Clin Transplant 2003: 17: 485–497

Clinical Transplantation

Review Article

Prediction of 3-yr cadaveric graft survival based on pre-transplant variables in a large national dataset

Goldfarb-Rumyantzev AS, Scandling JD, Pappas L, Smout RJ, Horn S. Prediction of 3-yr cadaveric graft survival based on pre-transplant variables in a large national dataset.
Clin Transplant 2003: 17: 485–497. © Blackwell Munksgaard, 2003

Abstract: Pre- and post-transplant predictive factors of graft survival for optimal and expanded criteria grafts have been studied in the past. The goal of our study was to evaluate the recent large set of United Network of Organ Sharing records (1990–1998) to generate a prediction algorithm of 3-yr graft survival based on pre-transplant variables alone. The dataset of patients with end-stage renal disease and cadaveric kidney or kidney–pancreas transplantation (1990–1998) used in the study consisted of 37 407 records. Logistic regression (LM) and a tree-based model (TBM) were used to identify predictors of 3-yr allograft survival and to generate prediction algorithm. Donor and recipient demographic characteristics (age, race, and gender) and body mass index showed non-linear, while human leukocyte antigen match showed strong linear relationships with 3-yr graft survival. Prediction of the probability of graft survival from the model, achieved a good match with the observed survival of the separate dataset, with a correlation of $r = 0.998$ for LM and $r = 0.984$ for TBM. The positive predictive value (PV) of allograft survival with LM and TBM was 76.0% and the negative PV was 63 and 53.8% for LM and TBM, respectively. Both LM and the TBM can potentially be used in clinical practice for long-term prediction of kidney allograft survival based on pre-transplant variables.

Alexander S Goldfarb-Rumyantzev[a], John D Scandling[b], Lisa Pappas[a], Randall J Smout[c] and Susan Horn[c]

[a] Division of Nephrology, University of Utah Health Sciences Center, Salt Lake City, UT, [b] Division of Nephrology, Kidney and Pancreas Transplant Program, Stanford University Medical Center, Stanford, CA and [c] Institute for Clinical Outcomes Research & ISIS Inc., Salt Lake City, UT, USA

Key words: graft survival – kidney transplant – logistic regression – outcome – prediction model – tree-based model – United Network of Organ Sharing

Corresponding author: Alexander Goldfarb-Rumyantzev, Division of Nephrology and Hypertension University of Utah Health Sciences Center, 85 North Medical Drive, East Rm 201, Salt Lake City, UT 84112, USA.
Tel.: +801 585 9455; fax: +801 585 3830;
e-mail: alex.goldfarb@hsc.utah.edu

Accepted for publication 6 December 2002

Improved immunosupression has reduced acute rejection, but has had little effect on chronic allograft nephropathy and late graft loss (1). Causes of long-term allograft failure are recurrent disease and chronic allograft nephropathy. Pre- and post-transplant predictive factors of graft survival for optimal and expanded criteria grafts have been extensively studied in adults (2, 3) and children (4–6), based on data from the United Network of Organ Sharing (UNOS) and North American Pediatric Renal Transplant Cooperative Study (NAPRTCS). Donor age (7, 8), hypertension (HTN) and diabetes (9), non-heartbeating donor (10), cold storage time (11), body mass index (BMI) of donor and recipient (12), and high degree of donor vascular pathology (13) were associated with worse outcomes in multivariate analyses. Recipient factors important to graft survival include recipient's general health (14), race (15), underlying kidney disease (16) and previous treatment modalities. Re-transplant (17), multiple (> 5) pre-transplant blood transfusions (18), human leukocyte antigen-B and DR mismatch and advanced recipient age (19) carry a higher risk of allograft failure. Pre-transplant dialysis modality may impact patient outcome (20), while pre-emptive transplantation of kidneys from living donors is associated with longer allograft survival (21). Larger centers (> 1000 transplants/10 yr) have a slightly higher graft survival than all other centers

(22, 23). Attempts have been made to develop prediction models of graft survival (mostly short-term) (24) based on data available using different statistical models, such as Cox regression (25), and artificial neural networks (26). The goal of our study was to evaluate the set of UNOS records (1990–1998) to identify pre-transplant factors affecting 3-yr allograft survival to generate a prediction models that would accurately identify patients at risk for 3-yr allograft failure using logistic regression (LM) and tree-based algorithm. To assure practical use of prediction model in pre-transplant evaluation and recipient counseling only variables available in pre-transplant period were used.

Patients and methods

Dataset

We selected patients with end-stage renal disease (ESRD) who underwent kidney or kidney–pancreas transplantation between 1990 and 1998 from the US Scientific Registry of Transplant Recipients supplied by UNOS. The dataset includes transplants performed in infants and young children (minimal age, <1 yr; maximum age 98 yr). To protect patient privacy follow-up dates and transplant dates were shifted randomly to $\pm$ 1–180 d. Dates were shifted by the same amount for any given record so that the difference between the dates is preserved. Independent variables available for analysis included: age, gender, race, height, and weight for both donor and recipient, recipient cause of ESRD, type of pre-transplant renal replacement therapy, number of previous kidney transplants and pre-transplant blood transfusions, recipients' most recent creatinine and donors' terminal creatinine, history and duration of diabetes and HTN in the donor, number of HLA match and mismatch, cold ischemia time, kidney or kidney–pancreas transplant, and transplant center code. The outcome variable was 3-yr graft survival, end-point was defined as allograft failure. Patient death with functioning graft was not included in the definition of graft failure. Information regarding most of the variables used in the analysis was collected by UNOS beginning October 1, 1987. However, several variables (donor history and duration of diabetes and HTN, recipients' most recent serum creatinine and donors' terminal creatinine) started to be collected by UNOS only since April 1, 1994. Furthermore, donors' most recent creatinine was collected only for non-dialyzed patients between April 1, 1994 and October 25, 1999 and started to be collected for all patients only after October 25, 1999.

Data cleaning and imputation

The initial dataset consisted of 102 686 records. Independent variables initially planned to be included in the analysis, but missing a large number of entries, were either eliminated (recipient creatinine missing 89.2% of entries) or categorized (previous number of kidney transplants, donor terminal creatinine, donor duration of diabetes, and donor duration of HTN were missing 87.3, 60.6, 63.7, and 60.7% respectively, and were converted into categorical variables with a separate code for missing values). The following variables were considered erroneous and were replaced with blank values: cold ischemia time $= 0$ ($n = 376$), cadaveric donor creatinine >3 mg/dL or 26.5 mmol/dL ($n = 923$), donor or recipient height <45 cm or >210 cm ($n = 211$, $n = 6$), donor or recipient weight <1 kg or >340 kg ($n = 21$, $n = 29$). A number of records with critical information being incomplete or deemed to be unreliable or erroneous had to be eliminated (Fig. 1). Records missing for both height and weight (for either donor or recipient) were eliminated. In the remaining missing records donor and recipient height and weight were imputed using a tree-based model (TBM) with height (for weight imputation), weight (for height imputation), age,

