Published in final edited form as:
Transplantation. 2019 April ; 103(4): 764–771. doi:10.1097/TP.0000000000002362.

Prospective Validation of Prediction Model for Kidney Discard

Sheng Zhou, MBBS, ScM[1], **Allan B. Massie, PhD, MHS**[1,2], **Courtenay M. Holscher, MD**[1], **Madeleine M. Waldram, BA**[1], **Tanveen Ishaque, ScM**[1], **Alvin G. Thomas, MSPH**[1], and **Dorry L. Segev, MD, PhD**[1,2,3]

[1]Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD

[2]Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD

[3]Scientific Registry of Transplant Recipients, Minneapolis, MN

Abstract

Background—Many kidneys are discarded every year, with 3631 kidneys discarded in 2016 alone. Identifying kidneys at high risk of discard could facilitate "rescue" allocation to centers more likely to transplant them. The Probability of Delay or Discard (PODD) model was developed to identify marginal kidneys at risk of discard or delayed allocation beyond 36 hours of cold ischemia time. However, PODD has not been prospectively validated, and patterns of discard may have changed following policy changes such as the introduction of Kidney Donor Profile Index and implementation of the Kidney Allocation System (KAS).

Methods—We prospectively validated the PODD model using SRTR data in the KAS era (1/1/15-3/1/18). C statistic was calculated to assess accuracy in predicting kidney discard. We assessed clustering in center's utilization of kidneys with PODD>0.6 ("high-PODD") using Gini coefficients. Using match run data 1/1/15-12/31/16, we examined distribution of these high-PODD kidneys offered to centers that never accepted a high-PODD kidney.

Results—PODD predicted discard accurately under KAS (C-statistic=0.87). Compared to utilization of low-PODD kidneys (Gini coefficient = 0.41), utilization of high-PODD kidneys was clustered more tightly among a few centers (Gini coefficient = 0.84 with >60% of centers never transplanted a high-PODD kidneys). In total 11,684 offers (35.0% of all high-PODD offers) were made to centers that never accepted a high-PODD kidney.

Conclusions—Prioritizing allocation of high-PODD kidneys to centers that are more likely to transplant them might help reduce kidney discard.

Correspondence Information: Dorry Segev, M.D., Ph.D., Marjory K. and Thomas Pozefsky Professor of Surgery and Epidemiology, Associate Vice Chair, Department of Surgery, Director, Epidemiology Research Group in Organ Transplantation, Johns Hopkins University, 2000 E. Monument Street, Baltimore, MD 21205, 410-502-6115 (tel) 410-614-2079 (fax), dorry@jhmi.edu.

Authorship

Dorry Segev and Allan Massie participated in the research design. Sheng Zhou, Allan Massie, Courtenay Holscher, Madeleine Waldram, Alvin Thomas, and Dorry Segev participated in the writing of the paper. Sheng Zhou and Tanveen Ishaque participated in the data analysis.

Disclosure

The authors declare no conflicts of interest.

Author Manuscript
Author Manuscript
Author Manuscript

Introduction

Increasing number of new patients were added to the kidney waiting list every year; between 2015 and 2016, this increase was 1.9%.[1] Based on Organ Procurement and Transplantation Network data as of April 2018, over 95 000 candidates were on the kidney waiting list. The increasing demand for organs far exceeded the supply of available kidneys, with only 19 849 waitlist candidates received a kidney transplant in 2017. Because of this organ shortage, over 16% candidates waited over 5 years for a transplant, and the number of patients who were removed for death or deteriorating medical conditions reached above 9000 in 2016.[2] Moreover, the number of living donor transplants has declined in the past decade, exacerbating the shortage.[2,3]

Despite the severe organ shortage, and numerous studies showing a survival benefit from transplantation even with marginal kidneys[4–7], the discard rate rose from 2127 (14.9%) in 2006 to 3631 (20%) in 2016, and remained high following the introduction of Kidney Donor Profile Index (KDPI) and the implementation of the Kidney Allocation System (KAS).[1–3,7–12] The increase in discard was mostly explained by the underlying changes in donor characteristics.[7] In Europe and the United Kingdom, allocation schemes such as a “rescue” allocation policy and Kidney Fast-Track Scheme (KFTS) have been enacted to reduce unnecessary kidney discard.[13–15] In the United States, although there has been discussion of policies that might reduce discard, none currently exist.[16] Identifying kidneys at high risk of discard could help facilitate development of a “rescue” allocation policy.

In an effort to help operationalize prioritized allocation of kidneys more likely to be discarded, in 2010 we created a model to predict the probability that a kidney will be discarded or transplanted with cold ischemia time (CIT) exceeding 36 hours (Probability of Discard/Delay, PODD).[17] This model seemed to function well, but has never been validated, and seems more needed than ever in light of markedly increased discard rates. Additionally, the mechanisms of kidney discard might have changed under KAS, further emphasizing the need for validation. Finally, the benefit of a “rescue” allocation policy requires that there be variation in center-level willingness to transplant an organ labeled likely to be discarded by the specified index. That is, there must be variation in rates of acceptance of high-PODD kidneys which can be used to direct such kidneys to a center that might be more willing to transplant them.[17,18]

We conducted a national registry study in the KAS era to prospectively validate the accuracy of PODD in predicting kidney discard. Additionally, we compared PODD with other scores including a reduced PODD score (PODD with terms for biopsy, glomerular sclerosis, and use of machine perfusion removed); KDPI; and a model developed by Marrero et al for predicting risk of discard.[19] Finally, we characterized the variation in utilization of hard-to-place kidneys predicted by PODD among centers.

Materials and Methods

Data Source

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 Organ Procurement and Transplantation Network (OPTN), and has been described elsewhere.[20] 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. Our study was classified by the Johns Hopkins Institutional Review Board as "exempt" – not human subjects research (NA_00042871).

