American Journal of Transplantation 2013; 13: 337–347
Wiley Periodicals Inc.

No claim to original US government works
© Copyright 2013 The American Society of Transplantation and the American Society of Transplant Surgeons

doi: 10.1111/ajt.12036

Optimizing the Program-Specific Reporting of Pancreas Transplant Outcomes

B. L. Kasiske[a,b,*], S. Gustafson[a], N. Salkowski[a], P. G. Stock[a,c], D. A. Axelrod[d], R. Kandaswamy[a,e], E. F. Sleeman[f], J. Wainright[f], A. K. Israni[a,b,g] and J. J. Snyder[a,g]

[a] *Scientific Registry of Transplant Recipients, Minneapolis Medical Research Foundation, Minneapolis, MN*
[b] *Department of Medicine, Hennepin County Medical Center, University of Minnesota, Minneapolis, MN*
[c] *Department of Surgery, University of California, San Francisco, CA*
[d] *Section of Solid Organ Transplant Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, NH*
[e] *Department of Surgery, University of Minnesota, Minneapolis, MN*
[f] *United Network for Organ Sharing, Richmond, VA*
[g] *Department of Epidemiology and Community Health, University of Minnesota, University of Minnesota, Minneapolis, MN*
* *Corresponding author: Bertram L. Kasiske, kasis001@umn.edu*

The Scientific Registry of Transplant Recipients is charged with providing program-specific reports for organ transplant programs in the United States. Monitoring graft survival for pancreas transplant programs has been problematic as there are three different pancreas transplant procedures that may have different outcomes, and analyzing them separately reduces events and statistical power. We combined two consecutive 2.5-year cohorts of transplant recipients to develop Cox proportional hazards models predicting outcomes, and tested these models in the second 2.5-year cohort. We used separate models for 1- and 3-year graft and patient survival for each transplant type: simultaneous pancreas–kidney (SPK), pancreas after kidney (PAK) and pancreas transplant alone (PTA). We first built a predictive model for each pancreas transplant type, and then pooled the transplant types within centers to compare total observed events with total predicted events. Models for 1-year pancreas graft and patient survival yielded C statistics of 0.65 (95% confidence interval, 0.63–0.68) and 0.66 (0.61–0.72), respectively, comparable to C statistics for 1-year patient and graft survival for other organ transplants. Model calibration (Hosmer–Lemeshow method) was also acceptable. We conclude that pooling the results of SPK, PAK and PTA can produce potentially useful models for reporting program-specific pancreas transplant outcomes.

Key words: Pancreas after kidney, pancreas transplant alone, Scientific Registry of Transplant Recipients, simultaneous pancreas-kidney transplant

Abbreviations: CI, confidence interval; CMS, Centers for Medicare and Medicaid Services; MPSC, Membership and Professional Standards Committee; OPTN, Organ Procurement and Transplantation Network; PAK, pancreas after kidney transplant; PDRI, pancreas donor risk index; PSR, program-specific report; PTA, pancreas transplant alone; SPK, simultaneous pancreas-kidney transplant; SRTR, Scientific Registry of Transplant Recipients; UNOS, United Network for Organ Sharing.

Received 16 May 2012, revised 14 August 2012 and accepted for publication 05 September 2012

Introduction

Pursuant to the Final Rule (1), the Scientific Registry of Transplant Recipients (SRTR) publishes program-specific reports (PSRs) for solid-organ transplant programs in the United States. The Membership and Professional Standards Committee (MPSC) of the Organ Procurement and Transplantation Network (OPTN) uses PSRs to identify underperforming programs that may need to undertake quality improvement measures to improve their outcomes. For most solid-organ transplant programs, adjusted 1-year patient and graft survival rates are regularly monitored by the MPSC. However, monitoring graft survival for pancreas transplant programs has been problematic, given that there are three different pancreas transplant procedures that may have different outcomes, and analyzing these procedures separately reduces events and statistical power. Additionally, of all the solid-organ transplants performed in the United States, pancreas is second only to intestine in lowest yearly volume. Therefore, the MPSC does not use observed and expected pancreas graft survival as criteria to flag programs for further scrutiny. However, the MPSC and the OPTN's Pancreas Transplantation Committee have been examining whether pancreas graft outcomes could be reliably monitored if the results of all three transplant procedures, simultaneous pancreas–kidney (SPK), pancreas after kidney (PAK) and pancreas transplant alone (PTA), were analyzed separately but then combined.

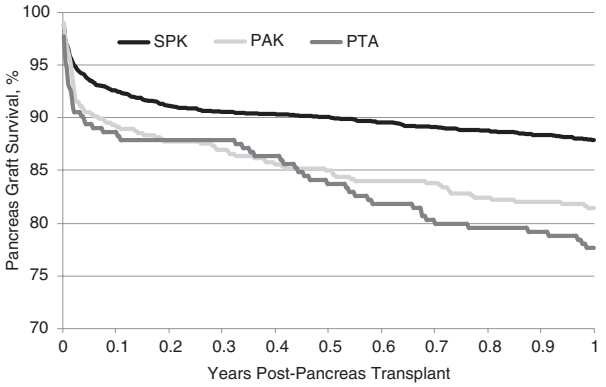

Figure 1: Unadjusted 1-year pancreas graft survival after pancreas transplant (Kaplan–Meier). PAK, pancreas after kidney transplant; PTA, pancreas transplant alone; SPK, simultaneous pancreas–kidney transplant.

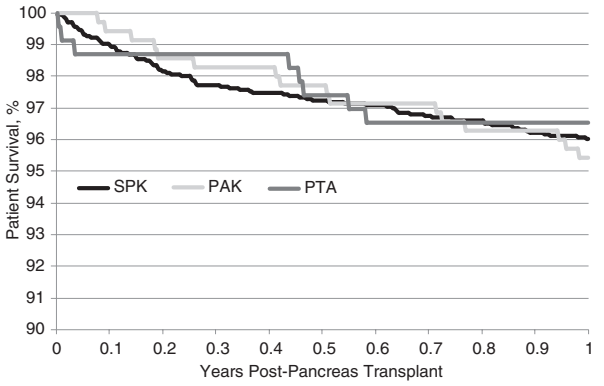

Figure 2: Unadjusted 1-year patient survival after pancreas transplant (Kaplan–Meier). PAK, pancreas after kidney transplant; PTA, pancreas transplant alone; SPK, simultaneous pancreas–kidney transplant.

Methods

Data source

This study used data from SRTR. The SRTR data system includes data on all donors, wait-listed candidates and transplant recipients in the United States, submitted by the members of the Organ Procurement and Transplantation Network (OPTN), and has been described elsewhere (2).The Health Resources and Services Administration (HRSA), US Department of Health and Human Services, provides oversight to the activities of the OPTN and SRTR contractors.

