Metabolism 167 (2025) 156174

Contents lists available at ScienceDirect

Metabolism

journal homepage: www.journals.elsevier.com/metabolism

Differential risk assessment in persons at risk of type 2 diabetes using urinary peptidomics

Anja Schork [a,b,c,*], Andreas Fritsche [a,b,c], Erwin D. Schleicher [d], Andreas Peter [d], Martin Heni [a,b,c,e], Norbert Stefan [a,b,c], Reiner Jumpertz von Schwartzenberg [a,b,c], Martina Guthoff [a,b,c], Harald Mischak [f], Justyna Siwy [f], Andreas L. Birkenfeld [a,b,c], Robert Wagner [a,b,c,g,h]

[a] Department of Internal Medicine IV, Division of Endocrinology, Diabetology and Nephrology, University Hospital Tübingen, Germany
[b] Institute of Diabetes Research and Metabolic Diseases (IDM) of the Helmholtz Center Munich at the University of Tübingen, Otfried-Müller-Strasse 10, 72076 Tübingen, Germany
[c] German Center for Diabetes Research (DZD), Otfried-Müller-Strasse 10, 72076 Tübingen, Germany
[d] Institute for Clinical Chemistry and Pathobiochemistry, Department for Diagnostic Laboratory Medicine, University Hospital of Tübingen, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany
[e] Division of Endocrinology and Diabetology, Department of Internal Medicine I, University of Ulm, Albert-Einstein-Allee 23, 89081 Ulm, Germany
[f] Mosaiques Diagnostics GmbH, Hannover, Germany
[g] Department of Endocrinology and Diabetology, Medical Faculty and University Hospital, Heinrich Heine University Düsseldorf, Moorenstr. 5, 40225 Düsseldorf, Germany
[h] Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Auf'm Hennekamp 65, 40225 Düsseldorf, Germany

ARTICLE INFO

Keywords:
Prediabetes
Cluster analysis
Urine peptidome
CKD273
HF2
CAD238
Chronic kidney disease
Cardiovascular mortality
Risk stratification

ABSTRACT

Objective: Individuals at increased risk of type 2 diabetes have recently been classified into six prediabetes clusters, which stratify the risk of progression to diabetes and diabetes complications. Clusters 1, 2 and 4 are low-risk clusters while clusters 3, 5 and 6 are high-risk clusters; individuals in cluster 6 have an elevated risk of nephropathy and all-cause mortality despite delayed onset of diabetes. The urinary peptidome classifiers CKD273 (chronic kidney disease, CKD), HF2 (heart failure, HF) and CAD238 (coronary artery disease, CAD) are based on unique urinary peptide patterns and have shown potential for identifying individuals at risk for CKD and cardiovascular pathologies. This observational study investigates whether peptidome classifiers can differentiate complication risks across the prediabetes clusters and if a novel combination of peptides can distinguish high-risk from low-risk prediabetes clusters.
Methods: Urine peptidome analysis was performed on spot urine samples from individuals across 6 prediabetes clusters (n = 249) and 19 individuals with screen-detected diabetes (study cohorts at University Hospital Tübingen, Germany from 11/2004 to 11/2012). Predefined urinary classifiers were calculated for each participant. Lasso regression analysis was used to identify an optimal combination of peptides distinguishing low-Schlesinger et al. (2022), Wagner et al. (2021) [1,2,4] and high-risk (Rooney et al., 2021; Wagner, 2023; Latosinska et al., 2021 [3,5,6]) clusters.
Results: The predefined urinary peptidome classifiers CKD273, HF2 and CAD238 differed significantly across prediabetes clusters, particularly with elevated values in cluster 6 compared to the healthiest cluster 2. CKD273, HF2 and CAD238 were inversely associated with insulin sensitivity indexes. Machine Learning identified a combination of 112 urinary peptides that differentiated low-risk from high-risk prediabetes clusters (AUC-ROC 0.868 (95 % CI 0.755–0.981)).
Conclusions: Urinary peptidome classifiers support the increased risk of CKD and suggest an elevated risk of heart failure and coronary artery disease in the high-risk prediabetes cluster 6. Urine peptidomics show promising potential as a tool for identifying high-risk prediabetes individuals and guiding early preventive interventions.

