Predicting Kidney Transplant Survival Using Tree-Based Modeling

Sergey Krikov,* Altaf Khan,† Bradley C. Baird,‡ Lev L. Barenbaum,§ Alexander Leviatov,§ James K. Koford,‖ and Alexander S. Goldfarb-Rumyantzev‡¶

Predicting the outcome of kidney transplantation is clinically important and computationally challenging. The goal of this project was to develop the models predicting probability of kidney allograft survival at 1, 3, 5, 7, and 10 years.

Kidney transplant data from the United States Renal Data System (January 1, 1990, to December 31, 1999, with the follow-up through December 31, 2000) were used (n = 92,844). Independent variables included recipient demographic and anthropometric data, end-stage renal disease course, comorbidity information, donor data, and transplant procedure variables.

Tree-based models predicting the probability of the allograft survival were generated using roughly two-thirds of the data (training set), with the remaining one-third left aside to be used for models validation (testing set). The prediction of the probability of graft survival in the independent testing dataset achieved a good correlation with the observed survival ($r = 0.94$, $r = 0.98$, $r = 0.99$, $r = 0.93$, and $r = 0.98$) and relatively high areas under the receiving operator characteristic curve (0.63, 0.64, 0.71, 0.82, and 0.90) for 1-, 3-, 5-, 7-, and 10-year survival prediction, respectively.

The models predicting the probability of 1-, 3-, 5-, 7-, and 10-year allograft survival have been validated on the independent dataset and demonstrated performance that may suggest implementation in clinical decision support system. *ASAIO Journal* 2007; 53:592–600.

Patients with end-stage renal disease (ESRD) who receive kidney transplants have better outcomes than those remaining on the waiting list.[1] However, graft failure terminates the patient survival benefit.[2] Whereas kidney transplantation remains the preferred modality of treatment of ESRD, there is a disparity between the number of patients on the waiting list and the number of kidneys available for transplant.[3] Therefore, efforts aimed at improving graft survival time are very important for clinical practice. One way to improve long-term allograft survival is to identify modifiable factors that negatively affect the outcome and include examining the interaction between and among them and designing the intervention to optimize the outcome. Using complex mathematical models to predict allograft outcome and identify factors affecting the survival may therefore make a significant impact on treatment strategies and improve overall clinical outcome.

Several significant predictors of suboptimal transplant outcome were previously identified in adults[4–8] and children,[9–11] based on data from the United Network of Organ Sharing (UNOS) and the North American Pediatric Renal Transplant Cooperative Study (NAPRTCS). Donor and recipient age,[12] preexisting donor hypertension and diabetes,[13,14] nonheart-beating donor,[15] prolonged cold storage time,[8] retransplantation,[16] pretransplant renal replacement therapy modality,[17] duration of pretransplant ESRD,[18] body mass index of donor and recipient,[4] and recipient marital status,[19] along with other factors, play important roles in the outcome. However, the probable interaction of these factors, plus their potential to act in various combinations, makes it difficult to predict the outcome in individual patients without using mathematical tools. Mathematical models would accurately predict graft survival duration as well as identify patients at risk and locate potentially modifiable risk factors. We previously described mathematical models predicting 3-year deceased graft survival.[20] However, that report was limited only to deceased donor kidney recipients, and the model was designed only to predict 3-year allograft survival. The attempt herein described is undertaken to develop a tree-based model predicting the probability of graft survival at posttransplant years 1, 3, 5, 7, and 10.

Methods

Dataset

We used the data collected by the United States Renal Data System (USRDS) and UNOS, which described all kidney allograft recipients (both pediatric and adults) who underwent kidney or kidney-pancreas transplantation during the period of January 1, 1990, through December 31, 1999. The follow-up period was extended through December 31, 2000. The variables for the prediction models were selected from the USRDS files: PATIENT (patients demographic data), RXHIST60 (patients ESRD history), TX (patient transplant data), TXUNOS (baseline detailed transplant data), and TXFUUNOS (transplant follow-up data). For recipients of multiple transplants, the most recent one was considered the target transplant (transplant of

From the *Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT; †Department of Family and Preventive Medicine, University of Utah, Salt Lake City, UT; ‡Division of Nephrology and Hypertension, University of Utah School of Medicine, Salt Lake City, UT; §RenalService.com, Inc., Salt Lake City, UT; ‖Division of Undergraduate Studies and University Writing Program University of Utah, Salt Lake City, UT; and ¶Veterans Affairs Salt Lake City Healthcare System, Salt Lake City, UT.

Submitted for consideration March 2007; accepted for publication in revised form June 2007.

Reprint Requests: Dr. Alexander Goldfarb-Rumyantzev, Division of Nephrology and Hypertension, University of Utah Health Sciences Center, 85 North Medical Drive, East Room 201, Salt Lake City, UT 84112.

DOI: 10.1097/MAT.0b013e318145b9f7

interest). Patient records with missing information regarding graft or patient survival were excluded from the study. A total of 92,844 patients with kidney transplants were identified. Censored data used for multivariate analysis to identify factors that have an association with the outcome were excluded from the prediction analysis, as described below. Separate datasets were generated for each of the five tree models. These datasets included only uncensored records that had specific information of graft survival at a given time period (*i.e.*, 1, 3, 5, 7, and 10 posttransplant years in the respective datasets). For example, for the 1-year prediction model, only patients with a known 1-year outcome were selected, whereas those who were censored because of insufficient duration of follow-up or other reasons were excluded. From each of the datasets, two-thirds of the data were randomly selected into the training dataset and the remaining one-third into the testing dataset. The training set was used for knowledge acquisition (to generate the model), whereas validation was performed using the records from the testing set.