Pancreas transplant study cohorts

The SRTR standard analysis files were used, with OPTN data current through November 15, 2011. Traditionally, 2.5-year (30-month) cohorts are used to develop risk adjustment models to estimate expected outcomes. To increase the pancreas transplant sample size in order to achieve improved stability of model estimates, we used cohorts of transplants occurring over a 5-year period. Although the models include 5 years of consecutive transplants, they were stratified by the first and second 2.5-year periods. The time periods were coordinated with the current PSR cycle so that the newest 2.5-year window coincides with the cohort being evaluated by the PSRs. Program effects were estimated for the latter window only. The primary reason for using a cohort split into two groups was to provide more stable estimates of model effects, not to use one group as a validation cohort and one as a test cohort.

The development cohort for the 1-year outcome models was all SPK, PAK and PTA transplants performed between July 1, 2005 and June 30, 2010. We divided this cohort into two transplant periods of 2.5 years each. Period 1 was July 1, 2005 through December 31, 2007; period 2 was January 1, 2008 through June 30, 2010. All transplants were included in the graft survival models. Patient survival models were limited to primary pancreas transplants only; patients who had undergone any previous pancreas transplant during or before the evaluation period were excluded. Pediatric transplants were included, but made up only 0.05% (n = 3) of the cohort. Multivisceral transplants (aside from kidney–pancreas) were excluded.

The development cohort for the 3-year outcome models was all SPK, PAK and PTA transplants performed between January 1, 2003 and December 31, 2007. We divided this cohort into two transplant periods of 2.5 years each. Period 1 was January 1, 2003 through June 30, 2005; period 2 was July 1, 2005 through December 31, 2007. As in the 1-year models, all transplants were included in the graft failure models; however, patient survival models were limited to primary pancreas transplants only. Pediatric transplants were included (0.08%, n = 5) and multivisceral transplants were excluded.

Modeling approach

Separate models were developed for each transplant type (SPK, PAK, PTA), and for each of four outcomes (1-year and 3-year graft and patient survival). We first built a predictive model for each pancreas transplant type (SPK, PAK, PTA) to predict expected events (graft failures and deaths). An alternative approach would be to develop one Cox model in the combined SPK–PAK–PTA cohort with stratification by transplant type; however, this approach assumes that covariate effects are constant across transplant types and does not allow for maximum flexibility in identifying covariates that may be predictive for one transplant type but do not apply to other transplant types (e.g. time since end-stage renal disease diagnosis is applicable to PAK and SPK but not PTA) or have different effects across transplant types. For these reasons, we chose to build separate predictive models for each transplant type.

Multiple covariates were examined separately and together to determine the predictors for each outcome for each transplant type. Covariates considered for inclusion in the final risk adjustment models are given in the Appendix. Covariates that by themselves were not significantly associated with the outcome ($p \geq 0.10$) were not considered in the final multivariate model. Only covariates that were known at the time of transplant and were independent of the transplanting center's practice choices were considered. Functional form for each predictor was assessed independently across the models; therefore, final parameterizations are not necessarily uniform across the models, and the same group of covariates is not necessarily used in the final models. In addition, we included the composite pancreas donor risk index (PDRI) in the 1-year graft survival models and in all SPK models. The PDRI combines ten pancreas donor variables and one transplant factor (3).

Observation time for all models began at the time of pancreas transplant. For the pancreas graft survival models, patients were followed until the first

Table 1: One-year pancreas graft survival

Covariate	% or mean (IQR)	HR (95% CI)	p
SPK			
PDRI, linear	1.1 (0.85–1.28)	1.6 (1.0–1.8)	< 0.0001
Recipient BMI (kg/m^2) at pancreas transplant (ref, 20–25)			
< 20	5.7	0.79 (0.56–1.1)	0.21
25–30	36.3	0.69 (0.54–0.87)	0.002
> 30	11.9	0.72 (0.56–0.92)	0.009
Unknown	0.2	0.35 (0.05–2.5)	0.30
Recipient race/ethnicity (ref, white)			0.04
African American	17.8	1.2 (0.96–1.5)	0.13
Hispanic	10.4	0.70 (0.51–0.97)	0.03
Other/multirace/unknown	2.5	1.1 (0.66–1.8)	0.71
Kidney received on ice versus pump	92.1	1.4 (1.0–2.1)	0.05
Recipient working for income at time of transplant (ref, no or unknown)	30.7	0.85 (0.71–1.02)	0.08
Recipient hospitalized in the 90 days preceding transplant (ref, no or unknown)	4.6	1.6 (1.2–2.2)	0.002
Cold ischemic time of organ(s), h (ref, > 20)			0.01
0–6	7.5	0.66 (0.42–1.0)	0.07
6–10	30.3	0.62 (0.45–0.84)	0.003
10–15	34.6	0.74 (0.55–0.99)	0.05
15–20	15.4	0.76 (0.55–1.05)	0.10
Unknown	5.8	0.97 (0.67–1.4)	0.88
Age at diabetes diagnosis > 35 years, linear	14.8 (9–20) age > 35 years: 2.2	1.08 (0.88–1.2)	0.05
Age at diabetes diagnosis: unknown	9.7	0.96 (0.74–1.3)	0.77
Recipient CPRA or PRA at transplant, per 10 points, linear	9.1 (0–3)	1.06 (1.03–1.09)	0.0006
PAK			
PDRI, linear	1.02 (0.80–1.15)	1.9 (1.2–2.8)	0.003
Age at transplant, years, linear	43.4 (37–49)	0.97 (0.96–0.99)	0.001
At least one previous pancreas transplant	11.3	1.6 (1.1–2.3)	0.01
Female donor	32.0	0.73 (0.55–0.97)	0.03
HLA mismatches at kidney transplant	3.2 (2–5)	1.07 (1.0–1.2)	0.05
HLA mismatches at kidney transplant, unknown[1]	$n = 5$	0 (0, inestimable)	0.98
Recipient hospitalized in the 90 days before transplant (ref, no or unknown)	10.5	1.6 (1.1–2.3)	0.01
Years of RRT < 10, linear	1.5, (0–1.6)	1.07 (0.99–1.1)	0.11
Years of RRT ≥ 10, linear	2.9	0.67 (0.44–1.03)	0.07
DR mismatches at pancreas transplant (ref, 2)			0.01
0	7.5	0.72 (0.42–1.2)	0.23
1	43.9	1.3 (1.02–1.7)	0.04
PTA			
Recipient BMI at transplant < 25 kg/m^2, linear	25.5 (22. 3–28.0)	0.91 (0.82–1.0)	0.06
Recipient age at transplant < 40 years, linear	41.6 (34–49)	0.95 (0.91–0.98)	0.003
Recipient age at transplant ≥ 40 years, linear		1.02 (0.98–1.06)	0.29
Recipient working for income at time of transplant (ref, no or unknown)	46.2	0.74 (0.51–1.1)	0.12
PDRI, linear	1.11 (0.84–1.21)	1.2 (0.73–1.8)	0.55
Total number of prior pancreas transplants (ref, 1)			< 0.0001
0	87.5	0.50 (0.31–0.81)	0.005
2	0.4	3.8 (1.2–11.9)	0.02
Years with diabetes > 40	26.3 (17–35)	1.06 (0.99–1.1)	0.09
Years with diabetes unknown	15.9	0.83 (0.47–1.4)	0.51
Type 1 diabetes (ref, no or other)	95.5	0.54 (0.28–1.0)	0.06

