The MEST score provides earlier risk prediction in IgA nephropathy

see commentary on page 19

Sean J. Barbour[1,2,3], Gabriela Espino-Hernandez[2], Heather N. Reich[4], Rosanna Coppo[5], Ian S.D. Roberts[6], John Feehally[7], Andrew M. Herzenberg[4,8] and Daniel C. Cattran[4]; for the Oxford Derivation, North American Validation and VALIGA Consortia

[1]*Division of Nephrology, University of British Columbia, Vancouver, BC, Canada;* [2]*BC Provincial Renal Agency, Vancouver, BC, Canada;* [3]*Centre for Health Evaluation and Outcome Sciences, St. Paul's Hospital, Vancouver, BC, Canada;* [4]*Division of Nephrology, University of Toronto, Toronto, ON, Canada;* [5]*Regina Margherita Children's University Hospital, Torino, Italy;* [6]*Oxford University Hospitals, Oxford, UK; and* [7]*The John Walls Renal Unit, Leicester General Hospital, Leicester, UK*

The Oxford Classification of IgA nephropathy (IgAN) includes the following four histologic components: mesangial (M) and endocapillary (E) hypercellularity, segmental sclerosis (S) and interstitial fibrosis/tubular atrophy (T). These combine to form the MEST score and are independently associated with renal outcome. Current prediction and risk stratification in IgAN requires clinical data over 2 years of follow-up. Using modern prediction tools, we examined whether combining MEST with cross-sectional clinical data at biopsy provides earlier risk prediction in IgAN than current best methods that use 2 years of follow-up data. We used a cohort of 901 adults with IgAN from the Oxford derivation and North American validation studies and the VALIGA study followed for a median of 5.6 years to analyze the primary outcome (50% decrease in eGFR or ESRD) using Cox regression models. Covariates of clinical data at biopsy (eGFR, proteinuria, MAP) with or without MEST, and then 2-year clinical data alone (2-year average of proteinuria/MAP, eGFR at biopsy) were considered. There was significant improvement in prediction by adding MEST to clinical data at biopsy. The combination predicted the outcome as well as the 2-year clinical data alone, with comparable calibration curves. This effect did not change in subgroups treated or not with RAS blockade or immunosuppression. Thus, combining the MEST score with cross-sectional clinical data at biopsy provides earlier risk prediction in IgAN than our current best methods.

Kidney International (2016) **89,** 167–175; http://dx.doi.org/10.1038/ki.2015.322

KEYWORDS: glomerular disease; IgA nephropathy; renal pathology

© 2015 International Society of Nephrology

Correspondence: *Daniel C. Cattran, Division of Nephrology, University of Toronto, 585 University Avenue, 11PMB 113, Toronto, ON M5G 2N2, Canada. E-mail: daniel.cattran@uhn.ca*

[8]Deceased.

Received 16 June 2015; revised 17 August 2015; accepted 3 September 2015; published online 28 October 2015

Until recently, there has not been a reproducible and validated histologic classification of IgA nephropathy (IgAN). The MEST score as part of the Oxford Classification overcomes these obstacles, and various components of its score have been validated in multiple studies worldwide to be associated with hard renal outcomes independent of kidney function, blood pressure, and proteinuria both at presentation and over time.[1–5] However, it remains largely unknown whether the MEST score can quantitatively improve the prediction of individual patient prognosis and guide management decisions at the time of biopsy.

The current approach to determining the risk of renal progression in IgAN using clinical data alone is challenging owing to the highly variable nature of the disease. Previous studies suggest that 2 years or longer of follow-up proteinuria and blood pressure data is needed before a clinically meaningful prediction can be achieved.[6–11] This approach has limited utility in clinical practice given current guidelines that recommend treatment decisions based mostly on clinical features near the time of biopsy.[12] We hypothesize that by adding the MEST score from the Oxford Classification to clinical data available at the time of biopsy, we can improve risk stratification earlier in the course of disease and predict the risk of renal outcome to the same degree as using longitudinal blood pressure and proteinuria over 2 years of follow-up. If the MEST score can achieve accurate risk stratification 2 years sooner than methods used in current clinical practice, it would allow earlier modification of patient treatment, which in turn may help preserve functioning nephron mass.

To address our hypothesis, we pooled cohorts from the VALIGA, Oxford, and North American validation studies in IgAN to compare the prediction of a hard renal outcome using the combination of renal function, blood pressure, and proteinuria at biopsy with and without the MEST score versus using only renal function and longitudinal changes in blood pressure and proteinuria over 2 years.[1,3,4] Because the use of renin-angiotensin system blockade (RASB) prior to biopsy and immunosuppression use during follow-up have the potential to impact the relationship between pathology and renal outcome, we repeated our analyses in *a priori* defined subgroups on the basis of the use of these medications.

RESULTS

Description of the cohort

There were 901 patients included in the analysis (Figure 1), and a description of the cohort is provided in Table 1. Overall, RASB was used in 38.4% at the time of biopsy and in 85.8% during follow-up starting a median of 0.6 months after biopsy (interquartile range 0, 11.5). Immunosuppression was used in 35.7% starting a median of 1.9 months after biopsy (interquartile range 0.1, 7.2). The primary renal outcome was a composite of end-stage renal disease (ESRD) or a 50% reduction in estimated glomerular filtration rate (eGFR) compared with baseline. This occurred in 18% ($N = 162$) of patients and was composed of 21.6% ($N = 36$) from the Oxford study, 16.1% ($N = 14$) from the North American validation study, and 17.3% ($N = 112$) from the VALIGA study. The 5- and 10-year risks of the composite renal outcome were 11.2% and 26.8%, respectively, as shown in the Kaplan–Meier curves in Figure 2.

Using MEST for risk prediction at the time of biopsy

The risks of the composite renal outcome associated with the combination of MEST plus clinical data at biopsy (eGFR, mean arterial blood pressure [MAP], and proteinuria) or with 2-year clinical data alone (eGFR at biopsy and 2-year averages of MAP and proteinuria) are shown in Table 2. MEST as a group ($P < 0.0001$) and T1, T2, and M1 scores individually ($P < 0.0001$ and 0.018) were significantly associated with the renal outcome independent of clinical data at biopsy.

When MEST was added to clinical data at biopsy, there was improvement in the prediction of the composite renal outcome compared with using clinical data alone. There was an increase in R^2 by 5.5% (12.6 vs. 18.1%) and a reduction in Akaike Information Criterion (AIC) by 41 (1777 down to 1736), demonstrating better model fit. There was also significant improvement in the ability to discriminate between those who did or did not experience the composite renal outcome 5 years after biopsy as measured by the change (Δ) in C-statistic (0.05), continuous net reclassification improvement (cNRI) (0.28), and integrated discrimination improvement (IDI) (0.06) with 95% confidence interval (CI) that were greater than the null value (see Table 3). Consistent with previous studies, 2-year clinical data predicted the composite renal outcome better than clinical data at the time of biopsy.[10,11] This was demonstrated by an increase in R^2 by 6.5% (12.6 vs. 19.1%), reduction in AIC by 60 (1777 down to 1717), and significant improvements in the ΔC-statistic (0.05), cNRI (0.38), and IDI (0.06) (see Table 3). The equations for each regression model that were used to calculate the probability of surviving 5 years without the composite renal outcome are provided in Table 3.

