JAMA | Original Investigation

Development and Validation of a Risk Score Predicting Death Without Transplant in Adult Heart Transplant Candidates

Kevin C. Zhang, MS; Nikhil Narang, MD; Carine Jasseron, MS, PhD; Richard Dorent, MD; Kevin A. Lazenby, BE; Mark N. Belkin, MD; Jonathan Grinstein, MD; Anoop Mayampurath, PhD; Matthew M. Churpek, MD, MPH, PhD; Kiran K. Khush, MD, MAS; William F. Parker, MD, MS, PhD

IMPORTANCE The US heart allocation system prioritizes medically urgent candidates with a high risk of dying without transplant. The current therapy-based 6-status system is susceptible to manipulation and has limited rank ordering ability.

OBJECTIVE To develop and validate a candidate risk score that incorporates current clinical, laboratory, and hemodynamic data.

DESIGN, SETTING, AND PARTICIPANTS A registry-based observational study of adult heart transplant candidates (aged ≥18 years) from the US heart allocation system listed between January 1, 2019, and December 31, 2022, split by center into training (70%) and test (30%) datasets. Adult candidates were listed between January 1, 2019, and December 31, 2022.

MAIN OUTCOMES AND MEASURES A US candidate risk score (US-CRS) model was developed by adding a predefined set of predictors to the current French Candidate Risk Score (French-CRS) model. Sensitivity analyses were performed, which included intra-aortic balloon pumps (IABP) and percutaneous ventricular assist devices (VAD) in the definition of short-term mechanical circulatory support (MCS) for the US-CRS. Performance of the US-CRS model, French-CRS model, and 6-status model in the test dataset was evaluated by time-dependent area under the receiver operating characteristic curve (AUC) for death without transplant within 6 weeks and overall survival concordance (c-index) with integrated AUC.

RESULTS A total of 16 905 adult heart transplant candidates were listed (mean [SD] age, 53 [13] years; 73% male; 58% White); 796 patients (4.7%) died without a transplant. The final US-CRS contained time-varying short-term MCS (ventricular assist–extracorporeal membrane oxygenation or temporary surgical VAD), the log of bilirubin, estimated glomerular filtration rate, the log of B-type natriuretic peptide, albumin, sodium, and durable left ventricular assist device. In the test dataset, the AUC for death within 6 weeks of listing for the US-CRS model was 0.79 (95% CI, 0.75-0.83), for the French-CRS model was 0.72 (95% CI, 0.67-0.76), and 6-status model was 0.68 (95% CI, 0.62-0.73). Overall c-index for the US-CRS model was 0.76 (95% CI, 0.73-0.80), for the French-CRS model was 0.69 (95% CI, 0.65-0.73), and 6-status model was 0.67 (95% CI, 0.63-0.71). Classifying IABP and percutaneous VAD as short-term MCS reduced the effect size by 54%.

CONCLUSIONS AND RELEVANCE In this registry-based study of US heart transplant candidates, a continuous multivariable allocation score outperformed the 6-status system in rank ordering heart transplant candidates by medical urgency and may be useful for the medical urgency component of heart allocation.

Editorial page 480

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: William F. Parker, MD, MS, PhD, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637 (wparker@uchicago.edu).

JAMA. 2024;331(6):500-509. doi:10.1001/jama.2023.27029

End-stage heart failure kills more than 300 000 US residents each year.[1] Heart transplant is a definitive lifesaving therapy for eligible patients with advanced heart failure, but the treatment is extremely scarce, with less than 4500 transplants performed annually in the US.[2] To make the best use of scarce donor hearts, the US Department of Health and Human Services mandates that the Organ Procurement and Transplantation Network (OPTN) prioritize medically urgent candidates with a high risk of death without transplant.[3,4] The OPTN currently determines medical urgency in heart candidates with a categorical 6-status system, based mainly on treatment intensity.[5] Waiting candidates are assigned a priority rank for each donor organ according to distance from the donor hospital, status, and waiting time. Despite a revision in October 2018 that expanded the number of statuses, the current heart allocation system is susceptible to manipulation from overtreatment and exception requests,[6-8] does not use important laboratory measures of illness severity,[9,10] and has only moderate predictive ability to rank order candidates according to medical urgency.[11] It is also unclear how the categorical 6-status system will be implemented in the forthcoming continuous distribution system, which will require a medical urgency score.[12]

In France, medical urgency is determined with the candidate risk score, a continuous score derived from a multivariable model that mainly relies on laboratory markers of illness severity.[9,13] Since the October 2018 heart allocation revision, these laboratory data are now collected routinely for US heart transplant candidates. We hypothesized that a multivariable predictive model similar to the French Candidate Risk Score (French-CRS) model could more accurately rank order candidates. Our study objective was to develop and validate a US Candidate Risk Score (US-CRS), a parsimonious medical urgency score designed to be part of the forthcoming continuous distribution system for adult donor hearts.

Key Points

Question Which medical urgency system identifies the adult US heart transplant candidates most likely to die without heart transplant?

Findings In this registry-based study of 16 905 heart transplant candidates, the novel US-Candidate Risk Score (US-CRS) based on required clinical and laboratory variables outperformed the current treatment-based categorical allocation system in identifying medically urgent candidates at the highest risk of death without transplantation.

Meaning Findings of this study suggest that the US-CRS may be useful in determining medical urgency, a major factor in the allocation of deceased donor hearts in the US.

Methods

Data Source and Study Population

This study used data from the Scientific Registry of Transplant Recipients (SRTR). The SRTR data system includes data on all donors, wait-listed candidates, and transplant recipients in the US, submitted by the members of the OPTN.

The study population included all adult (aged ≥18 years) heart transplant waiting list candidates initially listed between January 1, 2019, and December 31, 2022 (**Figure 1**). Candidates who were listed simultaneously for other solid organ transplants were included as the heart priority drives allocation in these cases. We used a location-based Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guideline type 2b split-sample development and validation approach, randomly selecting 97 centers (70%) for model training and 41 centers (30%) for model testing. This strategy enables testing for model generalizability across geographical and center-level variability in donor availability, candidate characteristics, and treatment practices.[6,14] This study followed the relevant portions of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. This study is a secondary analysis of a prospective cohort of deidentified data and was granted an exemption by the Institutional Review Board at the University of Chicago Biological Sciences Division to be performed without patient consent.

