Integrative Analysis of Prognostic Biomarkers Derived from Multiomics Panels for the Discrimination of Chronic Kidney Disease Trajectories in People with Type 2 Diabetes

Running title: Biomarkers predicting chronic kidney disease in diabetes

Michael Kammer, MS[1,2*], Andreas Heinzel, MS[1*], Jill A. Willency, BS[3], Kevin L. Duffin, PhD[3], Gert Mayer, MD[4], Kai Simons, PhD[5], Mathias J. Gerl, PhD[5], Christian Klose, PhD[5], Georg Heinze, PhD[2], Roman Reindl-Schwaighofer, MD[1], Karin Hu, BS[1], Paul Perco, PhD[4], Susanne Eder, PhD[4], Laszlo Rosivall, MD, PhD[6], Patrick B. Mark MB ChB, PhD[7], Wenjun Ju, PhD[8], Matthias Kretzler, MD[8], Mark I. McCarthy, MD[9,10], Hiddo L. Heerspink, PhD[11], Andrzej Wiecek, MD[12], Maria F. Gomez, PhD[13], and Rainer Oberbauer, MD, PhD[1] for the BEAt-DKD consortium

*equal authorship

1 Department of Nephrology, Medical University of Vienna, Vienna, Austria

2 Center for Medical Statistics, Informatics and Intelligent Systems (CeMSIIS), Section for Clinical Biometrics, Medical University of Vienna, Vienna, Austria

3 Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN, USA

4 Department of Internal Medicine IV (Nephrology and Hypertension), Medical University of Innsbruck, Innsbruck, Austria

5 Lipotype GmbH, Tatzberg, Dresden, Germany

6 International Nephrology Research and Training Centre, Institute of Pathophysiology, Semmelweis University, Budapest, Hungary

7 Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, United Kingdom

8 Department of Internal Medicine, Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

9 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.

10 Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

11 Clinical pharmacy & pharmacology, Faculty of Medical Sciences, University Medical Center Groningen, Groningen, Netherlands

12 Department of Nephrology, Transplantation and Internal Medicine, Medical University of Silesia, Katowice, Poland.

13 Department of Clinical Sciences in Malmö, Lund University Diabetes Centre, Lund, Sweden

Corresponding author

Rainer Oberbauer, MD, PhD

Medical University of Vienna

Währinger Gürtel 18-20, A-1090 Vienna, Austria

T: +43 1 4040043900

rainer.oberbauer@meduniwien.ac.at

Funding Sources

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 EFPIA with JDRF.

Word count: Abstract: 246, Translational Statement: 99, Manuscript including Abstract: 3981

Number of Tables: 2

Number of Figures: 2

Supplement: Number of Tables: 3, Number of Figures: 13

Keywords: multiomics, type 2 diabetes, biomarkers, chronic kidney disease, integrative analysis, prognosis

Abstract

Clinical risk factors explain only a fraction of the variability of estimated glomerular filtration rate (eGFR) decline in people with type 2 diabetes. Cross-omics technologies by virtue of a wide spectrum screening of plasma samples have the potential to identify biomarkers for the refinement of prognosis in addition to clinical variables.

We utilized proteomics, metabolomics and lipidomics panel assay measurements in baseline plasma samples from the multinational PROVALID study (PROspective cohort study in patients with type 2 diabetes mellitus for VALIDation of biomarkers) of people with incident or early chronic kidney disease (median follow-up 35 months, median baseline eGFR 84 mL/min/1.73m^2, urine albumin-to-creatinine ratio 8.1 mg/g). In an accelerated case-control study, individuals with stable eGFR course (n=258, median eGFR change 0.1 mL/min/year) were compared to individuals with rapid eGFR decline (n=223, median eGFR decline -6.75 mL/min/year) using Bayesian multivariable logistic regression models to assess the discrimination of eGFR trajectories.

The analysis included 402 candidate predictors and showed two protein markers (KIM1, NTproBNP) to be relevant predictors of the eGFR trajectory with baseline eGFR being an important clinical covariate. The inclusion of metabolomics and lipidomics platforms did not improve discrimination substantially. Predictions using all available variables were statistically indistinguishable from predictions using only KIM1 and baseline eGFR (area under the receiver operating characteristic curve 0.63).

In conclusion, the discrimination of eGFR trajectories in people with incident or early diabetic kidney disease and maintained baseline eGFR was modest and the protein marker KIM1 was the most important predictor.

Translational Statement

Considerable funds have been used in the search for prognostic biomarkers in the field of chronic kidney disease (CKD), specifically in people with diabetes mellitus type 2, but no single biomarker is used routinely for risk prediction yet due to low explanatory power. Therefore, given the widely available multiomics analyses nowadays, it is appealing to use a wider spectrum of biomarkers for prognosis. However, our study suggests that even using multiomics derived biomarker candidates, there is currently no clinically useful multivariable panel of plasma biomarkers that may be recommended for prediction of future eGFR loss in early CKD stages.

Introduction

The global burden of chronic kidney disease (CKD) in persons with type 2 diabetes is increasing (1). The progression of CKD in these individuals is highly inconsistent and clinical risk factors can only explain some of the variability (2; 3). Refinement of the prognostic precision using biomarkers would allow for a more individualized risk assessment and personalized therapy (4; 5).

Several large consortia have been set up to investigate the utility of such novel markers derived from omics experiments (SysKid, RHAPSODY, SUMMIT, BEAt-DKD) (6-8). Although several markers were univariably found to be statistically significantly associated with estimated glomerular filtration rate (eGFR) trajectories, the potential coregulation and collinearity have not been explicitly addressed in most of these studies. Additionally, the explained variability of eGFR prediction of most marker panels in addition to clinical risk factors was generally found to be low (9). The SUMMIT consortium addressed some of these issues and found that two protein biomarkers, Kidney Injury Molecule-1 (KIM1) and Beta-2-Microglobulin (B2M), performed equally well as a considerably larger panel adjusted for clinical risk factors in predicting rapid decline of eGFR (10).

Novel technologies nowadays allow for a simultaneous determination of a large group of potential progression risk markers derived from proteomics, metabolomics and lipidomics analyses (11; 12). However, the combined analysis of the different omics layers to study their association with renal function decline has not been performed yet. Therefore, the objective of the current study was to identify if and which type of biomarkers improve discrimination of different eGFR trajectories in early stage CKD patients with type 2 diabetes on top of clinical risk factors. To achieve this goal, innovative statistical methods had to be developed and applied to

account for the challenging nature of the data with relatively few study outcomes and the complex, largely unknown coregulation of the biomarker candidates. At the same time, clinically established risk factors needed to be considered in the analysis. We therefore applied a Bayesian approach for regularized regression models which integrated clinical covariates and explorative multiomics marker with focus on the metabolomics and lipidomics platforms to refine the discrimination of individual eGFR courses. This study was conducted by the BEAt-DKD investigators (Biomarker Enterprise to Attack DKD; https://www.beat-dkd.eu/).

Results

We assessed the discrimination of eGFR trajectories in an accelerated case-control study by comparing 223 fast progressing and 258 stable individuals selected from a cohort of people with type 2 diabetes and early or incident CKD (cohort selection is summarized in Supplementary Figure S1, eGFR trajectories are depicted in Supplementary Figure S2). Demographics are presented in Table 1, which shows balanced covariates between the two groups. In total, 402 predictors for 481 persons were evaluated in our study: 10 clinical covariates, 15 protein, 180 metabolite and 197 lipid biomarkers. Protein and metabolite markers had few missing values (below 10% missing for all biomarkers), lipids were restricted to those with at least 66% completeness. A breakdown of biomarkers used in this analysis and their availability is given in Supplement section 3. The overall correlation matrix is shown in Supplementary Figure S3. All analyses were based on Bayesian multivariable logistic regression models for the outcome of interest (fast progression versus stable course).

Clinical Model

As a baseline for the evaluation of more elaborate models we fitted a model comprising only clinical covariates using weakly informative Student-t prior

distributions, which corresponded to a standard frequentist analysis of the data (see Supplementary Table S1). The results indicated that the clinical risk factors by themselves were not sufficient to discriminate fast progressing and stable individuals (Area under the receiver operating characteristic curve AUROC 0.55, 95% Bayesian credible interval BCI [0.46, 0.64], Table 2 and Supplementary Figure S4).
To identify important predictors we employed shrinkage priors for all subsequent analyses. For the clinical model, this affected the coefficient estimates and shrank almost all standardized odds ratios for fast progression (sOR) to unity but did not change the discriminative ability (AUROC unchanged at 0.55 [0.46, 0.64], Table 2 and Supplementary Figure S5).

Single Biomarker Group Analysis

Regarding models which included only a single group of biomarkers, the pre-selected proteomics markers performed best in terms of AUROC (0.63 [0.51, 0.76], Table 2 and Supplementary Figure S6). KIM1 (sOR 1.52 [1.17, 1.92]) and NTproBNP (sOR 1.28 [0.99, 1.57]) had the largest marginal effects with high posterior probability for a sOR larger than unity. The metabolomics (AUROC 0.58 [0.47, 0.69], Table 2 and Supplementary Figure S7) and lipidomics panels (AUROC 0.54 [0.46, 0.63], Table 2 and Supplementary Figure S8) achieved similar performance as the clinical model. However, for these two large biomarker panels none of the candidate predictors showed large marginal effect sizes, likely due to correlation of the estimated variable effects.

Combined Analysis

The main structured model with shrinkage priors accommodating the group structure of the data and comprising all 402 candidate predictors (AUROC 0.63 [0.53, 0.74]) performed comparably to the protein subpanel and was robust regarding parameter settings (Supplementary Figure S9). We present the top 15 predictors in terms of

sOR in Figure 1 and Supplementary Table S2. Baseline eGFR (sOR 2.23 [1, 3.4]) and KIM1 (sOR 1.59 [1.13, 2.09]) were the only predictors for which the associated 95% BCI excluded unity. Across all sensitivity analyses and models, KIM1 was the single most robust marker, for which the effect was largely unaffected by the presence of other variables in the model. The protein marker NTproBNP (sOR 1.28 [0.97, 1.65]), even though its 95% BCI included unity, was also consistently found to have more than 90% posterior probability for a sOR larger than unity, regardless of choice of prior and other variables in the model. The role of baseline eGFR was subtle. Its apparent importance may partly be explained by our study design, as individuals with high eGFR at baseline possibly had a higher likelihood for declining eGFR trajectories and for having been selected as cases than individuals whose eGFR was already very low to begin with. Even though baseline eGFR was statistically indistinguishable between the person groups (Table 1) and did not have a marginal effect in the clinical model (sOR 1.01 [0.9, 1.2]), it became important when adjusted for correlated markers such as CystatinC and several metabolite markers. Each of these variables alone did not well discriminate fast progressing and stable individuals but mutual adjustment in the model led to increased effect sizes with large uncertainty, as observed in Figure 1. Nevertheless, the overall contribution of eGFR correlated markers to discrimination remained low. This is illustrated in Figure 2 which shows the change of AUROC while limiting the number of predictors in the model. KIM1 provided the largest increase in AUROC of any single marker. Together with baseline eGFR and NTproBNP the discrimination performance was comparable to the full panel of clinical covariates and biomarkers.

Sensitivity Analysis

Excluding baseline eGFR as a covariate resulted in a drop in effect size for most variables, including duration of diabetes at study entry, but KIM1 and NTproBNP

were largely unaffected (Supplementary Figure S10). The discrimination performance was slightly lower than for the main structured model (AUROC 0.61 [0.49, 0,72]). Changing the coefficient priors to be uninformed of the predictor grouping structure led to larger marginal effects for the metabolite and lipid markers, with little changes for the top proteomics and clinical variables (Supplementary Figure S11). The increased effect sizes were due to less regularization for the larger biomarker platforms, which, however, did not increase performance compared to the main structured model (AUROC 0.6 [0.48, 0.71]). Regarding the added value of the novel biomarker platforms on top of clinical variables and pre-selected proteomics biomarkers, none of the metabolite or lipid markers showed consistent effect sizes. These results were confirmed in the sensitivity analysis based on principal component analysis (Supplementary Figures S12 and S13). While some of the principal components showed larger effects than any of the single markers they were composed of, discrimination did not differ from the main structured model (AUROC 0.64 [0.55, 0.74]).

Discussion

In this innovative study we used a Bayesian approach to assess the added value of biomarkers derived from cross-omics experiments when used together with clinical covariates in baseline samples of a prospective cohort study of individuals with type 2 diabetes and incident or early diabetic CKD. The purpose of the study was to determine if lipid, metabolite, or protein biomarkers could enhance prognostication of future disease trajectories. We found that clinical variables by themselves were not sufficient to discriminate between fast progressing and stable individuals in this cohort, while two protein markers, KIM1 and to lesser extent NTproBNP, robustly contributed to the discrimination of the eGFR trajectories. Furthermore, when used together with the proteomics biomarkers and clinical variables, the metabolite and lipid biomarkers did

not increase discriminative ability of the models. Due to strong correlations, these biomarkers seem exchangeable – a model with a few of them performs as well as a model considering the entire platform. The overall discrimination and therefore the clinical utility for prediction of eGFR decline was low. Inclusion of urinary albumin excretion did not increase discrimination which may be explained by the low amount of albuminuria in this study cohort of people with early stage disease and maintained kidney function.

Other studies investigated the statistical performance of large panels of protein biomarkers for predicting disease progression. Gerstein and colleagues used a large protein panel in almost 7500 participants of the ORIGIN trial to identify suitable prognostic markers of eGFR slope. The investigators identified 13 markers that were independently associated with progression. The explained variance of the marker plus clinical parameters model was 0.155 (95% confidence interval [0.139, 0.172]) (Gerstein H, Pare G, McQueen M et al "Novel Biomarkers Predicting Renal Dysfunction in People with Dysglycemia in the ORIGIN Trial". Abstract for *American Diabetes Association's 79th Scientific Sessions* San Francisco, CA, 2019). However, this elegant study used only protein biomarker candidates to estimate eGFR slope. In general, only few studies investigated the capacity of multiomics panels to discriminate individuals with fast eGFR loss versus stable course.

Previously Bansal and colleagues measured two markers (NTproBNP and troponin T) in 3752 participants free of heart failure in the Cardiovascular Health Study (13). Participants in the highest quartile of NTproBNP concentrations (>237 pg/ml) exhibited a 67% higher risk of rapid eGFR decline and 38% higher adjusted risk of incident CKD compared to stable individuals (adjusted hazard ratio for incident CKD 1.38; 95% confidence interval [1.08, 1.76]). Troponin T was not statistically associated with this outcome.

KIM1 in combination with B2M, the latter not being measured in our study, were also found to be key predictors of eGFR loss in an analysis of the SUMMIT investigators (10).

Metabolomics technologies can measure thousands of small sized biochemicals, but its utility for biomarker discovery in CKD remains unclear. The US based CKD biomarkers consortium recently showed in a metabolomics analysis of 49 proteinuric patients with advanced CKD (eGFR <60 mL/min/1.73m^2) that several hundred metabolites were inversely correlated with eGFR in a cross-sectional study, but the intraindividual variability was high (14). Some of these investigators showed previously that many metabolites were associated with incident CKD in African American persons and identified two metabolites as best candidates to predict CKD risk (15).

Lipidomics is an emerging field and together with genomics and proteomics allows for a better understanding of cellular physiology and pathology. Prades and colleagues showed that this technology can be used to identify tissue specific patterns of panels (16). Therefore, it is appealing to identify lipid biomarker candidates for complex biological processes such as the progression of diabetic kidney disease. Here we applied all three technologies, proteomics, metabolomics and lipidomics, to the same baseline samples from individuals with early diabetic kidney disease using innovative analytical methods.

