

Urine proteomics for prediction of disease progression in patients with IgA nephropathy

Michael Rudnicki^{1,*}, Justyna Siwy^{2,*}, Ralph Wendt³, Mark Lipphardt⁴, Michael J. Koziol⁴, Dita Maixnerova⁵, Björn Peters^{6,7}, Julia Kerschbaum¹, Johannes Leierer¹, Michaela Neprasova⁵, Mirosław Banasik⁸, Ana Belen Sanz⁹, Maria Vanessa Perez-Gomez ⁹, Alberto Ortiz⁹, Bernd Stegmayr⁷, Vladimír Tesar⁵, Harald Mischak², Joachim Beige ^{3,10} and Heather N. Reich^{11,12}, on behalf of the PERSTIGAN working group[†]

¹Department of Internal Medicine IV, Nephrology and Hypertension, Medical University Innsbruck, Innsbruck, Austria, ²Mosaiques Diagnostics GmbH, Hannover, Germany, ³Division of Nephrology and KfH Renal Unit, Hospital St Georg, Leipzig, Germany, ⁴Department of Nephrology and Rheumatology, University Medical Centre Göttingen, Göttingen, Germany, ⁵Department of Nephrology, 1st School of Medicine and General University Hospital, Charles University, Prague, Czech Republic, ⁶Department of Nephrology, Skaraborg Hospital, Skövde, Sweden, ⁷Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden, ⁸Department of Nephrology and Transplantation Medicine, Wrocław Medical University, Wrocław, Poland, ⁹Research Health Institute, Fundación Jiménez Díaz University, Madrid, Spain, ¹⁰Martin-Luther-University Halle/Wittenberg, Halle/Saale, Germany, ¹¹Department of Medicine, Division of Nephrology, University Health Network, University of Toronto, Toronto, Canada and ¹²Nephrology Research, University of Toronto, Toronto, Ontario, Canada

Correspondence to: Michael Rudnicki; E-mail: michael.rudnicki@i-med.ac.at

*These authors contributed equally to this work.

[†]The members of the PERSTIGAN working group are provided in the Acknowledgements section.

ABSTRACT

Background. Risk of kidney function decline in immunoglobulin A (IgA) nephropathy (IgAN) is significant and may not be predicted by available clinical and histological tools. To serve this unmet need, we aimed at developing a urinary biomarker-based algorithm that predicts rapid disease progression in IgAN, thus enabling a personalized risk stratification.

Methods. In this multicentre study, urine samples were collected in 209 patients with biopsy-proven IgAN. Progression was defined by tertiles of the annual change of estimated glomerular filtration rate (eGFR) during follow-up. Urine samples were analysed using capillary electrophoresis coupled mass spectrometry. The area under the receiver operating characteristic curve (AUC) was used to evaluate the risk prediction models.

Results. Of the 209 patients, 64% were male. Mean age was 42 years, mean eGFR was 63 mL/min/1.73 m² and median proteinuria was 1.2 g/day. We identified 237 urine peptides

showing significant difference in abundance according to the tertile of eGFR change. These included fragments of apolipoprotein C-III, alpha-1 antitrypsin, different collagens, fibrinogen alpha and beta, titin, haemoglobin subunits, sodium/potassium-transporting ATPase subunit gamma, uromodulin, mucin-2, fractalkine, polymeric Ig receptor and insulin. An algorithm based on these protein fragments (IgAN237) showed a significant added value for the prediction of IgAN progression [AUC 0.89; 95% confidence interval (CI) 0.83–0.95], as compared with the clinical parameters (age, gender, proteinuria, eGFR and mean arterial pressure) alone (0.72; 95% CI 0.64–0.81).

Conclusions. A urinary peptide classifier predicts progressive loss of kidney function in patients with IgAN significantly better than clinical parameters alone.

Keywords: biomarker, glomerulonephritis, IgAN, progression, urine proteomics

Prospective clinical study

Urine proteomics for prediction of disease progression in patients with IgA nephropathy

Background



In IgA nephropathy (IgAN), predicting risk of rapid kidney disease progression is challenging



Current risk stratification tools require kidney biopsy; non-invasive markers are lacking

Methods



Multicenter Europe (n=6)/Canada (n=1)



Biopsy-proven IgAN (eGFR >20ml/min/1.73 m²)



IgAN237:

- Urine proteome profiling
- 237 discriminatory peptides



Outcome: eGFR slope

- Annual change in eGFR after baseline
- Divided into tertiles

Results

Group	Baseline eGFR (ml/min/1.73 m ²)	eGFR slope (ml/min/1.73 m ²)	Progression to ESKD
Fastest progression n=70	64.9 (30.8)	-6.4 (95% CI -8.8 to -5.2)	20%
Intermediate progression n=69	62.8 (29.5)	-1.4 (95% CI -1.8 to -1.1)	6%
Slowest progression n=70	60.7 (26.6)	+3.4 (95% CI 2.3 to 4.7)	0%

p=0.842 P<0.001 P<0.001

AUC for fast progression



Clinical parameters alone

0.72

(95% CI 0.64–0.81)



IgAN237 + clinical parameters

0.89

(95% CI 0.83–0.95)

Conclusion

In IgAN, a urinary peptide classifier (IgAN237) along with clinical parameters predicts progressive loss of kidney function better than clinical parameters alone.



Rudnicki, M., et al. NDT (2020) @NDTSocial

INTRODUCTION

Immunoglobulin A (IgA) nephropathy (IgAN) is the most common primary glomerulopathy with an incidence exceeding 2.5/100 000 patients per year [1]. A clinical challenge is the heterogeneity in kidney function decline, with some patients having an indolent course and others having a marked progressive trajectory with more rapid decline in kidney function. Between 15% and 60% of IgAN patients show a 50% reduction of estimated glomerular filtration rate (eGFR) or reach end-stage kidney disease (ESKD) in 5–10 years [2–4]. There is no targeted therapy for IgAN; corticosteroids are frequently used, however they are associated with substantial risk of toxicity and variable efficacy [5, 6].

The Kidney Disease: Improving Global Outcomes (KDIGO) guidelines recommend to stratify patients with IgAN based on risk of progression, so that immunosuppressive treatment can be targeted to high-risk patients [7]. At present, clinical variables associated with a higher risk of kidney function decline include proteinuria >1 g/day, decreased eGFR and the presence of hypertension at baseline and during follow-up [8]. In addition, the Oxford histological risk variables and presence of crescents add independent prognostic information [9–11]. Recently, clinical and histological characteristics were combined into a model to predict individual risk of disease progression [2]. While offering important information about prognosis,

this tool has been validated using clinical data at the time of biopsy, and requires biopsy scores. There is still a need for a tool that will enable individual risk stratification during the course of disease, that can be applied without concurrent kidney biopsy.

In several studies, the benefit and the added value of capillary electrophoresis (CE) coupled mass spectrometry (CE-MS)-based urinary proteome panels were demonstrated in the diagnosis of IgAN [12, 13]. The same technology also enabled the definition of specific peptides to discriminate IgAN from other causes of chronic kidney disease (CKD). Some similarities in the individual biomarkers between CKD aetiologies were observed, but also apparent IgAN-specific peptides that may be helpful to describe disease progression and development of kidney damage in IgAN on a molecular level [14].

