OPEN

Predictive Models for Recurrent Membranous Nephropathy After Kidney Transplantation

Edmund Y. M. Chung, MD,[1] Katrina Blazek, BMedSci,[2] Armando Teixeira-Pinto, PhD,[3] Ankit Sharma, PhD,[1,4] Siah Kim, PhD,[1,4] Yingxin Lin, PhD,[5] Karen Keung, PhD,[6] Bhadran Bose, MBBS,[7] Lukas Kairaitis, PhD,[8] Hugh McCarthy, PhD,[1,9] Pierre Ronco, PhD,[10–12] Stephen I. Alexander, MD,[1] and Germaine Wong, PhD[1,3,4]

Background. Recurrent membranous nephropathy (MN) posttransplantation affects 35% to 50% of kidney transplant recipients (KTRs) and accounts for 50% allograft loss 5 y after diagnosis. Predictive factors for recurrent MN may include HLA-D risk alleles, but other factors have not been explored with certainty. **Methods.** The Australian and New Zealand Dialysis and Transplant registry was used to develop 3 prediction models for recurrent MN (Group Least Absolute Shrinkage and Selection Operator [LASSO], penalized Cox regression, and random forest), which were tuned using tenfold cross-validation in a derivation cohort with complete HLA data. KTRs with MN but incomplete HLA data formed the validation cohort. Model performance was evaluated using area under the receiver operating characteristic curve (AUC-ROC). **Results.** One hundred ninety-nine KTRs with MN were included, and 25 (13%) had recurrent MN (median follow-up 5.9 y). The AUC-ROCs for Group LASSO, penalized Cox regression, and random forest models were 0.85 (95% confidence interval, 0.76-0.94), 0.91 (0.85-0.96), and 0.62 (0.57-0.69), respectively, in the derivation cohort, with moderate agreement in selected variables between the models (55%-70%). In their validation cohorts, the AUC-ROCs for Group LASSO and penalized Cox regression were 0.60 (0.49-0.70) and 0.73 (0.59-0.86), respectively. Variables of importance chosen by all models included recipient HLA-A2, donor HLA-DR12, donor-recipient HLA-B65, and HLA-DR12 match. **Conclusions.** A penalized Cox regression performed reasonably for predicting recurrent MN and was superior to Group LASSO and random forest models. These models highlighted the importance of donor-recipient HLA characteristics to recurrent MN, although validation in larger datasets is required.

(*Transplantation Direct* 2022;8: e1357; doi: 10.1097/TXD.0000000000001357).

INTRODUCTION

Primary membranous nephropathy (MN) is one of the most common causes of primary nephrotic syndrome in adults and is associated with progression to kidney failure in approximately one third of patients.[1] Although kidney transplantation is the preferred treatment modality in kidney failure,

Received 30 May 2022.

Accepted 20 June 2022.

[1] Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, NSW, Australia.

[2] School of Population Health, University of New South Wales, Kensington, NSW, Australia.

[3] School of Public Health, The University of Sydney, Camperdown, NSW, Australia.

[4] Department of Renal Medicine, Westmead Hospital, Westmead, NSW, Australia.

[5] School of Mathematics and Statistics, The University of Sydney, Camperdown, NSW, Australia.

[6] Department of Renal Medicine, Prince of Wales Hospital, Randwick, NSW, Australia.

[7] Department of Renal Medicine, Nepean Hospital, Kingswood, NSW, Australia.

[8] Department of Renal Medicine, Blacktown Hospital, Blacktown, NSW, Australia.

[9] Department of Renal Medicine, Sydney Children's Hospital, Randwick, NSW, Australia.

[10] Sorbonne Université, Université Pierre et Marie Curie, Paris, France.

[11] Institut National de la Santé et de la Recherche Médicale, Unité Mixte de Recherche, Paris, France.

[12] Department of Nephrology, Centre Hospitalier du Mans, Le Mans, France.

S.I.A. and G.W. are co-senior authors.

E.Y.M.C. is supported by the NHMRC Postgraduate Research Scholarship (APP 2014037). G.W. is supported by the NHMRC Fellowships (APP 1195414 and 1147657).

The authors declare no conflicts of interest.

E.Y.M.C., K.B., A.S., G.W., A.T.-P., and S.I.A. participated in research design. E.Y.M.C., K.B., S.K., G.W., A.T.-P., Y.L., S.I.A., A.S., K.K., B.B., L.K., H.M., and P.R. participated in the writing of the article. E.Y.M.C., K.B., S.K., G.W., A.T.-P, Y.L., S.I.A., and P.R. participated in data analysis. E.Y.M.C., K.B., S.K., G.W., A.T.-P., Y.L., S.I.A., K.K., B.B., L.K., H.M., and P.R. participated in the performance of the research. E.Y.M.C., K.B., A.T.-P., Y.L., A.S., and G.W. contributed analytic tools. G.W., A.T.-P., and S.I.A. were involved in supervision. Each author contributed important intellectual content during article drafting or revision and agrees to be personally accountable for the individual's own contributions and to ensure that questions pertaining to the accuracy or integrity of any portion of the work, even one in which the author was not directly involved, are appropriately investigated and resolved, including with documentation in the literature if appropriate.

Supplemental digital content (SDC) is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.transplantationdirect.com).

Correspondence: Edmund Y. M. Chung, MD, Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, NSW 2145, Australia. (yinmanchung@gmail.com).

Copyright © 2022 The Author(s). Transplantation Direct. Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 2373-8731

DOI: 10.1097/TXD.0000000000001357

recurrent MN occurs in 35% to 50% of kidney allografts and is associated with subsequent graft loss in half of these patients.[2-4] Autoantibodies against the podocyte antigen phospholipase-A2 receptor (PLA2R) underpins 70% of primary MN, and in 2 small case series, detectable anti-PLA2R antibodies at time of kidney transplantation were associated with recurrent MN with a positive predictive value of 57% to 83% and negative predictive value of 42% to 60%.[5,6]

Genome-wide association studies have also revealed the importance of the immune system in the pathogenesis of MN, identifying risk alleles associated with primary MN in the general population: HLA-DQA1*0501 (DQ2 by serology) in Europeans, DRB1*1501 (DR15 by serology) in East Asians, and HLA-DRB1*0301 (DR17 by serology) in both ethnicities.[7,8] In kidney transplant recipients with MN, a multicenter case series of 93 recipients (55 with recurrent MN) found an association with recipient HLA-A3 but not HLA-DR or HLA-DQ serotypes.[9] Targeted PLA2R1 and HLA-D loci sequencing in 145 kidney transplant recipients with primary MN (54 with recurrent MN) found an association with 2 noncoding HLA-D single nucleotide polymorphisms (SNPs) (rs9271550 and rs9271705) when present on the donor but not the recipient.[10] Prediction models for recurrent MN were improved using a genetic risk score built using these risk SNPs at the HLA-D locus and PLA2R1 locus in addition to clinical variables (area under the curve [AUC] 0.81 compared with AUC 0.71 with clinical variables alone).[10] These studies, however, are largely focused on Caucasian populations, and next-generation sequencing of HLA-D and PLA2R1 loci are not routinely available to assist in the prediction of recurrent MN after kidney transplantation. In this study, we aimed to develop prediction models for recurrent MN in a national cohort of kidney transplant recipients using routinely collected clinical data including donor-recipient HLA serotypes and HLA mismatch characteristics.