Outcome

The outcome was the time between the most recent kidney transplant and the failure of the graft. For the prediction model, the outcome was converted into 1-, 2-, 3-, 5-, 7-, and 10-year transplant recipients and graft survivals as a binary variable. The graft failure definition did not include patient death with a functioning graft (*i.e.*, death censored graft survival). In the event that the information regarding death with a functioning graft was missing in the dataset, and the patient death date has been found to be equal to the graft failure date, we assumed that the patient died with a functioning graft unless the cause of death specified in the UNOS file was coded (ICD-9) as one of the following: 3200, graft failure: primary failure; 3201, graft failure: rejection; 3202, graft failure: technical; 3299, graft failure: other; or 3903, miscellaneous: renal failure.

Independent Variables

The following independent variables were considered and evaluated for inclusion in the prediction models:

1. Recipient demographic and anthropometric data: age, race, gender, height, and weight. Information was obtained from USRDS files SAF.PATIENT and SAF.TXUNOS.
2. Variables describing recipient ESRD course were obtained from SAF.PATIENT and SAF.RXHIST60 files: age of onset of ESRD; total duration of pretransplant ESRD period (time between the first ESRD service and most recent transplant date); renal replacement therapy (RRT) modality immediately before current transplant; predominant RRT modality during ESRD course (defined as modality used for >50% of the ESRD period as previously described,[17] number of different RRT modalities used; the specific combination of RRT modalities; absolute time and percent time of the whole ESRD period that the patient was treated with specific RRT modality; history of transplants before the current one (yes/no); and total number of transplants (including the current one). Since preemptive transplantation was reported to be advantageous in terms of graft survival,[21–23] the binary variable defining preemptive transplant was considered for inclusion in the models. The definition of preemptive transplantation was based on the variable PRTXDIAL from the SAF.TXUNOS file, as was done by other researchers.[21] In addition, since the PRTXDIAL variable was not collected before 1995, we defined preemptive transplant from the SAF.RXHIST60 file, based on duration of ESRD and use of dialysis prior to the transplant of interest, as described before.[18] The recipient's dialysis network was used as a proxy for geographic location.
3. Recipient comorbidity status was described by a composite comorbidity index similar to the one proposed by Davies, which has been shown to be strongly associated with the outcome in ESRD patients.[24] Other comorbidity indices have been proposed in the literature, and since it has been demonstrated that Khan, Davies, and Charlson scores are appropriate for expressing the prognostic impact of comorbidity on mortality risk in patients with ESRD,[25,26] Davies's approach was selected for its simplicity. Also, the specific comorbid conditions used as separate variables were considered for the model: presence and duration of hypertension and diabetes mellitus; history of coronary artery disease; symptomatic cerebrovascular disease; symptomatic peripheral vascular disease; history of malignant tumors; recipient medical conditions at listing; and functional status before transplant. Information about coexisting conditions was obtained from the SAF.TXUNOS file, which was collected from the Transplant Candidate Registration Form before transplant (at the time of listing for the most recent transplant).
4. Donor variables, that is, type of donor (deceased or living), age, race, gender, height, weight, and donor health conditions before donation (*i.e.*, presence and duration of comorbidities: diabetes mellitus, hypertension, coronary artery disease; smoking history; heart beating or not; donor cause/mechanism of death), were obtained from SAF.TXUNOS file.
5. Transplant procedure variables were also obtained from SAF.TXUNOS file: cold ischemia time, transplant procedure type (*e.g.*, single-kidney, kidney-pancreas, double-kidney transplant), transplant center where surgery was done, donor and recipient HLA match, maintenance immunosuppressive therapy at the time of discharge from the hospital (latter was obtained from the SAF.TXIRUNOS file).

Variables Selection

We used several strategies to select the optimal combination of predictors for the model. The selection criteria were based on the predictive value of the variable weighted against the practicality of including it in the model. Even though the longer list of the predictors may potentially improve the outcome of the model, using too many variables may compromise the parsimony and practical usefulness of the model in the clinical setting. In particular, since the decision support tool might potentially be used in the pretransplant clinical environment, only variables available before transplantation were used in developing prediction algorithms.

Survival Analysis. We performed the survival analysis by using proportional hazards regression modeling for the pur-

pose of identifying the set of statistically significant predictors of graft ($p < 0.05$). For the survival analysis, where outcomes were analyzed as time to event, allograft outcome was censored at the earliest of the following events: loss to follow-up, patient death, or study completion date (December 31, 2000) and was analyzed as days to graft failure or censor. For the purpose of variables selection, the survival analysis was supplemented by the logistic regression models as described below.

Logistic Regression Modeling for Variables Selection. We generated five separate logistic regression models predicting the graft survival as a binary variable at 1, 3, 5, 7, and 10 years of the follow-up. We used a conservative approach to variable selection. Only variables that had significant association ($p < 0.05$) with the outcome in all of 5 models were included in the final tree-based analysis. In other words, variables that were not significant in at least one model were excluded.

Additional Variables. In addition to the variables selected by the algorithms described above, we also included several variables that were originally excluded. These variables were considered to be important for the graft outcome prediction: recipient history of unstable angina, predominant renal replacement therapy modality in the pretransplant course and percent time on peritoneal dialysis,[17] recipient history of hypertension, recipient gender, and donor gender.