Primary Medical Urgency Outcome: Death Without Transplant

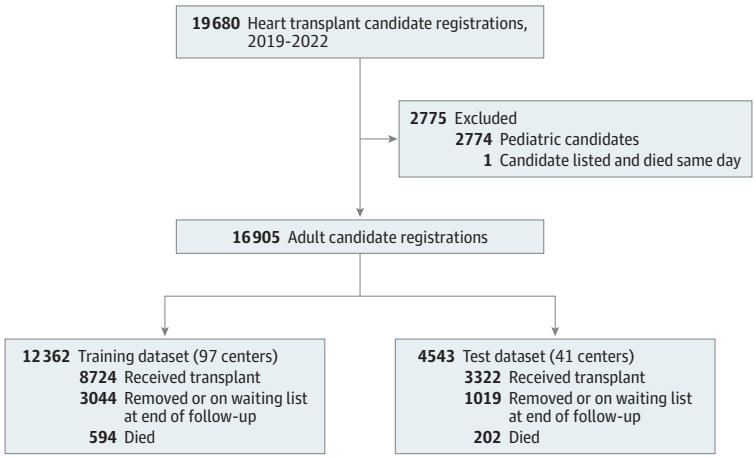
The primary study outcome was death without heart transplant, based on either OPTN-verified death[15] or death from the Social Security Administration death master file. We chose a 6-week prediction window for our primary analysis to include deaths that occur shortly after removal from the waiting list. Consistent with the approach taken in developing the Model for End-Stage Liver Disease with sodium score for liver allocation, patients who did not die before transplant were censored by transplant or last removal or by active or inactive status date on the waiting list.[16] Follow-up data were available through March 2023.

Predictor Variables for US-CRS and Data Structure

Potential predictor variables were obtained from multiple SRTR datasets, which included the candidate registration file and the status justification files. We collected 46 potential predictors (43 which varied over time) in total, which included age, diagnosis, diabetes, hemodynamics, level of inotropic support, mechanical circulatory support (MCS), and MCS complications (eTable 1 in Supplement 1). For the time-varying predictors, we generated treatment history (whether the patient had ever received the treatment), maximum and minimum, and change from the last value variables.

We formatted the data into a discrete-time survival analysis framework that accommodates time-dependent changes in clinical data. As opposed to continuous survival, the discrete-time survival framework breaks each patient's survival time into intervals, which allows for the calculation of the probability of death without transplant in the next 6 weeks, given that the patient has survived to the current interval.[17-20] We chose 14-day intervals because this was the most common treatment, laboratory, and hemodynamic data

Figure 1. Study Cohort From the United Network for Organ Sharing Scientific Registry of Transplant Recipients Files



There were 16 178 unique candidates with 16 905 registrations in the final study cohort.

update frequency.[21] eAppendix 1 in Supplement 1 presents details on dataset construction. Our approach to missing data followed the OPTN approach to calculating Model for End-Stage Liver Disease and Lung Allocation Score[22] by carrying forward values until updated. For missing hemodynamic variables, we performed single imputation with normal expected values[23] for advanced heart failure patients. eTable 2 in Supplement 1 presents imputation procedures.

Statistical Analysis

To create the US-CRS model, we performed a series of unpenalized logistic regressions with a set of preselected investigator-chosen predictor variables. All candidate US-CRS models included the variables in the French-CRS model (short-term MCS, bilirubin, B-type natriuretic peptide [BNP], estimated glomerular filtration rate [eGFR]) and durable left ventricular assist device (LVAD), given the known association of durable LVADs with lowering mortality rates.[24] We used the race-neutral CKD-EPI Creatinine Equation (2021) to estimate eGFR,[25] considered patients receiving dialysis to have an eGFR of 0, and log-transformed BNP and bilirubin. Our definition of short-term MCS included extracorporeal membrane oxygenation (ECMO), temporary surgical LVAD, and biventricular assist device without discharge, and excluded IABP and percutaneous endovascular left ventricular assist devices given the unclear association between the use of these devices and medical urgency.[10,26] We then added additional laboratory results (albumin, sodium, and aspartate aminotransferase) and hemodynamic predictors (cardiac power output [CPO], aortic pulsatility index [API], and pulmonary artery pulsatility index [PAPi]) that have known associations with mortality in patients with advanced heart failure.[10,11,27-31] We also included an interaction term for the model for whether the center reported N-terminal pro-B-type natriuretic peptide (NT-proBNP) or regular BNP. We selected the final US-CRS model based on anticipated ease of implementation, coefficient significance, the area under the receiver operating characteristic curve (AUC), and the value of the Akaike information criterion in the training dataset.[32] To account for each candidate potentially having multiple 2-week periods where the 6-week death outcome occurred, we computed robust standard errors clustered by candidate.

We compared the performance of the current US 6-status system, the French-CRS model, and our final US-CRS model in the test dataset using incident/dynamic time-dependent AUC for 6-week mortality evaluated at each 2-week interval following listing. We also computed survival concordance (c-index) in a continuous time version of the test dataset using the approach of Heagerty and Zhang.[33] We calculated standard errors for each statistic with bootstrapping with replacement at the candidate level. Calibration was tested by calculating the observed and predicted 6-week waiting list mortality by US-CRS deciles. Model calibration for 6-week mortality was assessed numerically by fitting a linear regression to the calibration plot to determine a calibration slope and intercept. To ensure that the US-CRS was not biased against specific sex or racial-ethnic groups, we performed algorithmic fairness tests by creating sex- and race-stratified calibration plots.[34] Race and ethnicity were reported as demographic characteristics based on patient self-report with multiple-choice options. We then constructed a 50-point medical urgency score from the final US-CRS by mapping 50 quantiles, following the procedure used by the OPTN in kidney allocation.[35,36]

To test the robustness of our approach, we developed and tested several alternative prediction models. First, we included IABP and percutaneous VAD in the definition of short-term MCS and refit the final selected US-CRS model. Next, we examined an alternative model that included 2 hemodynamic variables, CPO and PAPi. We then developed 2 models using machine learning methods (elastic net and gradient-boosted machine). We tuned the hyperparameters of each machine learning method using 5-fold cross-validation and hyperparameter tuning to maximize AUC. The procedure details are in eAppendix 2 in Supplement 1.