* Corresponding author at: Department of Internal Medicine, Division of Diabetology, Endocrinology and Nephrology, University Hospital Tübingen, Otfried-Müller-Str.10, 72076 Tübingen, Germany.
E-mail address: anja.schork@med.uni-tuebingen.de (A. Schork).

https://doi.org/10.1016/j.metabol.2025.156174
Received 19 July 2024; Accepted 24 February 2025
Available online 27 February 2025

1. Introduction

Prediabetes is an intermediary phenotype of hyperglycemia, characterized by elevated blood glucose levels that do not meet the diagnostic criteria for diabetes. However, the condition is already associated with an elevated risk of diabetes-related complications and comorbidities [1]. Prediabetes prevalence is as high as 40 % of the adult population in some countries, such as the United States [2]. Nevertheless, preventive interventions are challenging, as many individuals with prediabetes are asymptomatic, and there is a large interindividual variance in the risk of both progression to diabetes and associated complications [3]. Additionally, examination methods for prediabetes screening lack scalability. The heterogeneity of metabolism before diabetes development led to the identification of six prediabetes clusters that differentiate risk of diabetes or diabetic complications [4] (Fig. 1): Cluster 1, 2 and 4 represent low-risk clusters with low risk of developing diabetes and diabetic complications. Healthy individuals are found in clusters 1 and 2, with cluster 2 people also being of normal weight. Cluster 4 contains overweight people with comparatively healthy metabolism. In contrast, clusters 3, 5 and 6 are high-risk clusters. Individuals in clusters 3 and 5 have an elevated risk of progression to diabetes, with low insulin secretion in cluster 3 and fatty liver and insulin resistance in cluster 5. Individuals in cluster 6 have an elevated risk of chronic kidney disease and all-cause mortality although they are not at immanent risk of diabetes [4]. The classification of individuals with prediabetes into clusters has the potential to open new avenues for personalized risk assessment and targeted interventions to prevent the development of diabetes and its complications [4,5].

The present observational study aimed to further characterize prediabetes clusters using a urine peptidome analysis. There are predefined urinary classifiers based on unique urinary peptides that have shown potential in identifying individuals at risk of various renal and cardiovascular pathologies [6]. CKD273 is a urinary peptide classifier developed for early detection of chronic kidney disease (CKD) [7,8]. CKD273 is based on 273 urinary peptides that were identified by comparing a training set of $n = 230$ patients with CKD and $n = 379$ healthy controls, and validated in a separate test set containing $n = 110$ CKD patients and $n = 34$ healthy controls receiving an AUC of 0.96 [7]. The urinary classifier HF2 is a marker of heart failure (HF) and combines 671 urinary peptides identified in a cohort of $n = 98$ patients with left ventricular dysfunction and n = 98 matched controls [9]. HF2 correlated with LV dysfunction, measured by echocardiography, in a separate test set (odds ratio 1.38 (1.00–1.90; $p = 0.052$)) [9]. The classifier CAD238 combines 238 urine peptides to differentiate patients with coronary artery disease (CAD, $n = 204$ in the training set) from healthy controls ($n = 382$) [10]. CAD238 was validated in a test set of $n = 71$ CAD patients and $n = 67$ healthy controls, reaching an AUC of 0.87 (95 % CI 0.81–0.92) [10] and was reevaluated in 60 urine samples (32 cases; 28 controls) for the prediction of cardiovascular events (Kaplan-Meier $p = 0.021$) [11].