Additional Selection. Using the set of variables selected by these methods, we tested the tree-based model for convergence and demonstrated poor performance, which were thought to be due to potential collinearity in the data. To make the model more practical and parsimonious, we evaluated the performance of the model with the shorter list of variables, excluding the variables that were considered nonessential. Heartbeating donor variable was found to have significant missing information, whereas nonmissing data were collinear with donor type (living versus deceased). Variables describing cardiovascular disease history were collinear with the variable describing peripheral vascular disease history and therefore the latter was removed. The variable describing the use of antihypertensive medications by the donor was largely homogenous and was also removed. RRT modality immediately before transplant was not used, and instead predominant RRT modality during ESRD course was included in the model. We also excluded the variable describing dialysis network because the model did not converge in its presence. Based on R^2 statistics, the model based on the shorter list of variables (below) performed not worse than the one less parsimonious based on the longer list of predictors.

Final List of Predictors. The final list included the following recipient variables: recipient race, gender, age, height, weight, recipient having a transplant before the current one (yes/no), total number of transplants (including the current one), the time recipient has been on the list before transplant, predominant renal replacement therapy modality, percent time on peritoneal dialysis before transplant, number of renal replacement therapy modalities used before transplant, specific combination of renal replacement therapy modalities, recipient comorbidity score, history of cardiovascular disease, history of unstable angina, history of diabetes, history of hypertension, presence of hepatitis B core antibodies, presence of hepatitis C antibodies, peak and most recent level of panel reactive antibodies, and primary source of pay for medical services. In addition, the following donor variables were used in the final model: donor race, gender, age, height, weight, donor type (living or deceased).

Finally, we used the following transplant procedure variables: cold ischemia time and number of matched HLA antigens, using MMF in the immunosuppressive regimen (as a proxy for the transplant era).

Statistical Analysis and Prediction Models

Continuous variables were summarized by using means and standard deviations. A tree-based model analysis has been extensively described elsewhere[27] and was previously used by our group in the prediction of renal function of diabetics[28] and in the prediction of kidney allograft survival.[20] Briefly, tree-based modeling, also called classification and regression trees, or CART, is a form of binary recursive partitioning that systematically separates data into two groups by using regression of a single factor on the outcome. Unlike traditional methods, tree-building techniques are ideally suited for the development of a reliable clinical decision rule, which can be used to classify new patients into categories according to predicted allograft outcomes, where traditional statistical methods are sometimes cumbersome to use or of limited utility.[27] Tree-based modeling works well when the regression variables are a mixture of categorical and continuous variables. The algorithm is nonparametric, so no assumptions are made regarding the underlying distribution of values of the predictor variables. Tree-based modeling requires relatively little input from the analyst, as the outcome is presented in a form of binary trees and easy to interpret by a nonstatistician. However, the model is limited in that the partitioning method leads to the predicted value being presented in a discrete format, which may not make full use of the information that continuous variables can provide.[27]

Validation and Performance Testing. To test the performance of the models, prediction algorithms were applied to the testing dataset, and the values of the predicted probability of graft failure were generated and compared with actual values of graft outcome. Two measures were used for the validation of the prediction models. The probability of graft failure predicted on a testing set was categorized into deciles, and for each category the rate of graft failure was calculated and compared with the predicted value.[20] We also used receiver operating characteristic (ROC) curve analysis to evaluate and compare the performance of the models. ROC (probability that for a randomly chosen pair of patients the predicted and observed graft survival are concordant) analysis is a nonparametric method used to quantify the accuracy of the prediction. It is a plot of the true-positive rate against the false-positive rate for the different possible cut-points of a prediction algorithm. It shows the tradeoff between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity). The closer the curve follows the left-hand border and then the top border of the ROC space (resulting in large area under ROC curve: an area of 1 represents a perfect prediction), the more accurate the model. The closer the curve comes to the 45-degree diagonal of the ROC space (resulting in a smaller area under ROC curve: an area of 0.5 represents a worthless prediction), the less accurate the model. The procedure ROCCOM in the software package

Table 1. Baseline Characteristics of the of Kidney Transplant Recipients (n = 92,844) at the Time of the Most Recent Transplantation*

Recipient Characteristics	
Age (yr)	43.3 ± 14.2
Gender (male)	60.3%
Race (white, African American, Asian, Native American)	70.2%, 23.0%, 3.4%, 0.9%
Weight (kg)	72.6 ± 17.2
Height (cm)	169.0 ± 13.7
Primary cause of end-stage renal disease	
Diabetes mellitus	25.2%
Hypertension	17.2%
Glomerulonephritis	25.8%
Cystic disease	7.6%
Other	24.2%
Comorbidity score†	0.8 ± 0.8
History of diabetes	27.2%
History of hypertension	52.5%
Total duration of end-stage renal disease (yr)	3.1 ± 3.6
Percent of end-stage renal disease duration time on peritoneal dialysis‡	22.8 ± 38.0
Percent of end-stage renal disease duration time on hemodialysis‡	67.3 ± 41.5
Percent of total end-stage renal disease duration with transplant‡	6.1 ± 20.1
Renal replacement therapy modality immediately prior to transplant	
Hemodialysis	71.3%
Peritoneal dialysis	21.8%
Transplant (dialysis free re-transplant)	1.1%
Unknown	5.8%
Predominant renal replacement therapy modality§	
Hemodialysis	67.3%
Peritoneal dialysis	22.6%
Transplant	6.4%
None	3.6%
Total number of transplants (including the current one)	1.2 ± 0.4
Time on the transplant list (yr)	1.3 ± 1.1
Peak panel reactive antibody level (%)	12.1 ± 21.5
Most recent panel reactive antibody level (%)	5.3 ± 14.7
Number of matched HLA antibodies	1.8 ± 1.5
Cold ischemia time (h)	15.5 ± 8.7
Transplant day of the week‖	4.0 ± 1.8
History of previous kidney transplant(s)	12.6%
Donor characteristics	
Age (yr)	34.4 ± 15.5
Gender (male)	56.2%
Race (white, African American, Asian, Native American)	82.5%, 11.5%, 1.3%, 0.4%
Weight (kg)	72.8 ± 19.0
Height (cm)	164.3 ± 21.9
Terminal serum creatinine level (mg/dL)	0.9 ± 0.3
Terminal blood urea nitrogen level (mg/dL)	12.1 ± 6.1
Living donors	24.8%