MATERIALS AND METHODS

Study Population

This study was reported in adherence to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis statement.[11] Ethics approval was obtained from the Sydney Children's Hospitals Network Human Research Ethics Committee (HREC/ETH00021). We performed a cohort study in kidney transplant recipients with MN as the primary kidney disease from the Australian and New Zealand Dialysis and Transplant (ANZDATA) registry between 1963 and 2020. Indications for kidney allograft biopsy were determined by each individual transplant unit and were not standardized across the cohort. Recurrent MN was defined as biopsy-proven MN posttransplantation and nonrecurrent MN was defined as the absence of biopsy-proven MN posttransplantation because data on proteinuria and anti-PLA2R antibody titers were not available in ANZDATA. Because of the association between HLA and MN, donor and recipient pairs without complete HLA serotyping for HLA-A/B/DR/DQ were excluded from the derivation cohort used for building the prediction models. Although imputation based on HLA serological typing is potentially feasible using the Allele Frequencies Net Database,[12] Haplostats,[13,14] and the IMGT/HLA Database,[15] these are unlikely to be as accurate as using computational analysis with modern next-generation sequencing techniques.[16-19] For recipients who had recurrent MN in >1 kidney transplant, only the first episode of recurrence was included.

Data Collection

Biopsy-proven recurrent MN was the outcome of interest. Covariates of interest included recipient characteristics (age, ethnicity, weight, dialysis vintage, blood group, and comorbidities), donor characteristics (age and blood group), and transplant characteristics (HLA-A/B/DR/DQ serotypes from donor and recipient pairs, HLA-matched serotypes from donor and recipient pairs, total HLA mismatch, donor type, regraft status, maximum pretransplant panel reactive antibody, biopsy-proven rejection, induction immunosuppression regime, maintenance immunosuppression regime upon transplantation, and transplant era based on decade of transplantation). We also assessed HLA-matched serotypes from donor and recipient pairs in the same model because of an association between zero HLA mismatch and recurrent glomerulonephritis after kidney transplantation.[20,21]

Statistical Analysis

In the derivation cohort, Group Least Absolute Shrinkage and Selection Operator (LASSO) regression, a form of penalized logistic regression, was performed for variable selection using the R package "grpreg." In a simple regression, the model may put considerable weighting on certain variables that are deemed to be important, resulting in overfitting. In contrast, LASSO regression performs L1 regularization, where a penalization parameter (λ) is multiplied to the sum of the absolute values of the regression coefficients, imposing a penalty on the size of regression coefficients and shrinking some coefficients to zero during model building. As a result, regularization incorporates both model fitting and variable selection simultaneously, which is useful in high-dimensional datasets.[22,23] Multinomial variables (ethnicity, recipient and donor blood group, and smoking status) and ordinal variables (total HLA mismatch and HLA-A/B/DR/DQ loci mismatch) were grouped together. Group LASSO was chosen over ridge regression because LASSO performs variable selection and yields a sparse model by shrinking coefficient estimates of some variables to zero. In comparison, ridge regression performs L2 regularization, which also shrinks some coefficient estimates but not to zero and therefore does not eliminate variables. Group LASSO was chosen over stepwise logistic regression, which would be prone to false-positive variable selection because of multiple testing in a high-dimensional dataset. Furthermore, LASSO handles sparse events well and therefore is suitable for building prediction models for rare events such as recurrent MN.[24,25] The λ that minimized the binomial deviance was chosen by tenfold cross-validation, which involved dividing our dataset into 10 subsets (9 for training and 1 for validation). The error rates were then averaged across 10 trials to obtain the total efficiency of the modeling. Multiple imputation for missing data was performed using the R package "mice" before variable selection. We generated 100 imputed datasets in which cross-validated Group LASSO was performed. Model performance was assessed using the mean area under the receiver operating characteristic curve (AUC-ROC) and model accuracy (defined as the number of correct classifications divided by the total number of classifications) in the datasets created by tenfold cross-validation with their corresponding 95% confidence interval (CI), which

was calculated by the R package "pROC." Variable importance was determined by calculating its frequency of being selected in the tenfold cross-validation.

Sensitivity Analysis

In the subgroup of kidney transplant recipients in the derivation cohort with recorded time-to-recurrent MN, we performed group penalized Cox regression using the "cv.grpsurv" function in the R package "grpreg." Time-to-graft loss was chosen as the time variable for kidney transplant recipients without recurrent MN. The λ that minimized the partial likelihood deviance was chosen by tenfold cross-validation. Random forest was also performed in the derivation cohort to determine the sensitivity of algorithm variable selection using penalized regression. Random forest was chosen over single decision tree regression because it builds multiple decision trees (reducing high variance associated with single trees) on bootstrapped training samples using a random selection of variables at each split to decorrelate the trees (reducing correlation associated with multiple decision trees utilizing all variables) and therefore reduces overfitting. We performed tenfold cross-validation repeated 5 times to select the number of random variables used at each split and number of trees used in the forest, which maximized the AUC-ROC. Variable importance plots based on the mean decrease of accuracy over all out-of-bag cross-validated predictions were evaluated in the 100 imputed datasets using the tuned random forest model. Variables are ranked in descending order of importance in the variable importance plot, and variables above the inflection point when the curve of the variable importance plot flattens were selected as previously described.[26] Variables selected in at least 50 of the 100 variable importance plots were compared with variables selected using Group LASSO. Random forest, parameter tuning, and variable importance plots were performed using the R package "caret." Model performance of penalized Cox regression and random forest was assessed using the cross-validated AUC-ROC and model accuracy calculated by the R package "pROC."

Validation Cohort

To further evaluate the model performance of the Group LASSO, penalized Cox regression, and random forest models, we created separate validation cohorts from kidney transplant recipients initially excluded because of incomplete HLA serotyping, though retaining predictors driving each model. Model performance was assessed using AUC-ROC and model accuracy with their corresponding 95% CI calculated by the R package "pROC." Multiple imputation on non-HLA variables was performed in the validation cohort separate from the derivation cohort.

