Supplementary Materials

Supplementary Methods:

Study Population

Approval from the Institutional Review Board at the University of Texas Southwestern Medical Center, Dallas, Texas, was obtained before the study was undertaken. “The clinical and research activities being reported are consistent with the Principles of the Declaration of Istanbul as outlined in the 'Declaration of Istanbul on Organ Trafficking and Transplant Tourism".

This study identified a retrospective cohort of adult recipients who underwent HT alone between January 1, 2000, and September 30, 2019, using the OPTN national registry (N=49147). The exclusion criteria included pediatric age group [age<18 years old] (n=7199), multiorgan transplant including SHKT (n=1947), multiorgan listing including waitlisting for a kidney transplant (n=386), the candidates with pre-transplant missing creatinine levels (1439), and eGFR>60 ml/min at listing and pre-transplant (n=18315), see Supplemental Figure S1. The final study cohort included 19861 adult HT recipients with eGFR<60 ml/min/1.73 m^2 at listing and/or pre-transplant. Follow-up information for the cohort was collected through September 30, 2020.

Outcome Variable

The primary outcome, outcome to be predicted, was the development of CRO within one year following HT. Any patient with CRO in this period was labeled as a positive case. An HT patient who did not develop the CRO within one-year post-HT was labeled as a negative case without censoring for death. We performed two separate analyses for death censoring to check the robustness of our prediction against the assumed

outcome definition (more details provided under “Other Statistical Methods” subsection below).

The CRO variables were extracted from the UNOS-STAR Thoracic_Followup_Data file (chronic_dial, renal_tx, creat), reported at post-transplant 6 months, 12 months, and annually afterwards. The one-year post-HT duration was selected based on prior publications showing a steep decline in GFR within the first year, with the downtrend slowing beyond the first year post-HT. [S1,S2]

Predictor Variable Selection

In selecting the best subset of variables to be used for prediction, we relied first on expert knowledge and then on a data-driven approach. In the first phase, 39 variables (see Supplemental Table S3-4) were selected by three domain experts (a transplant nephrologist, cardiologist, and cardiac transplant surgeon) from a broader list of 533 demographic and pre-transplant variables available in the UNOS STAR dataset. The domain experts approved the list of variables for the prediction task based on whether the variable (i) was clinically relevant and (ii) is typically available for use before transplant (hence, practical to be used in decision making). A serum creatinine-based estimated glomerular filtration rate was also calculated using the CKD-EPI equation[S3] and included in the list of predictor variables. Because selecting the optimal subset of 39 variables for the best-performing prediction model required an exhaustive search and was computationally prohibitive, in the second phase, we relied on a data-driven optimization method. Specifically, a black-box optimization method was used to select the final RF model variables.[S4] The details of the algorithm for the optimization method are available for implementation in the open-source library RBFOpt. The selection of

final variables with RBFOpt was computationally efficient and also effective in selecting the best set of variables that will enhance the predictive performance.[S5] The method employed “feature importance score” in ranking a shorter list of variables (features) to be used in the final random forest prediction model (Supplemental Table S3). Note that the feature importance score does not identify the relationship between the independent (e.g., predictor) variables and the outcome (e.g., predicted) variable as typically done for associative modeling in clinical research. In contrast, the feature importance score helps in selecting variables to be used for prediction (the higher the feature importance score, the more important the feature for accurate prediction). We used two serum creatinine and their corresponding eGFR data points (at listing and at transplant, in average three months apart) and eGFR ratio to reflect trend in renal function. If the patient’s dialysis status at listing and/or pre-transplant is "YES" the corresponding eGFR is assumed to be 10 ml/min/1.73 m^2 regardless of serum creatinine level in the final RF model and web-based decision tool calculation.

Predictive Modeling

The prediction task entails the correct classification of patients into positive or negative case, i.e., developing CRO within one year of HT. The significant advantages of an RF model include measurability of variable importance for prediction, handling of a mixture of numerical and categorical variables, and accuracy comparable to other prominent methods.[S6,S7] Random forest is an ensemble method that crowdsources predictions from multiple trained decision trees[1] for a more accurate prediction.

[1] A decision tree, the constituent machine learning algorithm in an RF framework, produces the probability of a class by hierarchically splitting nodes based on independent variables into buckets of values (e.g., a split could be

The RF model was trained within an open-source software kit, Scikit-learn[S8], to identify heart transplant candidates at risk of developing CRO. Because the training of an RF requires first determining the number of iterations (i.e., number of embedded decision trees), number of randomly selected variables in each tree, and the depth of the decision trees (e.g., the number of splits), an optimization approach and performance validation is required to produce the final model to obtain the best performance.

