Validation of Plasma Biomarker Candidates for the Prediction of eGFR Decline in Patients With Type 2 Diabetes

Diabetes Care 2018;41:1947–1954 | https://doi.org/10.2337/dc18-0532

Andreas Heinzel,[1] Michael Kammer,[1,2] Gert Mayer,[3] Roman Reindl-Schwaighofer,[1] Karin Hu,[1] Paul Perco,[3] Susanne Eder,[3] Laszlo Rosivall,[4] Patrick B. Mark,[5] Wenjun Ju,[6] Matthias Kretzler,[6] Peter Gilmour,[7] Jonathan M. Wilson,[8] Kevin L. Duffin,[8] Moustafa Abdalla,[9,10,11] Mark I. McCarthy,[9,10,12] Georg Heinze,[2] Hiddo L. Heerspink,[13] Andrzej Wiecek,[14] Maria F. Gomez,[15] and Rainer Oberbauer[1] for the BEAt-DKD Consortium

OBJECTIVE

The decline of estimated glomerular filtration rate (eGFR) in patients with type 2 diabetes is variable, and early interventions would likely be cost-effective. We elucidated the contribution of 17 plasma biomarkers to the prediction of eGFR loss on top of clinical risk factors.

RESEARCH DESIGN AND METHODS

We studied participants in PROVALID (PROspective cohort study in patients with type 2 diabetes mellitus for VALIDation of biomarkers), a prospective multinational cohort study of patients with type 2 diabetes and a follow-up of more than 24 months (n = 2,560; baseline median eGFR, 84 mL/min/1.73 m^2; urine albumin-to-creatinine ratio, 8.1 mg/g). The 17 biomarkers were measured at baseline in 481 samples using Luminex and ELISA. The prediction of eGFR decline was evaluated by linear mixed modeling.

RESULTS

In univariable analyses, 9 of the 17 markers showed significant differences in median concentration between stable and fast-progressing patients. A linear mixed model for eGFR obtained by variable selection exhibited an adjusted R^2 of 62%. A panel of 12 biomarkers was selected by the procedure and accounted for 34% of the total explained variability, of which 32% was due to 5 markers. The individual contribution of each biomarker to the prediction of eGFR decline on top of clinical predictors was generally low. When included into the model, baseline eGFR exhibited the largest explained variability of eGFR decline (R^2 of 79%), and the contribution of each biomarker dropped below 1%.

CONCLUSIONS

In this longitudinal study of patients with type 2 diabetes and maintained eGFR at baseline, 12 of the 17 candidate biomarkers were associated with eGFR decline, but their predictive power was low.

The incidence of patients with type 2 diabetes is increasing worldwide, and diabetic kidney disease (DKD) is a major cause of premature disability and death. Interventions in later-stage chronic kidney disease (CKD) can only limit the damage, and thus it is necessary to risk-stratify incident patients according to their projected disease course (1,2). Unfortunately, the prediction of an individual's loss of estimated glomerular filtration rate (eGFR) based on clinical and demographic parameters is

[1]Department of Nephrology, Medical University of Vienna, Vienna, Austria
[2]Section for Clinical Biometrics, Center for Medical Statistics, Informatics and Intelligent Systems (CeMSIIS), Medical University of Vienna, Vienna, Austria
[3]Department of Internal Medicine IV (Nephrology and Hypertension), Medical University of Innsbruck, Innsbruck, Austria
[4]International Nephrology Research and Training Centre, Institute of Pathophysiology, Semmelweis University, Budapest, Hungary
[5]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, U.K.
[6]Department of Internal Medicine and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI
[7]Astellas Pharma Europe B.V., Leiden, the Netherlands
[8]Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN
[9]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, U.K.
[10]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.
[11]Computational Statistics and Machine Learning, Department of Statistics, University of Oxford, Oxford, U.K.
[12]National Institute for Health Research Oxford Biomedical Research Centre, Churchill Hospital, Oxford, U.K.

poor (3). Thus, research in the last decade focused on the discovery of molecular markers for the refinement of individual CKD progression (4).

Several candidate markers have been discovered that showed statistical associations with eGFR decline or progression of proteinuria (5). However, kidney disease in patients with type 2 diabetes is driven by a heterogeneous set of pathophysiological processes (6).

Consequently, it is unlikely that a sole marker can capture all of these different pathophysiological processes that lead to CKD progression. Investigators thus focused on parsimonious multimarker panels. Such a molecular selection was derived and experimentally tested in the European Union Seventh Framework Programme for Research project SysKid (Systems Biology toward Novel Chronic Kidney Disease Diagnosis and Treatment) (7). The biomarker panel added explained variability to a "clinical variable only" model but has not yet been thoroughly validated in an independent cohort. In addition, several other prognostic biomarkers for kidney disease progression in patients with diabetes have been identified but were never validated as a combined marker panel in a specifically designed prospective cohort.

The aim of our study was to integrate high-evidence biomarker candidates in a parsimonious panel of prognostic markers and to test their ability to predict eGFR loss when combined with commonly available clinical risk factors.

The BEAt-DKD consortium (Biomarker Enterprise to Attack DKD; http://www.imi.europa.eu/projects-results/project-factsheets/beat-dkd) was founded to identify targetable mechanisms and pathways underlying initiation and progression of DKD and to identify and validate biomarkers of disease progression and treatment responses. One of its first tasks is the validation of the best available biomarker candidates in a prospective cohort of patients with type 2 diabetes and early-stage kidney disease.