In addition, when MEST was added to the clinical data at biopsy, prediction of the composite renal outcome was similar to that using 2-year clinical data alone. This was evident by similar model fit with a difference in R^2 of only 1.0% (18.1 vs. 19.1%) and a difference in AIC of only 19 (1736 vs. 1717), and no significant changes in the ΔC-statistic (−0.007), cNRI (−0.08), and IDI (0.001) in Table 3 with 95% CI that included the null value. The receiver operating curves in Figure 3 demonstrate near superimposable curves for the two models. The calibration curves in Figure 4 show that there was similar and good calibration in both models, in that the observed and predicted risks were close to each other within the spectrum of predicted survival observed in the cohort (>80%). These results demonstrate that compared with using 2-year clinical data alone, the combination of MEST with clinical data at biopsy predicts the composite renal outcome with similar model fit and discrimination, and no loss in calibration.

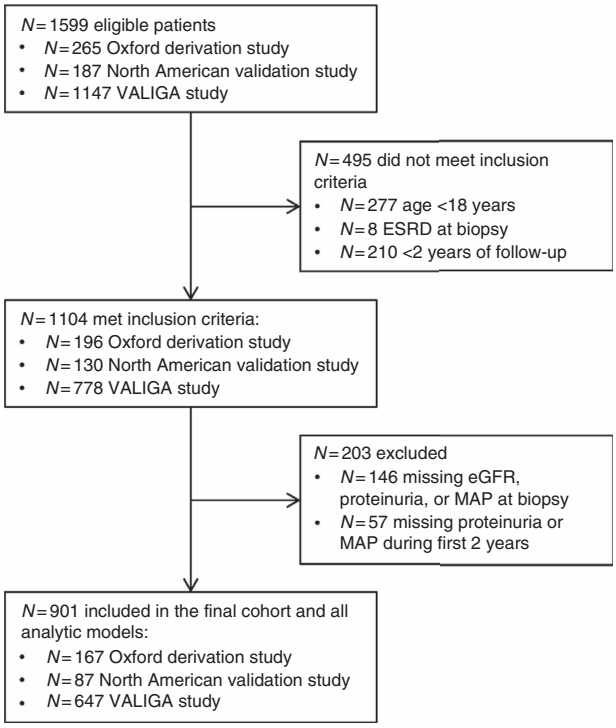

Figure 1 | Derivation of the cohort. eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; MAP, mean arterial blood pressure.

Sensitivity analyses within subgroups based on the use of RASB and immunosuppression

We performed sensitivity analyses in separate subgroups on the basis of RASB exposure at the time of biopsy, and on the basis of immunosuppression use during follow-up (see Supplementary Tables S1 and S2 online). In multivariable models that included clinical data at biopsy, the MEST score, and interaction terms between each MEST component and either RASB or immunosuppression exposure as appropriate, none of the interaction terms were significant (data not shown). This suggests that after adjusting for clinical data at biopsy, the association between each MEST component and the composite renal outcome did not depend on the prior use of RASB or subsequent use of immunosuppression. We repeated assessments of prediction performance in subgroups on the basis either of RASB or of immunosuppression exposure. In all subgroups, the primary findings regarding risk prediction in the overall cohort did not change (data not shown).

Table 1 | Description of the cohort

	Total $N = 901$	Oxford derivation $N = 167$	North American validation $N = 87$	VALIGA $N = 647$
Follow-up (years)	5.6 (3.8, 8.8)	6.8 (4.8, 9.2)	4.9 (3.9, 7.3)	5.4 (3.5, 8.8)
Age (years)	38.1 (29.2, 49.2)	36.1 (28.8, 46.8)	41.8 (31.3, 47.2)	38.4 (28.8, 50.2)
Male sex	640 (71%)	117 (70.1%)	51 (58.6%)	472 (73%)
Race				
Caucasian	780 (86.6%)	107 (64.1%)	45 (51.7%)	628 (97.1%)
Black	12 (1.3%)	4 (2.4%)	2 (2.3%)	6 (0.9%)
Asian	76 (8.4%)	52 (31.1%)	21 (24.1%)	3 (0.5%)
South Asian	9 (1%)	0 (0%)	0 (0%)	9 (1.4%)
Other	11 (1.2%)	4 (2.4%)	6 (6.9%)	1 (0.2%)
Creatinine at biopsy (μmol/l)	106.1 (85.0, 140.8)	105.6 (81.0, 130.0)	98.0 (80.0, 132.0)	106.1 (86.6, 147.0)
eGFR at biopsy (ml/min per 1.73 m^2)	68.4 (48.5, 88.6)	69.5 (55.3, 93.1)	69.8 (56.0, 85.8)	67.3 (45.9, 88.4)
MAP at biopsy (mm Hg)	100.0 (93.3, 106.7)	99.3 (90.7, 106.7)	98.7 (90.0, 106.0)	100 (93.3, 106.7)
MAP averaged over 2 years (mm Hg)	97.2 (91.3, 103.6)	97.0 (90.0, 103.3)	93.9 (88.4, 98.4)	97.9 (92.2, 104.4)
Proteinuria at biopsy (g per day)	1.5 (0.8, 2.6)	1.7 (1.1, 2.9)	1.6 (1.1, 2.8)	1.4 (0.7, 2.6)
Proteinuria averaged over 2 years (g per day)	1.1 (0.6, 2.1)	1.4 (0.8, 2.2)	1.2 (0.7, 2.0)	1.0 (0.5, 2.0)
Use of RASB at biopsy	346 (38.4%)	37 (22.2%)	27 (31%)	282 (43.6%)
Use of RASB during follow-up	773 (85.8%)	132 (79%)	75 (86.2%)	566 (87.5%)
Use of any immunosuppression during follow-up	322 (35.7%)	25 (15%)	32 (36.8%)	265 (41%)
Pathology				
M1	383 (42.5%)	131 (78.4%)	78 (89.7%)	174 (26.9%)
E1	163 (18.1%)	62 (37.1%)	27 (31%)	74 (11.4%)
S1	676 (75%)	132 (79%)	57 (65.5%)	487 (75.3%)
T1	161 (17.9%)	32 (19.2%)	15 (17.2%)	114 (17.6%)
T2	38 (4.2%)	8 (4.8%)	2 (2.3%)	28 (4.3%)
Crescents	154 (17.1%)	68 (40.7%)	28 (32.2%)	58 (9%)

eGFR, estimated glomerular filtration rate; IQR, interquartile range; MAP, mean arterial blood pressure; RASB, renin-angiotensin system blockade.
Data presented as median (IQR) or count (percentage).

