

## **Response to reviewers**

We thank the reviewers for the careful examination of our work, their positive summaries about the presented study and the constructive critiques provided in their comments. Please find our response to each comment in the table on the subsequent pages.

Reviewer comment	Author response
Reviewer 1	
<b>R1 C1:</b> Baseline table, given the case control status I would suggest stratifying the results according to stable and progressors (so use supl table 3 in main manuscript)	We thank the expert reviewer for the suggestion and agree that baseline characteristics stratified by stable and progressors are of interest to the reader. Therefore, we have added two separate columns for the stable and progressor groups to table 1 in the main manuscript. Since the main analysis in the manuscript makes use of actual eGFR slopes and not the two groups of patients we think that the reader will also benefit from presentation of the full cohort. Therefore, we hope the reviewer agrees with our choice to also include the characteristics for the full cohort as an additional column.
<b>R1 C2:</b> I am missing essential covariates in the baseline table, use of bloodpressure-lowering medication in particular	We thank the reviewer for pointing this out and have now included the medication status at baseline in table 1 and table S3.
<b>R1 C3:</b> Can the authors comment on the duration of follow-up? Do they think the markers would perform better at a longer FU time?	A longer follow up would stabilize the estimation of eGFR slopes but on the other hand during longer follow up more parameters not captured at baseline could potentially affect the eGFR slopes. Therefore, the marker performance would likely not improve with longer follow up.
<b>R1 C4:</b> I would suggest to perform the analysis with a hierarchical model of conventional predictors. I understand the current approach but downside of this is that the selection of markers that remain in the model is specific to this dataset, and since no external validation has been performed this makes the current results somewhat difficult to interpret and compare to other papers doing similar work.	As suggested, we performed two additional analyses by first fitting a model using only conventional predictors and markers on top of this model. The adjusted R2 is slightly lower than of the model from backward selection (see table S11 and S12). To investigate the stability of the selection we performed a bootstrap which is displayed in table S13. As expected, the strong predictors are stably selected while for some weaker predictors their role is more uncertain.
<b>R1 C5:</b> Can the authors comment on the clinical relevance of	We thank the reviewer for this question. As indicated in our

<p>the work, do they think the predictive value of the current markers holds true accros the full range of eGFR?</p>	<p>conclusions statement the clinical utility of the tested biomarkers is low in patients with maintained baseline eGFR.</p> <p>However, we already tested a similar set of markers in a later stage cohort (baseline eGFR &lt; 60ml/min, mean 37ml/min) and observed a higher explained variability. However, R2 was mainly driven by MMP7 and TNFR1 (Diabetes Care 2017). To address this point specifically we extended the paragraph in the main manuscript (page 19) to include also results from the group of patients with reduced baseline eGFR (mean 37ml/min SD 11ml/min/) from the previous study.</p>
<p>Reviwer 2</p>	
<p><b>R2 C1:</b> The authors do touch on possible treatment effects, but dismiss this as irrelevant because ‘all patients..... were optimally treated according to guidelines’. It would be helpful to include in Table 1 the proportions on ACEi or ARB therapy, and whether such therapy was added during the study. Initiation of ACEi or ARB treatment can result in a fairly acute drop in eGFR.</p>	<p>We thank the expert reviewer for the suggestion and added information on blood pressure, glucose medication as well as the use of erythropoietin stimulating agents to the manuscript. In order to keep the demographics table in the main manuscript concise only the number of blood pressure lowering agents, the number of glucose lowering drugs and use of ESA were added. Detailed information on the use of specific drugs has been included into table S3.</p> <p>In addition, we investigated as suggested how many patients were started on ACEi and ARB therapy during the study and didn’t find any significant difference between the stable and the progressor group (29% and 30% respectively). We did not, however, include information on treatment changes during follow-up in the models as we aimed at predicting change in eGFR using information available at baseline.</p>
<p><b>R2 C2:</b> Table 1 would also benefit from having two extra columns with the baseline data from the ‘stable’ and ‘progressive decline groups, respectively (i.e. combined with Supp Table 3). The glucose and cholesterol values are</p>	<p>We thank the reviewer for pointing this out and do agree that the reader will benefit from having both the overall demographics as well as demographics for each group available in table 1. Therefore, we have merged the separate columns for</p>