Our final RF model was optimized over a prediction search space of multiple parameters. The search space optimization aimed to achieve higher and lower performance bounds on sensitivity (≥ 80%) and specificity (≥ 50%), respectively. The relatively larger lower bound on specificity may permit higher false-positive predictions resulting in unwanted over-utilization of kidney allografts. However, the inverse approach, misclassifying CRO in patients at risk, would mean higher mortality within the first-year post-transplantation. After an initial parameter tuning, the search was performed by selecting the best-performing model by training RFs with 64 to 512 decision trees (in increments of 32), 4 to 40 predictor variables (using the RBFOpt based optimization algorithm), 1 to 4 features to consider when looking for the best split, 3 to 9 tree height (in increments of 1), and using Gini index as a splitting criterion.[S9] Because the data has a low base rate (3.9% of the study cohort developed CRO), known as unbalanced classification, more weight was assigned to the positive class. In the ensemble step, models with high accuracy were given more weight in deciding the

"eGFR <= 45 ml/min/1.73 m2) until a leaf node with a class label prediction is reached. The collection of splits used in reaching the leaf node constitutes the rule for a final probability assessment of the outcome variable, the prediction. The choice and order of nodes and splits used in a decision tree led to variation in the collection of rules, the associated prediction, and the overall performance.

final prediction. The feature importance score was calculated and reported for the final random-forest model variables.

Model Validation

To avoid overfitting and prevent data leakage when searching for the parameters of an optimal model, cross-validation was used. In this study, the standard ten-fold cross-validation[S10] was employed in measuring the model performance. Under ten-fold cross-validation, the dataset was divided into ten non-overlapping cohorts, with each cohort having a similar proportion of positive subjects. Based on cross-validated test subjects, the area under the curve (AUC) for Receiver Operating Characteristics (ROC) or C-statistic was calculated.[S11] A web-based decision tool was created online at https://faculty.tamuc.edu/mmete/esrd-risk.html for public use of the calculator. The time selection (within three months) between current and previous serum creatinine and eGFR in the web-based decision tool was based on observed median waitlisted time (87 days) among HT patients.

Other Statistical Methods:

Recipient characteristics were described as mean (± standard deviation) or median (interquartile range-IQR) for continuous variables and percent total for categorical variables. Comparisons between groups were made using the student t-test, Mann Whitney U test, one-way ANOVA, or the Kruskal-Wallis test by ranks for continuous variables, and Chi-squared tests categorical variables, as appropriate. Survival curves were plotted using the Kaplan-Meier method.

To show the robustness of our model, we conducted two separate analyses. First, we excluded patients who died in the no-CRO group, retrained the new sample's RF model, and reported its AUC performance. In the second analysis, we treated the death event in the first year as a competing event to CRO occurrence in the original study cohort, developed a random survival forest using randomForestSRC package in R (available in https://cran.r-project.org/web/packages/randomForestSRC/), and reported the accuracy of CRO prediction at one-year.[S12]

A *P*-value<0.05 was considered statistically significant. Statistical analyses were performed with Stata/MP16 (StataCorp LP, College Station, TX) and Python 3.7 Data Science Platform (Anaconda).

Handling missing data:

To account for missing continuous data, we imputed two variables: cardiac output (n=1390, 7.0%) and pulmonary artery mean pressure (n=1410, 7.1%) using multiple linear regression imputation methods (a flexible, 3-step simulation-based statistical method derived using the Bayesian paradigm) by generating five imputed datasets (can be accessed at https://www.stata.com/manuals/mi.pdf). A variable “unknown/missing data” is generated for VAD status (n=769, 3.9%) and functional status (n=1476, 7.4%) pre-transplant.

Supplementary References:

S1. Lachance K, White M, Carrier M, et al. Long-term evolution, secular trends, and risk factors of renal dysfunction following cardiac transplantation. *Transpl Int*. Aug 2014;27(8):824-37. doi:10.1111/tri.12340

S2. Lindelow B, Bergh CH, Herlitz H, Waagstein F. Predictors and evolution of renal function during 9 years following heart transplantation. *J Am Soc Nephrol*. May 2000;11(5):951-7.