Results of three multivariate Cox proportional hazards models, with the prevalence (%) or mean (IQR) of covariates, the HR (95% CI) and p value.

[1]This term allows all transplants to be included in the model despite missing data.

BMI = body mass index; CI = confidence interval; CPRA = calculated panel reactive antibody status (0–100%); HR = hazard ratio; IQR = interquartile range; PAK = pancreas after kidney transplant; PDRI = pancreas donor risk index; PRA = panel reactive antibody status (0–100%); PTA = pancreas transplant alone; RRT = renal replacement therapy; SPK = simultaneous pancreas–kidney transplant.

Table 2: One-year pancreas transplant patient survival

Covariate	% or mean (IQR)	HR (95% CI)	p
SPK			
Recipient BMI at transplant < 25 kg/m^2, linear	25.4 (22.5–27.8)	0.90 (0.83–0.98)	0.01
Recipient BMI at transplant ≥ 25 kg/m^2, linear		1.07 (1.02–1.13)	0.004
PDRI, linear	1.12 (0.85–1.28)	1.4 (1.06–1.9)	0.02
Years of RRT, < 6 years	2.0 (0.54–2.6)	1.1 (1.05–1.2)	0.003
Recipient working for income at time of transplant (ref, no or unknown)	30.9	0.68 (0.47–0.98)	0.04
Age at transplant, > 50 years, linear	42.5 (36–49)	1.1 (1.06–1.2)	< 0.0001
PAK			
Age at transplant, years, linear	43.3 (37–49)	1.05 (1.0–1.09)	0.04
Recipient working for income at time of transplant (ref, no or unknown)	39.7	0.27 (0.10–0.71)	0.008
Donor BMI (kg/m^2), linear	23.9 (21.3–26.4)	0.92 (0.83–1.02)	0.11
Recipient experienced delayed graft function at kidney transplant (ref, no or unknown)	4.9	3.5 (1.3–9.6)	0.01
Preemptive kidney transplant	28.3	2.5 (1.2–5.3)	0.02
PTA			
Recipient working for income at time of transplant (ref, no or unknown)	47.6	0.27 (0.07–1.0)	0.05
Recipient age > 50 years	41.7 (34–49)	1.1 (1.02–1.3)	0.02
Female recipient	32.5	0.28 (0.09–0.91)	0.03

Results of three multivariate Cox proportional hazards models, with the prevalence (%) or mean (IQR) of covariates, the HR (95% CI) and p value.
BMI = body mass index; CI = confidence interval; IQR = interquartile range; HR = hazard ratio; PAK = pancreas after kidney transplant; PDRI = pancreas donor risk index; PTA = pancreas transplant alone; RRT = renal replacement therapy; SPK = simultaneous pancreas-kidney transplant.

of: (1) pancreas graft failure recorded on the OPTN Transplant Recipient Registration Form or Transplant Recipient Follow-Up Form as the earliest of graft removal, graft failure, or pancreas retransplant; (2) patient death, as reported on the Transplant Recipient Registration Form or Transplant Recipient Follow-up Form, or as identified using the Social Security Death Master File; (3) 1 or 3 years posttransplant per the appropriate model or (4) loss to follow-up as reported on the Transplant Recipient Registration Form or Transplant Recipient Follow-up Form. For the patient survival models, patients were followed until the first of: (1) patient death, as defined above, or (2) 1 or 3 years posttransplant per the appropriate model.

The expected number of events (graft failures or patient deaths) for the cohorts was calculated as $\sum(-ln(S_i(t_i)))$, where $S_i(t_i)$ is the risk-adjusted survival probability for patient i at time t and t_i is the follow-up time for patient i. Observed and expected event counts were then summed across the transplant groups (SPK, PAK, PTA) within centers to compare total observed events with total expected events.

Assessing model performance

The performance of a model was evaluated by considering two domains of performance, discrimination and calibration. Discrimination is the model's ability to correctly distinguish higher-risk recipients from lower-risk recipients, i.e. to correctly rank patients based on their estimated risk. Discrimination was assessed by the C statistic. A model that correctly ranks all outcomes will have a C statistic of 1.0; a C statistic of 0.50 indicates that the model is not more effective than guessing or tossing a coin.

Another important feature of a prediction model is its ability to accurately predict outcomes at different levels of risk, or its calibration. Calibration was assessed by comparing the number of graft failures or deaths predicted by the model (expected) with the number that actually occurred (observed). Patients were ranked by risk, grouped into tertiles, quartiles, quintiles, sextiles or deciles based on the amount of data, i.e. the sample size available within each cohort. We used the Hosmer–Lemeshow statistic to test whether the overall differences between numbers of observed and expected events indicated a significant lack of fit (4,5).

Results

Study cohorts

Between July 1, 2005 and June 30, 2010, 4316 SPKs were performed; 86.3% of recipients ($n = 3724$) were censored at 1 year posttransplant and 13.7% (n = 592) experienced pancreas graft failure or died before the end of follow-up. Of 1230 PAK transplants performed, 80.3% of recipients (n = 988) were censored at 1 year posttransplant and 19.7% (n = 242) experienced pancreas graft failure or died before the end of follow-up. Of 532 PTAs performed, 77.8% of recipients (n = 414) were censored at 1 year posttransplant and 22.2% (n = 118) experienced pancreas graft failure or died before the end of follow-up.