RESULTS

Population Characteristics

Of the 32 858 kidney transplant recipients in the ANZDATA registry, 554 recipients had MN as their primary kidney disease of whom 199 donor-recipient pairs (36%) had complete HLA serotyping for donor and recipient HLA-A/B/DR/DQ and were included in the study (Figure 1). Nine variables (4%) had missing values (range, 2%–25%) (**Figure S1, SDC**, http://links.lww.com/TXD/A438). The median age of kidney transplant recipients was 48 y (interquartile range [IQR], 21 y), and the majority were male (73%) of Caucasian ethnicity (82%) (Table 1). The median donor age was 42 y (IQR, 24 y). The median follow-up time was 8.04 y (IQR, 14.90 y). The majority of kidney transplants occurred during the periods of 1990 to 1999 (42%) and 2010 to 2019 (41%). Living donation occurred in 27% and blood group incompatible transplants occurred in 3% of recipients. The proportion of recipients with a total HLA mismatch of 0 was 6%, 1 or 2 was 32%, 3 or 4 was 39%, and 5 or 6 was 23%. Biopsy-proven acute cellular rejection occurred in 19%, biopsy-proven acute antibody-mediated rejection occurred in 3%, and the median maximum pretransplant panel reactive antibody was 0% (IQR, 20). The most common induction immunosuppression was interleukin-2 (IL2) receptor blocker (37%), and the most common maintenance immunosuppressive agents were tacrolimus (50%) or cyclosporin (56%), mycophenolate (67%), and corticosteroids (99%).

A total of 25 kidney transplant recipients experienced recurrent MN (13%), predominantly during the period of 1990 to 1999 (68%). The time to biopsy-proven disease recurrence was reported in 17 of 25 kidney transplant recipients experiencing recurrent MN (median, 4.72 y [IQR, 4.13 y]). Two recipients had recurrent MN in >1 kidney transplant, and only the first episode of recurrence was included in the study. HLA serotypes in donor-recipient pairs are shown in **Table S1** (**SDC**, http://links.lww.com/TXD/A438).

Group LASSO

Model performance using Group LASSO for predicting recurrent MN in the derivation cohort was good (AUC-ROC, 0.85; 95% CI, 0.76-0.94 and model accuracy 88%, 84%-93%) (Figure 2). Tenfold cross-validation in the derivation cohort yielded an average error rate of 12.6%. Of the 222 variables (Table 2), Group LASSO selected 10 variables in at least 50 of the 100 imputed datasets as driving model predictions (Figure 3). Of these 10 selected variables, 6 were selected in at least 90 of the 100 imputed datasets. Variables driving model prediction were prior history of hepatitis C, IL2 receptor blocker induction, mycophenolate maintenance, azathioprine maintenance, other maintenance immunosuppression (medications included in this variable are detailed in Table 1), recipient HLA-A2, recipient HLA-B65, donor HLA-DR12, donor-recipient HLA-B65 match, and donor-recipient HLA-DR12 match. Six variables were chosen in all cross-validated folds and of equal importance: mycophenolate maintenance, other maintenance immunosuppression, recipient HLA-A2, recipient HLA-B65, donor-recipient HLA-B65 match, and donor-recipient HLA-DR12 match (Table 3).

Sensitivity Analysis

In the subgroup of kidney transplant recipients experiencing recurrent MN with recorded time-to-recurrence (17/25 recipients), penalized Cox regression performed well in predicting recurrent MN in the derivation cohort (AUC-ROC 0.91, 0.85-0.96 and model accuracy 69%, 63%-76%) (Figure 2). Nine variables, all chosen in at least 90 of the 100 imputed datasets, were selected as driving model predictions, including recipient HLA-A2, recipient HLA-B65, recipient HLA-DR17, donor HLA-DR12, donor HLA-DQ4, donor-recipient HLA-B51 match, donor-recipient HLA-B65 match, donor-recipient HLA-DR12 match, and donor-recipient HLA-DR17 match (Figure 3). Of these 9 variables, 5 (55%) were also chosen

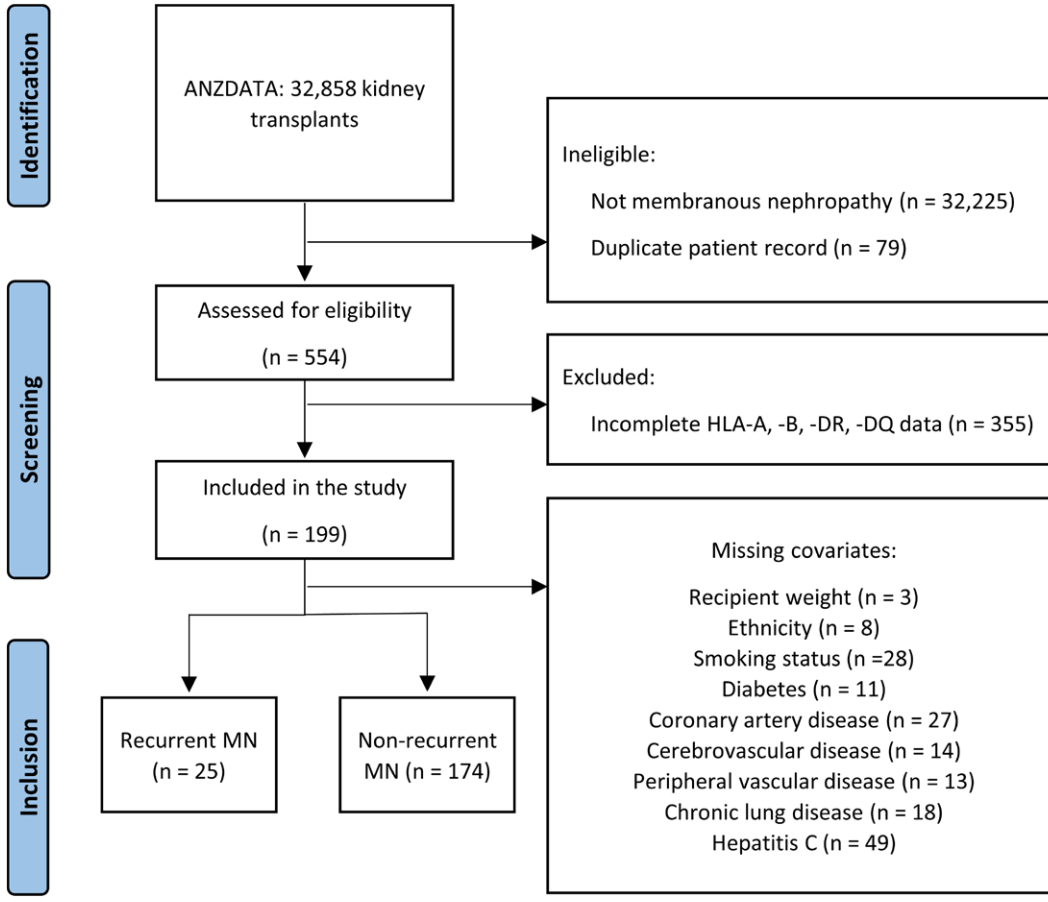

FIGURE 1. TRIPOD flow diagram: patient identification and selection. ANZDATA, Australian and New Zealand Dialysis and Transplant; MN, membranous nephropathy; TRIPOD, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis.

by Group LASSO. Eight variables were chosen in all cross-validated folds and of equal importance: recipient HLA-A2, recipient HLA-B65, recipient HLA-DR17, donor HLA-DR12, donor HLA-DQ4, donor-recipient HLA-B65 match, donor-recipient HLA-DR12 match, and donor-recipient HLA-DR17 match (Table 3).