S3. Levey AS, Stevens LA. Estimating GFR using the CKD Epidemiology Collaboration (CKD-EPI) creatinine equation: more accurate GFR estimates, lower CKD prevalence estimates, and better risk predictions. *Am J Kidney Dis*. Apr 2010;55(4):622-7. doi:10.1053/j.ajkd.2010.02.337

S4. Costa A, Di Buccio E, Melucci M, Nannicini G. Efficient parameter estimation for information retrieval using black-box optimization. *IEEE Transactions on Knowledge and Data Engineering*. 2017;30(7):1240-1253.

S5. Costa A, Nannicini G. RBFOpt: an open-source library for black-box optimization with costly function evaluations. *Math Program Comput*. Dec 2018;10(4):597-629. doi:10.1007/s12532-018-0144-7

S6. Breiman L. Random Forests. journal article. *Machine Learning*. October 01 2001;45(1):5-32. doi:10.1023/a:1010933404324

S7. Deo RC. Machine Learning in Medicine. *Circulation*. Nov 17 2015;132(20):1920-30. doi:10.1161/CIRCULATIONAHA.115.001593

S8. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research*. 2011;12(Oct):2825-2830.

S9. Gastwirth JL. The Estimation of the Lorenz Curve and Gini Index. *The Review of Economics and Statistics*. 1972;54(3):306-16.

S10. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*. Aug 1 2005;21(15):3301-7. doi:10.1093/bioinformatics/bti499

S11. Green DM, Swets JA. *Signal detection theory and psychophysics*. vol 1. Wiley New York; 1966.

S12. Ishwaran H, Lu M. Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Stat Med*. Feb 20 2019;38(4):558-582. doi:10.1002/sim.7803

Supplementary Tables:

Supplemental Table S1. Pre-transplant distribution of kidney function and dialysis status of recipients of adult heart transplant alone (excluding all multiorgan transplants, the ones waitlisted for other organs, pediatric group, and missing creatinine values at transplant) between 2000 and 2019 in the U.S.

N= 39,205	N (%)
Kidney function and dialysis status	
eGFR>60 ml/min/1.73 m^2	23579 (61.7)
eGFR 30-60 ml/min/1.73 m^2	12759 (33.4)
eGFR <30 ml/min/1.73 m^2 and not on dialysis	912 (2.4)
On dialysis	955 (2.5)

Supplemental Table S2. Characteristics of the deceased donors of the study cohort between 2000 and 2019 in the U.S.

Deceased donors*	**Whole cohort**	**No-CRO**	**CRO**	**P value****
N (%)	19,845*	19,062 (96.1)	783 (3.9)	
Age, median (IQR) years	31 (22, 42)	30 (22, 42)	32 (24, 44)	<0.001
Sex (male)	14,151 (71.3)	13,619 (71.4)	532 (67.9)	0.03
Race				
White	13,238 (66.7)	12,722 (66.7)	516 (65.9)	0.63
Black	2,850 (14.4)	2,736 (14.4)	114 (14.6)	
Hispanic	3,178 (16.0)	3,042 (16.0)	136 (17.4)	
Asian	306 (1.5)	297 (1.6)	9 (1.2)	
Other	273 (1.4)	265 (1.4)	8 (1.0)	
History of Hypertension				
No	16,794 (84.6)	16,159 (84.8)	635 (81.1)	0.002
Yes	2,938 (14.8)	2,791 (14.6)	147 (18.8)	
Unknown	113 (0.6)	112 (0.6)	1 (0.1)	
History of Diabetes				
No	19,209 (96.8)	18,457 (96.8)	752 (96.0)	0.35
Type I	331 (1.7)	313 (1.6)	18 (2.3)	
Type II	305 (1.5)	292 (1.5)	13 (1.7)	
Body mass index (kg/m^2)	27.2 ±5.8	27.2 ±5.8	27.5 ±5.7	0.72
Cause of Death				
Anoxia	4,349 (21.9)	4,143 (21.7)	206 (26.3)	0.01
Cerebrovascular/Stroke	4,261 (21.5)	4,082 (21.4)	179 (22.9)	
Head trauma	10,700 (53.9)	10,323 (54.2)	377 (48.2)	
CNS Tumor	141 (0.7)	138 (0.7)	3 (0.4)	
Other	394 (2.0)	376 (2.0)	18 (2.3)	
HLA Mismatch				
0	22 (0.1)	20 (0.1)	2 (0.3)	0.58
1-4	7141 (40.8)	6873 (40.9)	268 (38.0)	
5-6	10330 (59.1)	9894 (58.9)	436 (61.8)	
KDPI %, median (IQR)	20 (8, 39)	20 (8, 39)	20 (10, 40)	0.01
Cold Ischemia Time, median (IQR)	3.2 (2.5, 3.9)	3.2 (2.5, 3.8)	3.3 (2.5, 4.0)	0.42

*16 donor information is missing.**P value was calculated based on comparison of No-CRO and CRO groups.

