 857 (29.3%) |
| CHD: TOF/TOF Variant/DORV/RVOTO | - No: 2,737 (93.5%) - Yes: 190 (6.49%) |
| CHD: Single Ventricle | - No: 2,358 (64.9%) - Yes: 1,273 (35.1%) |
| Surgeries Prior to Listing | - No: 2,006 (53.1%) - Yes: 1,773 (46.9%) |
| Surgeries Prior to Listing: AP Shunt (including BT Shunt, Modified BT Shunt, Waterson Shunt, Pott's Shunt, Central Shunt, and MEE procedure) | - No: 3,643 (96.2%) - Yes: 144 (3.80%) |
| Surgeries Prior to Listing: Arterial switch operation | - No: 3,775 (99.7%) - Yes: 12 (0.32%) |
| Surgeries Prior to Listing: ASD Repair | - No: 3,733 (98.6%) - Yes: 54 (1.43%) |
| Surgeries Prior to Listing: Atrial switch (Senning/Mustard) | - No: 3,780 (99.8%) - Yes: 7 (0.18%) |
| Surgeries Prior to Listing: CABG | - No: 3,784 (99.9%) - Yes: 3 (0.08%) |
| Surgeries Prior to Listing: Complete AV Septal Defect Repair | - No: 3,764 (99.4%) - Yes: 23 (0.61%) |
| Surgeries Prior to Listing: Congenitally Corrected Transposition Repair (classic) | - No: 3,783 (99.9%) - Yes: 4 (0.11%) |
| Surgeries Prior to Listing: Congenitally Corrected Transposition Repair (double switch) | - No: 3,774 (99.7%) - Yes: 13 (0.34%) |
| Surgeries Prior to Listing: Damus Kaye Stansel (DKS) | - No: 3,738 (98.7%) - Yes: 49 (1.29%) |
| Surgeries Prior to Listing: d-Transposition of the Great Vessels Repair | - No: 3,766 (99.4%) - Yes: 21 (0.55%) |
| Surgeries Prior to Listing: Ebstein's Anomaly Repair | - No: 3,782 (99.9%) - Yes: 5 (0.13%) |
| Surgeries Prior to Listing: Fontan Procedure | - No: 3,279 (86.6%) - Yes: 508 (13.4%) |
| Surgeries Prior to Listing: Glenn Procedure | - No: 3,008 (79.4%) - Yes: 779 (20.6%) |
| Surgeries Prior to Listing: Hybrid Palliation | - No: 3,725 (98.4%) - Yes: 62 (1.64%) |
| Surgeries Prior to Listing: Norwood Stage I: BT Shunt | - No: 3,439 (90.8%) - Yes: 348 (9.19%) |
| Surgeries Prior to Listing: Norwood Stage I: Sano/RV-PA conduit | - No: 3,454 (91.2%) - Yes: 333 (8.79%) |
| Surgeries Prior to Listing: PA Banding | - No: 3,525 (93.1%) - Yes: 262 (6.92%) |
| Surgeries Prior to Listing: TOF/DORV/RVOTO Repair | - No: 3,744 (98.9%) - Yes: 43 (1.14%) |
| Surgeries Prior to Listing: Truncus Arteriosus Repair | - No: 3,773 (99.6%) - Yes: 14 (0.37%) |
| Surgeries Prior to Listing: Valve Replacement | - No: 3,639 (96.1%) - Yes: 148 (3.91%) |
| Surgeries Prior to Listing: VSD Repair | - No: 3,699 (97.7%) - Yes: 88 (2.32%) |
| Surgeries Prior to Listing: Other, specify | - No: 2,974 (78.5%) - Yes: 813 (21.5%) |
| Status at Listing: Combined Countries | - Priority: 3,072 (81.1%) - Routine: 715 (18.9%) |
| Was patient in or out of hospital at time of listing | - No: 974 (25.7%) - Yes: 2,813 (74.3%) |
| Listing: Continuous Invasive Mechanical Ventilation | - No: 2,570 (76.4%) - Yes: 795 (23.6%) |
| Did the patient require continuous inotropes at time of listing | - No: 1,444 (38.3%) - Yes: 2,329 (61.7%) |