The findings of the current study are in line with and extending our previous work (9), which used mixed linear models for the longitudinal eGFR values based on clinical covariates and protein biomarkers. There, KIM1 and NTproBNP were also found to contribute to the prediction of the eGFR slope. Baseline eGFR showed a dominating effect in terms of explained variation, as it provided the natural starting point for the future eGFR trajectory and because effects on the eGFR slope contributed less to the models' explained variation than effects on baseline eGFR. In this current analysis, we

were able to further dissect the role of baseline eGFR and found that its effect strongly correlated with effects of other biomarkers such as CystatinC and several metabolites, reflecting a persons' disease state at baseline.

Our study has some limitations. Missing biomarker measurements were mitigated by the use of multiple imputation. The early disease stage of the participants, which may limit the detection of progression specific signals, was chosen on purpose as risk prediction is likely most useful in early disease states when potential interventions exhibit the greatest benefit.

The definition of our outcome and the study design with exploratory aspects may be seen as both an advantage and a limitation of the study. By carefully investigating the patient's trajectories over the relatively short follow-up of around three and a half years in the PROVALID cohort, we attempted to select robustly identifiable eGFR courses on the opposite sites of the prognostic spectrum, which helped to mitigate the issue of limited sample size and to increase statistical power. However, by selecting based on the outcome, a possible source of bias is introduced towards larger effect sizes. Since our results yield low discrimination and effects which are in agreement with other studies conducted, we argue that this bias should be of little concern.

Strengths of our study include the precise clinical characterization of the prospective PROVALID cohort from which we selected individuals (17). Our statistical approach allowed us to integrate multiomics data. By foregoing a pure variable selection based approach we were able to carefully discuss the complex interplay of biomarkers and clinical variables.

In conclusion, the discrimination of eGFR trajectories using baseline circulating multiomics derived biomarkers in combination with clinical parameters was modest, with the two protein markers KIM1 and to lesser extent NTproBNP emerging as the most relevant biomarker candidates. The findings for metabolite and lipid biomarkers

may help to identify new targets and pathways for pharmaceutical intervention to treat individuals with early progressive CKD and mitigate their future renal decline to end stage renal disease and dialysis.

Methods

Study Design and Cohort

The motivation for this study was the careful evaluation of known protein biomarker as well as the exploration of further metabolite and lipid candidates. We employed an accelerated case-control design which selected individuals from the opposite ends of the prognostic spectrum and hence made associated variables easier to detect because the effects are more pronounced than in the full cohort. The associated increase in statistical power has been studied in (18). Our study cohort was derived from PROVALID, a prospective multinational cohort study of persons with type 2 diabetes and incident or early CKD with annual follow-up (19; 20). In total 2560 participants from three countries were available for the present analysis. After excluding people with less than 720 days of follow-up (FU) we estimated the annual change in eGFR via slopes from individual linear regressions. Persons were then grouped according to the slope quintiles. In a final screening step, individuals for whom the slope was deemed unsteady due to large uncertainty of the estimated slope and manual inspection for exceptional eGFR values were excluded. This resulted in 223 individuals with fast progression (first and parts of second quintile, median eGFR slope of -6.75 mL/min/year) and 258 individuals with stable course (fourth quintile, median eGFR slope of 0.1 mL/min/year) who were selected for this study. Further details of the cohort derivation may be found in Supplement section 1 and in (9).

Outcome of Interest

The study outcome was an individuals' course of renal function decline as determined by eGFR trajectory (fast progression versus stable course) where eGFR was computed according to the CKD-EPI equation (21).

Clinical Risk Factors

Clinical risk factors which were included in the analysis comprised age, sex, serum cholesterol, urine albumin-to-creatinine ratio, HbA1C, mean arterial pressure, BMI, smoking status and duration of type 2 diabetes at study entry. We also included eGFR at baseline as it was found to be an important covariate in our preceding analysis (9).

Biomarker Selection and Measurement

Baseline plasma samples were subjected to targeted proteomics and metabolomics profiling as well as untargeted lipidomics profiling. A complete list of all candidate biomarkers is provided in Supplement section 3. All markers were measured in K3 EDTA plasma that was stored at -80°C immediately after collection. All samples were handled in an identical manner.

Protein Biomarker Candidates

Details on the protein biomarker selection can be found in a previous publication (6). Protein marker measurements and quality control data were also described previously (9). In short, the following 15 plasma markers were selected and analyzed in this current study: UMOD, Endostatin, CystatinC, MMP1, MMP7, CCL2, GH, VCAM1, HGF, CHI3L1, TIE2, TNFR1, KIM1, FGF23 and NTproBNP. SOST and MMP8 were initially also selected but subsequently excluded from all multivariable analyses due to measurement issues as reported previously in (9).

Metabolomics Biomarker Candidates

Metabolomics analysis was performed using targeted hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry (HILIC-LC-MS/MS) methodology. A detailed description of the metabolomics analysis may be found in Supplement section 3. In brief, metabolites were extracted from plasma samples using acetonitrile:isopropanol:water. HILIC separation was achieved using an apHera NH2 column (Supelco Analytical, Bellefonte, PA, USA) and LC-MS/MS analysis was carried out on a Shimadzu Nexera UPLC system (Shimadzu, Kyoto, Japan) coupled with a SCIEX Triple Quadrupole 6500+ mass spectrometer (SCIEX Framingham, MA, USA).

Lipid Biomarker Candidates

Mass spectrometry-based lipid analysis was performed at Lipotype GmbH using high throughput Shotgun Lipidomics (Dresden, Germany) technology as described in (22-24). A detailed description of the lipid extraction and measurement process may be found in Supplement section 3. In total 637 lipids from all 16 lipid classes could be identified. Out of these, 197 lipids that could be determined in more than 66% of samples were included in further statistical analysis in this study.

Statistical Analysis

We estimated the effects of clinical risk factors and biomarkers on the binary outcome (fast progression versus stable course) using Bayesian multivariable logistic regression models with shrinkage priors to identify important risk factors. The Bayesian framework was chosen as it allows incorporation of the a-priori known data structure with four groups of variables (clinical covariates, proteomics, metabolomics, lipidomics) and provided a conceptually clear way of handling missing data. Furthermore, uncertainty in estimated variable effects could be assessed by their posterior distributions. All biomarker values were log2 transformed to achieve

symmetric distributions. For comparability, all variables were centered and scaled to have zero mean and unit variance. Results are presented as standardized odds ratios (sOR), such that values larger than unity correspond to faster eGFR decline. We summarize posterior distributions by means and 95% BCI (defined by highest posterior density intervals), which are contiguous regions containing the quantity of interest with 95% probability, given the prior assumptions and the data. The ability to discriminate the two patient groups was expressed in terms of AUROC, which was estimated using 5 times repeated 5-fold cross-validation.

We used shrinkage prior distributions based on the regularized horseshoe prior (25), which expressed our belief that only few biomarkers have a relevant individual association with the outcome. This prior specification shrank regression coefficients of weak predictors towards zero (corresponding to odds ratios of unity). Hyperparameters were specified following recommendations from the literature and were guided by model convergence (25-27).

To investigate the added value of using biomarkers on top of clinical covariates we fitted “single group” and “structured” models. “Single group” models included only one group of variables (i.e. clinical variables or a single biomarker platform). “Structured” models included all variables with coefficient priors accommodating the grouping of the variables by slightly extending the definition of the horseshoe prior. We added an additional shrinkage parameter for each group, which allowed hierarchical attribution of shrinkage to groups of variables and to variables within a group. This was motivated by the careful preselection of the protein biomarkers, whereas the full metabolite and lipid biomarker panels were of a more explorative nature, therefore likely requiring stronger regularization. As sensitivity analysis, we also fitted “combined” models which contained all variables but ignored their grouped data structure and used standard regularized horsehoe priors, as well as models based on

unsupervised principal components computed from each group.

Importance of predictors was assessed after model fit following methods outlined in (28; 29), which allowed computing predictions using only a subset of variables of a reference model. Using this technique, we obtained AUROC values corresponding to submodels of our presented reference models within the repeated cross-validation to investigate the change in discriminative accuracy when using only subsets of variables for predictions. Our simple strategy to choose these predictor subsets was to start with an intercept-only model, to which further variables were added in order of their absolute effect size (absolute posterior mean of the reference model).

Uncertainty due to missing data was accounted for using 20 multiply imputed datasets, extending our approach from (9) to the novel metabolite and lipid biomarkers. The Bayesian framework allowed pooling of the posteriors resulting from the multiple imputed datasets into a single pooled posterior, incorporating the additional uncertainty due to missing data. The pooled posterior of each model was used for further summarization.

Convergence of the Bayesian models was monitored using standard diagnostic tools (Supplementary Table S3). All analyses were implemented in R 3.4.1 (https://www.r-project.org/foundation, Vienna, AT), using Stan 2.18.0 (30) for fitting the Bayesian models. The Stan implementation of our models can be found in Supplement sections 11 and 12, based on code published in (25). Statistical methodology is fully described in Supplement section 2.

Sample Size Considerations

The sample size considerations were reported previously and were initially based on identification of single markers in a cohort study (9). Regarding the power of the current study, a sample size of 209 cases and 209 controls was sufficient to detect a standardized odds ratio of 1.5 of a normally distributed biomarker (or biomarker score)

adjusted for other covariates with a power of 80% at a two-sided significance level of 5%. This calculation further assumed that the biomarker exhibited a squared multiple correlation of 0.5 with the adjustment covariates. The power analysis was conducted using nQuery software version 8.2.

Disclosure

K.S. is CEO of Lipotype GmbH. K.S. and C.K. are shareholders of Lipotype GmbH. M.J.G. is an employee of Lipotype GmbH. No other potential conflicts of interest relevant to this article were reported.

Supplementary Material

Supplementary Figures

Supplementary Tables

Other Supplementary Material

Stan Implementation of Regularized Horseshoe Priors

Stan Implementation of the Structured Regularized Horseshoe Priors

Additional references

Supplementary information is available at Kidney International's website.

References

1. Jager KJ, Fraser SDS. The ascending rank of chronic kidney disease in the global burden of disease study. *Nephrol Dial Transplant* 2017;32:ii121-ii128.
2. Hallan SI, Ritz E, Lydersen S, et al. Combining GFR and albuminuria to classify CKD improves prediction of ESRD. *J Am Soc Nephrol* 2009;20:1069-1077.
3. Bruck K, Jager KJ, Zoccali C, et al. Different rates of progression and mortality in patients with chronic kidney disease at outpatient nephrology clinics across Europe. *Kidney Int* 2018;93:1432-1441.
4. Mann JFE, Orsted DD, Brown-Frandsen K, et al. Liraglutide and Renal Outcomes in Type 2 Diabetes. *N Engl J Med* 2017;377:839-848.
5. Wanner C, Inzucchi SE, Lachin JM, et al. Empagliflozin and Progression of Kidney Disease in Type 2 Diabetes. *N Engl J Med* 2016;375:323-334.
6. Mayer G, Heerspink HJ, Aschauer C, et al. Systems Biology-Derived Biomarkers to Predict Progression of Renal Function Decline in Type 2 Diabetes. *Diabetes Care* 2017;40:391-397.
7. Ahlqvist E, Storm P, Karajamaki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018;6:361-369.
8. van Zuydam NR, Ahlqvist E, Sandholm N, et al. A Genome-Wide Association Study of Diabetic Kidney Disease in Subjects With Type 2 Diabetes. *Diabetes* 2018;67:1414-1427.
9. Heinzel A, Kammer M, Mayer G, et al. Validation of Plasma Biomarker Candidates for the Prediction of eGFR Decline in Patients With Type 2 Diabetes. *Diabetes Care* 2018;41:1947-1954.
10. Colombo M, Looker HC, Farran B, et al. Serum kidney injury molecule 1 and beta2-microglobulin perform as well as larger biomarker panels for prediction of rapid decline in renal function in type 2 diabetes. *Diabetologia* 2019;62:156-168.
11. Mayer P, Mayer B, Mayer G. Systems biology: building a useful model from multiple markers and profiles. *Nephrol Dial Transplant* 2012;27:3995-4002.
12. Afshinnia F, Rajendiran TM, Wernisch S, et al. Lipidomics and Biomarker Discovery in Kidney Disease. *Semin Nephrol* 2018;38:127-141.
13. Bansal N, Katz R, Dalrymple L, et al. NT-proBNP and troponin T and risk of rapid kidney function decline and incident CKD in elderly adults. *Clin J Am Soc Nephrol* 2015;10:205-214.
14. Rhee EP, Waikar SS, Rebholz CM, et al. Variability of Two Metabolomic Platforms in CKD. *Clin J Am Soc Nephrol* 2019;14:40-48.
15. Yu B, Zheng Y, Nettleton JA, et al. Serum metabolomic profiling and incident CKD among African Americans. *Clin J Am Soc Nephrol* 2014;9:1410-1417.
16. Pradas I, Huynh K, Cabre R, et al. Lipidomics Reveals a Tissue-Specific Fingerprint. *Front Physiol* 2018;9:1165.
17. 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 nephrology* 2016;17:121.
18. Van Gestel S, Houwing-Duistermaat JJ, Adolfsson R, et al. Power of Selective Genotyping in Genetic Association Analyses of Quantitative Traits. *Behav Genet* 2000; 30(2):141-146.

19. 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:i482.
20. 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.
21. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604-612.
22. Surma MA, Herzog R, Vasilj A, et al. An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids. *Eur J Lipid Sci Technol* 2015;117:1540-1549.
23. Matyash V, Liebisch G, Kurzchalia TV, et al. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J Lipid Res* 2008;49:1137-1146.
24. Herzog R, Schwudke D, Schuhmann K, et al. A novel informatics concept for high-throughput shotgun lipidomics based on the molecular fragmentation query language. *Genome Biol* 2011;12:R8.
25. Piironen J, Vehtari A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics* 2017;11:5018-5051.
26. Carvalho CM, Polson NG, Scott JG. Handling Sparsity via the Horseshoe. In *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics* Hilton Clearwater Beach Resort, Clearwater Beach, Florida USA, Proceedings of Machine Learning Research 2009, p. 73-80.
27. Ghosh J, Li Y, Mitra R. On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression. *Bayesian Analysis* 2018;13:359-383.
28. Goutis C, Robert CP. Model Choice in Generalised Linear Models: A Bayesian Approach Via Kullback-Leibler Projections. *Biometrika* 1998;85:29-37.
29. Piironen J, Vehtari A. Comparison of Bayesian predictive methods for model selection. *Statistics and Computing* 2017;27:711-735.
30. Carpenter B, Gelman A, Hoffman MD, et al. Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 2017;76:32.

Acknowledgments

The study was conducted by the BEAt-DKD investigators (Biomarker Enterprise to Attack DKD). The BEAt-DKD consortium is a unique public private partnership funded by the Innovative Medicines Initiative (IMI), member companies from the European Federation of Pharmaceutical Industries and Associations (EFPIA), the Juvenile Diabetes Research Foundation (JDRF) and the Swiss state, with the goal to identify targetable mechanisms and pathways underlying initiation and progression of diabetic kidney disease, as well as to identify and validate biomarkers of disease progression and treatment responses. A full list of BeatDKD partners may be found on the website (https://www.beat-dkd.eu/).