PODD

We included 50 207 adult deceased donor kidneys recovered for transplantation between January 1, 2015 and March 1, 2018. Informed by the previous study, double/en-bloc kidneys were considered as 1 observation.[17] For each recovered kidney, we calculated the PODD score as described in the previous study.[17] Predictors of PODD included female, age spline at 40 years, blood type AB, donor Body Mass Index (BMI, kg/m^2), BMI spline at 23, cancer, smoking, hypertension, history of myocardial infarction, diabetes, diabetes over 10 years, diabetes requiring insulin, donation after cardiac death (DCD), death by cerebrovascular accident (CVA), death by head trauma, human T lymphocyte virus (HTLV) positive, cytomegalovirus (CMV) positive, Hepatitis B core antibody (HBcAb), Hepatitis B core antibody and surface antigen (HBsAg), hepatitis C antibody (HCV), CDC high-risk donor, machine perfusion, glomerulosclerosis over 20%, and serum creatinine (mg/dL). Because this study is an external (temporal) validation of the previous study, we did not refit the PODD model, rather, we kept the coefficients of the original PODD score and calculated the probabilities of discard/delay.[21] Candidate comorbidities were assumed to absent in the case of missing or unknown data. There was <0.1% missing data for continuous predictors for BMI and serum creatinine.

Comparing PODD to other Indices of Discard

For comparison, we tested the validity of several other indices in predicting discard. The Marrero score was calculated based on the odds ratios of the predictors and intercept described elsewhere.[19,21] Predictors included age over 50, biopsy of either kidney, CMV status, DCD, CVA, donor height (cm), donor weight (kg), tattoos, either kidney pumped, HBcAb, HBsAg, HCV, history of cigarette use, history of diabetes, history of drug use, history of hypertension, terminal lab creatinine over 1.5 (mg/L), blood type, and donor race. KDPI was calculated using 2017 as the reference year.

Both PODD and Marrero score might have an advantage in predicting discard since they include factors that are not typically known at time of initial allocation (glomerulosclerosis, use of machine perfusion, and whether biopsy was done). We calculated reduced scores for PODD (r-PODD) and Marrero score (r-Marrero). For r-PODD, we fit the model with glomerulosclerosis and machine perfusion removed. In order to conduct a strict validation, we fit this model on the dataset used to fit PODD by Massie et al in 2010.[17] We calculated r-

PODD for our current study population using the coefficients from the model. For Marrero score, because of the lack of access to the original data, we calculated the reduced score (r-Marrero) with the coefficients for machine perfusion and biopsy removed. We examined the correlation between all the indices.

Accuracy in Predicting Discard

Area Under the Curve (AUC)—Because PODD was developed to predict discard/delay while Marrero score was developed to predict discard only, we examined the validation on the same outcome – kidney discard. The fit of each model in predicting discard was evaluated and compared using C statistics and the area under the receiver operating characteristic curve (AUC).[22–24]

We calculated positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity for each of the indices. Cutoffs were informed by our previous study: a stringent cutoff (0.6) was close to 90th percentile of PODD.[17] This was applied to PODD, r-PODD, Marrero score, and r-Marrero. While we aim to validate the discrimination of PODD, the calibration power of PODD might change over time. In other words, the proportion of kidneys that were more likely to be discarded for a certain cutoff might change. We also calculated the proportion of kidneys with PODD above other cutoffs (0.5 and 0.7), and the discard rate for kidneys with PODD above these cutoffs. We used 85% as the cutoff for high KDPI.[11,25] Because KDPI was not designed to predict discard, we also explored the observed probabilities of discard of kidneys with KDPI over 90% and at 100%.

Net Reclassification Improvement (NRI) Index—We quantified the value of each model using the net reclassification improvement (NRI) index.[26–28] In a perfect model, all kidneys discarded would be classified as high PODD and all kidneys transplanted would be classified as low PODD. The NRI index offers an intuitive way to measure improvement offered by one model over the other by describing how well the new model reclassifies subjects with events into high or low likelihood categories.[29] In this setting, kidney discard may be considered an event, and cutoffs for high versus low likelihood of discard must be selected for each model. To make KDPI comparable to the other indices, we converted KDPI to predicted probability of discard using logistic regression and included KDPI as the only term as a linear predictor. Thus, each KDPI score corresponded to a probability of kidney discard. We used both a stringent cutoff (0.6) and a permissive cutoff (0.5), which was close to the 80th percentile of PODD, to evaluate the improvement from one index to the next.

Missing Data—For PODD and r-PODD, we first performed complete case analysis excluding the <0.1% missing data for BMI and serum creatinine. We then performed a sensitivity analysis using multiple imputation to handle the missing data. We conducted multiple imputation by chained equations (MICE) for 10 cycles.[30,31] We checked the normality of BMI, serum creatinine, and their log transformation. We specified linear regression model for the log transformation of BMI, and predictive mean matching (PMM) for serum creatinine.[32] PMM is a linear regression approach appropriate for nonnormally distributed continuous variables. This approach imputes missing values by randomly selecting a nonmissing value from 1 of the 5 nearest neighbors after linear regression. In other words, it calculates the regression-predicted values for each observation; then fills in a value randomly chosen from the observed values from the 5 observations with the closest regression-predicted values. For all other indices, we performed complete case analysis.

Clustering of Kidney Utilization among Centers—We assessed offer acceptance using match run data that were available to us through December 2016. Using match run data for each kidney recovered between January 1, 2015 and December 31, 2016, we determined the offer acceptance of kidneys in each transplant center. We defined "offer acceptance" as acceptance and subsequent transplantation of an offered kidney for any candidate on the center's match run list. That is to say, if a center declined the offer for 1 or more candidates but eventually accepted it for another candidates, we consider this an acceptance.[33] Kidneys that were accepted but eventually not transplanted were treated equivalent as being declined. Centers that received fewer than 100 kidney offers during the study period were excluded for statistical stability (38 out of 241 were excluded).

We defined a high-PODD kidney as one with a PODD >0.6 and a low-PODD kidney as one with a PODD ≤0.6. We examined the clustering among transplant centers in high-PODD versus low-PODD kidney utilization using Lorenz curves and the Gini coefficient.[34] Lorenz curves showed cumulative distribution of kidneys by center, sorted from those who transplanted the fewest kidneys to those who transplanted the most kidneys. The diagonal line represents perfectly equal utilization among the centers, the curves represent the actual distribution of utilization. Curves that are further away from the diagonal line represent more clustering. Gini coefficient is the area between the diagonal line and the curve. The Gini coefficient is a dimensionless measure of the degree of clustering of kidney utilization. It ranges from 0 to 1. A Gini coefficient of 0 means that all centers transplanted the same number of kidneys, while a Gini coefficient of 1 means that 1 center transplanted all kidneys.

In addition, we examined correlation between center practices over time. We identified the top 15 centers that transplanted the most high-PODD kidneys in 2015 and compared with the top 15 centers that transplanted the most of high-PODD kidneys in 2016. We assessed the correlation between these top 15 centers in 2015 and 2016. We repeatedly assessed the correlation between the top 20, 30, and 40 centers in 2015 and 2016.