The role of mesangial hypercellularity in reducing the need for 2-year clinical data

Because M1 was associated with the composite renal outcome when added to baseline clinical data (see Table 2), we explored whether mesangial hypercellularity may explain part of the benefit of the MEST score in predicting outcome at the time of biopsy. When we stratified the cohort on the basis of mesangial score, those with M1 compared with M0 had similar MAP but higher proteinuria at biopsy (1.8 vs. 1.2 g per day, $P < 0.001$), at 2 years (0.9 vs. 0.52 g per day, $P < 0.001$), and over the entire first 5 years of follow-up (see Supplementary Table S3 online and Supplementary Figure S1 online). In the model in Table 2 with MEST and clinical data at biopsy, the hazard ratio (HR) for M1 was 1.49 (95% CI

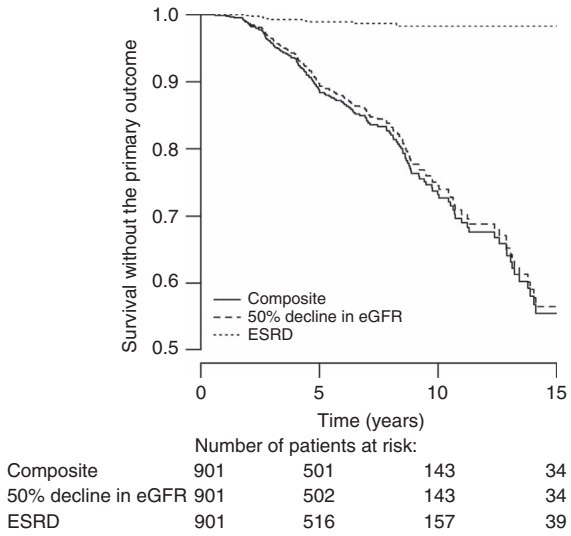

Figure 2 | The risk of the individual components that contribute to the primary composite renal outcome (first occurrence of either a 50% reduction in estimated glomerular filtration rate [eGFR] or end-stage renal disease [ESRD]). The 5-, 10-, and 15-year risk of survival without the composite outcome was 88.8%, 73.2%, and 55.4%, respectively.

Table 2 | The results of multivariable models for the risk of a 50% decline in eGFR or ESRD that included either the MEST score with clinical data at the time of biopsy, or clinical data over 2 years

	HR	95% CI	*P*-value
Model containing clinical data at biopsy			
eGFR at biopsy	0.99	0.98–0.99	0.020
Proteinuria at biopsy	1.59	1.29–1.96	<0.0001
MAP at biopsy	1.02	1.01–1.03	<0.0001
MEST score			<0.0001
M	1.49	1.07–2.07	0.018
E	1.15	0.78–1.71	0.483
S	1.31	0.81–2.12	0.267
T1	2.92	2.01–4.26	<0.0001
T2	4.21	2.28–7.78	<0.0001
Model containing 2-year clinical data			
eGFR at biopsy	0.98	0.97–0.99	<0.0001
Proteinuria over 2 years	2.62	2.14–3.22	<0.0001
MAP over 2 years	1.02	1.00–1.03	0.017

CI, confidence interval; eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; HR, hazard ratio; MAP, mean arterial blood pressure.
Proteinuria was log-transformed.

Table 3 | The discrimination performance of different models predicting the risk of a 50% reduction in eGFR or ESRD at 5 years after biopsy

	C-statistic (95% CI)	Δ C-statistic (95% CI)	cNRI (95% CI)	IDI (95% CI)
Models containing clinical data at biopsy				
Clinical data at biopsy[a]	0.75 (0.72, 0.79)	—	—	—
Clinical data at biopsy and MEST[b]	0.80 (0.77, 0.84)	0.05 (0.03, 0.08)	0.28 (0.16, 0.43)	0.06 (0.04, 0.11)
Models containing clinical data without MEST				
Clinical data at biopsy[a]	0.75 (0.72, 0.79)	—	—	—
2-year clinical data[c]	0.80 (0.77, 0.84)	0.05 (0.08, 0.03)	0.38 (0.5, 0.22)	0.06 (0.11, 0.03)
Models containing clinical data at biopsy with MEST, or 2-year clinical data alone				
2-year clinical data[c]	0.80 (0.77, 0.84)	—	—	—
Clinical data at biopsy and MEST[b]	0.80 (0.77, 0.84)	−0.007 (−0.03, 0.03)	−0.08 (−0.28, 0.17)	0.001 (−0.04, 0.07)

CI, confidence interval; cNRI, continuous net reclassification improvement; eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; IDI, integrated discrimination improvement; MAP, mean arterial blood pressure.
Clinical data at biopsy included eGFR, MAP, and proteinuria; 2-year clinical data included eGFR at biopsy and 2-year averages of MAP and proteinuria. For each regression model, we provide the equation to calculate the probability of survival without the composite renal outcome at 5 years.
Survival (5 years) = $S_o^{EL_P}$, where S_0 is the baseline survival at 5 years, and EL_P is the exponential of the linear predictor as below:
[a]So = 0.962; EL_P = exp(−0.019*eGFR at biopsy+0.553*ln(Proteinuria at biopsy)+0.020*MAP at biopsy)
[b]S_0 = 0.993; EL_P = exp(−0.009*eGFR at biopsy+0.464*ln(Proteinuria at biopsy)+0.022*MAP at biopsy+0.397*M+0.142*E+0.272*S+1.073*T1+1.438*T2)
[c]S_0 = 0.955; EL_P = exp(−0.020*eGFR at biopsy+0.964*ln(Proteinuria over 2 years)+0.018*MAP over 2 years).

1.07–2.07, $P = 0.018$). However, when proteinuria at biopsy was replaced by the average proteinuria over 2 years, the HR for M1 was substantially attenuated to 1.33 (95% CI 0.96–1.84, $P = 0.09$). In comparison, when MAP at biopsy was replaced MAP over 2 years, the HR for M1 was unchanged (HR = 1.47 95% CI 1.05–2.03, $P = 0.02$). These results suggest that M1 is associated with changes in proteinuria (but not MAP) over time, and the M1 score explains some of the same risk of the composite renal outcome as the average proteinuria (but not MAP) over 2 years.

Examples of using MEST to predict outcomes at the time of biopsy

To illustrate the clinical relevance of our results, we generated subgroups whose initial eGFR was ≥50 ml/min per 1.73 m^2 but with different proteinuria levels at biopsy. The addition of M and T scores identified patients at biopsy with similar risk as those with 2-year proteinuria values between 1 and 2 grams per day (who would qualify for steroid treatment in the KDIGO guidelines[12]) (see Figure 5). For example, compared with patients with 2-year proteinuria of 1–2 gram per day, the risk of renal outcome was similar in those with proteinuria at biopsy ≤ 1 gram per day but with M1 ($P = 0.56$). Conversely, the risk was much lower in patients with proteinuria at biopsy ≤ 1 gram per day but with M0 ($P = 0.008$), who had a similar risk as patients with proteinuria at biopsy of 1–1.5 gram per day but with M0 and T0 despite the higher degree of proteinuria ($P = 0.67$).