The combination of prediabetes clusters and urinary peptide markers

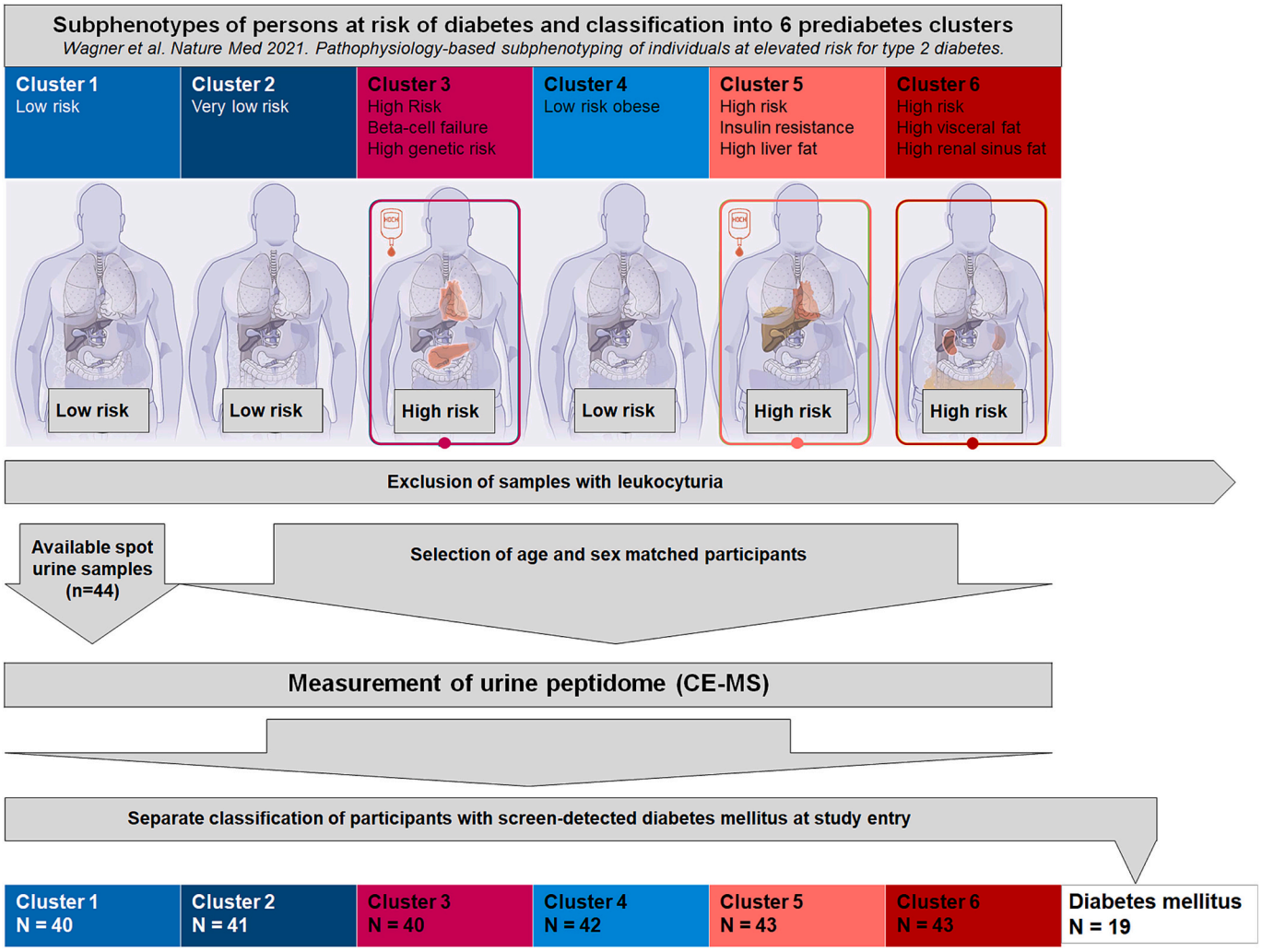


Fig. 1. Subphenotypes and classification into prediabetes clusters and flow chart of the study participants. Partly adapted from [36]. Abbreviations: CE-MS, capillary electrophoresis coupled mass spectroscopy.

presents a promising approach for early risk assessment in individuals with an increased risk of diabetes. Prediabetes clusters, characterized by diverse metabolic disturbances, also provide further opportunities to examine the utility of CKD273, HF2, and CAD238 as biomarkers for early risk identification.

2. Material and methods

2.1. Sample collection and study cohort

Individuals at increased risk of type 2 diabetes were metabolically phenotyped, and single spot urine samples were collected at the University Hospital Tübingen from 11/2004 to 11/2012 (TUEF/TULIP cohort, [12]). These cohorts included individuals with at least one of the following: a family history of type 2 diabetes, BMI > 27 kg/m^2, impaired fasting glucose (IGT, fasting blood glucose ≥7.78 mmol/l), or a previous diagnosis of gestational diabetes, but no diagnosed diabetes [12]. Phenotyping included assessment of insulin sensitivity by the Matsuda insulin sensitivity index (Matsuda-ISI) [13] and an insulin sensitivity index based on insulin and non-esterified fatty acids (NEFA-ISI) [14]. Cluster assignment was performed as described previously [4]. Urine samples with leukocyturia were excluded from the analysis. As cluster 1 comprised the lowest number of participants, participants from cluster 1 with available urine samples were selected ($n = 44$), and this number of patients was used as reference. A similar number of age and sex matched participants was selected from the other clusters (clusters 2, 3, 4, 5, and 6). The study cohort included also persons with screen-detected manifest diabetes at study entry, who were subsequently classified as participants with type 2 diabetes ($n = 19$, Fig. 1). Protein excretion was overall low with values below the threshold for microalbuminuria (<30 mg/g creatinine) in most of the cohort (Table 1).