| Ductal dependent pulmonary or systemic circulation -- by stent -- | - No: 3,643 (96.9%) - Yes: 116 (3.09%) |
| ABO Incompatible at Listing | - No: 2,729 (74.3%) - Yes: 945 (25.7%) |
| VAD at Listing | - No: 3,404 (89.9%) - Yes: 383 (10.1%) |
| ECMO at Listing | - No: 3,616 (95.5%) - Yes: 171 (4.52%) |
| No MCSD at Listing | - No: 554 (14.6%) - Yes: 3,233 (85.4%) |
| Listing Inf Dis Scr: HIV | - Negative: 3,616 (96.2%) - Not done: 135 (3.59%) - Positive: 7 (0.19%) |
| Listing Inf Dis Scr: CMV | - Negative: 2,163 (57.6%) - Not done: 311 (8.28%) - Positive: 1,284 (34.2%) |
| Listing Inf Dis Scr: CMV PCR | - Negative: 1,231 (33.6%) - Not done: 2,336 (63.7%) - Positive: 98 (2.67%) |
| Listing Inf Dis Scr: EBV | - Negative: 1,511 (40.3%) - Not done: 367 (9.79%) - Positive: 1,871 (49.9%) |
| Listing Inf Dis Scr: EBV PCR | - Negative: 1,180 (32.2%) - Not done: 2,342 (63.9%) - Positive: 141 (3.85%) |
| Listing Inf Dis Scr: IFA Toxo | - Negative: 2,814 (75.2%) - Not done: 809 (21.6%) - Positive: 119 (3.18%) |
| Listing Inf Dis Scr: HBs Ag | - Negative: 3,494 (93.0%) - Not done: 219 (5.83%) - Positive: 44 (1.17%) |
| Listing Inf Dis Scr: HB core Ab | - Negative: 3,070 (82.0%) - Not done: 584 (15.6%) - Positive: 90 (2.40%) |
| Listing Inf Dis Scr: HBs Ab | - Negative: 1,617 (43.2%) - Not done: 837 (22.4%) - Positive: 1,287 (34.4%) |
| Listing Inf Dis Scr: Hep C Ab | - Negative: 3,491 (93.2%) - Not done: 223 (5.95%) - Positive: 31 (0.83%) |
| Listing Inf Dis Scr: RPR/Syphilis | - Negative: 2,046 (55.0%) - Not done: 1,666 (44.8%) - Positive: 6 (0.16%) |
| Medical History at Listing | - No: 526 (16.2%) - Yes: 2,728 (83.8%) |
| History at Listing: Arrhythmia | - No: 2,737 (75.9%) - Yes: 868 (24.1%) |
| History at Listing: Cardiac Arrest/CPR | - No: 3,245 (90.0%) - Yes: 360 (9.99%) |
| History at Listing: Diabetes | - No: 3,597 (99.8%) - Yes: 8 (0.22%) |
| History at Listing: GI/Nutrition | - No: 2,802 (77.7%) - Yes: 803 (22.3%) |
| History at Listing: Malignancy | - No: 3,544 (98.3%) - Yes: 61 (1.69%) |
| History at Listing: Neurologic | - No: 3,439 (95.4%) - Yes: 166 (4.60%) |
| History at Listing: Pacemaker | - No: 3,150 (87.4%) - Yes: 455 (12.6%) |
| History at Listing: Peripheral Myopathy/Neuromuscular disease | - No: 3,579 (99.3%) - Yes: 26 (0.72%) |
| History at Listing: Prenatal Diagnosis | - No: 3,222 (89.4%) - Yes: 383 (10.6%) |
| History at Listing: Prior Transfusions | - No: 2,508 (69.6%) - Yes: 1,097 (30.4%) |
| History at Listing: Renal Insufficiency | - No: 3,463 (96.1%) - Yes: 142 (3.94%) |
| History at Listing: Shock | - No: 3,473 (96.3%) - Yes: 132 (3.66%) |
| History at Listing: Other | - No: 3,065 (85.0%) - Yes: 540 (15.0%) |
| HX Listing: Arrhythmia: Afib/flutter | - No: 3,210 (95.9%) - Yes: 137 (4.09%) |
| HX Listing: Arrhythmia: Complete heart block | - No: 3,195 (95.5%) - Yes: 152 (4.54%) |