To assess the dependence of our findings on the discrete-event framework, we also constructed a continuous-time

Table. Selected Characteristics of 16 905 US Heart Transplant Candidates at Listing, Stratified by Training and Test Datasets

Characteristic	Training (n = 12 362 from 97 centers)	Test (n = 4543 from 41 centers)
Age, median (IQR), y	57 (46-63)	56 (45-63)
Sex, No. (%)		
Female	3370 (27)	1162 (27)
Male	8992 (73)	3381 (74)
Race and ethnicity, No. (%)[a,b]		
American Indian or Alaska Native	55 (0.4)	12 (0.3)
Asian	478 (4.0)	161 (3.5)
Black or African American	3132 (25)	1331 (29)
Hispanic or Latino	1277 (10)	483 (11)
Native Hawaiian or Other Pacific Islander	53 (0.4)	<10 (<0.2)
White	7305 (59)	2525 (56)
Other	62 (0.5)	22 (0.5)
6-Status medical urgency, No. (%)[b]	N = 12 218	N = 4516
1	688 (5.6)	245 (5.4)
2	2975 (24)	1229 (27)
3	1274 (10)	534 (12)
4	4333 (35)	1701 (38)
5	527 (4.3)	126 (2.8)
6	2421 (20)	681 (15)
Laboratory measurements[c,d]		
Albumin, mean (SD), g/dL	3.80 (0.64)	3.77 (0.63)
Bilirubin, median (IQR), mg/dL	0.76 (0.50-1.20)	0.80 (0.50-1.20)
eGFR, mean (SD), mL/min/1.73m^2[,e]	56 (28)	57 (28)
Sodium, mean (SD), mEq/L	136.7 (4.2)	136.5 (4.0)
BNP, median (IQR), pg/mL[f]	1036 (332-2779)	1016 (365-2730)
Cardiovascular measurements[g,h]		
Cardiac output, mean (SD), L/min	4.31 (1.33)	4.42 (1.33)
Pulmonary artery pressure, mean (SD), mm Hg		
Systolic	41 (16)	42 (16)
Diastolic	20 (9)	20 (9)
Blood pressure, mean (SD), mm Hg		
Systolic	105 (18)	105 (17)
Diastolic	69 (12)	70 (13)
Central venous pressure, mean (SD), mm Hg	9 (6)	9 (6)
Cardiac power output, mean (SD), W	0.74 (0.28)	0.77 (0.27)
PAPi, mean (SD)	3.66 (4.09)	3.70 (3.81)
Cardiac devices		
Durable LVAD, No. (%)	3179 (26)	1340 (29)
IABP, No. (%)	1281 (10)	495 (11)
Percutaneous VAD, No. (%)	370 (3.0)	169 (3.7)
ECMO, No. (%)	382 (3.1)	130 (2.9)
BiVAD with no discharge, No. (%)	67 (0.5)	35 (0.8)
Temporary surgical LVAD, No. (%)	43 (0.3)	14 (0.3)

Abbreviations: BiVAD, biventricular assist device; BNP, B-type natriuretic peptide; ECMO, extracorporeal membrane oxygenation; eGFR, estimated glomerular filtration rate; IABP, intra-aortic balloon pump; LVAD, left ventricular assist device; PAPi, pulmonary artery pulsatility index; VAD, ventricular assist device.

SI conversion factors: To convert albumin to grams per liter, multiply by 10; bilirubin level to micromoles per liter, multiply by 17.104; sodium to millimoles per liter, multiply by 1; and BNP to nanograms per liter, multiply by 1.0.

[a] Clinically documented, multiple-selection allowed race categories. Race and ethnicity classifications are based on patient self-report with multiple-choice. No further break down of the classification of race and ethnicity as other is available.

[b] Percentages may not add up to 100% due to missing data.

[c] For the training set, N = 10 232; test set, N = 3848.

[d] The reference values for laboratory measurements are albumin: 3.4-5.4 g/dL; bilirubin: <1 mg/dL; eGFR: >90 mL/min/1.73 m^2; sodium: 135-145 mEq/L; and BNP: <100 pg/mL.

[e] Calculated using the race-neutral CKD-EPI Creatinine (2021) equation provided by the National Kidney Foundation and set to 0 if the patient is receiving dialysis.

[f] Includes both regular and N-terminal pro–B-type natriuretic peptide (NT-proBNP) values.

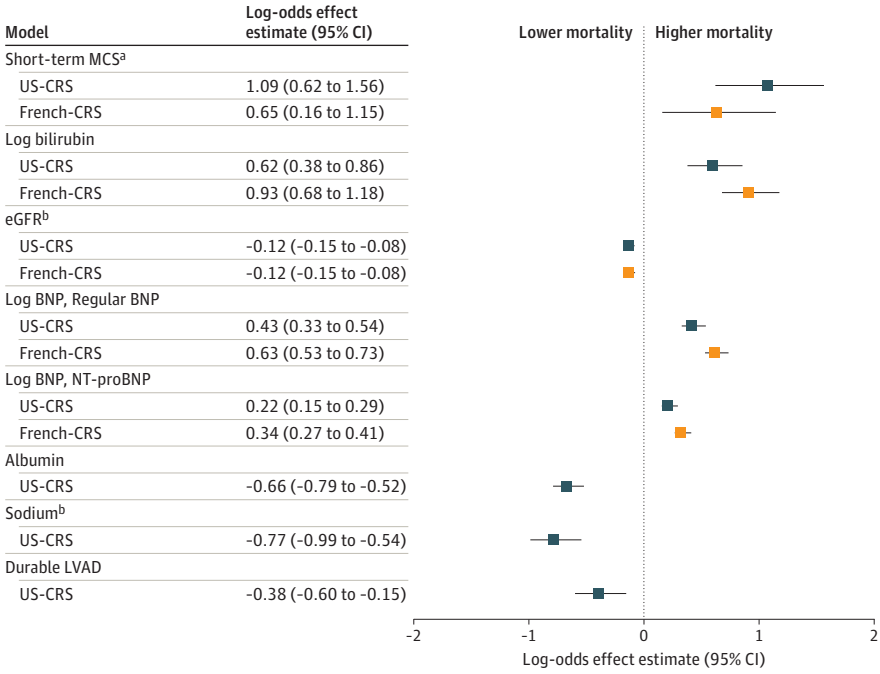
[g] For the training set, N = 9542; test set, N = 3517.

[h] The reference values for cardiovascular measurements are cardiac output: depends on body surface area (BSA), >2.5 times BSA; pulmonary artery systolic pressure: 15-30 mm Hg; pulmonary artery diastolic pressure: 4-12 mm Hg; blood pressure (systolic/diastolic): 120/80 mm Hg; central venous pressure: 8 mm Hg; cardiac power output: 1.0 W; and PAPi: >2.

version of the training and test datasets (each time interval determined by the exact day of data update for each patient). We used the same censoring approach as the discrete-time framework and included deaths up to 6-weeks after removal from the waiting list. We estimated a Cox proportional hazards regression model with time-dependent covariates in the training dataset using the final US-CRS parameters, and evaluated the result using a generalized c-index.[20] We then repeated our analysis with a 2-week death prediction interval so each candidate had at most 1 event.