* Continuous variables presented as mean ± SD.

† The comorbidity score used in our study was calculated on the basis of the following coexisting conditions, each of them contributing 1 point: cardiovascular disease (defined in USRDS as symptomatic cardiovascular disease or angina/coronary artery disease), symptomatic peripheral vascular disease, diabetes mellitus, and hypertension.

‡ Information obtained from USRDS RXHIST file; due to missing/unknown data and "60 d rule" convention adopted by USRDS (see text) the total is less than 100%.

§ Predominant renal replacement therapy modality defined as a modality used for >50% of the duration of end-stage renal disease.

‖ Transplant day of the week expressed in numbers starting with Sunday (1 = Sunday, 2 = Monday, *etc.*).

STATA (Stata Corporation, College Station, TX) was used to calculate and compare the area under the ROC curves.

Software. SAS (SAS Institute, Cary, NC) was used for descriptive statistics and survival analysis; S-Plus (Insightful, Seattle, WA) was used for logistic regression and tree-based modeling,[20,28] and STATA (Stata Corporation, College Station, TX) was used for ROC analysis.

Results

Descriptive Statistics

Data were collected from USRDS and included 92,844 records of patients receiving kidney or kidney-pancreas transplants starting January 1, 1990, and through December 31, 1999, with the follow-up period through December 31, 2000. The study population characteristics are presented in **Table 1**. The average age of patients was 43.3 years, of which 60.3% were male, 70.2% were white, 27.2% were diabetic, 77.1% were on hemodialysis before transplant, and 12.6% had another kidney transplant before the current transplant. During the 11 years of the study, the graft failed in 34.9% of the patients. Cold ischemia time was on average 15.5 hours.

Tree-Based Model Generation

Tree-Based Model Design. Five different tree-based models predicting the probability of the allograft survival for 1, 3, 5, 7, and 10 years were generated. Tree-based models were initially generated without restrictions by using a limited list of independent variables described above. To generate final, more parsimonious models, the optimal number of terminal nodes was determined for each model using the cross-validation procedure, where the deviance was plotted against the size of the tree to select the optimal tree size. The optimal size of the tree was identified as 93 for the model predicting 1-year survival, 40 for 3-year survival, 88 for 5-year survival, and 65 for 7-year survival. The cross-validation procedure did not indicate the optimal tree-size for the 10-year outcome model; therefore, we arbitrarily selected the model with 65 terminal nodes (the same as for 7-year outcome prediction). After that, the second set of tree models was generated and pruned to the size identified by the cross-validation procedure. After the models were created, the set of predicted outcome values was generated in the testing datasets. The residual mean deviance of the model and misclassification error rate are presented in **Table 2**.

Table 2. Residual Mean Deviance, Misclassification Error, and Area Under the Receiver Operating Characteristic Curve of the Prediction Models

Model	Residual Mean Deviance	Misclassification Error*	Area Under ROC (Standard Error)
One-year survival	0.76	0.14	0.626 (0.0048)
Three-year survival	1.13	0.27	0.640 (0.0041)
Five-year survival	1.18	0.32	0.717 (0.0038)
Seven-year survival	0.89	0.25	0.830 (0.0031)
Ten-year survival	0.35	0.08	0.901 (0.0043)

* The 50% cut-point of the predicted probability of graft survival was used to convert it into binary variable to calculate the misclassification error.

Table 3. Predicted Probabilities and Actual Graft Survival Rates

1-yr survival									
Predicted probability of survival	0–30%	31–50%	51–60%	61–70%	71–80%	81–90%	91–100%		
Observed percent survival	51.5	52.5	76.5	69.0	78.9	85.7	91.2		
n	33	139	162	924	4479	15598	11509		
3-yr survival									
Predicted probability of survival	0–40%	41–50%	51–60%	61–70%	71–80%	81–90%	91–100%		
Observed percent survival	38.8	47.4	54.9	67.1	77.0	82.7	89.3		
n	474	274	1914	7907	8598	3746	759		
5-yr survival									
Predicted probability of survival	0–10%	11–20%	21–30%	31–40%	41–50%	51–60%	61–70%	71–80%	81–100%
Observed percent survival	8.2	17.2	25.6	41.2	45.5	57.0	64.0	73.6	81.0
n	981	1123	520	1241	1777	4092	4262	2791	1218
7-yr survival									
Predicted probability of survival	0–10%	11–20%	21–30%	31–40%	41–50%	51–60%	61–70%	71–80%	81–100%
Observed percent survival	0.9	18.1	30.7	38.9	46.0	53.1	64.9	74.3	59.2
n	5307	548	807	792	2860	2537	2541	1323	76
10-yr survival									
Predicted probability of survival	0–10%	11–20%	21–30%	31–40%	41–50%	51–60%	61–100%		
Observed percent survival	1.2	16.1	26.4	34.9	36.9	47.0	62.9		
n	7488	1181	367	521	485	202	35		

