ORIGINAL ARTICLE

AJT

Variables of importance in the Scientific Registry of Transplant Recipients database predictive of heart transplant waitlist mortality

Eileen M. Hsich[1,2] | Lucy Thuita[3] | Dennis M. McNamara[4] | Joseph G. Rogers[5] |
Maryam Valapour[6] | Lee R. Goldberg[7] | Clyde W. Yancy[8] | Eugene H. Blackstone[1,2,3] |
Hemant Ishwaran[9] | Transplantation of HEarts to MaxImize Survival (THEMIS) Investigators

[1]The Heart and Vascular Institute, Cleveland Clinic, Cleveland, Ohio

[2]Cleveland Clinic Lerner College of Medicine of Case Western Reserve University School of Medicine, Cleveland, Ohio

[3]Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio

[4]University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

[5]Division of Cardiology, Duke University, Durham, North Carolina

[6]Department of Pulmonary Medicine, Cleveland Clinic, Cleveland, Ohio

[7]Division of Cardiology, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania

[8]Division of Cardiology, Northwestern University Medical Center, Chicago, Illinois

[9]Department of Public Health Sciences, Division of Biostatistics, University of Miami, Miami, Florida

Correspondence
Eileen Michelle Hsich
Email: Hsiche@ccf.org

The prelisting variables essential for creating an accurate heart transplant allocation score based on survival are unknown. To identify these we studied mortality of adults on the active heart transplant waiting list in the Scientific Registry of Transplant Recipients database from January 1, 2004 to August 31, 2015. There were 33 069 candidates awaiting heart transplantation: 7681 UNOS Status 1A, 13 027 Status 1B, and 12 361 Status 2. During a median waitlist follow-up of 4.3 months, 5514 candidates died. Variables of importance for waitlist mortality were identified by machine learning using Random Survival Forests. Strong correlates predicting survival were estimated glomerular filtration rate (eGFR), serum albumin, extracorporeal membrane oxygenation, ventricular assist device, mechanical ventilation, peak oxygen capacity, hemodynamics, inotrope support, and type of heart disease with less predictive variables including antiarrhythmic agents, history of stroke, vascular disease, prior malignancy, and prior tobacco use. Complex interactions were identified such as an additive risk in mortality based on renal function and serum albumin, and sex-differences in mortality when eGFR >40 mL/min/1.73 m. Most predictive variables for waitlist mortality are in the current tiered allocation system except for eGFR and serum albumin which have an additive risk and complex interactions.

KEYWORDS
artificial organs/support devices: heart/ventricular assist devices, clinical research/practice, gender, health services and outcomes research, heart transplantation/cardiology, organ procurement and allocation, organ transplantation in general, patient survival, Scientific Registry for Transplant Recipients (SRTR), waitlist management

1 | INTRODUCTION

Allocation of donor hearts in the United States is based on a tiered system that prioritizes candidates by risk of death while on the waiting list and has recently changed from 3 to 6 active tiers to better reflect medical urgency.[1] Although the 6-tiered system is more granular, there is still a desire to create an allocation score that incorporates all the important factors affecting waitlist mortality and early posttransplant mortality. To accomplish this goal, the Organ Procurement and Transplantation Network (OPTN)/UNOS Thoracic

---

Committee asked which variables are strong correlates predicting mortality.[1]

The current national OPTN/UNOS database has limited data elements and does not include serum sodium or prognostic biomarkers like natriuretic peptides.[2-4] To be able to collect more data and not over burden transplant centers, one needs to identify both the strong and weak correlates predictive of mortality in the national database. The objective of this study is to identify variables currently in the national database, but not necessarily crucial for UNOS Status, that are most important for heart transplant waitlist mortality and those that are least important.

2 | METHODS

2.1 | Patient population

All 33 069 adult candidates on the active waiting list for heart transplantation in the national Scientific Registry of Transplant Recipients (SRTR) database between January 1, 2004 and August 31, 2015 were included in this study. We excluded inactive adult candidates (N = 1428 UNOS Status 7 candidates) and candidates <18 years of age (N = 6234) because UNOS criteria for pediatric candidates differs from those for adults and the donor pools are distinguished by age.[5]

2.2 | SRTR database

This study used data from the SRTR. The SRTR data system includes data on all donor, waitlisted candidates, and transplant recipients in the United States, submitted by the members of the Organ Procurement and Transplantation Network (OPTN). The Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services provides oversight to the activities of the OPTN and SRTR contractors. Human error collecting data is minimized by edit checks, validation of data at time of entry, and internal verification of outliers.[6] The study was approved by the Cleveland Clinic Institutional Review Board and informed consent was waived because data obtained from routine care were completely de-identified by SRTR prior to their transmission to the investigators.

The data reported here have been supplied by the Hennepin Healthcare Research Institute (HHRI) as the contractor for the SRTR. The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy of or interpretation by the SRTR or the U.S. Government.

2.3 | Study design

The primary endpoint was all-cause mortality assessed as time from listing to death, right censored at time of heart transplantation or end of follow-up. This analysis was based on intent-to-treat such that deaths following removal from the waiting list were included in the primary analysis. SRTR mortality data were verified with the complete Social Security Death Master File recently available through a specific waiver granted to the SRTR. During a median heart transplant waitlist follow-up of 4.3 months, 5403 candidates died (1379 initially listed as UNOS Status 1A, 1876 UNOS Status 1B, and 2148 UNOS Status 2).