Random forest performed poorer for predicting recurrent MN in the derivation cohort than Group LASSO (AUC-ROC 0.62, 0.57-0.69 and model accuracy 51%, 44%-58%) (Figure 2). Random forest selected 7 variables in at least 50 of the 100 imputed datasets as driving model predictions (Figure 3), all of which were also selected by Group LASSO. The top 3 variables driving model predictions were donor-recipient HLA-DR12 match, recipient HLA-A2, and mycophenolate maintenance (Table 3).

Model Validation

The validation cohort for the Group LASSO model was created from kidney transplant recipients with MN and complete HLA serotyping for recipient HLA-A/B/DR and donor HLA-B/DR (required for model predictors) but missing data for donor HLA-A/DQ and recipient HLA-DQ (therefore excluded from the derivation cohort). This comprised 275 kidney transplant recipients, 34 (12%) of whom had recurrent MN (**Table S2, SDC**, http://links.lww.com/TXD/A438). Group LASSO model performance in the validation cohort was poor (AUC-ROC 0.60, 0.49-0.70 and model accuracy 61%, 55%-67%) (Figure 4). In comparison, the penalized Cox regression model performance in its validation cohort was reasonable (AUC-ROC 0.73, 0.59-0.86 and model accuracy 59%, 49%-68%) (Figure 4). Kidney transplant recipients in this validation cohort had complete HLA serotyping for recipient HLA-A/B/DR and donor HLA-B/DR/DQ but missing data for donor HLA-A and recipient HLA-DQ (99 kidney transplant recipients, 7 [7%] with recurrent MN [**Table S3, SDC**, http://links.lww.com/TXD/A438]). In the random forest model, variables from all recipient and donor HLA serotype groups contributed to model predictions, and a validation cohort could not be generated.

DISCUSSION

In this study of almost 200 kidney transplant recipients with MN, we developed cross-validated prediction models for recurrent MN using Group LASSO, penalized Cox regression, and random forest, which incorporated routinely collected clinical data from a nationwide transplant cohort (ANZDATA) including donor-recipient HLA serotypes and HLA mismatch characteristics. We found Group LASSO and penalized Cox regression performed well at predicting recurrent MN in the derivation cohort and had superior performance to random forest. However, Group LASSO and penalized Cox regression model performance in their respective validation cohorts were poor and reasonable, respectively, although differences in the validation cohorts for each model preclude direct comparison of model performance. There

TABLE 1.

Characteristics of kidney transplant recipients with recurrent membranous nephropathy and nonrecurrent membranous nephropathy in the derivation cohort.

Variables	Recurrent membranous nephropathy (n = 25)	Nonrecurrent membranous nephropathy (n = 174)
Age, median (IQR), y		
Recipient	42 (22)	49 (20)
Donor	37 (25)	44 (23)
Recipient sex: male, n (%)	18 (72)	127 (73)
Recipient weight, median (IQR), kg	75 (15)	76 (24)
Recipient ethnicity, n (%)		
Caucasian	20 (80)	143 (82)
Asian	3 (12)	11 (8)
Central or South American	0 (0)	2 (1)
South Pacific Island	0 (0)	3 (2)
Middle Eastern	0 (0)	4 (2)
African	0 (0)	1 (0.6)
Australian Aboriginal	1 (4)	3 (2)
Dialysis vintage, median (IQR), mo		
Transplant era, n (%)	22 (38)	23 (55)
1980–1989	3 (12)	5 (3)
1990–1999	17 (68)	67 (38)
2000–2009	2 (8)	24 (14)
2010–2019	3 (12)	78 (45)
Recipient blood group, n (%)		
A	10 (40)	71 (41)
AB	1 (4)	6 (3)
B	5 (20)	23 (13)
O	9 (36)	74 (43)
Donor blood group, n (%)		
A	8 (32)	66 (38)
AB	1 (4)	1 (0.6)
B	3 (12)	20 (11)
O	13 (52)	87 (50)
ABO blood group incompatible, n (%)	0 (0)	5 (3)
HLA-A loci mismatch, n (%)		
0	3 (12)	40 (23)
1	16 (64)	87 (50)
2	6 (24)	47 (27)
HLA-B loci mismatch, n (%)		
0	5 (20)	30 (17)
1	18 (72)	86 (49)
2	2 (8)	58 (33)
HLA-DR loci mismatch, n (%)		
0	10 (40)	57 (33)
1	11 (44)	75 (43)
2	4 (16)	42 (24)
HLA-DQ loci mismatch, n (%)		
0	14 (56)	78 (45)
1	10 (40)	75 (43)
2	1 (4)	21 (12)
Total HLA mismatch, n (%)		
0	2 (8)	10 (6)
1–2	7 (28)	56 (32)
3–4	12 (48)	66 (38)
5–6	4 (16)	42 (24)
Living donor, n (%)	9 (36)	44 (25)
Regraft, n (%)	3 (12)	32 (18)
Biopsy-proven rejection, n (%)		
Acute cellular rejection	2 (8)	30 (17)
Acute humoral rejection	0 (0)	6 (3)
Maximum panel reactive antibodies, median (IQR), %	3 (16)	0 (20)
Recipient comorbidities, n (%)		
Type 1 diabetes	0 (0)	1 (0.6)
Type 2 diabetes	0 (0)	15 (9)
Coronary artery disease	1 (4)	18 (10)
Cerebrovascular disease	1 (4)	2 (1)
Peripheral vascular disease	0 (0)	6 (3)
Chronic lung disease	0 (0)	7 (4)
Hepatitis C virus	0 (0)	1 (0.6)
Cancer	2 (8)	26 (15)
Smoking status, n (%)		
Never	13 (52)	81 (47)
Former	5 (20)	57 (33)
Current	1 (4)	14 (8)
Induction immunosuppression, n (%)		
IL2 receptor blocker	2 (8)	72 (41)
T-cell depletion	3 (12)	17 (6)
B-cell depletion	0 (0)	1 (0.5)
IVIG induction	0 (0)	5 (2)
Any induction	5 (20)	90 (52)
Maintenance immunosuppression, n (%)		
Tacrolimus	5 (20)	95 (55)
Cyclosporin	20 (80)	91 (52)
Mycophenolate	7 (28)	126 (72)
Azathioprine	19 (76)	63 (36)
mTOR inhibitor	1 (4)	23 (13)
Corticosteroids	25 (100)	172 (99)
Other maintenance therapy[a]	2 (8)	2 (1)

[a]Other maintenance immunosuppression includes chlorambucil, cyclophosphamide, leflunomide, sotrastaurin, and janus kinase 3 inhibitor.
IL2, interleukin-2; IQR, interquartile range; mTOR, mammalian target of rapamycin.