Based on these encouraging findings, the multicentre project ‘Personalized Treatment in IgA Nephropathy’ (PERSTIGAN) was initiated. The aim of this study is to identify urinary peptides that enable prediction of kidney function decline and, in a subsequent step, may predict response to immunosuppressive treatment on an individual (i.e. personalized) level. These findings could help to identify those patients with IgAN who require increased medical attention and who may benefit from immunosuppressive therapy. In this article, we present the PERSTIGAN cohort and describe a new urinary peptide classifier identifying patients with rapid eGFR loss during follow-up.

KEY LEARNING POINTS

What is already known about this subject?

- the clinical challenge in patients with immunoglobulin A (IgA) nephropathy (IgAN) is to identify those patients who show rapid progression to end-stage kidney disease and need closer medical attention;
- current risk prediction tools have been validated at time of biopsy and include clinical data and kidney biopsy scores; and
- there is still a need for a non-invasive tool that enables individual risk stratification during the course of disease, and that can be applied without concurrent kidney biopsy.

What this study adds?

- this article describes the discovery of a new urinary proteomics-based biomarker classifier IgAN237 for prediction of risk in patients with IgAN; and
- IgAN237 demonstrates a significantly better prediction of progressive decline of kidney function than clinical parameters alone.

What impact this may have on practice or policy?

- application of this novel classifier may allow individual risk stratification in patients with IgAN; and
- this may be done at any timepoint and independently of a kidney biopsy, and potentially may influence and improve patient management.

MATERIALS AND METHODS

Study population

Urine samples and clinical data were obtained from independent projects from six centres in Europe (Prague, Innsbruck, Göttingen, Wrocław, Madrid and Leipzig; $n = 134$) and from one centre in Canada (Toronto; $n = 75$). All $n = 209$ patients included in this study had biopsy-proven IgAN, a baseline visit and follow-up after collection of the urine sample of at least 12 months, and an eGFR at time of inclusion of at least $20 \text{ mL/min/1.73 m}^2$. The mean number of eGFR measurements per patient was 5.1 ± 4.3 . This study was approved by the local Institutional Review Boards of all participating centres.

Study design and definition of progression

Urine protein excretion per day (i.e. urinary protein–creatinine ratio or 24-h proteinuria, whichever was available), eGFR [calculated using the CKD Epidemiology Collaboration (CKD-EPI) formula] [15], body mass index, age, gender, presence of hypertension, blood pressure, prior use of medications that block the renin–angiotensin–aldosterone system and prior use of immunosuppressive agents (such as steroids) were

determined at time of baseline urine sample collection and during follow-up. Annual eGFR slope after the baseline visit (i.e. during follow-up) was calculated by fitting a straight line through the calculated eGFR using linear regression and the principle of least squares. The total cohort of 209 subjects was divided into tertiles according to the annual eGFR slope. Patients in the lowest eGFR slope tertile (i.e. highest loss of kidney function after baseline) were defined as progressors, patients in the highest eGFR slope tertile (i.e. nearly no loss of kidney function after baseline) were defined as non-progressors, all other were defined as intermediate. When planning the study, we performed a power calculation, which revealed that 29 samples per group were sufficient to detect a 30% change in peptide abundance, with a Type I error of 0.05 and 80% power. We randomly divided the 140 samples from progressors and non-progressors into a training set ($n = 94$, 47 progressors and 47 non-progressors) for biomarker discovery and a test set ($n = 46$, 23 progressors and 23 non-progressors) for the validation study (Figure 1A). Randomization was performed using the Excel randomization (RAND) function (Microsoft, Redmond, WA, USA). There were no significant differences in the clinical characteristics between the subjects from the training set and from the test set (data not shown). The CE-MS data of the training set were used for the definition of peptides being significantly differentially abundant between progressors and non-progressors (Figure 1B) and for the generation of the classifier. The generated classifier was then validated in the test set.

Urine proteome profiling

Proteome analysis was performed in baseline urine samples. The urine samples were thawed immediately before use, prepared as described previously [16] and analysed using a CE coupled online to a micro-time of flight MS, following the sample injection and acquisition protocol as previously described [17]. CE-MS data assessment was performed using an approach based on clustering of the identified compounds to an underlying matrix of previously known identified peptides as described by Latosinska *et al.* [18]. To correct for variability, a linear regression algorithm was applied for normalization, using internal standard peptides as reference [19]. After normalization, all proteomics datasets were deposited, matched and annotated in a Microsoft Structured Query Language (SQL) database and used as input in this study.

Peptide sequencing and matching

The amino acid sequence was determined by CE-MS/MS or liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis as previously described [20]. Protein matching and data analysis were based on Proteome Discoverer version 1.2 (activation type: HCD; precursor mass tolerance: 5 p.p.m.; fragment mass tolerance: 0.05 Da). No fixed modifications were selected, oxidation of methionine and proline were selected as variable modifications. The data were searched against the UniProt human database without enzyme specificity. In the case of LC-MS/MS, matching of the amino acid sequences with the CE-MS acquired ion peaks was based on mass correlation