Data at time of initial wait listing was used for analysis. Continuous variables were expressed as medians and accompanied by 25th and 75th (Q1, Q3) percentiles. These included year, age, body mass index (BMI), pulmonary artery mean, pulmonary capillary wedge pressure (PCWP) mean, cardiac index, total albumin, and estimated glomerular filtration rate (eGFR) calculated with the Modification of Diet in Renal Disease (MDRD) Study equation.[7] Categorical variables were expressed as number of candidates and frequency. These variables included race (white, black, Hispanic, Asian, other), insurance (private, Medicare, Medicaid, other), ABO blood type, cardiac diagnosis (dilated cardiomyopathy, ischemic cardiomyopathy, congenital heart disease, hypertrophic cardiomyopathy, restrictive cardiomyopathy, valvular cardiomyopathy, and other), history of tobacco use, diabetes mellitus, hypertension, malignancy, peripheral vascular disease, stroke, antiarrhythmic agents, inotropes, implantable cardioverter defibrillator (ICD), mechanical ventilators, intra-aortic balloon pump (IABP), extracorporeal membrane oxygenation (ECMO), and type of ventricular assist device (VAD) including left VAD (LVAD), right VAD (RVAD) ± LVAD, total artificial hearts (TAH), and unspecified mechanical circulatory device.

2.4 | Statistical analysis

The primary reason for using Random Survival Forest (RSF), a robust, nonparametric machine learning algorithmic method, instead of conventional statistical methods was to identify risk factors without prior knowledge of any possible parametric relationship (linear or nonlinear), identify complex interactions, and overcome barriers such as high amounts of missing data. RSF also has the potential to improve the allocation scheme since it has successfully been utilized to pick variables to create esophageal staging[8,9] and can be used to rapidly assess an individual's prognosis as a step towards precision medicine for advanced heart failure therapy.[10]

2.4.1 | Data stratification

Candidates were stratified by UNOS Status because prior research showed sex-specific differences not easily identified when analyzed over the entire cohort such as higher mortality in women compared to men awaiting transplant as UNOS Status 1A and lower mortality awaiting transplant as UNOS Status 2.[11] UNOS Status was based on the three-tier system (UNOS Status 1A, 1B, and 2) because this was the OPTN/UNOS policy at time of listing.[12]

2.4.2 | Missing data

All variables had low level of missingness (<10%) except for PCWP (12%), hypertension (18%), albumin (20%), peripheral

vascular disease (21%), anti-arrythmic agents (25%), and peak oxygen consumption (64%). Missing data were imputed using missForest imputation.[13-15]

2.4.3 | Survival analyses

Preimputed data were stratified by UNOS Status and separate RSF analyses[16] were applied to each UNOS Status using the random-ForestSRC R-software package (additional supporting information for the RSF analysis may be found online in the supporting information tab for this article). Prognostic factors were ranked by their predictive importance using observations not used for constructing a given tree (out-of-bag data). Specifically, bootstrap sampling from the original cohort was used to generate a given tree. The bootstrap sampling procedure was similar to leave-one-out cross-validation but superior because it generally has lower variance than leave-one-out. Each bootstrap sample left out 37% of the data on average, which is referred to as the out-of-bag data. Use of out-of-bag data to evaluate each tree provides a consistent estimate of external prediction accuracy with multiple internal validation cohorts.[15]

Variable importance (VIMP) was calculated using Harrell's concordance index. VIMP measures the difference in prediction error for an RSF with a variable in the model and with the variable randomly permuted (generating noise). Positive VIMP values indicate variables that are predictive, adjusted for all other variables. Thus under Harrell's concordance index, a VIMP of 5% indicates a variable that improves by 5% the ability of RSF to rank two new candidates by their survival.[16] Strong correlates were based on an alpha = 0.05 level of confidence for VIMP, where confidence was determined using delete-d jackknife confidence intervals.[17]

To derive valid standard errors and confidence regions for VIMP, each RSF procedure was subsampled 1000 times using a subsampling rate of 0.5% (one over the square root of the sample size). Confidence regions were determined using the deleted jackknife under the assumption of asymptotic normality.[18]

2.4.4 | Partial plots

Partial plots[19] were used to investigate the relationship between survival and candidate variables. Partial plots display adjusted survival versus target patient variables, where adjusted survival is defined as out-of-bag survival for a patient adjusted by integrating out all patient variables other than the targeted variable of interest. Integration is approximated using the data by averaging over variables.[19]

3 | RESULTS

3.1 | Study population

Most heart transplant candidates were male, white, over 50 years old, blood type O, and had a dilated or ischemic cardiomyopathy (Table 1). Median eGFR was above 60 mL/min/1.73 m^2 and few were dialysis dependent. Implantable cardioverter defibrillators were present in most candidates (73%) while mechanical circulatory support and inotropes were almost exclusively used in UNOS Status 1A and 1B candidates. Hemodynamics were notable for higher filling pressures among UNOS Status 1A candidates compared to UNOS Status 2 candidates. Median serum albumin was lowest among UNOS Status 1A candidates and highest among UNOS Status 2 candidates with no clinically significant differences between groups regarding median eGFR.

3.2 | Relationship of survival to important variables

Survival was best for UNOS Status 2 candidates and worse for UNOS Status 1A compared to UNOS Status 1B candidates. With respect to renal function, which is not currently considered in the heart transplant allocation system, waitlist survival declined substantially for all heart transplant candidates when eGFR fell below about 80 mL/min/1.73 m^2 (Figure 1). Although eGFR calculation adjusted for sex and race, we identified some important sex differences. Among UNOS Status 1A candidates, there was worse survival in women compared to men with eGFR between 40-80 mL/min/1.73 m^2, with survival curves continuing to diverge until they plateaued with an eGFR <80 mL/min/1.73 m^2. Among UNOS Status 1B candidates, the associations between renal function and sex were similar to UNOS Status 1A except the magnitude of the sex difference was less and the overall survival was better for UNOS Status 1B compared to UNOS Status 1A. Among UNOS Status 2 candidates, women had better survival than men with sex differences in survival increasing the longer candidates awaited for a heart transplant.

Another variable not considered in the current heart transplant allocation system is serum albumin. We identified a near linear association between albumin and survival with worse survival among those with lower albumin when compared to higher albumin (Figure 2). UNOS Status 1A candidates had lower survival for any given albumin value than those whose UNOS Status was 1B or 2.

