Validation of the Oxford classification of IgA nephropathy

Andrew M. Herzenberg[1,6,7], Agnes B. Fogo[2,6], Heather N. Reich[1,6], Stéphan Troyanov[3,6], Nuket Bavbek[2], Alfonso E. Massat[4], Tracy E. Hunley[2], Michelle A. Hladunewich[2], Bruce A. Julian[5], Fernando C. Fervenza[4] and Daniel C. Cattran[2]

[1]Division of Nephrology, Toronto General Hospital, University of Toronto, Ontario, Canada; [2]Department of Pathology, Vanderbilt University, Nashville, Tennessee, USA; [3]Department of Nephrology, Hôpital du Sacré-Coeur de Montréal, University of Montreal, Montreal, Quebec, Canada; [4]Division of Nephrology and Hypertension, Mayo Clinic, Rochester, Minnesota, USA and [5]Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA

The Oxford classification of IgA nephropathy (IgAN) identified four pathological elements that were of prognostic value and additive to known clinical and laboratory variables in predicting patient outcome. These features are segmental glomerulosclerosis/adhesion, mesangial hypercellularity, endocapillary proliferation, and tubular atrophy/interstitial fibrosis. Here, we tested the Oxford results using an independent cohort of 187 adults and children with IgAN from 4 centers in North America by comparing the performance of the logistic regression model and the predictive value of each of the four lesions in both data sets. The cohorts had similar clinical and histological findings, presentations, and clinicopathological correlations. During follow-up, however, the North American cohort received more immunosuppressive and antihypertensive therapies. Identifying patients with a rapid decline in the rate of renal function using the logistic model from the original study in the validation data set was good (*c*-statistic 0.75), although less precise than in the original study (0.82). Individually, each pathological variable offered the same predictive value in both cohorts except mesangial hypercellularity, which was a weaker predictor. Thus, this North American cohort validated the Oxford IgAN classification and supports its utilization. Further studies are needed to determine the relationship to the impact of treatment and to define the value of the mesangial hypercellularity score.

Kidney International (2011) **80,** 310–317; doi:10.1038/ki.2011.126; published online 4 May 2011

KEYWORDS: clinical epidemiology; glomerulonephritis; IgA nephropathy; pathology; statistics

Correspondence: *Daniel C. Cattran, Department of Pathology, University Health Network, NCSB 11-1256 585 University Avenue, Toronto General Hospital, Toronto, Ontario, Canada M5G 2N2.*
E-mail: daniel.cattran@uhn.on.ca

[6]These authors contributed equally to this work.
[7]This article is dedicated to the memory of Dr Andrew M. Herzenberg who was a major contributor to this work and died in a tragic accident before its completion.

Received 19 December 2010; revised 12 March 2011; accepted 15 March 2011; published online 4 May 2011

IgA nephropathy (IgAN) is the most common glomerular disease worldwide. The outcome of patients with IgAN varies greatly.[1–4] Clinical variables available to physicians treating patients with IgAN account for less than half of the variability in outcome.[1] Assessing prognosis is challenging and has contributed to the lack of consensus regarding the management of these patients.

The additive value of pathology features to the known clinical parameters of proteinuria, serum creatinine, and blood pressure in predicting long-term outcome has long been debated.[5–9] A recent international effort made by a group of expert nephrologists and nephropathologists using clinical, laboratory, and pathology features identified a set of distinct pathological variables that demonstrated an independent prognostic value in patients with IgAN, known as the Oxford classification.[10,11] In 265 cases from centers across continents, four pathology features were found to be of independent value in terms of predicting the outcome of renal function: mesangial hypercellularity score (M; M0 $\leqslant$ 0.5, M1 $>$ 0.5), the presence of endocapillary proliferation (E; E0: absent, E1: present), segmental glomerulosclerosis/adhesion (S; S0: absent, S1: present), and severity of tubular atrophy/interstitial fibrosis (T; T0 $\leqslant$ 25%, T1: 26–50%, T2 $>$ 50%). The MEST score remained statistically significant for prediction of the outcome even after taking into account the clinical indicators available both at the time of biopsy and during the observation period.

These findings have the potential to greatly affect nephrologists' approach to IgAN by improving the capacity to identify, at the time of biopsy, high-risk-of-progression patients in contrast to other indicators that require an observation period before helping with prognosis.[12] Definitive pathological features would significantly modify not only the approach by the physician to the individual patient but

also the framework for designing and interpreting therapeutic trials. However, before widespread adoption of this new Oxford classification system, validation of its findings is essential.[13–15] Validation is often neglected in medical prognostic models and has been rarely performed in kidney disease despite its obvious importance.[15] Our objective was to assemble an independent cohort of IgAN patients to test the validity of the Oxford pathology classification.

RESULTS

A total of 187 patients were recruited from 4 different centers in North America (NA). In all, 44 patients (24%) were younger than 18 years at the time of biopsy, a proportion similar to that in the Oxford data set (Table 1). The median year of biopsy was 2002 for the NA validation cohort versus 1997 in the Oxford derivation data set ($P<0.001$).

Clinical characteristics

The clinical characteristics of the validation cohort are presented in Table 2. Compared with the Oxford derivation cohort, there were no significant differences in age, estimated glomerular filtration rate (eGFR), proteinuria, and mean arterial pressure (MAP) at the time of biopsy. The NA validation cohort contained a greater proportion of female subjects (42 compared with 28%, $P=0.001$).