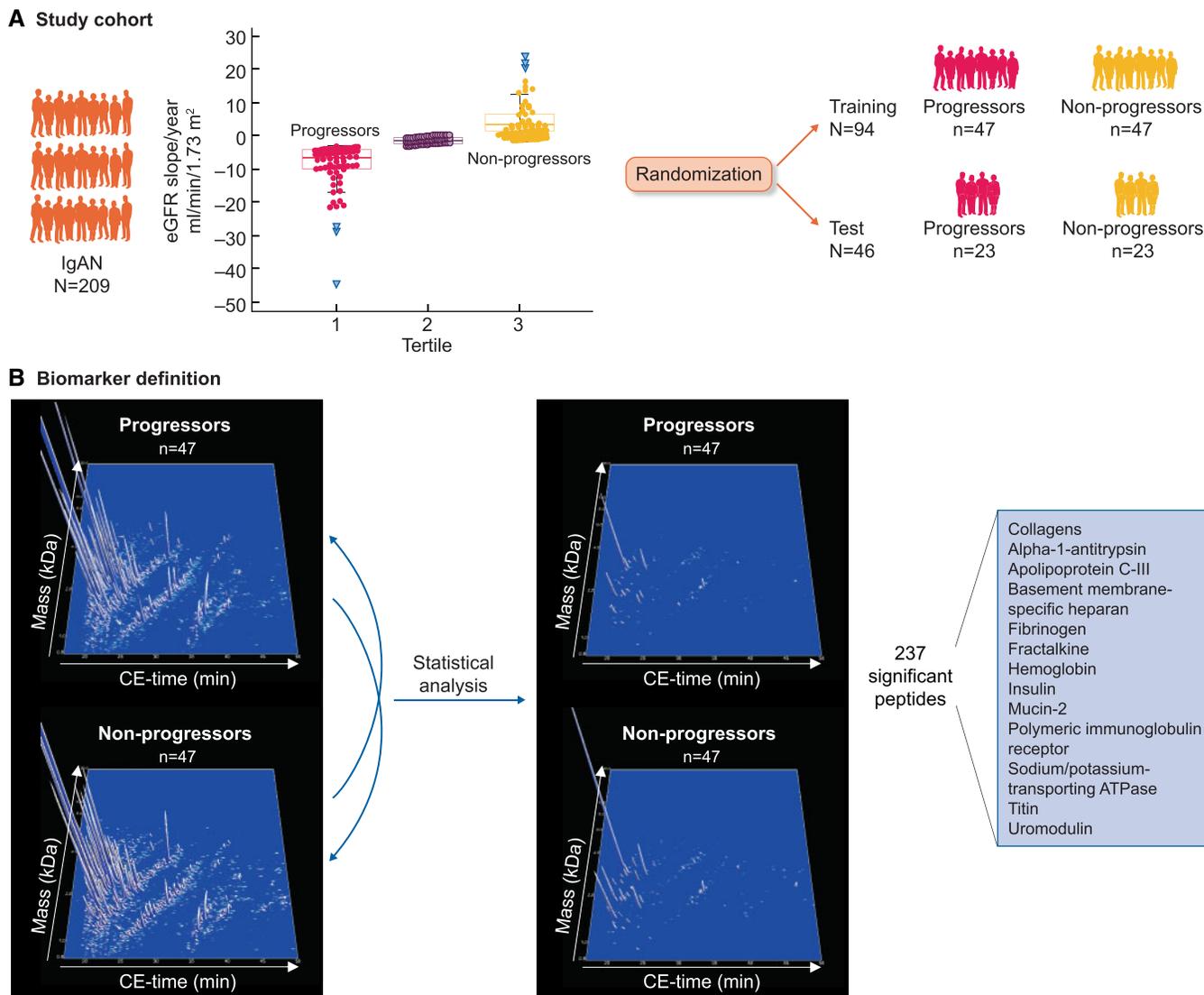


FIGURE 1: Study design of the PERSTIGAN project. (A) We collected urine samples and clinical data from $n = 209$ patients with biopsy-proven IgAN. Patients were stratified by eGFR slope tertiles based on annual eGFR slope during follow-up. Patients in the highest eGFR slope tertile were defined as non-progressors, and patients in the lowest eGFR slope tertile were defined as progressors (patients in the middle tertile were excluded for biomarker definition and validation). $N = 140$ patients were randomized into a training set and into a test set in a 2:1 ratio. (B) For the biomarker definition all detected urinary peptides of the training cohort were compared between progressors and non-progressors (left compiled 3D depiction of all peptide signals in the training cohort), resulting in definition of 237 discriminating peptides (right compiled 3D depiction of 237 significant peptide signals in the training cohort). Protein names of the identified significant peptides are shown in the lower right corner.

between CE-MS and LC-MS/MS analysis and on the theoretical migration time as a result of the assessment of the peptide charge at the working pH of 2.2 [21].

Statistical analysis and classifier generation

Clinical and demographical data were compared using Kruskal–Wallis test MedCalc version 12.7.5.0 (MedCalc Software bvba, Ostend, Belgium). For the definition of specific biomarkers for IgAN progression data of the training set were used and only peptides detected with a frequency of $>30\%$ in at least one of the two groups (progressors or non-progressors) were included. P-values were calculated using non-parametric Wilcoxon test (R-based statistic software, version 2.15.3) and

corrected using the false-discovery rate procedure by Benjamini and Hochberg [22]. A corrected two-tailed $P < 0.05$ was set as the significance level. Potential biomarkers were combined in a support vector machine (SVM)-based classifier.

The area under the receiver operating characteristic (ROC) curve (AUC) of the generated classifier was calculated using MedCalc software. In the case of training set data from the $n-1$ cross-validation were evaluated. The association of baseline and follow-up clinical data and eGFR slope with proteomics classification score was analysed using Spearman's rank correlation test. To analyse the relationship between progressors and non-progressors, proteomics classification score and baseline clinical variables, and to find the

Table 1. Clinical and laboratory parameters at baseline and during follow-up. RAAS, renin–angiotensin–aldosterone system

Characteristics	All, <i>n</i> = 209	eGFR slope tertile 1, <i>n</i> = 70	eGFR slope tertile 2, <i>n</i> = 69	eGFR slope tertile 3, <i>n</i> = 70	P-value
Male gender, %	64	71	64	57	0.113
Caucasian, <i>n</i> (%)	169 (81)	57 (81)	52 (75)	60 (86)	0.299
Age at diagnosis, mean (SD), years	42 (15)	39 (14)	41 (13)	45 (16)	0.113
eGFR, mean (SD), mL/min/1.73 m ²	62.8 (29)	64.9 (30.8)	62.8 (29.5)	60.7 (26.6)	0.842
Proteinuria, median, g/day (95% CI)	1.2 (0.1 to 1.5)	1.8 (1.4–2.2)	1.0 (0.8–1.2)*	1.0 (0.6–1.6)*	<0.001
eGFR slope, median, mL/min/ 1.73 m ² /year (95% CI)	−1.4 (−2.1 to −0.8)	−6.4 (−8.8 to −5.2)	−1.4 (−1.8 to −1.1)*	3.4 (2.3–4.7)*	<0.001
ESKD, <i>n</i> (%)	18(8.6)	14 (20.0)	4 (5.8)	0 (0)*	<0.001
Follow-up time, mean (SD), months	39.4 (27.0)	34.5 (19.0)	48.0 (38.1)*	35.8 (17.1)	0.046
Received RAAS blockade, %	87	87	84	90	0.580
Received IS treatment, %	38	39	31	43	0.306

The P-values were calculated using the Kruskal–Wallis test. *P < 0.05 versus Tertile 1 in the *post hoc* analysis.

best fitting model, the binomial logistic regression analysis was performed. The comparison of the ROC curves was performed using DeLong tests.

Data availability statement

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD020288 [23].

RESULTS

Clinical characteristics of the PERSTIGAN cohort

Baseline urine samples from 209 biopsy-proven IgAN patients were analysed. Detailed clinical characteristics of the whole cohort and of each of the eGFR slope tertiles are listed in Table 1. In brief, baseline mean eGFR was 63 mL/min/1.73 m² and median proteinuria was 1.2 g/day. Twenty-six (12%) of the subjects had a baseline eGFR <30 mL/min/1.73 m² and 110 (53%) of the subjects had a baseline eGFR <60 mL/min/1.73 m². During a mean follow-up time of 39 months median annual eGFR decline was −1.4 mL/min/1.73 m²/year. Eighteen patients (8.6%) progressed to ESKD. Progressors had a median annual eGFR decline of −6.4 mL/min/1.73 m², and 14 (20%) of these patients progressed to ESKD, while non-progressors had a median annual eGFR change of +3.4 mL/min/1.73 m² and none progressed to ESKD. Notably, mean baseline eGFR of progressors was not significantly different from non-progressors (64.9 versus 60.7 mL/min/1.73 m², P = 0.385).

Development of the urine peptide-based classifier

We identified 237 discriminatory peptides showing a significantly different abundance between progressors and non-progressors of the training set after correction for multiple testing (Supplementary data, Table S1 and Figure 1). These peptides were combined into a single analysis covariate using an SVM algorithm, and this was designated as ‘IgAN237’. Sequence information was obtained for 84 of the 237 peptides associated with progression of IgAN. Here, fragments of different collagens, apolipoprotein C-III, alpha-1 antitrypsin, fibrinogen alpha and beta, titin, haemoglobin subunits, sodium/potassium-transporting ATPase subunit gamma, uromodulin, mucin-2, fractalkine, polymeric Ig receptor and insulin were

included. As shown in Supplementary data, Table S1, most of the collagen fragments displayed decreased abundance in progressive subjects. The most significant 25 sequenced peptides are listed in Table 2.

Comparison with CKD273

Next, we investigated whether the IgAN237 classifier was specific for IgAN or reflected more general changes during CKD progression. For this purpose, the 87 sequenced peptides were compared with urinary peptides of the CKD273 biomarker, which has been shown to identify patients with progressive CKD irrespective of a specific aetiology [24]. Only 15 peptides (14 collagen and 1 fibrinogen fragment) were found in both, the IgAN237 and the CKD273 classifiers, suggesting that IgAN237 may be disease-specific.