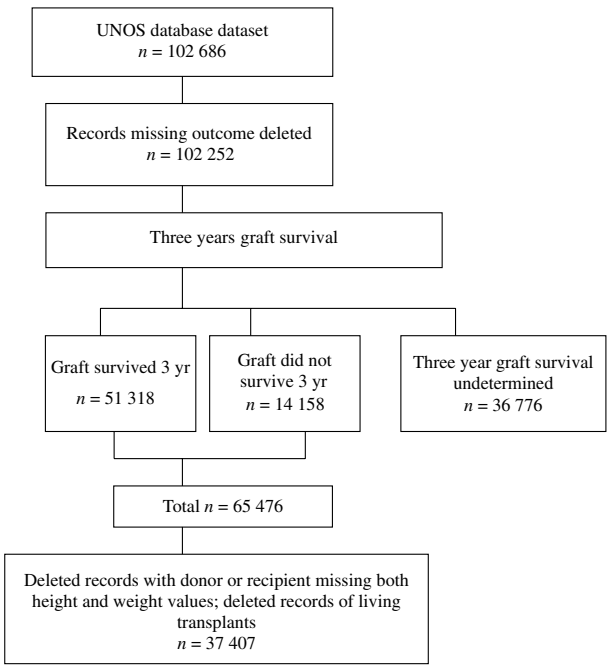

Fig. 1. Process of data elimination from the initial dataset ($n = 102\,686$). Records with missing or undetermined 3-yr allograft survival, missing both height and weight for donor or recipient, records of living donors were deleted. Final dataset consisted of 37 407 records.

gender, and race as an independent variables. The imputation algorithm for recipient height (missing 90% of the values) was tested using 8282 United States Renal Data System (USRDS) patients with complete data for age, weight, height, gender, and race from Dialysis Morbidity and Mortality Study (DMMS) Waves 3 and 4 studies. The dataset consists of a random sample of all ESRD patients on January 1, 1994. As a test for reliability, the intraclass correlation coefficient was calculated using SPSS (SPSS Inc., Chicago, IL, USA). Based on the Landis and Koch guideline for evaluation of the reliability coefficient (27), our correlation coefficient of 0.76, the predicted height values have substantial agreement with the observed heights. Table 1 further illustrates the performance of the imputation algorithm. Records with missing and indeterminate 3-yr outcome were deleted; therefore, the dataset was biased towards a higher proportion of failed grafts. As a result, values of the percent survival have only relative meaning and are used for the purpose of comparison between groups studied. Two of the categorical variables had more than 34 levels (transplant center code, and cause of ESRD). Based on the possibility that transplant center volume may have an effect on the outcome, a five-category variable was created: transplant center codes were grouped into quintiles according to total number of transplants performed between 1990 and 1998 (1 = 1–83; 2 = 84–209; 3 = 210–355; 4 = 356–615; 5 = 616–2529). Causes of ESRD were grouped into deciles by total number of transplants with known 3-yr survival (Appendix). The final dataset consisted of 37 407 records.

Statistical analysis

Bivariate analysis was performed using cross-tabulation and comparison of graft survival in the subgroup using the chi-square test. Friedman supersmoothing method was used to fit the curve in bivariate analysis. Discrimination was determined by area under receiver operating characteristic (ROC) curve and chi-square for LM models. Model calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test. For the purpose of prediction analysis, all records were randomly assigned either to a 'training' set (n = 25 000), used for knowledge acquisition and model development, or to a 'testing' set (n = 12 407), used to validate the models. Predicted probabilities of 3-yr graft survival were generated on a testing set and was compared with the actual patient outcomes. The predicted probability of the graft survival with group-average observed graft survival was used to compare the performance of the models. Also 2 × 2 contingency tables were used to determine positive and negative predictive values (PV). Statistical models used in the analysis included LM and classification trees. In certain situations, traditional statistical methods are poorly suited for complex interactions or detecting patterns in the data. Many possible predictor variables may violate the normality assumptions necessary for parametric analysis. In addition, the results of traditional methods sometimes may be difficult to use. Therefore, along with a traditional regression model that assumes linear relationship between predictors and the outcome, we decided to use a less-commonly used TBM, which does not require the linearity assumption, and was used in clinical prediction before (28). TBM is an exploratory technique for uncovering structure in data which generates a collection of many rules displayed in the form of binary tree (29). We used S-Plus statistical software package (MathSoft, Inc., Seattle, WA) for bivariate analysis and TBM and SAS (SAS Institute, Cary, NC) for LM.

Table 1. Performance of the imputation algorithm for recipient height validated on a separate dataset (USRDS, Wave 3 and 4, n = 8282). Average absolute difference = 6.44 cm, minimum absolute difference = 0 cm, maximum absolute difference = 49.23 cm

Absolute difference between predicted and observed values (cm)	Frequency	Percentage
=5	4015	48.48
>5 and =10	2629	31.74
>10 and =15	1083	13.08
>15 and =20	332	4.01
>20 and =25	119	1.44
>25	104	1.26

Results

Comparison between initial and final datasets

The elimination of the large number of records could potentially bias the dataset therefore, after completing the data cleaning described in the previous section, we compared the final dataset (n = 37 407) with the initial one (n = 102 686). The final dataset had the same donor and recipient mean age, height, weight, number of matched and mismatched antigens, and cold ischemia time as the initial dataset and the same distribution of donors and recipients by gender, dialysis type (Table 2), race, presence of diabetes mellitus (DM) and HTN in donors, number of pretransplant transfusions, and transplant procedure (data not shown). Compared with the initial dataset number of transplant centers in the final datasets has not changed.

Table 2. Descriptive statistics of the initial and final datasets

	Initial dataset[a] (mean ± SD)	Final dataset
Donor age (years)	34.4 ± 15.6 (1067)	32.3 ± 28.7
Donor height (cm)	163 ± 29.7 (45 384)	164.6 ± 25.2
Donor weight (kg)	70.5 ± 23.9 (30 550)	69.8 ± 21.8
Recipient age (years)	41.4 ± 14.5 (3)	42.0 ± 13.7
Recipient height (cm)	168.1 ± 17.1 (79 889)	168.3 ± 11.9
Recipient weight (kg)	71.1 ± 19.6 (28 550)	71.7 ± 18
Number of mismatched HLA	3 ± 1.7 (1613)	3.3 ± 1.6
Number of matched HLA	2.6 ± 1.6 (1613)	2.3 ± 1.5
Cold ischemia time (h)	16.5 ± 12.3 (8789)	22.0 ± 9.7
Donor gender: percent males (%)	56.3	62
Recipient gender percent males (%)	59.8 (1)	60.0
Dialysis type (no dialysis, HD, PD)	8.3, 36.5, 13.3 (43 014)	5.2, 30.9, 12.27
Organ (kidney, kidney–pancreas)	93.8, 6.2 (0)	91.3, 8.7
Diagnosis (percent DM, HTN) (%)	24.3, 15.1	25.0, 16.2
n	102 686	37 407

[a]Values within parenthesis represent number of records missing or reported unknown.
DM, diabetes mellitus; HTN, hypertension; HLA, human leukocyte antigen; HD, hemodialysis; PD, peritoneal dialysis.

Therefore it was concluded that after the elimination of a large number of records the final dataset is still representative of the initial sample.

Bivariate analysis

Donor and recipient characteristics. Young and old donors and recipients have lower 3-yr graft survival ($p < 0.001$) (Fig. 2, panels A and B). There were differences in outcome associated with donor and recipient gender (Table 3) and race (Table 4) (p values presented in the tables). Kidneys from the donors with both DM and HTN had the worst 3-yr survival (59.3%), while those from the donors without either had the best outcome (76.3%). Kidneys from either diabetic or hypertensive donors were roughly in the middle (66.2 and 64.3%, respectively) ($p < 0.001$). Increased duration of HTN and/or diabetes (from 1 to 5 yr by 1-yr increments) in the donor was associated with worse outcome ($p < 0.001$ for both). There is no relationship between donors' terminal creatinine and graft survival. There were differences in outcome associated with different etiologies of renal failure (data not shown). Patients with no dialysis history (pre-emptive transplant) had the best 3-yr graft survival (81.3%, n = 1940) followed by those with history of peritoneal dialysis (76.1%, n = 4,591) and then hemodialysis (73.0%, n = 11 542) ($p < 0.001$). A previous transplant history worsened 3-yr survival in almost a linear fashion with 76.7% survival in recipient with no previous transplant history, 70.9, 62.1, and 56.9% in those with one, two and more than two previous transplants, respectively ($p < 0.001$). Number of pre-transplant transfusions did not significantly affect graft survival in bivariate analysis.