Supplemental Table S3. Variable (feature) ranking in the final random forest prediction model based on the feature importance score (the higher the score, the more important the feature is for accurate prediction).

Random Forest Feature Importance Rank	Feature	Feature Importance Score
1	eGFR (ml/min/1.72 m^2) pre-transplant according to CKD-EPI equation	1.61
2	eGFR (ml/min/1.72 m^2) at listing according to CKD-EPI equation	0.74
3	eGFR ratio (eGFR pre-transplant/ eGFR at listing)	0.70
4	Dialysis pre-transplant (Yes, No)	0.66
5	Cardiac index (L/min/m^2) pre-transplant	0.54
6	BMI (kg/m^2) pre-transplant	0.51
7	Functional status (Karnofsky scale: 80-100%, 51-79%, 0-50%) pre-transplant	0.41
8	Pulmonary capillary wedge pressure (mmHg) pre-transplant	0.33
9	Diabetes (no, type I, type II)	0.31
10	Pulmonary artery mean pressure (mmHg) pre-transplant	0.30
11	Age (year)	0.29
12	Race/Ethnicity (White, Black, Hispanic, Asian, Other)	0.24
13	UNOS Region (1-11) *	0.18
14	VAD status (None, LVAD, RVAD/BiVAD/TAH) pre-transplant	0.18
15	Dialysis status at listing (Yes, No)	0.12

Abbreviations: eGFR= estimated glomerular filtration rate; CKD-EPI= Chronic Kidney Disease-Epidemiology Collaboration; BMI= body mass index; *UNOS= United Network of Organ Sharing, (https://unos.org/community/regions/); VAD=Ventricular assist device; LVAD= left ventricular assist device; RVAD= right ventricular assist device; BiVAD= Biventricular assist device; TAH=total artificial heart.

Supplemental Table S4. The variables (pre-transplant recipient characteristics, total of 39) selected by the domain experts from the UNOS-STAR Dataset (Thoracic_Data).

1. Age at listing
2. Sex at listing
3. Ethnicity / Race at listing
4. History of diabetes at listing
5. Highest education level at listing
6. Previous heart transplant at listing
7. Primary insurance at listing
8. History of cigarette use at listing
9. Primary diagnosis at listing
10. Symptomatic cerebrovascular disease at listing
11. ABO blood type at listing
12. PRA at listing
13. UNOS Region at listing
14. Serum creatinine (mg/dl) at listing and pre-transplant
15. eGFR (ml/min/1.72 m^2) according to the CKD-EPI equation at listing and pre-transplant
16. Dialysis at listing and pre-transplant
17. Functional status (Karnofsky score, %) at listing and pre-transplant
18. Weight (kg) at listing and pre-transplant
19. Height (cm) at listing and pre-transplant
20. BMI (kg/m^2) at listing and pre- transplant
21. Cardiac output (L/min) at listing and pre-transplant
22. Previous cardiac surgery (non-heart transplant) at listing and pre-transplant
23. Mechanical ventilation at listing and pre-transplant
24. ECMO at listing and pre-transplant
25. IABP at listing and pre-transplant
26. ICD at listing and pre-transplant
27. VAD type at listing and pre-transplant
28. Pulmonary capillary wedge pressure, mmHg at listing and pre-transplant
29. Pulmonary artery systolic pressure at listing and pre-transplant

30. Pulmonary artery diastolic pressure at listing and pre-transplant
31. Pulmonary artery mean pressure at listing and pre-transplant
32. Cardiac index at listing and pre-transplant (calculated from cardiac output and body surface area)
33. Infection requiring IV antibiotics two weeks pre-transplantation
34. Patient location (ICU, non-ICU, not hospitalized) pre-transplant
35. EBV serostatus pre-transplant
36. CMV serostatus pre-transplant
37. HCV status pre-transplant
38. HIV status pre-transplant
39. Bilirubin level pre-transplant

Supplemental Table S5. Characteristics of the adult heart transplant alone patients who died within first year of transplantation in the study cohort (N=2525).