A similar proportion of patients in both cohorts had received immunosuppression at the time of renal biopsy (10 NA compared with 14% Oxford, $P>0.1$). At the time of biopsy, 58% of patients in the NA validation data set were on an antihypertensive medication versus 31% in the Oxford derivation data set ($P=0.006$). Moreover, at entry, a higher percentage of patients was receiving renin angiotensin system blockade (RASB) in the NA cohort (31 versus 20%, $P=0.007$).

The NA validation cohort also received more therapy during the follow-up period with regard to immunosuppression (41% NA versus 29% Oxford, $P=0.02$), RASB (87 versus 74%, $P=0.001$), and number of anti-hypertensive medications (median of 1.1 versus 0.9, $P<0.001$). Perhaps related to these differences, the NA validation cohort achieved a lower-average MAP than did the Oxford cohort: 92 ± 11 mm Hg compared with 95 ± 10 mm Hg ($P=0.002$). Time-average proteinuria was not different: 1.0 (0.6–1.9) g/24 h compared with 1.1 (0.6–2.0) g/24 h. The outcomes were similar with a rate of renal-function decline in the NA set of -2.6 ± 6.8 ml/min per 1.73 m^2 per year compared with -3.5 ± 8.5 ml/min per 1.73 m^2 per year in the Oxford set ($P>0.1$). Kidney survival from a combined event was not statistically different.

Pathology characteristics

The agreements between the two pathologists (ABF and AMH) are shown in Table 3. The intraclass correlation coefficients (ICCs) were higher for each of the pathology variables assessed in comparison with the initial study, with the exception of the mesangial score.

Table 1 | Age distribution and geographical origin of the NA validation cohort

	Total	Adults	Children at biopsy
n	187	143	44
Vanderbilt University	55	19	36
Mayo Clinic	38	32	6
University of Alabama at Birmingham	22	21	1
University of Toronto	72	71	1

Abbreviation: NA, North America.

Table 3 | Intraclass correlation coefficients comparing the derivation Oxford and the validation North American cohorts

	Oxford cohort	NA cohort
Mesangial hypercellularity score	0.64	0.37
Percentage of total glomeruli showing segmental sclerosis	0.46	0.94
Percentage of total glomeruli showing adhesions	0.20	0.76
Percentage of total glomeruli showing endocapillary hypercellularity	0.57	0.87
Percentage of cortex showing tubular atrophy	0.79	0.98
Percentage of cortex showing interstitial fibrosis	0.78	0.98

Abbreviation: NA, North America.

Table 2 | NA validation cohort: clinical characteristics at the time of biopsy and follow-up

At the time of biopsy		Follow-up	
Age (years)	34 (18–45)	Duration of follow-up (months)	53 (36–77)
Female	42%	MAP (mm Hg)	92 ± 11
Pediatric at biopsy (<18 years)	24%	No. antihypertensive medications	1.1 (0.8–1.9)
Ethnicity (Caucasian/African American/Asia/Others)	69, 5, 20, 4%	Treated with RASB	87%
MAP (mm Hg)	96 ± 14	Proteinuria (g/24 h)	1.0 (0.6–1.9)
Proteinuria (g/24 h)	1.7 (1.0–2.9)	Immunosuppression (any)	41%
Previous immunosuppression	10%	Prednisone/other	37%/24%
eGFR (ml/min per 1.73 m^2)	82 ± 37	Fish oil during follow-up	61%
Previous RASB	31%	Slope (ml/min per 1.73 m^2/years)	-2.6 ± 6.8
		50% decline in eGFR	14%
		End-stage renal disease	11%

Abbreviations: eGFR, estimated glomerular filtration rate; MAP, mean arterial pressure; NA, North America; RASB, renin-angiotensin-system blockade.
Variables were expressed as mean ± s.d., median with interquartile range, or percentage.