RESEARCH DESIGN AND METHODS

Biomarker Selection

Biomarkers for the current study were selected from biomarker candidates generated by the SysKid and SUMMIT (SUrrogate markers for Micro- and Macrovascular hard endpoints for Innovative diabetes Tools) consortia (8,9). We integrated diverse sources of information relevant to the relationship with DKD, including evidence from literature (4), transcriptomic analyses from microdissected renal tissue ascertained from subjects with DKD (www.nephroseq.org), whole-blood methylation profiles from patients with type 1 diabetes with and without DKD, and genetic association data. Priority biomarkers from this integration were assessed for availability of Luminex and ELISA assays and combined to maximize the number of markers that could be measured in a single sample aliquot. A list of candidates from which the current 17 markers were selected is provided in Supplementary Table 1.

Study Cohort and Selection of Study Participants

The study cohort was derived from PROVALID (PROspective cohort study in patients with type 2 diabetes mellitus for VALIDation of biomarkers), a multinational cohort study of patients with type 2 diabetes and incident or early CKD (10–12). A flowchart of patient selection is provided in Supplementary Fig. 1. There were 4,065 subjects recruited in five countries, and 2,560 subjects, from Austria, Hungary, and Scotland, were available for this study. After excluding subjects with less than 720 days of follow-up (FU), patients were grouped by CKD stage and by quintiles based on their individual eGFR slopes (Supplementary Table 2). PROVALID recruited subjects at the primary health care level, and thus the number of patients in stages G4 and G5 is low. For the remaining stages (G1 to G3b), samples in the fourth quintile (eGFR slope [−0.79, 1.39] mL/min/1.73 m^2/year) and first quintile (eGFR slope [−24.9, −5.2] mL/min/1.73 m^2/year) were deemed to be stable and fast progressors, respectively. Within the first quintile, stage G3 was underrepresented compared with the stable group and was therefore supplemented from the second quintile [−5.2, −2.58]. This selection yielded 258 patients in the stable group (median eGFR slope, 0.1 mL/min/1.73 m^2/year) and 223 patients in the fast progressors group (median eGFR slope, −6.75 mL/min/1.73 m^2/year). The two groups were closely matched for age, sex, BMI, blood pressure, and baseline eGFR. Demographics of the study population and medication details are provided in Table 1 and Supplementary Table 3, respectively.

Outcome of Interest

The outcome of interest was renal function decline over time, which was determined annually by eGFR, estimated according to the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (13).

Clinical Risk Factors

The following baseline clinical risk factors served as candidate predictors: age, sex, serum cholesterol, urine albumin-to-creatinine ratio, HbA_{1c}, mean arterial pressure, and BMI. eGFR at baseline was part of the dependent variable or included as a predictor. Because anemia does not present a problem in early-stage CKD, hemoglobin levels were omitted from the models.

Biomarker Selection and Measurement

All markers were measured in K3 EDTA plasma. A custom Human Premixed Multiplex Luminex (catalog no. CUST0I704; R&D Systems, Minneapolis, MN) was used to measure 11 markers with 1:2 sample dilution: chitinase-3-like protein 1 (CHI3L1), chemokine receptor

[13]Clinical Pharmacy and Pharmacology, Faculty of Medical Sciences, University Medical Center Groningen, Groningen, the Netherlands
[14]Department of Nephrology, Transplantation and Internal Medicine, Medical University of Silesia in Katowice, Katowice, Poland
[15]Department of Clinical Sciences in Malmö, Lund University Diabetes Centre, Lund, Sweden

Corresponding author: Rainer Oberbauer, rainer.oberbauer@meduniwien.ac.at.

Received 10 March 2018 and accepted 10 June 2018.

This article contains Supplementary Data online at http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc18-0532/-/DC1.

A.H. and M.Ka. contributed equally to this work.

*A full list of BEAt-DKD partners may be found at http://www.imi.europa.eu/projects-results/project-factsheets/beat-dkd.

© 2018 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at http://www.diabetesjournals.org/content/license.

Table 1—Baseline characteristics of the study cohort overall and stratified by stable or fast progression of eGFR decline