Continued next page

was reasonable agreement in variables driving predictions across the Group LASSO and random forest models, which were recipient HLA-A2, donor HLA-DR12, donor-recipient HLA-B65 match and donor-recipient HLA-DR12 match, mycophenolate maintenance, azathioprine maintenance, and IL2 receptor blocker induction. Caution is needed in interpreting the importance of immunosuppression variables chosen, as it may reflect differences in the transplant era between kidney transplant recipients with recurrent MN and those without recurrent MN in the ANZDATA registry. We investigated whether the association between zero HLA mismatch and recurrent glomerulonephritis was driven by specific HLA and found donor-recipient HLA-B65 match and donor-recipient HLA-DR12 match contributed to prediction models for recurrent MN.[21] Although this suggests that mechanisms independent of allograft rejection contribute to recurrent MN, the potential mechanism by which HLA matching contributes to recurrent glomerulonephritis remains poorly understood,

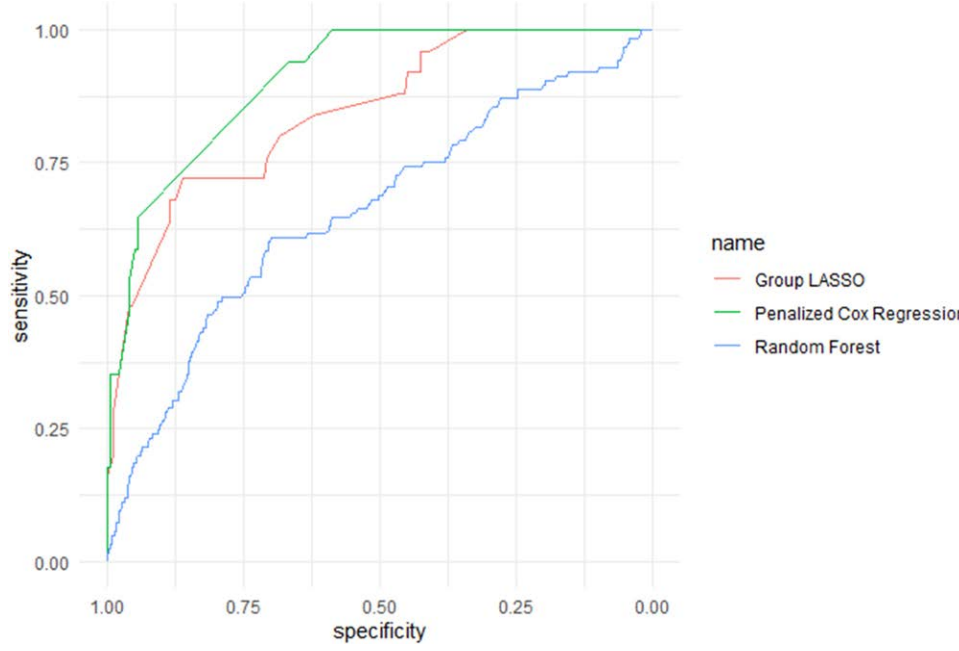

FIGURE 2. Receiver operating characteristic curves of the Group LASSO, penalized Cox regression, and random forest models in the derivation cohort. LASSO, Least Absolute Shrinkage and Selection Operator.

and its role should be interpreted with caution because this association has been mostly reported in retrospective studies. Variables were chosen by supervised feature selection using penalized regression and ensemble tree machine learning techniques, which addressed the issues of overfitting and multiple testing associated with stepwise feature selection in our high-dimensional dataset.[27,28]

Although multiple studies have identified risk factors for recurrent MN after kidney transplantation such as detectable pretransplant anti-PLA2R autoantibody, steroid-free immunosuppression, and recipient HLA-A3 and donor risk HLA-D and PLA2R1 alleles,[3,6,9,10,29] prediction models to identify patients at high risk of recurrent MN (or other forms of glomerulonephritis) using routine pretransplant clinical data have not been comprehensively evaluated. These predictive approaches may inform clinicians on patients requiring closer monitoring of proteinuria, anti-PLA2R antibody titer, and/or protocol allograft biopsies to detect recurrent MN posttransplantation. Early diagnosis of recurrent MN may facilitate timely treatment with rituximab.[30] However, external and prospective validations of our prediction models are required.

Agreement between the Group LASSO, penalized Cox regression, and random forest models in our study regarding the importance of donor-recipient HLA in predicting recurrent MN offers potential insights into the pathogenesis of recurrent MN, which remains incompletely understood. Class I HLA, such as HLA-A2, may play a role in directly activating cytotoxic CD8 T cells, which mediate glomerular injury downstream of antibody deposition in Heymann nephritis, a rat model of MN.[31] However, the role of CD8 T cells in MN is less clear with conflicting reports on the proportion of CD8 T cells in MN compared with healthy.[32,33]

The mechanism underpinning class II HLA-mediated autoimmune effector mechanisms or immune tolerance is better understood.[34] In recurrent MN, class II HLA-DR12 may activate autoreactive CD4 T cells. However, HLA-DR12 is different from MN risk alleles in the general population, which is unlikely explained by differences in podocyte antigens driving recurrent MN because of positive glomerular staining for PLA2R in recipients with recurrent PLA2R-associated MN.[6]

Alternatively, HLA-A2 and HLA-DR12 may contribute indirectly to recurrent MN via molecular mimicry, whereby donor HLA-recipient HLA complexes are also recognized by PLA2R-specific recipient T cells.[35] Finally, HLA-A2 and HLA-DR12 may not be mechanistically involved in recurrent MN but represent haplotypes associated with recurrent MN via linkage disequilibrium.[36] HLA-DR12 is in linkage disequilibrium with HLA-DQA1*0501, which is not consistent with a study by Berchtold et al[10] that identified 2 donor noncoding HLA-D SNPs associated with recurrent MN, none of which were in linkage disequilibrium with MN risk alleles such as HLA-DQA1*0501. Although Berchtold et al[10] performed molecular HLA typing and validated their results in a replication cohort, their analysis did not include class I HLA loci. Our results also disagree with a multicenter case series by Batal et al,[9] which analyzed HLA-A/B/DR/DQ serotypes and identified an association between recurrent MN and older recipient age, living related donors, steroid-free immunosuppression, and recipient HLA-A3. These discrepancies may be because of different diagnostic criteria for recurrent MN.