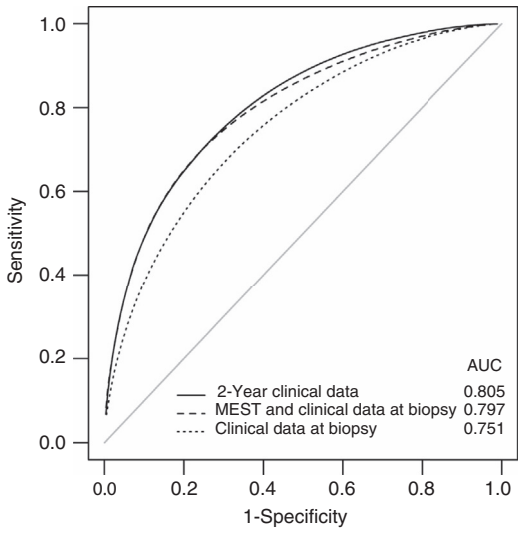

Figure 3 | The receiver operating curves for models predicting the 5-year risk of a 50% reduction in estimated glomerular filtration rate (eGFR) or end-stage renal disease (ESRD) using clinical data at biopsy with or without MEST, and 2-year clinical data alone. AUC, area under the curve.

DISCUSSION

We used a large well-characterized cohort of patients with IgAN to demonstrate that the addition of MEST to baseline eGFR, blood pressure, and proteinuria substantially improves prediction of the patient-level risk of a 50% decline in renal function or ESRD. More particularly, it has comparable accuracy with the currently established method using 2 years of blood pressure and proteinuria measurements to predict the subsequent risk of renal function decline.[10] In addition, the quantitative benefit of using MEST to predict long-term outcomes was unchanged in subgroups based on the prior use of RASB or subsequent use of immunosuppression. This is the first study to demonstrate that the MEST score improves the clinical utility of risk stratification in IgAN by allowing accurate prediction at the time of biopsy. This potentially eliminates the need for 2 years of follow-up data before making patient-related treatment decisions.

IgAN is an extremely heterogeneous disease with highly variable risk of kidney function decline, and currently, there is no established prediction tool widely used in clinical practice.[6–9] Previously, it has been demonstrated that the Lee histology classification was not independently associated with

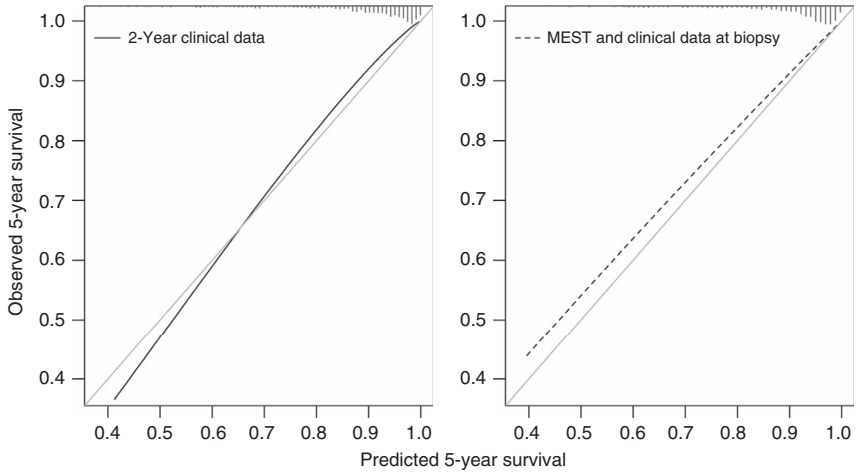

Figure 4 | The calibration plots of predicted versus observed survival without a 50% reduction in estimated glomerular filtration rate (eGFR) or end-stage renal disease (ESRD) at 5 years, as determined using either MEST and clinical data at biopsy, or 2-year clinical data alone. The line depicting perfect calibration is shown in light gray. The hash marks along the top axis represent the number of patients at each level of predicted risk and show that the majority of the cohort had a predicted survival greater than 80%. The calibration plots for the two models are very similar in the range of predicted risk observed in the cohort.

outcome, that baseline clinical data were insufficient to predict the rate of renal function decline, and that at least 2 years of blood pressure and proteinuria measurements were required.[10,11] The major limitation of this approach in clinical practice is the 2-year wait time. Earlier prediction of outcomes would be preferable, as it would allow the introduction of effective treatments 2 years sooner, which may improve the preservation of nephron mass and potentially delay the progression of IgAN. To overcome this limitation, newer prediction models have combined pathology features with clinical data at biopsy, but most have used histology classifications that have not been validated or that are not independently correlated with outcome, and none have compared their results with using blood pressure and proteinuria measured over time.[7,13–17] Since the publication of the original Oxford Classification, multiple studies have shown an association between some or all of the MEST components and renal outcome independent of blood pressure and proteinuria during follow-up.[1–3,5] Our results expand upon these observations by using newer statistical techniques to evaluate the prediction of individual patient-level outcomes in a way that is not possible from simple multivariable analysis. We demonstrate that when the MEST score is combined with cross-sectional data available at renal biopsy (i.e., eGFR, blood pressure, and proteinuria), it is possible to predict renal outcome with similar accuracy as if one had 2 years of follow-up data. This effect cannot be observed from our multivariable model results alone, but is instead apparent from our more detailed prediction analysis. We show, as examples, that patients with mesangial hypercellularity (M1) and proteinuria at biopsy ≤ 1 gram per day have a similar renal prognosis as those with persistent proteinuria between 1 and 2 grams per day, suggesting that despite lower protein excretion, this is a high-risk group that may benefit from earlier immunosuppression potentially improving the preservation of nephron mass. Conversely, patients with higher grades of proteinuria of 1–1.5 grams per day at biopsy but with both minimal mesangial hypercellularity (M0) and minimal interstitial fibrosis (T0) have a favorable prognosis, suggesting prolonged observation and RASB is all that is required thereby sparing the potential toxicity of immunosuppressive therapy. Although individual risk of immunosuppression and specific patient characteristics were not included in our models, these illustrations demonstrate the potential benefits to clinical treatment decisions that arise from accurate risk prediction at the time of biopsy using the MEST score.