Sensitivity Analysis—As a sensitivity analysis, we calculated Gini coefficients for high- and low-PODD kidneys for all centers (N=241).

Surplus Offers—Surplus offers were defined as offers (kidneys being offered to 1 or more recipients at a single transplant center) of high-PODD kidneys to centers that have never accepted a high-PODD kidney between January 1, 2015 and December 31, 2016.[18] We plotted the distribution of surplus offers among centers that received at least 100 total offers during the study period.[18]

Statistical analysis—We used Wilcoxon rank-sum test to compare differences in continuous variables. We used Pearson's chi-squared test to compare differences in categorical variables. We calculated Pearson's correlation coefficients between different indices. The DeLong test was used to compare differences in AUCs.[22] The Gini coefficient were calculated based on the INEQUAL7 module for Stata by Philippe Van Kerm. Lorenz curves were created based on sg30 module for Stata by Edward Whitehouse. We used Q-Q plot to check the normality of the distributions. We used Kappa coefficient to assess the correlation between the top 15, 20, 30, and 40 centers in different years.[35] Confidence intervals were reported as per the method of Louis and Zeger.[36] All analyses were performed using Stata/SE 14.1 for Windows (College Station, Texas).

Results

Characteristics of Donors

Of the 50 207 kidneys recovered during the study period, 10 381 (20.7%) were discarded (Table 1). Compared to kidneys that were transplanted, kidneys that were discarded came from donors with a higher median age (54 vs. 38 years, p<0.001) and body mass index (BMI, 28 vs. 27, p<0.001). Discarded kidneys were more likely to be from donors who were African American (17.5% vs. 14.1%, p<0.001), female (46.0% vs. 38.2%, p<0.001), DCD (20.2% vs. 18.8%, p<0.001), Extended Criteria Donor (ECD, 49.5% vs. 13.2%, p<0.001), died from CVA (45.3% vs. 25.7%, p<0.001), had a history of smoking (32.5% vs. 19.0%, p<0.001), hypertension (60.7% vs. 27.4%, p<0.001), diabetes (23.8% vs. 6.7%, p<0.001), myocardial infarction (7.7% vs. 2.4%, p<0.001), and a history of infectious disease (Table 1). Discarded kidney were less likely to be from donors who died of head trauma (14.8% vs. 32.6%) and be labeled CDC increased risk donors (21.5% vs. 25.1%, p<0.001). Discarded kidneys were less likely to be machine perfused (27.3% vs. 36.6%, p<0.001), and more likely to have sclerosis >20% (23.2% vs. 1.4%, p<0.001, Table 1).

Correlation between indices

Predictors of each index were listed in Table 2. Correlation was good between PODD, r-PODD, Marrero score, and r-Marrero, with all the correlation coefficients between 0.8 and 1.0 (Table 3). KDPI was correlated less well with the other indices, with all correlation coefficients between 0.7 and 0.8 (Table 3).

Accuracy in Predicting Discard

Of kidneys studied, 9.5% had PODD over 0.6, 13.0% had Marrero score over 0.6, 9.3% had r-PODD over 0.6, 8.2% had r-Marrero over 0.6, 12.7% had KDPI over 85% (Table 4). The observed probability of discard increased from 63.6% for kidneys with KDPI ≥85% to 70.5% for kidneys with KDPI ≥90%, and 92.0% for kidneys with KDPI of 100%. Among all the indices, PODD had the highest positive predictive value of 78.5%. This means if a kidney has a PODD greater than 0.6, then there is a 78.5% chance that this kidney will be discarded. The simpler r-PODD still had a higher positive predictive value than the Marrero score, r-Marrero, or KDPI (Table 4). When looking at other PODD cutoffs, 13.7% had PODD over 0.5, 71.7% of these kidneys were discarded; 6.2% had PODD over 0.7, 85.2% of these kidneys were discarded. AUC was the largest for PODD (0.87), followed by r-PODD (0.85), Marrero (0.84), r-Marrero (0.82), and KDPI (0.82) (all p<.001) (Figure 1). Using 0.6 as the cutoff, Net Reclassification Improvement (NRI) was 15.6% from KDPI to PODD, and 4.1% from Marrero to PODD. NRI was 11.4% from KDPI to r-PODD, and

11.9% from r-Marrero to r-PODD. Using 0.5 as the cutoff, NRI was 9.3% from KDPI to PODD, and 4.2% from Marrero to PODD. NRI was 7.0% from KDPI to r-PODD, and 14.7% from r-Marrero to r-PODD. In summary, by all 3 metrics – AUC, NRI, and positive predictive value – PODD predicted discard most accurately. After multiple imputation, the conclusions stayed the same.

Clustering of Kidney Utilization among Centers

There were 203 centers that received at least 100 kidney offers over the study period. The median number of low-PODD kidneys transplanted by centers was 66 (range=2-375, IQR=36-119). However, the median number of high-PODD kidneys transplanted by centers during the study period was 0 (IQR=0-1, range=0-54). There was markedly greater clustering among centers in use of high-PODD kidneys as compared to use of low-PODD kidneys (Gini=0.84 vs 0.41, Figure 2). In total 122 (60.1%) centers did not use a single high-PODD kidney, and the top 15 centers (ranked according to number of high-PODD kidneys accepted) used the majority (63.0%) of high-PODD kidneys (Figure 2). When including all 241 centers, Gini coefficients remained the same (0.84 vs 0.41). In addition, among the top 15 centers that used the most high-PODD kidneys in 2015, 13 remained to use the most high-PODD kidneys in 2016 (kappa coefficient=0.77). The consistency between 2015 and 2016 was less well among the top 20, 30, and 40 centers (kappa coefficient=0.61, 0.52, 0.38, respectively).

Surplus Offers

Among the 203 centers offered ("offer" was defined as a donor kidney being offered to 1 or more recipients at a single transplant center) at least 100 kidneys during the study period, 122 (60.1%) never accepted a high-PODD kidney. There were 11 684 of surplus offers (35.0% of all 33 386 high-PODD offers) to these centers. Distribution of these surplus offers were shown in Figure 3. The median (IQR) number of surplus offers per center was 78 (27-125). Surplus offers per center ranged from 0 to 629.