Model Validation

Correlation Analysis. The predicted variable in this study is the probability of graft survival, which is a continuous variable. However, the actual outcome for each individual patient is binary. All records were divided into 10 groups, based on predicted probability of graft survival using the following cut-points: 0% to 10%, >10% to 20%, >20% to 30%, >30% to 40%, >40% to 50%, >50% to 60%, >60% to 70%, >70% to 80%, >80% to 90%, and >90% to 100%. The observed graft survival was calculated for each group and compared with the predicted probability. If the number of patients in a particular group was low (arbitrarily selected value of <30), it was merged with next group up, except for the very last group, and that was merged with the next group down. In particular, for the 1-year prediction group, the models did not make any predictions with the probability of graft survival 0% to 10% and 11% to 20%; therefore, this group was merged with the group where the predicted probability of graft survival was 21% to 30%. Similarly, the 31% to 40% group had only 7 patients and therefore was merged with the 41% to 50% group. In the 3-year prediction model, none of the groups with predicted probability between 0% and 30% had any patients and therefore were merged with the 31% to 40% group. In the 5-year model, the last group 91% to 100% had 13 patients and was merged with the 81% to 90% group. For the 7-year prediction, the 91% to 100% group had only 21 patients and was merged with the 81% to 90% group. Finally, for the 10-year prediction, none of the groups >60% had enough patients and were merged together in the 61% to 100% group.

The results of the analysis are presented in **Table 3**, where the percent of actual graft survival and number of patients for each of the groups of predicted probability of graft survival are presented. These results are illustrated in **Figure 1**.

The midpoint of each group's probability range was used as the predicted percent survival for the group and compared with observed graft survival for the group by correlation analysis. The prediction of the probability of graft survival from the training model achieved a good correlation with the observed survival of the testing set with $r = 0.94$ for 1-year survival prediction, $r = 0.98$ for 3-year survival prediction, $r = 0.99$ for 5-year survival prediction, $r = 0.93$ for 7-year survival prediction, and $r = 0.98$ for 10-year survival prediction.

Receiver Operator Characteristics Curve Analysis. The ROC analysis was performed for each model by using the predictions generated on the testing dataset. The ROC curves are presented in **Figure 2**. The area under the ROC curve was calculated for each model by using the prediction data generated on the testing dataset. All models achieved a reasonable prediction accuracy on the independent testing dataset. For 1-year prediction, the area under the ROC curve was 0.63; for 3-year prediction, 0.64; for 5-year prediction, 0.71; for 7-year prediction, 0.82; and for 10-year prediction, 0.90.

Discussion

Factors affecting kidney allograft survival were evaluated previously, based on both local and national databases. Other authors attempted to generate prediction models of the transplant outcome. A neural network model was used to predict the outcome of liver transplant[29,30] and delayed graft function after renal transplantation.[31] Multivariate modeling was used to predict living graft recipients' creatinine, based on four parameters: recipient age, body mass index, creatinine clearance, and degree of relationship.[32]

As far as previous studies are concerned, we found one paper in which investigators used multivariate modeling to predict the outcome of the transplantation to optimize deceased kidney allocation decision making in a northern Italy transplant program.[33] However, to the best of our knowledge, aside from our report in 2003,[20] no other investigators have used working prediction models to study long-term renal allograft outcomes.