	Missing	Baseline value		
		Overall	Stable	Fast
Patients (*n*)		481	258	223
Age (years)	0	64 ± 9.3	64 ± 10	65 ± 9
Female sex	0	232 (47)	117 (45)	115 (50)
Smoking status (never)	0	250 (51)	134 (52)	116 (50)
Duration of diabetes (years)	0	10.8 ± 8.8	10 ± 8	12 ± 9
BMI (kg/m^2)	0	31 ± 5.5	31 ± 5	32 ± 6
Blood pressure (mmHg)				
Systolic	0	138 ± 17.4	138 ± 16	139 ± 19
Diastolic	0	79 ± 10.3	79 ± 10	79 ± 10
HbA_{1c} (%)	4 (<1)	6.8 (6.3, 7.6)	6.8 (6.3, 7.7)	6.8 (6.2, 7.6)
HbA_{1c} (mmol/mol)		51 (45, 60)	51 (45, 61)	51 (44, 60)
Hemoglobin (mmol/L)	9 (1)	8.6 (8.1, 9.3)	8.8 (8.2, 9.3)	8.5 (7.9, 9.1)
Serum glucose (mmol/L)	1 (<1)	7.4 (6.2, 9)	7.5 (6.3, 9)	7.4 (6, 8.9)
Serum cholesterol (mmol/L)	1 (<1)	4.6 (4, 5.5)	4.6 (4, 5.4)	4.6 (4, 5.6)
Serum creatinine (μmol/L)	0	77 (66, 95)	77 (67, 95)	77 (65, 95)
UACR (mg/g)	14 (3)	8.8 (4.7, 26.5)	8.2 (4.6, 21)	9.2 (5, 36.5)
Glucose-lowering agents*	0			
None		59 (12)	35 (14)	24 (11)
1–2 agents		355 (74)	192 (74)	163 (73)
>2 agents		67 (14)	31 (12)	36 (16)
Blood pressure–lowering agents†	0			
None		77 (16)	52 (20)	25 (11)
1–2 agents		195 (41)	106 (41)	89 (40)
>2 agents		209 (43)	100 (39)	109 (49)
ESA therapy‡	0	11 (2)	4 (2)	7 (3)
eGFR CKD-EPI (mL/min/1.73 m^2)	0	84 (64, 94)	85 (65, 96)	82 (63, 94)
Decline per year (mL/min/1.73 m^2/year)	0	−0.71 (−6.3, 0.2)	0.14 (−0.44, 0.68)	−6.75 (−9.04, −5.48)

Data are reported as mean ± SD, median (first quartile, third quartile), and absolute frequency (relative frequency) where appropriate. Supplementary Table 3 gives more details on medication of the study patients. None of the differences between the two groups are significant after adjusting for multiple comparisons by Holm's method (except for eGFR decline, which is the outcome of the study). We compared medians of continuous variables with Mann-Whitney *U* tests and proportions of categorical variables with χ^2 tests. ESA, erythropoietin-stimulating agents; UACR, urinary albumin-to-creatinine ratio. *Agent classes: biguanides, insulin, sulfonylureas, dipeptidyl peptidase 4 inhibitors/glucagon-like peptide 1 agonists, glinides, glitazones, α-glucosidase inhibitors, sodium–glucose cotransporter 2 inhibitors. †Agent classes: ACE inhibitors/angiotensin 2 receptor blockers, β-blockers, calcium antagonists (including direct vasodilators), α-blockers, diuretics (thiazide diuretics/loop diuretics). ‡Includes darbepoetin-α, epoetin-α, epoetin-β, epoetin-θ, epoetin-ζ, and others.

ligand 2 (CCL2), growth hormone (GH), hepatocyte growth factor (HGF), matrix metalloproteinase 1 (MMP1), matrix metalloproteinase 7 (MMP7), matrix metalloproteinase 8 (MMP8), sclerostin (SOST), tyrosine-protein kinase receptor (TIE2), tumor necrosis factor receptor 1 (TNFR1), and vascular cell adhesion molecule 1 (VCAM1). A second Human Premixed Multiplex Luminex Kit (catalog no. LXSAH-03; R&D Systems) was used to measure three markers with 1:50 sample dilution: uromodulin (UMOD), endostatin, and cystatin C. Samples were diluted using the calibrator diluent provided in the kit, processed according to the manufacturer's specification, and measured on a Luminex 200 (Luminex, Austin, TX) with xPONENT 3.1.971.0 software. Instrument settings were set according to assay protocol.

For the calibration and verification of the Luminex, a Luminex 200 Performance Verification Kit and a Calibration Kit (catalog no. 40-276 and 40-275; Merck Millipore, Billerica, MA) were used.

Kidney injury molecule-1 (KIM1) was measured by ELISA (catalog no. DSKM100; R&D Systems). Samples were diluted 1:2, processed according to assay procedure, and measured on a TriStar^2 LB 942 Modular Multimode Microplate Reader (Berthold Technologies, Bad Wildbad, Germany) using wavelength settings as instructed in the assay procedure. Absolute concentrations were determined using MikroWin 2010 v5.21 software (Berthold Technologies).

For quality control, pooled normal human plasma K3 EDTA (catalog no. IPLA-N-100ml-K3 EDTA; Innovative Research, Novi, MI) was spiked with recombinant proteins (R&D Systems) to create low-, medium-, and high-level controls.

All samples were measured as two technical replicates and required to have a coefficient of variation (%CV) below 12%. In addition, 10% of all samples were remeasured on a different plate to perform U.S. Food and Drug Administration (FDA)–recommended incurred sample reanalysis. More than 72% of incurred sample reanalysis showed a percentage difference below 20%, and measurements were therefore in concordance with FDA guidelines (Supplementary Table 4). Values out of quantifiable range were set to 0.5 and 1.5 times the lower and upper quantification limits, respectively (Supplementary Table 5 and Supplementary Fig. 2).