The interaction of survival among UNOS Status, eGFR, and albumin demonstrates that for any given UNOS Status, survival was worse over time (about a 20% decline in survival from years 1 to 5) and the effect of eGFR and serum albumin on survival was cumulative with renal function having a greater influence than albumin on survival (Figure 3). Survival began to plateau around an eGFR of 80 mL/min/1.73 m^2 and was worse among candidates with eGFR <40 mL/ min/1.73 m^2 and albumin <3.0 mg/dL. Candidates in UNOS Status 1A had a lower survival than UNOS Status 1B and UNOS Status 2 candidates. Among the most urgently listed candidates (UNOS Status 1A and 1B), the effect of serum albumin on survival diminished over time (less effect when comparing 1- to 5-year survival). Among UNOS Status 2 candidates, the relationship between eGFR and albumin remained essentially unchanged over time

**TABLE 1** Baseline characteristics of patients awaiting heart transplantation

	UNOS status 1A N = 7681	UNOS status 1B N = 13 027	UNOS status 2 N = 12 361
Variable			
Female, n (%)	1965 (26)	3226 (25)	3197 (26)
Age (y), median (Q1, Q3)	54 (43,61)	55 (44,62)	55 (46,62)
Race, n (%)			
White	4952 (64)	8235 (63)	9155 (74)
Black	1715 (22)	3321 (25)	1920 (16)
Hispanic	643 (8)	1024 (8)	847 (7)
Asian	292 (4)	329 (3)	284 (2)
Other	79 (1)	118 (1)	155 (1)
BMI, median (Q1, Q3)	26 (23,30)	27 (24,31)	28 (24,31)
Insurance, n (%)			
Private	4222 (55)	6716 (52)	6841 (55)
Medicare	2018 (26)	4033 (31)	3780 (31)
Medicaid	1066 (14)	1685 (13)	1181(10)
Other	375 (5)	593 (5)	559 (5)
ABO blood type, n (%)			
A	2938 (38)	4813 (37)	4857 (39)
B	1161 (15)	1802 (14)	1491 (12)
O	3187 (41)	5821 (45)	5490 (44)
AB	395 (5)	591 (5)	523 (4)
Diagnosis, n (%)			
Dilated	3819 (50)	6997 (54)	4900 (40)
Ischemic	2859 (37)	4689 (36)	4911 (40)
Congenital	126 (2)	298 (2)	620 (5)
Hypertrophic	119 (2)	196 (2)	362 (3)
Restrictive	166 (2)	254 (2)	491 (4)
Valvular	129 (2)	214 (2)	269 (2)
Other	463 (6)	379 (3)	808 (7)
ICD, n (%)	4911 (64)	10 085 (77)	9215 (75)
Diabetes mellitus, n (%)	2064 (27)	3887 (30)	3286 (27)
Dialysis at listing, n (%)	329 (4)	263 (2)	344(3)
Hypertension, n (%)	2984 (39)	492 (42)	5422 (44)
Tobacco usage, n (%)	3453 (45)	6310 (48)	5708 (46)
Malignancy, n (%)	504 (7)	953 (7)	836 (7)
PVD, n (%)	179 (2)	346 (3)	368 (3)
Prior stroke, n (%)	402 (5)	715 (6)	559 (5)
Antiarrhythmic agent, n (%)	2338 (30)	3774 (29)	3393 (27)
eGFR mLs/min/1.73 m^2, median (Q1, Q3)	66 (48,89)	67 (50,86)	65 (50,81)
Serum albumin g/dL, median (Q1, Q3)	3.4 (2.9,3.9)	3.7 (3.2,4.1)	4.0 (3.6,4.3)
Mean PAP mm Hg, median (Q1, Q3)	31 (24,38)	30 (24,38)	28 (21,35)
PCWP mm Hg, median (Q1, Q3)	22 (15,28)	20 (14,27)	18 (13,24)
Cardiac index L/min, median (Q1, Q3)	2.1 (1.7,2.6)	2.1 (1.7,2.5)	2.1 (1.8,2.5)

(Continues)

TABLE 1 (Continued)

	UNOS status 1A N = 7681	UNOS status 1B N = 13 027	UNOS status 2 N = 12 361
PVO$_2$ mL/kg/min, median (Q1, Q3)	11 (9,14)	11 (9,14)	12 (10,14)
Ventilator, n (%)	656 (9)	142 (1)	71 (1)
Inotrope, n (%)	3309 (43)	6571 (50)	688 (6)
LVAD, n (%)	2122 (28)	3315 (25)	313 (3)
RVAD ± LVAD, n (%)	555 (7)	199 (2)	20 (0)
MCS unspecified, n (%)	253 (3)	236 (2)	75 (1)
TAH, n (%)	104 (1)	37 (0)	16 (0)
ECMO, n (%)	321 (4)	13 (0)	15 (0)
IABP, n (%)	1266 (16)	244 (2)	130 (1)