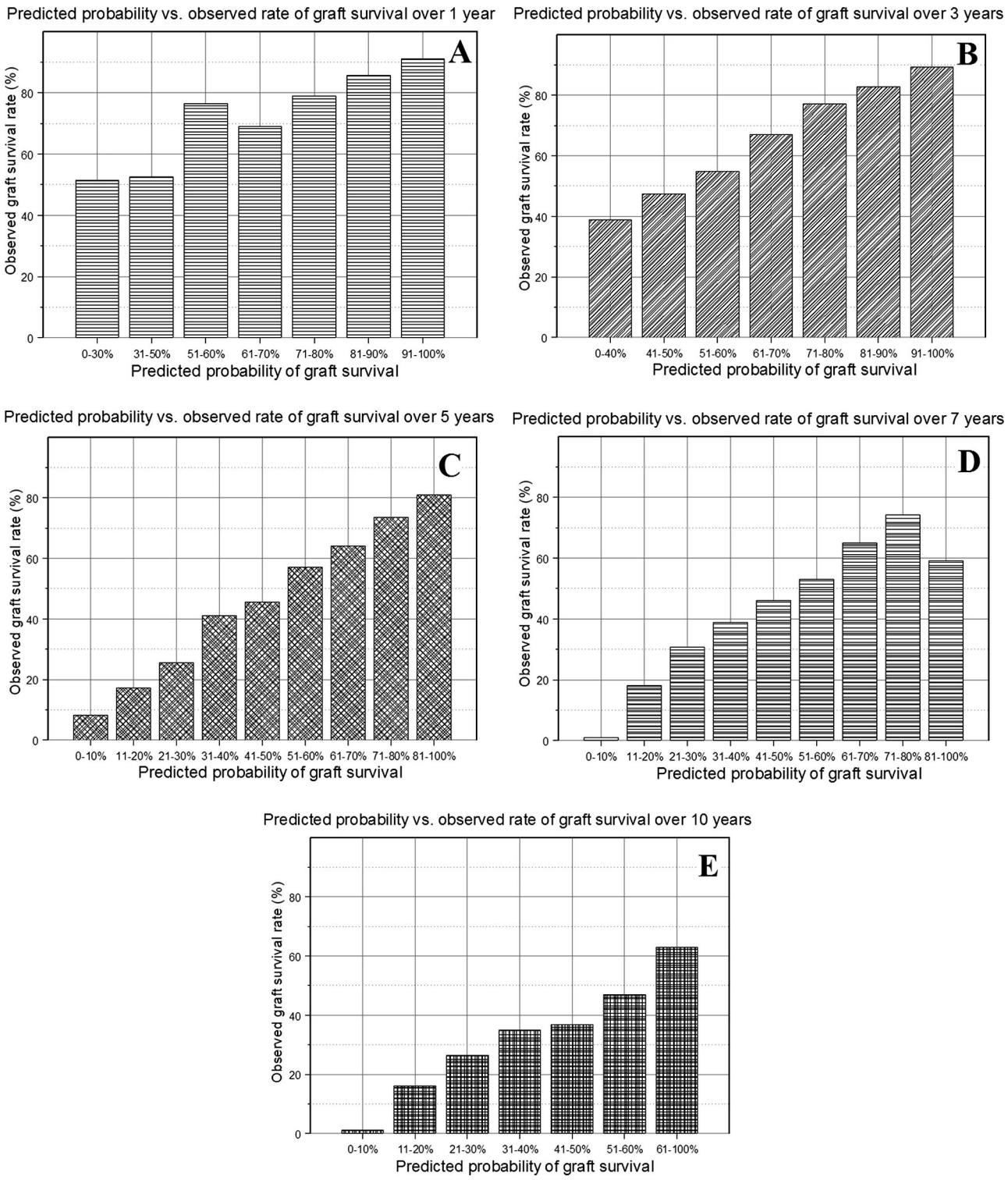

Figure 1. Bar plots of the graft survival rates versus predicted probability of graft survival for 1 (A), 3 (B), 5 (C), 7 (D), and 10 (E) years of graft survival. Predictions were generated in the independent testing dataset, separate from the training dataset on which the models were created.

We previously presented mathematical models predicting the probability of 3-year kidney allograft survival from a deceased donor based on UNOS data.[20] That model, however, had several limitations. The model was based on the deceased donors only (as opposed to the current model based on both deceased and living donor kidney transplants), it used a very limited set of predictors, and it was based on the relatively old dataset. In addition, the previously reported model predicted only 3-year allograft survival, whereas the currently reported model predicts the probability of 1-, 3-, 5-, 7-, and 10-year