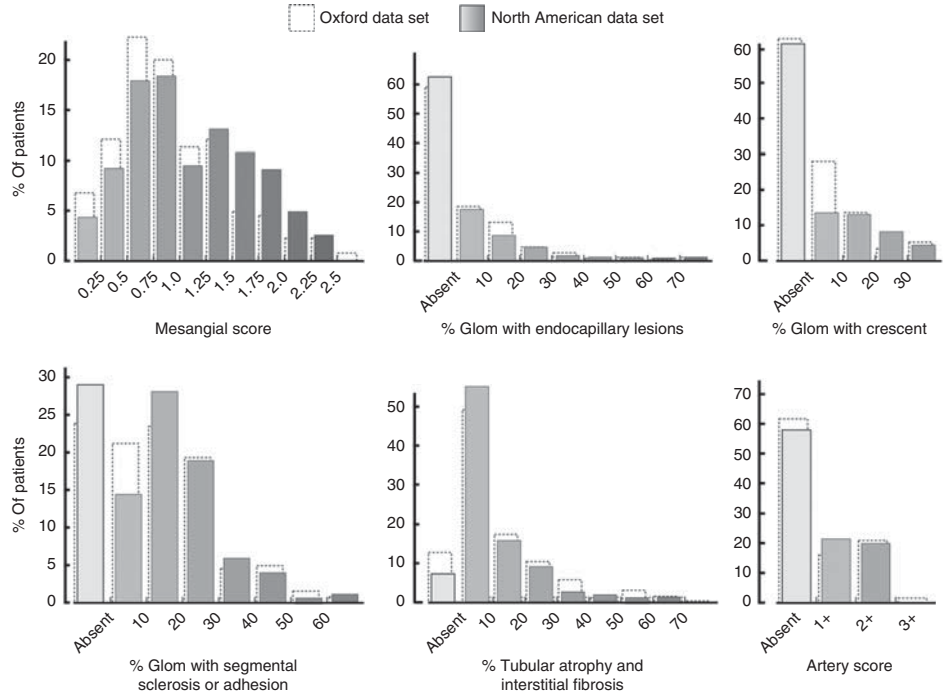

Figure 1 | Frequency of pathological features. Percentage of patients with each pathological feature. The six pathological features illustrated were those with sufficient reproducibility from the Oxford study. Glom, glomeruli.

The prevalence of lesions in the validation compared with the derivation data set is shown in Figure 1. Overall, the distributions of the pathology findings were similar in both cohorts with the exception of the mesangial score, with a median score of 1.07 (0.73–1.59) in the NA validation data set compared with 0.89 (0.60–1.30) in the Oxford derivation data set ($P<0.001$). All other comparisons were not statistically different.

Clinicopathological correlations at the time of biopsy

The NA validation study found similar correlations as in the Oxford set between the pathology findings and the clinical assessment (eGFR, MAP, proteinuria) at the time of biopsy. Exceptions were the lack of a statistically significant association between proteinuria at the time of biopsy with either the M- or the T-score, as shown in Figure 2.

Impact of pathology findings on immunosuppressive therapy during follow-up

The administration of immunosuppression seemed to be strongly related to the pathology findings in the NA cohort, similar to that seen in the Oxford derivation data set (Figure 3). In both cohorts, patients with endocapillary and extracapillary lesions were exposed to greater immunosuppression than were patients without such findings, whereas the presence of focal glomerulosclerosis/adhesion, severity of tubulointerstitial lesions, and vascular lesions were not correlated with treatments (data not shown). This relationship applied somewhat differently in the two cohorts with regard to the mesangial hypercellularity score (Figure 3).

Applicability of the Oxford model prediction rule to the validation cohort

To validate the results in our NA data set, we derived for every patient in each cohort the predicted probability of a rapid rate of renal disease progression based on the β-coefficients from the logistic regression model of the Oxford derivation data set. The model included M-, S-, and T-scores in addition to the known clinical predictors, initial eGFR, follow-up time-average MAP, and time-average proteinuria. The E-score was assessed in terms of interaction with immunosuppressive therapy as was done in the derivation Oxford data set (originally, the important interaction between the E lesion and clinical severity/treatment allocation precluded its use in the multivariate model; see below). To facilitate visualization, the calculated predicted probabilities were used to create quintiles of predicted risk. Figure 4 illustrates for each quintile the observed rate of renal-function decline (with 95% confidence interval (95% CI) of the mean) in each data set. Overall, the model performed well in both cohorts validating the conclusions from the Oxford study.

Comparisons of individual pathology predictors

We also derived a new logistic regression using the same variables from the original study. Figure 5 compares both models and illustrates a striking resemblance for the S- and

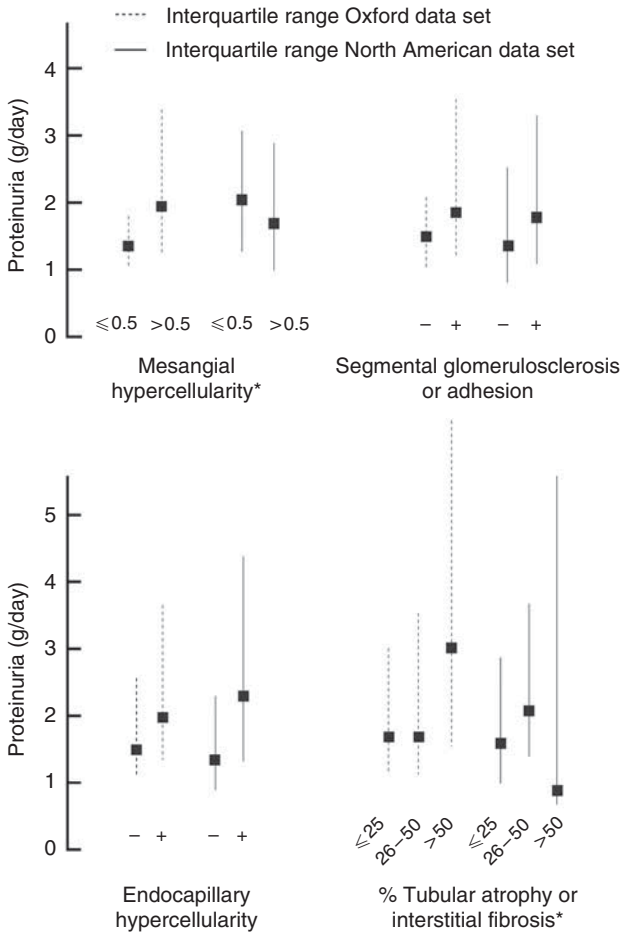

Figure 2 | Correlations between pathological features and proteinuria at the time of biopsy. Median with interquartile range. *The relationship between proteinuria at the time of biopsy and pathological findings is not seen for either the mesangial score or the tubulointerstitial score for the North American data set, in contrast to the Oxford study.