Discussion

In this national study of kidney discard in the KAS era, we prospectively validated PODD in predicting discard of deceased donor kidneys. The accuracy of PODD was excellent, and better than the other indices using several evaluation metrics including C-statistic, PPV, and NRI. We also found substantial clustering in center's use of high-PODD kidneys: while 122 centers did not use a single high-PODD kidneys, 15 centers used the majority of these kidneys. The centers that used the most high-PODD kidneys remained consistent over time. Of all high-PODD kidney offers, 35.0% were arguably unnecessarily made to centers that never accepted such kidneys.

The high rate of kidney discard has been a concern for decades and has continued despite recent allocation changes in KAS.[7,37–40] A "rescue" strategy has been proposed to reduce discard in which kidneys at high risk of discard are identified prospectively and preferentially offered to centers more likely to transplant such kidneys. Such a strategy requires a reliable method for identified kidneys at high risk of discard; the number of kidneys thus identified should be small enough to ensure minimal disruption to the overall allocation system. In this study, we have validated that PODD would be an appropriate tool for such a system. Only 9% of kidneys had PODD exceeding 0.6; over 79% of these kidneys were discarded. Over 13% of kidneys have PODD exceeding 0.5; 71.7% of these kidneys were discarded. Only 6.2% of kidneys had PODD exceeding 0.7; 85.2% of these kidneys were discarded. The OPTN could implement rescue allocation by allocating kidneys differently if PODD exceeded a certain threshold. The higher the threshold, the higher the chance that a "rescued" kidney would otherwise be discarded – but the fewer kidneys would be "rescued". In addition, we have also validated the usefulness of reduced PODD, which does not depend on factors that may not be known at time of initial allocation, thus would be a more practical tool. Concerns have been raised that such a system might direct all "rescue" offers to a few centers, denying other centers the opportunity to show that they are willing to utilize these marginal kidneys. Applying "rescue" allocation only to kidneys refused at the local OPO level, or allowing centers to petition for a temporary opportunity to prove themselves, could partially alleviate such concerns.

Our findings are consistent with previous reports that KDPI inadequately identifies kidneys at high risk of being discarded.[17,19] While KDPI was designed to describe organ quality, it was not designed to predict kidney discard.[41] The current findings were consistent with our initial report of PODD, showing that PODD was substantially better at predicting discard than KDPI with an improvement of 15.6%.[17] The ability of PODD to discriminate between acceptance and discard has been consistently high, with an AUC of 0.83 in our prior study and 0.87 in our current validation study. [17] After excluding machine perfusion and glomerulosclerosis as predictor variables which are often not available at the time of initial organ offer, our reduced PODD (r-PODD) model remained more accurate in predicting discard than KDPI with an improvement of 11.4%. Compared to the Marrero score, for kidney discard prediction, PODD still performed significantly better with an improvement of 4.1%. Reduced PODD also performed significantly better than reduced Marrero with an improvement of 11.9%.

In considering center-level behavior, we found that the use of high-PODD kidneys varied widely in the US. This confirms several other reports of center-level variation in utilization of kidneys. Wey et al found that lower offer acceptance in a donation service area (DSA) was associated both with higher discard rates in that DSA and with higher rates of export of organs from that DSA.[42] Our group previously demonstrated that there was clustering of the utilization of suboptimal grafts among transplant centers.[43] We also found the clustering of use of high-PODD kidneys over time were the most consistent among the top 15 centers. In addition, we found that 35.0% high-PODD kidney offers went to centers that have never accepted a single high-PODD kidney. This is also consistent with previous findings that marginal kidneys offers were concentrated among a small amount of centers.[18] Expedited offer of high-PODD kidneys to centers that are willing to accept them could be an effective way to reduce discard and improve access to transplantation.

Internationally, there have been efforts to target distribution of kidneys that were more likely to be discarded to those centers more likely to transplant them. In Europe, a policy was introduced in 2006 that transitioned organ offers to a center-oriented "rescue allocation" following a series of 5 rejections in a patient-oriented offer scheme.[13] This was effective at decreasing discard rate[13], and following changes to the policy in 2013, the cold ischemia time of transplanted kidneys decreased as well.[15] In another example, a policy was implemented in the UK in 2006 to allocate kidneys to centers agreeing to participate in the “Declined Kidney Scheme” which then have the discretion to use the offered kidney locally for one of their patients, following a series of 5 rejections in the national allocation policy.[44] Later in 2012, the “Kidney Fast-Track Scheme” further expanded upon this by allowing these kidneys to be offered simultaneously to multiple centers.[14,45] Studies have not only shown decreased kidney discard following both policies, but perhaps more importantly, have demonstrated posttransplant outcomes that are comparable to those of organs transplanted through standard allocation schemes.[13,14,45–47]

Our study validated the effectiveness of PODD to identify kidneys that are more likely to be discarded, and demonstrated existing variation in utilization of kidneys with high PODD among centers. Use of such a validated index to improve the distribution of kidneys that are at higher risk of being discarded could decrease kidney discard rates and potentially improve candidate survival. In the setting of an ongoing dialogue about the high rates of kidney discard, we recommend developing a policy using a validated metric to improve distribution of kidneys likely to be discarded.

Acknowledgments

Funding

This work was supported by grants number K24DK101882 (Segev) and F32DK109662 (Holscher) from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and an American College of Surgeons Resident Research Scholarship (Holscher). The analyses described here are the responsibility of the authors alone and do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government.

The data reported here have been supplied by the Minneapolis Medical Research Foundation (MMRF) as the contractor for the Scientific Registry of Transplant Recipients (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.

Abbreviations

AUC	area under the curve
BMI	body mass index
CIT	cold ischemia time
CMV	cytomegalovirus
CVA	cerebrovascular accident
DCD	donation after cardiac death
ECD	extended criteria donor
HBcAb	hepatitis B core antibody

HBsAg	hepatitis B surface antigen
HCV	hepatitis C virus
HRSA	health resources and services administration
HTLV	human T lymphotropic virus
IRD	infectious risk donor
KAS	kidney allocation system
KDPI	kidney donor profile index
KFTS	kidney fast-track scheme
MICE	multiple imputation by chained equation
NPV	negative predictive value
NRI	net reclassification improvement
OPTN	organ procurement and transplantation network
PMM	predictive mean matching
PODD	probability of discard/delay
PPV	positive predictive value
ROC	receiver operating curve
r-PODD	reduced probability of discard/delay
SRTR	scientific registry of transplant recipients