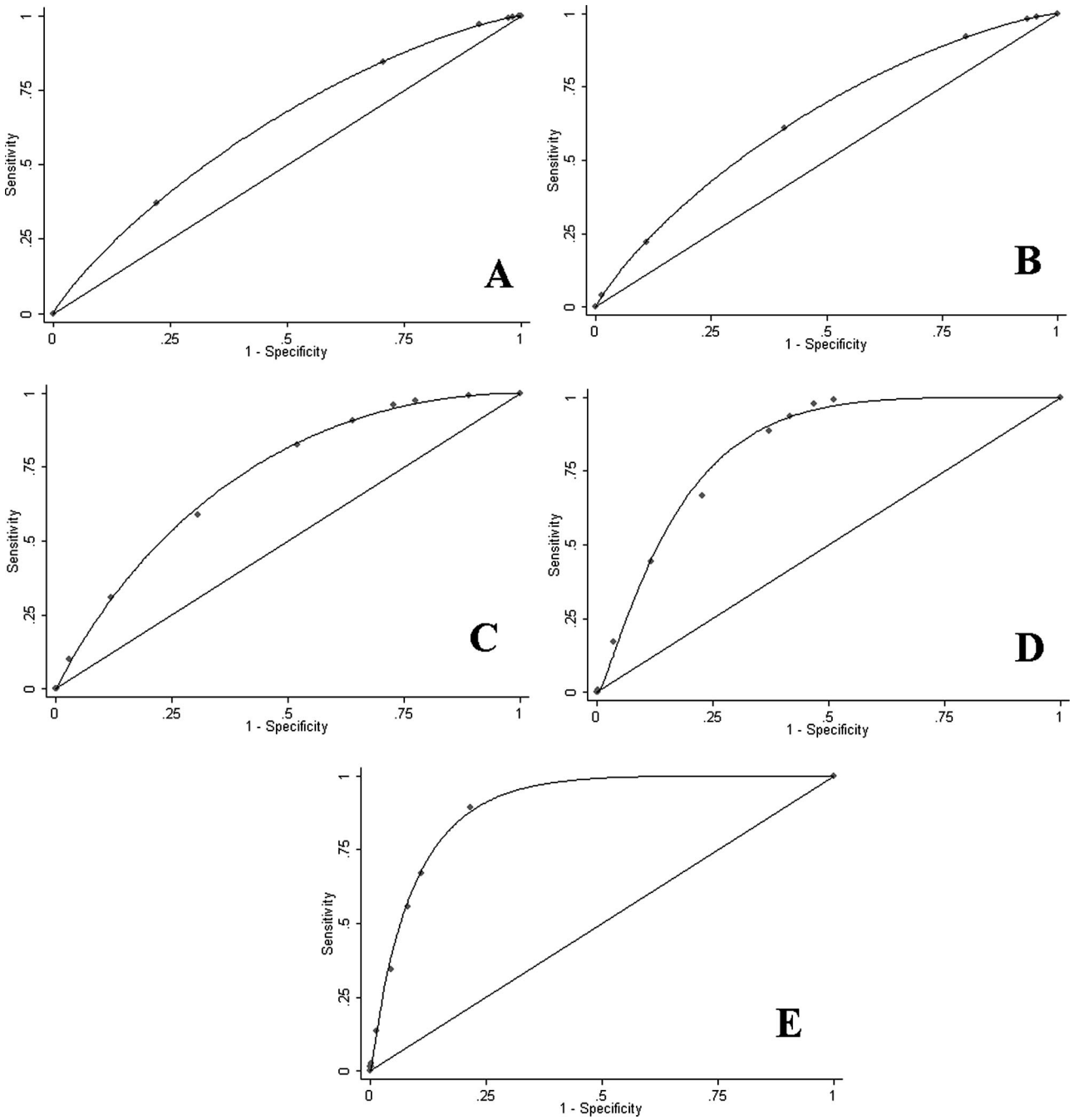

Figure 2. ROC curves for the prediction models of the 1 (A), 3 (B), 5 (C), 7 (D), and 10 (E) years of graft survival. ROC curves were generated in the independent-testing dataset, separate from the training dataset on which the models were created.

graft survival. The previous analysis was also challenging in the face of a relatively large amount of missing data. The current model is based on the more recent dataset, representing clinical practice modalities of the late 1990s. In recent years, the quality of data has improved, especially since the introduction of UNet, the online transplant data entry system that was implemented in October 1999. In the current study, we used the data supplied by USRDS, which, in addition to the UNOS data, has information regarding patient dialysis course, more detailed patient comorbidity data, and more comprehensive information on patients' demographics. In designing the study, our intent was to develop a prediction model to be used in the pretransplant setting; therefore, we excluded posttransplant variables that were not available until after the transplant procedure. We also did not analyze the impact of immunosuppressive therapy, immediate posttransplant graft function, and episodes of acute rejection, since this information was not available before the transplantation procedure.

The tree-based modeling used in this study represents a relatively new approach compared with conventional regression analysis of the data. This nonparametric modeling works when the regression variables are a mixture of categorical and

continuous variables in that it identifies "splitting" variables based on an exhaustive search of all possibilities, even in problems with many hundreds of possible predictors. Simultaneously, it requires relatively little input from the analyst. This graphical algorithm, presented as a collection of simple binary rules, is much simpler to interpret by a nonstatistician than the multivariate logistic regression. Prediction algorithms evaluated in this study can potentially be used in recipient counseling and decision making processes regarding renal transplants. Tree-based modeling is easy to implement in the computer-based decision support system to be used in the pretransplant clinic. In addition, it can be used as a tool to identify patients at risk for premature graft failure and to model different clinical situations, where the modifiable factors of the recipient, donor, and transplant procedures can be optimized. The identification of factors that play an important role in graft survival helps to focus efforts of transplant programs on certain individual aspects of patient care.

Certain limitations should be considered while interpreting the results of this project. The dataset used covers the time period of the last 11 years of the last century. One should realize that there will always be a time gap, and the data cannot be very recent, as a certain period of follow-up is necessary. Although this is the case, changes in clinical practice should be considered by the reader and potential users of the model. Another limitation is the discrete type of the tree-based model output. Although the number of terminal nodes was relatively high, the output information is still limited due to the noncontinuous nature of the predicted probability. Finally, as in every other analysis of the large registry data, the quality of the data is of concern. That concern has been alleviated, however, by the recent improvement in the UNOS data collection techniques. Also, the relatively good performance of the models indirectly indicates the reasonable quality of the input data.

In conclusion, we developed and validated tree-based models to predict allograft survival in patients with kidney transplants. The models predicting the probability of 1-, 3-, 5-, 7-, and 10-year allograft survival have been validated on the independent dataset and demonstrated performance that may suggest implementation in the clinical decision support system. Evaluating these models in a prospective study may be the subject of a future project.

Acknowledgments

This study was funded by the Dialysis Research Foundation of Utah. The data reported here have been supplied by the USRDS. The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as official policy or interpretation of the US government.