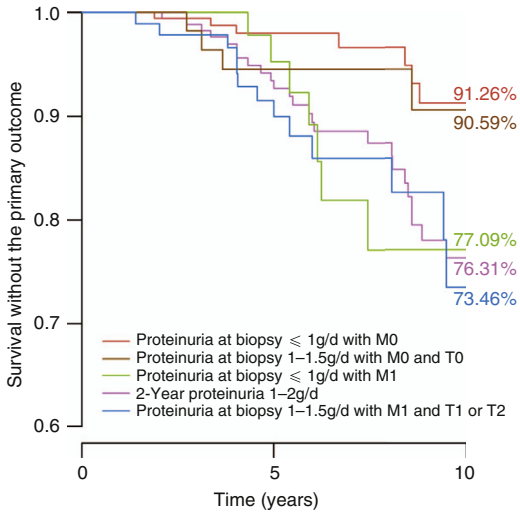

Figure 5 | The risk of the primary composite renal outcome (50% reduction in estimated glomerular filtration rate [eGFR] or end-stage renal disease) among those with eGFR at biopsy ≥50 ml/min per 1.73 m^2 and 2-year proteinuria 1–2 gram per day, proteinuria at biopsy ≤ 1 gram per day with different M scores, or proteinuria at biopsy 1–1.5 gram per day with different M and T scores. The 10-year survival is provided for each subgroup.

Immunosuppression and RASB are both thought to have structural effects on glomerular pathology, and clinical trials have shown that they decrease the risk of renal function decline in select patients with IgAN.[18–26] As such, it is important to consider whether the significance of the MEST score as a predictor of outcome is dependent on the prior use of RASB or the subsequent use of immunosuppression. We therefore repeated our analyses in subgroups based on treatment exposures. Our primary findings were unchanged irrespective of prior RASB or subsequent use of immunosuppression. However, owing to substantial confounding by indication, these results most accurately apply to patients who *happen to have been* exposed to these therapies. To truly evaluate the more clinically relevant situation of predicting patient outcomes under different treatment scenarios would require comparing *otherwise similar* groups of patients. This type of analysis could only be carried out in much larger cohorts to allow sufficient matching on treatment indications. Future international research collaborations that merge existing Oxford validation datasets may provide the infrastructure capable of such analyses.

This study provides novel insights into the importance of mesangial hypercellularity within the overall Oxford Classification. Whereas most studies demonstrate a clear association between tubulointerstitial fibrosis and renal outcome independent of clinical data either at biopsy or over time, there have been less consistent findings with mesangial hypercellularity.[1–3,27–29] We observed that an M1 score was associated with the composite renal outcome independent of clinical data at biopsy; that M1 was correlated with higher levels of proteinuria but not blood pressure both at biopsy and during follow-up; and that the risk of renal outcome explained by an M1 score was also explained by 2-year proteinuria but not 2-year blood pressure. Although the use of RASB or immunosuppression therapies may have altered proteinuria and were statistically different between mesangial score groups, there was no clear treatment pattern favoring one group over the other, and the absolute differences were small. This suggests that mesangial hypercellularity provides important information at the time of biopsy about the likelihood of higher levels of proteinuria over time, and that it therefore has an important and independent role in the prediction performance of the MEST score.

There are several limitations that should be considered when interpreting our results. Owing to different eras and inclusion criteria of the studies, there were differences in geography, treatment, and pathology between the Oxford and VALIGA cohorts. To account for any potential data clustering within cohorts that may have resulted, we performed a sensitivity analysis using shared frailty models. There was no substantial change in our multivariable model results and none of the frailty estimates were statistically significant (data not shown), suggesting it is unlikely that our primary findings depend on systematic differences between the cohorts. By necessity, we required at least 2 years of follow-up, which excluded patients who progressed to ESRD within that time potentially limiting generalizability. However, predicting outcome within this rapidly progressive group is clinically obvious and not representative of the overall IgAN population. We specifically did not consider changes in eGFR over 2 years in our analysis in order to be consistent with other studies that used only renal function at biopsy, and because our emphasis was on improving risk prediction at the time of biopsy.[1–4,10] Similar to other research investigating the MEST score in IgAN, treatment in our cohort was not standardized but was instead according to local practice. The majority (85.8%) were given RASB during follow-up starting a median of 0.6 months after biopsy; as such, our results most accurately generalize to this group of patients. Given the KDIGO GN guidelines that endorse RASB and the near universal use in clinical practice, this strengthens the real-world applicability of our results.[12,30] We did not have data on the dose of RASB or the duration of use prior to biopsy, and hence cannot exclude the possibility that these impacted our results. The largest R^2 for our models was only 19.1%, suggesting substantial variability in outcome not explained by pathology, renal function, proteinuria, and blood pressure. This may be improved in future research that considers other demographic and prognostic variables that by design were not included in our analysis, with the goal of systematically deriving and validating a comprehensive prediction model in IgAN suitable for use in diverse populations. However, our calibration curves do show sufficiently comparable predicted and observed risks that our results are nonetheless applicable in the clinical care of patients similar to our cohort.

In conclusion, using a large international cohort, we have shown that the combination of MEST score with readily available blood pressure, proteinuria, and eGFR at the time of biopsy predicted the composite renal outcome similar to using clinical data over 2 years of follow-up. We provide illustrations of the clinical utility and potential therapeutic implications of the earlier risk stratification provided by the Oxford Classification.

MATERIALS AND METHODS

Study design

This is a *post hoc* analysis of patients from the Oxford derivation ($N = 265$), North American validation ($N = 187$), and VALIGA ($N = 1147$) studies, which have been previously described in detail.[1,3,4] In brief, the Oxford derivation and North American validation studies included patients with biopsy-proven IgAN and initial proteinuria ≥ 0.5 g per day, eGFR ≥ 30 ml/min per 1.73 m^2, and follow-up ≥ 1 year. The VALIGA study included patients with biopsy-proven IgAN and either ≥ 1-year follow-up or progression to ESRD within 1 year, without any restriction on proteinuria or eGFR. From the combined cohort, we included patients in our analysis with age ≥ 18 years without ESRD at the time of biopsy and ≥ 2 years of follow-up. We excluded those missing eGFR, blood pressure, and proteinuria at the time of biopsy, or blood pressure or proteinuria over the first 2 years. These criteria were specifically chosen to ensure that all patients in the analysis had the necessary clinical variables to be included in all prediction models studied.

Definitions

Proteinuria, MAP, and eGFR at biopsy were the closest values within 6 months of biopsy, whereas 2-year proteinuria and MAP were the average of all values over the first 24 months after biopsy. GFR was estimated using the 4-variable Modified Diet in Renal Disease formula.[31] The use of RASB at biopsy was any exposure within 6 months prior to biopsy, and RASB or immunosuppression use during follow-up was any exposure after biopsy. The primary outcome was a composite of either ESRD (eGFR to <15 ml/min per 1.73 m^2), or a permanent reduction in eGFR to below 50% of the value at biopsy.