All analyses were performed using R version 4.2.3 (R Foundation for Statistical Computing). Statistical significance was set at 2-tailed $P < .05$. Complete statistical code necessary to reproduce the study results is available online.[37]

Figure 2. Model Effect Estimates From the French Candidate Risk Score (French-CRS) and the US Candidate Risk Score (US-CRS)



Effect estimates on log-odds for all predictor variables in the US-CRS and current French-CRS, which represent the effect estimate of a unit change in each predictor on 6-week waiting list mortality, or medical urgency, when the other predictors are held constant. The 6-status system is not a multivariable predictor model, so it is not included. The dashed line denotes no effect, dots denote the effect size from the relevant model, and error bars denote 95% CIs. BNP indicates-type natriuretic peptide; eGFR, estimated glomerular filtration rate; LVAD, left ventricular assist device; MCS, mechanical circulatory support; NT-proBNP, N-terminal pro-B-type natriuretic peptide.

[a] The French-CRS version of short-term MCS includes extracorporeal membrane oxygenation only in the current interval. The US-CRS version includes prior or current extracorporeal membrane oxygenation, prior or current temporary surgical left ventricular assist device, and prior or current biventricular assist device without discharge.

[b] Multiplied by a factor of 10, representing a 10-unit change.

Results

Overall Study Population

Between 2019 and 2022, there were 16 905 adult candidates (mean [SD] age, 53 [13] years; 73% male; 0.4% American Indian or Alaska Native, 4% Asian, 26% Black or African American, 10% Hispanic or Latino, 0.4% Native Hawaiian or Other Pacific Islander, 58% White, 0.5% other) listed for heart transplant with a total of 2 854 138 days at risk available for analysis. Among these, 796 (4.7%) candidates experienced the primary outcome of death on the waiting list or within 6 weeks of removal from the waiting list, 12 046 (71.3%) were transplanted, and the remaining 4063 candidates (24.0%) had been removed or were still on the waiting list in March 2023. The most common reason for removal was given as other, followed by condition deterioration and condition improvement. Of the candidates who died after being removed from the waiting list, 183 (44.9%) died within 6 weeks of removal. The median time between events (listing, status change, transplant, death, removal, or last follow-up) was 12 days.

Candidate Characteristics in the Training and Test Datasets

Data from 97 transplant centers (70.3%) served as the training dataset, and data from the remaining 41 (29.9%) as the test dataset. Details of selected variables in both datasets can be found in the **Table**. Most candidates were listed as status 4 (36%) initially, followed by status 2 (25%) and status 6 (19%). Most candidates with short-term MCS were treated with ECMO (82.9%). Bilirubin, BNP, and PAPi were right skewed (eFigure 1 in Supplement 1). The BNP (13.1%), systolic blood pressure (16.4%), and diastolic blood pressure (19.1%) had a large proportion of patients without any recorded values. All other variables had less than 10% of patients without recorded values.

US-CRS and French-CRS Results

Model coefficients of the current French-CRS and the US-CRS are shown on the log-odds scale in **Figure 2**. Our final US-CRS model included prior or current short-term MCS, durable LVAD, log bilirubin, eGFR, albumin, sodium, and a log BNP and NT-proBNP type interaction term. In the US-CRS, higher bilirubin, BNP (both regular and NT-proBNP), and prior or current short-term MCS were associated with higher 6-week mortality, while higher sodium, eGFR, albumin, and durable LVAD were associated with lower 6-week mortality. The associations between predictors were similar to the current French-CRS score. An alternative version of the US-CRS that included hemodynamics had a coefficient for CPO that was not significant (eTable 3 in Supplement 1). Classifying IABP and

percutaneous VADs as short-term MCS reduced the effect size of the term from 1.09 to 0.50 (54%).

Performance of the 6-Status Model, French-CRS, and US-CRS in the Test Dataset

Figure 3 compares AUC values for 6-week mortality in the test dataset. The 6-status model had an AUC for death within 6 weeks of listing of 0.68 (95% CI, 0.62-0.73), the current French-CRS had an AUC of 0.72 (95% CI, 0.67-0.76), and the novel US-CRS had an AUC of 0.79 (95% CI, 0.75-0.83). Plots of time-dependent AUCs for each interval over the first year are shown in eFigure 2 of Supplement 1. Overall c-index was 0.67 (95% CI, 0.63-0.71) for the 6-status model, 0.69 (95% CI, 0.65-0.73) for the current French-CRS, and 0.76 (95% CI, 0.73-0.80) for the US-CRS. eTable 4 in Supplement 1 details the sensitivity and specificity gains of the US-CRS relative to each status threshold. The US-CRS had significantly higher sensitivity and specificity than each of the possible status system allocation thresholds. The largest improvements were a 14.2% absolute sensitivity improvement in status 2 and a 31.8% reduction in the false positive rate in status 5. When comparing the US-CRS with the French-CRS, the US-CRS had numerically higher sensitivity and specificity at all points on the AUC. At a 0.50 sensitivity, the French-CRS had a specificity of 0.78 (95% CI, 0.72-0.86), and the US-CRS had a specificity of 0.84 (95% CI, 0.79-0.88). At a 0.75 sensitivity, the French-CRS had a specificity of 0.61 (95% CI, 0.54-0.68), and the US-CRS had a specificity of 0.70 (95% CI, 0.63-0.76).

The calibration plot for the US-CRS showed good calibration across each decile of the 50-point medical urgency score (eFigure 3 in Supplement 1). Most deciles had no difference between observed and predicted mortality (calibration slope = 0.87 and intercept = 0.001).

Distribution of US-CRS by Medical Urgency Status Category

The distribution of the 50-point US-CRS medical urgency score by 6-status at initial listing is shown in **Figure 4**. For status 1, the median (IQR) score was 48 (43-50), and for status 2, 44 (38-48). Status 3 and status 4 had particularly large ranges, with a median (IQR) of 37 (20-46) for status 3 and 26 (13-39) for status 4. Status 5 had a median (IQR) of 42 (31-47), and status 6 had 30 (19-39). A total of 25.8% of candidates currently categorized at low priority status (statuses 3-6) were at high risk of death (US-CRS >40).

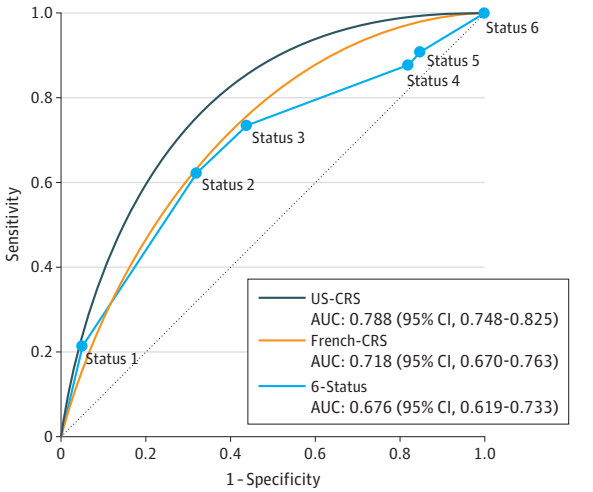
Sensitivity Analyses

When including IABP and percutaneous VAD in the definition of US-CRS, the effect size decreased from 1.09 to 0.50 (eTable 5 in Supplement 1). The results were similar in the alternate model that included hemodynamics.