References

1. Ojo AO, Hanson JA, Meier-Kriesche HU, *et al*: Survival in recipients of marginal cadaveric donor kidneys compared with other recipients and wait-listed transplant candidates. *J Am Soc Nephrol* 12: 589–597, 2001.
2. Ojo AO, Wolfe RA, Agodoa LY, *et al*: Prognosis after primary renal transplant failure and the beneficial effects of repeat transplantation: multivariate analyses from the United States Renal Data System. *Transplantation* 66: 1651–1659, 1998.
3. United Network for Organ Sharing Web site. Organ procurement and transplantation network data. Available at: http://www.unos.org. Last accessed December 12, 2005.
4. Remuzzi G, Grinyo J, Ruggenenti P, *et al*: Early Experience with dual kidney transplantation in adults using expanded donor criteria. *J Am Soc Nephrol* 10: 2591–2598, 1999.
5. Lu AD, Carter JT, Weinstein RJ, *et al*: Excellent outcome in recipients of dual kidney transplants: a report of the first 50 dual kidney transplants at Stanford University. *Arch Surg* 134: 971–975, 1999.
6. VanBiesen, Vanholder R, Lameire N: Impact of pretransplantation dialysis modality on patient outcome after renal transplantation: the role of peritoneal dialysis revisited. *Peritoneal Dialysis International* 19: 103–106, 1999.
7. McLaren AJ, Jassem W, Gray DW, *et al*: Delayed graft function: risk factors and the relative effects of early function and acute rejection on long-term survival in cadaveric renal transplantation. *Clin Transplant* 13: 266–272, 1999.
8. Jordan ML, Shapiro R, Vivas CA, *et al*: High-risk donors: expanding donor criteria. *Transplant Proc* 31: 1401–1403, 1999.
9. Gjertson DW, Cecka JM: Determinants of long-term survival of pediatric kidney grafts reported to the United Network for Organ Sharing kidney transplant registry. *Pediatr Transplant* 5: 5–15, 2001.
10. Ishitani M, Isaacs R, Norwood V, *et al*: Predictors of graft survival in pediatric living-related kidney transplant recipients. *Transplantation* 70: 288–292, 2000.
11. Tejani A, Stablein DM, Donaldson L, *et al*: Steady improvement in short-term graft survival of pediatric renal transplants: the NAPRTCS experience, in *Clin Transpl* 1999: 95–110, 1999.
12. Alexander JW, Bennett LE, Breen TJ: Effect of donor age on outcome of kidney transplantation. *Transplantation* 57: 871–876, 1994.
13. Ojo AO, Leichtman AB, Punch JD, *et al*: Impact of pre-existing donor hypertension and diabetes mellitus on cadaveric renal transplant outcomes. *Am J Kidney Dis* 36: 153–159, 2000.
14. Cho YW, Cecka JM, Gjertson DW, *et al*: Prolonged hypertension (>10 years) is a significant risk factor in older cadaver donor renal transplants. *Transplant Proc* 31: 1283, 1999.
15. Sánchez-Fructuoso AI, Prats D, Torrente J, *et al*: Renal transplantation from non-heart beating donors: a promising alternative to enlarge donor pool. *J Am Soc Nephrol* 11: 350–358, 2000.
16. Mouquet C, Benalia H, Chartier-Kastler E, *et al*: Renal retransplantation in adults: comparative prognostic study. *Progres en Urologie* 9: 239–243, 1999.
17. Goldfarb-Rumyantzev AS, Hurdle JF, Scandling JD, *et al*: The role of pretransplantation renal replacement therapy modality in kidney allograft and recipient survival. *Am J Kidney Dis* 46: 537–549, 2005.
18. Goldfarb-Rumyantzev A, Hurdle JF, Scandling J, *et al*: Duration of end-stage renal disease and kidney transplant outcome. *Nephrol Dial Transplant* 20: 167–175, 2005.
19. Naiman N, Baird BC, Isaacs RB, *et al*: Role of pre-transplant marital status in renal transplant outcome. *Clin Transplant* 21: 38–46, 2007.
20. Goldfarb-Rumyantzev AS, Scandling JD, Pappas L, *et al*: Prediction of 3-year cadaveric graft survival based on pre-transplant variables in a large national dataset. *Clin Transplant* 17: 485–497, 2003.
21. Kasiske BL, Snyder JJ, Matas AJ, *et al*: Preemptive kidney transplantation: the advantage and the advantaged. *J Am Soc Nephrol* 13: 1358–1364, 2002.
22. Meier-Kriesche H, Port FK, Ojo AO, *et al*: Deleterious effect of waiting time on renal transplant outcome. *Transplant Proc* 33: 1204–1206, 2001.
23. Mange KC, Joffe MM, Feldman HI: Effect of the use or nonuse of long-term dialysis on the subsequent survival of renal transplants from living donors. *N Engl J Med* 344: 726–731, 2001.
24. Davies SJ, Russell L, Bryan J, *et al*: Comorbidity, urea kinetics, and appetite in continuous ambulatory peritoneal dialysis patients: their interrelationship and prediction of survival. *Am J Kidney Dis* 26: 353–361, 1995.
25. Van Manen JG, Korevaar JC, Dekker FW, *et al*: How to adjust for comorbidity in survival studies in ESRD patients: a comparison of different indices. *Am J Kidney Dis* 40: 82–89, 2002.

26. Van Manen JG, Korevaar JC, Dekker FW, *et al*: Adjustment for comorbidity in studies on health status in ESRD patients: which comorbidity index to use? *J Am Soc Nephrol* 14: 478–485, 2003.
27. Breiman L, Friedman JH, Olshen RA, *et al*: *Classification and Regression Trees.* Monterey, CA, Wadsworth and Brooks/Cole, 1984.
28. Goldfarb-Rumyantzev AS, Pappas L: Prediction of renal insufficiency in Pima Indians with nephropathy of type 2 diabetes mellitus. *Am J Kidney Dis* 40: 252–264, 2002.
29. Doyle HR, Dvorchik I, Mitchell S, *et al*: Predicting outcomes after liver transplantation: a connectionist approach. *Ann Surg* 219: 408–415, 1994.
30. Parmanto B, Doyle HR: Recurrent neural networks for predicting outcomes after liver transplantation: representing temporal sequence of clinical observations. *Methods Inf Med* 40: 386–391, 2001.
31. Brier ME, Ray PC, Klein JB: Prediction of delayed renal allograft function using an artificial neural network. *Nephrol Dial Transplant* 18: 2655–2659, 2003.
32. Zapletal C, Lorenz MW, Woeste G, *et al*: Predicting creatinine clearance by a simple formula following live-donor kidney transplantation. *Transpl Int* 17: 490–494, 2004.
33. Poli F, Scalamogna M, Cardillo M, *et al*: An algorithm for cadaver kidney allocation based on a multivariate analysis of factors impacting on cadaver kidney graft survival and function. *Transpl Int 13 Suppl* 1: S259–S262, 2000.