Transplant procedure, matching donor and recipient. The 3-yr survival improves and declines in linear fashion with increasing number of matched and mismatched antigens, respectively ($p < 0.001$). Donor/recipient BMI vs. 3-yr graft survival looks almost like a bell-shape curve with the best outcome associated with the donor/recipient BMI = 1 (Fig. 3, panel A). The worst survival was in grafts from relatively small donors to large recipients ($p < 0.001$). Transplant centers with a low volume of transplants had variable outcome, while in those with high number of transplants the outcome was relatively uniform (Fig. 3, panel B). There was slight downward trend in relation of 3-yr graft survival to cold ischemia time (Fig. 3, panel C). Recipients of kidney–pancreas transplants had better 3-yr kidney survival (82.5%, n = 3243) than those receiving a single (75.7%, n = 33 526) or *en-bloc* kidneys (68.2%, n = 638) ($p < 0.001$).

Mulitvariate analysis

Logistic regression The whole dataset (n = 37 407) was initially included in a LM model predicting 3-yr graft survival. Using stepwise forward selection, we set a significance level of 0.05 for independent variables to enter the model. The variables and model information are presented in Table 5. Odds ratios with 95% confidence intervals (CI) for the binary variables identified by the model are presented graphically (Fig. 4). Deciles 6 and 7 of ESRD causes were identified as having a significantly higher risk of allograft failure, the odds ratios of 3-yr graft survival were 0.75 (95% CI 0.6–0.9) and 0.78 (95% CI 0.7–0.9) respectively. Causes of ESRD in these categories that demonstrated <70% 3-yr survival are: membranous nephropathy (66.2%), cyclosporine nephrotoxicity (68.3%), analgesic nephropathy (68.8%), type II insulin-dependent DM (65.6%), Henoch–Schönlein purpura (69.7%), mesangio-capillary type 1 glomerulonephritis (68.5%), hemolytic uremic syndrome (54.8%).

Model discrimination using the *c* index (area under the receiver operating characteristic curve) was 0.653. This is the probability that for a randomly chosen pair of patients, the predicted and observed graft survival are concordant. Model

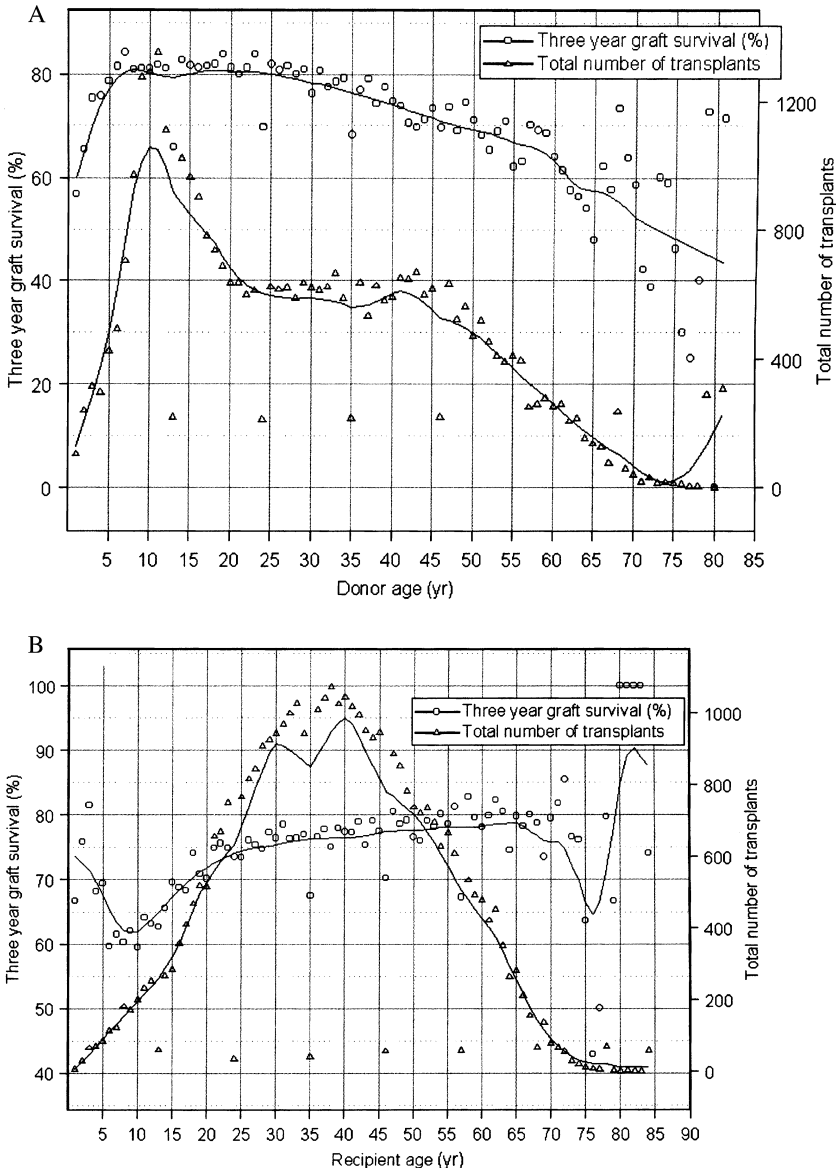

Fig. 2. Three-year graft survival and donor and recipient age. Panel A: 3-yr graft survival (%) and total number of kidney transplants vs. cadaver donor age ($\chi^2 = 980.1$, $p < 0.001$). Panel B: 3-yr graft survival (%) and total number of transplants vs. recipient age ($\chi^2 = 349$, $p < 0.001$).

Table 3. Donor ($\chi^2 = 90.5$, $p < 0.001$) and recipient ($\chi^2 = 4.3$, $p < 0.05$) gender and 3-yr graft survival (%)

	Recipient		Donor	
Gender	Total number	Percent survival	Total number	Percent survival
Female	14 961	75.7	14 202	73.6
Male	22 446	76.7	23 205	77.9

Table 4. Donor ($\chi^2 = 118.7$, $p < 0.001$) and recipient ($\chi^2 = 559.5$, $p < 0.001$) race and 3-yr graft survival (%)

	Donor		Recipient	
Race	Total number	Percent survival	Total number	Percent survival
White	29 796	77.2	23 322	79.2
Black	3968	69.7	8852	67.0
Hispanic	2943	75.4	3493	78.2
Asian	432	73.6	1194	81.0

calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test. As the p value, $p = 0.63$, of this test was not significant, the model's estimated probabilities of 3-yr graft survival are not significantly different from the actual survival of patients over groups spanning the entire range of probabilities.