| HX Listing: Arrhythmia: V Fibrillation | - No: 3,269 (97.7%) - Yes: 78 (2.33%) |
| HX Listing: Arrhythmia: V Tachycardia | - No: 2,963 (88.5%) - Yes: 384 (11.5%) |
| HX Listing: Arrhythmia: Unknown | - No: 3,341 (99.8%) - Yes: 6 (0.18%) |
| HX Listing: Arrhythmia: Other | - No: 3,073 (91.8%) - Yes: 274 (8.19%) |
| HX Listing: GI: Failure to thrive | - No: 2,707 (81.2%) - Yes: 626 (18.8%) |
| HX Listing: GI: Infectious hepatitis | - No: 3,327 (99.8%) - Yes: 6 (0.18%) |
| HX Listing: GI: Protein losing Enteropathy | - No: 3,188 (95.6%) - Yes: 145 (4.35%) |
| HX Listing: Malignancy: Other | - No: 3,210 (98.7%) - Yes: 41 (1.26%) |
| HX Listing: Neurologic: Anoxic brain injury | - No: 3,268 (99.8%) - Yes: 5 (0.15%) |
| HX Listing: Neurologic: Hemorrhagic /thromboembolic stroke | - No: 3,144 (96.1%) - Yes: 129 (3.94%) |
| HX Listing: Neurologic: Other | - No: 3,235 (98.8%) - Yes: 38 (1.16%) |
| HX Listing: Pacemaker: Defibrillator/AICD | - No: 3,080 (94.6%) - Yes: 176 (5.41%) |
| HX Listing: Pacemaker: CRT/biventricular pacing | - No: 3,129 (95.2%) - Yes: 158 (4.81%) |
| HX Listing: Pacemaker: not CRT and not ICD | - No: 3,237 (98.5%) - Yes: 50 (1.52%) |
| HX Listing: Respiratory: Asthma | - No: 3,198 (97.4%) - Yes: 86 (2.62%) |
| HX Listing: Respiratory: Plastic bronchitis | - No: 3,194 (98.2%) - Yes: 58 (1.78%) |
| Primary Insurance | - Charitable dona: 2 (0.05%) - Free: 134 (3.54%) - Government: 2,084 (55.1%) - Other: 96 (2.54%) - Private: 1,433 (37.9%) - Self pay: 32 (0.85%) |
| Listed for prospective crossmatch | - No: 3,204 (88.8%) - Yes: 403 (11.2%) |
| Were hemodynamics done prior to listing | - No: 1,572 (42.3%) - Yes: 2,147 (57.7%) |
| Prior to Listing: Hemodynamic Agents Used | - No: 2,323 (79.5%) - Yes: 599 (20.5%) |
| Prior to Listing: Hemodynamic Agent: 100% O2 | - No: 3,148 (94.1%) - Yes: 198 (5.92%) |
| Prior to Listing: Hemodynamic Agent: Dobutamine | - No: 3,319 (99.2%) - Yes: 27 (0.81%) |
| Prior to Listing: Hemodynamic Agent: Dopamine | - No: 3,257 (97.3%) - Yes: 89 (2.66%) |
| Prior to Listing: Hemodynamic Agent: Isoproterenol | - No: 3,346 (100%) |
| Prior to Listing: Hemodynamic Agent: Nitroglycerin | - No: 3,338 (99.8%) - Yes: 8 (0.24%) |
| Prior to Listing: Hemodynamic Agent: Nitroprusside | - No: 3,335 (99.7%) - Yes: 11 (0.33%) |
| Prior to Listing: Hemodynamic Agent: Norepinephrine | - No: 3,345 (100%) - Yes: 1 (0.03%) |
| Prior to Listing: Hemodynamic Agent: Other | - No: 3,267 (97.6%) - Yes: 79 (2.36%) |
| School at Listing: In School | - No: 99 (2.66%) - Not routinely done: 2,005 (53.8%) - Yes: 1,623 (43.5%) |
| Was exercise test performed | - No: 3,315 (88.5%) - Not routinely done: 262 (6.99%) - Yes: 169 (4.51%) |
| NYHA Class at Listing | - Not done: 1,980 (56.7%) - Nyha 1: 22 (0.63%) - Nyha 2: 220 (6.29%) - Nyha 3: 497 (14.2%) - Nyha 4: 776 (22.2%) |