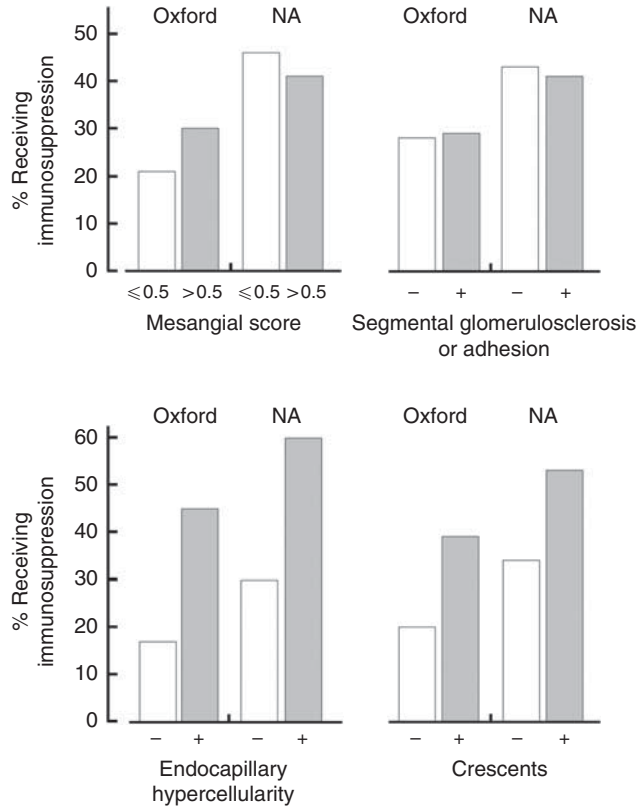

Figure 3 | Immunosuppressive therapy received during follow-up in relation to pathological features. NA, North America.

T-scores in both data sets, but the M-score performed differently. To further investigate the predictive value of the M-score, we performed a survival analysis and did not find a statistically significant reduction of a combined outcome (hazard ratio of 0.62, 95% CI: 0.14–2.74, $P>0.1$ in the validation cohort compared with the Oxford derivation hazard ratio of 0.07, 95% CI: 0.01–0.52, $P=0.009$).

Interactions between the E lesion and immunosuppression from both cohorts are illustrated in Figure 6. The relationship between E lesions and the rate of renal-function decline is evident in those who did not receive immunosuppression and is lost in those who did. When both derivation and validation data sets with E lesions are combined, patients who never received immunosuppression during follow-up had a rate of renal-function decline of -5.0 ± 9.6 ml/min per 1.73 m^2 compared with -2.3 ± 5.4 ml/min per 1.73 m^2 in those who were treated ($P=0.02$). This difference remained significant after adjustment for eGFR, follow-up time-average proteinuria, and MAP.

Added value of pathological variables on the predictive rule

The discriminative strength of the logistic model was high in both cohorts with an area under the receiver-operating characteristic curve (or *c*-statistic) of 0.75 (95% CI: 0.68–0.82) in the NA validation cohort (0.70 using clinical variables only; 0.63 using pathology variables only). The *c*-statistic was originally 0.82 (95% CI: 0.77–0.87) in the Oxford data set (0.78 using clinical variables only; 0.70 using pathology variables only) (Figure 7).

DISCUSSION

A recent international effort identified a set of distinct pathological variables with prognostic value, independent of all clinical and laboratory parameters in patients with IgAN.[10,11] Although these findings are currently insufficient to recommend specific treatment recommendations, they offer the possibility of a targeted approach to management based on an improved and semi-quantified risk assessment. Given the potential relevance of this proposed classification and its distinctive nature compared with standard classifications, we believed that validation was essential before recommending its implementation on a global scale. Three of the four pathological variables identified in the Oxford derivation cohort predicted a rapid rate of renal-function decline in the NA validation data set, supporting the original findings. Additional support of its validity is the closeness of

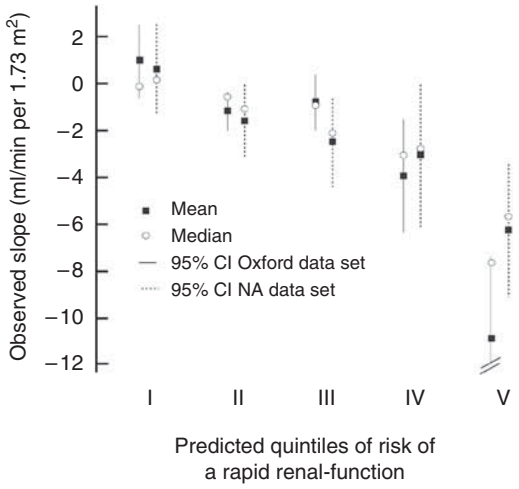

Figure 4 | Observed rate of renal-function decline according to predicted quintiles of risk derived from the Oxford multivariate logistic model. In each patient from both cohorts, the predicted probability of a rapid rate of renal-disease progression was calculated based on the β-coefficients from the logistic regression model of the Oxford study. Patients were categorized into quintiles using these predicted probabilities, and the actual observed rate of renal-function decline for each quintile is illustrated with 95% confidence intervals (95% CIs).