Between January 1, 2003 and December 31, 2007, 4408 SPKs were performed; 79.1% of recipients (n = 3486) were censored at 3 years posttransplant and 20.9% (n = 922) experienced pancreas graft failure or died before the end of follow-up. Of 1668 PAKs performed, 65.4% of recipients (n = 1091) were censored at 3 years posttransplant and 34.6% (n = 577) experienced pancreas graft failure or died before the end of follow-up. Of 568 PTAs performed, 59.3% of recipients (n = 337) were censored at 3 years posttransplant and 40.7% (n = 231) experienced pancreas graft failure or died before the end of follow-up.

Pancreas graft and patient survival models

The 1-year unadjusted pancreas graft survival was greater after SPK than after PAK or PTA (Figure 1). Patient survival was similar after SPK, PAK and PTA (Figure 2). Results of the final 1-year and 3-year multivariate models

Table 3: Three-year pancreas graft survival

Covariate	% or mean (IQR)	HR (95% CI)	p
SPK			
PDRI, linear	1.12 (0.83–1.28)	1.6 (1.4–1.8)	< 0.0001
Recipient BMI (kg/m^2) at pancreas transplant (ref, 20–25)			
< 20	8.2	1.2 (0.94–1.5)	0.15
25–30	32.8	1.2 (1.0–1.4)	0.03
> 30	12.3	1.5 (1.2–1.8)	< 0.0001
Unknown	0.5	1.07 (0.50–2.2)	0.87
Female recipient	37.8	1.2 (1.02–1.3)	0.03
Age at transplant ≤ 50 years, linear	42.3 (36–48)	0.99 (0.98–1.0)	0.07
Age at transplant > 50 years, linear		1.03 (1.0–1.07)	0.02
Years of RRT at transplant, linear	2.12 (0.46–2.65)	1.02 (0.99–1.04)	0.06
Recipient hospitalized in the 90 days before transplant (ref, no or unknown)	5.4	1.4 (1.1–1.9)	0.01
Recipient CPRA or PRA at transplant, per 10 points, linear, ≤ 80	8.13 (0–4)	1.05 (1.0–1.1)	0.02
Recipient CPRA or PRA at transplant, per 10 points, linear, > 80		0.72 (0.51–1.01)	0.06
PAK			
Age at transplant, ≤ 60 years, linear	43.4 (38–49)	0.98 (0.97–0.99)	0.0007
Age at transplant, > 60 years, linear		1.2 (0.94–1.6)	0.11
Donor age, linear	24.9 (18–30)	1.01 (1.0–1.02)	0.004
At least one previous pancreas transplant		1.5 (1.1–1.9)	0.003
Peripheral vascular disease at listing (ref, no)			0.02
Yes	6.6	1.4 (1.1–1.9)	0.01
Unknown	6.8	1.3 (0.97–1.8)	0.08
Years of RRT at kidney transplant, ≤ 10, linear	4.1 (0.47–6.4)	1.02 (0.99–1.05)	0.16
Years of RRT at kidney transplant, > 10, linear		0.94 (0.87–1.0)	0.09
Donor BMI linear	24.1 (21.5–26.4)	1.02 (0.99–1.04)	0.08
Recipient BMI at transplant ≤ 22 kg/m^2, linear	25.1 (22.2–27.6)	0.92 (0.85–1.0)	0.07
Recipient BMI at transplant > 22 kg/m^2, linear		1.01 (0.99–1.04)	0.41
Recipient working for income at time of transplant (ref, no or unknown)	37.9	0.83 (0.68–1.02)	0.08
Recipient hospitalized in the 90 days before transplant (ref, no or unknown)	9.5	1.3 (0.97–1.8)	0.07
Recipient eGFR at pancreas transplant, per 10 units, linear	62.6 (49.2–75.2)	0.96 (0.93–1.0)	0.08
Type 1 diabetes (ref, not type 1)		1.4 (0.99–2.05)	0.07
Recipient eGFR at discharge postkidney transplant, ≤ 30 mL/min/1.73m^2, per 10 mL/min/1.73m^2	61.5 (45.1–79.0)	0.83 (0.70–0.98)	0.03
Recipient eGFR at discharge post-kidney transplant, > 30 mL/min/1.73m^2, per 10 mL/min/1.73m^2		1.0 (0.97–1.05)	0.70
Deceased kidney donor (ref, living)	38.1	1.1 (0.89–1.3)	0.42
PTA			
Age at transplant, ≤ 50 years, linear	41.7 (35–49)	0.97 (0.95–0.98)	< 0.0001
Age at transplant, > 50 years, linear		1.05 (0.99–1.1)	0.06
Recipient CPRA or PRA at transplant, per 10 points, linear	12.3 (0–7)	1.08 (1.03–1.1)	0.0008
Donor cause of death: cerebrovascular accident	16.8	1.5 (1.06–2.0)	0.02
Recipient working for income at time of transplant (ref, no or unknown)	55.6	0.70 (0.52–0.95)	0.02
Type 1 diabetes (ref, not type 1)	93.7	0.59 (0.36–0.97)	0.04
Female recipient	60.8	1.4 (1.04–1.9)	0.03
Donor BMI, kg/m^2 (ref, 20–25)			0.08
< 20	13.1	0.89 (0.60–1.3)	0.56
25–30	28.7	1.4 (1.04–1.9)	0.03
> 30	6.0	1.1 (0.69–1.8)	0.63

Results of three multivariate Cox proportional hazards models, with the prevalence (%) or mean (IQR) of covariates, the HR (and 95% CI) and p value.

BMI = body mass index; CI = confidence interval; CPRA = calculated panel reactive antibody status (0–100%); e GFR = estimated glomerular filtration rate; HR = hazard ratio; IQR = interquartile range; PAK = pancreas after kidney; PDRI = pancreas donor risk index; PRA = panel reactive antibody status (0–100%); PTA = pancreas transplant alone; RRT = renal replacement therapy; SPK = simultaneous kidney–pancreas.