Fibroblast growth factor 23 (FGF23) was measured using a FGF-23 (C-Term)

ELISA (catalog no. 60-610; Quidel, San Diego, CA) with a dilution of 1:2 according to the manufacturer's recommendations. Signal was measured using EnVision plate reader (PerkinElmer, Waltham, MA) using the optical density wavelength instructed by the procedure. N-terminal prohormone of brain natriuretic peptide (NTproBNP) was measured using an NTproBNP ELISA (catalog no. K151JKC; Mesoscale Discovery, Gaithersburg, MD) with a dilution of 1:10. Samples were processed according to the manufacturer's recommendations, and electrochemiluminescence signal was measured on a MESO QuickPlex SQ 120 (Mesoscale Discovery). Samples were measured in technical replicates. Interassay %CV was estimated using internal controls. Due to limited available sample volume, NTproBNP and FGF23 concentrations were determined in only 480 and 437 of the 481 samples, respectively, and FGF23 concentration could only be determined in single measurements for 86 samples. The %CV were required to be less than 20% and average inter- and intra-assay %CV for FGF23 and NTproBNP were 5.5 and 3.2 and 13.2 and 7.0, respectively.

Samples Size Estimate

Because all 17 biomarkers were preselected from previous projects, we estimated that 500 samples would be sufficient to reach more than 80% power to detect at least a single biomarker with a statistically significant effect on the outcome renal function decline (Supplementary Fig. 4). Key assumptions were derived from Mayer et al. (7).

Statistical Analysis

Patient characteristics are described by mean and SD or median and first and third quartile for continuous variables or by frequency and percentage for binary variables. Biomarker levels between the stable and fast-progressing patient groups were compared by Mann-Whitney *U* tests. To estimate the effects of clinical risk factors and protein biomarkers on the outcome, we used univariable and multivariable linear mixed models. eGFR baseline measurements take on a special role in such analysis because they can be understood as part of the outcome or as a clinical covariate. Our main goal was to validate the biomarkers as predictors for renal function decline; therefore, eGFR levels at baseline were considered as part of the outcome and thus included in the dependent variable because they are subject to the same random variation as later values. To compare the contribution of biomarkers and baseline eGFR to the prediction of future eGFR levels, we repeated the same modeling procedure as described in the following but added baseline eGFR to the set of covariates and removed it from the dependent variable. Random intercepts and random slopes were used to model the patient-specific eGFR trajectories, imposing no restrictions on their covariance. Interaction terms with time were included to model the effect on the eGFR slope.

Results are reported as coefficients and associated *P* values. The baseline coefficient (main effect) can be interpreted as association with mean eGFR levels and the slope coefficient (interaction effect) as association with the eGFR change over time. We investigated the importance of predictors by applying backward elimination based on the Akaike information criterion (AIC) on a model containing all protein biomarkers and clinical predictors. Hierarchy of interactions and main effects were kept intact: a baseline effect was only dropped if no associated slope effect was present in the model.

The adjusted R^2 of the fixed effects part was obtained by multiplying the unadjusted R^2 with a correction factor of $(N - K - 1)/(N - 1)$, where N and K denote the number of patients and the number of fixed effects in the model, respectively. To further assess the contribution of specific covariates to the prediction of the outcome, we decomposed the adjusted R^2 by computing the drop in R^2 when excluding a specific covariate from the model and scaling the resulting values to add up to the total adjusted model R^2.

Biomarker levels were log2 transformed to normalize their distributions. The model results presented here were pooled by Rubin's rules from multiply imputed data sets to account for uncertainty due to missing data in predictors. Thus, all 481 samples were included in each model. By applying the variable selection procedure to each imputation, we obtained selection frequencies facilitating assessment of model instability due to missing data. Our final model comprised predictors chosen in at least half of the imputations. Model instability due to general sampling variation was assessed by drawing bootstrap resamples in each imputed data set.

Complete-case-only analyses, the number of available samples per predictor, and a description of the multiple imputation procedure is provided in the Supplementary Data.

Logistic regression models were applied to obtain classification models for progression status based on the predictor values.

P values of <0.05 were considered statistically significant, and all *P* values are two-sided. We used R statistical software (https://www.r-project.org/foundation, Vienna, Austria) for all analyses.

RESULTS

In total, 481 patient baseline plasma samples were measured for 17 biomarkers. A detailed breakdown of availability by marker is provided in Supplementary Table 6. Figure 1 shows biomarker levels grouped by speed of progression of renal function decline. When median levels were compared using Mann-Whitney *U* tests, several markers showed a significant difference between the two groups of patients (no adjustment for multiple testing). However, the marker level distributions overlapped, and the observed differences were small, with an average of 25%. Nonparametric Spearman correlation coefficients between biomarkers and clinical data are visualized as a heat map in Supplementary Fig. 3.

Supplementary Table 7 reports the results from univariable mixed model analysis. Reported baseline coefficients correspond to the prediction of mean eGFR levels at $t = 0$ and slope coefficients to the prediction of the change of eGFR levels over time. Except for SOST, MMP8, and CCL2, all markers show a significant association with eGFR levels via the baseline coefficients. Most of these associations remain significant, even after adjustment for clinical risk factors (Supplementary Table 8). Only few biomarkers show a significant association with eGFR change over time via their slope coefficients.