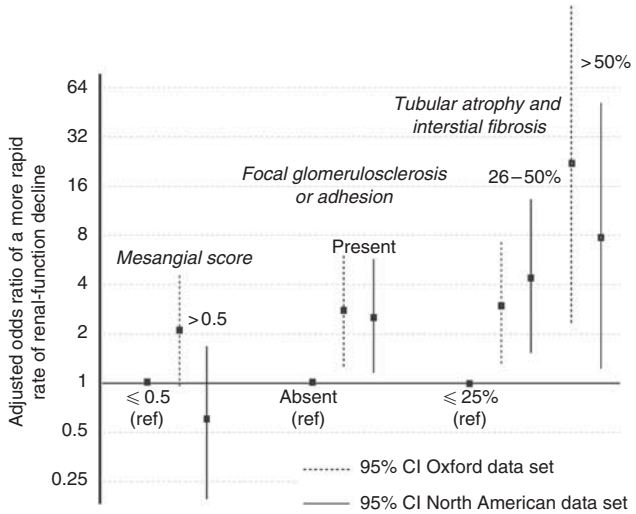

Figure 5 | Association between pathological variables and more rapid rate of renal-function decline (defined categorically). Multivariate logistic model using these three pathological features + initial eGFR and follow-up MAP and proteinuria. Adjusted odds ratios are presented with 95% confidence intervals (95% CIs) in reference (ref) to an absent lesion. eGFR, estimated glomerular filtration rate; MAP, mean arterial pressure.

fit across the quintiles of risk when the observed-to-predicted rate of renal-function decline is compared in both cohorts (Figure 4). The exception was the mesangial hypercellularity score (M-score), with a predictive value that underperformed in the NA data set. Although the added value of pathology to longitudinal clinical assessment may seem moderate

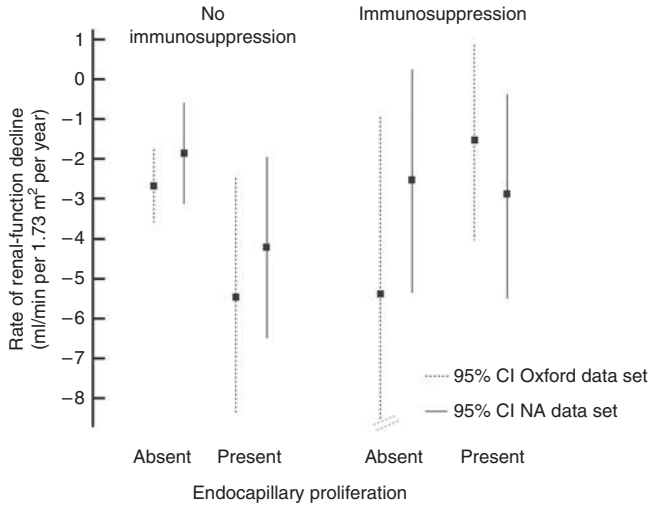

Figure 6 | Interaction between immunosuppression and endocapillary proliferation on the rate of renal-function decline. CI, confidence interval; NA, North America.

(Figure 7), biopsy findings have the advantage of immediate long-term predictive value.

The M-score deserves additional comment. Not only was its long-term predictive value reduced but also the cross-sectional correlation with proteinuria at the time of biopsy was absent. Possible explanations, other than a spurious relationship identified in the derivation study, are multiple. The lower ICC observed for this variable in this study may have resulted in error in measurement, leading to a non-significant finding. Alternative explanations include the different therapeutic strategies in the two cohorts, such as the greater anti-hypertensive treatment and RASB in the NA cohort, both during follow-up and before renal biopsy. Certainly others studies have shown that the effect of the renin-angiotensin system is associated with mesangial cell inhibition and prevention of renal-function decline in both human and experimental glomerulonephritis.[16,17] As in the derivation data set, we found an interaction between the E-score and the use of immunosuppressive therapy and the rate of renal-function decline in the validation cohort. In the original derivation study, because of the small number of subjects, we could not determine the independent predictive value of the E-score, given this important potential confounder of treatment effect. However, by combining the two cohorts, we were able to provide further evidence that supports the independent predictive value of endocapillary proliferation (Figure 6). Our findings strongly corroborate the value of the inclusion of the E-score in the Oxford classification and address previous questions in this regard.[13] These results also illustrate the potential role of renal biopsy findings in the selection of IgA patients requiring aggressive therapy. Although a recent study did not show a predictive value of endocapillary lesions, the prevalence of such findings was not reported, and potential interaction with immunosuppression was not assessed.[18]