Table 4: Three-year pancreas transplant patient survival

Covariate	% or mean (IQR)	HR (95% CI)	p
SPK			
PDRI, linear	1.12 (0.83–1.27)	1.6 (1.3–1.9)	< 0.0001
Age at transplant, years, linear	42.3 (36–48)	1.03 (1.01–1.04)	<0.0001
Years of RRT, linear, ≤ 6	2.0 (0.44–2.56)	1.1 (1.05–1.2)	0.0003
Recipient BMI at transplant, ≤ 24 kg/m^2, linear	25.1 (22.2–27.6)	0.92 (0.86–0.98)	0.01
Recipient BMI at transplant, > 24 kg/m^2, linear		1.05 (1.0–1.1)	0.03
Recipient working for income at time of transplant (ref, no or unknown)	31.6	0.70 (0.52–0.94)	0.02
Recipient hospitalized in the 90 days before transplant (ref, no or unknown)	5.4	1.5 (0.96–2.3)	0.08
Peripheral vascular disease (ref, no or unknown)	7.5	1.3 (0.93–1.9)	0.13
PAK			
Recipient BMI at pancreas transplant ≤ 24 kg/m^2, linear	24.3 (22.5–27.5)	0.82 (0.72–0.93)	0.002
Recipient BMI at pancreas transplant 24–30 kg/m^2, linear		1.1 (0.99–1.25)	0.05
Recipient BMI at pancreas transplant > 30 kg/m^2, linear		0.91 (0.72–1.15)	0.42
Deceased kidney donor at most recent kidney transplant (ref, living)	29.3	1.9 (1.2–2.8)	0.003
Peripheral vascular disease at listing (ref, no or unknown)	6.5	1.7 (1.0–3.0)	0.05
Pancreas donor eGFR, per 10 mL/min/1.73m^2, linear	93.3 (71.7–121.1)	0.94 (0.89–0.99)	0.04
Recipient eGFR at pancreas transplant, per 10 mL/min/1.73m^2, linear	62.4 (49.5–74.7)	0.87 (0.81–0.97)	0.01
Recipient working for income at time of transplant (ref, no or unknown)	37.3	0.53 (0.30–0.94)	0.03
Donor BMI, < 20 kg/m^2, linear	24.1 (21.5–26.5)	1.4 (0.94–2.2)	0.10
Donor BMI 20–30 kg/m^2, linear		0.94 (0.88–1.0)	0.09
Donor BMI > 30 kg/m^2, linear		1.26 (1.1–1.4)	0.0003
Recipient experienced delayed graft function at kidney transplant (ref, no or unknown)	6.3	1.7 (0.98–3.1)	0.06
PTA			
Age at transplant, linear	41.8 (35–49)	1.03 (1.0–1.08)	0.04
Type 1 diabetes (ref, not type 1)	92.9	0.31 (0.12–0.79)	0.01
Recipient working for income at time of transplant (ref, no or unknown)	56.9	0.44 (0.19–0.99)	0.05
Recipient hospitalized in the 90 days preceding transplant (ref, no or unknown)	5.4	2.9 (0.93–9.1)	0.07
Female donor	33.9	0.58 (0.26–1.29)	0.18
Age at diabetes diagnosis < 25 years, linear	15.9 (8–22)	0.95 (0.90–0.99)	0.04
Age at diabetes diagnosis > 25 years, linear	14.6 > 25 years	1.03 (0.99–1.09)	0.21
Age at diabetes diagnosis: unknown	11.7	0.36 (0.09–1.5)	0.15

Results of three multivariate Cox proportional hazards models, with the prevalence (%) or mean (IQR) of covariates, the HR (and 95% CI) and p value.

BMI = body mass index; CI = confidence interval; eGFR = estimated glomerular filtration rate; IQR = interquartile range; PAK = pancreas after kidney; PDRI = pancreas donor risk index; PTA = pancreas transplant alone; RRT = renal replacement therapy; SPK = simultaneous kidney-pancreas transplant.

are displayed in Tables 1–4, which include hazard ratios, confidence intervals and pvalues. Older age was associated with worse 1-year graft survival (PAK, PTA) and 1-year (SPK, PAK, PTA) and 3-year (SPK, PTA) patient survival. However, older ages less than 60 years (SPK, PTA) and less than 50 years (PAK) were associated with *improved* 3-year pancreas graft survival, while older ages above these thresholds were associated with *worse* 3-year graft survival. Functional status, reflected by ability to work for income at the time of transplant, was an important risk factor for graft survival and for patient survival. Increasing PDRI was associated with worse outcomes, especially for 1-year pancreas graft survival. Hospitalization in the 90 days before pancreas transplant was associated with worse 1- and 3-year pancreas graft survival for SPK and PAK.

Model performance

Calibration and discrimination were assessed in the 2.5-year PSR evaluation cohort, i.e. the second half of the 5-year development cohort. For each of the models, calibration was generally good, as reflected by low chi-square values (less than 10) and high p values (greater than 0.20) (Table 5). The calibration of models predicting pancreas graft survival was similar to the calibration of models predicting patient survival. Likewise, calibration was similar for models predicting 1- and 3-year outcomes (Table 5, Figures 3 and 4).

Discrimination, as assessed by the C statistic, ranged from 0.59 to 0.78 (Table 5). In general, the C statistics for 1-year graft survival models were similar (0.63–0.68) for SPK, PAK and PTA. C statistics were also similar (0.62–0.78), albeit

Table 5: Model calibration and accuracy

Model	Calibration[1]	C statistic (95% CI)[2]	Combined C statistic (95% CI)[3]
1 year			
SPK graft survival	$\chi^2(8) = 7.88$, $p = 0.44$	0.63 (0.60–0.67)	0.65 (0.63–0.68)
PAK graft survival	$\chi^2(8) = 9.29$, $p = 0.32$	0.63 (0.58–0.69)	
PTA graft survival	$\chi^2(8) = 6.44$, $p = 0.60$	0.68 (0.61–0.75)	
SPK patient survival	$\chi^2(8) = 4.86$, $p = 0.77$	0.62 (0.57–0.69)	0.66 (0.61–0.72)
PAK patient survival	$\chi^2(3) = 0.88$, $p = 0.83$	0.75 (0.62–0.88)	
PTA patient survival	$\chi^2(1) = 0.11$, $p = 0.74$	0.78 (0.58–0.98)	
3-year			
SPK graft survival	$\chi^2(8) = 8.60$, $p = 0.38$	0.59 (0.56–0.62)	0.62 (0.60–0.64)
PAK graft survival	$\chi^2(8) = 4.51$, $p = 0.81$	0.60 (0.56–0.64)	
PTA graft survival	$\chi^2(8) = 5.74$, $p = 0.68$	0.66 (0.61–0.71)	
SPK patient survival	$\chi^2(8) = 3.28$, $p = 0.92$	0.64 (0.60–0.68)	0.66 (0.62–0.69)
PAK patient survival	$\chi^2(4) = 2.14$, $p = 0.71$	0.68 (0.59–0.77)	
PTA patient survival	$\chi^2(2) = 2.84$, $p = 0.24$	0.76 (0.66–0.86)	

CI = confidence interval; PAK = pancreas after kidney; PTA = pancreas transplant alone; SPK = simultaneous pancreas-kidney.
[1]Hosmer-Lemeshow test statistic; a p value ≤ 0.05 indicates significant misfit of the model to the observed data; likewise, a p value > 0.05 indicates that the model fits the observed data reasonably well.
[2]The C statistic reflects the model's ability to correctly rank patients based on their estimated risk. C = 1.0 is perfect ranking; C = 0.50 is roughly equal to a coin toss.
[3]C statistic of the combined models for the three transplant types (SPK, PAK, PTA).

higher, for 1-year SPK, PAK and PTA patient survival models. The C statistics for 3-year graft survival models were lower (0.59–0.66) than those for 1-graft survival models. However, the 3-year patient survival model C statistics were similar (0.64–0.76) to those for the 1-year patient survival models (Table 5).