Our study has several strengths and limitations. We developed prediction models for recurrent MN in a large national transplant registry, which minimizes the potential risk of selection bias associated with case-series studies. These models utilized routinely collected pretransplant clinical data including donor-recipient HLA serotypes and HLA mismatch characteristics, which improves the potential translation of these models. We addressed the issue of missing data using

TABLE 2.

Summary of variables included in Group LASSO, penalized Cox regression, and random forest models

Type of variable	Number of variables (N = 222)
Recipient characteristics	
Age	1
Sex	1
Ethnicity	1
Weight	1
Dialysis vintage	1
Blood group	1
Comorbidities	8
Smoking status	1
Donor characteristics	
Age	1
Blood group	1
Transplant characteristics	
HLA mismatch (A, B, DR, DQ, total)	5
Donor status (living or deceased)	1
Blood group incompatibility status	1
Regraft status	1
Maximum panel reactive antibodies	1
Biopsy-proven rejection (acute cellular, acute antibody-mediated)	2
Induction immunosuppression	5
Maintenance immunosuppression	7
Transplant era	1
HLA characteristics	
Recipient HLA-A serotype	16
Donor HLA-A serotype	16
Recipient HLA-B serotype	25
Donor HLA-B serotype	27
Recipient HLA-DR serotype	18
Donor HLA-DR serotype	19
Recipient HLA-DQ serotype	9
Donor HLA-DQ serotype	9
Donor-recipient HLA-A match	10
Donor-recipient HLA-B match	14
Donor-recipient HLA-DR match	11
Donor-recipient HLA-DQ match	7

LASSO, Least Absolute Shrinkage and Selection Operator.

multiple imputation and the issue of high-dimensional data using Group LASSO penalized regression. We demonstrated the robustness of our results with another machine learning method (random forest) and used penalized Cox regression to analyze the subgroup of our cohort with time-to-recurrent MN data. Despite the relatively small derivation cohort used to build the prediction models, we utilized patients initially excluded for incomplete HLA data to form validation cohorts for our models, representing an efficient use of our limited dataset.

However, there were many limitations in this study. Firstly, our study sample size was small despite analyzing a national cohort of kidney transplant recipients, and the certainty of our results were low with wide CIs. Therefore, the findings of our study are solely exploratory. Furthermore, our study population was limited to Australia and New Zealand and was predominantly of Caucasian ethnicity, limiting the generalizability of our results. Second, molecular HLA typing was not available for any kidney transplant recipients with MN in the ANZDATA registry. It is likely that recurrent MN is driven by specific HLA allelic variants, and the lack of granular HLA data for all recipients and donors in transplant registries may potentially bias our findings because even single amino-acid substitutions in the HLA-DR alpha-chain can modify the peptide-binding groove and affect T-cell responses.[37] Furthermore, HLA-DQ serotyping was unavailable in 267 (48%) of the 554 kidney transplant recipients with MN, and HLA-DP serotyping was mostly unavailable. This may have limited our models' predictive performance because class II HLA is frequently associated with autoimmunity.[38,39] We also did not perform HLA imputation because of the inaccuracy of HLA serotype imputation compared with HLA allelic imputation. Third, detection and misclassification bias is likely present with a lower rate of recurrent MN in ANZDATA (13%) than in existing literature (35%–50%).[2-4] This may represent incomplete reporting and/or differences in clinical practice between treatment sites and time periods such as protocol biopsies posttransplantation, which have been associated with higher rates and earlier detection of disease recurrence (median time to recurrence 4–15 mo in studies implementing protocol biopsies compared with 56 mo in ANZDATA).[2,5,6] This suggests that the clinical practice of protocol biopsies was limited in our cohort and that biopsy-proven diagnosis of recurrent MN occurred late after detection of relatively high levels of proteinuria, although proteinuria at time of diagnosis was not available in our cohort. As a result, it is likely that mild forms of recurrent MN were not diagnosed, representing misclassification bias. Fourth, our findings were not confirmed in an external validation cohort. Despite evaluating a large transplant registry, cases of recurrent MN were small and HLA serotyping incomplete, which prevented us from dividing into separate derivation and validation cohorts. We addressed this issue by using tenfold cross-validation to balance efficient data usage and avoid individual splits in our dataset that may have been poor choices. Furthermore, we developed separate validation cohorts for the Group LASSO and penalized Cox regression models using kidney transplant recipients initially excluded for incomplete HLA data but retained predictors driving each model. However, validation cohorts differed between models, preventing direct comparison of model performance. Fifth, anti-PLA2R antibody titer at transplantation, pretransplant proteinuria, and pretransplant serum albumin, known predictors of recurrent MN posttransplantation, were not available in the ANZDATA registry.[3] However, anti-PLA2R antibody data would be unavailable for most recipients because PLA2R was discovered 46 y after the inception of the ANZDATA registry. Furthermore, discovery of other podocyte antigens targeted by autoantibodies in MN over the past decade suggest MN is composed of different disease entities with distinct drivers of disease recurrence.[40,41] Lastly, our relatively small sample size with a high proportion of missingness may affect the statistical validity of using multiple imputation to handle missing data.

In conclusion, we developed Group LASSO, penalized Cox regression, and random forest prediction models for recurrent MN in kidney transplant recipients, which remains