Prediction analysis

To generate the prediction model we randomly selected 25 000 records as the training set, while the remaining 12 407 records were designated as a testing set and were used to compare predicted and

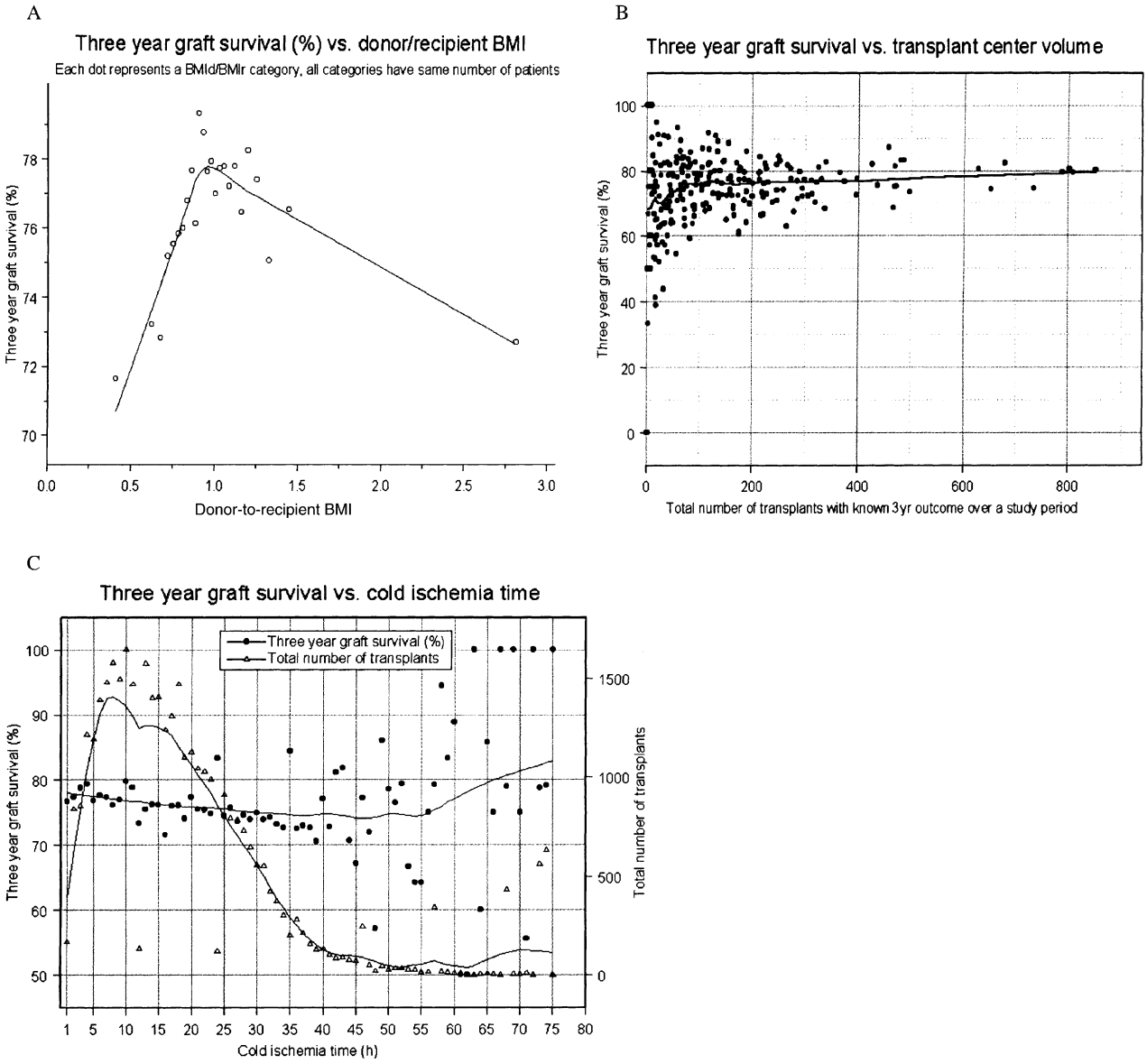

Fig. 3. Bivariate analysis of 3-yr graft survival (%) and relative body mass index (BMI), transplant center volume, and cold ischemia time. Panel A: relationship between donor-to-recipient BMI category and 3-yr graft survival ($\chi^2 = 88.74798$, $p < 0.001$). Each dot represents a donor/recipient BMI category. Each of the 25 categories have same number of patients ($n = 1496$). Scale indicates the mean for each category. Panel B: 3-yr graft survival and transplant center volume (total number of transplants with known outcomes over a study period) ($\chi^2 = 43.3$, $p < 0.001$) Panel C: cold ischemia time and 3-yr graft survival.

observed 3-yr allograft survival. A LM model was again generated on the training set only. This model was 65% concordant, 34.5% discordant, and the *c* index was 0.653. Using the variables and parameter estimates generated with the training set, we calculated the probability of 3-yr graft survival in the testing set. All records were divided into 10 groups based on deciles of predicted probability of graft survival (0–10%, >10–20%, >20–30%, etc.). The observed percentage of 3-yr graft survival was calculated for each group, and the observed graft survival was compared with the expected survival. As there was only one patient in the >10–20% group, that group was combined with the >20–30% group to produce a >10–30% group. The midpoint of each group's probability range was used as the expected percent survival. As shown in Fig. 5 the prediction of the probability of graft survival from the training model achieved a very good match with the observed survival of the testing set, with a chi-square value of 6.15 and $p = 0.63$, which shows no significant difference between observed and predicted category, and a correlation of $r = 0.998$.

Table 5. Predictors of the outcome (3-yr graft survival) identified by logistic regression for the whole dataset (n = 37 407)

Independent variable	Coefficient	χ^2	p	Odds ratio	95% CI
Intercept	1.332	89.474	<0.0001		
Donor age	−0.0145	297.87	<0.0001		
Donor BMI	0.0015	9.0748	0.0026		
Recipient BMI	−0.0121	42.774	<0.0001		
Recipient age	0.0146	231.46	<0.0001		
HLA match	0.1336	206.65	<0.0001		
Cold ischemia time	−0.0079	35.701	<0.0001		
Recipient is male	0.0648	6.4246	0.0113	1.067	1.015–1.122
Donor is male	0.1467	30.611	<0.0001	1.158	1.099–1.22
Terminal donor creatinine 0.1–0.5	−0.2087	10.343	0.0013	0.812	0.715–0.922
Terminal donor creatinine >1.5–2	−0.2389	12.579	0.0004	0.787	0.69–0.899
Terminal donor creatinine >2–2.5	−0.4012	8.8319	0.003	0.67	0.514–0.872
Previous number of transplants = 1	−0.4078	11.241	0.0008	0.665	0.524–0.844
Previous number of transplants = 2	−0.8534	35.723	<0.0001	0.426	0.322–0.564
Previous number of transplants > 2	−1.1078	25.17	<0.0001	0.33	0.214–0.509
Previous number of transplants unknown	−0.0454	0.1503	0.6982	0.956	0.76–1.202
Donor is Black	−0.3229	66.57	<0.0001	0.724	0.67–0.782
Donor is Hispanic	−0.1247	7.1664	0.0074	0.883	0.806–0.967
Recipient is Black	−0.4726	263.48	<0.0001	0.623	0.589–0.66
Recipient is Asian	0.2201	8.065	0.0045	1.246	1.071–1.451
Recipient was never dialyzed	0.2001	9.7585	0.0018	1.222	1.077–1.385
Recipient dialysis modality is unknown	0.1754	33.774	<0.0001	1.192	1.123–1.264
Donor: HTN (but not DM)	−0.3701	32.775	<0.0001	0.691	0.608–0.784
Donor: no DM	−0.571	13.845	0.0002	0.565	0.418–0.763
Donor: duration of DM ≥5 yr	−0.5702	14.815	0.0001	0.565	0.423–0.756
Donor: duration of HTN ≥5 yr	0.1856	4.7968	0.0285	1.204	1.02–1.421
Simultaneous kidney–pancreas transplant	0.3052	30.044	<0.0001	1.357	1.217–1.513
Transplant procedure: *en-bloc* transplant	−0.6445	47.954	<0.0001	0.525	0.437–0.63
Transplant procedure: double kidney	−12.727	0.021	0.8849	<0.001	>999.99
Transplant procedure: whole pancreas/right kidney	−1.413	3.9032	0.0482	0.243	0.06–0.989
Transplant center volume (>83–209)	−0.1436	8.2045	0.0042	0.866	0.785–0.956
Transplant center volume (>355–615)	−0.1115	14.812	0.0001	0.895	0.845–0.947
Number of transplants for this diagnosis >46–77 (6th decile)	−0.2942	7.2995	0.0069	0.745	0.602–0.922
Number of transplants for this diagnosis >77–196 (7th decile)	−0.2435	7.9364	0.0048	0.784	0.662–0.929