	Overall	eGFR trajectory		Baseline eGFR	
		Stable	Fast	Above median (84 ml/min)	Below or equal to median (84 ml/min)
Number of persons	481	258	223	240	241
Age (Years)	64 ± 9.3	64 ± 10	65 ± 9	60 ± 8	68 ± 8
Sex (Female)	232 (47%)	117 (45%)	115 (50%)	95 (40%)	133 (55%)
Smoking status (Never)	250 (51%)	134 (52%)	116 (50%)	100 (42%)	145 (60%)
Duration of type 2 diabetes (Years)	10.8 ± 8.8	10 ± 8	12 ± 9	8 ± 7	13 ± 10
BMI (kg/m^2)	31 ± 5.5	31 ± 5	32 ± 6	32 ± 6	31 ± 5
Systolic blood pressure (mmHg)	138 ± 17.4	138 ± 16	139 ± 19	139 ± 18	138 ± 17
Diastolic blood pressure (mmHg)	79 ± 10.3	79 ± 10	79 ± 10	81 ± 10	77 ± 10
HbA1c (mmol/mol)	51 (45 / 60)	51 (45 / 61)	51 (44 / 60)	51 (44 / 61)	52 (45 / 60)
Hemoglobin (mmol/l)	8.6 (8.1 / 9.3)	8.8 (8.2 / 9.3)	8.5 (7.9 / 9.1)	8.9 (8.3 / 9.4)	8.4 (7.8 / 9.1)
Serum glucose (mmol/l)	7.4 (6.2 / 9)	7.5 (6.3 / 9)	7.4 (6 / 8.9)	7.4 (6.2 / 9.1)	7.4 (6.1 / 8.9)
Serum cholesterol (mmol/l)	4.6 (4 / 5.5)	4.6 (4 / 5.4)	4.6 (4 / 5.6)	4.5 (3.9 / 5.4)	4.7 (4 / 5.4)
Serum creatinine (µmol/l)	77 (66 / 95)	77 (67 / 95)	77 (65 / 95)	95 (81 / 111)	67 (60 / 75)
UACR* (mg/g)	8.8 (4.7 / 26.5)	8.2 (4.6 / 21)	9.2 (5 / 36.5)	7.6 (4.2 / 17.4)	10.8 (5.6 / 40.4)
Glucose lowering agents†					
None	59 (12%)	35 (14%)	24 (11%)	38 (16%)	21 (9%)
1 – 2 agents	355 (74%)	192 (74%)	163 (73%)	175 (73%)	180 (75%)
> 2 agents	67 (14%)	31 (12%)	36 (16%)	27 (11%)	40 (16%)
Blood pressure lowering agents‡					
None	77 (16%)	52 (20%)	25 (11%)	59 (25%)	18 (7%)
1 – 2 agents	195 (41%)	106 (41%)	89 (40%)	111 (46%)	84 (35%)
> 2 agents	209 (43%)	100 (39%)	109 (49%)	70 (29%)	139 (58%)

	Overall	eGFR trajectory		Baseline eGFR	
		Stable	Fast	Above median (84 ml/min)	Below or equal to median (84 ml/min)
ESA therapy§	11 (2%)	4 (2%)	7 (3%)	0 (0%)	11 (5%)
eGFR CKD-EPI (ml/min/1.73m^2)	84 (64 / 94)	85 (65 / 96)	82 (63 / 94)	94 (90 / 100)	64 (50 / 74)
eGFR CKD-EPI decline per year (ml/min/1.73m^2/year)	-0.7 (-6.3 / 0.2)	0.1 (-0.4 / 0.7)	-6.8 (-9 / -5.5)	-0.7 (-6.7 / 0)	-0.7 (-6 / 0.4)
Follow-up (months)	35	35	34	34	35

*UACR: Urine albumin-to-creatinine ratio

† Agent classes: Biguanides, Insulin, Sulfonylureas, DPPIV inhibitors/GLP1 agonists, Glinides, Glitazones, Alpha-Glucosidase-inhibitors, SGLT2

‡ Agent classes: ACE inhibitors / ARBs, β-blockers, Calcium antagonists (including direct vasodilators), α-blockers, Diuretics (Thiazide diuretics / Loop diuretics)

§ Including Darbepoetin alfa, Epoetin alfa, Epoetin beta, Epoetin theta, Epoetin zeta, Others

Table 1. **Cohort characteristics**

Characteristics of the study cohort, overall and stratified by pre-defined eGFR trajectory as well as by baseline eGFR. This table presents the same data as Table 1 in (9), which found baseline eGFR to be important for prediction of future eGFR trajectories. For baseline eGFR we chose the median of the cohort (84 ml/min/1.73m^2) as threshold to accommodate the early disease stage of the individuals. Data are reported as mean ± standard deviation, median (1st quartile, 3rd quartile) and absolute frequency (relative frequency). Clinical data was almost completely available (number of missing values: HbA1C 4, hemoglobin 9, serum glucose 1, serum cholesterol 1, urine albumin-to-creatinine ratio 14).

Model	AUROC			
	Mean	Median	95% BCI	
Clinical	0.55	0.55	0.46	0.64
Clinical with Shrinkage	0.55	0.54	0.46	0.64
Proteomics	0.63	0.64	0.51	0.76
Metabolomics	0.58	0.58	0.47	0.69
Lipidomics	0.54	0.54	0.46	0.63
Main Structured	0.63	0.64	0.53	0.74

Table 2. Model discrimination of fast progressing versus stable individuals

Discrimination of eGFR trajectory groups (fast progression versus stable course) with the presented models. Performance is measured in terms of cross-validated area under the receiver operatic characteristic curve (AUROC) and presented by posterior means, medians and 95% Bayesian credible interval (BCI). “Main Structured” indicates the main model of the study, using all variables with a prior adapted to the group structure of the data.

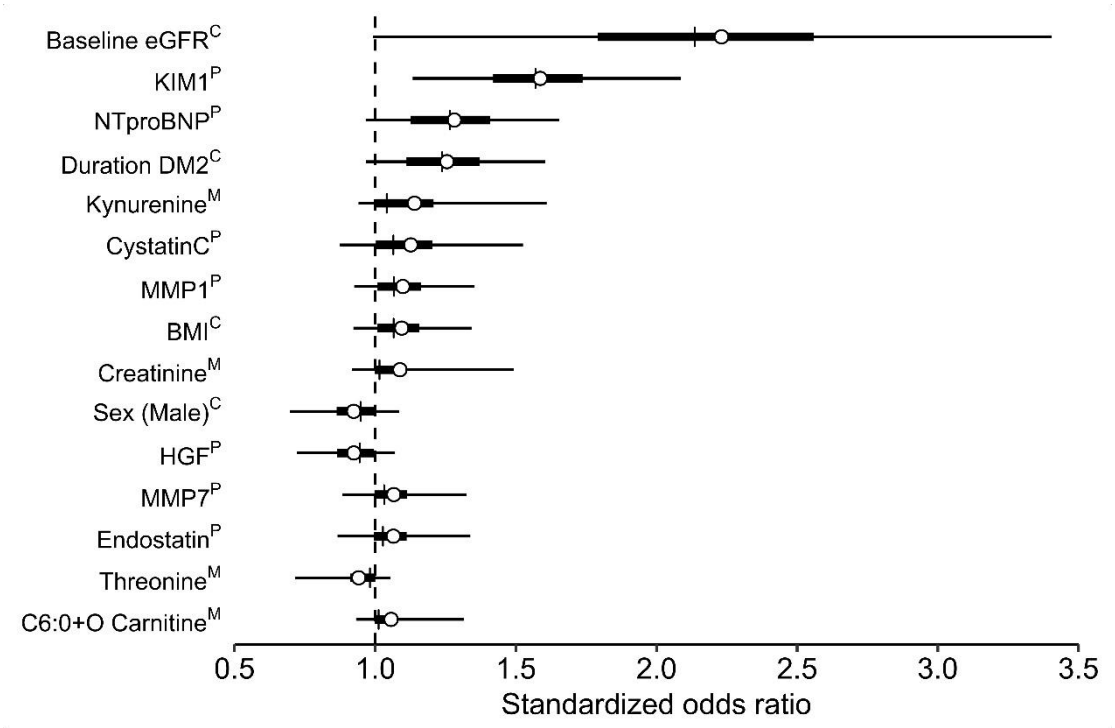


Figure 1. Standardized odds ratios from the main structured model

Coefficient posteriors from the main structured model including 402 candidate predictors and priors reflecting the group structure of the variables. The top 15 predictors in terms of effect size (absolute marginal posterior mean) are shown. Circles depict a point estimate (posterior mean), the posterior median is indicated by a vertical bar. The thick line represents the interquartile range and the thin line a 95% Bayesian credible interval. DM2 denotes type 2 diabetes mellitus. Group assignments are indicated by superscript initials (C for Clinical, P for Proteomics, M for Metabolomics, L for Lipidomics). Exact values for the standardized odds ratios are given in Supplementary Table S2.

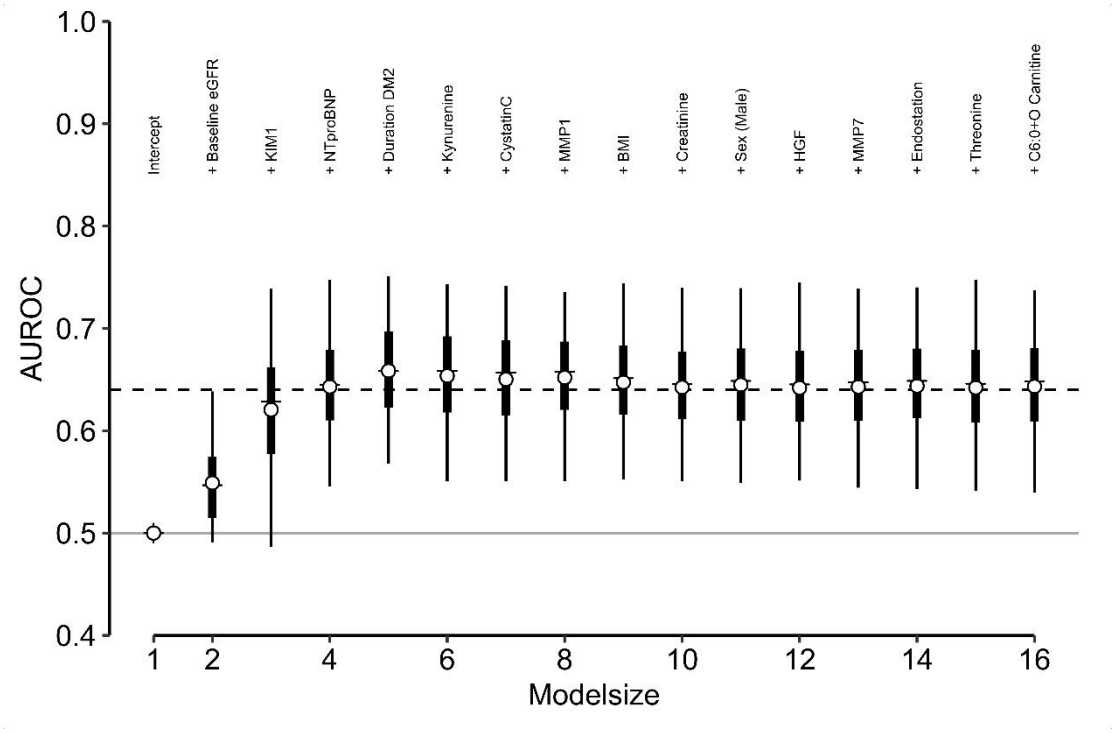


Figure 2. Change in model performance when using a limited number of predictors in the main structured model

Change in model performance when only a limited number of predictors is used to discriminate eGFR trajectory groups. Circles depict a point estimate of the cross-validated AUROC (posterior mean), the posterior median is indicated by a horizontal bar. The thick line represents the interquartile range and the thin line a 95% Bayesian credible interval. The horizontal dashed line indicates the AUROC of the main structured model using all 402 variables, which serves as a reference. Modelsize denotes the number of variables in the model, starting with the intercept-only model, then adding variables in order of their effect size as shown in Figure 1 (starting at the top with baseline eGFR). The predictions of the reference model were than projected onto the corresponding subset of its variables via the methodology outlined in the Statistical Methods section in order to obtain the AUROC values depicted.

Supplementary material

Integrative Analysis of Prognostic Biomarkers Derived from Multiomics Panels for the Discrimination of Chronic Kidney Disease Trajectories in People with Type 2 Diabetes

Michael Kammer, MS[1,2*], Andreas Heinzel, MS[1*], Jill A. Willency, BS[3], Kevin L. Duffin, PhD[3], Gert Mayer, MD[4], Kai Simons, PhD[5], Mathias J. Gerl, PhD[5], Christian Klose, PhD[5], Georg Heinze, PhD[2], Roman Reindl-Schwaighofer, MD[1], Karin Hu, BS[1], Paul Perco, PhD[4], Susanne Eder, PhD[4], Laszlo Rosivall, MD, PhD[6], Patrick B. Mark MB ChB, PhD[7], Wenjun Ju, PhD[8], Matthias Kretzler, MD[8], Mark I. McCarthy, MD[9,10], Hiddo L. Heerspink, PhD[11], Andrzej Wiecek, MD[12], Maria F. Gomez, PhD[13], and Rainer Oberbauer, MD, PhD[1] for the BEAt-DKD consortium

*equal authorship

1 Department of Nephrology, Medical University of Vienna, Vienna, Austria

2 Center for Medical Statistics, Informatics and Intelligent Systems (CeMSIIS), Section for Clinical Biometrics, Medical University of Vienna, Vienna, Austria

3 Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN, USA

4 Department of Internal Medicine IV (Nephrology and Hypertension), Medical University of Innsbruck, Innsbruck, Austria

5 Lipotype GmbH, Tatzberg, Dresden, Germany

6 International Nephrology Research and Training Centre, Institute of Pathophysiology, Semmelweis University, Budapest, Hungary

7 Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, United Kingdom

8 Department of Internal Medicine, Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

9 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.

10 Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

11 Clinical pharmacy & pharmacology, Faculty of Medical Sciences, University Medical Center Groningen, Groningen, Netherlands

12 Department of Nephrology, Transplantation and Internal Medicine, Medical University of Silesia, Katowice, Poland.

13 Department of Clinical Sciences in Malmö, Lund University Diabetes Centre, Lund, Sweden

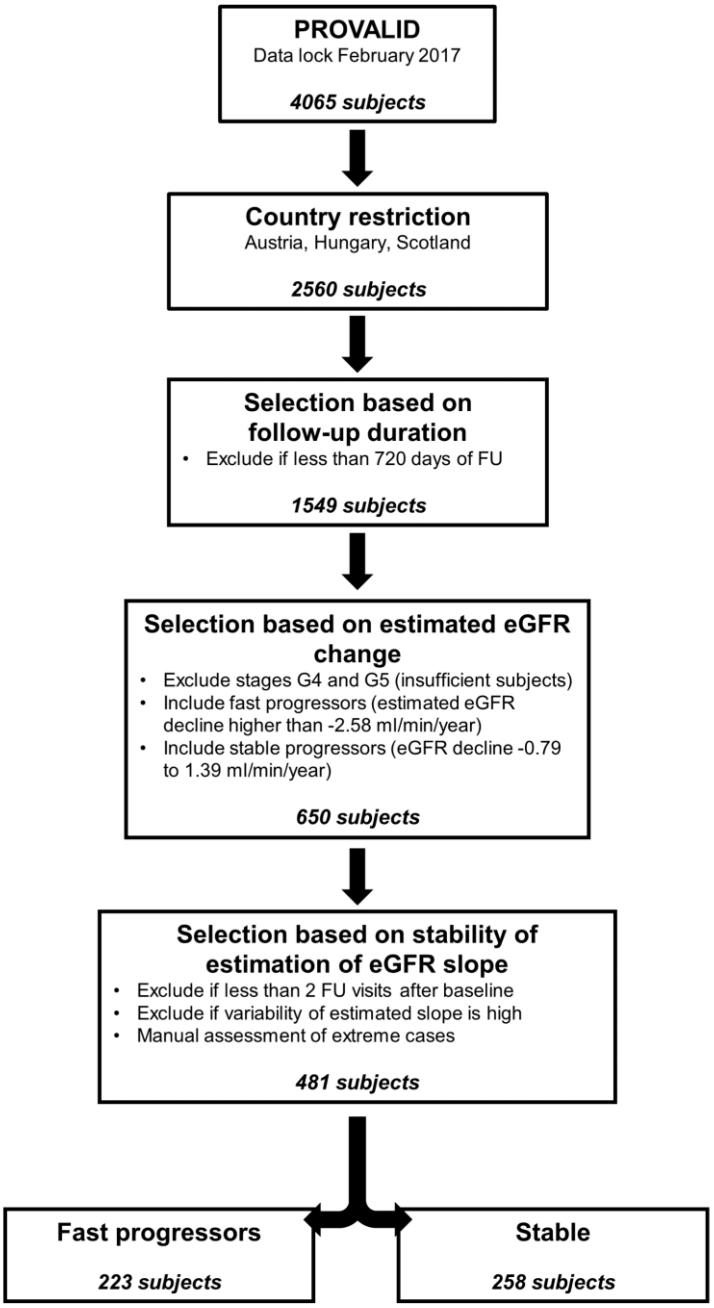
Corresponding author:

Rainer Oberbauer, MD, PhD
Medical University of Vienna
Währinger Gürtel 18-20, A-1090 Vienna, Austria
T: +43 1 4040043900
rainer.oberbauer@meduniwien.ac.at

Table of contents

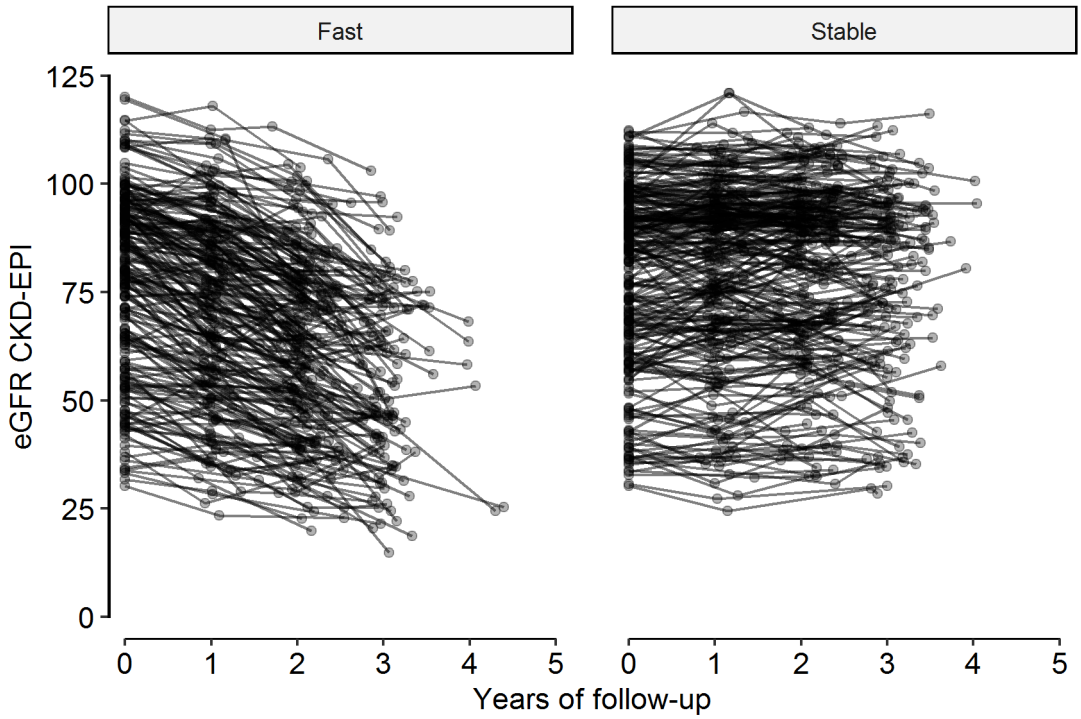
Cohort Selection and Grouping

In this paragraph we add some further insight into the rationale behind our definition of the eGFR trajectory groups, which was motivated by our choice to use an accelerated case-control study. As described in the main manuscript and summarized by the flowchart below, we selected 223 fast progressing and 258 stable individuals from the full PROVALID cohort. This selection was based on the quintiles of estimated annual eGFR slopes. After screening for individuals for whom the eGFR slopes were unsteady (i.e. likely to be uninformative of their true disease status) and excluding them from the study, the remaining individuals from the first slope quintile (-24.9 to -5.2 mL/min/year change of eGFR) represented the fast progressing group. These were complemented by persons from the second quintile (-5.2 to -2.6 mL/min/year) in order to balance out the fraction of patients in advanced disease stages to mitigate selection bias (KDIGO stage G3a/b at baseline), which would otherwise have been underrepresented in the fast-progression sample. The stable individuals were the individuals from the fourth quintile (-0.8 to 1.4 mL/min/year). We chose to use the fourth quintile instead of the fifth as the latter represented persons with atypical eGFR trajectories, i.e. mostly later stage individuals (according to KDIGO stage at baseline) who were e.g. influenced by treatment changes which led to a slight recovery of their eGFR values and thus increasing eGFR slopes.



Supplementary Figure S2: Flowchart of study cohort selection from the PROVALID study (S1)

Flowchart of study cohort selection from the PROVALID study, as previously reported in (S1). The full PROVALID cohort recruited persons in five countries, of which only three were available for the current analysis. The outcome groups were based on manual and statistical aspects to provide steady eGFR trajectories. Detailed information were previously reported in Supplementary Figure 1 and Supplementary Table 2 of (S1).



Supplementary Figure S3: eGFR trajectories of individuals selected for this study

Each individual is represented in the above plot by a line, actual eGFR measurements are indicated as dots. The purpose of this graphic is to give an indication of the difference between the two selected groups of persons and on the general steadiness of their eGFR slopes.

Extended Statistical Methods

We estimated the effects of clinical risk factors and biomarkers on the binary outcome (fast progression versus stable course) using Bayesian multivariable logistic regression models with shrinkage priors to identify important risk factors. Odds ratios larger than unity correspond to faster eGFR decline. Cohort characteristics are described by mean and standard deviation or median and first and third quartile for continuous variables or by frequency and percentage for categorical variables.

Bayesian Framework

The reason why we decided to model the data in a Bayesian framework were three-fold: first, to deal with missing data in a conceptually sound and practical way. Second, we could incorporate a-priori known information into the modelling approach. Specifically, this allowed us to adapt the models to the structure of our data, which comprised different groups of variables (clinical data, proteomics, metabolomics, lipidomics). Lastly, posterior distributions and credible intervals provided insights regarding the uncertainty of coefficient estimates and model performance results. Furthermore, we are able to report the models in a fully transparent manner (see code for the implementation of our analysis in Supplement sections 11 and 12). Alternative commonly used techniques such as the Lasso, Elastic net and other penalized regression approaches (adaptive Lasso, IPF Lasso) were considered but found to be lacking regarding the ability to provide interval estimates, had issues with available software or were more difficult to handle within the multiple imputation setup.
The Bayesian framework required several decisions to be made before fitting the models, namely specifying the prior distributions for the model parameters. These priors are then combined with the data to yield posterior distributions, which correspond to the probability of the parameters given the data observed and the prior assumptions. The resulting coefficient posterior distributions were transformed to the odds ratio scale and summarised by their means and 95% highest posterior density intervals, which we call Bayesian credible interval (BCI) from now on. Point estimates (means) admit the usual interpretation as in a frequentist analysis: the expected increase in the odds of being classified as a fast progressing individual per unit (standard deviation in this work) of the corresponding variable. Credible intervals with 95% coverage represent a contiguous region which contain the unobserved, true value of interest with at least 95% probability, given the model structure, prior assumptions and the data. They result in a more fine grained evaluation of uncertainty compared to confidence intervals due to the associated posterior distribution of the quantity of interest. While confidence intervals only yield the information whether a specific value is contained or not, the allocation of posterior probability mass allows to study how likely a specific value actually is. Predictions in this framework can be obtained using the coefficient posteriors and yield a distribution of predictions for each observation.

Model Structures

As a baseline for subsequent evaluation of more elaborate models, we fitted a clinical model using weakly informative Student-t prior distributions, which corresponded to a standard frequentist analysis of the data (see Supplementary Table S1 in Supplement section 4).
To elucidate the added benefit of using biomarker on top of clinical covariates, we present several further types of models, differing by the variables contained and the prior choices. First, models which included only one group of variables (single group models). This allowed comparisons between the clinical covariates and the individual biomarker platforms. Second, domain knowledge inspired models which included all variables and had coefficient priors adapted to the structure of the data (structured models). Third, models which contained all of the variables (combined models) but with similar prior choices as the single group models, which were uninformed about the structure of the data. Lastly, models which were based on

principal components (PCA) in order to make the number of variables in each group equal. The last two model types served mainly as sensitivity analysis and are reported in Supplement sections 7 and 8.
To make the effects of the variables comparable, we log2 transformed all of the biomarkers to achieve symmetric distributions and then centered and scaled all of the candidate predictors (including clinical) to have zero mean and unit variance. Hence, all odds ratios are to be interpreted in units of standard deviations and are denoted as standardized odds ratios (sOR).
In contrast to the mixed model approach of our previous analysis on a subset of the data (S1) we decided to use logistic regression in the current setting. The main reason for that was the large number of covariates, which made the computational cost and complexity of mixed model inference infeasible. Furthermore, by resorting to a logistic regression we could focus on the association with the eGFR trajectories, which are less influenced by baseline eGFR values. Lastly, we could avoid the re-estimation of possibly unstable eGFR slopes and make use of the carefully established individual eGFR trajectory classification.

Choice of Priors

The choice of priors expressed our belief that only few biomarkers have a relevant individual association with the outcome of interest. This sparsity assumption was crucial in the setting of our current analysis, as we had more variables than "events" (i.e. the size of the smaller person group) – hence, some form of prior information was necessary to make the estimation of odds ratios robust. This was achieved by the use of coefficient priors which induced shrinkage or regularization of coefficient estimates towards zero (corresponding to an odds ratio of unity). Our main analyses were based on regularized horseshoe prior distributions (S2). These so called global local shrinkage priors assign a normal distribution to each of the coefficients, centered around zero and with a variance depending on the data. The variance parameter consists of two parts: a global parameter τ that shrinks all of the predictor effects towards zero and a local parameter λ_i which allows certain variables to escape this global shrinkage. This prior was used by the single group, combined and PCA-based models. Each coefficient β_i, for all variables numbered from 1 to p, were assigned the same prior defined by

$$\beta_i \sim N\left(0, \tau^2 \tilde{\lambda_i}^2\right), \qquad \tilde{\lambda_i}^2 = \frac{c^2 \lambda_i^2}{c^2 + \tau^2 \lambda_i^2},$$
$$\lambda_i \sim Student - t_\nu^+\left(0, s_\lambda^2\right),$$
$$c^2 \sim inv - gamma(a, b),$$
$$\tau^2 \sim Student - t_1^+(0, s_\tau^2)$$

Here $Student - t_\nu^+$ denotes a half Student-t distribution with degrees of freedom ν. Several hyperparameters needed to be defined to complete the specification. For the global scale τ we followed the recommendations in (S3) and set $s_\tau = 1$ (note that in that case the assumed half-Student-t prior distribution is equal to the half-Cauchy distribution as found in the literature). For the local scale λ_i, we set $s_\lambda = 1$ and $\nu = 3$, the latter was done to improve convergence properties and speed of the Marko chain Monte Carlo (MCMC) sampling while having virtually no impact on predictive abilities. For the auxiliary parameter c, we set $a = b = 3$, again to improve sampling robustness. Alternative prior parameter settings were based on considerations in (S2) and are reported as sensitivity analysis (Supplement section 0). The structured models used a different shrinkage prior, which was similar to the approach presented in (S4). We added a further layer of shrinkage parameters between the global and local scale. It corresponds to a group scaling, as it is dependent on the group to which the predictor belongs to.

Hence, we defined

$$\beta_i \sim N\left(0, \tau^2 \tau_{G_i}^2 \tilde{\lambda}_i^{\,2}\right), \qquad \tilde{\lambda}_i^{\,2} = \frac{c^2 \lambda_i^2}{c^2 + \tau^2 \tau_{G_i}^2 \lambda_i^2},$$
$$\tau_G^2 \sim Student - t_1^+(0, s_\tau^2),$$

Where G represents the variable group (i.e. clinical, proteomics, metabolomics or lipidomics) and G_i denotes the group to which predictor i belongs to. The additional parameters for the group scale were set in the same way as for the global scale. This middle layer of shrinkage acts similar to “just another global” scale that modulates the overall shrinkage strength per group. The motivation was that we expected that discriminative power differs between the variable groups. The protein biomarkers were already selected in earlier studies and targeted at predicting the disease outcome, while the metabolomics and lipidomics did not undergo such a thorough selection process. Thus, they likely required stronger regularization to identify a small number of strong predictors. This idea is similar to the way IPF Lasso introduces modifiers for penalty factors depending on variable modality. The remaining hyperparameters were set as in the unstructured models.
The base clinical model without regularization used weakly informative Student-t priors with 3 degrees of freedom as priors, based on suggestions made in (S5). Its agreement with a frequentist analysis of the data is shown in

Supplementary Table S1 in Supplement section 4.

Model Performance

The ability to discriminate the two eGFR trajectory groups is expressed in terms of area under the receiver operating curve (AUROC), which was corrected for optimism using 5 times repeated 5-fold cross-validation. We approximated the full AUROC posterior distribution by doing 100 simulations from the posterior predictive densities for each fold in the repeated cross-validation.

Identification of Important Predictors

The drawback of using a Bayesian approach is that even though shrinkage priors shrink variable effects towards zero, they do not automatically select variables to produce truly sparse models, as the probability for a variable to have a coefficient larger than zero is always non-zero. Thus, to identify a smaller number of biomarkers which are important to discriminate between fast progressing and stable individuals, we elucidated the performance of smaller models as a subsequent step. We followed a projection approach, based on the methodology outlined in (S6; S7). In this feature selection step, the initial model fitted on all variables served as a reference model. A submodel is a model that contains only a subset of the variables of the full model. The goal was then to search for submodels whose predictive ability was as good as the reference model. A very simple search strategy was to define submodels by including variables ranked by their absolute marginal effect strength (absolute value of mean of coefficient posterior). The coefficient posteriors for this submodel were then obtained from a linear regression where the linear predictors of the reference model served as outcome and only the specified subset of variables entered the model as independent variables. This essentially corresponds to a projection of the predictions from the full model to the restricted parameter space of the submodel, and thus only includes a subset of the variables. We computed the AUROC for submodels up to size 16 (i.e. starting with the intercept only model up to including the top 15 predictors) in the cross-validation to identify how the discriminative accuracy changes when making predictions using only the projected posteriors of the strongest predictors. More sophisticated search strategies to identify submodels were considered and tested, but lead to similar conclusions, as we were only interested in the change of AUROC in the submodels and not their exact constitution. Furthermore, the generally low AUROC values and clear results regarding predictor importance were further reasons for the similar results concerning the first few predictors, regardless of method used. Thus, and because of the simplicity of the approach which can be easily reported, we present the “naïve” approach of ranking by marginal effects here.

Multiple Imputation

As reported in Table 1 and the availability breakdown in Supplement section 3, several of the predictors had missing data issues, especially the lipidomics data. Therefore we extended our multiple imputation procedure from (S1) to also include the novel metabolomics and lipidomics biomarkers. Multiple imputation was conducted using multivariate imputation by chained equations, as implemented in the package “mice” for the R statistical software. The imputation model included all candidate predictors as well as the outcome eGFR and eGFR slope to impute the data. The outcome variable itself was completely available. Since measurements of biomarkers out of quantifiable range were set to suitable values prior to this imputation procedure, we assumed the remaining data was missing at random. We chose to use 20 imputed datasets as a compromise between computational effort and statistical accuracy. Convergence of the multiple imputation procedure was checked using graphical methods provided by the “mice” package and found to be satisfactory. The Bayesian framework admits to incorporate the multiple imputed datasets naturally by

pooling the posteriors from models fitted to each imputed dataset. Thus, we obtained a final model which incorporates uncertainty due to missing data into the posterior. This model could then be used for further analysis.
We decided not to incorporate the imputation model into the Bayesian logistic models themselves for three reasons: first, we use the same imputed data for several model structures. This ensured comparability. Second, the computational effort required to impute the data in each model individually would have been prohibitive. Lastly, the model complexity would have increased, making it more difficult to sample robustly from these models.