MMP8 (66% of samples below the limit of quantification) and SOST (30% missing) were removed from all further multivariable analyses due to the high

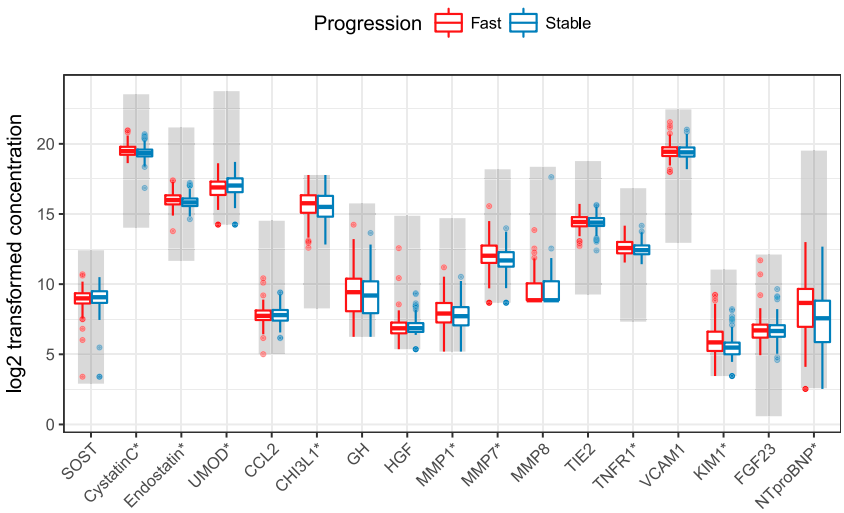

Figure 1—Concentration levels (log2 transformed) for all 17 biomarkers grouped by speed of progression of renal function decline. All concentrations are given in pg/mL except for FGF23, which is given in relative units. The quantifiable range of assays is indicated by areas shaded in gray. *These biomarkers show a significant difference in median levels between the two groups (Mann-Whitney *U* test, no *P*-value adjustment for multiple testing).

number of missing values and no apparent association with eGFR decline in univariable analysis. The results from the final model obtained from backward elimination are reported in Table 2. Overall, the biomarker predictors account for 34.4% of the explained variability of eGFR levels. According to their contribution to explained variability, the biomarkers can be split into two groups (Supplementary Table 9). Biomarkers in the first group are mainly useful for prediction of mean eGFR values via their baseline coefficients, and their effect stays constant over time. The second group also adds to the prediction of eGFR change over time via their slope coefficients. However, the contribution of the second group to the model R^2 is comparatively small (30.9% for group one and 3.5% for group two), indicating that the biomarkers contribute primarily through the prediction of mean eGFR levels. Figure 2 shows the model fit via predicted median eGFR trajectories for stable and fast-progressing patients from the final mixed model together with observed eGFR distributions at each FU visit. An additional bootstrap procedure shows that almost all predictors from our final model are selected with high frequency, indicating satisfactory model stability (Supplementary Table 10). Furthermore, multivariable models without variable selection including clinical covariates only and biomarkers on top of clinical covariates give an indication of the achievable predictive performance in this cohort (Supplementary Tables 11 and 12).

The weak association with eGFR slopes is further demonstrated by low discriminative power when using logistic regression models to discriminate between stable and fast-progressing patients. The resulting low area under the curve values corroborate our findings that the biomarkers are mainly associated with mean eGFR baseline levels (Supplementary Tables 13 and 14).

Results from the analysis including baseline eGFR levels as a covariate further indicate that the added value of the biomarkers on top of these measurements is low for the prediction of future eGFR levels. In corresponding univariable analyses, only six biomarkers remained significantly associated with mean eGFR levels when baseline eGFR was included in the models (Supplementary Table 15).

The final model selected by backward elimination from the pool of candidate predictors including baseline eGFR is provided in Supplementary Table 16. Compared with the model without baseline eGFR as a covariate, a further four biomarkers are eliminated from the model; otherwise, biomarker selection remains similar (Supplementary Table 17). The dominating influence of baseline eGFR levels on prediction of eGFR levels after baseline is expressed by their high adjusted R^2 measure, which is essentially equal to the model's R^2—dropping any other predictor leaves the model's

Table 2—Multivariable linear mixed model for prediction of eGFR levels (with baseline eGFR as part of the dependent variable) obtained from AIC-based backward elimination on all candidate predictors (log2 transformed biomarker and clinical)

	Baseline		Slope		
Predictor	Coefficient	*P* value	Coefficient	*P* value	R^2 decomposition
Constant	407.630	<0.001	−2.102	0.396	—
Cystatin C	−11.661	<0.001	n.s.	n.s.	9
Endostatin	−2.957	0.133	n.s.	n.s.	<1
UMOD	2.977	<0.001	n.s.	n.s.	4.3
CHI3L1	1.124	0.037	n.s.	n.s.	<1
HGF	0.047	0.949	0.463	0.044	<1
MMP1	0.731	0.187	−0.307	0.082	<1
MMP7	−1.233	0.064	n.s.	n.s.	<1
TIE2	4.622	<0.001	n.s.	n.s.	3.3
TNFR1	−10.888	<0.001	n.s.	n.s.	12.9
KIM1	−0.064	0.922	−1.084	<0.001	3
FGF23	−0.983	0.269	0.456	0.086	<1
NTproBNP	0.071	0.793	−0.253	0.006	<1
Age (years)	−0.684	<0.001	0.048	0.039	27
Current or former smoker	2.460	0.024	n.s.	n.s.	<1
Mean arterial pressure	0.085	0.087	n.s.	n.s.	<1
Total cholesterol	−0.762	0.096	n.s.	n.s.	<1

The model had an adjusted R^2 of 62.5%. Biomarkers had a total contribution of 34.4%, and clinical risk factors had a total contribution of 28.1%. The decomposition of the model R^2 combines the contributions of baseline and slope coefficients for each predictor. n.s., not selected.