CMP, cardiomyopathy; CAD, coronary artery disease; CO, cardiac output; ECMO, extracorporeal membrane oxygenation; eGFR, estimated glomerular filtration rate; IABP, intra-aortic balloon pump; ICD, implantable cardioverter-defibrillator; IQR, interquartile range; LVAD, left ventricular assist device; MCS, mechanical circulatory support; OHT, orthotopic heart transplant; PAP, pulmonary arterial pressure; PVD, peripheral vascular disease; PVO$_2$, peak oxygen consumption; RVAD, right ventricular assist device; TAH, total artificial heart.
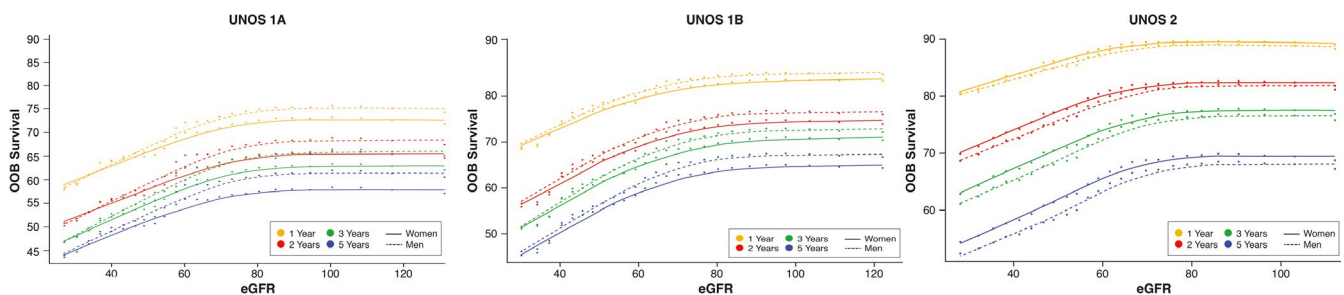
FIGURE 1 Sex differences in heart transplant waitlist survival based on estimate glomerular filtration rate and UNOS status at time of listing. Risk-adjusted relationship of 1-, 2-, 3-, and 5-year survival (OOB = out-of-bag survival) on the transplant waitlist and eGFR at listing for UNOS Status 1A, 1B, and 2 candidates. The shape of all curves is estimated nonparametrically without model assumptions. Note that there is near linear decrease in survival as eGFR falls below 80 mL/min/1.73 m^2, but for eGFR greater than this, there is no relationship of survival to eGFR. This holds for all UNOS statuses. Although eGFR adjusts creatinine levels for age, sex, and race, there remains a notable small interaction of the relation on survival to eGFR with respect to sex, which is also depicted on these curves. eGFR, estimated glomerular filtration rate; UNOS, United Network for Organ Sharing

3.3 | Strong and weak correlates predicting waitlist mortality

The variables at time of listing predicting waitlist mortality based on UNOS Status were ranked by variable importance (Figure 4). In general, eGFR was the most predictive of survival followed by a group of highly predictive variables and others with less ability to predict survival. The strong correlates predicting waitlist survival included albumin, especially among UNOS Status 1A and UNOS Status 2 candidates. Hemodynamics and functional capacity were most predictive of mortality among UNOS Status 2 candidates, but were still important predictors of mortality among those listed as UNOS Status 1A and 1B. Therapy that defined urgent UNOS Status 1A, such as ECMO, mechanical ventilation, and VADs, were highly predictive of mortality and more important than hemodynamics and peak oxygen. Use of VADs was also an important predictor of mortality for UNOS Status 1B and more important risk factors than hemodynamics and peak oxygen. Inotropes and IABP were mostly utilized for urgent status (UNOS Status 1A and 1B) and were

predictive of mortality. However, these therapies were not as important prognostic risk factors as ECMO among UNOS Status 1A candidates, or VADs among UNOS Status 1A or 1B candidates. Finally, diagnosis of heart disease and age were variables of importance for all candidates awaiting transplantation, and were more predictive of mortality than sex or race.