	Whole Study Cohort	**Deaths within one-year post transplant in No-CRO group**	**Deaths within one-year post-transplant in CRO group**	***P*-value***
N (%)	19861	2298/19708 (11.7)	227/783 (29.0)	<0.001
Age (years), median (IQR) Mean ± SD	59 (52, 64) 56.5 ± 10.4	59 (52, 64) 56.6 ± 11.0	60 (54,64) 57.6 ± 10.1	0.19
Sex (male)	14761 (74.3)	1681 (73.2)	171 (75.3)	0.48
Race				0.38
White	14395 (72.5)	1634 (71.1)	154 (67.8)	
Black	3415 (17.2)	401 (17.5)	48 (21.2)	
Hispanic	1302 (6.6)	176 (7.7)	14 (6.2)	
Asian	524 (2.6)	59 (2.6)	9 (4.0)	
Other	225 (1.1)	28 (1.2)	2 (0.9)	
Recipient height (cm)	173.9 ± 9.8	172.7 ± 10.1	172.9 ± 9.8	0.75
Recipient weight (kg)	83.4 ± 17.4	82.7 ± 17.9	84.5 ± 17.2	0.15
Body mass index (kg/m^2)	27.5 ± 4.8	27.6 ± 5.0	28.2 ± 5.0	0.10
History of diabetes				<0.001
No	13932 (70.2)	1603 (69.8)	121 (53.3)	
Type I	382 (1.9)	44 (1.9)	8 (3.5)	
Type II	5547 (27.9)	651 (28.3)	98 (43.2)	
Etiology of cardiomyopathy				0.34
Ischemic	7815 (39.4)	780 (33.9)	80 (35.2)	
Non-ischemic	7433 (37.4)	938 (40.8)	87 (38.3)	
Congenital	416(2.1)	83 (3.6)	4 (1.8)	
Other**	4197 (21.1)	497 (21.6)	56 (24.7)	
Previous heart transplant, n (%)	695 (3.5)	119 (5.2)	12 (5.3)	0.95
Cardiac index, $L/min/m^2$	2.31 ± 0.70	2.31 ± 0.71	2.28 ± 0.62	0.59
Pulmonary capillary wedge pressure, mmHg	18.8 ± 8.6	19.3 ± 8.7	18.5 ± 8.6	0.22
Pulmonary artery mean pressure, mmHg	28.3 ± 9.9	29.1 ± 10.0	28.2 ± 9.3	0.21
Mechanical ventilation requirement, %	462 (2.3)	160 (7.0)	12 (5.3)	0.34
ECMO, %	207 (1.0)	38 (1.7)	7 (3.1)	0.12
IABP, %	1536 (7.7)	186 (8.1)	21 (9.3)	0.54
VAD, %				0.75