<p>presumably given in SI units - so it should be 'mmol/l' not mmol/dl, surely? And why not give the creatinine results in SI units, too?</p>	<p>progressors and stable patients from original table S3 into table 1.</p> <p>We are also very grateful that the reviewer pointed out our typo in units for glucose and cholesterol and welcome the suggestion to also give creatinine concentrations in SI units. We corrected the typo in mmol/l in table 1 and converted the creatinine measurements from mg/dl to <math>\mu</math>mol/l.</p>
<p><b>R2 C3:</b> It would be important to mention that renal disease in type 2 diabetes is a complex phenomenon with many potential disease processes going on (see, for example, BMJ Open Diab Res Care 2017;5:e000412). Is it reasonable to expect that any single marker will accurately reflect all such pathologies?</p>	<p>We thank the reviewer for the comment and the suggestion of an appropriate reference. We rephrased the specific part of the second paragraph in the introduction to put more emphasize on this issue and included the reference.</p>
<p>Reviewer 3</p>	
<p><b>R3 C1:</b> The authors have undertaken an interesting approach, using a variety of potential biomarkers of renal disease in a cohort of persons with type 2 diabetes, but report an essentially negative set of findings, in as much as the major predictor of low eGFR was the baseline eGFR, with the biomarkers adding little to this finding. Furthermore, although some clinically important measurements appear to have been included in the model, it does not appear that the authors have taken into account the potentially differing effects of various glucose-lowering agents, as there certainly is some evidence of benefit of the dpp4i, glp-1ra, sglt2i, and TZD classes, and similarly potentially differing effects of various BP-lowering agents (ACE/ARB vs diuretics etc) do not seem to be factored in the model.</p>	<p>We thank the expert reviewer for pointing out that various glucose medications and blood pressure lowering agents may have different effects on the outcome parameter and agree that inclusion of the medication status at baseline will strengthen the manuscript. While we initially didn't account for medication at baseline as all patients were treated according to clinical guidelines and as such no differences between the groups were to be expected we have now added information on blood pressure and glucose lowering medication as well as the use of erythropoietin stimulating agents to the baseline demographics table (table 1 and table S3). In addition, we included the medication status into the multivariable models. The majority of patients received one or more of the following drugs biguanides (72%), sulfonylureas (28%) insulines (25%), DDP4-anta/GLP1 analogs (20%) and only 5% were treated with thiazolidinediones and none received SGLT2-inhibitors (Table S3). We included the number of orally administered glucose lowering</p>

	<p>medications, insulin and the number of blood pressure medications into the models. (table S20 and S21). The model remained virtually unchanged with the exception that antidiabetic medication was selected into the model and blood pressure medication replaced MAP as predictor.</p>
<p><b>R3 C2:</b> Anemia is another clinical parameter which the authors omit, but there may be a relationship between hemoglobin level and renal outcome. In any event, fig 1 shows tremendous overlap between progressors and non-progressors, even in the few biomarkers for which there is a difference in means that is statistically significant (without adjustment for multiple testing).</p>	<p>We thank the reviewer for pointing out the importance of anemia in DKD and its potential impact on the outcome parameter. We initially omitted hemoglobin levels and the use erythropoietin stimulating agents from the manuscript as we did not expect any anemia in this very early stage of DKD. Nevertheless, we agree with the reviewer that it still is important to statistically assess the relationship with hemoglobin levels and the eGFR decline. Therefore, we made the following changes to the manuscript. Hemoglobin levels and use of ESA were included into the baseline demographics (table 1 and table S3). In addition, the hemoglobin level was included as clinical co-variate into the multivariable models. (table S20 and S21) Use of ESA therapy was not included into the multivariable models as the majority (98%) of study participants didn't receive any ESA treatment.</p> <p>We agree with the expert reviewer's observation that there exists a tremendous overlap between the groups (hence that the absolute difference of markers between groups is minimal) and thus we extended the sentence in the result section about the small differences between groups to also explicitly state this fact. (see page 13)</p>
<p><b>R3 C3:</b> Another issue is the loss of sample values in the measurement procedure. Certainly for MMP8 most of the results were unmeasurable, and the authors correctly decided to remove this from subsequent analysis, but table S6 shows that three more of the values have &gt;100 lost measurements, and one has 84 lost measurements – imputation for those four variables might be questionable, and in general this reviewer is uncomfortable with imputing values for variables in a study of</p>	<p>We thank the expert reviewer for this comment and agree that handling of failed measurements / missing values is crucial. We appreciate that the reviewer agrees with our decision to exclude MMP8 from all further analysis due to the high number of missing values. We hope that the reviewer also approves that we did the same for SOST which was only available for a subset of the samples. Regarding the number of lost measurements in table S6 we would like to point out that this table did only hold</p>