BMI, body mass index; DM, diabetes mellitus; HTN, hypertension; HLA, human leukocyte antigen; CI, confidence interval.

We converted predicted allograft failure probability into a binary variable (graft survival = 'yes' or 'no') using a cut-point of 50% probability. The results were compared by means of a 2×2 contingency table. The positive PV of allograft survival with the model was 76.0% and the negative PV was 63%.

Tree-based model. We used TBM to identify predictors of 3-yr graft survival and develop a prediction model. The outcome of cross-validation procedure in the form of deviance plotted against number of terminal nodes (tree size) was analyzed and the optimal size of the tree was determined to be equal to 54 terminal nodes. To identify predictors of the outcome the initial tree was constructed on the whole dataset and pruned to 54 terminal nodes. The following 17 predictors of outcome (in order from the root of the tree to the terminal nodes) were identified by the TBM: recipient race, donor age, recipient weight, cold ischemia time, recipient height, previous number of transplants, recipient age, number of matched HLA antigens, donor race, cause of ESRD, recipient gender, number of mismatched HLA antigens, recipient BMI, recipient weight, presence of diabetes and/or HTN, donor height, donor/recipient BMI. The residual mean deviance of the model is 1.03 and misclassification error rate was 0.23.

The new TBM was built upon a training set and validated on the testing set. Using the model generated with the training set, we calculated the probability of 3-yr graft survival in the testing set. All records were divided into 10 groups based on deciles of predicted probability of graft survival (0–10%, > 10–20%, > 20–30%, etc.). The observed percentage of 3-yr graft survival was calculated for each group. The observed graft survival was compared with the expected survival. As there were only six patients in the 0–10% and > 10–20%

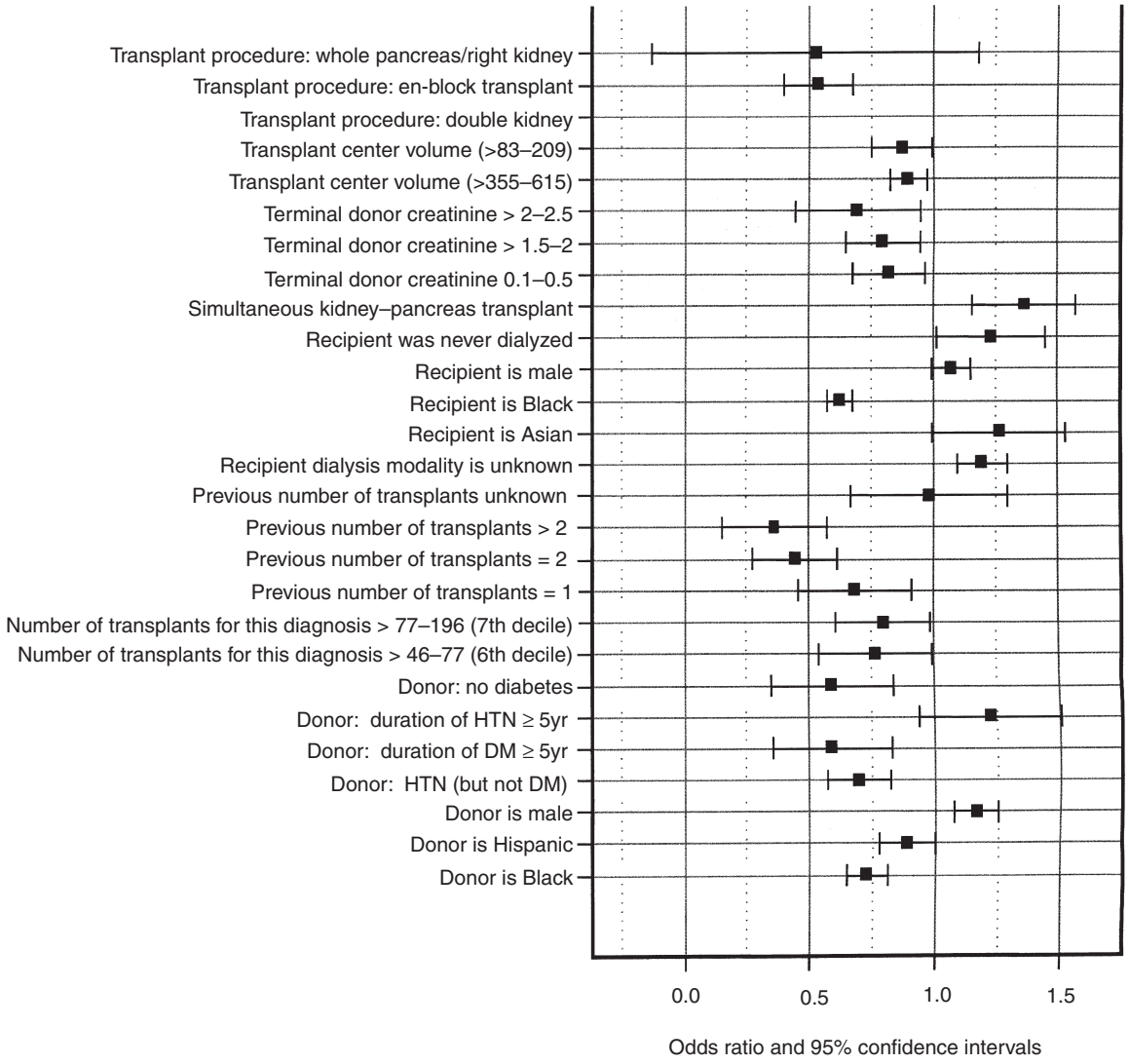

Fig. 4. Odds ratios of the 3-yr graft survival.

groups together, those groups were combined with the >20–30% group to produce a 0–30% group. For the same reason groups >30–40% and >40–50% were combined to produce >30–50% group. The midpoint of each group's probability range was used as the observed percent survival (Fig. 6). The prediction of the probability of graft survival from the training model achieved a good correlation with the observed survival of the testing set ($r = 0.984$). We converted predicted allograft failure probability into a binary variable (graft survival = 'yes' or 'no') using a cut-point of 50% probability (Fig. 7). The graph represents the model in a form of dichotomous tree, where each node presents a question regarding the value of a single independent variable. If the answer to the question is 'yes' users move to the next node by way of the left branch (or right branch, if the answer is 'no') until it reaches the terminal node, which predicts 3-yr graft survival (Y or N). The results were compared by means of a 2×2 contingency table. The positive PV of the allograft survival with the model was 76.0% and the negative PV was 53.8%.