Sensitivity Analysis for Prior Parameters

The sensitivity to the prior assumptions were thoroughly investigated. We found the results for the strong predictors to be very robust. Our choice of hyperparameters was guided mostly by recommendations in the literature and prior knowledge gained in (S1). We checked model robustness and aimed at preventing sampling issues as reported by the sampler diagnostics implemented in Stan (namely divergent transitions) as well as traditional MCMC diagnostics ($\hat{R}$ statistics, ratio of effective samples and total posterior samples).

Software and Code Used

Our analysis was implemented in R 3.4.1 (https://www.r-project.org/foundation, Vienna, AT), using Stan 2.18.0 (S8) for fitting the Bayesian models. The code to run our analysis for logistic regression models using regularized horseshoe priors can be found in Supplement section 11. Our implementation of the structured horseshoe models can be found in Supplement section 12. The code was based on the implementation examples in (S2).

Markov chain Monte Carlo Sampling

We used the NUTS algorithm of the Stan software package, accessed via the R interface rstan, to draw samples from the posterior distributions of our models. Each model for each imputed dataset used 3 chains for parallel sampling. We used 2000 warmup and 1000 sampling iterations. For the main tuning parameter adapt_delta we used a value of 0.9 to prevent issues during model sampling when using the regularized horseshoe priors, even if lower settings also led to robust sampling in most cases. Other settings were left at their default values.

Biomarker Measurement and Data Overview

Detailed Metabolomics Measurement Procedure

Metabolomics analysis was performed using targeted hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry (HILIC-LC-MS/MS) methodology. Plasma samples were thawed on ice, after which 300 µL of a pre-chilled mixture of acetonitrile:isopropanol:water (3:3:2 v/v) was added to 50 µL of each plasma sample. The samples were mixed thoroughly and incubated at -20°C overnight. The samples were then centrifuged for 15 min at 3200 x g at 4°C, and the supernatants were removed for LC-MS/MS analysis. A pooled QC sample was created by taking an aliquot of the same volume from all the extracted samples in the study. This pooled sample was then serially diluted to form a standard curve. The study samples were then further diluted 1:1 with the same pre-chilled acetonitrile:isopropanol:water mixture prior to analysis.
LC-MS/MS was performed using a Shimadzu Nexera UPLC system (Shimadzu, Kyoto, Japan) coupled with a SCIEX Triple Quadrupole 6500+ mass spectrometer (SCIEX Framingham, MA, USA). HILIC-LC separation was achieved using an apHera NH2 column (15cm × 2mm, 5 µm particle size, Supelco Analytical, Bellefonte, PA, USA) with an apHera NH2 guard column (1cm × 2 mm, 5 µm particle size, Supelco Analytical, Bellefonte, PA, USA). The mobile phases were 50 mM ammonium bicarbonate (pH 9.5, adjusted with ammonium hydroxide) (A) and acetonitrile (B). After a 3-min isocratic run at 90% B, gradients to 85%, 75%, 45%, 30% and 2% B were concluded at 3, 11.5, 15, 20 and 21 min, respectively. A gradient to 90% B was completed at 23.5 min and was followed by a 4-min column equilibration. The flow rate started at 0.25 mL/min and increased to 0.35 mL/min at 20 min and 0.4 mL/min at 21 min and then went back to 0.25 mL/min at the end of the equilibration.
Positive and negative scheduled MRM transitions were monitored for a total of 215 metabolites. The metabolite peaks were integrated, and the AUCs and relative intensities of the metabolites were calculated using SCIEX MultiQuant 3.0.2 software (SCIEX Framingham, MA, USA).

Detailed Lipidomics Measurement Procedure

Mass spectrometry-based lipid analysis was performed at Lipotype GmbH using high throughput Shotgun Lipidomics (Dresden, Germany) technology as described in (S9). Samples were processed in batches of 84 samples and each batch was accompanied by identical reference samples and blank controls. Lipids were extracted from an equivalent of 1 µL of undiluted plasma using methyl tert-butyl ether/methanol (7:2, V:V) as in (S10). Known amounts of internal standards were pre-mixed with the organic solvent mixture. All extraction steps were carried out on the Hamilton Robotics STARlet robotic platform with the Anti Droplet Control feature for organic solvents pipetting.
Samples were analyzed by direct infusion in a QExactive mass spectrometer (Thermo Scientific) equipped with a TriVersa NanoMate ion source (Advion Biosciences). Lipid identification and quantification from unprocessed mass spectra data was carried out using Lipotype's proprietary LipotypeXplorer (S11). For quantification lipid intensities were normalized to lipid-class specific standards. Only lipids at least five-times higher than in blank samples were kept for measurement samples. Correction for batch-effects and analytical drift was performed using the reference samples.
Analysis of reference samples verified high analytical quality and reproducibility. The median coefficient of variation across all classes and batches was 10.8%.

Biomarker Measurement Availability

Availability for protein marker measurements is reported in more detail in the Supplement of (1). In summary, each protein biomarker was available for at least 90% of the participants; 12 out of 15 markers were available for at least 95% of the participants.

Each metabolite biomarker was available for at least 98% of participants.

For lipid biomarkers the number of missing measurements was higher. We selected 197 out of 637 lipid markers for which at least 66% of participants had measurements available. This threshold is similar to other studies in the field. It presents a compromise between using as many markers as possible while providing the multiple imputation procedure with enough information for stable results and convergence. The amount of missing data was comparable between the eGFR trajectory groups.

Biomarker lists

The following lists give the names of all biomarker candidates used in the current analysis.

Proteomics (15 biomarkers)

CystatinC	MMP7
Endostatin	TIE2
UMOD	TNFR1
CCL2	VCAM1
CHI3L1	KIM1
GH	FGF23
HGF	NTproBNP
MMP1	

Metabolomics (180 biomarkers)

1-Methyladenosine	C3+O Carnitine	Kynurenic Acid
1-Methyl-Histidine	C4 Carnitine	Kynurenine
2,3-Dihydroxybenzoic Acid	C5:0 C4+O Carnitine	Lactate
2-Aminobutyrate 58	C5:1 Carnitine	Leucine
2-Aminooctanoic Acid	C5 Sa 89	Lpc 16:0
2-Hydroxybutyric Acid	C5+0 Carnitine	Lysine
2-Isopropylmalic Acid	C6:0 Carnitine	Malate
3-Hydoxy-Dl-Kynurenine 162	C6:0+O Carnitine	Methionine
3-Hydroxybutyric Acid	C6:1 Carnitine	Methionine Sulfoxide
3-Phosphoglycerate	C6:1+O Carnitine	Methylnicotinamide
4-Pyridoxic Acid	C6 Sa 89	Mevalonic Acid 59.1
6-Phospho-D-Glucono-1,5-Lactone	C7:0+O Carnitine	Myo-Inositol
7-Methylguanosine	C8:0 Carnitine	N-Acetyl-Dl-Serine
Acetoacetate	C8:0+O Carnitine	N-Acetyl-Glutamine
Acetylcarnitine Dl	C8:1 Carnitine	N-Acetyl-L-Alanine
Acetyllysine	C8:1+O Carnitine	Nadp+ Pos
Acetylphosphate	Carnitine	Nadph
Adenine	Cdp	N-Carbamoyl-L-Aspartate Neg
Adp	Chenodeoxycholate	N-Formyl Kynurenine 146
A-Ketoglutarate	Cholesteryl Sulfate	Ng,Ng-Dimethyl-L-Arginine

Alanine	Choline	Nicotinamide
Allantoate	Citrulline	O-Acetyl-L-Serine
Allantoin	Creatine	Octulose-1,8-Bisphosphate (Obp)
Aminoadipic Acid	Creatinine	O-Phosphorylethanolamine
Amp	Cyclic Bis(3->5) Dimeric Gmp	Ornithine
Arginine	Cyclic-Amp	Orotate
Asparagine	Cysteine	Pantothenate
Aspartate	Cystine	Phenylalanine
Betaine	Datp	Phenyllactic Acid
Betaine Aldehyde	Dctp	Phenylpropiolic Acid
C0 Carnitine	Deoxyribose-Phosphate	Phosphoenolpyruvate
C10:0 Carnitine	D-Glucarate	Phosphorylcholine
C10:0+O Carnitine	D-Gluconate	P-Hydroxybenzoate
C10:1 Carnitine	D-Glucosamine-6-Phosphate	Proline
C10:2 Carnitine	D-Glyceraldehdye-3-Phosphate	Purine
C10:3 Carnitine	Dimethylglycine	Pyroglutamic Acid
C12:0 Carnitine	Dl-3-Aminobutyrate	Riboflavin
C12:1 Carnitine	Dl-Pipecolic Acid	Ribose-Phosphate
C12:1+O Carnitine	Dttp	S1p
C14:0 Carnitine	Erythritol 89	S1p-Neg
C14:0+O Carnitine	Ethanolamine	Sarcosine
C14:1 Carnitine	Glutamate	Serine
C14:2 Carnitine	Glutamine	Shikimate
C16:0 Carnitine	Glutathione Neg	Shikimate-3-Phosphate
C16:0+O Carnitine	Glycerate	Succinate
C16:1 Carnitine	Glycerophosphocholine	Taurine
C16:1+O C17:0 Carnitine	Glycine	Taurodeoxycholate
C18:0 Carnitine	Glycolate	Tetradecanedioic Acid 239
C18:0+O Carnitine	Guanidoacetic Acid	Threonine
C18:1 Carnitine	Guanine	Trans, Trans-Farnesyl Diphosphate
C18:1+O Carnitine	Histidine	Tryptophan
C18:2 Carnitine	Hydroxyisocaproic Acid	Tyrosine
C18:3 Carnitine	Hydroxyproline	Uracil
C2 Carnitine	Hypoxanthine	Urea
C2+O Carnitine	Imidazoleacetic Acid	Uric Acid
C20:0 Carnitine	Indole	Valine
C20:0+O Carnitine	Indole-3-Carboxylic Acid	Xanthine
C20:3 Carnitine	Indoleacrylic Acid	Xanthosine
C20:4 Carnitine	Inosine	Xanthurenic Acid
C3 Carnitine	Isoleucine	Ztp 78.9

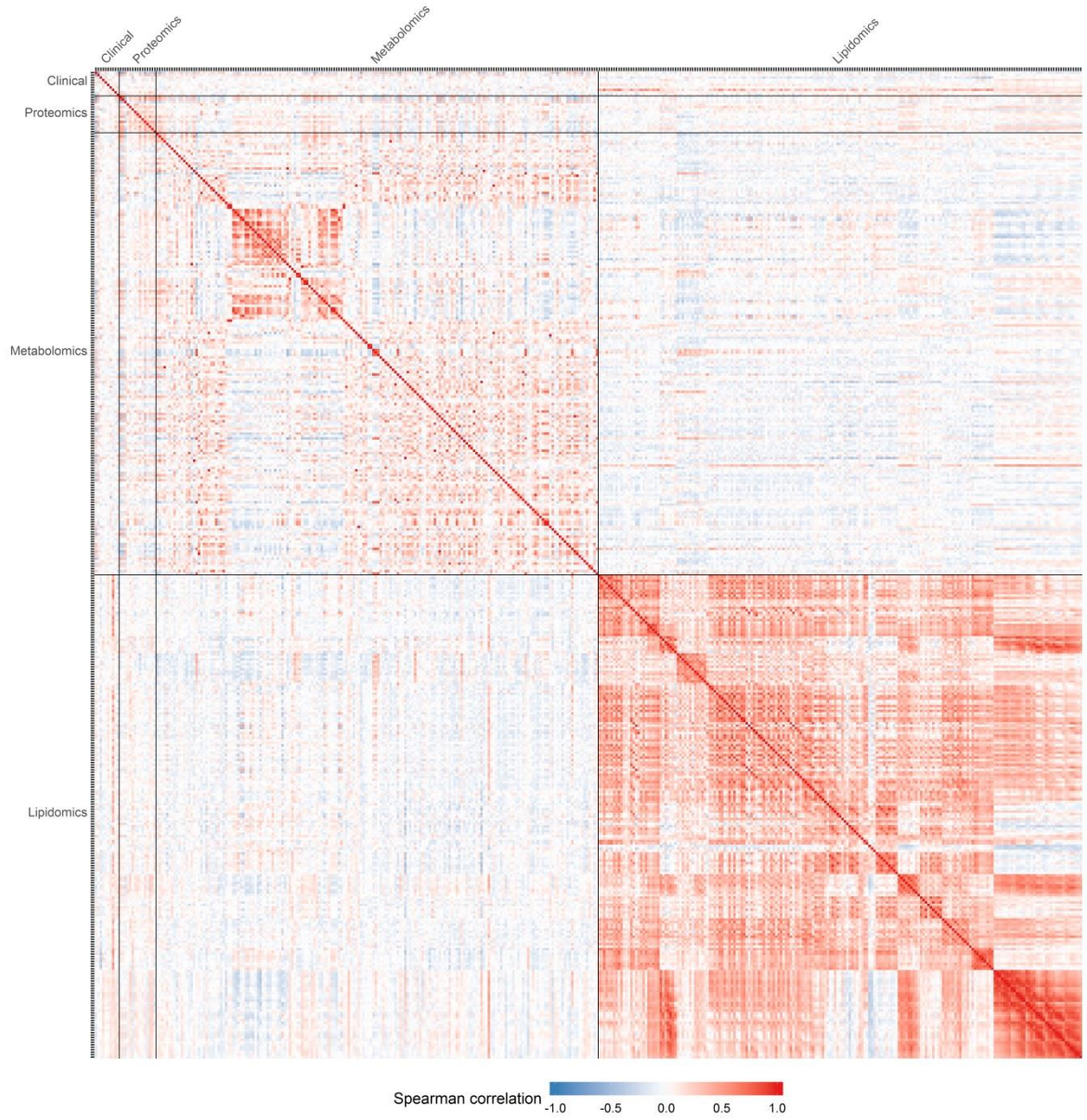
Lipidomics (197 biomarkers)

CE 14:0;0	PC 16:1;0_18:1;0	PE O-16:1;0/20:3;0
CE 15:0;0	PC 16:1;0_18:2;0	PE O-16:1;0/20:4;0
CE 16:0;0	PC 16:1;0_20:4;0	PE O-16:1;0/22:5;0
CE 16:1;0	PC 17:0;0_18:1;0	PE O-18:1;0/18:2;0
CE 17:0;0	PC 17:0;0_18:2;0	PE O-18:1;0/20:4;0