When refitting US-CRS as a proportional hazards model with time-dependent covariates in the continuous training dataset, the effect sizes of the predictors were similar (eTable 6 in Supplement 1).

Elastic net models and gradient-boosted models had AUC at listing of 0.81 (95% CI, 0.76-0.85) and 0.69 (95% CI, 0.64-0.74) (eFigure 4 in Supplement 1). In the elastic net model, the highest predictor coefficients included ECMO, eGFR, albumin,

Figure 3. Performance of 6-Status, Current French Candidate Risk Score (French-CRS), and US Candidate Risk Score (US-CRS) Models for 6-Week Mortality Without a Transplant at Initial Listing



Receiver operating characteristic curves for 3 models (6-status, current French-CRS, US-CRS) evaluated in the test dataset at listing. AUC represents the area under the curve; higher AUC means the model can better differentiate between those who experience the outcome and those who do not. The dots denote the values of 1-specificity and sensitivity for each 6-status category, and the dotted line represents a model that randomly guesses, with an AUC of 0.5. The 95% CIs were calculated using 2000 stratified bootstrap replicates.

nondischargeable biventricular assist devices, and temporary surgical LVAD, which are predictors contained in the US-CRS. An examination of variable importance also confirmed that the US-CRS predictors had high importance (eFigure 5 in Supplement 1), as albumin, BNP, and eGFR had very high importance in the gradient-boosted model.

An additional analysis to determine whether the US-CRS could lead to gender or racial disparities found that US-CRS calibration plots stratified by sex and race and ethnicity showed good calibration (eFigure 6 in Supplement 1).

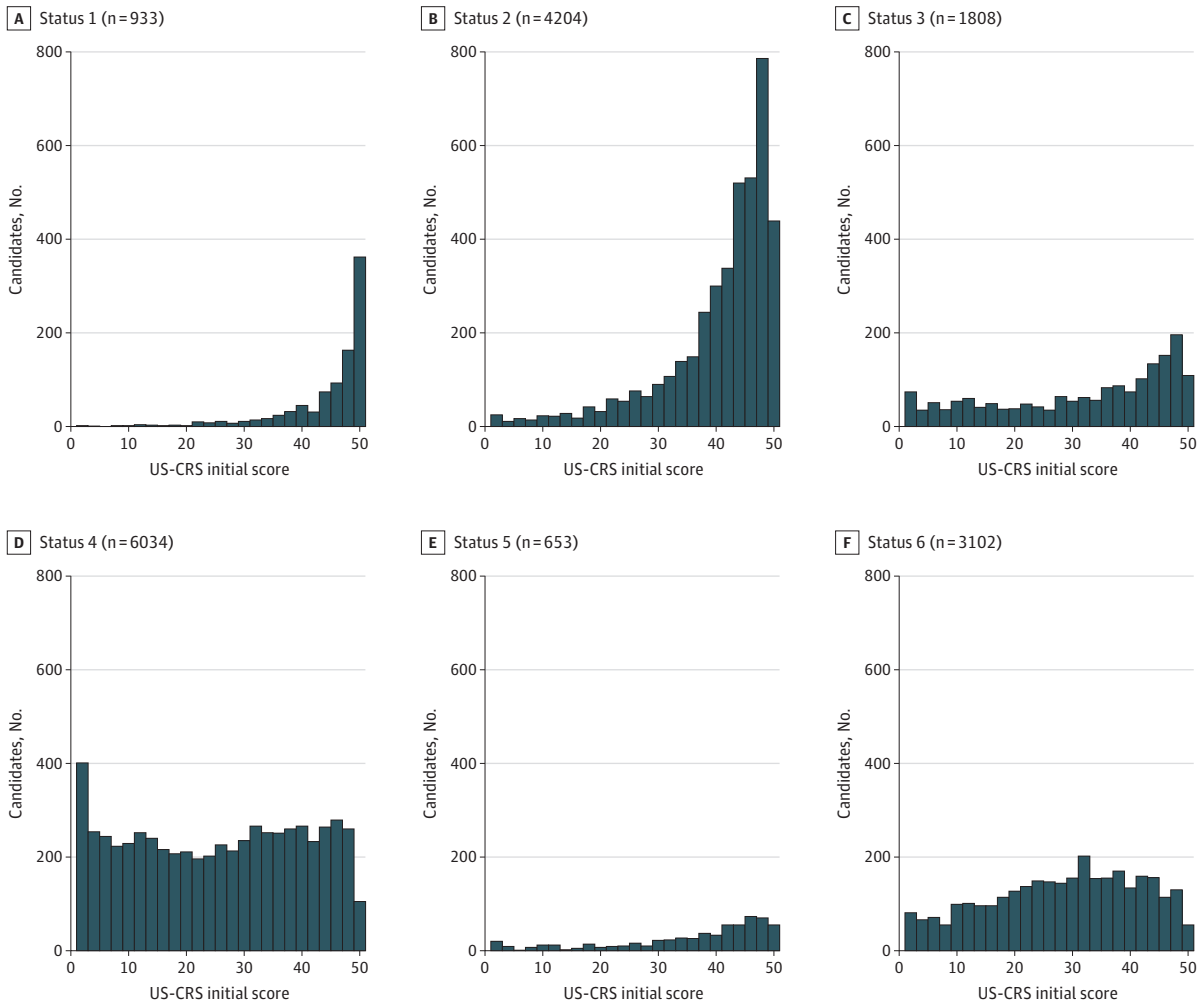
A sensitivity analysis was performed using an alternative longitudinal train-test split strategy. The first 11 834 (70.0%) patients (listed from January 1, 2019, to November 5, 2021) were used to construct the training dataset, while the 5071 (30.0%) recent patients (listed from November 5, 2021, to December 31, 2022) were used to construct the test dataset. We found that the temporal split-derived US-CRS performed well, with an AUC at initial listing of 0.79 (95% CI, 0.77-0.81) compared with AUC of 0.71 (95% CI, 0.69-0.73) for 6-status at listing and AUC of 0.73 (95% CI, 0.70-0.75) for the current French-CRS.

In sensitivity analyses with a 2-week mortality prediction window instead of a 6-week mortality window, we also found similar results (eFigure 7 in Supplement 1).