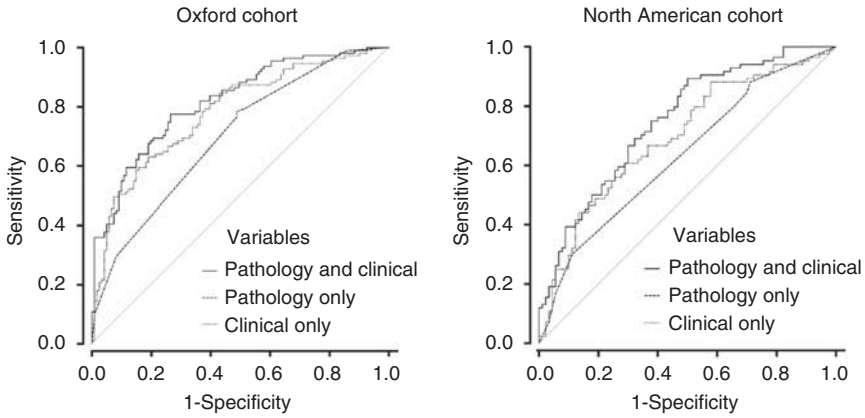

Figure 7 | Added value of pathology variables in predicting a more rapid rate of renal-function decline. Pathology variables include the presence of a mesangial score >0.5, the presence of segmental glomerulosclerosis or adhesion, and the severity of tubulointerstitial disease. Clinical variables include eGFR at onset, time-average MAP, and time-average proteinuria. eGFR, estimated glomerular filtration rate; MAP, mean arterial pressure.

Prediction models estimate an outcome using a mathematical equation combining multiple variables. Three steps guide their clinical utility: (1) development of a prediction equation, (2) validation of the accuracy and generalizability of findings, and (3) studying the clinical impact on physician behavior and patient outcome after implementation in daily practice. Although articles proposing predictive models based on risk factors are numerous, very few have been validated or have addressed how their implementation has affected patient care.[19] This failure to validate has potential serious consequences, because premature implementation of an imprecise model or inappropriate generalization of conclusions can cause harm to patients, and can misdirect treatment algorithms and clinical trials.

True validation of any type is complex and has its own set of rules often not understood or followed. There are also various methodologies to test validity.[20] Internal validation can be performed by splitting a study cohort into two groups and using one for validation. More sophisticated methods repeat this procedure (bootstrap) many times on random data subsets to identify a more robust model.[21] External validation is a superior method but does require the use of a new database, ideally from different centers and scientists.[15] A validation study classically requires an adequate sample of a different but related patient population as that found in the original derivation cohort. We followed our original Oxford inclusion and exclusion criteria to maintain this element of validation in our study. We recognize that, as discussed in the original paper, the elimination of the subsets of the IgAN patient population displaying extremes of the histological and/or clinical features, that is, those with very rapid progression or with no significant proteinuria may mean that our results cannot be applied to these subsets. It may also not apply to other ethnicities, such as African Americans, Hispanics, and Asians.

There are two additional elements that can be examined to strengthen the validation process; these include studying the predictive value of a model across time and across geographic boundaries. There was a 5-year difference between the time of biopsy in the Oxford derivation cohort (1997) versus the North American validation cohort (2002), thus fulfilling this criterion. It is likely that this resulted in the previously discussed differences in terms of blood pressure medications, RASB usage, and immunosuppression; however, despite these changes, the predictive capacity of the Oxford classification remained largely intact. We also fulfill the different origins of the data elements as our NA cohort is geographically different from the derivation data set being restricted to a North American base versus a more varied origin in the Oxford derivation cohort.

In general, any classification needs to be reproducible. This capacity for replication of the interpretation of the pathology findings to span a wide range of nephropathologists, from academic specialists to those in the community, has been a notable limitation in many previous studies that have attempted to quantify the added value of pathology in renal diseases.[22–25] In all, 10 of the 24 initial variables from the original study were judged to be insufficiently reproducible for further analysis; after assessing co-linearity, only 6 were sufficiently independent from one another to warrant predictive testing.[11] In our NA cohort, the ICCs were very good and higher than those in the Oxford derivation data set. This finding supports its reproducibility. The agreement based on the cutoffs proposed by the Oxford classification was also high. Part of this is because of the simplicity of the chosen variables. Two, namely the presence of endocapillary proliferation or segmental glomerulosclerosis, are based on a simple 'present' or 'absent' assessment. Thus, a single glomerulus with either of these lesions correlated with a worse prognosis during follow-up. Tubulointerstitial fibrosis, assessed with a semi-quantitative score with the Oxford classification, was also robustly scored in this study. Different morphometric approaches and various staining parameters have been advocated,[26] but the semi-quantitative 'eyeballing'

assessment on a light microscopy slide by an experienced pathologist is a straightforward, reliable, and robust morphological parameter.[27] The ICC for mesangial proliferation was lower than that in the derivation data set. This could have occurred because the scoring process is more time-consuming and complex compared with the other features as it involves counting mesangial cell nuclei in each glomerulus.[11]

In large part, this study validates the Oxford classification in IgAN and provides strong support for a more widespread adoption of this model in routine pathology practice. In addition, the validation process indicates that utilization of this new classification will improve the ability to identify early those patients who are most likely to progress to renal failure and may respond to immunosuppressive treatment. The classification should significantly modify not only our approach to the individual patient but also our framework for designing future clinical trials in IgAN.

MATERIALS AND METHODS

Design

This is a retrospective study of 187 patients with biopsy-proven IgAN followed in 1 of 4 North American centers: the University of Toronto Glomerulonephritis Registry, the University of Alabama at Birmingham, the Vanderbilt University Medical Center in Tennessee, and the Mayo Clinic in Minnesota.

