Supplementary Information

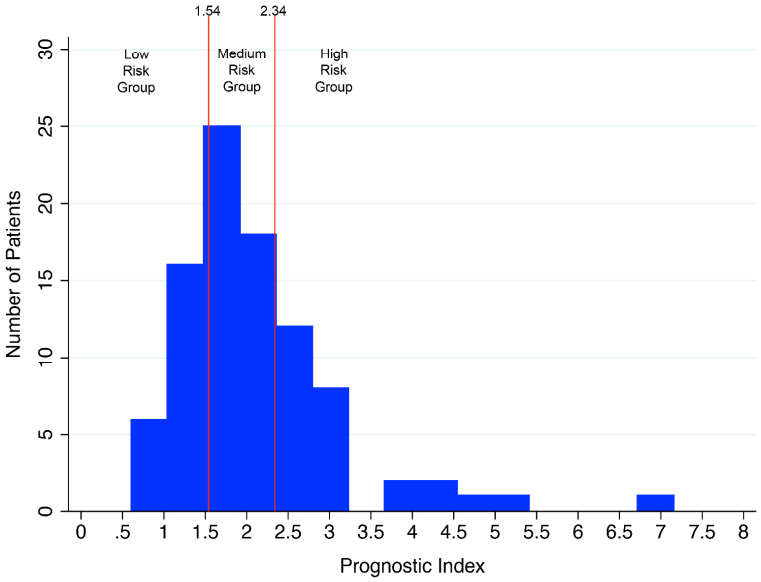
Development and Validation of a Prognostic Index for Allograft Outcome in Kidney Recipients with Transplant Glomerulopathy

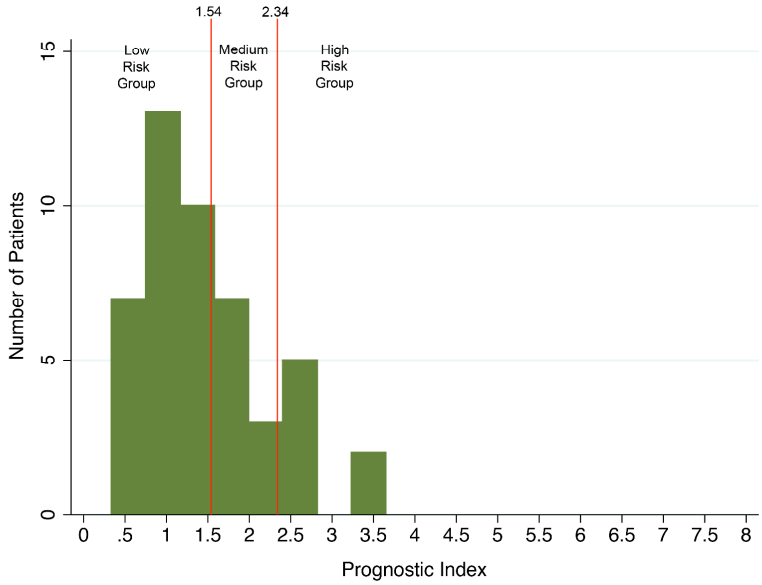
Supplementary Table 1. Characteristics of kidney allograft recipients of the Validation cohort.

Variables	N=47 patients
At the time of transplant	
Age (years), mean (SD)	42 (15)
Women, *N (%)*	19 (40)
Racial categories (Black), *N (%)*	6 (13)
Cause of end stage kidney disease, *N (%)*	
Diabetes	3 (6)
Hypertension	1 (2)
Polycystic kidney disease	3 (6)
IgA nephropathy	10 (22)
Lupus nephritis	1 (2)
Focal and segmental glomerulosclerosis	3 (6)
Other glomerular diseases	6 (13)
Others or Unknown	20 (43)
Deceased donor organ, *N (%)*	41 (87)
Cold ischemia time, hours, deceased donor, mean (SD)	24 (7)
Human leukocyte antigen mismatches, mean (SD)	4 (2)
Donor information available, *N (%)*	42 (89)
Age (years), mean (SD)	50 (16)
Women, *N (%)*	19 (45)
Previous transplants, *N (%)*	13 (28)
CDC cross match results, Data available, *N (%)*	47 (100)
T-cell positive	0 (0)
B-cell positive	2 (4)
Flow Cytometry cross match, Data available, *N (%)*	6 (13)
T-cell positive	0 (0)
B-cell positive	0 (0)
Luminex platform DSA, Data available, *N (%)*	7 (15)
DSA negative (MFI of the highest rank donor-specific bead <1000)	6 (86)
DSA positive (MFI of the highest rank donor-specific bead >1000)	1 (14)
Received desensitization therapy, *N (%)*	1 (2)
Induction Immunosuppression, *N (%)*	29 (62)
Antithymocyte globulin	14 (48)
Interleukin receptor-2 antibodies	15 (52)
After transplant and before the index allograft biopsy	
Delayed graft function, *N (%)*	13 (28)
Calcineurin inhibitor based maintenance immunosuppression, *N (%)*	46 (98)