References

1. OPTN/SRTR 2016 Annual Data Report: Introduction. Am J Transplant. 2018; 18(Suppl 1):10–17. [PubMed: 29292601]
2. Hart A, Smith JM, Skeans MA, et al. OPTN/SRTR 2016 Annual Data Report: Kidney. Am J Transplant. 2018; 18(Suppl 1):18–113. [PubMed: 29292608]
3. U.S. Department of Health & Human Services. SRTR/OPTN Annual Data Report Archives, ADR. 2007. https://srtr.transplant.hrsa.gov/archives.aspx. Accessed February 5, 2018
4. Massie AB, Luo X, Chow EK, Alejo JL, Desai NM, Segev DL. Survival benefit of primary deceased donor transplantation with high-KDPI kidneys. Am J Transplant. 2014; 14:2310–2316. [PubMed: 25139729]
5. Chow EK, Massie AB, Muzaale AD, et al. Identifying appropriate recipients for CDC infectious risk donor kidneys. Am J Transplant. 2013; 13:1227–1234. [PubMed: 23621162]
6. Merion RM, Ashby VB, Wolfe RA, et al. Deceased-donor characteristics and the survival benefit of kidney transplantation. JAMA. 2005; 294:2726–2733. [PubMed: 16333008]
7. Stewart DE, Garcia VC, Rosendale JD, Klassen DK, Carrico BJ. Diagnosing the Decades-Long Rise in the Deceased Donor Kidney Discard Rate in the United States. Transplantation. 2017; 101:575–587. [PubMed: 27764031]

8. Formica RN Jr. A critical assessment on kidney allocation systems. Transplant Rev (Orlando). 2017; 31:61–67. [PubMed: 27919635]

9. Massie AB, Luo X, Lonze BE, et al. Early Changes in Kidney Distribution under the New Allocation System. J Am Soc Nephrol. 2016; 27:2495–2501. [PubMed: 26677865]

10. Stewart DE, Kucheryavaya AY, Klassen DK, Turgeon NA, Formica RN, Aeder MI. Changes in Deceased Donor Kidney Transplantation One Year After KAS Implementation. Am J Transplant. 2016; 16:1834–1847. [PubMed: 26932731]

11. Bae S, Massie AB, Luo X, Anjum S, Desai NM, Segev DL. Changes in Discard Rate After the Introduction of the Kidney Donor Profile Index (KDPI). Am J Transplant. 2016; 16:2202–2207. [PubMed: 26932575]

12. White SL. Nudging the Organ Discard Problem. Transplantation. 2017; 101:1518–1519. [PubMed: 28272289]

13. Vinkers MT, Smits JM, Tieken IC, de Boer J, Ysebaert D, Rahmel AO. Kidney donation and transplantation in Eurotransplant 2006-2007: minimizing discard rates by using a rescue allocation policy. Prog Transplant. 2009; 19:365–370. [PubMed: 20050462]

14. White AD, Roberts H, Ecuyer C, et al. Impact of the new fast track kidney allocation scheme for declined kidneys in the United Kingdom. Clin Transplant. 2015; 29:872–881. [PubMed: 26094680]

15. Wahba R, Suwelack B, Arns W, et al. Rescue Allocation and Recipient Oriented Extended Allocation in kidney transplantation - Influence of the EUROTRANSPLANT allocation system on recipient selection and graft survival for initially non-accepted organs. Transpl Int. 2017; 30:1226–1233. [PubMed: 28574620]

16. Reese PP, Harhay MN, Abt PL, Levine MH, Halpern SD. New Solutions to Reduce Discard of Kidneys Donated for Transplantation. J Am Soc Nephrol. 2016; 27:973–980. [PubMed: 26369343]

17. Massie AB, Desai NM, Montgomery RA, Singer AL, Segev DL. Improving distribution efficiency of hard-to-place deceased donor kidneys: Predicting probability of discard or delay. Am J Transplant. 2010; 10:1613–1620. [PubMed: 20642686]

18. Massie AB, Stewart DE, Dagher NN, Montgomery RA, Desai NM, Segev DL. Center-level patterns of indicated willingness to and actual acceptance of marginal kidneys. Am J Transplant. 2010; 10:2472–2480. [PubMed: 20977638]

19. Marrero WJ, Naik AS, Friedewald JJ, et al. Predictors of Deceased Donor Kidney Discard in the United States. Transplantation. 2017; 101:1690–1697. [PubMed: 27163541]

20. Massie AB, Kucirka LM, Segev DL. Big data in organ transplantation: registries and administrative claims. Am J Transplant. 2014; 14:1723–1730. [PubMed: 25040084]

21. Osborne JW. Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. Pract Assess Res Eval. 2006; 11:1–6.

22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics. 1988; 44(3):837–845. [PubMed: 3203132]

23. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978; 8:283–298. [PubMed: 112681]

24. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982; 143:29–36. [PubMed: 7063747]

25. Snyder JJ, Salkowski N, Wey A, et al. Effects of High-Risk Kidneys on Scientific Registry of Transplant Recipients Program Quality Reports. Am J Transplant. 2016; 16:2646–2653. [PubMed: 26954720]

26. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008; 27:157–172. [PubMed: 17569110]

27. Moqueet N, Kanagaratham C, Gill MJ, et al. A prognostic model for development of significant liver fibrosis in HIV-hepatitis C co-infection. PLoS One. 2017; 12:e0176282. [PubMed: 28467457]