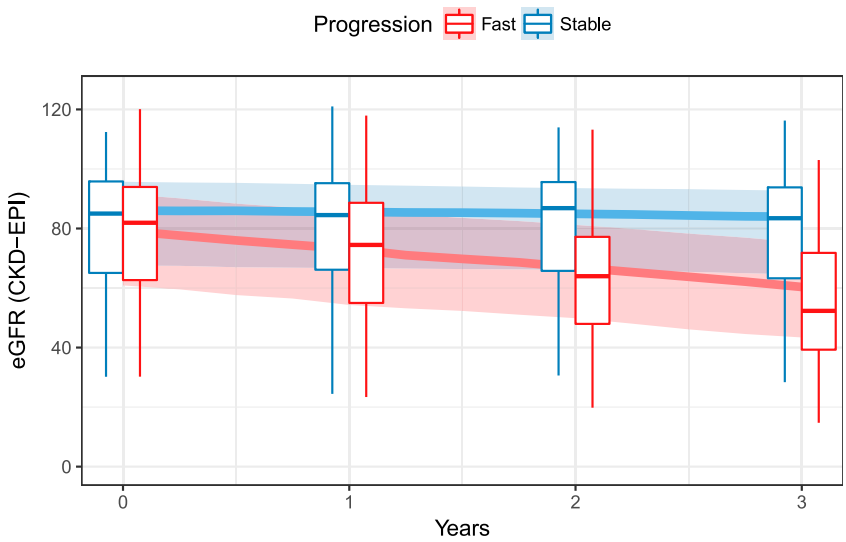

Figure 2—Predicted median eGFR trajectories from the multivariable linear mixed model for eGFR levels (with baseline eGFR as part of the dependent variable) obtained by AIC-based backward elimination (solid line). Shaded areas indicate the interquartile range of predictions. Superimposed box plots show the observed values summarized at each yearly FU visit.

predictions virtually unchanged. However, even for baseline eGFR levels, the predictive power is mostly due to prediction of mean eGFR levels after baseline rather than through the prediction of the eGFR slope.

Corresponding complete-case-only analyses and the number of available samples per predictor are provided in Supplementary Tables 18 and 19. The results remained largely unchanged, with an adjusted R^2 of 62% of the mixed model for eGFR prediction, thus supporting the validity of the results of the multiply imputed analysis.

CONCLUSIONS

We performed a validation study of 17 preselected plasma protein markers with reported high evidence for the prediction of eGFR decline in patients with type 2 diabetes and incident or early-stage CKD. We showed in an univariable analysis that 9 of the markers had significantly different concentration levels between patients with stable eGFR and fast progression of eGFR decline and that 14 biomarkers significantly contributed to the prediction of eGFR levels. However, most of the predictive ability was attributable to the association with baseline GFR. In the multivariable analysis of eGFR decline over time, only five markers (KIM1, FGF23, NTproBNP, HGF, and MMP1) remained significant but exhibited only a modest predictive power on top of clinical covariates. Furthermore, if the longitudinal analysis was adjusted for baseline eGFR, none of the biomarkers were able to contribute a relevant portion of explained variability, suggesting that baseline eGFR is the key variable in prediction of renal function at a very early CKD stage. Interestingly, urinary albumin excretion was only of limited value for predicting eGFR loss, which may be explained by the minute amount of albuminuria in the patients investigated.

Other studies of patients with type 2 diabetes and incident or early-stage CKD have shown that even well-established clinical risk factors of later-stage disease do not perform well in discriminating progressors at early stages. Dunkler et al. (14) showed by using only clinical variables that the discrimination for the progression of CKD in patients with type 2 diabetes is actually very low on an individual basis. eGFR and to some extent albuminuria were the most important factors for predicting progression, but their predictive ability in total was modest.

Niewczas et al. (15) showed that, on top of clinical covariates, elevated concentrations of serum TNFR1 in 410 patients with later-stage disease and long-term FU was strongly associated with baseline GFR and predicted end-stage renal disease (ESRD) that happened in 59 patients after a median FU of 12 years. This is in line with our current and previous findings that TNFR1 exhibited the highest explained variability in the longitudinal analysis for eGFR loss. However, death as a competing risk factor in analysis of the progression of kidney disease needs to be considered here. Differences in lead-time bias between studies can possibly explain the discrepancies in biomarker prediction of eGFR decline between the present and other studies (15). For example, patients who reached ESRD in the Niewczas et al. study exhibited macroalbuminuria of 623 μg/min already at baseline compared with very low-grade albuminuria of 20 μg/min in patients without progression. In addition, supporting the argument of a lead time is that patients who progressed to ESRD over 12 years exhibited an eGFR reduction by almost half, to 61 mL/min/1.73 m^2 at baseline, compared with nonprogressors.

A multinational consortium from France investigated serum TNFR1 as predictor of eGFR slope in 522 patients (median FU of 4 years) with type 2 diabetes at later CKD stages (i.e., albuminuria of more than 30 mg/mmol creatinine) (16). The investigators applied logistic regression and found a statistically significant increase of ESRD discrimination in each TNFR1 baseline quartile, but parameter estimates for a multivariable model with clinical covariates were missing.

Similar to our previous study and in Saulnier et al. (17), we used linear mixed models to separate the marker contributions into association with baseline eGFR and eGFR change over time (7). Our finding of higher levels of TNFR1 and NTproBNP in the fast progressor group is in line with the association with renal function loss observed by Saulnier et al. (17).