Pathology review

In all three studies, renal biopsies were scored according to the Oxford MEST scoring system by pathologists experienced in the classification given their involvement in the original Oxford study, the North American validation study, and the VALIGA study. They were all blinded to patient outcomes at the time of pathology review. Details of the histologic classification have been previously described, and all cases had a MEST score defined as follows: M0/M1 as a mesangial score $\leq$ or $>$ 0.5, or $\leq$ or >50% of glomeruli with $\geq$ 4 mesangial cells per mesangial area, E0/E1 as the presence or absence of endocapillary hypercellularity, S0/S1 as the presence or absence of segmental sclerosis or tuft adhesions, and T0/T1/T2 as the degree of tubular atrophy or interstitial fibrosis (<25%, 25–50%, >50%, respectively).[1,4]

Statistical analysis

Time from kidney biopsy to the composite renal outcome (censored at death or loss to follow-up) was analyzed using Cox proportional hazards models. The analyses were chosen to identify the additional predictive value of adding MEST to clinical data at biopsy compared with using clinical data over 2 years of follow-up. Three different statistical models were created: (i) clinical data at biopsy alone (eGFR, proteinuria, and MAP); (ii) clinical data at biopsy with MEST; and (iii) 2-year clinical data alone (eGFR at biopsy with 2-year proteinuria and MAP). Renal function at 2 years was not included in the models to be consistent with the methods used in the Bartosik, Oxford derivation, and North American validation, and VALIGA studies.[1–4,10] The functional form of all continuous variables was assessed, with log-transformation of proteinuria to improve linearity. The proportional hazards assumption was assessed using the global test for Schoenfeld partial residuals, and by adding time-dependent covariates obtained as an interaction between each variable and a heavyside step function of time. In both tests, there was insufficient evidence to reject the null hypothesis of proportional hazards. Overall model fit was assessed using Nagelgerke's R^2 and the AIC, with an increase in R^2 and reduction in AIC suggesting better model fit. Discrimination was assessed using the C-statistic, the cNRI (which does not require the *a priori* specification of probability cutoffs to define cases from non-cases), and the IDI, all adapted to account for censoring and evaluated using the 5-year risk of the composite outcome (as this corresponded to the median duration of follow-up of the cohort).[32–34] When comparing two models, a ΔC-statistic, cNRI, and IDI significantly greater than zero suggest improvement in discrimination. CIs were generated using 1000 bootstrap samples. Calibration plots were generated using lowess regression to plot predicted versus observed 5-year survival without the composite renal outcome, with observed risks taken from Kaplan–Meier estimates. To determine whether our results differed within *a priori* defined subgroups based on the use of RASB prior to biopsy or immunosuppression use during follow-up, we evaluated interaction effects between each MEST component and the RASB or immunosuppression variables, and repeated our analyses within each subgroup. To account for any effect that data clustering within cohorts may have had on our findings, we repeated the multivariable models using shared frailty survival analysis. The cluster effects were added as independent and identically distributed random variables assuming a normal distribution with mean zero and a common variance (i.e., a lognormal distribution for the frailty components).[35,36] The estimated common variance of the random effects were close to zero (<0.013) for all models, indicating no within-cluster correlation. Similarly, the estimates of the lognormal frailties were close to one for all clusters, and none were significant ($P > 0.19$). The HRs and CIs for all covariates in the shared frailty models were very similar compared with those obtained from the Cox models that did not account for cluster effects.

Descriptive variables were presented as median (interquartile range, owing to non-normal distributions) or frequency (count) and compared across relevant groups using the Kruskal–Wallis test for continuous variables and Fisher exact test for categorical variables. All tests were two-sided with $P < 0.05$ considered statistically significant. Analyses were performed using SAS Version 9.3 (SAS Institute, Cary, NC) and figures were generated using R 3.1.2 (R Core Team 2014, Vienna, Austria).

DISCLOSURE

All the authors declared no competing interests.

ACKNOWLEDGMENTS

We would like to acknowledge the investigators who contributed to the Oxford derivation, North American validation, and VALIGA studies. SJB is funded by the Carraresi Foundation and is a Scholar with the Michael Smith Foundation for Health Research. The VALIGA study was supported by a grant from the first research call of the European Renal Association—European Dialysis and Transplant Association in 2009. The Oxford derivation and North American validation studies were supported by the International IgA Nephropathy Network, the Toronto GN Registry, and the Toronto General Hospital Foundation.

THE VALIGA STUDY

The VALIGA steering committee includes R. Coppo (Torino, Italy), S. Troyanov (Montreal, Canada), D.C. Cattran (Toronto, Canada), H.T. Cook (London, UK), J. Feehally (Leicester, UK), and I. Roberts (Oxford, UK). The VALIGA nephrology investigators include V. Tesar, D. Maixnerova (Nephrology, First Faculty of Medicine and General University Hospital, Prague, Czech Republic); S. Lundberg (Nephrology, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden); L. Gesualdo (Nephrology, Emergency and Organ Transplantation, University of Bari "Aldo Moro", Foggia-Bari, Italy); F. Emma, L. Fuiano (Nephrology, Pediatrico Bambino Gesù Hospital, Rome, Italy); G. Beltrame, C. Rollino (Nephrology, San Giovanni Bosco Hospital, Turin, Italy); RC, A. Amore, R. Camilla, L. Peruzzi (Nephrology, Regina Margherita Children's Hospital, Turin, Italy); M. Praga (Nephrology, Hospital 12 de Octubre, Madrid, Spain); S. Feriozzi, R. Polci, (Nephrology, Belcolle Hospital, Viterbo, Italy); G. Segoloni, L. Colla (Nephrology, S. Giovanni Battista University Hospital, Turin, Italy); A. Pani, A. Angioi, L. Piras (Nephrology, G. Brotzu Hospital, Cagliari, Italy); JF (John Walls Renal Unit, Leicester General Hospital, Leicester, United Kingdom); G. Cancarini, S. Ravera (Nephrology, Spedali Civili University Hospital, Brescia, Italy); M. Durlik