<p>this sort, where the basic notion is that a sturdy set of biomeasures are desired to predict the outcome - here of renal function worsening.</p>	<p>information on individual measurements including multiple failed attempts for the same sample. (We now expanded the table to provide a better overview of how many samples were actually lost due to measurement issues.) We would like to direct the reviewer's attention to table S5 in which the number of samples with an available biomarker measurement are given. Nevertheless, also on this level the three markers (ntProBNP, HGF and GH) have close to 100 missing samples. However, after truncation of values below and above detection limit 10% of samples were missing for the marker with the highest missingness (NTProBNP). At this note we would like to mention that we realized after careful consideration of the reviewers comment that the wording "imputation of out of range values" was misleading in this context since values above and below detection limit were included as .5 * min or 1.5 max values (left and right truncated values), respectively. Thus, we changed the wording throughout the manuscript. Furthermore, we want to point out that we adhered to strict requirements during the measurement (in terms of %CV for replicates) which in part may contributed to higher rates of missingness and that the approach of multiple imputation was employed to account for the additional uncertainty resulting from the missing values. However, we in addition also provide corresponding complete case only analysis for the main analysis in the manuscript (Table S18 and S19).</p>
<p><b>R3 C4:</b> A further question from table S6 is that the "total" column is in 3 cases more than the sum of the "%CV between duplicates &gt;11.2" and the "Out of quantifiable range" columns – for UMOD, TIE2, and TNFR1. Perhaps the authors need to proofread this table more carefully!</p>	<p>We thank the reviewer for pointing out this apparent discrepancy. We would like to assure the reviewer that the numbers in the table are correct but must admit that crucial information was lacking in the table caption. Table S6 includes detailed numbers for the two main reasons (out of range and high CV%) of sample failure, but there were also a small number of instances were a measurement failed due to some other technical reasons like not enough beads or a cloaked filter plate. While in such cases it was not possible to check quantifiable range or CV% the measurement still had to be</p>