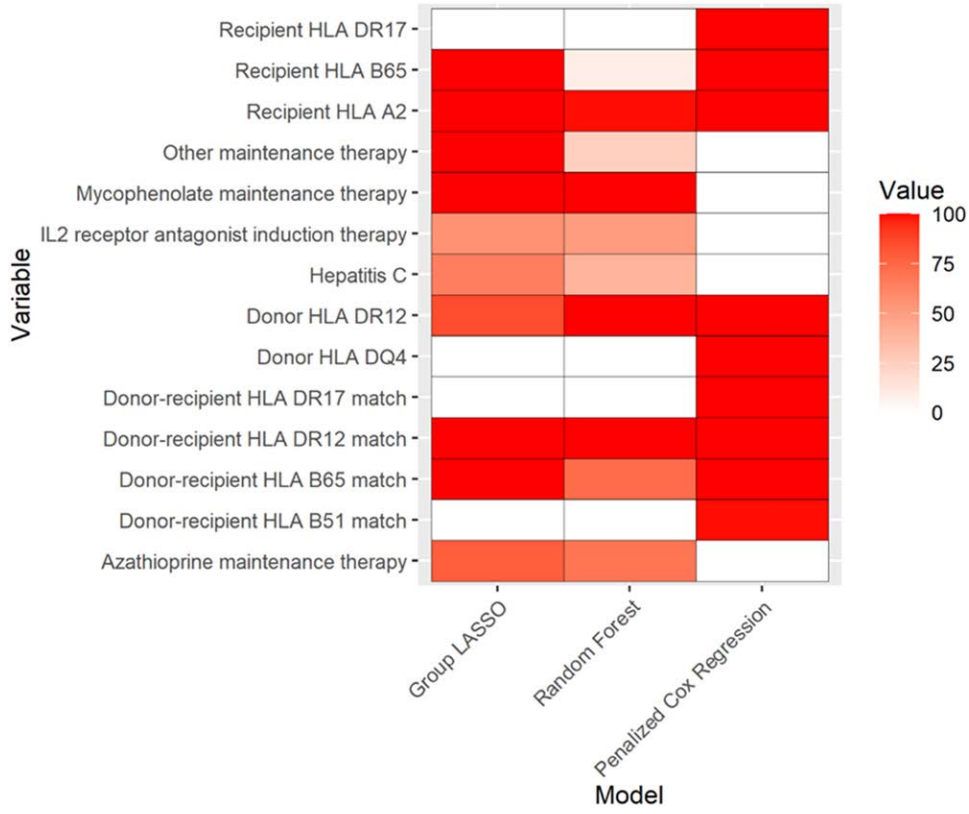

FIGURE 3. Variables selected by Group LASSO, penalized Cox regression, and random forest models as predictive of recurrent membranous nephropathy in kidney transplant recipients. IL2, interleukin 2; LASSO, Least Absolute Shrinkage and Selection Operator.

TABLE 3.

Variable importance in Group LASSO, penalized Cox regression, and random forest models

Variable ranking (1 = highest importance, 10 = lowest importance)	Group LASSO	Penalized Cox regression	Random forest
1	Mycophenolate maintenance Other maintenance immunosuppression[a] Recipient HLA-A2 Recipient HLA-B65 Donor-recipient HLA-B65 match Donor-recipient HLA-DR12 match	Recipient HLA-A2 Recipient HLA-B65 Recipient HLA-DR17 Donor HLA-DR12 Donor HLA-DQ4 Donor-recipient HLA-B65 match Donor-recipient HLA-DR12 match Donor-recipient HLA-DR17 match	Donor-recipient HLA-DR12 match
2			Recipient HLA-A2
3			Mycophenolate maintenance
4			Donor HLA-DR12
5			Donor-recipient HLA-B65 match
6			Azathioprine maintenance
7	Donor HLA-DR12		IL2 receptor blocker induction
8	Azathioprine maintenance		
9	Hepatitis C		
10	IL2 receptor blocker induction		

[a]Other maintenance immunosuppression includes chlorambucil, cyclophosphamide, leflunomide, sotrastaurin, and janus kinase 3 inhibitor.
IL2, interleukin 2; LASSO, Least Absolute Shrinkage and Selection Operator.

incompletely understood, challenging to predict, and an important cause of allograft loss in recipients with MN. Donor-recipient HLA characteristics were major drivers of model predictions. Our findings are exploratory and require further external and prospective validation in larger datasets. Future studies evaluating prediction models for recurrent MN utilizing molecular HLA typing, stratified by underlying disease antigen and associated autoantibody status before transplantation, may be possible using data linkage between transplant registries and existing biorepositories, which may

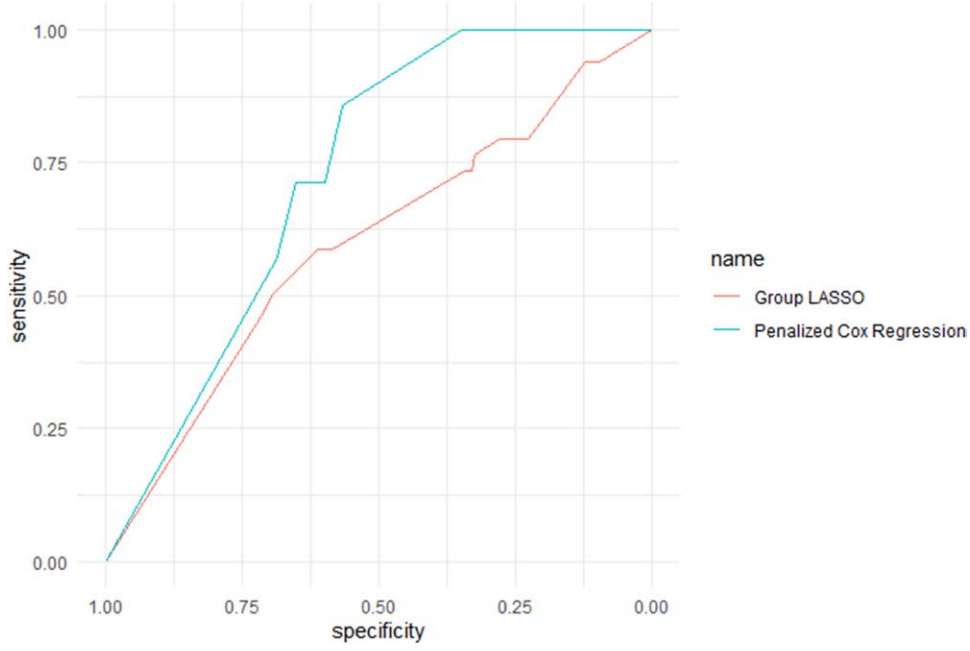

FIGURE 4. Receiver operating characteristic curves of the Group LASSO and penalized Cox regression models in their respective validation cohorts. LASSO, Least Absolute Shrinkage and Selection Operator.

assist clinical decision making and management of kidney transplant recipients with MN.

ACKNOWLEDGMENTS

We thank Anne Taverniti for her advice on HLA imputation techniques.