Prediction of progression using the IgAN237 classifier

We used ROC statistics to validate that the classification score of the IgAN237 biomarker discriminates IgAN progressors versus non-progressors. When applied to the n-1 cross-validated training set ROC resulted in an AUC of 0.91 [95% confidence interval (CI) 0.85–0.97]. To validate the generated IgAN237 classifier we applied it to the independent test set of 23 progressors and 23 non-progressors. This resulted in a significant discrimination of IgAN progressors with an AUC of 0.72 (95% CI 0.57–0.87) (Figure 2).

Correlation of IgAN237 with clinical parameters

To characterize the correlation of IgAN237 with the phenotype we performed a correlation analysis within the whole cohort of *n* = 209 subjects with baseline clinical data at the time of urine sampling. The IgAN237 score showed a significant correlation with proteinuria (ρ = 0.181, P = 0.011) and a borderline significant correlation with baseline eGFR (ρ = −0.136, P = 0.05) (Figure 3A). No other baseline variables were significantly associated with the IgAN237 score.

Next, we analysed the association of the IgAN237 classification score with kidney function decline, follow-up eGFR and progression to ESKD. The IgAN237 scores were significantly negatively correlated with eGFR slope (ρ = −0.484, P < 0.001) and with the last follow-up eGFR value (ρ = −0.319, P < 0.001; Figure 3B). Patients who progressed to ESKD (*n* = 18) showed

Table 2. List of 25 most significant sequenced peptides of IgAN237 with lowest adjusted P-values

Peptide ID	Corr. P-value	Mean amplitude progressors (training)	Mean amplitude non-progressors (training)	Fold change	Sequence	Protein name	Protein symbol	Start AA	Stop AA
99915883	0.0105	8.96	58.32	0.154	NSGEpGApGSK GDTGAKGep GPVGVQGPpGPAG	Collagen alpha-1(I) chain	COL1A1	432	464
99910985	0.0105	43.83	86.99	0.504	NSGEKGDQGFQ GQPGFPGPPGP	Collagen alpha-1(XVI) chain	COL16A1	1145	1166
99907573	0.0105	92.15	292.94	0.315	ESVVLEPEAT	Fractalkine	X3CL1	123	132
99910668	0.0123	1497.39	2131.3	0.703	GpPGEAGKpGEQ GVpGDLGAPGp	Collagen alpha-1(I) chain	COL1A1	650	672
99914456	0.0123	28.16	450.67	0.062	DLQVGVQVELGGPGA GSLQLALEGSLQ	Insulin	INS	60	87
99904786	0.0125	69.59	122.18	0.570	SpGSPGPDGKTGPPGP	Collagen alpha-1(I) chain	COL1A1	543	559
99917505	0.0157	122.02	283.96	0.430	AGRpGEVGPpGPpGPAGEKG SPGADGPAGAPGTpGPQG	Collagen alpha-1(I) chain	COL1A1	916	953
99917388	0.0157	99.88	204.59	0.488	GpPpGpGRDGEDGPTGP pGPPGPPGPPGLGGNFAAQ	Collagen alpha-2(I) chain	COL1A2	45	80
99916194	0.0167	26.62	623.15	0.043	EAEDLQVGVQVELGGG PGAGSLQLALEGSLQ	Insulin	INS	57	87
99910573	0.0168	987.11	1499.34	0.658	GpPGEAGKpGEQ GVpGDLGAPGP	Collagen alpha-1(I) chain	COL1A1	650	672
99912693	0.0175	31.22	51.49	0.606	GPAGpPGKAGEDGH PGKpGRpGERG	Collagen alpha-2(I) chain	COL1A2	133	157
99908185	0.0179	108.3	412.91	0.262	EEAPSLRPAPPPISGGGY	Fibrinogen beta chain	FGB	54	71
99914501	0.0179	98.36	152.89	0.643	pPGADGQPgAKGepGD AGAKGDAGpPGPAGP	Collagen alpha-1(I) chain	COL1A1	816	846
99909057	0.0186	3.49	35.74	0.098	GpAGATGDRGEAG AAGPAGpAGP	Collagen alpha-2(I) chain	COL1A2	685	707
99903148	0.0234	1279.9	275.75	4.642	FMGKVVNPTQK	Alpha-1-antitrypsin	SERPINA1	408	418
99910838	0.0234	25.77	61.51	0.419	GpTGpIGPpGpAGQ PGDKGEGGAP	Collagen alpha-1(III) chain	COL3A1	762	785
99908885	0.0240	3223.72	6243.43	0.516	AAHLPAEFTPAVHASLTK	Haemoglobin subunit alpha	HBA1	111	128
99917047	0.0242	2828.47	4810.36	0.588	PpGESGREGApGAEGSp GRDGSpGAKGDRGETGP	Collagen alpha-1(I) chain	COL1A1	1008	1041
99906068	0.0242	326.45	545.87	0.598	NDGApGKNGERGGpGGP	Collagen alpha-1(III) chain	COL3A1	586	602
99909564	0.0242	18.51	96.54	0.192	GEKpSGEAGTA GPpGTpGPQG	Collagen alpha-2(I) chain	COL1A2	844	865
99913909	0.0242	902.04	5019.41	0.180	EKSAVTALWGVNV DEVGGEALGRL	Haemoglobin subunit beta	HBB	8	32
99906402	0.0254	11.3	57.83	0.195	WQGVVEGEGAGQKDF	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	4245	4259
99912198	0.0254	134.56	344.73	0.390	AGppGEAGKPEQGV pGDLGApGPPG	Collagen alpha-1(I) chain	COL1A1	649	674
99913650	0.0266	20.96	63.72	0.329	GpTGATGDKGPPG PVGPPGSNGpVGEpGP	Collagen alpha-2(V) chain	COL5A2	1020	1048
99905463	0.0276	60.81	258.26	0.235	ApGDKGESGSPGAPGT	Collagen alpha-1(I) chain	COL1A1	777	793

a significantly higher IgAN237 score than patients who did not (0.968; 95% CI 0.643–1.529 versus 0.195; 95% CI –0.038 to 0.353, $P < 0.001$; [Figure 3C](#)). The median IgAN237 score was 0.649 in the progressors, 0.409 in the intermediate group and –0.465 in the non-progressors, respectively ([Figure 4](#)). Significant differences were observed between progressors and non-progressors ($P < 0.001$), between progressors

and the intermediate tertile ($P = 0.025$) and between the intermediate tertile and non-progressors ($P < 0.001$).