Weak correlates for predicting waitlist mortality included antiarrythmic therapy, history of stroke, history of vascular disease, history of malignancy, and blood type. Despite the fact that blood type and body mass index had low predictive value for UNOS Status 1A candidates when compared to ECMO and eGFR, there were interesting relationships between these variables. In Figure 5, UNOS Status 1A candidates with low BMI had slightly worse survival for any given blood type. When comparing different blood types, blood type A had a better waitlist survival for any given BMI than blood type AB which had better survival than blood type B and O. Some other variables had predictive importance depending on the UNOS Status at time of listing. For instance, implantable defibrillators, history of malignancy, and prior tobacco usage had no predictive value
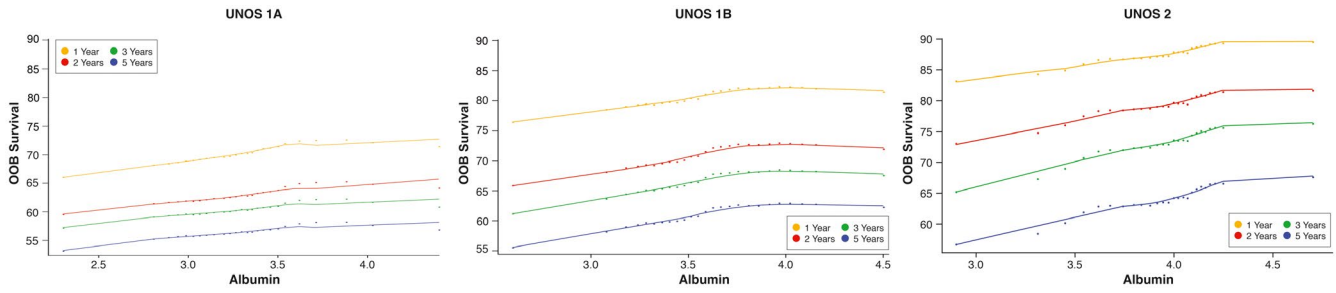
FIGURE 2 Heart transplant waitlist survival based on serum albumin and UNOS status at time of listing. Risk-adjusted relationship of 1-, 2-, 3-, and 5-year survival (OOB = out-of-bag survival) on the heart transplant waiting list and serum album at listing for UNOS Status 1A, 1B, and 2 candidates. The shape of all curves is estimated nonparametrically without model assumptions. Note the near linear decrease in survival with progressively lower albumin levels. UNOS, United Network for Organ Sharing
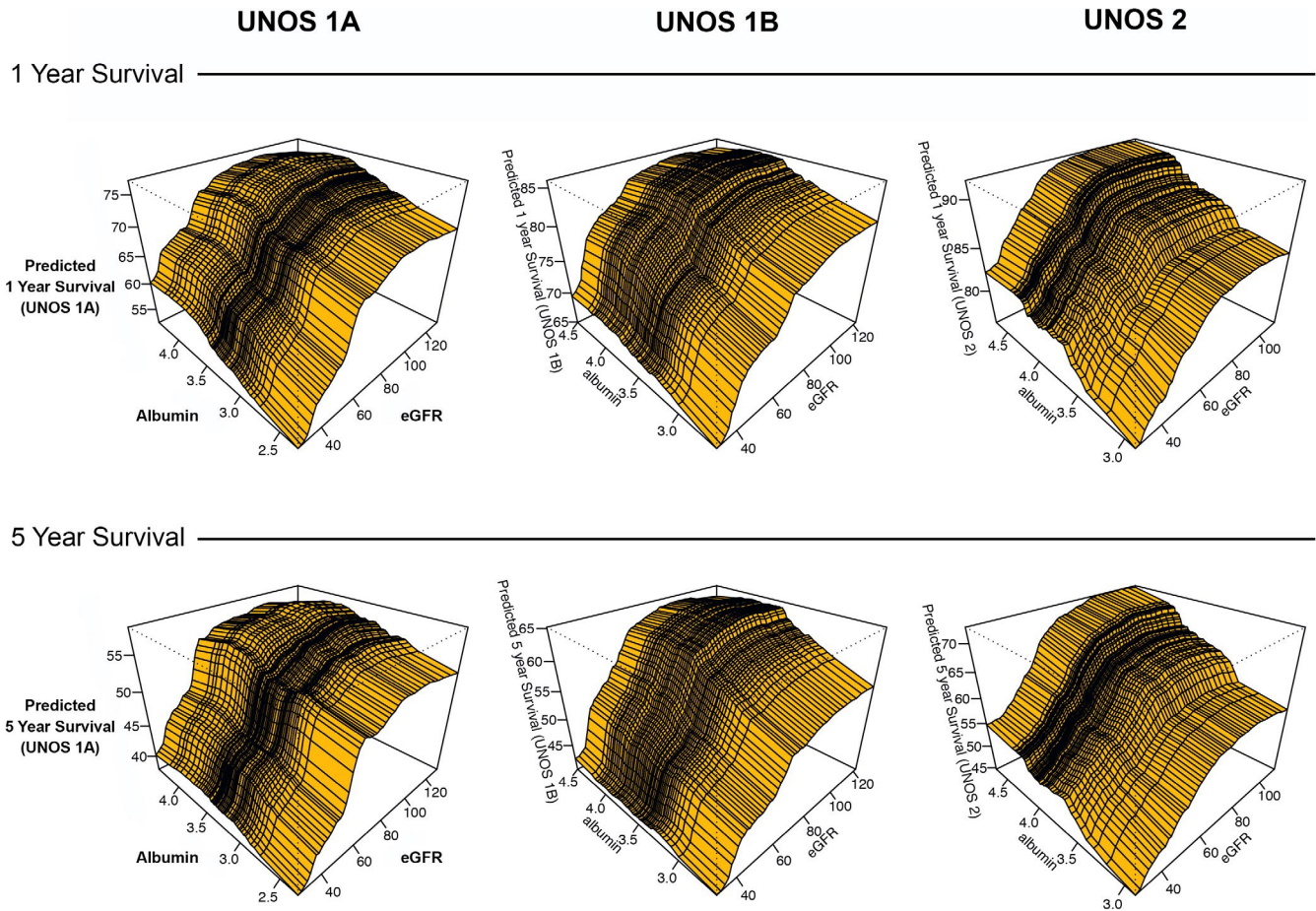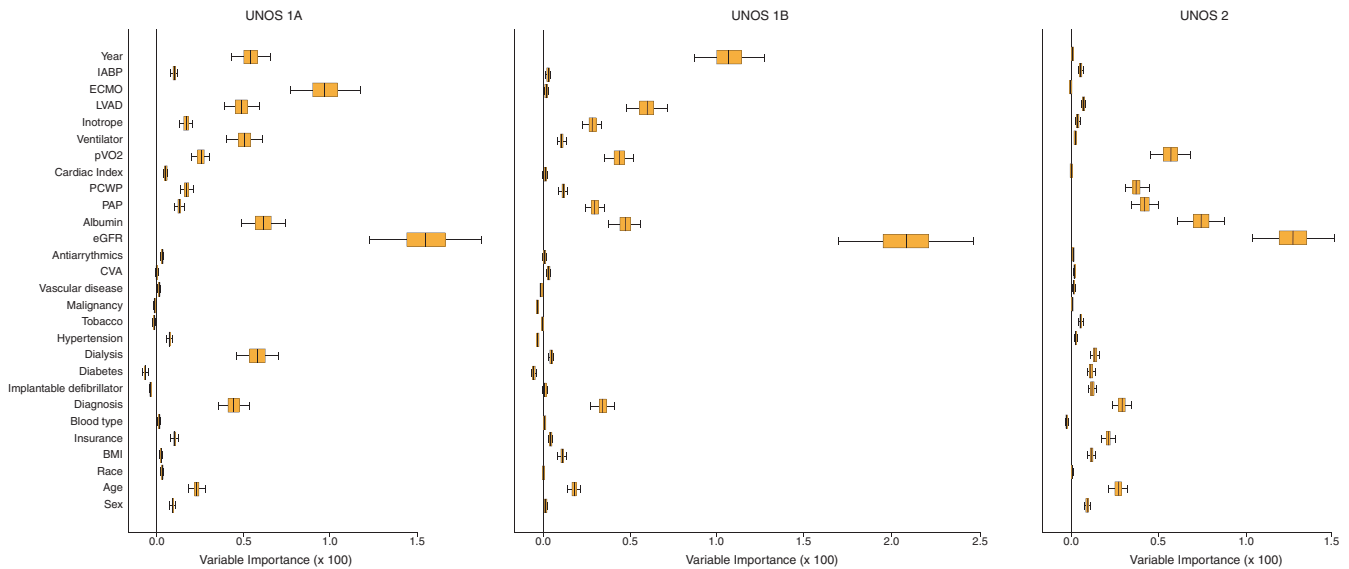