Discussion

In this study of 16 905 adult heart transplant candidates listed between January 2019 and December 2022, we developed

Figure 4. US Candidate Risk Score (US-CRS) Medical Urgency Score at Listing by 6-Status



The 50-point medical urgency score based on the US-CRS, at initial listing, separated by waiting list status. The score is generated by dividing the predicted log-odds from the US-CRS on the entire 14-day interval dataset into 50 equal quantiles. Lower 6-status and higher 50-point scores represent greater medical urgency. Missing initial status (not shown): 171 (1%).

and validated the US-CRS, a multivariable prediction model that significantly outperformed the current 6-status allocation system in rank ordering candidates by medical urgency. The US-CRS is a parsimonious, easy-to-implement model that captures the key clinical predictors of medical urgency. Machine learning approaches did not outperform the US-CRS. A summary of both the 6-status approach and the US-CRS is shown in eAppendix 3 in Supplement 1. After converting the resulting predicted risk into a 50-point medical urgency score, we found that more than 1 in 4 candidates currently categorized at low priority status (statuses 3-6) were at high risk of death (top decile of risk, US-CRS>40). By assigning these high-risk candidates higher scores and less urgent candidates lower scores, an improvement in AUC in the US-CRS model translates into clinically relevant differences in allocating a scarce resource.

The OPTN is currently transitioning US deceased donor organ allocation to a continuous distribution framework. Continuous distribution will give each donor-candidate pair a composite continuous score that will be based on medical urgency, expected posttransplant survival, placement considerations, and several other factors.[4,38-40] We believe that medical urgency, as estimated by the US-CRS, will be the largest component of the forthcoming continuous distribution score that will need to adhere to the US Department of Health and Human Services mandate to prioritize candidates in order of decreasing medical urgency status.[4] We demonstrate how the US-CRS could be converted into a medical urgency score by mapping quantiles of risk, similar to how OPTN calculates the Kidney Donor Profile Index and estimated posttransplant survival scores. The high calibration accuracy of US-CRS means differences between scores are meaningful. The US-CRS also has significantly higher sensitivity and specificity than all of the 6-status system thresholds, indicating a stronger ability to identify candidates who will die in the short term while also being less likely to

prioritize candidates who will not die. The US-CRS had a smaller but consistently better performance than the French-CRS at likely allocation thresholds. It remains unclear how the current statuses could be converted into a certain number of medical urgency allocation points, and our results show that such an effort would lead to many urgent candidates being assigned inappropriately low priority.

While our aim was to build a prediction model, the association between our predictors and mortality is consistent with physiology and known causal effects. The exception is the increased risk associated with short-term MCS, which likely reflects the unmeasured medical urgency of candidates treated with ECMO rather than a treatment effect. Our decision to follow the French-CRS and not to include IABP and percutaneous VAD in the definition of short-term MCS is supported by our sensitivity analysis. Including these support devices in the definition meant that 43% of short-term MCS would be IABP, which diluted the risk of this population.

The final US-CRS presented does not include hemodynamics. We made this decision because including hemodynamics did not result in major improvements, and hemodynamics are especially susceptible to manipulation. Centers can use varying measurement techniques and manipulation of existing inotropes or vasoactive drugs to obtain worse hemodynamic values for a patient.[6,41] However, if measured fairly, CPO, API, and PAPi, combinations of hemodynamic parameters that represent left and right ventricular function, have potential as they have been shown in other advanced heart failure contexts to be important outcome predictors.[41-46] Future versions of the US-CRS may consider adding these hemodynamic variables under specific contexts to better reflect medical urgency.

The US-CRS would assign patients with durable LVADs lower medical urgency scores because of the powerful treatment effect of the device relative to other medical or mechanical support strategies.[47-49] While durable LVAD candidates will receive low US-CRS and medical urgency scores, LVAD equity points should be added to the patient access component of the continuous distribution score to allow scarce hearts to go to other candidates who cannot receive a durable LVAD.[38,47-49] Similarly, additional points could be added for other groups found to be at a systematic disadvantage in the allocation system, such as blood type O and highly sensitized candidates.[50]

Limitations

Our study has some limitations. First, variables in SRTR are audited by the US Department of Health and Human Services but ultimately center-reported, introducing the potential for recall or misclassification bias. Second, we used only OPTN or Social Security Administration–verified deaths in this analysis, which may underreport deaths.[14,51] Third, the imputation methods we used to conform to other organ allocation scores may have reduced model performance. Fourth, we did not validate US-CRS on an external dataset, but rather with a split by center. This approach is not random at the patient level, ensuring that the US-CRS is not overfit and should work well for a range of candidates. Fifth, because a large number of patients were censored by transplant, informative censoring may bias our results. However, an examination of 100 patients who received transplant and 100 patients who died while waiting, restricted to the first interval, found that of all the US-CRS score components, only eGFR was significantly different between the patients who were transplanted and those who died. Future iterations of the US-CRS, and organ allocation scores in general, may benefit from advanced causal inference approaches such as marginal structural models that can account for informative censoring.[52] In addition, because the choice of predictors was preselected, we may not have included some predictors that will eventually be shown to be important. For example, the French-CRS was shown to have poor discrimination among patients with ventricular tachycardia, an important risk factor.[13] However, the US-CRS score was significantly higher in patients with life-threatening arrhythmias.

Conclusion

In this study of adult heart transplant candidates in the US, we developed a novel US-CRS incorporating prior or current short-term MCS (excluding IABP and percutaneous LVAD), log bilirubin, eGFR, log BNP, NT-pro BNP, albumin, sodium, and current LVAD to predict 6-week mortality for patients on the waiting list. The US-CRS has better discrimination than the current 6-status ranking system used for donor heart allocation in the US, suggesting that it may be useful for ranking patients by medical urgency. The OPTN should consider implementing the US-CRS to quantify medical urgency in the upcoming continuous distribution system for donor heart allocation.

ARTICLE INFORMATION

Accepted for Publication: December 11, 2023.

Author Affiliations: Department of Medicine, University of Chicago, Chicago, Illinois (Zhang, Belkin, Grinstein, Parker); Advocate Heart Institute, Advocate Christ Medical Center, Oak Lawn, Illinois (Narang); Department of Medicine, University of Illinois-Chicago (Narang); Agence de la Biomédecine, Direction Prélèvement Greffe Organes-Tissus, Saint-Denis La Plaine, France (Jasseron, Dorent); Pritzker School of Medicine, University of Chicago, Chicago, Illinois (Lazenby); Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison (Mayampurath); Department of Medicine, University of Wisconsin, Madison (Churpek); Division of Cardiovascular Medicine, Department of Medicine, Stanford University, Stanford, California (Khush); Department of Public Health Sciences, University of Chicago, Chicago, Illinois (Parker); MacLean Center for Clinical Medical Ethics, University of Chicago, Chicago, Illinois (Parker).