REFERENCES

1. Lerner GB, Virmani S, Henderson JM, et al. A conceptual framework linking immunology, pathology, and clinical features in primary membranous nephropathy. *Kidney Int*. 2021;100:289–300.
2. Dabade TS, Grande JP, Norby SM, et al. Recurrent idiopathic membranous nephropathy after kidney transplantation: a surveillance biopsy study. *Am J Transplant*. 2008;8:1318–1322.
3. Grupper A, Cornell LD, Fervenza FC, et al. Recurrent membranous nephropathy after kidney transplantation: treatment and long-term implications. *Transplantation*. 2016;100:2710–2716.
4. Moroni G, Gallelli B, Quaglini S, et al. Long-term outcome of renal transplantation in patients with idiopathic membranous glomerulonephritis (MN). *Nephrol Dial Transplant*. 2010;25:3408–3415.
5. Debiec H, Martin L, Jouanneau C, et al. Autoantibodies specific for the phospholipase A2 receptor in recurrent and de novo membranous nephropathy. *Am J Transplant*. 2011;11:2144–2152.
6. Kattah A, Ayalon R, Beck LH Jr, et al. Anti-phospholipase A_2 receptor antibodies in recurrent membranous nephropathy. *Am J Transplant*. 2015;15:1349–1359.
7. Stanescu HC, Arcos-Burgos M, Medlar A, et al. Risk HLA-DQA1 and PLA(2)R1 alleles in idiopathic membranous nephropathy. *N Engl J Med*. 2011;364:616–626.
8. Xie J, Liu L, Mladkova N, et al. The genetic architecture of membranous nephropathy and its potential to improve non-invasive diagnosis. *Nat Commun*. 2020;11:1600.
9. Batal I, Vasilescu ER, Dadhania DM, et al. Association of HLA typing and alloimmunity with posttransplantation membranous nephropathy: a multicenter case series. *Am J Kidney Dis*. 2020;76:374–383.
10. Berchtold L, Letouzé E, Alexander MP, et al. HLA-D and PLA2R1 risk alleles associate with recurrent primary membranous nephropathy in kidney transplant recipients. *Kidney Int*. 2021;99:671–685.
11. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162:55–63.
12. Gonzalez-Galarza FF, McCabe A, Santos EJMD, et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res*. 2020;48:D783–D788.
13. Gragert L, Madbouly A, Freeman J, et al. Six-locus high resolution HLA haplotype frequencies derived from mixed-resolution DNA typing for the entire US donor registry. *Hum Immunol*. 2013;74:1313–1320.
14. Paunić V, Gragert L, Madbouly A, et al. Measuring ambiguity in HLA typing methods. *PLoS One*. 2012;7:e43585.
15. Robinson J, Halliwell JA, Hayhurst JD, et al. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res*. 2015;43:D423–D431.
16. Cook S, Choi W, Lim H, et al. Accurate imputation of human leukocyte antigens with CookHLA. *Nat Commun*. 2021;12:1264.
17. Jia X, Han B, Onengut-Gumuscu S, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One*. 2013;8:e64683.
18. Zheng X, Shen J, Cox C, et al. HIBAG–HLA genotype imputation with attribute bagging. *Pharmacogenomics J*. 2014;14:192–200.
19. Naito T, Suzuki K, Hirata J, et al. A deep learning method for HLA imputation and trans-ethnic MHC fine-mapping of type 1 diabetes. *Nat Commun*. 2021;12:1639.
20. Allen PJ, Chadban SJ, Craig JC, et al. Recurrent glomerulonephritis after kidney transplantation: risk factors and allograft outcomes. *Kidney Int*. 2017;92:461–469.
21. Jeon HJ, Kim YS, Lee SM, et al. The effect of recurrent glomerulonephritis and acute rejection episodes in zero human leukocyte antigen-mismatched kidney transplantation. *Transplant Proc*. 2012;44:600–606.
22. Au EH, Francis A, Bernier-Jean A, et al. Prediction modeling-part 1: regression modeling. *Kidney Int*. 2020;97:877–884.
23. Coorey CP, Sharma A, Muller S, et al. Prediction modeling-part 2: using machine learning strategies to improve transplantation outcomes. *Kidney Int*. 2021;99:817–823.
24. Huang J, Ma S, Zhang CH. Adaptive Lasso for sparse high-dimensional regression models. *Stat Sin*. 2008;18:1603–1618.
25. Simon N, Friedman J, Hastie T, et al. A sparse-group lasso. *J Comput Graph Stat*. 2013;22:231–245.
26. Genuer R, Poggi JM, Tuleau-Malot C. Variable selection using random forests. *Pattern Recognit Lett*. 2010;31:2225–2236.
27. Emmert-Streib F, Dehmer M. High-dimensional LASSO-based computational regression models: regularization, shrinkage, and selection. *Mach Learn Knowl Extr*. 2019;1:359–383.
28. Borboudakis G, Tsamardinos I. Forward-backward selection with early dropping. *J Mach Learn Res*. 2019;20:276–314.

29. Quintana LF, Blasco M, Seras M, et al. Antiphospholipase A2 receptor antibody levels predict the risk of posttransplantation recurrence of membranous nephropathy. *Transplantation*. 2015;99:1709–1714.
30. Sprangers B, Lefkowitz GI, Cohen SD, et al. Beneficial effect of rituximab in the treatment of recurrent idiopathic membranous nephropathy after kidney transplantation. *Clin J Am Soc Nephrol*. 2010;5:790–797.
31. Penny MJ, Boyd RA, Hall BM. Permanent CD8(+) T cell depletion prevents proteinuria in active Heymann nephritis. *J Exp Med*. 1998;188:1775–1784.
32. Cagnoli L, Tabacchi P, Pasquali S, et al. T cell subset alterations in idiopathic glomerulonephritis. *Clin Exp Immunol*. 1982;50:70–76.
33. Rosenzwajg M, Languille E, Debiec H, et al. B- and T-cell subpopulations in patients with severe idiopathic membranous nephropathy may predict an early response to rituximab. *Kidney Int*. 2017;92:227–237.
34. Ooi JD, Petersen J, Tan YH, et al. Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells. *Nature*. 2017;545:243–247.
35. Robson KJ, Kitching AR. Recurrent membranous nephropathy after transplantation: donor antigen and HLA converge in defining risk. *Kidney Int*. 2021;99:545–548.
36. Morris DL, Taylor KE, Fernando MM, et al; International MHC and Autoimmunity Genetics Network; Systemic Lupus Erythematosus Genetics Consortium. Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. *Am J Hum Genet*. 2012;91:778–793.
37. Zeliszewski D, Golvano JJ, Gaudebout P, et al. Implication of HLA-DR residues at positions 67, 71, and 86 in interaction between HLA-DR11 and peptide HA306-320. *J Immunol*. 1993;151:6237–6247.
38. Gough SC, Simmonds MJ. The HLA region and autoimmune disease: associations and mechanisms of action. *Curr Genomics*. 2007;8:453–465.
39. Gebe JA, Swanson E, Kwok WW. HLA class II peptide-binding and autoimmunity. *Tissue Antigens*. 2002;59:78–87.
40. Tomas NM, Beck LH Jr, Meyer-Schwesinger C, et al. Thrombospondin type-1 domain-containing 7A in idiopathic membranous nephropathy. *N Engl J Med*. 2014;371:2277–2287.
41. Sethi S. New 'antigens' in membranous nephropathy. *J Am Soc Nephrol*. 2021;32:268–278.