Added value of proteomics IgAN237 classifier

In the next step, we evaluated if the IgAN237 classifier improves discrimination between progressors and non-progressors compared with clinical data alone [age, gender,

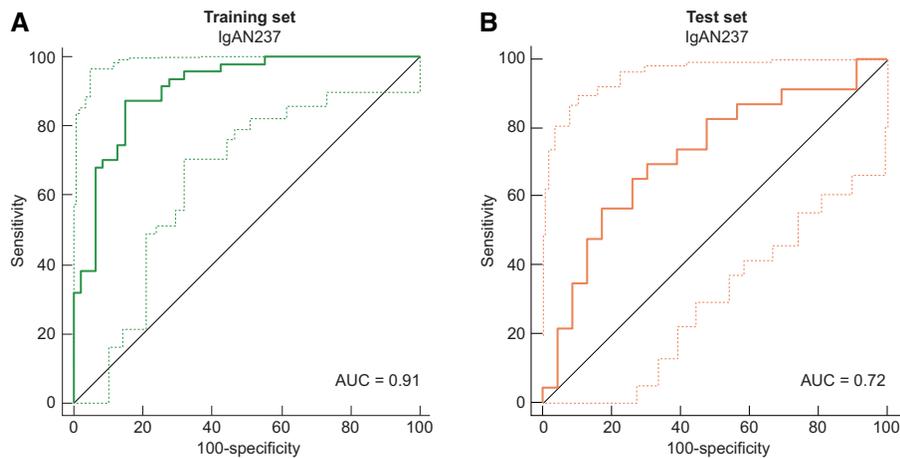


FIGURE 2: ROC for prediction of IgAN progression by the IgAN237 biomarker panel. (A) *n*-1 cross-validated training set (*n* = 94). (B) Test set (*n* = 46). Light lines: 95% CI.

eGFR, proteinuria and mean arterial pressure (MAP) at urine sampling]. For this analysis, we used progressors and non-progressors of the whole patient cohort. The combined baseline clinical data resulted in an AUC of 0.72 (95% CI 0.64–0.81) for the prediction of fast progression. In the next step, the IgAN237 classifier was added to the clinical parameters. Logistic regression analysis was performed (i) to analyse the relationship between the progression and non-progression outcome and the individual clinical variables together with IgAN237 score and (ii) to find the best fitting model to predict the progression and non-progression outcome. The resulting regression coefficients showed the highest value for IgAN237 score (1.98072) in comparison with the regression coefficients of the clinical parameters, of which only gender reached statistical significance (Table 3), indicating highly significant association of IgAN237 with outcome.

All parameters together with their individual regression coefficients were combined to the following regression equation for a clinical data-adjusted IgAN237 score:

$$\text{IgAN237}_{\text{adj}} = -2.8438 + (-0.034667 \cdot \text{age}) + (0.016583 \cdot \text{eGFR}) + (0.019597 \cdot \text{MAP}) + (1.26778 \cdot \text{gender}) + (0.14726 \cdot \text{urine protein}) + (1.98072 \cdot \text{IgAN237 score}).$$

(age in years; eGFR calculated by the CKD-EPI formula in mL/min/1.73 m²; MAP in mmHg; male = 1, female = 0; urinary protein/24 h in g/day).

Using this formula the combination of clinical parameters with the IgAN237 scores resulted in AUC of 0.89 (95% CI 0.83–0.95), demonstrating significantly better performance than the clinical parameters alone (*P* < 0.001) (Figure 5).

DISCUSSION

In this study, we developed a urinary biomarker-based classifier to predict a progressive or a non-progressive clinical course in patients with IgAN. The IgAN237 classifier correlated significantly with eGFR slope during follow-up, eGFR at the end of follow-up and was significantly different between patients who progressed to ESKD and those who did not. Further, we showed that addition of the IgAN237 classifier to clinical variables

improves discrimination between progressors and non-progressors beyond clinical data alone.

Progressors were defined based upon tertiles of the annual eGFR decline during follow-up. The urine peptide profiles of the lowest (i.e. progressors) and the highest (i.e. non-progressors) eGFR slope tertile were compared. We chose this approach for several reasons: By analysing extremes, i.e. patients with well-separated phenotypical characteristics such as eGFR slopes of $-6.4 \text{ mL/min/1.73 m}^2$ versus $+3.4 \text{ mL/min/1.73 m}^2/\text{year}$, it is more probable that biomarkers with ‘true’ predictive value for a given clinical outcome will be identified. Furthermore, the median annual eGFR decline of $-6.4 \text{ mL/min/1.73 m}^2$ is in line with current CKD KDIGO guidelines in which fast progression is defined as a sustained decline in eGFR of $>5 \text{ mL/min/1.73 m}^2/\text{year}$ [25]. A dichotomous separation of the cohort using a clear cut-off such as an eGFR decline of below or above $-5 \text{ mL/min/1.73 m}^2/\text{year}$ or $-30/-40\%$ may have resulted in a large number of patients with a more vague or borderline definition of progression. This might have impacted the validity of the urinary peptide classifier to identify truly progressive patients with IgAN. Although the number of patients with ESKD was relatively low (*n* = 18) for the purpose of classifier development, subjects who reached ESKD showed a significantly higher IgAN237 score than patients who did not. About 14 of the 18 ESKD subjects were part of the progressor group. These results suggest that the relative definition of CKD progression used in this project is in line with current guidelines and represents progressive and non-progressive patients seen in routine clinical practice.

A major constituent of the IgAN237 classifier was collagen fragments, representing 69% of the identified peptides. Particularly fragments from collagen Type I have been reported as an integral part of diagnostic classifiers in many studies of CKD [14, 24, 26], and also recently in a study on obesity-related nephropathy [27]. Most collagen fragments are decreased in subjects with progressive CKD, including patients with diabetes and obesity [27, 28]. The negative correlation of urinary fragments of collagen Type I with the degree of renal fibrosis has been demonstrated in kidney biopsies from various CKD aetiologies [29]. This reduced abundance of collagen fragments in the