	<p>counted as failed measurement.</p> <p>In consequence we have extended table S6 and the corresponding table caption to report the number of measurements failed due to other technical issues.</p>
<p><b>R3 C5:</b> A final problem pertains to the choice of the investigators to use persons with diabetes and relatively intact renal function. While the notion that finding predictors of early kidney disease might be of greater value than predictors for persons already having decreased eGFR, the finding that only 223/2560 persons in the cohort were rapid-decliners might itself lead the statistical power of the analysis to be intrinsically low, so that including more persons with stage 3 (and perhaps stage 4) CKD to observe the effects of the biomarkers in that population would have been useful.</p>	<p>As stated in the manuscript we were specifically interested to test markers in incident and early prevalent kidney disease because a potential therapeutic intervention is likely most effective in early stage disease patients.</p> <p>The 223 fast progressors were selected from the 1549 patients with more than 720 days follow up which were selected for the study (figure S1). In PROVALID being a cohort of T2D patients with maintained eGFR the majority (80%) of patients are G1/G2.</p>
Reviewer 4	
<p><b>R4 Maj1:</b> One of the key question in the present work is the selection of the studied biomarkers. The authors claim to have a validation strategy but do not consider several papers which previously examined the same research question (such as but not limited to Peters et al Diabetes Care 2017, Saulnier et al Diabetes Care 2017). It is easy to understand that the strategy was preplanned but the inclusion of the evidence from existing literature must clearly be taken into account and additional biomarkers added in the set studied here.</p>	<p>We thank the expert reviewer for the suggestion of literature describing further marker candidates. As described in the methods section we focused our validation specifically on candidates generated by the SYSKID and SUMMIT consortia and expanded the set with diverse sources of information. One of the key strategies employed was building molecular models based on omics data sets and diabetic nephropathy associated genes/proteins from literature for subsequent biomarker candidate selection. In this approach NCBI MeSH and gene2pubmed annotation was exploited and text mining techniques were employed to systematically record disease gene association available from articles in pubmed/medline. Please see also ref 4 in the main manuscript.</p> <p>Regarding the work performed by Peters et al. The manuscript reporting the results was published (epub) in August 2017 and</p>

	<p>at this time we had completed the marker selection process and already started measuring the samples.</p> <p>Regarding the work of Saulnier et al.: Two out of the three markers namely TNFR1 and NTProBNP studied by Saulnier and colleagues were included in this study. The third marker Adrenomedullin was in our short list (see table S1) but could not be included as no suitable analyte was available for the LXSAH assay. The work by Saulnier et al. has now been included into the discussion.</p>
<p><b>R4 Maj2:</b> The enrichment of patients with patients with low eGFR for fast progressors is an issue. I do not see the point here. If the slope representing renal function decline is the key phenotype so the authors should not derive from their data, no matter the baseline eGFR. This point is surely critical to the paper, as the reasons why to match on baseline eGFR are not (well-) exposed or exposed in the statistical section after the point is addressed in the methods section. Please address.</p>	<p>We thank the reviewer for this comment. Generally, we agree that for the primary outcome of interest (GFR slope, continuous) it would not have been necessary to match on baseline GFR. Nevertheless, we at the same time aimed at generating a data set for a nested case control study for which a balanced distribution of covariates facilitates analysis. Baseline GFR is known to be a strong predictor for kidney function decline and was therefore balanced between the groups. The inclusion of patients with less pronounced GFR decline into the fast progressors group reduces the difference between the two groups (regarding GFR slope), however, it is not in violation with the group assignment one would get by dichotomizing with a fixed cutoff.</p>
<p><b>R4 Maj3:</b> A key finding is the fact that the biomarkers studied here are helpful for eGFR at baseline rather than for eGFR decline at variance with the title. This is really an issue as the selection of patients was made to maximize difference in eGFR decline, taking an extreme phenotype strategy. Whether this is related to the enrichment of group 2 patients should be deeply searched for very carefully as it questions the analysis strategy proposed here. Please address.</p>	<p>We thank the reviewer for this critical comment. We performed an additional analysis to follow up on the reviewer's concern that the inclusion of patients with less pronounced GFR decline into the fast progressors group caused the markers to be only associated with the general trend of GFR decline and not the individual slopes. Patients from quintile two that were originally added to the fast progressors group were removed and as illustration the main model (presented in Table 2) was refitted (see additional table 1 in this letter). Marker selection into models, performance and association with baseline eGFR remained virtually unchanged and the overall conclusion that the</p>