| Ross Heart Failure Class at Listing | - Not done: 2,226 (71.9%) - Ross 1: 13 (0.42%) - Ross 2: 126 (4.07%) - Ross 3: 254 (8.21%) - Ross 4: 475 (15.4%) |
| Transplant simultaneous organ: None | - No: 17 (0.45%) - Yes: 3,763 (99.6%) |
| Transplant type | - Heterotopic: 12 (0.32%) - Orthotopic: 3,772 (99.7%) |
| Status at Transplant: Country | - Brazil: 51 (1.35%) - Canada: 164 (4.33%) - United kingdom: 252 (6.65%) - United states: 3,320 (87.7%) |
| Patient in Hospital at Transplant | - No: 1,001 (26.4%) - Yes: 2,786 (73.6%) |
| Status at Transplant: Combined Countries | - Priority: 3,464 (92.1%) - Routine: 297 (7.90%) |
| Continuous Invasive Mechanical Ventilation at Transplant | - No: 2,635 (80.7%) - Yes: 629 (19.3%) |
| Continuous Inotropes at Transplant | - No: 1,503 (39.9%) - Yes: 2,265 (60.1%) |
| ABO Incompatible at Transplant | - No: 3,437 (91.4%) - Yes: 323 (8.59%) |
| VAD at Transplant | - No: 2,802 (74.0%) - Yes: 985 (26.0%) |
| ECMO at Transplant | - No: 3,624 (95.7%) - Yes: 163 (4.30%) |
| Transplant: Donor Specific or Retrospective Crossmatch | - No: 458 (12.3%) - Yes: 3,255 (87.7%) |
| Donor Specific Antibodies (DSA) | - No: 3,201 (87.8%) - Yes: 445 (12.2%) |
| Hemodynamics Performed | - No: 2,726 (73.3%) - Yes: 995 (26.7%) |
| Hemodynamic Agent: 100% O2 | - No: 3,179 (97.7%) - Yes: 76 (2.33%) |
| Hemodynamic Agent: Dobuatmine | - No: 3,245 (99.7%) - Yes: 10 (0.31%) |
| Hemodynamic Agent: Dopamine | - No: 3,227 (99.1%) - Yes: 28 (0.86%) |
| Hemodynamic Agent: Isoproterenol (Isuprel) | - No: 3,252 (99.9%) - Yes: 3 (0.09%) |
| Hemodynamic Agent: Milrinone | - No: 2,930 (90.0%) - Yes: 325 (9.98%) |
| Hemodynamic Agent: Nesiritide | - No: 3,254 (100%) - Yes: 1 (0.03%) |
| Hemodynamic Agent: Nitric Oxide | - No: 3,199 (98.3%) - Yes: 56 (1.72%) |
| Hemodynamic Agent: Nitroglycerine | - No: 3,255 (100%) |
| Hemodynamic Agent: Nitroprusside (Nipride) | - No: 3,247 (99.8%) - Yes: 8 (0.25%) |
| Hemodynamic Agent: PGE (Alprostadil) | - No: 3,232 (99.3%) - Yes: 23 (0.71%) |
| Hemodynamic Agent: PGI (Flolan) | - No: 3,254 (100%) - Yes: 1 (0.03%) |
| Was patient on inoropes, pressors, or thyroid hromones at time of transplant? | - No: 1,337 (35.8%) - Yes: 2,397 (64.2%) |
| Inotropes, Pressors, or Thyroid Hormones: Epinephrine | - No: 2,850 (87.2%) - Yes: 418 (12.8%) |
| Inotropes, Pressors, or Thyroid Hormones: Milrinone | - No: 1,123 (33.3%) - Yes: 2,247 (66.7%) |
| Inotropes, Pressors, or Thyroid Hormones: Neosynephrine | - No: 3,234 (99.4%) - Yes: 18 (0.55%) |
| Inotropes, Pressors, or Thyroid Hormones: T3 (Tri-iodothyronine) | - No: 3,185 (97.8%) - Yes: 70 (2.15%) |
| Inotropes, Pressors, or Thyroid Hormones: T4 (Levothyroxine) | - No: 3,230 (99.2%) - Yes: 27 (0.83%) |
| Technique of transplant: Bicaval | - No: 453 (12.9%) - Yes: 3,049 (87.1%) |
| MSCD at Transplant | - No: 2,639 (69.7%) - Yes: 1,148 (30.3%) |