Early corticosteroid withdrawal	0 (0)
Thrombotic microangiopathy, *N(%)*	6 (13)
Hepatitis C virus, *N (%)*	0 (0)
Acute rejection, *N (%)*	14 (30)
Acute rejection episodes, *N*	20
Acute T-cell mediated rejection episodes, *N (%)*	15 (75)
Acute antibody-mediated rejection episodes, *N (%)*	5 (25)
At the time of index allograft biopsy	
Age, mean (SD)	49 (14)
Time from transplantation to biopsy (months), median (IQR)	60 (28-115)
Luminex platform DSA, Data available, *N (%)*	29 (62)
DSA negative (MFI of the highest rank donor-specific bead <1000)	11 (38)
DSA positive (MFI of the highest rank donor-specific bead >1000)	18 (62)
Serum creatinine (mg/dl), median (IQR)	2.04 (1.67-3.17)
Proteinuria >1 g/day, *N (%)*	27 (57)

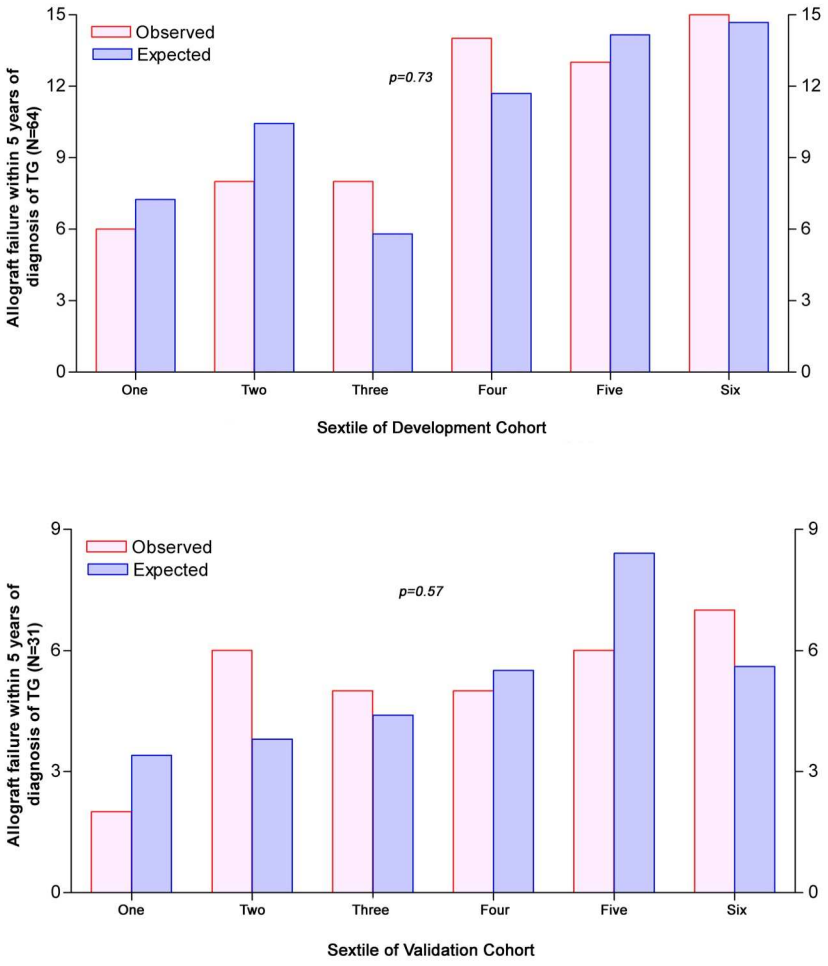
Supplementary Figure 1. Distribution of the prognostic index in the Development and Validation cohorts





We derived the ~~The~~ prognostic index (PI) ~~was derived~~ from the final Cox model ~~and was~~ represented by the equation: (0.29*serum creatinine)+(0.48*proteinuria)+(0.12*chronic-inflammation score). Shown in blue (top) is the histogram of the PI in the Development cohort composed of 92 patients. We did internal validation of the model fit by 10-fold cross validation method. Based on the cross validated estimates of the PI, we divided the entire cohort arbitrarily into three risk groups of allograft failure; low risk (<30th percentile of the PI, cut off: 1.54), medium risk (30th-70th percentile) and high risk (>70th percentile of the PI, cut off: 2.34). Shown in green (bottom) is the histogram of the PI in the independent external Validation cohort that included 47 clinically indicated kidney allograft biopsies from 47 kidney transplant recipients with a diagnosis of TG from the Henri Mondor Hospital, Créteil, France. We applied the same statistical model (equation) generated in the D~~d~~evelopment cohort to the patients in the V~~v~~alidation cohort and derived the PI. We used the same PI cut off values to define the three risk groups.

Supplementary Figure 2. Calibration of the prognostic index model in the Development and Validation cohorts



Calibration (prediction accuracy) reflects how survival from a model compares to that in the observed data. We used the Likelihood ratio statistic of the added variable version of Grønnesby-Borgan test implemented in Stata, to assess the calibration of the model. This is a goodness of fit test for Cox proportional hazards model and is similar to the Hosmer-Lemeshow test for logistic regression. This test uses martingale residuals to compare the count of events to the semi-parametric estimates from the Cox proportional hazards model on a cumulative hazards scale. Patients in the Development cohort (top panel) and Validation cohort (bottom panel) were divided into sextile based on their risk for allograft failure. For each quantile, the observed (pink) and expected (blue) number of patients who developed allograft failure within five years of diagnosis of TG is shown. The p-values of 0.73 and 0.57 imply no statistically significant difference between expected and observed number of patients in each quantile, thus suggesting good calibration of the model in both the cohorts.