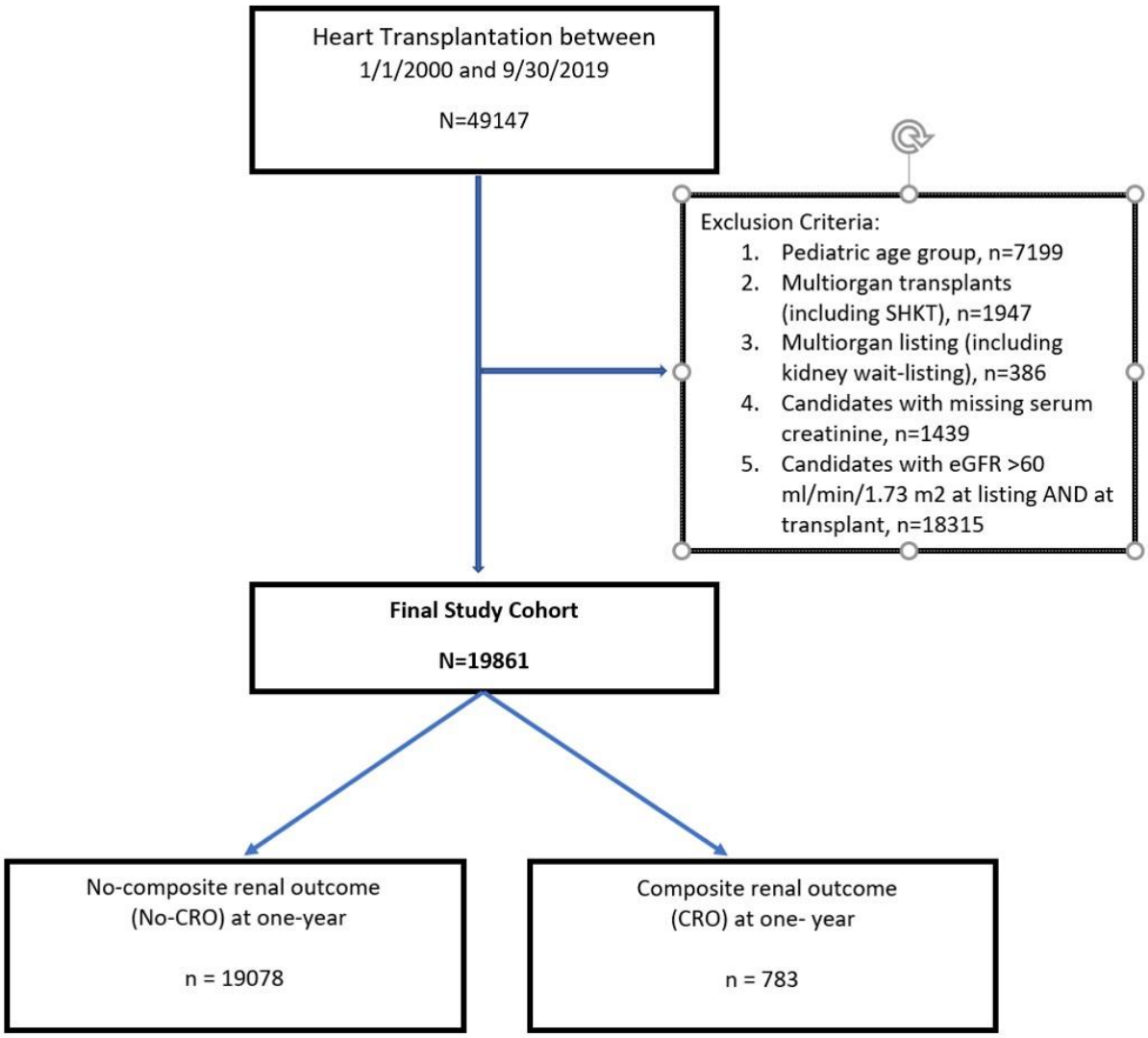
None	11698 (58.9)	1200 (52.2)	120 (52.9)	
LVAD	6530 (32.9)	783 (34.1)	76 (33.5)	
RVAD/BiVAD/TAH	864 (4.4)	192 (8.4)	22 (9.7)	
Unknown	769 (3.9)	123 (5.4)	9 (4.0)	
eGFR ml/min/1.73 m^2 at listing (if not on dialysis)	54.3 ± 17.6	53.7 ± 19.6	51.3 ± 16.8	0.08
eGFR ml/min/1.73 m^2 pre-transplant (if not on dialysis)	53.3 ± 17.6	50.9 ± 19.8	48.0 ± 18.6	0.06
eGFR ratio (transplant/wait listing)	1.10 ± 0.90	1.10 ± 1.12	1.01 ± 0.50	0.29
Dialysis at listing, n (%)	370 (1.9)	88 (3.8)	12 (5.3)	0.28
Dialysis at transplant, n (%)	1038 (5.2)	240 (10.4)	37 (16.3)	0.01
Functional status by Karnofsky score, %				0.003
80-100	2825 (14.2)	253 (11.0)	22 (9.7)	
51-79	9062 (45.6)	1096 (47.7)	85 (37.4)	
0-50	7974 (40.2)	949 (41.3)	120 (52.9)	
UNOS Region				0.05
1	989 (4.5)	100 (4.4)	10 (4.4)	
2	2371 (11.9)	319 (13.9)	30 (13.2)	
3	2233 (11.4)	252 (11.0)	26 (11.5)	
4	2340 (11.8)	297 (12.9)	25 (11.0)	
5	3136 (15.8)	316 (13.8)	42 (18.5)	
6	656 (3.3)	61 (2.7)	3 (1.3)	
7	1900 (9.6)	190 (8.3)	18 (7.9)	
8	1107 (5.6)	121 (5.3)	9 (4.0)	
9	1202 (6.1)	147 (6.4)	27 (11.9)	
10	1626 (8.2)	202 (8.8)	18 (7.9)	
11	2392 (12.0)	293 (12.8)	19 (8.4)	

Abbreviations: CRO= composite renal outcome; ECMO= extracorporeal membrane oxygenation; IABP= Intra-aortic balloon pump; VAD= ventricular assist device; LVAD= left ventricular assist device; BiVAD= biventricular assist device; TAH= total artificial heart; eGFR= estimated glomerular filtration rate; UNOS= United Network of Organ Sharing.