28. Canales MT, Blackwell T, Ishani A, et al. Renal Function and Death in Older Women: Which eGFR Formula Should We Use? Int J Nephrol. 2017; 2017:8216878. [PubMed: 28465840]
29. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011; 30:11–21. [PubMed: 21204120]
30. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011; 30:377–399. [PubMed: 21225900]
31. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 2011; 20:40–49. [PubMed: 21499542]
32. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Med Res Methodol. 2014; 14:75. [PubMed: 24903709]
33. Bowring MG, Massie AB, Craig-Schapiro R, Segev DL, Nicholas LH. Kidney offer acceptance at programs undergoing a Systems Improvement Agreement. Am J Transplant. May 2.2018 doi: 10.1111/ajt.14907
34. Gastwirth JL. The Estimation of the Lorenz Curve and Gini Index. Rev Econ Stat. 1972; 54:306–316.
35. Cyr L, Francis K. Measures of clinical agreement for nominal and categorical data: the kappa coefficient. Comput Biol Med. 1992; 22:239–246. [PubMed: 1643847]
36. Louis TA, Zeger SL. Effective communication of standard errors and confidence intervals. Biostatistics. 2009:1–2.
37. Tanriover B, Mohan S, Cohen DJ, et al. Kidneys at higher risk of discard: expanding the role of dual kidney transplantation. Am J Transplant. 2014; 14:404–415. [PubMed: 24472195]
38. Heilman RL, Green EP, Reddy KS, Moss A, Kaplan B. Potential Impact of Risk and Loss Aversion on the Process of Accepting Kidneys for Transplantation. Transplantation. 2017; 101:1514–1517. [PubMed: 28640014]
39. Pruett TL, Clark MA, Taranto SE. Deceased Organ Donors and PHS Risk Identification: Impact on Organ Usage and Outcomes. Transplantation. 2017; 101:1670–1678. [PubMed: 28252560]
40. Volk ML, Wilk AR, Wolfe C, Kaul DR. The "PHS Increased Risk" Label Is Associated With Nonutilization of Hundreds of Organs per Year. Transplantation. 2017; 101:1666–1669. [PubMed: 28196048]
41. Heilman RL, Mathur A, Smith ML, Kaplan B, Reddy KS. Increasing the Use of Kidneys From Unconventional and High-Risk Deceased Donors. Am J Transplant. 2016; 16:3086–3092. [PubMed: 27172238]
42. Wey A, Salkowski N, Kasiske BL, Israni AK, Snyder JJ. Influence of Kidney Offer Acceptance Behavior on Metrics of Allocation Efficiency. Clin Transplant. 2017; 31doi: 10.1111/ctr.13057
43. Garonzik-Wang JMJN, Weatherspoon KC, Deshpande NA, Berger JA, Hall EC, Montgomery RA, Segev DL. The aggressive phenotype: center-level patterns in the utilization of suboptimal kidneys. Am J Transplant. 2012; 12:400–408. [PubMed: 21992578]
44. Callaghan CJ, Harper SJ, Saeb-Parsy K, et al. The discard of deceased donor kidneys in the UK. Clin Transplant. 2014; 28:345–353. [PubMed: 24506794]
45. Mittal S, Adamusiak A, Horsfield C, et al. A Re-Evaluation of Discarded Deceased Donor Kidneys in the UK: Are Usable Organs Still Being Discarded? Transplantation. 2017; 101:1698–1703. [PubMed: 27779574]
46. Callaghan CJ, Mumford L, Pankhurst L, Baker RJ, Bradley JA, Watson CJE. Early Outcomes of the New UK Deceased Donor Kidney Fast-track Offering Scheme. Transplantation. 2017; 101:2888–2897. [PubMed: 28640070]
47. Wahba R, Teschner S, Stippel DL. Results of kidney transplantation after rescue allocation. Transpl Int. 2011; 24:e46–47. [PubMed: 21208298]

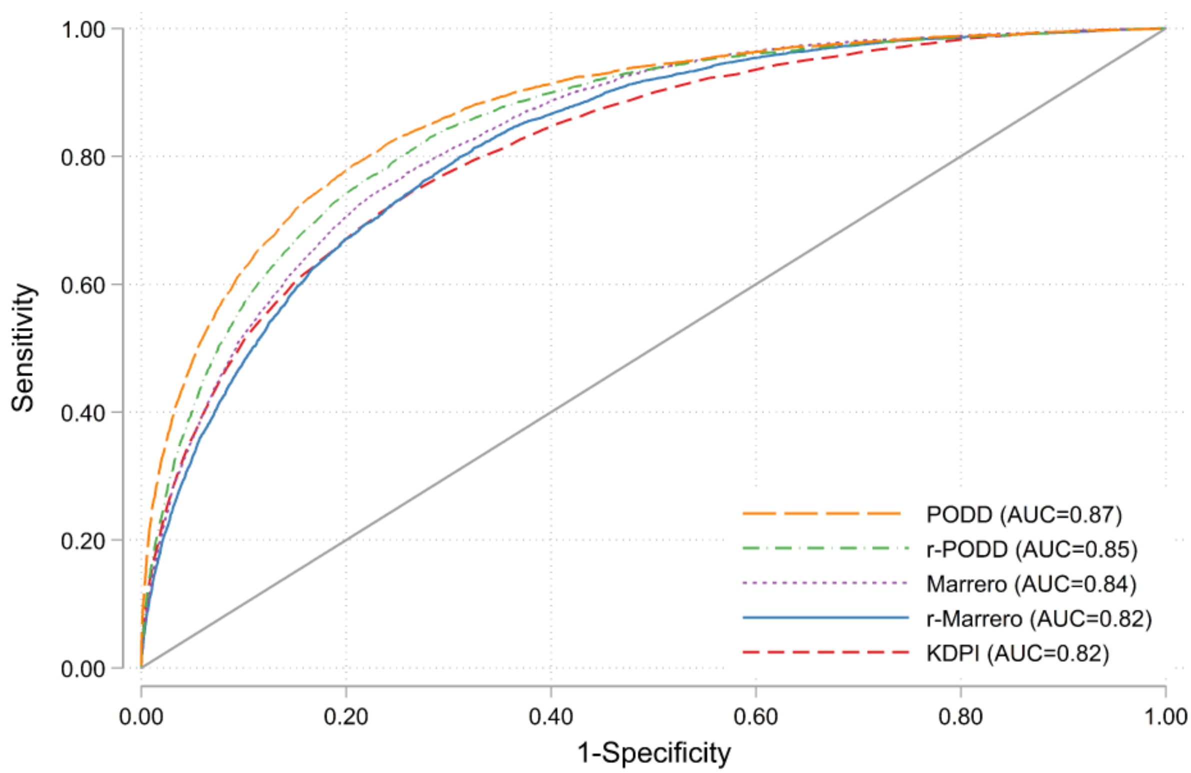

Figure 1. ROC curves showing accuracy in predicting discard

Curves closer to the diagonal line represent worse accuracy while curves closer to the upper left represent better accuracy in predicting discard. PODD have greater accuracy in predicting discard than all other scores ($p<0.001$).