Author Contributions: Mr Zhang and Dr Parker had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Zhang, Narang, Khush, Parker.
Acquisition, analysis, or interpretation of data: Zhang, Jasseron, Dorent, Lazenby, Belkin, Grinstein, Mayampurath, Churpek, Parker.
Drafting of the manuscript: Zhang, Narang, Lazenby, Parker.
Critical review of the manuscript for important intellectual content: All authors.
Statistical analysis: Zhang, Jasseron, Lazenby, Mayampurath, Churpek, Parker.
Obtained funding: Parker.
Administrative, technical, or material support: Grinstein, Parker.
Supervision: Narang, Belkin, Churpek, Parker.

Conflict of Interest Disclosures: Dr Narang reported receiving personal fees from Boehringer Ingelheim and Abbott outside the submitted work. Dr Grinstein reported receiving speaking fees from Abbott, Abiomed, Medtronic, and CH Biomedical

outside the submitted work; in addition, Dr Grinstein reported having a patent for Virtual Patient Simulator pending, broadly related to this as it involves hemodynamics for prognostication but this is a separate entity to the current work. Dr Mayampurath reported receiving grants from the National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI) outside the submitted work. Dr Churpek reported grants from NIH R01 grants from the NHLBI and the National Institute of Diabetes and Digestive and Kidney Diseases, R35 grant from the National Institute of General Medical Sciences and grants from the Department of Defense outside the submitted work; in addition, Dr Churpek had a patent 11,410,777 with royalties paid for risk stratification of hospitalized patients and receives less than $5k per year for the above noted patent from the University of Chicago. Dr Parker reported receiving grants from National Institutes of Health R01 LM014263 and grants from Greenwall Foundation outside the submitted work. No other disclosures were reported.

Funding/Support: Dr. Parker has received funding for this study from National Institutes of Health (grant K08HL150291).

Role of the Funder/Sponsor: The NIH had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy of or interpretation by the SRTR or the US government.

Data Sharing Statement: See Supplement 2.

REFERENCES

1. Tsao CW, Aday AW, Almarzooq ZI, et al; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2023 Update: A Report From the American Heart Association. *Circulation*. 2023;147(8):e93-e621. doi:10.1161/CIR.0000000000001123

2. Colvin MM, Smith JM, Ahn YS, et al. OPTN/SRTR 2021 Annual Data Report: Heart. *Am J Transplant*. 2023;23(2)(suppl 1):S300-S378. doi:10.1016/j.ajt.2023.02.008

3. Organ Procurement and Transplant Network. OPTN Policies. Accessed June 21, 2023. https://optn.transplant.hrsa.gov/media/1200/optn_policies.pdf

4. Health Resources and Services Administration (HRSA), Department of Health and Human Services (HHS). Organ procurement and transplantation network: final rule. *Fed Regist*. 2013;78(128):40033-40042.

5. Shore S, Golbus JR, Aaronson KD, Nallamothu BK. Changes in the United States adult heart allocation policy: challenges and opportunities. *Circ Cardiovasc Qual Outcomes*. 2020;13(10):e005795. doi:10.1161/CIRCOUTCOMES.119.005795

6. Parker WF, Anderson AS, Hedeker D, et al. Geographic variation in the treatment of US adult heart transplant candidates. *J Am Coll Cardiol*. 2018;71(16):1715-1725. doi:10.1016/j.jacc.2018.02.030

7. Parker WF, Chung K, Anderson AS, Siegler M, Huang ES, Churpek MM. Practice changes at US transplant centers after the new adult heart allocation policy. *J Am Coll Cardiol*. 2020;75(23):2906-2916. doi:10.1016/j.jacc.2020.01.066

8. Johnson DY, Ahn D, Lazenby K, et al. Association of high-priority exceptions with waitlist mortality among heart transplant candidates. *J Heart Lung Transplant*. 2023;42(9):1175-1182. doi:10.1016/j.healun.2023.05.009

9. 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. doi:10.1097/TP.0000000000001724

10. Hsich EM, Thuita L, McNamara DM, et al; Transplantation of HEarts to MaxImize Survival (THEMIS) Investigators. Variables of importance in the Scientific Registry of Transplant Recipients database predictive of heart transplant waitlist mortality. *Am J Transplant*. 2019;19(7):2067-2076. doi:10.1111/ajt.15265

11. Pelzer KM, Zhang KC, Lazenby KA, et al. The Accuracy of Initial U.S. Heart Transplant Candidate Rankings. *JACC Heart Fail*. 2023;11(5):504-512. doi:10.1016/j.jchf.2023.02.005

12. OPTN. Continuous distribution. Accessed July 10, 2023, https://optn.transplant.hrsa.gov/policies-bylaws/a-closer-look/continuous-distribution

13. Dorent R, Jasseron C, Audry B, et al. New French heart allocation system: Comparison with Eurotransplant and US allocation systems. *Am J Transplant*. 2020;20(5):1236-1243. doi:10.1111/ajt.15816

14. Akintoye E, Shin D, Alvarez P, Briasoulis A. State-level variation in waitlist mortality and transplant outcomes among patients listed for heart transplantation in the US from 2011 to 2016. *JAMA Netw Open*. 2020;3(12):e2028459. doi:10.1001/jamanetworkopen.2020.28459

15. OPTN. Additional data on long-term outcomes incorporated into OPTN living donor data. Accessed June 26, 2023. https://optn.transplant.hrsa.gov/news/additional-data-on-long-term-outcomes-incorporated-into-optn-living-donor-data

16. Kim WR, Biggins SW, Kremers WK, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med*. 2008;359(10):1018-1026. doi:10.1056/NEJMoa0801209

17. Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. *J Am Stat Assoc*. 1988;83(402):414-425. doi:10.1080/01621459.1988.10478612

18. Suresh K, Severn C, Ghosh D. Survival prediction models: an introduction to discrete-time modeling. *BMC Med Res Methodol*. 2022;22(1):207. doi:10.1186/s12874-022-01679-6

19. Singer JD, Willett JB. It's about time: using discrete-time survival analysis to study duration and the timing of events. *J Educ Stat*. 1993;18(2):155-195.

20. Bansal A, Heagerty PJ. A tutorial on evaluating the time-varying discrimination accuracy of survival models used in dynamic decision making. *Med Decis Making*. 2018;38(8):904-916. doi:10.1177/0272989X18801312

21. OPTN. OPTN Policies Effective as of Mar 16 2023. Accessed June 27, 2023. https://optn.transplant.hrsa.gov/media/eavh5bf3/optn_policies.pdf

22. OPTN. Improving Liver Allocation: MELD, PELD, Status 1A, Status 1B. Accessed June 27, 2023. https://optn.transplant.hrsa.gov/media/3idbp5vq/policy-guid-change_impr-liv-alloc-meld-peld-sta-1a-sta-1b_liv.pdf

23. Libby P, Bonow RO, Mann DL, et al, eds. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*. 12th ed. Elsevier; 2022.