Inclusion criteria and clinical data set

The inclusion criteria as required in validation studies were identical to those of the initial study. Patients of any age with native-kidney biopsy-proven IgAN were included, provided their initial eGFR was $\geqslant 30$ ml/min per 1.73 m^2, proteinuria was $\geqslant 0.5$ g/24 h ($\geqslant 0.5$ g/24 h per 1.73 m^2 in children), follow-up was at least 12 months, and an adequate renal biopsy was available. Patients with isolated hematuria, Henoch-Schoenlein purpura, or co-existing conditions such as diabetes mellitus were excluded.

Demographics collected included gender, ethnicity, and age at the time of biopsy. Children were defined as patients aged <18 years at biopsy. Clinical parameters collected initially (within 3 months of the date of biopsy) and during follow-up included systolic and diastolic blood pressures, weight, height, serum creatinine, and 24-h urine protein or urine protein-to-creatinine ratio. To provide consistency between measurements in adults and children, proteinuria was expressed in g/24 h per 1.73 m^2 in children and g/24 h in adults. Drug exposure was also collected, including immunosuppressive agents, fish oil, and the number and type of antihypertensive medications, including angiotensin-converting enzyme inhibitor and angiotensin receptor blocker. Data verification occurred by communication between the two of the writing committee authors (ST and HNR) and the contributing centers.

Pathology data set

Biopsy adequacy was defined as a minimum of eight glomeruli available for examination by light microscopy. IgAN was confirmed as predominant or codominant immunoglobulin in the mesangial deposits. Two experienced renal pathologists, both members of the original Oxford study, independently scored every feature according to the full Oxford score sheet.[11] The first pathologist scoring a given case chose the most representative section within biopsy slides. This section was marked and independently scored by the second pathologist. We derived an MEST score based on these results. Discordant scores were observed in a minority of cases and they were resolved by a head-to-head meeting between the pathologists (ABF and AMH).

Definitions

The eGFR was calculated using the four-variable Modification of Diet in Renal Disease Study formula in adults and the Schwartz formula in children (using the constant 0.55). End-stage renal disease was defined as eGFR <15 ml/min per 1.73 m^2. A combined event was defined by a 50% reduction in eGFR or end-stage renal disease. MAP was defined as diastolic pressure plus a third of the pulse pressure. For each patient, average MAP and proteinuria were determined for each year of observation. Time-average MAP and proteinuria represent the average of these annual values. Immunosuppressive treatment is reported as intent to treat, regardless of the type or duration of therapy. RASB indicates any exposure to angiotensin-converting enzyme inhibitor, angiotensin receptor blocker, or both.

Statistical analysis

Normally distributed variables were expressed as mean ± s.d. and compared using Student's *t*-test or one-way ANOVA (analysis of variance). Non-parametric variables were expressed as median with interquartile range and compared using either Mann–Whitney or Kruskal–Wallis test. Categorical variables were expressed in percentages and compared using the Pearson χ^2 test.

Reproducibility was assessed for each variable of the extended pathology data set using the ICC, a measure of reproducibility applicable to multiple raters. By convention, an ICC of 0.40–0.59 is moderate inter-rater reliability; 0.60–0.79, substantial; and $\geqslant 0.80$, outstanding.[28]

The rate of renal-function decline (slope) was used as the main outcome. It was determined by fitting a straight line through available data for eGFR using the principle of least squares. This line was plotted and visually examined in each patient. Obvious outliers were censored.

The original study used three multivariate methodologies to assess the independent value of pathology findings: linear regression (slope), logistic regression (slope categorized by slow or rapid rate of renal-function decline based on the median value), and Cox regression (survival from either end-stage renal disease or a 50% decrease in eGFR). However, the slope distribution in this cohort was flatter and more skewed than the normal 'bell-shaped' curve; hence, the assumption of normality could not be met and therefore linear regression analysis was not applied. We had also assessed kidney survival from a 50% decrease in eGFR or ERSD but could not adjust for multiple variables in the NA cohort because the number of these events was too low in comparison with the Oxford cohort. Therefore, we could only use the logistic model to test whether the added value of pathology features remained independent of the clinical assessment.

Validation of the Oxford clinicopathological classification using this new cohort was tested in a series of predefined successive steps:

(1) Using the logistic regression equation from the Oxford data set, we calculated for each patient from both cohorts the predicted probability of a rapid rate of renal-disease progression.
(2) We then categorized these predicted probabilities into five groups with increasing risk (quintiles) for each data set.
(3) We compared the observed rate of renal-function decline among quintiles. A consistent relationship between predicted risk and

observed rate of renal-function decline in both cohorts validates the Oxford logistic model.

(4) To further address differences between the two data sets and to determine whether any of the specific pathological findings differed in the predictive value compared with the original, we performed a new logistic regression on the validation data set using the same variables as in the initial Oxford model and compared graphically the adjusted odds ratio of each pathological variable with those reported in the original classification paper.

(5) Finally, we calculated the area under the receiver-operating characteristic curves (*c*-statistic) from the Oxford and North American logistic regression models.

Survival from a 50% reduction in renal function or end-stage renal disease was performed using the log-rank test.