For 1- and 3-year graft and patient survival models, the pooled C statistics for the SPK, PAK and PTA models were similar, but slightly higher than those for the corresponding isolated SPK models (Table 5).

The C statistic for the pooled (SPK, PAK and PTA) pancreas 1-year graft survival model was comparable to the C statistics for 1-year graft survival models for other organ transplants (Table 6). The C statistic for the pooled pancreas 1-year patient survival model was slightly lower than those for kidney transplant models (Table 6).

Discussion

SRTR currently provides two pancreas transplant PSRs, one for kidney/pancreas (SPK) and one for pancreas that includes PAK and PTA, but not SPK. OPTN/United Network for Organ Sharing (UNOS) does not designate a kidney/pancreas program, but separately designates programs that perform kidney transplants as kidney programs and programs that perform pancreas transplants as pancreas programs; to perform combined kidney and pancreas transplants, a program must be OPTN/UNOS approved to provide both kidney and pancreas transplants. The SRTR kidney/pancreas reports include 1-month, 1-year and 3-year observed and expected kidney graft survival, pancreas graft survival and patient survival statistics. The pancreas reports include 1-month, 1-year and 3-year observed, *but not expected*, pancreas graft and patient survival rates, unlike reports for other organ transplants. The models in this study were developed in response to requests from the Pancreas Transplantation Committee and OPTN's MPSC to align monitoring of posttransplant outcomes for pancreas recipients with monitoring of outcomes for other organ transplants. Currently, the MPSC does not flag programs for pancreas outcomes, although programs may be flagged for kidney outcomes. The MPSC has been studying ways to combine pancreas transplant recipients across transplant types to increase statistical power and produce PSRs and methods to flag underperforming pancreas transplant programs. Use of an expanded and stratified cohort is an innovation with respect to PSR methodology and allows more stable covariate estimates with few disadvantages. This approach could be used to advantage in other small-volume transplant cohorts, such as pediatric transplants.

The goal of the current study was to find the best way to compare pancreas transplant results of different transplant centers. The challenge is to incorporate the fact that patient characteristics and outcomes are very different for SPK, PAK and PTA. It is possible to compare the results of transplant programs for each transplant procedure type separately, but flagging multiple programs within each center would be problematic for several reasons.

We constructed separate models for SPK, PAK, and PTA outcomes. We then used these models to determine expected outcomes by pooling the outcome results for these three separate transplant types. This allows for a comparison between observed and expected outcomes for all pancreas transplant recipients, SPK, PAK or PTA. We performed this maneuver for both 1- and 3-year

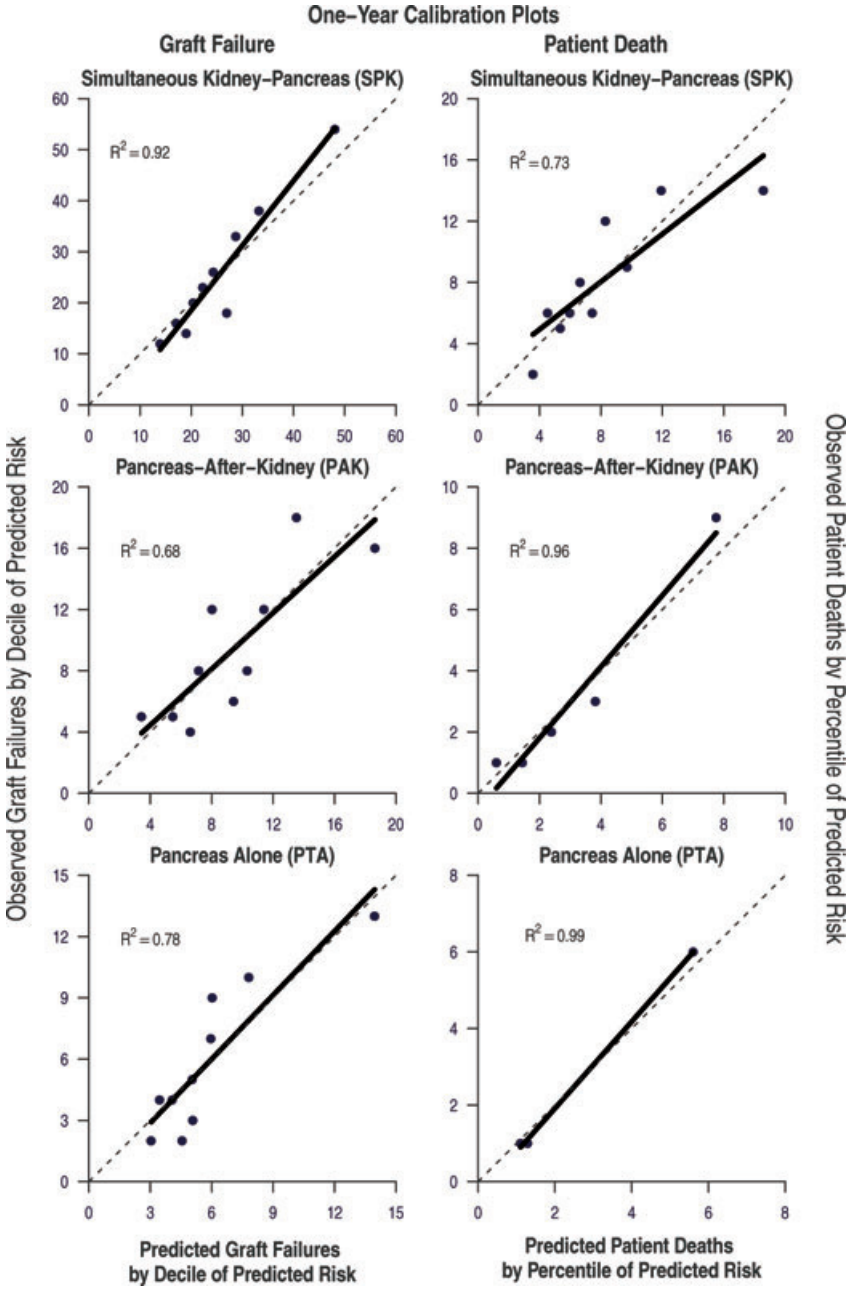

Figure 3: Calibration plots for 1-year models. Bold line is linear regression of observed versus expected events; dashed line is 45°.

outcomes. However, current OPTN policy uses only 1-year graft and patient survival outcomes to identify ('flag') underperforming programs for further scrutiny. Only large-volume programs (defined as those performing 10 or more transplants during a 30-month window) are assessed using risk-adjusted models. One hundred and thirty-three pancreas programs performed at least one transplant during the development cohort (transplants occurring between January 1, 2008 and June 30, 2010); 88 (66.2%) of these were large-volume and 45 (33.8%) were small-volume programs. Of the 88 large volume programs, 6 (6.8%) would have been flagged for 1-year pancreas graft outcomes and 0 would have been flagged for 1-year patient survival outcomes. Of note, the proposed PSR models have no bearing on small-volume centers that are flagged using different methods.