**Table S3**: Predicted risk for graft loss or mortality at 1 year following transplant according to variables with rank 11th or higher importance selected for inclusion in the final prediction model.

|  | **1-year predicted risk for graft loss or mortality** | | |
| --- | --- | --- | --- |
| **Characteristic** | **Risk (95% CI)** | **Ratio (95% CI)** | **Difference (95% CI)** |
| PRA Max List | | | |
| 25th percentile: 0.00 | 4.51 (4.37, 4.70) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 0.00 | 4.51 (4.34, 4.73) | 1.00 (1.00, 1.00) | 0.00 (0.00, 0.00) |
| 75th percentile: 10.0 | 4.98 (4.78, 5.16) | 1.10 (1.08, 1.11) | 0.47 (0.38, 0.51) |
| Transplant Lab: Total Protein g/dL | | | |
| 25th percentile: 5.80 | 4.76 (4.54, 4.99) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 6.40 | 4.60 (4.27, 4.87) | 0.97 (0.95, 0.98) | -0.16 (-0.25, -0.11) |
| 75th percentile: 6.90 | 4.87 (4.70, 5.04) | 1.02 (1.00, 1.04) | 0.11 (0.02, 0.18) |
| CHD: Hypoplastic Left Heart | | | |
| No | 4.61 (4.43, 4.74) | 1 (Reference) | 0 (Reference) |
| Yes | 6.08 (5.95, 6.21) | 1.32 (1.29, 1.35) | 1.46 (1.40, 1.54) |
| Surgeries Prior to Listing: Norwood Stage I: BT Shunt | | | |
| No | 4.49 (4.25, 4.70) | 1 (Reference) | 0 (Reference) |
| Yes | 7.30 (7.06, 7.66) | 1.62 (1.59, 1.66) | 2.80 (2.71, 2.97) |
| Recipient Height (inches) at Transplant | | | |
| 25th percentile: 26.8 | 6.31 (6.11, 6.60) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 40.2 | 5.05 (4.74, 5.27) | 0.80 (0.77, 0.81) | -1.26 (-1.43, -1.19) |
| 75th percentile: 59.1 | 4.60 (4.40, 4.74) | 0.73 (0.71, 0.75) | -1.71 (-1.85, -1.52) |
| Recipient Age (years) at Listing | | | |
| 25th percentile: 0.50 | 5.75 (5.50, 6.02) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 4.31 | 5.18 (4.88, 5.44) | 0.90 (0.88, 0.91) | -0.57 (-0.67, -0.52) |
| 75th percentile: 12.8 | 4.81 (4.62, 4.95) | 0.84 (0.83, 0.85) | -0.94 (-0.98, -0.82) |
| BSA (m2) at Transplant | | | |
| 25th percentile: 0.38 | 5.39 (5.18, 5.53) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 0.68 | 4.92 (4.63, 5.16) | 0.91 (0.89, 0.94) | -0.48 (-0.57, -0.33) |
| 75th percentile: 1.33 | 4.86 (4.58, 5.15) | 0.90 (0.89, 0.93) | -0.53 (-0.62, -0.40) |
| Transplant Lab: eGFR | | | |
| 25th percentile: 76.4 | 5.05 (4.81, 5.26) | 1 (Reference) | 0 (Reference) |
| 50th percentile: 97.1 | 4.44 (4.32, 4.62) | 0.88 (0.86, 0.90) | -0.61 (-0.73, -0.51) |
| 75th percentile: 122 | 4.51 (4.32, 4.74) | 0.89 (0.88, 0.91) | -0.54 (-0.64, -0.43) |
| Inotropes, Pressors, or Thyroid Hormones: T3 (Tri-iodothyronine) | | | |
| No | 4.54 (4.35, 4.82) | 1 (Reference) | 0 (Reference) |
| Yes | 14.5 (14.3, 14.6) | 3.18 (3.04, 3.29) | 9.92 (9.68, 10.1) |
| Surgeries Prior to Listing: PA Banding | | | |
| No | 4.54 (4.41, 4.69) | 1 (Reference) | 0 (Reference) |
| Yes | 7.15 (6.96, 7.38) | 1.58 (1.53, 1.61) | 2.61 (2.50, 2.73) |