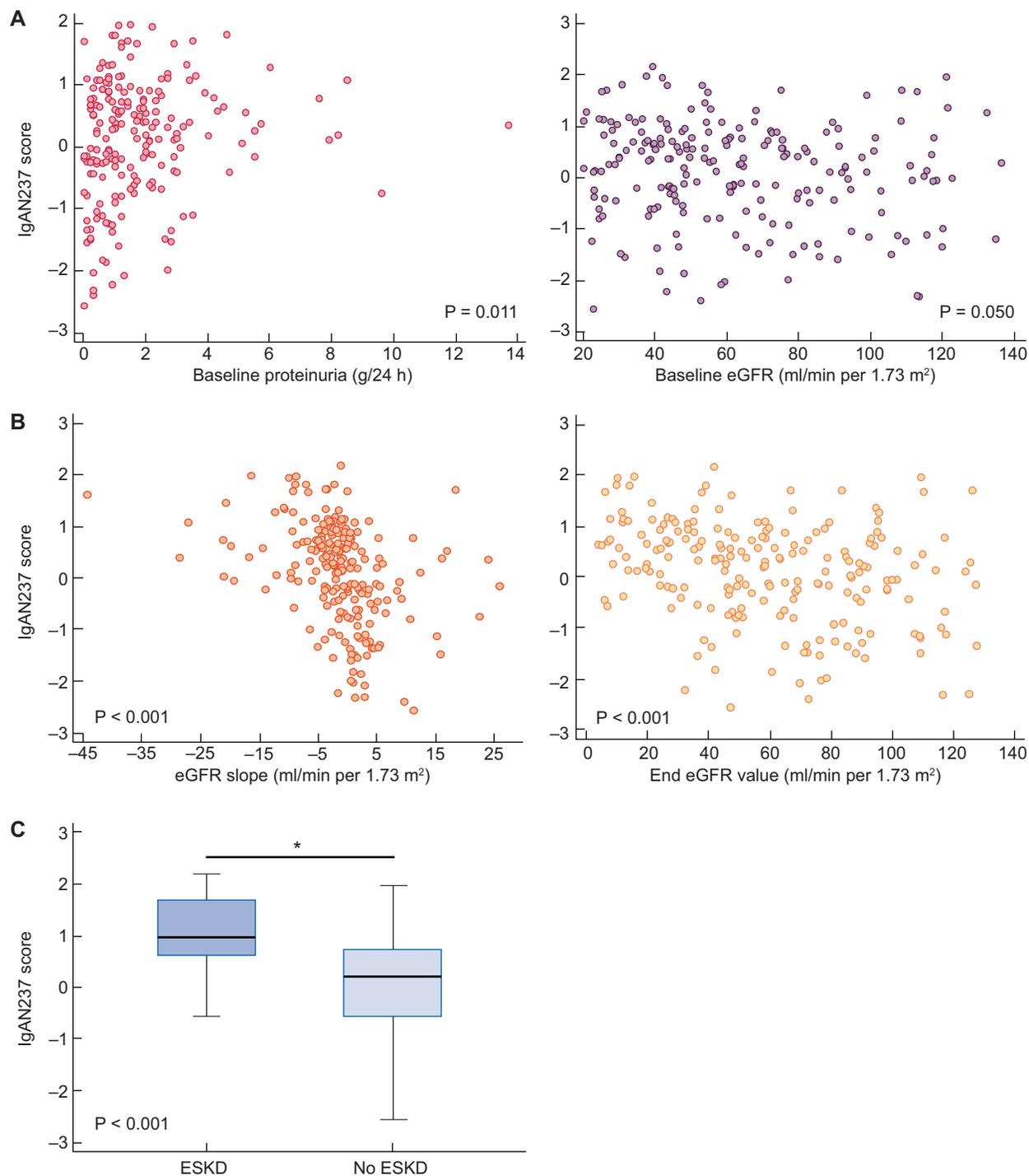


FIGURE 3: Association of IgAN237 scores with clinical parameters at baseline and during follow-up. (A) Scatter diagram for correlation between IgAN237 score and baseline proteinuria and IgAN237 score and eGFR, respectively. (B) Scatter diagram for correlation of IgAN237 score with eGFR slope and IgAN237 score and last eGFR value, respectively. (C) IgAN237 scores in form of box-and-whisker plot (median, 25–75 percentile) in patients who progressed to ESKD and who did not, * $P < 0.05$.

urine might mirror reduced degradation of collagen and an increase in extracellular matrix accumulation in the kidney in progressive patients, consequently leading to kidney fibrosis. Thus the urinary proteome appears to a substantial degree reflect the turnover of extracellular matrix in the kidney [30].

A surprising finding of our study was the decreased abundance of several fragments of the haemoglobin alpha and beta

chains in progressive patients. In previous studies, higher urinary levels of various haemoglobin fragments have been described in vasculitis patients [26] and also in IgAN patients [14], most probably reflecting prominent glomerular inflammation. Thus, a decrease in haemoglobin peptides in the urine of progressive subjects may indicate a less inflammatory and more fibrotic phenotype. These findings are further supported by the

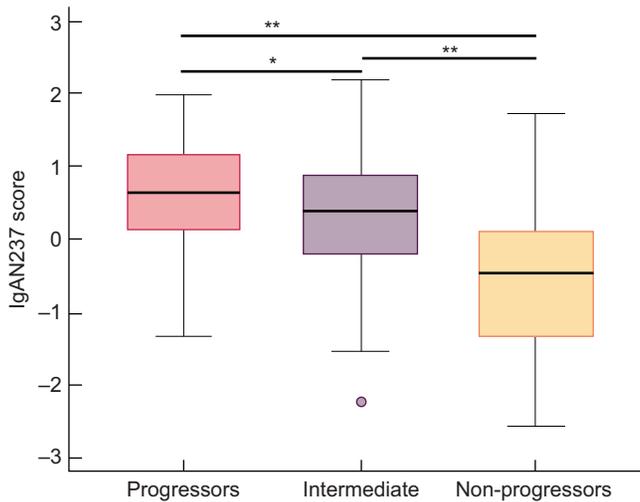


FIGURE 4: IgAN237 classification score. IgAN237 score in the progressor ($n = 70$), intermediate ($n = 69$) and non-progressor ($n = 70$) groups, * $P < 0.05$, ** $P < 0.001$.

Table 3. Logistic regression coefficients for prediction of outcome

Variable	Coefficient	SE	P-value
Age	-0.034667	0.021400	0.105
eGFR	0.016583	0.010744	0.123
Male gender	1.26778	0.57323	0.027
Proteinuria	0.14726	0.13789	0.286
MAP	0.019597	0.017673	0.268
IgAN237	1.98072	0.39590	<0.001

SE, standard error.

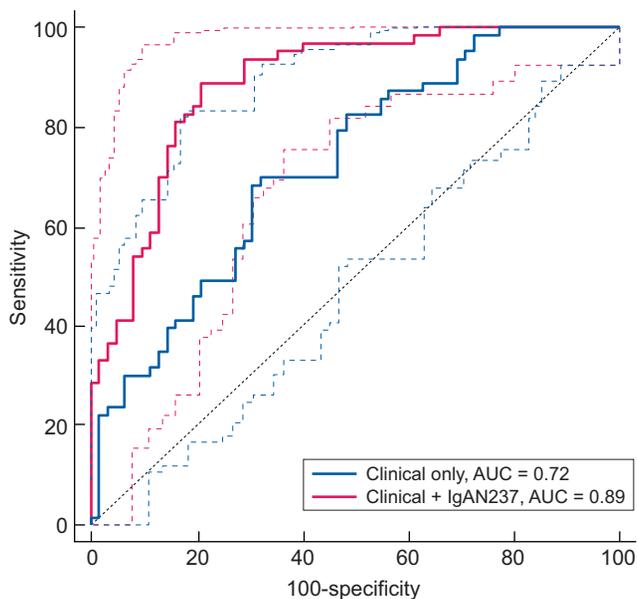


FIGURE 5: ROC for prediction of IgAN progression based on clinical characteristics or a combination of clinical characteristics and the IgAN237 biomarker panel in all progressor and non-progressor patients ($n = 140$). Clinical characteristics were age, gender, eGFR, proteinuria and MAP. Dashed lines: 95% CI.

decrease of peptide fragments from tubular proteins (mucin, uromodulin), possibly indicating disturbed tubular function, due to an increase in tubular atrophy and, again, interstitial fibrosis. Alpha-1 antitrypsin was highly increased in progressive subjects. Alpha-1 antitrypsin is a major protease inhibitor controlling tissue damage by inhibiting proteases, including elastase and proteinase-3 released by leucocytes infiltrating the kidney as well as collagenases [31]. An increased abundance of alpha-1 antitrypsin has been reported in various types of glomerulonephritis, such as IgAN, vasculitis and membranous nephropathy [14, 26, 32]. Smith *et al.* identified urinary alpha-1 antitrypsin excretion in primary glomerulonephritis and showed localization of this protein in the podocytes within sclerotic glomeruli by immunohistochemistry [33]. Thus, the highly increased abundance of alpha-1 antitrypsin peptide fragments in the urine is compatible with the observed decrease in collagen fragments, potentially related to inhibition of collagenase activity, and to sclerotic glomeruli in progressive patients. In summary, the components of the IgAN237 classifier seem to reflect several pathologic processes known to be associated with progressive IgAN, such as inflammation, matrix turnover and fibrosis.