CE 17:1;0	PC 17:0;0_20:3;0	PE O-18:2;0/18:1;0
CE 18:0;0	PC 17:0;0_20:4;0	PE O-18:2;0/18:2;0
CE 18:1;0	PC 18:0;0_18:1;0	PE O-18:2;0/20:4;0
CE 18:2;0	PC 18:0;0_18:2;0	PI 16:0;0_18:1;0
CE 18:3;0	PC 18:0;0_18:3;0	PI 16:0;0_18:2;0
CE 19:2;0	PC 18:0;0_20:2;0	PI 16:0;0_20:3;0
CE 20:1;0	PC 18:0;0_20:3;0	PI 16:0;0_20:4;0
CE 20:2;0	PC 18:0;0_20:4;0	PI 18:0;0_18:1;0
CE 20:3;0	PC 18:0;0_20:5;0	PI 18:0;0_18:2;0
CE 20:4;0	PC 18:0;0_22:4;0	PI 18:0;0_20:3;0
CE 20:5;0	PC 18:0;0_22:5;0	PI 18:0;0_20:4;0
CE 22:6;0	PC 18:0;0_22:6;0	PI 18:1;0_18:1;0
CE 23:2;0	PC 18:1;0_18:1;0	PI 18:1;0_18:2;0
CE 24:1;0	PC 18:1;0_18:2;0	PI 18:1;0_20:4;0
CE 24:2;0	PC 18:1;0_20:2;0	PI 18:2;0_18:2;0
Cer 40:1;2	PC 18:1;0_20:3;0	SM 32:1;2
Cer 40:2;2	PC 18:1;0_20:4;0	SM 34:1;2
Cer 42:1;2	PC 18:2;0_18:2;0	SM 34:2;2
Cer 42:2;2	PC 18:2;0_20:1;0	SM 36:1;2
Chol	PC 18:2;0_20:3;0	SM 36:2;2
DAG 16:0;0_18:1;0	PC 18:2;0_20:4;0	SM 38:2;2
DAG 16:0;0_18:2;0	PC O-16:0;0/16:0;0	SM 40:1;2
DAG 16:1;0_18:1;0	PC O-16:0;0/16:1;0	SM 40:2;2
DAG 18:1;0_18:1;0	PC O-16:0;0/18:1;0	SM 42:2;2
DAG 18:1;0_18:2;0	PC O-16:0;0/18:2;0	TAG 46:0;0
DAG 18:1;0_18:3;0	PC O-16:0;0/20:3;0	TAG 46:1;0
DAG 18:2;0_18:2;0	PC O-16:0;0/20:4;0	TAG 46:2;0
LPC 16:0;0	PC O-16:0;0/22:4;0	TAG 48:0;0
LPC 16:1;0	PC O-16:0;0/22:5;0	TAG 48:1;0
LPC 18:0;0	PC O-16:1;0/16:0;0	TAG 48:2;0
LPC 18:1;0	PC O-16:1;0/18:0;0	TAG 48:3;0
LPC 18:2;0	PC O-16:1;0/18:1;0	TAG 49:1;0
LPC 20:3;0	PC O-16:1;0/18:2;0	TAG 49:2;0
LPC 20:4;0	PC O-16:1;0/20:3;0	TAG 50:1;0
LPE 16:0;0	PC O-16:1;0/20:4;0	TAG 50:2;0
LPE 18:0;0	PC O-16:2;0/18:0;0	TAG 50:3;0
LPE 18:1;0	PC O-16:2;0/18:1;0	TAG 50:4;0
LPE 18:2;0	PC O-17:0;0/15:0;0	TAG 51:1;0
LPE 20:4;0	PC O-17:0;0/17:1;0	TAG 51:2;0
LPE 22:6;0	PC O-17:2;0/17:0;0	TAG 51:3;0
PC 14:0;0_16:0;0	PC O-18:0;0/14:0;0	TAG 52:2;0
PC 14:0;0_18:1;0	PC O-18:0;0/16:1;0	TAG 52:3;0
PC 14:0;0_18:2;0	PC O-18:0;0/20:4;0	TAG 52:4;0
PC 15:0;0_18:1;0	PC O-18:1;0/16:0;0	TAG 52:5;0
PC 15:0;0_18:2;0	PC O-18:1;0/18:2;0	TAG 52:6;0
PC 16:0;0_16:0;0	PC O-18:1;0/20:3;0	TAG 53:2;0
PC 16:0;0_16:1;0	PC O-18:1;0/20:4;0	TAG 53:3;0
PC 16:0;0_17:1;0	PC O-18:2;0/16:0;0	TAG 53:4;0
PC 16:0;0_18:0;0	PC O-18:2;0/18:1;0	TAG 54:3;0
PC 16:0;0_18:1;0	PC O-18:2;0/18:2;0	TAG 54:4;0
PC 16:0;0_18:2;0	PC O-18:2;0/20:4;0	TAG 54:5;0
PC 16:0;0_18:3;0	PE 16:0;0_18:1;0	TAG 54:6;0
PC 16:0;0_20:1;0	PE 16:0;0_18:2;0	TAG 54:7;0

PC 16:0;0_20:2;0	PE 16:0;0_20:4;0	TAG 56:3;0
PC 16:0;0_20:3;0	PE 18:0;0_18:1;0	TAG 56:4;0
PC 16:0;0_20:4;0	PE 18:0;0_18:2;0	TAG 56:5;0
PC 16:0;0_20:5;0	PE 18:0;0_20:4;0	TAG 56:6;0
PC 16:0;0_22:4;0	PE 18:1;0_18:1;0	TAG 56:7;0
PC 16:0;0_22:5;0	PE 18:1;0_18:2;0	TAG 56:8;0
PC 16:0;0_22:6;0	PE 18:1;0_20:4;0	TAG 58:8;0
PC 16:1;0_18:0;0	PE O-16:1;0/18:2;0	

Correlation Structure



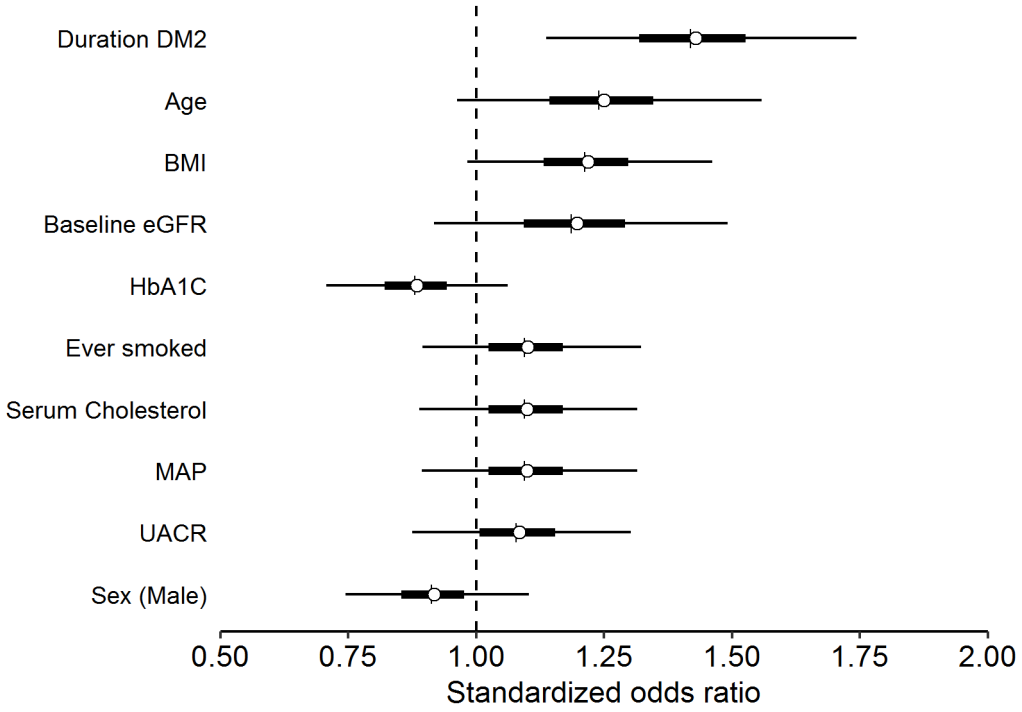
Supplementary Figure S4: Correlation matrix of all candidate predictors

Correlation matrix with Spearman correlation coefficients for all variables. Variable groups are represented by blocks separated by black lines. The matrix comprises 10 clinical, 15 protein, 180 metabolite and 197 lipid variables, respectively.

Single Group Models

This section includes the results for models comprising only a single group of variables. We report coefficients of the top 15 predictors in terms of absolute effect strength (absolute marginal posterior means). Model performance in terms of cross-validated AUROC is given in the main manuscript (Table 2).

Clinical Model



Supplementary Figure S5: Standardized odds ratios for the clinical model

Coefficient posteriors for clinical model with weakly informative Student-t priors, equivalent to a frequentist analysis, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval.

As a comparison to the Bayesian approach we also fitted frequentist logistic regression models to the multiple imputed datasets. Final coefficient estimates were pooled using Rubin's rules.

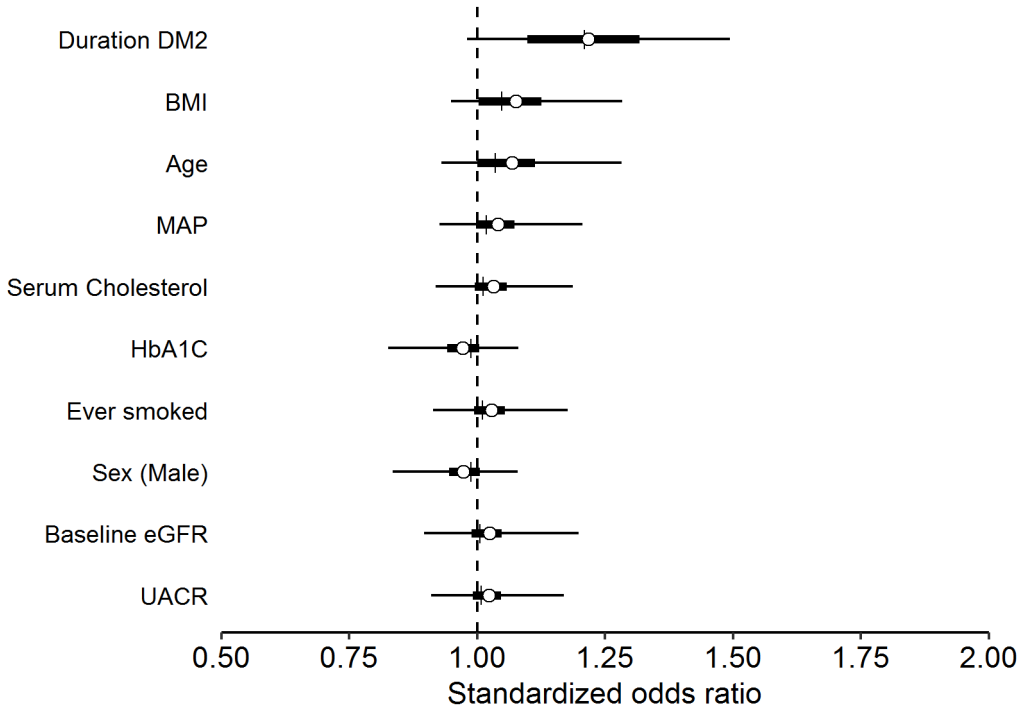
Supplementary Table S1 shows that the two modelling approaches led to virtually identical results in this simple baseline case.

Supplementary Table S1: Standardized odds ratios for the clinical model (Frequentist and Bayes estimates)

Standardized odds ratios (sOR) of progression for clinical risk factors. The Bayesian model used weakly informative Student-t priors and the reported odds ratio is the mean of the model posterior.

Variable	Frequentist Model			Bayesian Model		
	sOR	95% Confidence interval		sOR	95% Bayesian credible interval	
		Lower	Upper		Lower	Upper
Duration DM2	1.41	1.14	1.74	1.43	1.14	1.74
Age	1.23	0.97	1.56	1.25	0.96	1.56
BMI	1.21	0.99	1.47	1.22	0.98	1.46
Baseline eGFR	1.18	0.93	1.5	1.2	0.92	1.49
HbA1C	0.88	0.73	1.08	0.88	0.71	1.06
Ever smoked	1.09	0.9	1.32	1.1	0.89	1.32
Serum Cholesterol	1.09	0.9	1.32	1.1	0.89	1.31
MAP	1.09	0.9	1.32	1.1	0.89	1.31
UACR	1.08	0.89	1.31	1.08	0.87	1.3
Sex (Male)	0.91	0.75	1.11	0.92	0.74	1.10

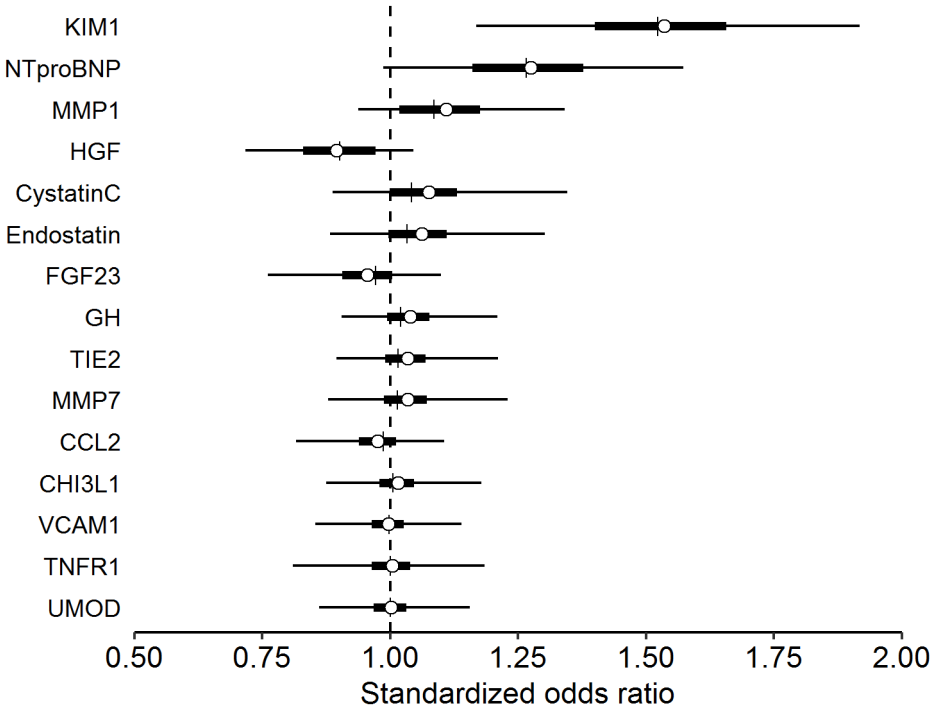
Clinical Model with Shrinkage



Supplementary Figure S6: Standardized odds ratios for the clinical model with shrinkage priors

Coefficient posteriors for clinical model with shrinkage priors, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval.

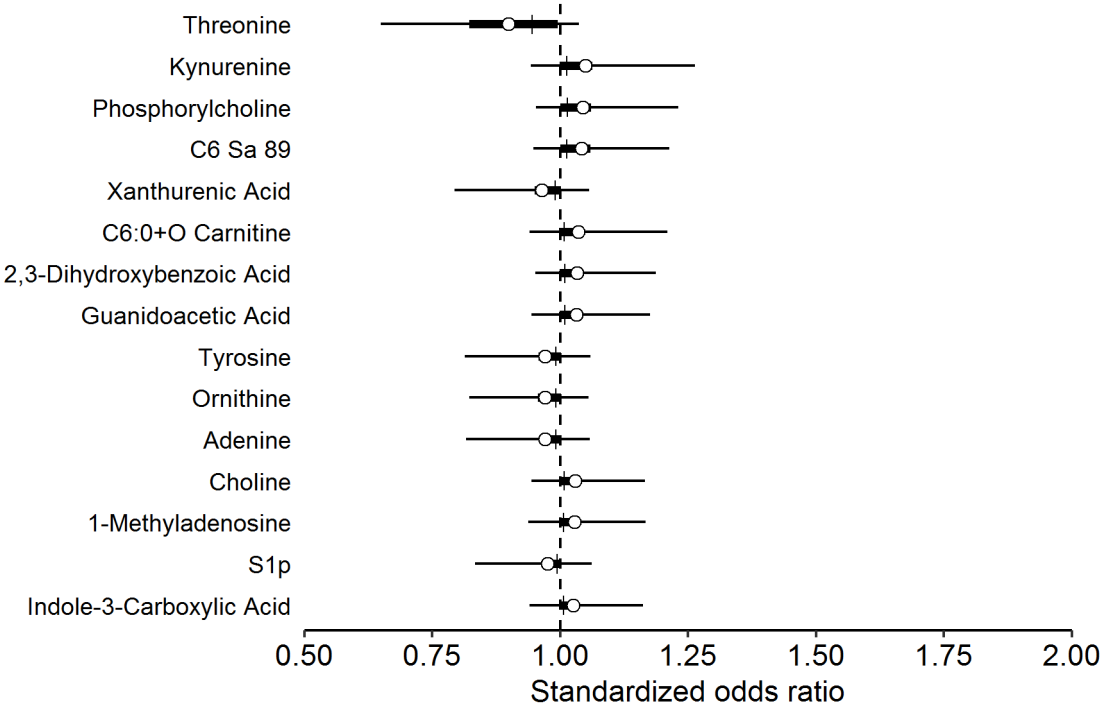
Proteomics Model



Supplementary Figure S7: Standardized odds ratios for the proteomics model

Coefficient posteriors for proteomics model, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval.