(Transplantation Medicine and Nephrology, Warsaw Medical University, Warsaw, Poland); E. Moggia (Nephrology, Santa Croce Hospital, Cuneo, Italy); J. Ballarin (Nephrology, Fundacion Puigvert, Barcelona, Spain); S. Di Giulio (Nephrology, San Camillo Forlanini Hospital, Rome, Italy); F. Pugliese, I. Serriello (Nephrology, PoliclinicoUmberto I University Hospital, Rome, Italy); Y. Caliskan, M. Sever, I.Kilicaslan (Nephrology, Internal Medicine, Istanbul Faculty of Medicine, Istanbul, Turkey); F. Locatelli, L. Del Vecchio (Nephrology, A. Manzoni Hospital, Lecco, Italy); J.F.M. Wetzels, H. Peters (Nephrology and Pathology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands); U. Berg (Pediatrics, Department of Clinical Science, Intervention and Technology, Huddinge, Sweden); F. Carvalho, A.C. da Costa Ferreira (Nephrology, Hospital de Curry Cabral, Lisbon, Portugal); M. Maggio (Nephrology, Hospital Maggiore di Lodi, Lodi, Italy); A. Wiecek (Nephrology, Endocrinology and Metabolic Diseases, Silesian University of Medicine, Katowice, Poland); M. Ots-Rosenberg (Nephrology, Tartu University Clinics, Tartu, Estonia); R. Magistroni (Nephrology, Policlinic of Modena and Reggio Emilia; Modena, Italy); R. Topaloglu, Y. Bilginer (Pediatric Nephrology and Rheumatology, Hacettepe University, Ankara, Turkey); M. D'Amico (Nephrology, S. Anna Hospital, Como, Italy); M. Stangou (Nephrology, Hippokration General Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece); F. Giacchino (Nephrology, Ivrea Hospital, Ivrea, Italy); D. Goumenos, P. Kalliakmani, M. Gerolymos (Nephrology, University Hospital of Patras, Patras, Greece); K. Galesic (Nephrology, University Hospital Dubrava, Zagreb, Croatia); C. Geddes (Renal Unit, Western Infirmary Glasgow, Glasgow, United Kingdom); K. Siamopoulos, O. Balafa (Nephrology, Medical School University of Ioanina, Ioannina, Greece); M. Galliani (Nephrology, S. Pertini Hospital, Rome, Italy); P. Stratta, M. Quaglia (Nephrology, Maggiore della Carità Hospital, Piemonte Orientale University, Novara, Italy); R. Bergia, R. Cravero (Nephrology, Degli Infermi Hospital, Biella, Italy); M. Salvadori, L. Cirami (Nephrology, Careggi Hospital, Florence, Italy); B. Fellstrom, H. Kloster Smerud (Renal Department, University of Uppsala, Uppsala, Sweden); F. Ferrario, T. Stellato (Nephropathology, San Gerardo Hospital Monza, Italy); J. Egido, C. Martin (Nephrology, Fundacion Jimenez Diaz, Madrid, Spain); J. Floege, F. Eitner (Nephrology and Immunology, Medizinische Klinik II, University of Aachen, Aachen, Germany); A. Lupo, P. Bernich (Nephrology, University of Verona, Verona, Italy); P. Menè (Nephrology, S. Andrea Hospital, Rome, Italy); M. Morosetti (Nephrology, Grassi Hospital, Ostia, Italy); C. van Kooten, T. Rabelink, M.E.J. Reinders (Nephrology, Leiden University Medical Centre, Leiden, The Netherlands); J.M. Boria Grinyo (Nephrology, Hospital Bellvitge, Barcelona, Spain); S. Cusinato, L. Benozzi (Nephrology, Borgomanero Hospital, Borgomanero, Italy); S. Savoldi, C. Licata (Nephrology, Civile Hospital, Ciriè, Italy); M. Mizerska-Wasiak (Pediatrics, Medical University of Warsaw, Warsaw, Poland); G. Martina, A. Messuerotti (Nephrology, Chivasso Hospital, Chivasso, Italy); A. Dal Canton, C. Esposito, C. Migotto (Nephrology Units, S. Matteo Hospital and Maugeri Foundation, Pavia, Italy); G. Triolo, F. Mariano (Nephrology CTO, Turin, Italy); C. Pozzi (Nephrology, Bassini Hospital, Cinisello Balsamo, Italy); R. Boero (Nephrology, Martini Hospital, Turin, Italy). The VALIGA pathology investigators include S. Bellur (Oxford, UK), G. Mazzucco (Turin, Italy); C. Giannakakis (Rome, Italy); E. Honsova (Prague, Czech Republic); B. Sundelin (Stockholm, Sweden); A.M. Di Palma (Foggia-Bari, Italy); F. Ferrario (Monza, Italy); E. Gutiérrez (Madrid, Spain); A.M. Asunis (Cagliari, Italy); J. Barratt (Leicester, United Kingdom); R. Tardanico (Brescia, Italy); A. Perkowska-Ptasinska (Warsaw, Poland); J. Arce Terroba (Barcelona, Spain); M. Fortunato (Cuneo, Italy); A. Pantzaki (Thessaloniki, Greece); Y. Ozluk (Istanbul, Turkey); E. Steenbergen (Nijmegen, The Netherlands); M. Soderberg (Huddinge, Sweden); Z. Riispere (Tartu, Estonia); L. Furci (Modena, Italy); D. Orhan (Ankara, Turkey); D. Kipgen (Glasgow, United Kingdom); D. Casartelli (Lecco, Italy); D. Galesic Ljubanovic (Zagreb, Croatia); H Gakiopoulou (Athens, Greece), E. Bertoni (Florence, Italy); P. Cannata Ortiz (Madrid, Spain); H. Karkoszka (Katowice, Poland), H.J. Groene (Heidelberg, Germany); A. Stoppacciaro (Rome, Italy); I. Bajema, J. Bruijn (Leiden, The Netherlands); X. Fulladosa Oliveras (Barcelona, Spain); J. Maldyk (Warsaw, Poland); and E. Ioachim (Ioannina, Greece).

THE OXFORD DERIVATION AND NORTH AMERICAN VALIDATION STUDIES

N. Bavbek, T. Cook, S. Troyanov, C. Alpers, A. Amore, J. Barratt, F. Berthoux, S. Bonsib, J. Bruijn, V. D'Agati, G. D'Amico, S. Emancipator, F. Emmal, F. Ferrario, F. Fervenza, S. Florquin, A. Fogo, C. Geddes, H. Groene, M. Haas, P. Hill, R. Hogg, S. Hsu, T. Hunley, M. Hladunewich, C. Jennette, K. Joh, B. Julian, T. Kawamura, F. Lai, C. Leung, L. Li, P. Li, Z. Liu, A. Massat, B. Mackinnon, S. Mezzano, F. Schena, Y. Tomino, P. Walker, H. Wang, J. Weening, N. Yoshikawa, and H. Zhang.

SUPPLEMENTARY MATERIAL

Table S1. Description of the cohort based on RASB exposure within 6 months prior to biopsy. Data presented as median (IQR) or count (percentage).

Table S2. Description of the cohort based on exposure to immunosuppression after biopsy. Data presented as median (IQR) or count (percentage). N/A, not applicable.

Table S3. Description of the cohort based on M0 versus M1 status. Data presented as median (IQR) or count (percentage).

Figure S1. Median proteinuria values over time in subgroups based on M1 versus M0 mesangial scores. Proteinuria both at biopsy and over time was lower in those with M0.