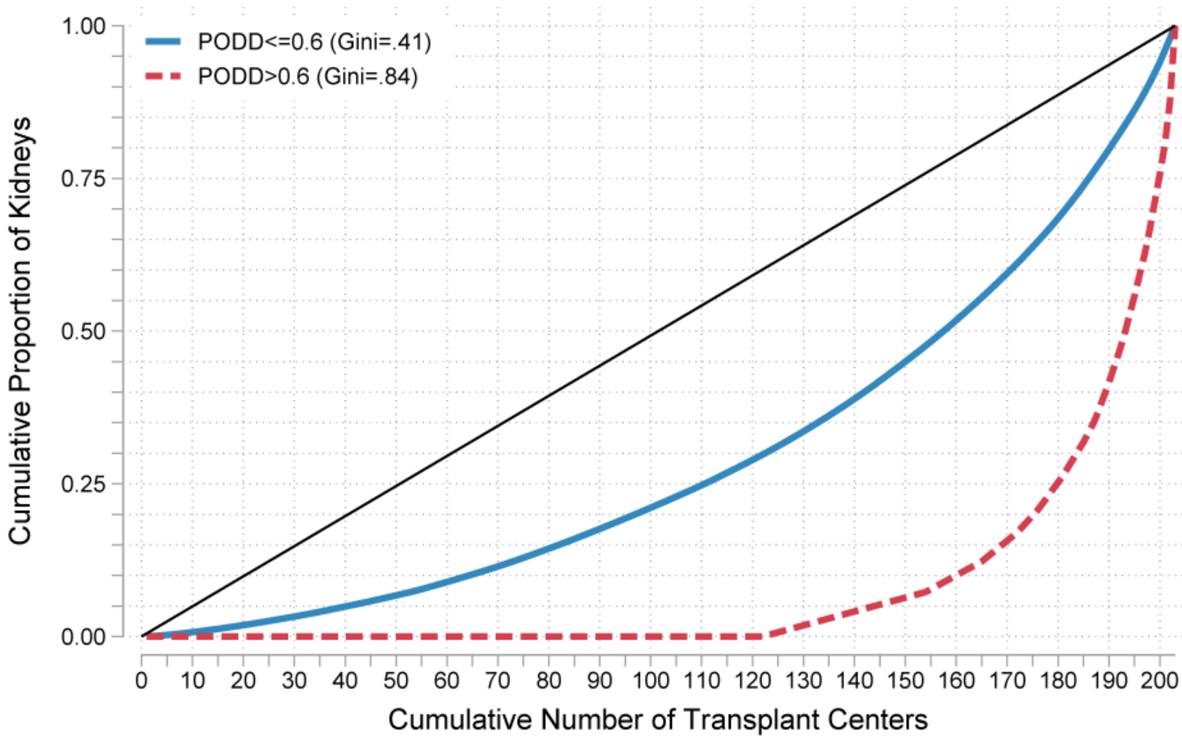

Figure 2. Clustering in kidney utilization among centers

Curves closer to the diagonal line represent more equity and those closer to the lower right indicate more clustering and less equity of distribution. The Gini coefficient denotes the degree of clustering. There was less equitably distributed utilization of high-PODD kidneys than for low-PODD kidneys (Gini=0.84 vs 0.41). In total 122 (60.1%) centers did not use a single high-PODD kidney, 15 centers have used the majority (63.0%) of high-PODD kidneys.

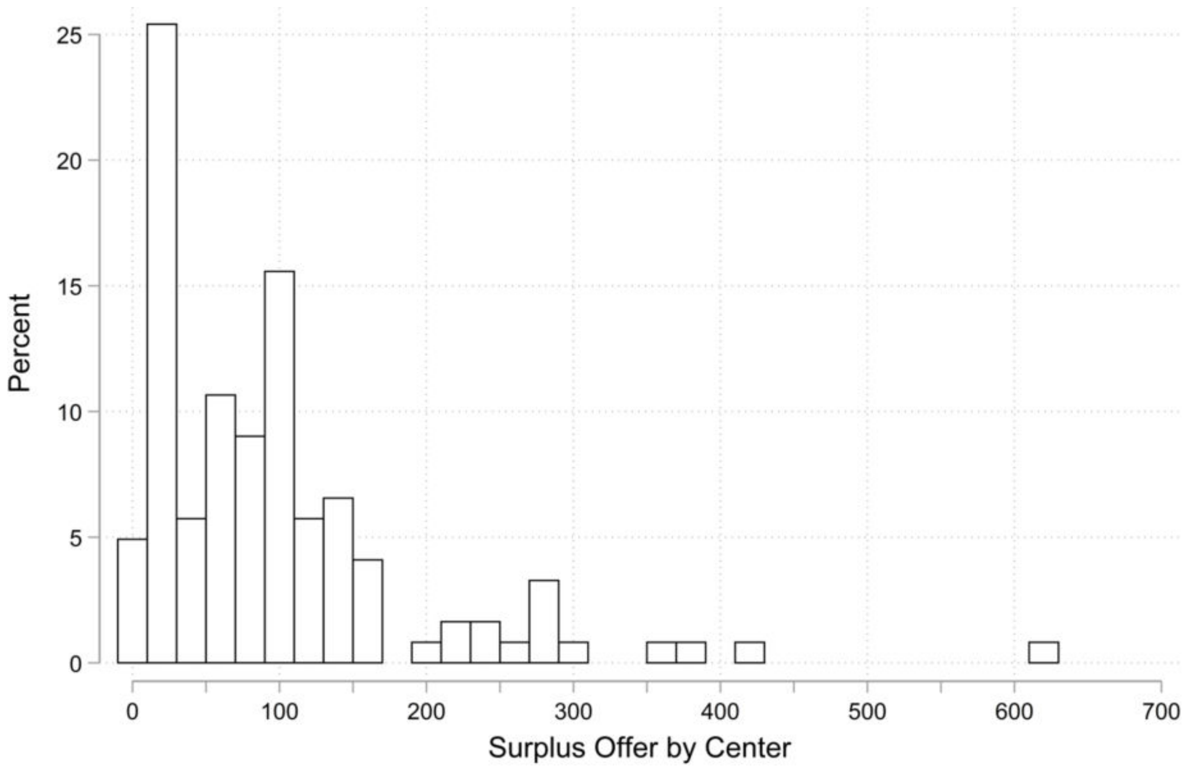

Figure 3. Distribution of Surplus Offers

Surplus offers were defined as offers (kidneys being offered to 1 or more recipients at a single transplant center) of high-PODD kidneys to centers that never accepted a high-PODD kidney. This figure shows the distribution of surplus offers among 122 centers that never accepted a high-PODD kidney during the study period. In total, 11684 surplus offers were made to these centers; a median (IQR) of 78 (27-125) surplus offers were made per center. At maximum 629 surplus offers were made to a center.