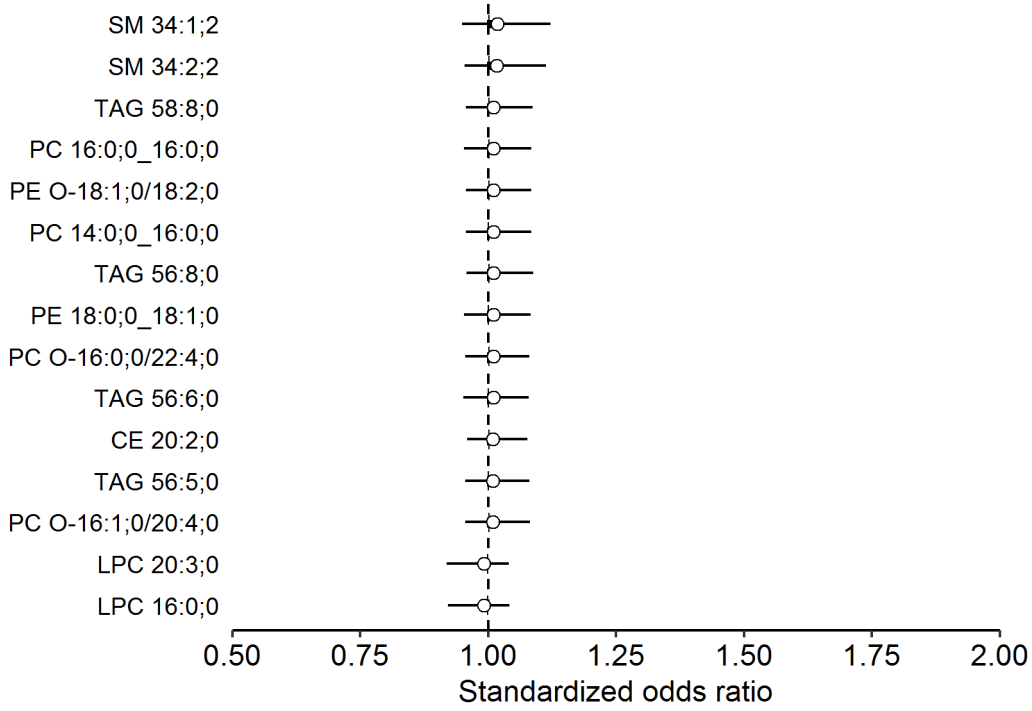
Metabolomics Model



Supplementary Figure S8: Standardized odds ratios for the metabolomics model

Coefficient posteriors for the top 15 predictors of the metabolomics model, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval.

Lipidomics Model



Supplementary Figure S9: Standardized odds ratios for the lipidomics model

Coefficient posteriors for the top 15 predictors of the lipidomics model, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval.

Main Structured Model Coefficients

Table with coefficient posteriors for the top 15 predictors from the main structured model (Figure 1 in main manuscript).

Supplementary Table S2: Standardized odds ratios for the main structured model

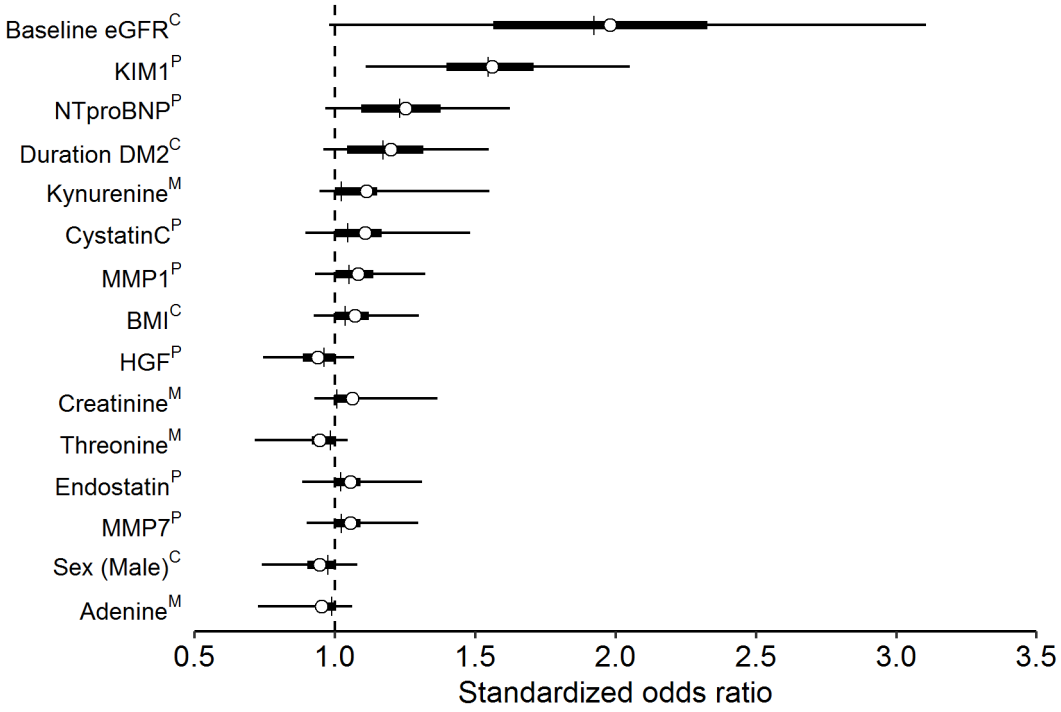
Summary of posteriors for standardized odds ratios of progression for top predictors of the main structured model. We report mean, median, quantiles and 95% Bayesian credible interval (BCI) of the model posterior.

Variable	Group	Mean	Median	Quantiles		95% BCI	
				25%	75%	Lower	Upper
Baseline eGFR	Clinical	2.23	2.14	1.79	2.56	1	3.4
KIM1	Proteomics	1.59	1.57	1.42	1.74	1.13	2.09
NTproBNP	Proteomics	1.28	1.26	1.13	1.41	0.97	1.65
Duration DM2*	Clinical	1.25	1.24	1.11	1.37	0.97	1.61
Kynurenine	Metabolomics	1.14	1.04	1	1.21	0.94	1.61
CystatinC	Proteomics	1.13	1.06	1	1.2	0.87	1.53
MMP1	Proteomics	1.1	1.07	1.01	1.16	0.93	1.35
BMI	Clinical	1.09	1.07	1.01	1.16	0.92	1.34
Creatinine	Metabolomics	1.09	1.02	1	1.09	0.92	1.49
Sex (Male)	Clinical	0.92	0.95	0.86	1	0.70	1.09
HGF	Proteomics	0.92	0.94	0.86	1	0.72	1.07
MMP7	Proteomics	1.07	1.03	1	1.11	0.88	1.32
Endostatin	Proteomics	1.06	1.03	0.99	1.11	0.87	1.34
Threonine	Metabolomics	0.94	0.98	0.91	1	0.72	1.05
C6:0+O Carnitine	Metabolomics	1.06	1.01	1	1.07	0.93	1.31

* denotes duration of type 2 diabetes mellitus at study entry

Sensitivity Analysis – Prior Parameter

Results for a model using same structured prior as the main structured model, but with different hyperparameter settings. In this model we set the global scale to a much smaller value following the suggestions in (S2). We set the expected number of large coefficients to be p_0 = 15 and used formula 3.12 in the logistic regression setting to derive a global scale which results in much stronger expected shrinkage than in the setting of the main model. The coefficient posteriors for the top predictors were largely unaffected by this change in prior settings. They were insensitive regarding other settings of p_0 as well. The discriminative performance was also virtually identical (AUROC 0.62 [0.5, 0.73]).



Supplementary Figure S10: Standardized odds ratios for alternative structured model (different hyperparameter)

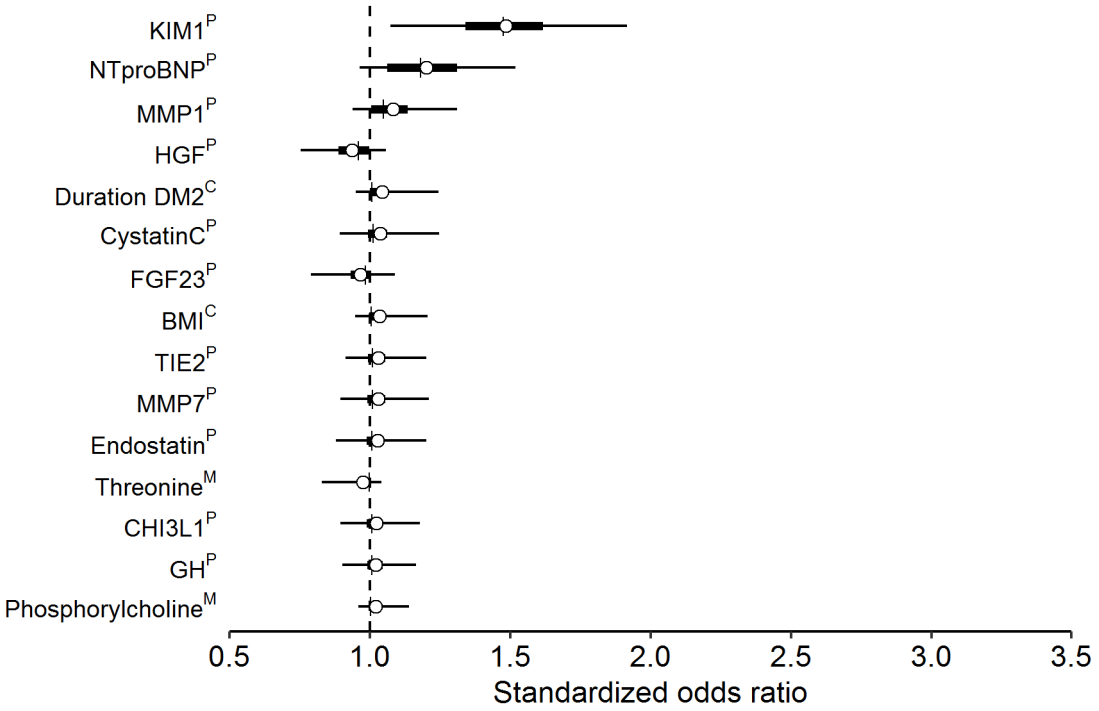
Coefficient posteriors for the top 15 predictors of the structured model using prior parameter which lead to much stronger expected shrinkage, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval. Group assignments are indicated by superscript initials (C for Clinical, P for Proteomics, M for Metabolomics, L for Lipidomics).

We studied other prior parameter settings as well (different group and local scales) and found no large differences between the models for the top predictors, as long as convergence was not hindered by too stringent prior parameters.

Sensitivity Analysis – Excluded Baseline eGFR

Results for a model of the same structure and prior settings as the main structured model but excluding baseline eGFR as a variable. Comparing the coefficient posteriors from the main model, the large effect of eGFR on the other variables is demonstrated. Excluding this predictor led to generally lower effect sizes for most of the other variables, except for KIM1 and NTproBNP which were relatively unaffected. This indicated that their contribution to the discrimination of eGFR trajectories was largely independent of the presence of baseline eGFR in the model. Other protein biomarkers such as HGF and MMP1 showed similar results compared to the main model, while the effect of the clinical covariate duration of diabetes decreased.

The discriminative performance of the model was slightly lower than for the main structured model (AUROC 0.61 [0.49, 0,72]).

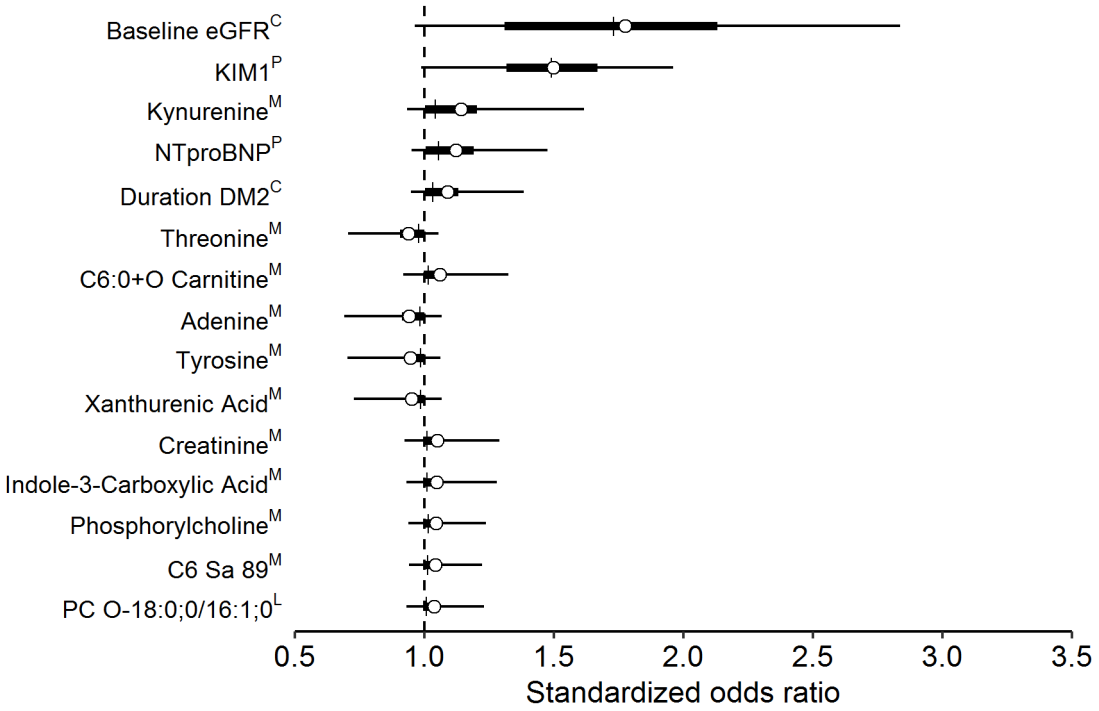


Supplementary Figure S11: Standardized odds ratios for the structured model excluding baseline eGFR

Coefficient posteriors for the top 15 predictors of the structured model excluding baseline eGFR as covariate, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval. Group assignments are indicated by superscript initials (C for Clinical, P for Proteomics, M for Metabolomics, L for Lipidomics).

Sensitivity Analysis – Prior Structure

Results for a model using priors that were uninformed regarding the group structure of the data. We set the global scale to one, as in the main structured model. The coefficient posteriors for baseline eGFR and KIM1 were largely unaffected by this change. Some metabolite and lipid markers increased in marginal effect size which can be explained by weaker regularization of these platforms in this model. However, the discriminative performance was slightly lower than for the main structured model (AUROC 0.6 [0.48, 0.71]). A possible explanation is that these markers showed increased effect sizes due to their correlation with the effect of baseline eGFR, while not contributing much to the discrimination performance, as indicated by the single group models. Hence, we favoured the main structured model over this approach.