The major finding of this study is that pancreas transplant PSRs could be modeled by pooling outcomes from SPK, PAK and PTA at the program level. Importantly, the accuracy of pooled 1-year pancreas graft survival model results, as measured by the C statistic, was comparable to the accuracy of graft survival results for other organs included in the PSRs (Table 6). Whether OPTN and/or the Centers for Medicare & Medicaid Services (CMS) decide to implement this approach for monitoring pancreas transplant programs

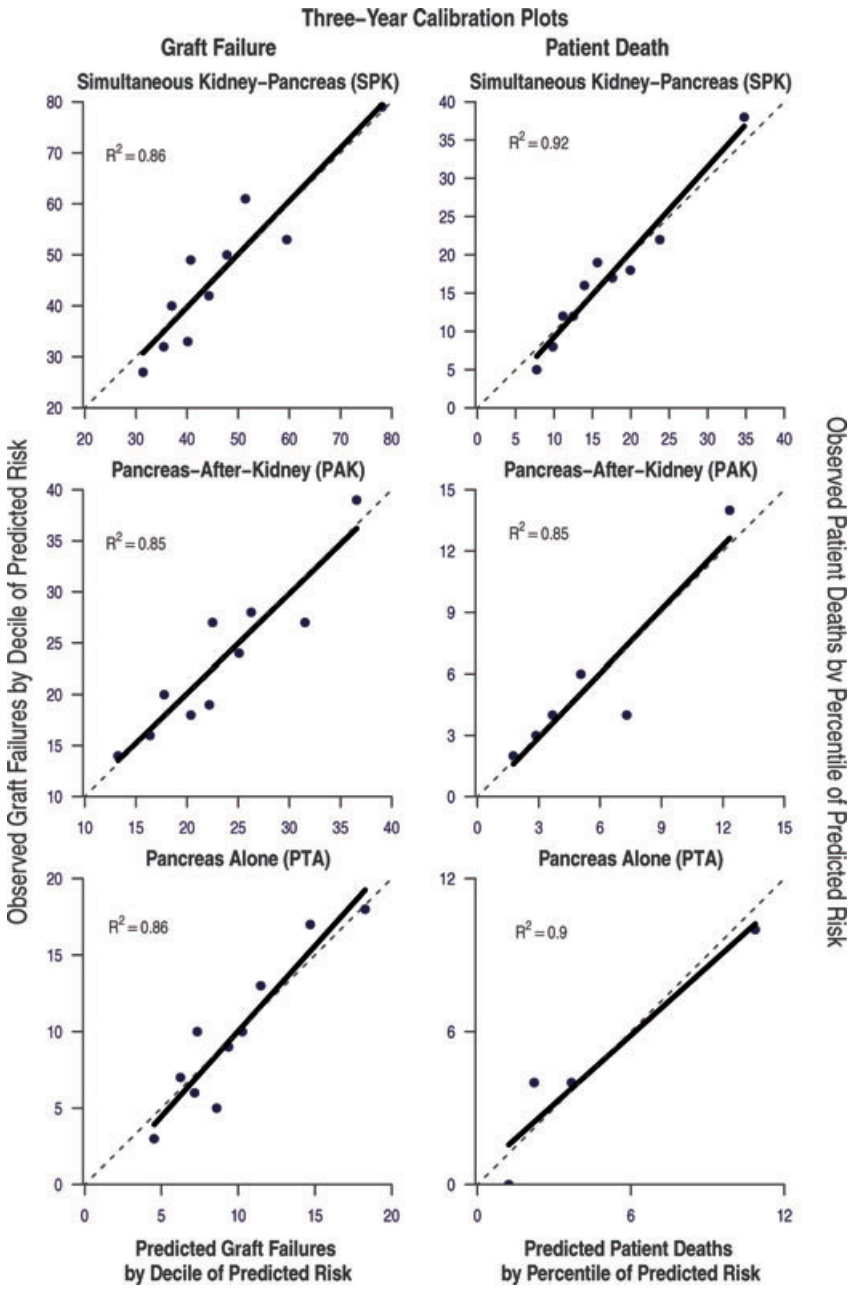

Figure 4: Calibration plots for 3-year models. Bold line is linear regression of observed versus expected events; dashed line is 45°.

is yet to be determined. If this approach is adopted, the models can be updated as needed.

The PDRI was the strongest predictor of 1- and 3-year pancreas graft failure in kidney-pancreas recipients. Furthermore, only one of the 10 individual elements (cold ischemic time) in the PDRI was an independent risk factor for pancreas graft failure when the PDRI was included in the model (Table 1); this suggests that the PDRI identifies the donor characteristics well. The PDRI also predicted 1-year pancreas outcomes after PAK, and to a lesser degree after PTA. However, it was not a strong predictor of mortality after PAK or PTA, nor was it a uniform predictor of 3-year graft failure after PAK or PTA. This may be due to the relatively small numbers of events used to define these two models. It may also be due to the development of the PDRI itself, which was subject to many of the same limitations as these PSR models. Primary among the limitations is that SPKs greatly outnumber PAKs and PTAs; consequently, power to detect predictors of failure is disproportionate between cohorts and can allow the predictors for SPK outcomes to dominate the model. Age and functional status (working for income at the time of transplant) were understandably consistent predictors of mortality after SPK, PAK and PTA. Recent hospitalization was also a consistent predictor in most of the patient survival models (Tables 2 and 4).