Ban et al. (18) reported serum MMP7 association with proteinuria and GFR in a cross-sectional analysis.

In 2017, a group from Denver (U.S.) investigated the test characteristics of selected plasma biomarkers for predicting eGFR $<$60 mL/min/1.73 m^2 and urinary albumin-to-creatinine ratio $\geq$30 mg/g in patients with type 1 diabetes using principal component analysis and Cox proportional hazards models (19). The main finding was that after adjustment for traditional risk factors, only KIM1 and cystatin C exhibited a significant but modest improvement in discrimination.

The principal component holding the most promising markers increased the area under the curve by only 2%. These results are well in line with our finding that candidate biomarkers may not be useful for eGFR slope prediction in patients with type 2 diabetes.

Recently, Garlo et al. (20) observed results similar to our findings in their biomarker evaluation in more than 5,000 enrolled patients with type 2 diabetes. eGFR decline occurred in 98 patients over a median of 1.5 years. Established markers such as cystatin C or biomarkers of tubular injury did not substantially improve the prediction of eGFR loss on top of clinical predictors.

Our previous analysis of nine biomarkers in two cohorts of patients with different baseline eGFR showed that explained variability of eGFR loss in patients with eGFR $<$60 mL/min/1.73 m^2 was mainly driven by MMP7 and TNFR1 (7). In patients with baseline eGFR $\geq$60 mL/min/1.73 m^2, contribution of all markers was modest, with an adjusted R^2 of 15% and 35% for a combination of biomarkers and clinical predictors. The inclusion of a further eight well-investigated biomarkers did not substantially increase the predictability of eGFR loss in our current analysis. However, 9 of the 17 markers showed statistically significant differences in concentration levels between the group of patients with stable kidney function and the group with fast renal function decline, which supports the initial marker selection for this study. Yet, our analysis showed that the main contribution of these biomarkers is their association with baseline eGFR values rather than eGFR slopes.

The individual slope of eGFR loss is highly variable in patients with diabetes and may be modified by medication. However, the aim of the current study was to predict the slope from baseline biomarkers independently of subsequent interventions such as comedication, lifestyle changes, or any other factors. Therefore, we did not use medication in our main model on purpose, because it would be a baseline adjustment for interventions that occurred afterward, hence using information not available at time of prediction. In addition, all patients in the PROVALID study were optimally treated according to guidelines for patients with diabetes (21,22). An analysis including treatment status at baseline, corroborating the negligible effect of medication on the performance of the biomarkers, can be found in Supplementary Tables 20 and 21.

A key strength of PROVALID is that the study was specifically designed for the validation of biomarkers in patients with type 2 diabetes (10). However, our study has a few limitations. The selection of patients from the PROVALID cohort was based on the outcome of eGFR. Although this likely leads to overoptimistic results, our rationale was to maximize power to validate the utility of the candidate biomarkers to predict eGFR loss independently of clinical parameters. This limitation is irrelevant, however, because performance was poor in this cohort of early-stage CKD patients even with the preselection marker. A potential limitation is the relative short FU of 3 years. However, all patients had baseline and annual eGFR determinations, which led to a robust slope estimation and thus stable marker performance estimates.

The strengths of our study are the careful analysis of biomarkers according to European Medicines Agency and FDA standards in a multinational prospective study (http://academy.gmp-compliance.org/guidemgr/files/UCM368107.PDF). Sample aliquots were stored at -80°C immediately after collection and never thawed until analysis. The percentage difference between reruns was well within the FDA recommended range. A further asset is the thorough statistical analysis in which we dissociated the effects of markers on the prediction of baseline eGFR and slope alone and in combination with clinical covariates known to be key risk factors for CKD progression.

In conclusion, the prediction of eGFR slope using baseline circulating biomarkers in combination with clinical parameters was modest. Most of the predictive power was generated by the association of markers with baseline eGFR, which was by far the strongest predictor of future eGFR levels. Given the inferior performance of this highly selected set of biomarkers in early-stage CKD patients to predict future eGFR loss, these markers are not likely to be useful for clinical decision making. Nevertheless, their assessment might be useful to identify individual biological processes that may contribute to the progression of very early-stage renal disease.

Funding. This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement no. 115974. This Joint Undertaking receives support from the European Union's Horizon 2020 Research and Innovation Programme and European Federation of Pharmaceutical Industries and Associations. This work was partially supported by the National Institute of Diabetes and Digestive and Kidney Diseases (P30DK081943 to M.Kr.). M.I.M. is a Wellcome Senior Investigator supported by Wellcome grants 090532, 098381, and 203141. This work was also supported by JDRF award 2-SRA-2014-276-Q-R and by Innovative Medicines Initiative funding to the SUMMIT consortium.