Although current risk stratification of patients with IgAN based on the dynamics of proteinuria, eGFR and blood pressure control is well established, it depends upon availability of detailed histologic data obtained concurrent to the clinical data, and importantly it is not validated using clinical data at other points during the follow-up [2]. We endeavoured to derive a non-invasive tool that did not require pathologic data. In our study, the addition of the IgAN237 score to clinical variables including eGFR, proteinuria, gender and MAP resulted in better discrimination of patients at the highest risk of progressive disease (AUC of 0.89 compared with 0.720, $P < 0.001$). This proteomic classifier, adjusted for baseline clinical parameters derived from the presented cohort, was further developed into an online calculator of risk of progression in IgAN (www.perstigan.eu) to be used for baseline risk prediction to potentially contribute to therapy decisions. As a hypothetical example, considering a patient with an IgAN237 score in the range of the progressor group (median 0.64), only minor clinical signs of chronicity, preserved kidney function and histological signs of only limited tubular atrophy (MEST T0 or T1), one could decide to apply immunosuppression (IS), having in mind the discussed guidelines and evidence applicable in this field. These combined modelling approaches must be taken to next steps, including prospective therapy studies using such newly developed diagnostic approaches.

The ability of CE-MS-based urinary proteomics to discriminate between IgAN and healthy controls [12, 13] or other kidney diseases [14] has been demonstrated previously. Many of the biomarker peptides identified in our study were also identified in other studies on IgAN, further supporting the significance of these peptides as specific markers of disease activity or progression. The PERSTIGAN study was designed to generate a urinary peptide profile for personalized risk stratification. A central element in personalized medicine is the assessment of the individual health status of a patient, which is a significant

challenge particularly in heterogeneous diseases such as IgAN. In comparison with single biomarkers, omics-multimarker-based approaches offer the possibility to describe a dynamic phenotype more comprehensively and more accurately, as a result of simultaneous assessment of multiple parameters. Applications of proteomics biomarkers in personalized medicine in the context of kidney disease provided already successful examples [34], with CKD273 for the diagnosis and prediction of CKD being the most mature example [24]. Especially in early disease, CKD273 showed a significantly better predicted value for CKD progression than albuminuria [35–37], also in non-albuminuric patients [38], and it is predictive of death in two independent cohorts [39, 40].

Although peptide contents between CKD273 and IgAN237 are moderately similar, CKD273 did not identify progressive disease in the PERSTIGAN cohort (data not shown). The overlap between IgAN237 and CKD273 was small, with 15 of the 84 (18%) sequenced peptides being present in both classifiers. These data further suggest that IgAN237 serves as a disease-specific classifier.

The most important strengths of the PERSTIGAN project and this study are the relatively large number of IgAN subjects, the multicentre design and the clinical characteristics of the patients, which are quite comparable to patients included in large randomized controlled trials such as the Supportive Versus Immunosuppressive Therapy for the Treatment of Progressive IgA Nephropathy (STOP-IgAN) and the Therapeutic Evaluation of Steroids in IgA Nephropathy Global (TESTING) study trial. Furthermore, many of the peptides identified can be associated with pathological processes known to be associated with worsening of kidney function.

This study has several limitations, the lack of histological scores (MEST-C) being the most important one. Therefore, we cannot compare the value of the IgAN237 score with an integrated risk prediction tool applied at the time of biopsy [2]. While this represents an area for future studies, a prediction tool that depends entirely upon data derived non-invasively presents a distinct advantage. Also, this study is not a prospective study, however it is based on prospectively collected samples. Furthermore, the IgAN237 score was developed in a mainly Caucasian population, so it remains speculative whether our findings can be extended to other ethnicities. The number of subjects ($n = 209$) may also limit the generalizability of our findings, so the results of this study need to be validated in larger and independent cohorts.

In conclusion, the IgAN237 classifier identifies IgAN patients who are at high risk for disease progression, independently of a concurrent kidney biopsy or traditional clinical parameters such as eGFR, proteinuria or blood pressure. This may allow individual risk stratification and potentially may influence patient management.

SUPPLEMENTARY DATA

Supplementary data are available at [ndt](http://ndt.oxfordjournals.org/) online.

ACKNOWLEDGEMENTS

Members of the PERSTIGAN working group. Division of Nephrology and KfH Renal Unit, Hospital St Georg, Leipzig,

Germany: Joachim Beige, Ralph Wendt. Mosaiques Diagnostics GmbH, Hannover, Germany: Justyna Siwy, Petra Zürbig, Harald Mischak, Annika Durban, Julia Raad, Igor Golovko. University of Toronto and Toronto General Research Institute, Toronto, Canada: Heather Reich, Ping Lam, Stuart Yang. Research Health Institute-Fundación Jiménez Díaz (IIS-FJD UAM), Madrid, Spain: Ana Belen Sanz, Beatriz Fernandez-Fernandez, Jorge Enrique Rojas-Rivera, Maria Vanessa Perez-Gomez, Alberto Ortiz, Maria Dolores Sanchez-Niño, Jinny Sanchez-Rodriguez. Medical University Innsbruck, Innsbruck, Austria: Michael Rudnicki, Julia Kerschbaum, Johannes Leierer, Gert Mayer. Umea University, Department Public health and clinical medicine, Umea, Sweden: Bernd Stegmayr, Björn Peters.

FUNDING

This work was supported by the ERA-NET PerMed programme (Ref. No. ERAPERMED2018-217) co-funded by the European Commission and the national funding agencies (see below). Germany (J.S., R.W., H.M. and J.B.): this project was supported by the Federal Ministry of Education and Research (BMBF) under grant number 01KU1922A for H.M. and J.S. and 01KU1922B 9899073 for R.W. and J.B. Spain (A.O., A.B.S. and M.V.P.-G.): these authors are supported by FIS/FEDER funds under grant number AC18/00071 and AC18/00064. Salary support ISCIII Miguel Servet to ABS and RETIC-REDINREN RD 016/0009. Austria (M.R., J.K. and J.L.): this project was funded by the Austrian Science Fund under grant number I 4239-B to M.R. Sweden (B.P. and B.S.): this project was funded by the Swedish Research Council (2018-05615) and the Research Fund (FoU) at Skaraborg Hospital, Skövde, Sweden (VGSKAS-930079). Canada (H.N.R.): this project was supported by the Canadian Institutes of Health Research through the ERA-Net PerMed initiative to H.N.R. H.N.R.'s work is funded by the Gabor Zellerman Chair in Nephrology Research at the University of Toronto.

AUTHORS' CONTRIBUTIONS

M.R., J.S., R.W., B.P., A.B.S., A.O., B.S., H.M., J.B. and H.N.R. designed the study; M.R., J.K., J.L., R.W., M.L., M.J.K., D.M., M.N., M.B., A.B.S., A.O., M.V.P.-G., V.T. and H.N.R. collected urine samples and/or clinical data; J.S. carried out the experiments; J.S., H.M. and H.N.R. analysed the data; M.R., J.S. and H.N.R. made the tables and the figures and drafted the manuscript; all authors revised the manuscript; all authors approved the final version of the manuscript.

CONFLICT OF INTEREST STATEMENT

J.S. is employed by Mosaiques Diagnostics. A.O. reports grants and personal fees from Sanofi, Mundipharma, personal fees from Otsuka, Amgen, Fresenius, Amicus and AstraZeneca, outside the submitted work. H.M. reports grants from Bundesministerium für Bildung und Forschung (BMBF) during the conduct of the study; he is the co-founder and co-owner of Mosaiques Diagnostics.

H.N.R. reports other financial activities from Calliditas and from Omeros outside the submitted work. All other authors have nothing to disclose. The results presented in this article have not been published previously in whole or part, except in abstract format.