Table 1

Characteristics of donors of all kidneys recovered for transplantation in the United States 1/1/15-3/1/18.

	Transplanted N= 39 826	**Discarded N=10 381**	**p value**
Age, median (IQR)	38 (27, 50)	54 (44, 62)	<0.001
Race/ethnicity			<0.001
White	68.18%	67.67%	
African American	14.07%	17.46%	
Hispanic	14.08%	11.25%	
Asian	2.45%	2.75%	
Other	1.22%	0.87%	
Female	38.20%	46.00%	<0.001
BMI, median (IQR)	27.02 (23.55, 31.50)	28.47 (24.61, 33.41)	<0.001
Blood type			0.06
A	35.14%	34.06%	
B	11.92%	12.26%	
AB	3.33%	3.73%	
O	49.61%	49.95%	
Medical History			
smoking	19.04%	32.46%	<0.001
hypertension	27.36%	60.69%	<0.001
diabetes	6.71%	23.81%	<0.001
diabetes >10 yrs	1.52%	8.30%	<0.001
Insulin-dependent diabetes	2.28%	9.42%	<0.001
myocardial infarction	2.44%	7.69%	<0.001
Donation Characteristics			
DCD	18.76%	20.20%	<0.001
ECD	13.17%	49.49%	<0.001
Cause of death: CVA	25.73%	45.29%	<0.001
Cause of death: head trauma	32.56%	14.83%	<0.001
History of Infection			
HTLV positive	0.03%	0.05%	0.3
CMV positive	60.72%	65.07%	<0.001
HBcAb positive	3.56%	7.42%	<0.001
HBsAg positive	0.04%	0.23%	<0.001
HCV positive	4.31%	11.21%	<0.001
IRD	25.09%	21.47%	<0.001
Machine Perfusion	36.62%	27.29%	<0.001

	Transplanted N= 39 826	Discarded N=10 381	p value
Biopsy done	49.79%	87.13%	<0.001
Sclerosis >20%	1.40%	23.22%	<0.001
Terminal serum creatinine, median (IQR)	0.93 (0.70, 1.32)	1.30 (0.90, 2.20)	<0.001

*BMI: Body Mass Index; DCD: Donation after Cardiac Death; ECD: Extended Criteria Donor; CVA: Cerebrovascular accident; HTLV: Human T lymphotropic virus; CMV: cytomegalovirus; HBcAb: Hepatitis B Core Antibody; HBsAg: Hepatitis B Antigen; HCV: Hepatitis C Virus; IRD: Infectious Risk Donor

Table 2

Predictors of indices

PODD and Marrero indices were similar in donor demographic predictors including age, blood type, and BMI, though different functional forms were used. PODD had more predictors of comorbidities including cancer and myocardial infarction, and infection history including IRD and HTLV. Marrero score had additional predictors of drug use and tattoos. PODD included donor sex while Marrero score included donor race. Both PODD and Marrero score had predictors that are not known at the time of allocation including machine perfusion, biopsy, and glomerulosclerosis. These were excluded in the reduced indices.

Index	**PODD**	**r-PODD**	**Marrero**	**r-Marrero**
Original Outcome	**Delay or discard**	**Delay or discard**	**Discard**	**Discard**
Predictors	Female	Female		
			Race category	Race category
	Age spline at 40	Age spline at 40	Age > 50	Age > 50
	Blood type AB	Blood type AB	Blood type category	Blood type category
	BMI	BMI	Height	Height
	BMI spline at 23	BMI spline at 23	Weight	Weight
	Smoking	Smoking	Smoking	Smoking
			Drug use	Drug use
			Tattoos	Tattoos
	Hypertension	Hypertension	Hypertension	Hypertension
	Diabetes	Diabetes	Diabetes	Diabetes
	Diabetes>10yrs	Diabetes>10yrs		
	Insulin-dependent diabetes	Insulin-dependent diabetes		
	Cancer	Cancer		
	Myocardial infarction	Myocardial infarction		
	Death by head trauma	Death by head trauma		
	DCD	DCD	DCD	DCD
	CVA	CVA	CVA	CVA
	CMV	CMV	CMV	CMV
	HCV	HCV	HCV	HCV
	HBcAb	HBcAb	HBcAb	HBcAb
	HBcAg+HBsAg	HBcAg+HBsAg	HBsAg	HBsAg
	HTLV	HTLV		
	IRD	IRD		
	Machine perfusion		Machine perfusion	
	Creatinine	Creatinine	Creatinine > 1.5	Creatinine > 1.5
	Sclerosis > 20%			
			Biopsy	

*BMI: Body Mass Index; DCD: Donation after Cardiac Death; ECD: Extended Criteria Donor; CVA: Cerebrovascular accident; HTLV: Human T lymphotropic virus; CMV: cytomegalovirus; HBcAb: Hepatitis B Core Antibody; HBsAg: Hepatitis B Antigen; HCV: Hepatitis C Virus; IRD: Infectious Risk Donor

Table 3

Correlation between indices

Correlation coefficients that are closer to 1 indicate good correlation while correlation coefficients that are closer to 0 indicate no correlation. Correlations between PODD, Marrero score, and the reduced scores were good; while KDPI was moderately correlated with the other indices.

	PODD	r-PODD	Marrero score	r-Marrero	KDPI
PODD	1.00				
r-PODD	0.94	1.00			
Marrero score	0.84	0.85	1.00		
r-Marrero	0.81	0.85	0.95	1.00	
KDPI	0.71	0.76	0.78	0.76	1.00

Table 4

Evaluation of Indices

Positive predictive value (PPV) is the probability of discard for a kidney with a score that is above the threshold. For example, 78.5% of kidneys with PODD>0.6 were discarded. Negative predictive value (NPV) is the probability that a kidney with a score that is below the threshold is being transplanted. Sensitivity is the probability that a discarded kidney has a score that is above the threshold. Specificity is the probability that a transplanted kidney has a score that is below the threshold. PODD and r-PODD had the highest PPV; over 70% of kidneys flagged as high-PODD or high r-PODD were discarded.

Indice	% of all kidneys	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)
PODD>0.6	9.46	78.45	85.36	35.90	97.43
Marrero>0.6	13.04	55.51	84.55	35.02	92.69
r-PODD>0.6	9.33	71.98	84.60	32.46	96.71
r-Marrero>0.6	8.23	52.19	82.15	20.77	95.04
KDPI≥85%	12.71	63.60	85.57	39.08	94.17