The funding sources did not have a role in the design, conduct, or analysis of the study.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. A.H. supervised data acquisition and contributed to interpretation of data and writing of the manuscript. A.H., M.Ka., G.M., R.R.-S., K.H., P.P., S.E., L.R., P.B.M., W.J., M.Kr., P.G., J.M.W., K.L.D., M.A., M.I.M., G.H., H.L.H., A.W., M.F.G., and R.O. revised the manuscript for important intellectual content. M.Ka. performed data processing and analysis and contributed to writing of the manuscript. G.M. conceived the design of the study cohort and contributed data. R.R.-S. contributed to analysis of the data. K.H. performed most laboratory measurements and contributed to data quality control. P.P. contributed to study design and writing of the manuscript. S.E. contributed to data acquisition. J.M.W. and K.L.D. contributed additional laboratory measurements. M.A. contributed to initial biomarker selection. G.H. advised on the statistical analysis plan. R.O. wrote the draft of the manuscript. R.O. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Wanner C, Inzucchi SE, Lachin JM, et al.; EMPA-REG OUTCOME Investigators. Empagliflozin and progression of kidney disease in type 2 diabetes. N Engl J Med 2016;375:323–334
2. Mann JFE, Ørsted DD, Brown-Frandsen K, et al.; LEADER Steering Committee and Investigators. Liraglutide and renal outcomes in type 2 diabetes. N Engl J Med 2017;377:839–848
3. Hallan SI, Ritz E, Lydersen S, Romundstad S, Kvenild K, Orth SR. Combining GFR and albuminuria to classify CKD improves prediction of ESRD. J Am Soc Nephrol 2009;20:1069–1077
4. Heinzel A, Mühlberger I, Stelzer G, et al. Molecular disease presentation in diabetic nephropathy. Nephrol Dial Transplant 2015;30 (Suppl. 4):iv17–iv25
5. Looker HC, Colombo M, Hess S, et al.; SUMMIT Investigators. Biomarkers of rapid chronic kidney disease progression in type 2 diabetes. Kidney Int 2015;88:888–896
6. Tan J, Zwi LJ, Collins JF, Marshall MR, Cundy T. Presentation, pathology and prognosis of renal

disease in type 2 diabetes. BMJ Open Diabetes Res Care 2017;5:e000412
7. Mayer G, Heerspink HJ, Aschauer C, et al.; SysKid Consortium. Systems biology-derived biomarkers to predict progression of renal function decline in type 2 diabetes. Diabetes Care 2017;40:391–397
8. Martini S, Nair V, Keller BJ, et al.; European Renal cDNA Bank; C-PROBE Cohort; CKDGen Consortium. Integrative biology identifies shared transcriptional networks in CKD. J Am Soc Nephrol 2014;25:2559–2572
9. Sambo F, Trifoglio E, Di Camillo B, Toffolo GM, Cobelli C. Bag of Naïve Bayes: biomarker selection and classification from genome-wide SNP data. BMC Bioinformatics 2012;13(Suppl. 14):S2
10. Mayer G, Eder S, Rosivall L, et al. Baseline data from the multinational prospective cohort study for validation of biomarkers (PROVALID). Nephrol Dial Transplant 2016;31(Suppl. 1):i482
11. Dienemann T, Fujii N, Orlandi P, et al. International Network of Chronic Kidney Disease cohort studies (iNET-CKD): a global network of chronic kidney disease cohorts. BMC Nephrol 2016;17:121
12. Eder S, Leierer J, Kerschbaum J, et al. A prospective cohort study in patients with type 2 diabetes mellitus for validation of biomarkers (PROVALID) - study design and baseline characteristics. Kidney Blood Press Res 2018;43:181–190
13. Levey AS, Stevens LA, Schmid CH, et al.; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604–612
14. Dunkler D, Gao P, Lee SF, et al.; ONTARGET and ORIGIN Investigators. Risk prediction for early CKD in type 2 diabetes. Clin J Am Soc Nephrol 2015;10:1371–1379
15. Niewczas MA, Gohda T, Skupien J, et al. Circulating TNF receptors 1 and 2 predict ESRD in type 2 diabetes. J Am Soc Nephrol 2012;23:507–515
16. Saulnier PJ, Gand E, Ragot S, et al.; SURDIAGENE Study Group. Association of serum concentration of TNFR1 with all-cause mortality in patients with type 2 diabetes and chronic kidney disease: follow-up of the SURDIAGENE cohort. Diabetes Care 2014;37:1425–1431
17. Saulnier P-J, Gand E, Velho G, et al.; SURDIAGENE Study Group. Association of circulating biomarkers (adrenomedullin, TNFR1, and NT-proBNP) with renal function decline in patients with type 2 diabetes: a French prospective cohort. Diabetes Care 2017;40:367–374
18. Ban CR, Twigg SM, Franjic B, et al. Serum MMP-7 is increased in diabetic renal disease and diabetic diastolic dysfunction. Diabetes Res Clin Pract 2010;87:335–341
19. Bjornstad P, Pyle L, Cherney DZI, et al. Plasma biomarkers improve prediction of diabetic kidney disease in adults with type 1 diabetes over a 12-year follow-up: CACTI study. Nephrol Dial Transplant 2018;33:1189–1196
20. Garlo KG, White WB, Bakris GL, et al. Kidney biomarkers and decline in eGFR in patients with type 2 diabetes. Clin J Am Soc Nephrol 2018;13:398–405
21. Qaseem A, Barry MJ, Humphrey LL, Forciea MA; Clinical Guidelines Committee of the American College of Physicians. Oral pharmacologic treatment of type 2 diabetes mellitus: a clinical practice guideline update from the American College of Physicians. Ann Intern Med 2017;166:279–290
22. National Kidney Foundation. KDOQI clinical practice guideline for diabetes and CKD: 2012 update. Am J Kidney Dis 2012;60:850–886