All *P*-values were two-tailed, and values <0.05 were considered statistically significant. CIs included 95% of predicted values. Analyses were carried out using SPSS software (version 16, SPSS, Chicago, IL).

DISCLOSURE

The authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All the authors declared no competing interests.

ACKNOWLEDGMENTS

HNR is the recipient of a KRESCENT clinician scientist fellowship from the Kidney Foundation of Canada and the Canadian Institute of Health Research. ST's research efforts are supported by the Fonds de la Recherche en Santé du Québec. This study was supported in part by a grant from the Canadian Institutes of Health Research, Net Grant on Genes, Gender and Glomerulonephritis (no. 452773L), the Fonds de la Recherche en Santé du Québec (no. 14395), and the NIH grants DK082753, DK078244, and DK080301. We thank the following coordinators who were of substantial help in collecting the data: from Toronto, N. Ryan and P. Ling; and from Birmingham, C.V. Barker and C. Snyder.

REFERENCES

1. Bartosik LP, Lajoie G, Sugar L *et al.* Predicting progression in IgA nephropathy. *Am J Kidney Dis* 2001; **38**: 728–735.
2. Woo KT, Edmondson RP, Wu AY *et al.* The natural history of IgA nephritis in Singapore. *Clin Nephrol* 1986; **25**: 15–21.
3. Droz D, Kramar A, Nawar T *et al.* Primary IgA nephropathy: prognostic factors. *Contrib Nephrol* 1984; **40**: 202–207.
4. Geddes CC, Rauta V, Gronhagen-Riska C *et al.* A tricontinental view of IgA nephropathy. *Nephrol Dial Transplant* 2003; **18**: 1541–1548.
5. D'Amico G. Natural history of idiopathic IgA nephropathy: role of clinical and histological prognostic factors. *Am J Kidney Dis* 2000; **36**: 227–237.
6. Donadio JV, Bergstralh EJ, Offord KP *et al.* Clinical and histopathologic associations with impaired renal function in IgA nephropathy. Mayo Nephrology Collaborative Group. *Clin Nephrol* 1994; **41**: 65–71.
7. Johnston PA, Brown JS, Braumholtz DA *et al.* Clinico-pathological correlations and long-term follow-up of 253 United Kingdom patients with IgA nephropathy. A report from the MRC Glomerulonephritis Registry. *Q J Med* 1992; **84**: 619–627.
8. Lee SM, Rao VM, Franklin WA *et al.* IgA nephropathy: morphologic predictors of progressive renal disease. *Hum Pathol* 1982; **13**: 314–322.
9. Haas M. Histologic subclassification of IgA nephropathy: a clinicopathologic study of 244 cases. *Am J Kidney Dis* 1997; **29**: 829–842.
10. 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.
11. 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.
12. Reich HN, Troyanov S, Scholey JW *et al.* Remission of proteinuria improves prognosis in IgA nephropathy. *J Am Soc Nephrol* 2007; **18**: 3177–3183.
13. Yamamoto R, Imai E. A novel classification for IgA nephropathy. *Kidney Int* 2009; **76**: 477–480.
14. Bleeker SE, Moll HA, Steyerberg EW *et al.* External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003; **56**: 826–832.
15. Toll DB, Janssen KJ, Vergouwe Y *et al.* Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008; **61**: 1085–1094.
16. Minutolo R, Balletta MM, Catapano F *et al.* Mesangial hypercellularity predicts antiproteinuric response to dual blockade of RAS in primary glomerulonephritis. *Kidney Int* 2006; **70**: 1170–1176.
17. Miyake-Ogawa C, Miyazaki M, Abe K *et al.* Tissue-specific expression of renin-angiotensin system components in IgA nephropathy. *Am J Nephrol* 2005; **25**: 1–12.
18. 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.
19. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006; **144**: 201–209.
20. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000; **19**: 453–473.
21. Schumacher M, Hollander N, Sauerbrei W. Resampling and cross-validation techniques: a tool to reduce bias caused by model building? *Stat Med* 1997; **16**: 2813–2827.
22. Weening JJ, D'Agati VD, Schwartz MM *et al.* The classification of glomerulonephritis in systemic lupus erythematosus revisited. *J Am Soc Nephrol* 2004; **15**: 241–250.
23. Solez K, Colvin RB, Racusen LC *et al.* Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant* 2008; **8**: 753–760.
24. Furness PN, Taub N. Interobserver reproducibility and application of the ISN/RPS classification of lupus nephritis-a UK-wide study. *Am J Surg Pathol* 2006; **30**: 1030–1035.
25. Furness PN, Taub N. International variation in the interpretation of renal transplant biopsies: report of the CERTPAP Project. *Kidney Int* 2001; **60**: 1998–2012.
26. Hunter MG, Hurwitz S, Bellamy CO *et al.* Quantitative morphometry of lupus nephritis: the significance of collagen, tubular space, and inflammatory infiltrate. *Kidney Int* 2005; **67**: 94–102.
27. Mengel M, Chapman JR, Cosio FG *et al.* Protocol biopsies in renal transplantation: insights into patient management and pathogenesis. *Am J Transplant* 2007; **7**: 512–517.
28. Armstrong GD. The intraclass correlation as a measure of interrater reliability of subjective judgments. *Nurs Res* 1981; **30**: 314–315, 320A.