Supplementary material is linked to the online version of the paper at www.kidney-international.org

REFERENCES

1. Cattran DC, Coppo R, Cook HT, et al. The Oxford classification of IgA nephropathy: rationale, clinicopathological correlations, and classification. *Kidney Int*. 2009;76:534–545.
2. Coppo R, Troyanov S, Bellur S, et al. Validation of the Oxford classification of IgA nephropathy in cohorts with different presentations and treatments. *Kidney Int*. 2014;86:828–836.
3. Herzenberg AM, Fogo AB, Reich HN, et al. Validation of the Oxford classification of IgA nephropathy. *Kidney Int*. 2011;80:310–317.
4. Roberts IS, Cook HT, Troyanov S, et al. The Oxford classification of IgA nephropathy: pathology definitions, correlations, and reproducibility. *Kidney Int*. 2009;76:546–556.
5. Lv J, Shi S, Xu D, et al. Evaluation of the Oxford Classification of IgA nephropathy: a systematic review and meta-analysis. *Am J Kidney Dis*. 2013;62:891–899.
6. Reich HN, Troyanov S, Scholey JW, et al. Remission of proteinuria improves prognosis in IgA nephropathy. *J Am Soc Nephrol*. 2007;18:3177–3183.
7. Tanaka S, Ninomiya T, Katafuchi R, et al. Development and validation of a prediction rule using the Oxford classification in IgA nephropathy. *Clin J Am Soc Nephrol*. 2013;8:2082–2090.
8. Le W, Liang S, Hu Y, et al. Long-term renal survival and related risk factors in patients with IgA nephropathy: results from a cohort of 1155 cases in a Chinese adult population. *Nephrol Dial Transplant*. 2012;27:1479–1485.
9. Barbour SJ, Reich HN. Risk stratification of patients with IgA nephropathy. *Am J Kidney Dis*. 2012;59:865–873.
10. Bartosik LP, Lajoie G, Sugar L, et al. Predicting progression in IgA nephropathy. *Am J Kidney Dis*. 2001;38:728–735.
11. Mackinnon B, Fraser EP, Cattran DC, et al. Validation of the Toronto formula to predict progression in IgA nephropathy. *Nephron Clin Pract*. 2008;109:c148–c153.
12. Kidney Disease: Improving Global Outcomes (KDIGO) Glomerulonephritis Work Group. KDIGO Clinical Practice Guideline for Glomerulonephritis. *Kidney Int Suppl*. 2012;2:139–274.
13. Goto M, Kawamura T, Wakai K, et al. Risk stratification for progression of IgA nephropathy using a decision tree induction algorithm. *Nephrol Dial Transplant*. 2009;24:1242–1247.
14. Goto M, Wakai K, Kawamura T, et al. A scoring system to predict renal outcome in IgA nephropathy: a nationwide 10-year prospective cohort study. *Nephrol Dial Transplant*. 2009;24:3068–3074.

15. Wakai K, Kawamura T, Endoh M, et al. A scoring system to predict renal outcome in IgA nephropathy: from a nationwide prospective study. *Nephrol Dial Transplant*. 2006;21:2800–2808.
16. Berthoux F, Mohey H, Laurent B, et al. Predicting the risk for dialysis or death in IgA nephropathy. *J Am Soc Nephrol*. 2011;22:752–761.
17. Walsh M, Sar A, Lee D, et al. Histopathologic features aid in predicting risk for progression of IgA nephropathy. *Clin J Am Soc Nephrol*. 2010;5:425–430.
18. Cherney DZ, Zinman B, Kennedy CR, et al. Long-term hemodynamic and molecular effects persist after discontinued renin-angiotensin system blockade in patients with type 1 diabetes mellitus. *Kidney Int*. 2013;84:1246–1253.
19. Reiser J, Mundel P. Dual effects of RAS blockade on blood pressure and podocyte function. *Curr Hypertens Rep*. 2007;9:403–408.
20. Ruggenenti P, Perna A, Gherardi G, et al. Chronic proteinuric nephropathies: outcomes and response to treatment in a prospective cohort of 352 patients with different patterns of renal injury. *Am J Kidney Dis*. 2000;35:1155–1165.
21. Praga M, Gutierrez E, Gonzalez E, et al. Treatment of IgA nephropathy with ACE inhibitors: a randomized and controlled trial. *J Am Soc Nephrol*. 2003;14:1578–1583.
22. Li PK, Leung CB, Chow KM, et al. Hong Kong study using valsartan in IgA nephropathy (HKVIN): a double-blind, randomized, placebo-controlled study. *Am J Kidney Dis*. 2006;47:751–760.
23. Manno C, Torres DD, Rossini M, et al. Randomized controlled clinical trial of corticosteroids plus ACE-inhibitors with long-term follow-up in proteinuric IgA nephropathy. *Nephrol Dial Transplant*. 2009;24:3694–3701.
24. Lv J, Zhang H, Chen Y, et al. Combination therapy of prednisone and ACE inhibitor vs. ACE-inhibitor therapy alone in patients with IgA nephropathy: a randomized controlled trial. *Am J Kidney Dis*. 2009;53:26–32.
25. Pozzi C, Andrulli S, Del Vecchio L, et al. Corticosteroid effectiveness in IgA nephropathy: long-term results of a randomized, controlled trial. *J Am Soc Nephrol*. 2004;15:157–163.
26. Tesar V, Troyanov S, Bellur S, et al. Corticosteroids in IgA nephropathy: a retrospective analysis from the VALIGA Study. *J Am Soc Nephrol*. 2015;26:2248–2258.
27. Alamartine E, Sauron C, Laurent B, et al. The use of the oxford classification of IgA nephropathy to predict renal survival. *Clin J Am Soc Nephrol*. 2011;6:2384–2388.
28. Shi SF, Wang SX, Jiang L, et al. Pathologic predictors of renal outcome and therapeutic efficacy in IgA nephropathy: validation of the oxford classification. *Clin J Am Soc Nephrol*. 2011;6:2175–2184.
29. Katafuchi R, Ninomiya T, Nagata M, et al. Validation study of oxford classification of IgA nephropathy: the significance of extracapillary proliferation. *Clin J Am Soc Nephrol*. 2011;6:2806–2813.
30. Barbour S, Beaulieu M, Gill J, et al. The need for improved uptake of the KDIGO glomerulonephritis guidelines into clinical practice in Canada: a survey of nephrologists. *Clin Kidney J*. 2014;7:538–545.
31. Levey AS, Bosch JP, Lewis JB, et al. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med*. 1999;130:461–470.
32. Chambless LE, Cummiskey CP, Cui G. Several methods to assess improvement in risk prediction models: extension to survival analysis. *Stat Med*. 2011;30:22–38.
33. Pencina MJ, D'Agostino RB, Pencina KM, et al. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol*. 2012;176:473–481.
34. Pencina MJ, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27:157–172, discussion 207-12.
35. Amrhein J. *Introduction to frailty models. Proceedings of the SAS Global Forum 2014 Conference*. Available at: http://support.sas.com/resources/papers/proceedings14/1492-2014.pdf. Washington, DC: SAS Institute Inc; 2014:1492–2014 pp.
36. Kleinbaum D, Klein M. *Survival Analysis*. 2nd ed. New York: Springer Science; 2005.