FIGURE 3 Heart transplant waitlist survival based on serum albumin, estimated glomerular filtration rate, and UNOS status at time of listing. Three-dimensional plots using random survival forest analysis were constructed to depict the association between serum albumin, eGFR, and UNOS Status at time of listing. One-year survival is compared to 5-year survival for UNOS Status 1A, 1B, and 2 candidates. eGFR, estimated glomerular filtration rate; UNOS, United Network for Organ Sharing

among those listed at urgent UNOS Status 1A or 1B, but were predictive of mortality among ambulatory UNOS Status 2 candidates. Year of listing for heart transplant was not predictive of mortality among UNOS Status 2 candidates, but was predictive of mortality among UNOS Status 1 candidates. Finally, race was predictive of waitlist mortality among UNOS Status 1A candidates, but not important among UNOS Status 1B or UNOS Status 2 candidates.

4 | DISCUSSION

The THEMIS Investigators (Transplantation of HEarts to MaxImize Survival) were recently awarded NIH support to reduce heart transplant waitlist mortality and minimize organ wastage by identifying risk factors for disparities in survival. Utilizing the SRTR national database we found many strong and some weak correlates predictive

FIGURE 4 Variables of importance predicting heart transplant waitlist mortality. Random Survival Forest investigation of variables predicting heart transplant waitlist mortality based on initially listing candidates as UNOS Status 1A, 1B, or 2. Boxes encompass median (line) and 25th and 75th percentile confidence limits, and whiskers 95% confidence limits. A value of 1.5% as reported for eGFR among UNOS Status 1A candidates means that without eGFR in the survival model we would misclassify 1.5% of new candidates. Thus, given two new candidates we would incorrectly identify which has worse survival on average 1.5% of the time. eGFR, estimated glomerular filtration rate; UNOS, United Network for Organ Sharing

of waitlist mortality. Among the most predictive variables were eGFR, serum albumin, ECMO, VADs, mechanical ventilation, peak oxygen capacity, year, hemodynamic parameters and inotrope support. Weaker correlates for predicting waitlist mortality included antiarrhythmic agents, history of stroke, vascular disease, prior malignancy, history of tobacco, blood type, and race. Our analysis supports usage of the 3-tiered and recently accepted 6-tiered heart allocation system, which prioritizes allocation to those most at risk for waitlist mortality. Our study also emphasizes the fact that other variables like renal function and serum albumin are highly predictive of waitlist mortality and not included in the heart transplant allocation system. Finally, with RSF this analysis rapidly identified variables of importance and can be utilized for assessment of complex interactions.

The current and recently approved tiered allocation system defines medical urgency for heart transplantation based on devices, inotropes, mechanical ventilation, functional capacity, and heart disease. All of these variables were highly predictive of waitlist mortality in our study. In fact, need for ECMO support currently places candidates into the highest tier for transplantation and was more predictive of mortality in our analysis compared to other devices. Inotropes and intra-aortic balloon pumps were less predictive of mortality than VADs and ECMO but still were important. It is possible that candidates in these subgroups had fewer deaths than those having VADs because they were transplanted faster (reduced outcomes) or bridged with a device after a period of time (delayed outcome). Year of waitlist was very predictive of mortality for UNOS Status 1A and 1B but not Status 2 candidates most likely because of the FDA approval over time of smaller and better mechanical
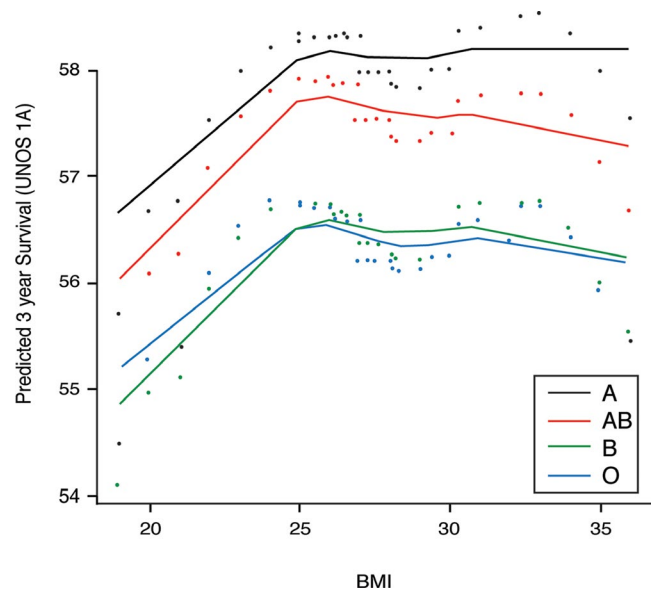
FIGURE 5 Heart transplant waitlist survival based on blood type and body mass index. Risk-adjusted 3-year survival (OOB = out-of-bag survival) on the heart transplant waiting list based on blood type (A, B, AB, O) and body mass index for UNOS Status 1A candidates. The shape of all curves is estimated nonparametrically without model assumptions

circulatory support devices.[11] Most interesting was the discovery that eGFR and serum albumin were more predictive of waitlist mortality than most devices but are not currently among the criteria for heart transplantation.