Discussion

Factors affecting kidney allograft survival were evaluated previously based on local datasets and national databases. We looked at a large national dataset which includes relatively new collection of data covering all renal and kidney–pancreas transplants between 1990 and 1998 and, using strict criteria, eliminated records with incomplete information and made careful imputation of some variables. While cleaning the data, we encountered

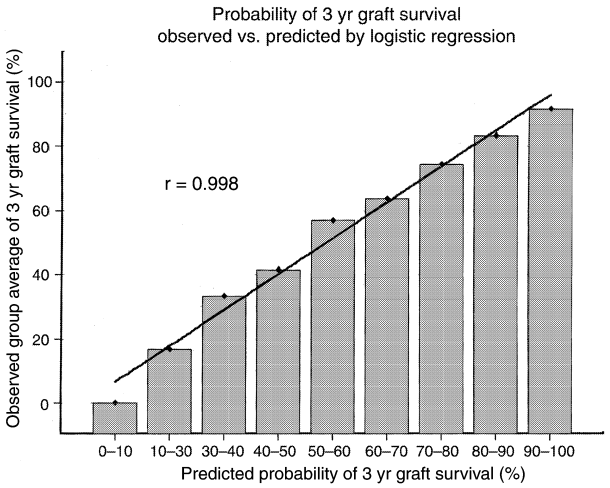

Fig. 5. Results of the prediction of 3-yr graft survival using logistic regression model on the testing dataset. All patients were divided in 10 groups based on predicted probability of graft survival. The observed group averaged graft survival is compared with the predicted probability.

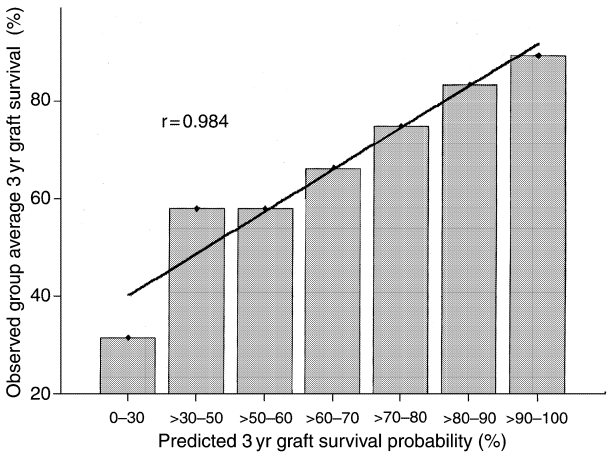

Fig. 6. Results of the prediction of 3-yr graft survival using a tree-based model on the testing dataset. All patients are divided in seven groups based on predicted probability of graft survival. The observed group averaged graft survival is compared with the predicted probability.

certain problems with missing and poorly reported values. The extent of the missing data for some variables is presented in Table 2. The amount of misreported or missing information in 1990–1998 UNOS dataset can be explained by several factors that need to be considered by researchers analyzing the data. As we mentioned in the methods section, certain variables may not have been collected over the entire time period of the cohort. For example, donor history and duration of diabetes and HTN, recipient most recent serum creatinine and donor terminal creatinine were collected only since April of 1994. Additionally, voluntary data submission via paper form (with some fields not mandatory) account for much of the missing information. It has been speculated that this is the reason that height and weight are not populated well (e.g. recipient height is missing approximately 90% of the values). In some cases, the field may not be relevant in that particular instance, so the member may choose to leave it blank [e.g. data for panel reactive antibodies (PRA) tends to be entered only if the patient is sensitized: PRA 80 or above, otherwise it is left blank]. One may expect improved quality of data in the future. The number of outstanding forms has been steadily declining in recent years (there were almost 150 000 outstanding forms in April of 2000 and 56 000 in September 2002) (UNOS, personal communications). UNet, the online transplant data entry system, was implemented in October 1999. The new system of on-line data entry employed real-time data quality control, forcing the user to enter the data in a correct and unified format. Therefore, the quality of information should substantially improve. Thus far, the effect of the new system on the individual variables over time has not been studied closely. In this study, missing categorical variables were coded as new categories and missing continuous variables were replaced using appropriate data imputation methods. In particular, a tree-based algorithm was used for height and weight imputation. The algorithm that we developed has been shown to have good precision when validated on a separate database derived from USRDS DMMS Wave 3 and 4 study patients. Tree-based imputation can be a useful tool for the researchers analyzing the datasets with missing values of the anthropometric characteristics. After careful imputation the results and conclusion of our analysis should not be affected by various causes of missing data (UNOS not collecting it vs. poor reporting). During cleaning of the initial dataset we tried to preserve as much useful information as possible and at the same time eliminate potentially erroneous, incomplete, or unreliable information. A significant number of records had to be eliminated as some critical information was missing or deemed unreliable (Fig. 1).

Our bivariate and multivariate analyses demonstrated the importance of several pre-transplant donor, recipient, and procedure variables in predicting 3-yr graft survival: the number of previous kidney transplants in recipients has a direct relationship with the transplant failure rate, diabetes and HTN worsen the outcome. The relationship between donor and recipient age, race, gender, and 3-yr graft survival previously reported (7, 30), and is non-linear. Number of HLA matched/mismatched