REFERENCES

- McGrogan A, Franssen CF, de Vries CS. The incidence of primary glomerulonephritis worldwide: a systematic review of the literature. *Nephrol Dial Transplant* 2011; 26: 414–430
- Barbour SJ, Coppo R, Zhang H *et al.*; for the International IgA Nephropathy Network. Evaluating a new international risk-prediction tool in IgA nephropathy. *JAMA Intern Med* 2019; 179: 942–952
- Barbour SJ, Espino-Hernandez G, Reich HN *et al.* The MEST score provides earlier risk prediction in IgA nephropathy. *Kidney Int* 2016; 89: 167–175
- Reich HN, Troyanov S, Scholey JW *et al.* Remission of proteinuria improves prognosis in IgA nephropathy. *J Am Soc Nephrol* 2007; 18: 3177–3183
- Lv J, Zhang H, Wong MG *et al.*; for the TESTING Study Group. Effect of oral methylprednisolone on clinical outcomes in patients with IgA nephropathy: the TESTING randomized clinical trial. *JAMA* 2017; 318: 432–442
- Rauen T, Eitner F, Fitzner C *et al.* Intensive supportive care plus immunosuppression in IgA nephropathy. *N Engl J Med* 2015; 373: 2225–2236
- Kidney Disease: Improving Global Outcomes (KDIGO) Glomerulonephritis Work Group. KDIGO clinical practice guideline for glomerulonephritis. *Kidney Int Suppl* 2012; 2: 139–274
- Barbour SJ, Reich HN. Risk stratification of patients with IgA nephropathy. *Am J Kidney Dis* 2012; 59: 865–873
- Haas M, Verhave JC, Liu ZH *et al.* A multicenter study of the predictive value of crescents in IgA nephropathy. *J Am Soc Nephrol* 2017; 28: 691–701
- Roberts ISD, Cook HT, Troyanov S *et al.*; Working Group of the International Ig ANN, the Renal Pathology S. The Oxford classification of IgA nephropathy: pathology definitions, correlations, and reproducibility. *Kidney Int* 2009; 76: 546–556
- Cattran DC, Coppo R, Cook HT *et al.*; Working Group of the International IgA Nephropathy Network and the Renal Pathology Society. The Oxford classification of IgA nephropathy: rationale, clinicopathological correlations, and classification. *Kidney Int* 2009; 76: 534–545
- Julian BA, Wittke S, Novak J *et al.* Electrophoretic methods for analysis of urinary polypeptides in IgA-associated renal diseases. *Electrophoresis* 2007; 28: 4469–4483
- Haubitz M, Wittke S, Weissinger EM *et al.* Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. *Kidney Int* 2005; 67: 2313–2320
- Siwy J, Zurbig P, Argiles A *et al.* Noninvasive diagnosis of chronic kidney diseases using urinary proteome analysis. *Nephrol Dial Transplant* 2017; 32: 2079–2089
- Levey AS, Stevens LA, Schmid CH *et al.*; for the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; 150: 604–612
- Theodorescu D, Schiffer E, Bauer HW *et al.* Discovery and validation of urinary biomarkers for prostate cancer. *Prot Clin Appl* 2008; 2: 556–570
- Zimmerli LU, Schiffer E, Zurbig P *et al.* Urinary proteomic biomarkers in coronary artery disease. *Mol Cell Proteomics* 2008; 7: 290–298
- Latosinska A, Siwy J, Mischak H *et al.* Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: the past, the present, and the future. *Electrophoresis* 2019; 40: 2294–2308
- Jantos-Siwy J, Schiffer E, Brand K *et al.* Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. *J Proteome Res* 2009; 8: 268–281
- Klein J, Papadopoulos T, Mischak H *et al.* Comparison of CE-MS/MS and LC-MS/MS sequencing demonstrates significant complementarity in natural peptide identification in human urine. *Electrophoresis* 2014; 35: 1060–1064
- Zurbig P, Renfrow MB, Schiffer E *et al.* Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. *Electrophoresis* 2006; 27: 2111–2125
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 1995; 57: 289–300
- Perez-Riverol Y, Csordas A, Bai J *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 2019; 47: D442–D450
- Good DM, Zurbig P, Argiles A *et al.* Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Mol Cell Proteomics* 2010; 9: 2424–2437
- KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl* 2013; 3: 1–150
- Haubitz M, Good DM, Woywodt A *et al.* Identification and validation of urinary biomarkers for differential diagnosis and evaluation of therapeutic intervention in anti-neutrophil cytoplasmic antibody-associated vasculitis. *Mol Cell Proteomics* 2009; 8: 2296–2307
- Wendt R, He T, Latosinska A *et al.* Proteomic characterization of obesity-related nephropathy. *Clin Kidney J* 2020; 13: 684–692
- Rossing K, Mischak H, Dakna M *et al.* Urinary proteomics in diabetes and CKD. *J Am Soc Nephrol* 2008; 19: 1283–1290
- Magalhaes P, Pejchinovski M, Markoska K *et al.* Association of kidney fibrosis with urinary peptides: a path towards non-invasive liquid biopsies? *Sci Rep* 2017; 7: 16915
- Decramer S, Gonzalez de Peredo A, Breuil B *et al.* Urine in clinical proteomics. *Mol Cell Proteomics* 2008; 7: 1850–1862
- Eisen AZ, Bauer EA, Jeffrey JJ. Human skin collagenase. The role of serum alpha-globulins in the control of activity in vivo and in vitro. *Proc Natl Acad Sci USA* 1971; 68: 248–251
- Prikryl P, Vojtova L, Maixnerova D *et al.* Proteomic approach for identification of IgA nephropathy-related biomarkers in urine. *Physiol Res* 2017; 66: 621–632
- Smith A, L'Imperio V, De Sio G *et al.* Antitrypsin detected by MALDI imaging in the study of glomerulonephritis: its relevance in chronic kidney disease progression. *Proteomics* 2016; 16: 1759–1766
- Siwy J, Mischak H, Zurbig P. Proteomics and personalized medicine: a focus on kidney disease. *Expert Rev Proteomics* 2019; 16: 773–782
- Pontillo C, Zhang ZY, Schanstra JP *et al.* Prediction of chronic kidney disease stage 3 by CKD273, a urinary proteomic biomarker. *Kidney Int Rep* 2017; 2: 1066–1075
- Pontillo C, Jacobs L, Staessen JA *et al.* A urinary proteome-based classifier for the early detection of decline in glomerular filtration. *Nephrol Dial Transplant* 2017; 32: 1510–1516
- Schanstra JP, Zurbig P, Alkhalaf A *et al.* Diagnosis and prediction of CKD progression by assessment of urinary peptides. *J Am Soc Nephrol* 2015; 26: 1999–2010
- Zurbig P, Mischak H, Menne J *et al.* CKD273 enables efficient prediction of diabetic nephropathy in nonalbuminuric patients. *Diabetes Care* 2019; 42: e4–e5
- Verbeke F, Siwy J, Van Biesen W *et al.* The urinary proteomics classifier chronic kidney disease 273 predicts cardiovascular outcome in patients with chronic kidney disease. *Nephrol Dial Transplant* 2019; doi: 10.1093/ndt/gfz242 (14 December 2019, date last accessed)
- Currie GE, von Scholten BJ, Mary S *et al.* Urinary proteomics for prediction of mortality in patients with type 2 diabetes and microalbuminuria. *Cardiovasc Diabetol* 2018; 17: 50

Received: 21.7.2020; Editorial decision: 9.10.2020