Renal function is known to be an important predictor of mortality among patients with heart failure.[20-24] In a large meta-analysis that included 43 heart failure survival models predicting mortality, renal function was one of the most predictive variables.[20] eGFR using the MDRD equation was also shown to predict post–heart transplant mortality.[21] Despite these findings, the major issue with using renal function for heart transplant allocation in the future is that it is a dynamic variable affected by diuretic usage and hydration. Nevertheless, dynamic changes in renal function substantially alters instantaneous estimated mortality on the waitlist.[25]

Serum albumin has been shown in several studies to be an independent predictor of mortality.[24,26-30] In a single center study involving 438 patients admitted for acute decompensated heart failure, serum albumin <3.4 g/dL was one of the strongest predictors of 1 year mortality (aHR = 2.05, 95% CI 1.10-3.81, P = .001).[26] Hypoalbuminemia was also predictive of length of postoperative stay and acute renal failure among LVAD recipients.[24,27,28] Serum albumin is a reflection of nutritional state, hepatic synthetic function, catabolic state, inflammation, and protein-losing conditions. The significance of these factors needs further evaluation, but it is interesting to note that a few studies have shown that the predictive power of hypoalbuminemia is independent of the presence of cachexia or malnutrition.[26,27]

Several variables in our study were deemed less predictive of waitlist mortality. Should they be eliminated from the OPTN/UNOS database to make room for more essential variables? Not based on this analysis alone. The decision to retain or eliminate a variable depends on its importance for ensuring a proper donor/recipient match, predictive value for posttransplant outcome, necessity to ensure fair distribution of organs, and any significant interactions with other variable that affect outcome. For instance, blood type and BMI were not important predictors of waitlist mortality in our analysis but are essential for matching donor organs and recipient. History of tobacco use was also not highly predictive of waitlist mortality but has been shown to predict coronary allograft vasculopathy, graft dysfunction, and mortality post–heart transplantation.[31,32] Race was among the weak correlates predicting waitlist mortality in our study but is important for tracking fair distribution of organs and outcome for specific races.[33] Finally, the importance of a variable also depends on its ability to modify the effect of another variable. For instance, sex was modestly predictive of waitlist mortality, but more substantially affected survival when associated with renal function among candidates with eGFR > 40 mL/min/1.73 m^2. Sex has also been shown to interact with many variables of importance affecting waitlist survival including serum albumin, hemodynamics, VAD, and peak oxygen consumption.[11] Among candidates bridged to transplant with VADs, factors affecting posttransplant survival included sex, mechanical ventilation, history of hemodialysis, history of coronary artery bypass surgery, serum creatinine, and serum bilirubin obtained around the time of transplantation.[34] In fact, death is often the result of many factors, making it imperative that we use statistical methods like RSF to handle complex interactions and to create allocation scores.

This analysis has several limitations. The SRTR database is a large national database that is limited to the data elements prospectively entered and subject to human error during data entry. Although each transplant center routinely collects a lot of information and rigorously evaluates potential candidates, only selected data at specific time points are entered into the national database. These data elements do not include natriuretic peptides, serum sodium, heart rate, and blood pressure that are known to have prognostic significance.[2,3,35,36] It also does not include serum bilirubin at time of listing although this variable is collected at time of transplant. In addition, values for known prognostic risk factors like peak oxygen consumption were often not entered into the database. Despite concerns, human error is minimized in SRTR by edit checks, validation of data at time of entry, and internal verification when there are outliers. Potential problems are reviewed by data quality specialists who resolve discrepant data by verifying the information with the involved transplant center. Data quality is further improved by UNOS and CMS audit checks done routinely every few years at every transplant center.[6] Missing data are also less of a concern when using RSF. RSF performs excellently even with heavy missingness and when missing data are not missing completely at random.[14] RSF also can identify complex interactions but cannot determine how much the integrative interaction effect improves prediction performance. However, this limitation does not preclude developing a coarse tiered allocation score derived from the random survival forest prognostic model,[8,9] nor does it prevent using RSF to quickly derive a survival curve for an individual patient.[10] Finally, it is important to mention data quality. As an example, due to lack of standardization, until recently we were not able to utilize panel of reactive antibody in this analysis.

In conclusion, we found many strong and weak correlates predicting heart transplant waitlist mortality in a large national registry with RSF machine learning statistical methods. Most variables highly predictive of waitlist mortality are incorporated into the current and/or new tiered heart allocation system except for eGFR and serum albumin. To create an allocation score for the future, variables not available in the OPTN/UNOS database will need to be compared with existing variables to determine best predictors of waitlist mortality and the factors affecting posttransplant survival. Allocation may also need to vary with demographics, type of heart disease, type of devices, and other factors because complex interactions predicting survival exist. Machine learning technology like RSF can be used to develop a data-driven allocation system similar to what was created for staging of esophageal cancer.[8] RSF is a robust, nonparametric algorithmic statistical method that can identify risk factors without prior knowledge of any possible parametric relationship (linear or nonlinear) to mortality and can handle complex interactions and large amounts of missingness unlike conventional statistical methods like Cox proportional hazards models. Although RSF is not an additive procedure, variables derived from RSF can be used to create an additive allocation score to better predict survival and is worthy of future research.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. Restrictions may apply to the availability of these data based on current data use agreements with HHRI and SRTR.

DISCLOSURE

The authors have no conflicts of interest to disclose as described by the *American Journal of Transplantation*.

REFERENCES

1. Zannad F, McMurray JJ, Krum H, et al. Eplerenone in patients with systolic heart failure and mild symptoms. *N Engl J Med*. 2011;364(1):11-21.

2. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure. A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2013;62:e147-e239.

3. Levy WC, Mozaffarian D, Linker DT, et al. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation*. 2006;113(11):1424-1433.

4. Jasseron C, Legeai C, Jacquelinet C, et al. Prediction of waitlist mortality in adult heart transplant candidates: the candidate risk score. *Transplantation*. 2017;101(9):2175-2182.

5. Organ Procurement and Transplantation Network: Policies. http://optn.transplant.hrsa.gov/PoliciesandBylaws2/policies/pdfs/policy_9.pdf. Accessed January 29, 2017.

6. Leppke S, Leighton T, Zaun D, et al. Scientific Registry of Transplant Recipients: collecting, analyzing, and reporting data on transplantation in the United States. *Transplant Rev (Orlando)*. 2013;27(2):50-56.