2.2. Peptidome analysis

The urine peptides in spot urine samples were investigated using capillary electrophoresis coupled mass spectroscopy (CE-MS). The procedures for CE-MS analysis, encompassing sample preparation, measurement, peptide sequencing, data processing and calibration, have been extensively documented previously [15–17]. In summary, a P/ACEMDQCE system (Beckman Coulter, Fullerton, CA, USA) coupled with a Compact Q-TOF-MS (Bruker Daltonic, Bremen, Germany) was employed for CE-MS analysis. Peptide separation was achieved by reverse polarity at 25 kV. A running buffer comprising 20 % acetonitrile (Sigma-Aldrich, Taufkirchen, Germany) in high-performance liquid chromatography-grade water (Roth, Karlsruhe, Germany), supplemented with 0.94 % formic acid (Sigma-Aldrich), was utilized. The ESI sprayer (Agilent Technologies, Palo Alto, CA) was grounded, and the ion spray interface potential was set between −4 and − 4.5 kV. Spectra were recorded over an m/z range of 350 to 3000 and accumulated every 3 s. After the CE-MS analysis, mass spectral ion peaks representing identical molecules at different charge states were deconvoluted into single masses using MosaFinder software [18]. Reference signals of 29 abundant peptides served as internal standards for signal intensity calibration using linear regression [15]. This calibration procedure is highly reproducible and addresses both analytical and dilution variances in a single step. The resulting peak list characterized each polypeptide by its calibrated molecular mass (Da), calibrated capillary electrophoresis migration time (minutes), and normalized signal intensity. All detected peptides were deposited, matched, and annotated in a Microsoft SQL database, facilitating further statistical analysis. Detailed technical aspects of the analytical process have been elucidated in prior publications [17,19,20]. The following urinary classifiers as previously defined and validated combinations of urinary peptides for detection of chronic kidney disease and cardiovascular conditions were applied: CKD273 (chronic kidney disease), HF2 (heart failure) and CAD238 (coronary artery disease). The classification scores for each classifier were calculated using a support vector machine (SVM) algorithm, integrated into the MosaCluster software.

2.3. Statistical analysis

Mean and standard deviation (SD) are reported for continuous variables after test for normal distribution. Number and percentage are reported for categorical variables. ANOVA and Chi^2 test were used to test for differences between all groups of participants, and t-test with Bonferroni correction was used for pairwise comparisons. Correlations of continuous variables were analyzed with Pearson's product moment correlation and linear regression model. Outliers defined as values outside 1.5 * IQR of Q1 and Q3 were excluded for NEFA-ISI and Matsuda-ISI.

Machine learning was performed using lasso regression to identify a new combination of urinary peptides differentiating between high and low-risk clusters. Lasso (Least Absolute Shrinkage and Selection Operator) regression is a type of a linear regression model that uses a regularization with a penalty term called lambda to shrink the coefficients of less important features towards zero. For the lasso regression, participants were dichotomized into low-risk (cluster 1, 2 and 4) and high-risk (cluster 3, 5 and 6) clusters. A stratified random split into a 80 % training

Table 1
Characteristics of the study cohort.