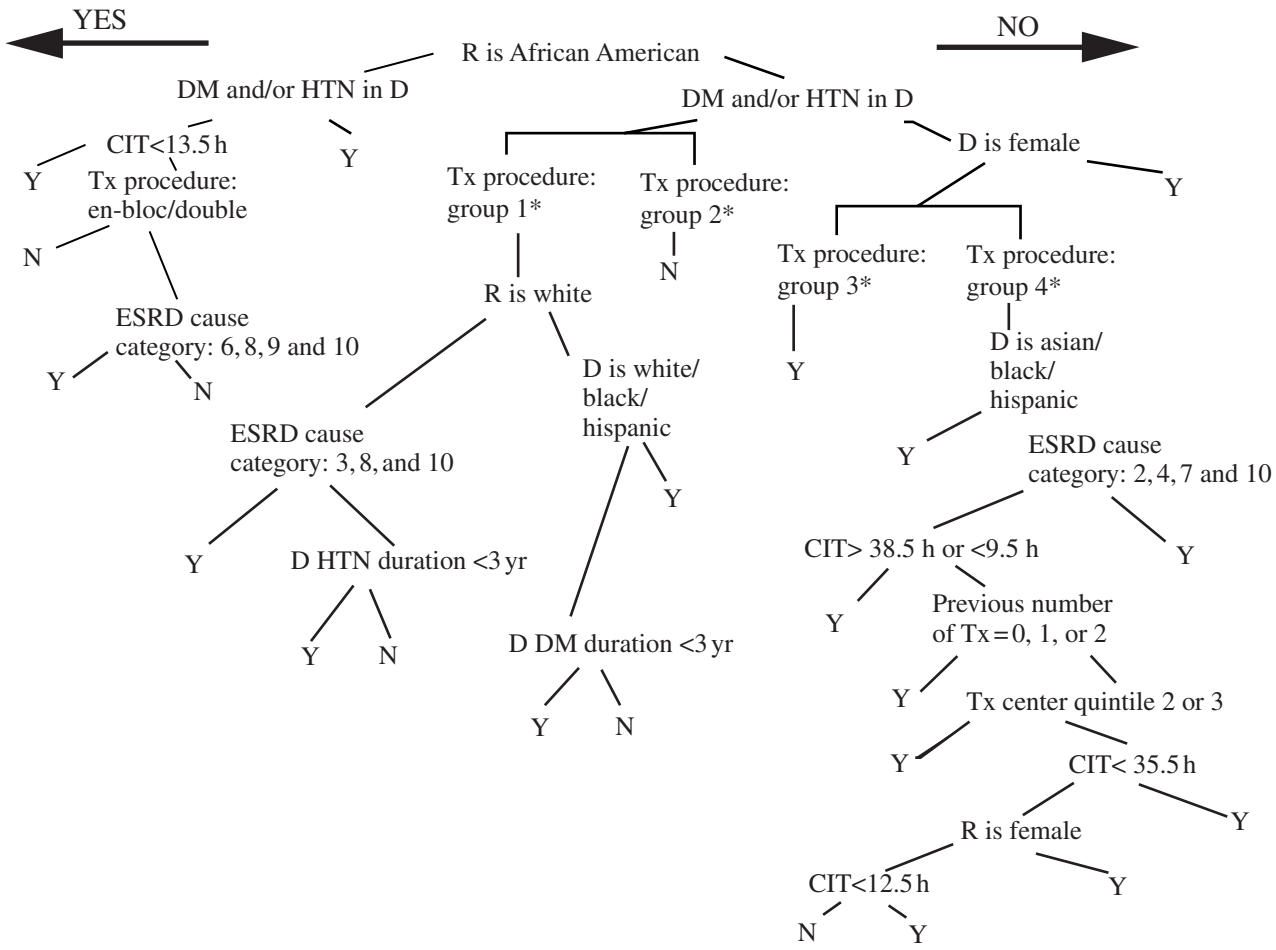

Fig. 7. Tree-based model build on the training dataset. D – donor, R – recipient, Tx – transplant, CIT – cold ischemia time, 'Y' and 'N' at the terminal nodes of the tree correspond to predicted 3-yr graft survival (yes/no). The list of end-stage renal disease cause categories presented in the Appendix. *Transplant procedure group 1: left/right/*en-bloc* kidney with/without the whole pancreas with duodenum or whole pancreas with duodenal patch/left kidney. Transplant procedure group 2: double kidneys, pancreas segment with left kidney, whole pancreas with duodenum and double kidneys, whole pancreas with right kidney. Transplant procedure group 3: pancreas segment and left kidney, whole pancreas with duodenum and right/*en-bloc* kidney, whole pancreas and left kidney Transplant procedure group 4: left/right/*en-bloc* kidney, whole pancreas with duodenum and left kidney, whole pancreas with duodenal patch and left kidney, whole pancreas and right kidney.

antigens has a very strong linear relationship with the percent 3-yr graft survival. The effect of cold ischemia time on the other hand is much less dramatic by bivariate analysis than we initially expected and that was previously reported (11). The transplant center effect was studied before (22) and showed only a very slight difference between the large and small centers. In our study centers with higher number of transplants have more similar outcome, while the outcome of the smaller centers has a lot of variability. This may represent either a regression to the mean or true phenomenon of more uniform outcome that comes with greater experience. LM model selected transplant center volume as a predictor of the outcome, centers with less experience increasing the risk of 3-yr allograft failure. Some of the causes of ESRD in a recipient as a predictor of the outcome were described before (16). In addition, we evaluated all the diagnoses with known 3-yr outcome that we could derive from UNOS database. We confirmed previously reported beneficial effect of preemptive transplantation (21, 31, 32). BMI of donor and recipient as well as recipient obesity in relation to outcome discussed in literature and found to have an important role in the prediction of the kidney allograft outcome in some studies (12, 33), while in others obese (high BMI) transplant recipients have similar outcomes to non-obese patients (34). In our study, both donor and recipient BMI found to be good predictors of the outcome by means of bivariate and multivariate analysis. As successful transplantation of adult living donor kidneys into infants and small children with good long-term

outcome has been shown in a small study (35), the unexpected almost bell-shaped curve (Fig. 3, panel A) describing relationship between donor-to-recipient BMI and graft survival is surprising. That may represent either the deleterious effect of donor obesity (33) or the impact of poor recipient nutritional status. This relationship needs to be further evaluated in prospective study and may be an important factor affecting the selection of the donor.

The novel part of this study is the predictive model. The time period of interest covers the 'post-cyclosporine era', however the 1990s were associated with changes in immunosupression protocols and surgical technique, and therefore the database represents very heterogenous population. This heterogeneity may potentially affect the performance of the prediction models, especially as we included in the analysis only a limited number (26) of pre-transplant independent variables. In designing the study, our intent was to develop a prediction model for use prior to transplantation, therefore we excluded post-transplant variables, that are not available until after the transplant procedure, we did not analyze the impact of immunosuppressive therapy, immediate post-transplant graft function and episodes of acute rejection as this information is not available prior to transplant. Along with conventional LM model, we used a TBM, never before used to analyze transplant outcome. This model represents a relatively new approach compared with conventional regression analysis of the data. The interest in this statistical approach has been increasing over the last 10 yr. Several features make TBM a powerful tool for building a prediction algorithm that can be successfully used in practice. TBM works when the regression variables are a mixture of categorical and continuous variables, it is often able to uncover complex interactions between predictors which may be difficult or impossible to do using traditional multivariate techniques. The algorithm is non-parametric, so no assumptions are made regarding the underlying distribution of values of the predictor variables. TBM 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 non-statistician than the multivariate LM. Thus can be used in the decision making without doing any additional calculations, and therefore is more likely to be followed in clinical practice.

Prediction models using LM and tree-based algorithms are developed in this study on the large set of data, and can potentially be used in the recipient counseling and decision making regarding cadaveric renal transplants. Relatively low area under ROC curves of the initial models suggests that longer list of the potential predictors should be evaluated. However, the prediction algorithms generated on the training dataset can be successfully used in practice to identify the probability of 3-yr kidney allograft survival, as both models achieved good precision in predicting the probability of the graft survival on the separate set of data. There is an experience of using similar data derived from univariate and multivariate analyses in a smaller study in a cadaveric kidney allocation decision making in North Italy Transplant Program (36). 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. The implementation of the models that were generated in this study in a form of software to make it available for transplant program and prospective transplant recipients may be a subject of future projects.

Acknowledgements

The authors thank Jude Maghirang, MS (Department of Research, United Network for Organ Sharing, Richmond, VA) for his assistance in cleaning data and preparation of this manuscript.

This study was funded in part by the US Department of Health and Human Services, Health Resources and Services Administration, Office of Special Programs, Division of Transplantation, under contract numbers 240-97-0001 and 240-97-0002, for the operation of the Organ Procurement and Transplantation Network and the US Scientific Registry of Transplant Recipients.