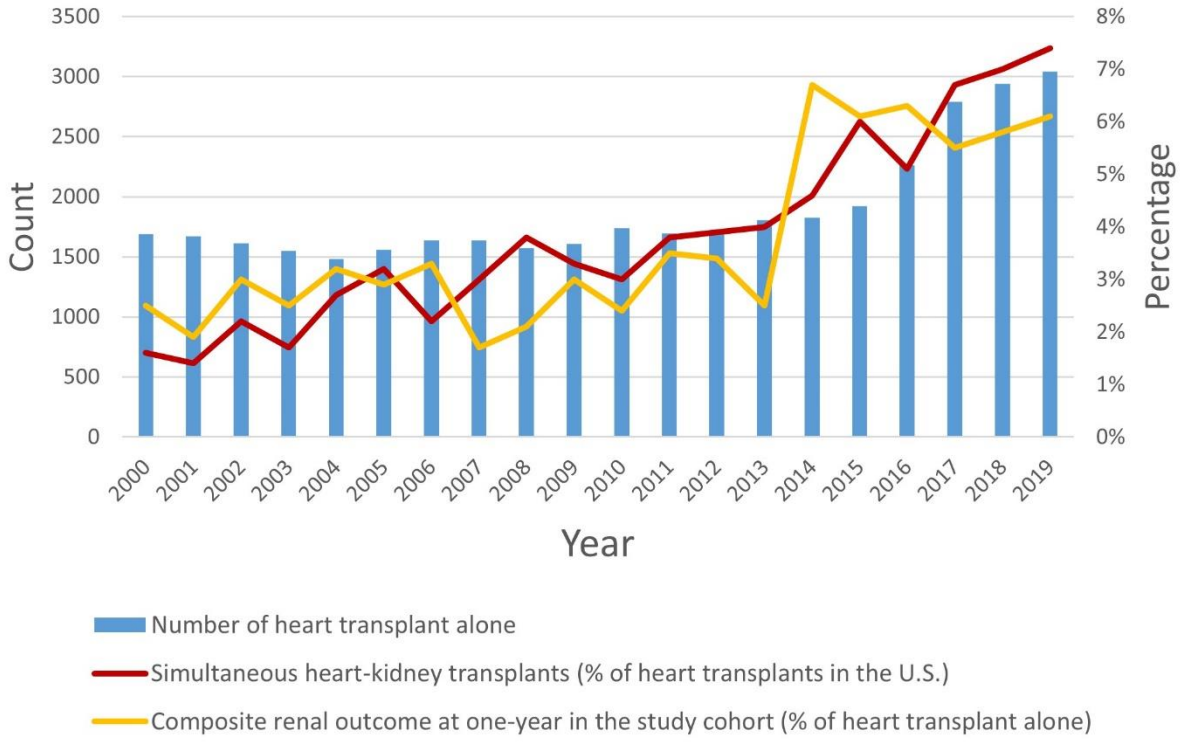
Data are presented as number (%), median (interquartile range) as appropriate. *p-value applies to the comparison of no-CRO and CRO groups. **Other: restrictive cardiomyopathy, congenital, arrhythmia, valvular, and heart transplant-related diagnosis. *** Previous cardiac surgery; CABG, valve replacement/repair, congenital, LV remodeling, other non-transplant surgeries.

Supplementary Figures:

Supplemental Figure S1. Flow chart showing the final study cohort selection (SHKT= simultaneous heart-kidney transplant; eGFR= estimated glomerular filtration rate).



Supplemental Figure S2: The number of adult heart transplants, the incidence of simultaneous heart kidney transplants, and the composite renal outcome (defined as dependence on chronic dialysis, estimated glomerular filtration rate [eGFR] < 20 ml/min/1.73 m^2, or received kidney transplantation) at one-year in the United States between 2000 and 2019.



Supplemental Figure S3. C-statistic for the ten-fold cross-validation study cohort.