7. Levey AS, Coresh J, Greene T, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Intern Med*. 2006;145(4):247-254.

8. Rice TW, Rusch VW, Ishwaran H, Blackstone EH. Cancer of the esophagus and esophagogastric junction: data-driven staging for the seventh edition of the American Joint Committee on Cancer/International Union Against Cancer Cancer Staging Manuals. *Cancer*. 2010;116(16):3763-3773.

9. Rice TW, Gress DM, Patil DT, Hofstetter WL, Kelsen DP, Blackstone EH. Cancer of the esophagus and esophagogastric junction-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017;67(4):304-317.

10. Lu M, Sadiq S, Feaster DJ, Ishwaran H. Estimating individual treatment effect in observational data using random forest methods. *J Comput Graph Stat*. 2018;27(1):209-219.

11. Hsich EM, Blackstone EH, Thuita L, et al. Sex differences in mortality based on united network for organ sharing status while awaiting heart transplantation. *Circ Heart Fail*. 2017;10(6):pii: e003635.

12. Meyer DM, Rogers JG, Edwards LB, et al. The future direction of the adult heart allocation system in the United States. *Am J Transplant*. 2015;15:44-54.

13. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28(1):112-118.

14. Tang F, Ishwaran H. Random forest missing data algorithms. *Stat Anal Data Min*. 2017;10(6):363-377.

15. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.

16. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. Ann. *Appl Stat*. 2008;2:841-860.

17. Ishwaran H, Lu M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival [published online ahead of print 2018]. *Stat Med*. https://doi.org/10.1002/sim.7803.

18. Random Forests for Survival, Regression, and Classification (RF-SRC). R-Package Version 2.6.0. Ishwaran H and Kagalur, U. 2018. http://cran.r-project.org. Accessed March 15, 2018

19. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29(5):1189-1232.

20. Rahimi K, Bennett D, Conrad N, et al. Risk prediction in patients with heart failure: a systematic review and analysis. *JACC Heart Fail*. 2014;2(5):440-446.

21. Habib PJ, Patel PC, Hodge D, et al. Pre-orthotopic heart transplant estimated glomerular filtration rate predicts post-transplant mortality and renal outcomes: an analysis of the UNOS database. *J Heart Lung Transplant*. 2016;35(12):1471-1479.

22. Trivedi JR, Cheng A, Ising M, Lenneman A, Birks E, Slaughter MS. Heart transplant survival based on recipient and donor risk scoring: a UNOS database analysis. *ASAIO J*. 2016;62(3):297-301.

23. Deo SV, Al-Kindi SG, Altarabsheh SE, et al. Model for end-stage liver disease excluding international normalized ratio (MELD-XI) score predicts heart transplant outcomes: evidence from the registry of the United Network for Organ Sharing. *J Heart Lung Transplant*. 2016;35(2):222-227.

24. Cowger J, Sundareswaran K, Rogers JG, et al. Predicting survival in patients receiving continuous flow left ventricular assist devices: the HeartMate II risk score. *J Am Coll Cardiol*. 2013;61(3):313-321.

25. Blackstone EH, Rajeswaran J, Cruz VB, et al. Continuously updated estimation of heart transplant waitlist mortality. *J Am Coll Cardiol*. 2018;72(6):650-659.

26. Uthamalingam S, Kandala J, Daley M, et al. Serum albumin and mortality in acutely decompensated heart failure. *Am Heart J*. 2010;160(6):1149-1155.

27. Go PH, Hodari A, Nemeh HW, et al. Effect of preoperative albumin levels on outcomes in patients undergoing left ventricular device implantation. *ASAIO J*. 2015;61(6):734-737.

28. Kato TS, Kitada S, Yang J, et al. Relation of preoperative serum albumin levels to survival in patients undergoing left ventricular assist device implantation. *Am J Cardiol*. 2013;112(9):1484-1488.

29. Biegus J, Hillege HL, Postmus D, et al. Abnormal liver function tests in acute heart failure: relationship with clinical characteristics and outcome in the PROTECT study. *Eur J Heart Fail*. 2016;18(7):830-839.

30. Horwich TB, Kalantar-Zadeh K, MacLellan RW, Fonarow GC. Albumin levels predict survival in patients with systolic heart failure. *Am Heart J*. 2008;155(5):883-889.

31. Gali K, Spaderna H, Smits JM, Bramstedt KA, Weidner G. Smoking status at time of listing for a heart transplant predicts mortality on the waiting list: a multicenter prospective observational study. *Prog Transplant*. 2016;26(2):117-121.

32. Corbett C, Armstrong MJ, Neuberger J. Tobacco smoking and solid organ transplantation. *Transplantation*. 2012;94(10):979-987.

33. Singh TP, Almond CS, Taylor DO, Milliren CE, Graham DA. Racial and ethnic differences in wait-list outcomes in patients

listed for heart transplantation in the United States. *Circulation*. 2012;125(24):3022-3030.

34. Healy AH, Stehlik J, Edwards LB, McKellar SH, Drakos SG, Selzman CH. Predictors of 30-day post-transplant mortality in patients bridged to transplantation with continuous-flow left ventricular assist devices–an analysis of the International Society for Heart and Lung Transplantation Transplant Registry. *J Heart Lung Transplant*. 2016;35(1):34-39.

35. Fonarow GC, Adams KF Jr, Abraham WT, Yancy CW, Boscardin WJ. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. *JAMA*. 2005;293(5):572-580.

36. Hsich EM, Grau-Sepulveda MV, Hernandez AF, et al. Sex differences in in-hospital mortality in acute decompensated heart failure with reduced and preserved ejection fraction. *Am Heart J*. 2012;163(3):430-437 e431-433.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Hsich EM, Thuita L, McNamara DM, et al. Variables of importance in the Scientific Registry of Transplant Recipients database predictive of heart transplant waitlist mortality. *Am J Transplant*. 2019;19:2067-2076. https://doi.org/10.1111/ajt.15265