# REFERENCES

1. Friedman JH. Greedy function approximation: A gradient boosting machine. Annals of statistics 2001:1189–1232.

2. Breiman L. Random forests. Machine learning 2001;45:5–32.

3. Zhang Z. Variable selection with stepwise and best subset approaches. Annals of translational medicine 2016;4.

4. Freedman L, Pee D, Midthune D. The problem of underestimating the residual error variance in forward stepwise regression. Journal of the Royal Statistical Society: Series D (The Statistician) 1992;41:405–412.

5. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.

6. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining., 2016:785–794.

7. Breiman L. Bagging predictors. Machine learning 1996;24:123–140.

8. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS, others. Random survival forests. The annals of applied statistics 2008;2:841–860.

9. Jaeger BC, Long DL, Long DM, et al. Oblique random survival forests. The Annals of Applied Statistics 2019;13:1847–1883.

10. Xu Q-S, Liang Y-Z. Monte carlo cross validation. Chemometrics and Intelligent Laboratory Systems 2001;56:1–11.

11. Steyerberg EW, Harrell Jr FE, Borsboom GJ, Eijkemans M, Vergouwe Y, Habbema JDF. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of clinical epidemiology 2001;54:774–781.

12. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: A framework for some traditional and novel measures. Epidemiology (Cambridge, Mass.) 2010;21:128.

13. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: Seven steps for development and an abcd for validation. European heart journal 2014;35:1925–1931.

14. Blanche P, Dartigues J-F, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Statistics in medicine 2013;32:5381–5397.

15. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Statistics in medicine 2015;34:1659–1680.

16. Pencina MJ, D’Agostino Sr RB, D’Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the roc curve to reclassification and beyond. Statistics in medicine 2008;27:157–172.

17. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. Springer; 2013.

18. Josse J, Prost N, Scornet E, Varoquaux G. On the consistency of supervised learning with missing values. arXiv preprint arXiv:1902.06931 2019.

19. Benavoli A, Corani G, Demšar J, Zaffalon M. Time for a change: A tutorial for comparing multiple classifiers through bayesian analysis. The Journal of Machine Learning Research 2017;18:2653–2688.

20. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): Explanation and elaboration. Annals of internal medicine 2015;162:W1–W73.

21. Molnar C. Interpretable machine learning. Lulu. com; 2020.

22. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available at: <https://www.R-project.org/>.

23. Jaeger B. Table.glue: Make and apply customized rounding specifications for tables. Available at: <https://github.com/bcjaeger/table.glue>.

24. Landau WM. The drake R package: A pipeline toolkit for reproducibility and high-performance computing. Journal of Open Source Software 2018;3. Available at: <https://doi.org/10.21105/joss.00550>.

25. Kuhn M, Wickham H. Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.; 2020. Available at: <https://www.tidymodels.org>.

26. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. Journal of Open Source Software 2019;4:1686.