References

1. Halloran PF, Melk A, Barth C. Rethinking chronic allograft nephropathy: the concept of accelerated cenescence. J Am Soc Nephrol 1999: 10: 167.
2. Cecka JM. The UNOS Scientific Renal Transplant Registry. In: Cecka JM, Terasaki PI, eds. Clinical Transplants. Los Angeles: UCLA Immunogenetics Center, 1999.
3. Gjertson DW. Determinants of long-term survival of adult kidney transplants: a 1999 UNOS update. In: Cecka JM, Terasaki PI, eds. Clinical Transplants. Los Angeles: UCLA Immunogenetics Center, 1999.
4. 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 2001: 5: 5.
5. Ishitani M, Isaacs R, Norwood V, Nock S, Lobo P. Predictors of graft survival in pediatric living-related kidney transplant recipients. Transplantation 2000: 70: 288.
6. Tejani A, Stablein DM, Donaldson L et al. Steady improvement in short-term graft survival of pediatric renal transplants: the NAPRTCS experience. In: Cecka JM,

Terasaki PI, eds. Clinical Transplants. Los Angeles: UCLA Immunogenetics Center, 1999.
7. Terasaki PI, Gjertson DW, Cecka JM, Takernoto S, Cho YW. Significance of the donor age effect on kidney transplants. Clin Transplant 1997: 11: 366.
8. Alexander JW, Bennett LE, Breen TJ. Effect of donor age on outcome of kidney transplantation. Transplantation 1994: 57: 871.
9. Ojo AO, Leichtman AB, Punch JD et al. Impact of preexisting donor hypertension and diabetes mellitus on cadaveric renal transplant outcomes. Am J Kidney Dis 2000: 36: 153.
10. 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 2000: 11: 350.
11. Jordan ML, Shapiro R, Vivas CA et al. High-risk donors: expanding donor criteria. Transplant Proc 1999: 31: 1401.
12. Remuzzi G, Grinyo J, Ruggenenti P et al. Early experience with dual kidney transplantation in adults using expanded donor criteria. J Am Soc Nephrol 1999: 10: 2591.
13. Karpinski J, Lajoie G, Cattran D et al. Outcome of kidney transplantation from high-risk donors is determined by both structure and function. Transplantation 1999: 67: 1162.
14. Hariharan S. Long-term kidney transplant survival. Am J Kidney Dis 2001: 38 (Suppl. 6): S44.
15. Young CJ, Gaston RS. African Americans and renal transplantation: disproportionate need, limited access, and impaired outcomes. Am J Med Sci 2002: 323: 94.
16. Bleyer AJ, Donaldson LA, McIntosh M, Adams PL. Relationship between underlying renal disease and renal transplantation outcome. Am J Kidney Dis 2001: 37: 1152.
17. Mouquet C, Benalia H, Chartier-Kastler E et al. Renal retransplantation in adults. Comparative prognostic study. Progres en Urologie 1999: 9: 239.
18. Chavers BM, Sullivan EK, Tejani A, Harmon WE. Pretransplant blood transfusion and renal allograft outcome: a report of the North American Pediatric Renal Transplant Cooperative Study. Pediatr Transplant 1997: 1: 22.
19. Kim ST, Lee S, Kim JS, Ha J, Chung JK, Kim SJ. Analysis of the factors affecting survival of renal transplant recipients. Transplant Proc 1998: 30: 3086.
20. VanBiesen, Vanholder R, Lameire N. Impact of pretransplantation dialysis modality on patient outcome after renal transplantation: the role of peritoneal dialysis revisited. Perit Dial Int 1999: 19: 103.
21. Mange KC, Joe MM, Feldman H. Effect of the use or nonuse of long-term dialysis on the subsequent survival of renal transplants from living donors. N Engl J Med 2001: 344: 726.
22. Terasaki PI, Cecka JM. The center effect: is bigger better? In: Cecka JM, Terasaki PI, eds. Clinical Transplants. Los Angeles: UCLA Immunogenetics Center, 1999.
23. Lin HM, Kauman HM, McBride MA et al. Center-specific graft and patient survival rates: 1997 United Network for Organ Sharing (UNOS) report. JAMA 1998: 280: 1153.
24. Nyberg SL, Matas AJ, Rogers M et al. Donor scoring system for cadaveric renal transplantation. Am J Transpl 2001: 1: 162.
25. Hennige M, Köhler CO, Opelz G. Multivariate prediction model of kidney transplant success rates. Transplantation 1986: 42: 491.
26. Shoskes DA, Ty R, Barba L, Sender M. Prediction of early graft function in renal transplantation using a computer neural network. Transplant Proc 1998: 30: 1316.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977: 33: 159.
28. Goldfarb-Rumyantzev AS, Pappas L. Prediction of renal insufficiency in Pima Indians with nephropathy of type 2 diabetes mellitus. Am J Kidney Dis 2002: 40: 252.
29. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Monterey, CA: Wadsworth and Brooks/Cole, 1984.
30. Degawa H, Nemoto T, Uchiyama M et al. Effect of donor age on renal allograft survival. Transplant Proc 1998: 30: 3660.
31. Meier-Kriesche HU, Port FK, Ojo AO et al. Effect of waiting time on renal transplant outcome. Kidney Int 2000: 58: 1311.
32. Kasiske BL, Snyder JJ, Matas AJ, Ellison MD, Gill JS, Kausz AT. Preemptive kidney transplantation: the advantage and the advantaged. J Am Soc Nephrol 2002: 13: 1358.
33. Meier-Kriesche HU, Vaghela M, Thambuganipalle R, Friedman G, Jacobs M, Kaplan B. The effect of body mass index on long-term renal allograft survival. Transplantation 1999: 68: 1294.
34. Howard RJ, Thai VB, Patton PR et al. Obesity does not portend a bad outcome for kidney transplant recipients. Transplantation 2002: 73: 53.
35. Healey PJ, McDonald R, Waldhausen JH, Sawin R, Tapper D. Transplantation of adult living donor kidneys into infants and small children. Arch Surg 2000: 135: 1035.
36. Poli F, Scalamogna M, Cardillo M, Porta E, Sirchia G. An algorithm for cadaver kidney allocation based on a multivariate analysis of factors impacting on cadaver kidney graft survival and function. Transpl Int 2000: 13 (Suppl. 1): S259.

Appendix

End-stage renal disease cause categories, based on the total transplant number (n) for the specific diagnosis

Category 1 (n = 1–5): Radiation nephritis, lymphoma;

Category 2 (n = 6–12): Progressive systemic sclerosis, Wilms' tumor, myeloma, antibiotic-induced nephritis, cancer chemotherapy induced nephritis, polyarteritis, urolithiasis;

Category 3 (n = 13–22): Nephrophthisis, gout, incidental carcinoma, cortical necrosis, heroin nephrotoxicity, renal artery thrombosis;

Category 4 (n = 23–27): Mesangio-capillary type 2 glomerulonephritis, cystinosis, Fabry's disease, sickle cell anemia, Goodpasture's syndrome, sarcoidosis;

Category 5 (n = 28–46): Oxalate nephropathy, amyloidosis, renal cell carcinoma, acute tubular necrosis, scleroderma, nephrolithiasis;

Category 6 (n = 47–77): Familial nephropathy, Henoch–Schönlein purpura, prune belly syndrome, type 2 diabetes (insulin-dependent adult onset), membranous nephropathy, analgesic nephropathy, cyclosporin nephrotoxicity;

Category 7 (n = 78–196): Mesangio-capillary type 1 glomerulonephritis, anti-GBM disease, hemolytic uremic syndrome, medullary cystic disease, chronic glomerulosclerosis unspecified, Wegeners granulomatosis;

Category 8 (n = 197–541): Idiopathic/post-infectious crescentic glomerulonephritis, membranous glomerulonephritis, hypoplasia/dysplasia/dysgensis/agenesis, acquired obstructive nephropathy, Alport's syndrome, chronic nephrosclerosis-unspecified, congenital obstructive uropathy;

Category 9 (n = 542–1176): IgA nephropathy, chronic pyelonephritis/reflux nephropathy, systemic lupus erythematosus, malignant HTN, retransplant/graft failure;

Category 10 (n = 1177–7777): Focal glomerularscherosis, polycystic kidneys, type 1 diabetes (insulin-dependent juvenile onset), type 2 diabetes (non-insulin-dependent adult onset), hypertensive nephrosclerosis, chronic glomerulonephritis unspecified, other.