	Total cohort	Cluster 1 + 2 + 4	Cluster 3	Cluster 5	Cluster 6	Diabetes mellitus	p-value
Cluster definition		Low-risk	Beta-cell failure	High-risk insulin-resistant fatty liver	High-risk visceral fat Nephropathy		
Number	268	123	40	43	43	19	
Age, years	46 ± 11	45 ± 11	48 ± 11	45 ± 10	45 ± 12	49 ± 11	n.s.
Gender, male	152 (61 %)	71 (66 %)	20 (50 %)	27 (63 %)	22 (54 %)	12 (67 %)	n.s.
BMI, kg/m^2	34 ± 9	29 ± 5	29 ± 3	41 ± 7	44 ± 9	37 ± 10	<0.001
HbA1c, %	5.5 ± 0.4	5.4 ± 0.3	5.6 ± 0.4	5.8 ± 0.4	5.6 ± 0.4	6.5 ± 1.0	<0.001
eGFR (CKD-EPI), ml/min/1.73m^2	95 ± 16	94 ± 15	88 ± 18	103 ± 10	98 ± 15	91 ± 18	<0.001
UACR, mg/g Crea	12.65 ± 14.5	9.31 ± 12.6	9.82 ± 8.0	11.82 ± 9.1	23.74 ± 27.59	16.96 ± 21.78	0.001
medication with RAAS inhibitor	63 (23.5 %)	25 (20.3 %)	7 (17.5 %)	13 (30.2 %)	12 (27.9 %)	6 (31.6 %)	0.142
Matsuda-ISI	11.7 ± 6.7	16.0 ± 5.8	12.2 ± 4.7	5.1 ± 2.6	6.8 ± 3.0	7.5 ± 5.1	<0.001
NEFA-ISI	3.2 ± 1.6	4.2 ± 1.3	3.5 ± 1.0	1.5 ± 0.6	1.7 ± 0.5	2.5 ± 1.7	<0.001
CKD273	−0.72 ± 0.34	−0.78 ± 0.32	−0.62 ± 0.32	−0.74 ± 0.36	−0.66 ± 0.39	−0.66 ± 0.37	0.034
HF2	−0.85 ± 0.45	−0.93 ± 0.41	−0.84 ± 0.51	−0.85 ± 0.39	−0.62 ± 0.49	−0.85 ± 0.57	0.005
CAD238	−0.37 ± 0.30	−0.41 ± 0.26	−0.45 ± 0.30	−0.29 ± 0.26	−0.26 ± 0.39	−0.44 ± 0.28	0.002

Abbreviations: UACR, urinary albumin to creatinine ratio; RAAS, Renin angiotensin aldosterone system; CKD, chronic kidney disease; HF, heart failure; CAD, coronary artery disease; Matsuda-ISI, Matsuda insulin sensitivity index; NEFA-ISI, insulin sensitivity index based on insulin and non-esterified fatty acids.
Values are mean and SD or number and percentage. P values are from ANOVA or chi^2 test.

and a 20 % test set was performed. Lasso regression was performed and the optimal lambda with the highest accuracy was selected for the final model. Performance of the model was tested with Receiver operating characteristic (ROC, with reporting area under the curve, AUC with 95 % confidence interval, CI) and a confusion matrix to calculate precision score, recall score, F1 score, sensitivity and specificity. To analyze the number of individuals needed to be screened to identify one true positive case, the Screening Efficiency Ratio (SER) was calculated as follows: true positive [TP] + false positive [FP] + true negative [TN] + false negative [FN]) / TP; reversely, the False Negative Burden Ratio (FNBR) was calculated as (TP + FP + TN + FN) / FN to define the number of cases screened while missing one case, as suggested by Kokkorakis et al. [21]. A complementary statistical approach of variable selection was additionally performed based on the work of Guo et al. [22] as follows: Variable importance scores were obtained using the varimp function of the caret package [23] and features were ordered by importance. Starting with the most important feature, logistic regression models were trained on the training set and tested on the test set, and iteratively