24. Slaughter MS, Rogers JG, Milano CA, et al; HeartMate II Investigators. Advanced heart failure treated with continuous-flow left ventricular assist device. *N Engl J Med*. 2009;361(23):2241-2251. doi:10.1056/NEJMoa0909938

25. National Kidney Foundation. CKD-EPI Creatinine Equation (2021). Accessed July 14, 2023. https://www.kidney.org/content/ckd-epi-creatinine-equation-2021

26. Barge-Caballero G, Castel-Lavilla MA, Almenar-Bonet L, et al. Venoarterial extracorporeal membrane oxygenation with or without simultaneous intra-aortic balloon pump support as a direct bridge to heart transplantation: results from a nationwide Spanish registry. *Interact Cardiovasc Thorac Surg*. 2019;29(5):670-677. doi:10.1093/icvts/ivz155

27. Goldwasser P, Feldman J. Association of serum albumin and mortality risk. *J Clin Epidemiol*. 1997;50(6):693-703. doi:10.1016/S0895-4356(97)00015-2

28. Shinagawa H, Inomata T, Koitabashi T, et al. Prognostic significance of increased serum bilirubin levels coincident with cardiac decompensation in chronic heart failure. *Circ J*. 2008;72(3):364-369. doi:10.1253/circj.72.364

29. Smith GL, Lichtman JH, Bracken MB, et al. Renal impairment and outcomes in heart failure: systematic review and meta-analysis. *J Am Coll Cardiol*. 2006;47(10):1987-1996. doi:10.1016/j.jacc.2005.11.084

30. Cadnapaphornchai MA, Gurevich AK, Weinberger HD, Schrier RW. Pathophysiology of sodium and water retention in heart failure. *Cardiology*. 2001;96(3-4):122-131. doi:10.1159/000047396

31. Cowie MR, Mendez GF. BNP and congestive heart failure. *Prog Cardiovasc Dis*. 2002;44(4):293-321. doi:10.1053/pcad.2002.24599

32. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19(6):716-723. doi:10.1109/TAC.1974.1100705

33. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92-105. doi:10.1111/j.0006-341X.2005.030814.x

34. Mitchell S, Potash E, Barocas S, et al. Algorithmic Fairness: Choices, Assumptions, and Definitions. *Annu Rev Stat Appl*. 2020;8:141-163. doi:10.1146/annurev-statistics-042720-125902

35. OPTN. A Guide to Calculating and Interpreting the Kidney Donor Profile Index (KDPI). Accessed June 28, 2023. https://optn.transplant.hrsa.gov/media/j34dm4mv/kdpi_guide.pdf

36. OPTN. A Guide to Calculating and Interpreting the Estimated Post-Transplant Survival (EPTS) Score Used in the Kidney Allocation System (KAS). Accessed June 28, 2023. https://optn.transplant.hrsa.gov/media/1511/guide_to_calculating_interpreting_epts.pdf

37. Zhang KC. Heart-Continuous-Score. Accessed November 24, 2023. https://github.com/kevinz1194/Heart-Continuous-Score

38. Khush KK, Sandhu AT, Parker WF. How to make the transplantation allocation system better. *JACC Heart Fail*. 2023;11(5):516-519. doi:10.1016/j.jchf.2022.11.029

39. Kasiske BL, Pyke J, Snyder JJ. Continuous distribution as an organ allocation framework. *Curr Opin Organ Transplant*. 2020;25(2):115-121. doi:10.1097/MOT.0000000000000733

40. UNOS. Organ distribution. Accessed June 21, 2023. https://unos.org/policy/organ-distribution

41. Saxena A, Garan AR, Kapur NK, et al. Value of hemodynamic monitoring in patients with cardiogenic shock undergoing mechanical circulatory support. *Circulation*. 2020;141(14):1184-1197. doi:10.1161/CIRCULATIONAHA.119.043080

42. Fincke R, Hochman JS, Lowe AM, et al; SHOCK Investigators. Cardiac power is the strongest hemodynamic correlate of mortality in cardiogenic shock: a report from the SHOCK trial registry. *J Am Coll Cardiol*. 2004;44(2):340-348. doi:10.1016/j.jacc.2004.03.060

43. Morine KJ, Kiernan MS, Pham DT, Paruchuri V, Denofrio D, Kapur NK. Pulmonary artery pulsatility index is associated with right ventricular failure after left ventricular assist device surgery. *J Card Fail*. 2016;22(2):110-116. doi:10.1016/j.cardfail.2015.10.019

44. Belkin MN, Kalantari S, Kanelidis AJ, et al. Aortic Pulsatility Index: a novel hemodynamic variable for evaluation of decompensated heart failure. *J Card Fail*. 2021;27(10):1045-1052. doi:10.1016/j.cardfail.2021.05.010

45. Belkin MN, Alenghat FJ, Besser SA, et al. Aortic pulsatility index predicts clinical outcomes in heart failure: a sub-analysis of the ESCAPE trial. *ESC Heart Fail*. 2021;8(2):1522-1530. doi:10.1002/ehf2.13246

46. Belkin MN, Shah J, Neyestanak ME, Burkhoff D, Grinstein J. Should we be using aortic pulsatility index over cardiac power output in heart failure cardiogenic shock? *Circ Heart Fail*. 2022;15(7):e009601. doi:10.1161/CIRCHEARTFAILURE.122.009601

47. Rose EA, Gelijns AC, Moskowitz AJ, et al; Randomized Evaluation of Mechanical Assistance for the Treatment of Congestive Heart Failure (REMATCH) Study Group. Long-term use of a left ventricular assist device for end-stage heart failure. *N Engl J Med*. 2001;345(20):1435-1443. doi:10.1056/NEJMoa012175

48. Mehra MR, Uriel N, Naka Y, et al; MOMENTUM 3 Investigators. A fully magnetically levitated left ventricular assist device - final report. *N Engl J Med*. 2019;380(17):1618-1627. doi:10.1056/NEJMoa1900486

49. Chung K, Parker WF. A bridge to nowhere: the durable left ventricular assist device dilemma in the new heart allocation system. *J Heart Lung Transplant*. 2023;42(1):87-88. doi:10.1016/j.healun.2022.10.012

50. Kim WR, Mannalithara A, Heimbach JK, et al. MELD 3.0: The Model for End-Stage Liver Disease Updated for the Modern Era. *Gastroenterology*. 2021;161(6):1887-1895.e4. doi:10.1053/j.gastro.2021.08.050

51. Wilk AR, Edwards LB, Edwards EB. The effect of augmenting OPTN data with external death data on calculating patient survival rates after organ transplantation. *Transplantation*. 2017;101(4):836-843. doi:10.1097/TP.0000000000001448

52. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-560. doi:10.1097/00001648-200009000-00011