Table 6: Comparison C statistics for outcome models

Transplant type	1-year graft survival C statistic	1-year patient survival C statistic
Kidney deceased donor	0.659	0.706
Kidney living donor	0.640	0.742
Liver deceased donor	0.664	0.669
Liver living donor	0.593	0.629
Heart	0.700	0.681
Lung	0.663	0.659
Proposed pancreas	0.652	0.665

A major strength of this study is use of separate cohorts to develop and test the models, which greatly increased the validity of the test results. However, there are also some important limitations to this study. The models can only be as accurate as the data used to generate them; it is likely that better data would yield better models. For example, the OPTN data include no clear definition of what constitutes a pancreas graft failure. Unlike other organs, for which graft failure may be marked by death, retransplant, or (in the case of kidney grafts) initiation of dialysis, the definition of pancreas graft failure is left to the subjective assessment of the transplant center. Some centers define pancreas graft failure as return to insulin dependence, whereas other centers report pancreas graft failure after complete loss of c-peptide. This lack of a uniform definition will be problematic if the reporting of pancreas graft failures has potentially adverse consequences for the center. Similarly, diabetes type can be defined based on the presence of c-peptide levels, but this data element is not currently reported by centers. In the approved (but not yet implemented) new pancreas allocation system, c-peptide levels will be required for an SPK candidate to accrue waiting time, but not for a PTA candidate.

Another approach to pooling outcome data from the different pancreas transplant types would be to pool all patients registered on the waiting list for deceased donor pancreas transplants. The starting time ("time 0") for analyzing outcomes becomes the time of listing for pancreas transplant, not the time the transplant occurs. Pancreas transplant outcomes by transplant type (SPK, PAK or PTA) are then determined by using three separate time-dependent covariates, respectively, in a nonproportional hazards analysis. This approach has been used by others to compare differences in outcomes between SPK, PAK and PTA, and to compare waiting list outcomes (6,7). While this is arguably the best approach for comparing outcomes between the different types of pancreas transplants using observational data, it includes time, events and outcomes that occur while the patient is on the waiting list. Comparisons of transplant centers therefore include more than just their posttransplant outcomes. In addition, when patients are listed is somewhat arbitrary, and varies from center to center. Moreover, this approach to pooling pancreas transplant outcomes would be fundamentally different from how the MPSC currently monitors other organ transplant programs.

Finally, as with all PSRs, there are concerns that their use by the MPSC could make transplant programs unduly risk-adverse, if risk adjustment in the models is not adequate. In the case of pancreas transplant, concerns have been expressed that including small numbers of PAKs with SPKs could create a disincentive for programs to perform living donor kidney transplants followed by PAKs. However, in creating separate models for SPK, PAK and PTA before combining their results, we have attempted to optimally account for differences in risk for these three different procedures, and thereby avoid unfair disincentives for one procedure versus another. However, no models are perfect, and SRTR intends to revisit PSR models with input from the corresponding organ-specific OPTN committees every 3 years.

In summary, we developed a method that combines the results of SPK, PAK and PTA models to predict pancreas graft failure and patient mortality. The accuracy of the models was comparable to the accuracy of other models used in SRTR PSRs. Whether this approach is used by OPTN and/or CMS to scrutinize program results and thereby aid quality improvement efforts has yet not been determined. These methods, like other methods used to flag large transplant programs, applies only to programs that transplanted more than 10 pancreas grafts in the 30-month period analyzed. In addition, in the future a better clinical definition of pancreas graft failure and type of diabetes may improve the ability of the models to predict outcomes.

Acknowledgments

This work was conducted under the auspices of the Minneapolis Medical Research Foundation, contractor for the Scientific Registry of Transplant Recipients, as a deliverable under contract no. HHSH250201000018C (US Department of Health and Human Services, Health Resources and Services Administration, Healthcare Systems Bureau, Division of Transplantation). As a US Government-sponsored work, there are no restrictions on its use. The views expressed herein are those of the authors and not necessarily those of the US Government. The authors thank SRTR colleague Nan Booth, MSW, MPH, ELS, for manuscript editing.

Disclosure

The authors of this manuscript have no conflicts of interest to disclose as described by the *American Journal of Transplantation*.

References

1. Department of Health and Human Services 42 CFR Part 12Organ Procurement and Transplantation Network; Final Rule. Available at: www.gaonet gov/special pubs/organ/appendd pdf. Accessed May 8, 2012.
2. Levine GN, McCullough KP, Rodgers AM, Dickinson DM, Ashby VB, Schaubel DE. Analytical methods and database design: Implications

for transplant researchers, 2005. Am J Transplant 2006; 6(Pt 2): 1228–1242.
3. Axelrod DA, Sung RS, Meyer KH, Wolfe RA, Kaufman DB. Systematic evaluation of pancreas allograft quality, outcomes and geographic variation in utilization. Am J Transplant 2010; 10: 837–845.
4. Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol 1982; 115: 92–106.
5. Parzen M, Lipsitz SR. A global goodness-of-fit statistic for Cox regression models. Biometrics 1999; 55: 580–584.
6. Venstrom JM, McBride MA, Rother KI, Hirshberg B, Orchard TJ, Harlan DM. Survival after pancreas transplantation in patients with diabetes and preserved kidney function. JAMA 2003; 290: 2817–2823.
7. Gruessner AC, Sutherland DE, Dunn DL, et al. Pancreas after kidney transplants in posturemic patients with type I diabetes mellitus. J Am Soc Nephrol 2001; 12: 2490–2499.

Appendix

Covariates considered for inclusion in the final Models

Recipient characteristics

AGE at transplant; race; sex; body mass index; estimated glomerular filtration rate at transplant; creatinine at transplant; diabetes type; years of diabetes at time of transplant; age at diabetes diagnosis; pretransplant insulin requirements; calculated panel-reactive antibody at transplant; previous transplants; previous pancreas transplants; previous kidney donor type (SPK/PAK only); HLA mismatches at kidney transplant (PAK only); delayed graft function at kidney transplant (PAK only); kidney donor eGFR (PAK only); post-kidney transplant eGFR (PAK only); time on waiting list; time between kidney and pancreas transplants (PAK only); medical condition at time of listing; medical condition at time of offer; functional status (Karnofsky scale); working for income at time of transplant; hospitalized in 90 days before transplant; primary payer at transplant; dialysis type (SPK, PAK only); angina; drug-treated hypertension; symptomatic peripheral vascular disease; symptomatic cerebrovascular disease; serum albumin at listing; pretransplant dialysis (SPK, PAK only); years of renal replacement therapy at index transplant (SPK and PAK only); years of renal replacement therapy at first kidney transplant (SPK and PAK only).

Donor characteristics

Pancreas donor risk index; age; body mass index; race; sex; donation after circulatory death; expanded criteria donor; estimated glomerular filtration rate; creatinine; cause of death; history of hypertension; history of diabetes.

Transplant characteristics

Local versus shared organ; cold ischemic time; HLA mismatches; kidney received on pump versus ice (SPK only).