	markers are of low clinical utility remains the same.
<b>R4 Maj4:</b> The authors claim that the originality of the paper is related to its capacity to distinguish between biomarkers involved in baseline eGFR and in eGFR decline. However, Saulnier et al in a previous article on the same topic already performed such an analysis, even though it was on a limited number of biomarkers.	We thank the reviewer for this comment. Yes, we appreciate the statistically sound analysis performed by Saulnier et al on the three markers under investigation in the study. We rephrased the paragraph about splitting contribution into involvement in baseline and slope and refer now explicitly to the Saulnier et al.
<b>R4 Maj5:</b> The literature must be more carefully examined and relevant papers in this field must be confronted and referenced (see previous point in Major Concern #1).	We thank the reviewer for sharing the concern regarding the examination of existing literature. We hope that the systematic inclusion of gene disease association from literature eliminates major concern 1 and also partial addresses this one. For all claims in the manuscript appropriate references are given and results from selected previous studies have been included. For the given set of markers a large body of literature exists as these markers candidates are already well known proteins in the context of DKD. However, a complete review of the existing literature for those markers would be beyond the scope of this article.
<b>R4 Min1:</b> Table 1 is interesting but could include data comparing stable vs fast progressors as in supplementary table 3 so that the baseline characteristics of each of the study population could be easy to understand by the reader.	We thank the reviewer for this suggestion and have included separated columns for fast and stable groups into table 1.
<b>R4 Min2:</b> The authors used the term CONSORT flowchart which is simply a flow chart. As CONSORT refers to randomized trials the term looks inappropriate.	We thank the reviewer for pointing out our mistake and have now removed the word CONSORT from the manuscript.
<b>R4 typographical 1:</b> Unit for cholesterol and serum glucose in table 1/ Supplementary table 3 is mmol/l (not /dl)	We are very grateful that the reviewer spotted our typo in units for glucose and cholesterol and changed it appropriately.

**Additional table 1:** Mixed model for prediction of eGFR levels (with baseline eGFR as part of the dependent variable) obtained from AIC-based backward elimination applied to all candidate predictors (log2 transformed biomarker and clinical) but excluding patients which were selected from the second slope quintile to balance eGFR distributions across the two patients groups (see supplement table 2). The overall adjusted R<sup>2</sup> for this model is 59%.

	Baseline		Slope		
Predictor	Coefficient	p-value	Coefficient	p-value	R <sup>2</sup> decomposition
Constant	395.609	<0.001	-1.575	0.559	-
Cystatin C	-10.728	<0.001	n.s.	n.s.	7.2
Endostatin	-3.390	0.089	n.s.	n.s.	<1
UMOD	3.196	<0.001	n.s.	n.s.	5.3
CHI3L1	1.340	0.014	n.s.	n.s.	1.3
HGF	-0.040	0.957	0.472	0.056	<1
MMP1	0.805	0.154	-0.318	0.097	<1
MMP7	-1.116	0.098	n.s.	n.s.	<1
TIE2	3.872	0.003	n.s.	n.s.	1.9
TNFR1	-10.110	<0.001	n.s.	n.s.	10
KIM1	-0.202	0.766	-1.179	<0.001	3.3
FGF23	-0.816	0.349	0.427	0.161	<1
NTproBNP	0.114	0.671	-0.290	0.004	<1
Age (Years)	-0.706	<0.001	0.056	0.028	25.5
Current or former smoker	2.953	0.007	n.s.	n.s.	<1
MAP	0.096	0.057	n.s.	n.s.	<1
HbA1C	-0.754	0.131	n.s.	n.s.	<1
Total cholesterol	-1.022	0.026	n.s.	n.s.	<1

The results show that by enriching the dataset with the additionally selected patients, the results stay essentially the same (only HbA1C enters the model as additional, but as weak predictor with low R<sup>2</sup> contribution) but power increases with the more equalized distribution and increases in sample size.

## **Response to formatting comments**

Formatting comment 1: Abstracts are limited to 250 words. You currently have 270.

Response: The abstract has been shortened to 250 words.

Formatting comment 2: Original Articles are limited to 4,000 words. You currently have 4,339.

Response: The main text has been shortened to 3995 words.

Formatting comment 3: In this current version, your table(s) contains hard returns. This is when the "Enter" or "Return" key is used at the end of a line or within the cell. We ask that the text wrap naturally or be included in a new row.

Response: Hard returns have been removed.