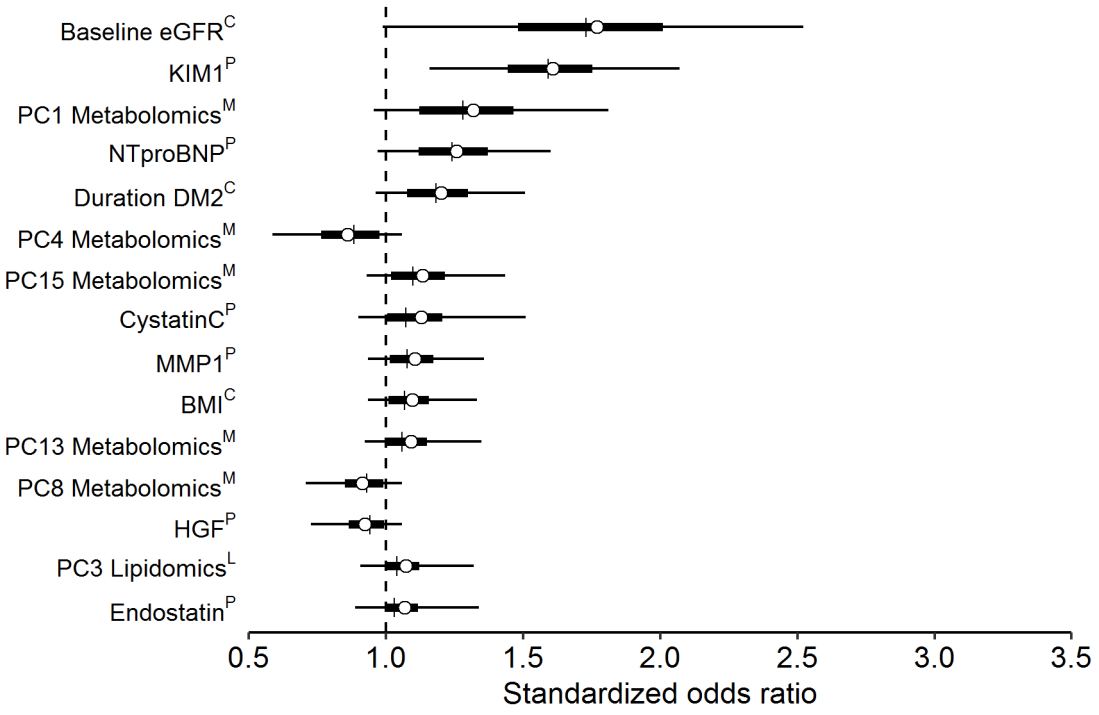


Supplementary Figure S12: Standardized odds ratios for model with alternative prior

Coefficient posteriors for the top 15 predictors of the combined model using a prior uninformed of the structure of the data, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval. Group assignments are indicated by superscript initials (C for Clinical, P for Proteomics, M for Metabolomics, L for Lipidomics).

Sensitivity Analysis – Principal Components Analysis

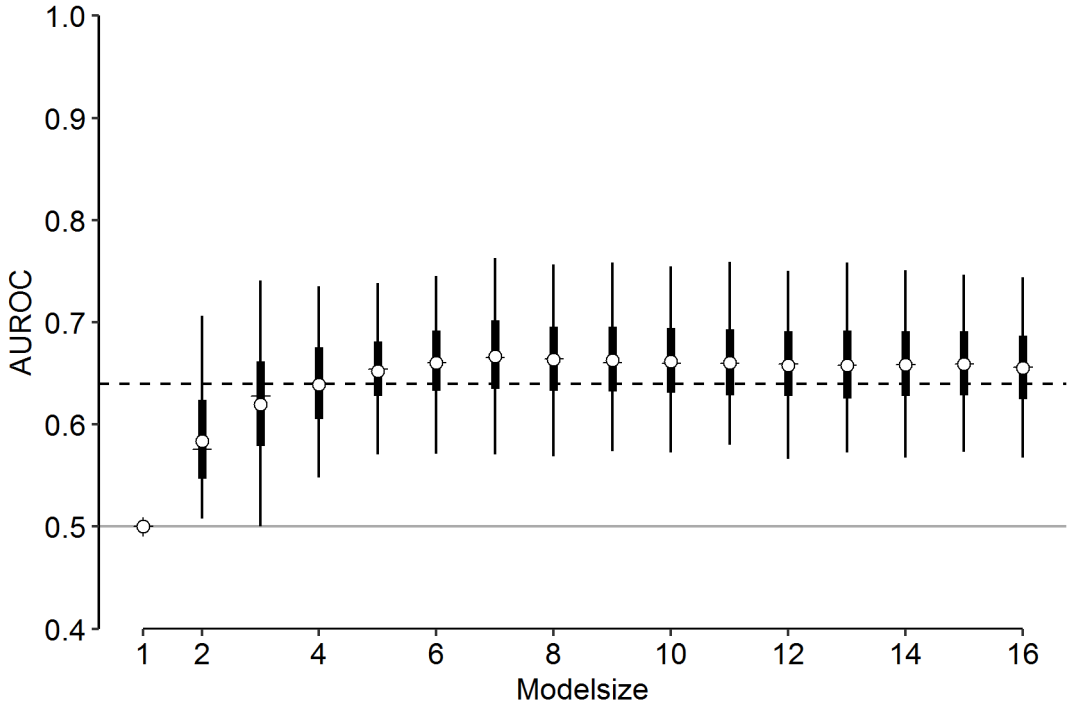
Result for a model based on principal component analysis (PCA) derived predictors. First, a PCA was applied to the large metabolite and lipid biomarker platforms. We chose the 15 principal components (PC) with largest explained variability for both platforms and used these instead of individual biomarkers in a model using a prior uninformed of the data structure. The rationale was to make the number of predictors for each variable type comparable to ensure that this did not influence the results – now all biomarker platforms contributed 15 predictors. For metabolites, around 60% of the total variability were captured by the 15 PCs, for lipid markers around 75%.



Supplementary Figure S13: Standardized odds ratios for model based on principal components

Coefficient posteriors for the top 15 predictors of the PCA model, ranked by marginal means. Point estimate is given by the mean (circle), the median is indicated by a vertical line. The thick horizontal line represents the interquartile range, while the thin line represents a 95% Bayesian credible interval. Group assignments are indicated by superscript initials (C for Clinical, P for Proteomics, M for Metabolomics, L for Lipidomics).

While the clinical covariates and protein markers were largely unaffected by the replacement of individual markers with PCA derived predictors, some of the metabolite PCs showed up with large marginal effects. These strongly correlated with baseline eGFR and were thus reliant on its presence in the model. Overall, the discrimination did not improve when using this approach than compared to the individual marker based modelling (AUROC 0.64 [0.55, 0.74]). Neither did the change in AUROC for submodels (**Supplementary Figure S14**) including fewer markers led to different conclusions than for the main structured model – again, only a few predictors sufficed to achieve performance comparable to the full model.
We therefore favoured the structured model and report it as our main result as its predictions were based on individual variables, which eases interpretation and handling of the data. Furthermore, it led to more parsimonious results because the PCA derived predictors incorporate the whole platform in each single PC.



Supplementary Figure S14: Change in model performance using limited predictors in the model based on principal components

Change in model performance when only a limited number of predictors is used to discriminate eGFR trajectory groups in the PCA model. Circles depict a point estimate of the cross-validated AUROC (posterior mean), the posterior median is indicated by a horizontal bar. The thick line represents the interquartile range and the thin line a 95% Bayesian credible interval. The horizontal dashed line indicates the AUROC of the main structured model using all 55 variables, which serves as a reference. Modelsize denotes the number of variables in the model, starting with the intercept only model, then adding variables in the order as shown in Supplementary Figure S13 (starting at the top with baseline eGFR). The predictions of the reference model were than projected onto the corresponding subset of its variables via the methodology outlined in the Statistical Methods section in order to obtain the AUROC values depicted.

Model Convergence

All of the Bayesian models were found to converge satisfactorily by manual inspection of several common convergence and efficiency diagnostics at a per-parameter level. The parameters most difficult to sample were related to baseline eGFR due to high correlations with other variables. Nevertheless, acceptable diagnostics were reached for all parameters. Below we give high-level summaries for the main statistics which indicate robust sampling from the model posteriors:

- $\hat{R}$ statistics for all parameters summarised by mean and upper 97.5% quantile. A value of one indicates convergence.
- Ratio of estimated effective sample size and number of draws from the posterior distribution (N_{eff} ratio) for all variables summarised by mean and lower 2.5% quantile. Higher values are better and indicate low autocorrelation of the MCMC draws.
- Percentage of divergent transitions of the Hamiltonian MCMC procedure implemented in Stan. Ideally there are no divergent transitions.

Supplementary Table S3: Bayesian model diagnostics

High-level summary of standard diagnostics for all of the models presented in this work.

Model	$\widehat{R}$		N_{eff} ratio		% divergent
	Mean	97.5% quantile	Mean	2.5% quantile	
Clinical	1.000	1.001	1.517	0.637	0
Clinical with shrinkage	1.001	1.004	0.778	0.316	0.002
Proteomics	1.000	1.003	0.838	0.409	0.002
Metabolomics	1.000	1.002	0.972	0.605	0
Lipidomics	1.000	1.002	0.848	0.519	0
Main Structured	1.000	1.002	0.863	0.489	0
Structured – alternative prior	1.000	1.002	0.869	0.465	0
Structured – excluding eGFR	1.000	1.002	0.933	0.605	0
Combined	1.000	1.002	0.909	0.510	0
Combined - PCA	1.000	1.003	0.815	0.392	0.002

Stan Implementation of Regularized Horseshoe Priors

Below is the code used for the combined models with a prior which is uninformed of the grouped data structure. The implementation is following the one given in (S2).

```
data {
    int<lower=0> n;
    int<lower=0> p;
    int<lower=0, upper=1> group[n];
    matrix[n,p] X;

    real<lower=0> scale_intercept;

    real<lower=0> scale_global;
    real<lower=1> df_global;
    real<lower=1> df_local;
    real<lower=0> scale_slab;
    real<lower=0> df_slab;
}
parameters {
    real intercept;
    vector[p] beta_tilde;
    real<lower=0> aux1_global;
    real<lower=0> aux2_global;
    vector<lower=0>[p] aux1_local;
    vector<lower=0>[p] aux2_local;
    real<lower=0> caux;
}
transformed parameters {
    real<lower=0> tau;
    vector<lower=0>[p] lambda;
    vector<lower=0>[p] lambda_tilde;
    real<lower=0> c;

    vector[p] beta;
    vector[n] mu;
```

```
    tau = aux1_global * sqrt(aux2_global) * scale_global;
    c = scale_slab * sqrt(caux);

    lambda = aux1_local .* sqrt(aux2_local);
    lambda_tilde = sqrt( square(c) * square(lambda) ./
                (square(c) + square(tau) * square(lambda)) );
    beta = beta_tilde .* lambda_tilde * tau;
    mu = intercept + X * beta;
}
model {
    target += std_normal_lpdf(beta_tilde);
    target += std_normal_lpdf(aux1_local);
    target += inv_gamma_lpdf(aux2_local | 0.5 * df_local, 0.5 *
df_local);
    target += std_normal_lpdf(aux1_global);
    target += inv_gamma_lpdf(aux2_global | 0.5 * df_global, 0.5 *
df_global);
    target += inv_gamma_lpdf(caux | 0.5 * df_slab, 0.5 * df_slab);

    target += normal_lpdf(intercept | 0, scale_intercept);
    target += bernoulli_logit_lpmf(group | mu);
}
generated quantities {
    vector[n] log_lik;

     for (i in 1:n) {
        log_lik[i] = bernoulli_logit_lpmf(group[i] | mu[i]);
  }
}
```

Stan Implementation of the Structured Regularized Horseshoe Priors

Below is the code used for the combined models with a prior which is adapted to the grouped data structure by the use of additional parameters, as described in the extended statistical methods (Supplement section 2).

```
data {
    int<lower=0> n;
    int<lower=0> p;
    int<lower=0, upper=1> group[n];
    matrix[n,p] X;

    int<lower=0> p_vgroup;
    int<lower=0> vgroup[p];

    real<lower=0> scale_intercept;

    real<lower=0> scale_global;
    real<lower=1> df_global;
    real<lower=1> df_local;
    real<lower=0> scale_slab;
    real<lower=0> df_slab;
    real<lower=0> df_vgroup;
    real<lower=0> scale_vgroup;
}
parameters {
    real intercept;
    vector[p] beta_tilde;
    real<lower=0> aux1_global;
    real<lower=0> aux2_global;
    vector<lower=0>[p_vgroup] aux1_vgroup;
    vector<lower=0>[p_vgroup] aux2_vgroup;
    vector<lower=0>[p] aux1_local;
    vector<lower=0>[p] aux2_local;
    real<lower=0> caux;
}
transformed parameters {
```

```
    real<lower=0> tau;

    vector<lower=0>[p_vgroup] tau_vgroup;

    vector<lower=0>[p] lambda;
    vector<lower=0>[p] lambda_tilde;
    real<lower=0> c;

    vector[p] beta;
    vector[n] mu;

    tau = aux1_global * sqrt(aux2_global) * scale_global;

    tau_vgroup = aux1_vgroup .* sqrt(aux2_vgroup) * scale_vgroup;

    c = scale_slab * sqrt(caux);
    lambda = aux1_local .* sqrt(aux2_local);
    lambda_tilde = sqrt( square(c) * square(lambda) ./
                (square(c) + square(tau) *
square(tau_vgroup[vgroup]) .* square(lambda)) );

    beta = beta_tilde .* lambda_tilde .* tau_vgroup[vgroup] * tau;
    mu = intercept + X * beta;
}
model {
    target += std_normal_lpdf(beta_tilde);
    target += inv_gamma_lpdf(caux | 0.5 * df_slab, 0.5 * df_slab);
    target += std_normal_lpdf(aux1_local);
    target += inv_gamma_lpdf(aux2_local | 0.5 * df_local, 0.5 *
df_local);
    target += std_normal_lpdf(aux1_vgroup);
    target += inv_gamma_lpdf(aux2_vgroup | 0.5 * df_vgroup, 0.5 *
df_vgroup);
    target += std_normal_lpdf(aux1_global);
    target += inv_gamma_lpdf(aux2_global | 0.5 * df_global, 0.5 *
df_global);
```

```
    target += normal_lpdf(intercept | 0, scale_intercept);
    target += bernoulli_logit_lpmf(group | mu);
}
generated quantities {
    vector[n] log_lik;

     for (i in 1:n) {
        log_lik[i] = bernoulli_logit_lpmf(group[i] | mu[i]);
  }
}
```

References

S1. Heinzel A, Kammer M, Mayer G, et al. Validation of Plasma Biomarker Candidates for the Prediction of eGFR Decline in Patients With Type 2 Diabetes. *Diabetes Care* 2018;41:1947-1954.

S2. Piironen J, Vehtari A. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics* 2017;11:5018-5051.

S3. Carvalho CM, Polson NG, Scott JG. Handling Sparsity via the Horseshoe. In *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics Hilton Clearwater Beach Resort*, Clearwater Beach, Florida USA, Proceedings of Machine Learning Research 2009, p. 73-80.

S4. Xu Z, Schmidt DF, Makalic E, et al. Bayesian Grouped Horseshoe Regression with Application to Additive Models. In *AI 2016: Advances in Artificial Intelligence* 2016, Springer International Publishing, p. 229-240.

S5. Ghosh J, Li Y, Mitra R. On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression. *Bayesian Analysis* 2018;13:359-383

S6. Goutis C, Robert CP. Model Choice in Generalised Linear Models: A Bayesian Approach Via Kullback-Leibler Projections. *Biometrika* 1998;85:29-37.

S7. Piironen J, Vehtari A. Comparison of Bayesian predictive methods for model selection. *Statistics and Computing* 2017;27:711-735.

S8. Carpenter B, Gelman A, Hoffman MD, et al. Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 2017;76:32.

S9. Surma MA, Herzog R, Vasilj A, et al. An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids. *Eur J Lipid Sci Technol* 2015;117:1540-1549.

S10. Matyash V, Liebisch G, Kurzchalia TV, et al. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J Lipid Res* 2008;49:1137-1146.

S11. Herzog R, Schwudke D, Schuhmann K, et al. A novel informatics concept for high-throughput shotgun lipidomics based on the molecular fragmentation query language. *Genome Biol* 2011;12:R8.