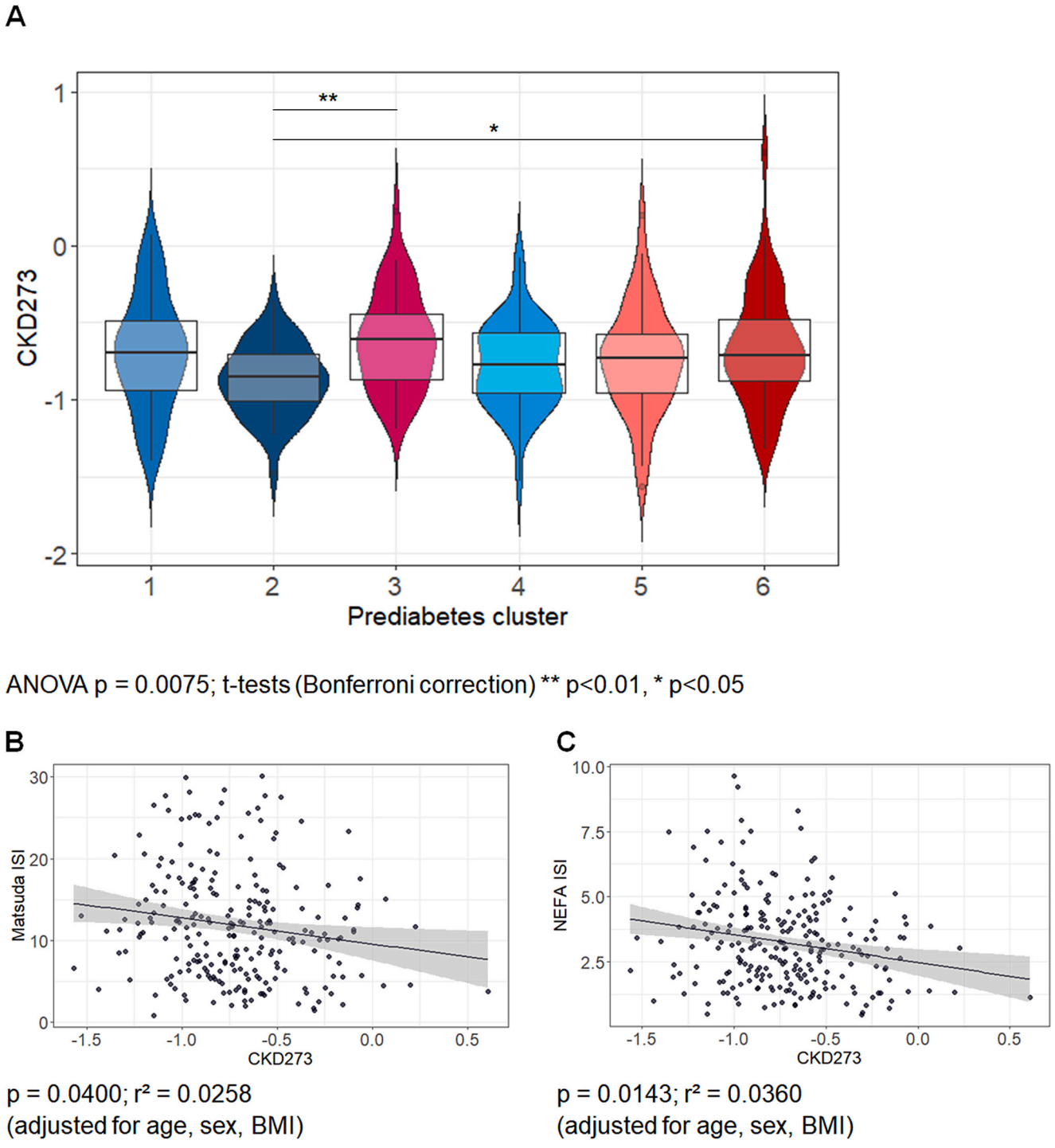


Fig. 2. Urinary peptidome classifier CKD273 in prediabetes clusters (A) and correlation with insulin sensitivity indices in prediabetes persons (B, Matsuda-ISI and C, NEFA-ISI).

Abbreviations: CKD, chronic kidney disease; Matsuda-ISI, Matsuda insulin sensitivity index; NEFA-ISI, insulin sensitivity index based on insulin and non-esterified fatty acids.

extended by the next most important feature from the ordered list. The AUC-ROC of each new model was compared to the previous one using DeLong's test to determine if the improvement was statistically significant. All analyses were performed using the statistical software R version 4.1.2. Statistical significance was defined as a significance threshold of $p < 0.05$.

3. Results

3.1. Characteristics of study cohort and urinary classifiers

Mean age of the entire cohort was 46 ± 11 years and 61 % of the participants were males, with no significant difference across the prediabetes clusters, as participants were age- and sex-matched (Table 1). BMI, HbA1c and insulin sensitivity differed across groups, as expected from the cluster definitions (Table 1).

To better interpret the investigation of urinary peptide-based classifiers in the present study cohort, the overlap between the classifiers was examined by comparing the peptides included in the classifiers and assessing whether the change (up- or down-regulation) occurred in the same direction. The predefined classifiers CKD273, CAD238, and HF2 include 273, 238, and 671 peptides, respectively. Across all three classifiers, an overlap of 17 peptides was observed. CKD273 and CAD238 share 26 common peptides, CAD238 and HF2 share 102 peptides, and CKD273 and HF2 share 100 peptides (Suppl. Table 1).

3.2. CKD273 in prediabetes clusters

The urinary peptide-based classifier CKD273 was significantly different across prediabetes clusters ($p = 0.023$) and marginally significant in the total cohort, which included individuals with diabetes ($p = 0.058$, Table 1). Higher values of CKD273 were found in clusters 3 and 6 and in persons with diabetes (Table 1). CKD273 was significantly higher in cluster 3 and had a trend of being significantly higher in cluster 6, all compared to the healthiest cluster 2 ($p = 0.018$ and $p = 0.078$, Fig. 2 A). CKD273 inversely associated with Matsuda-ISI ($p = 0.0400$; $r^2 = 0.0258$)

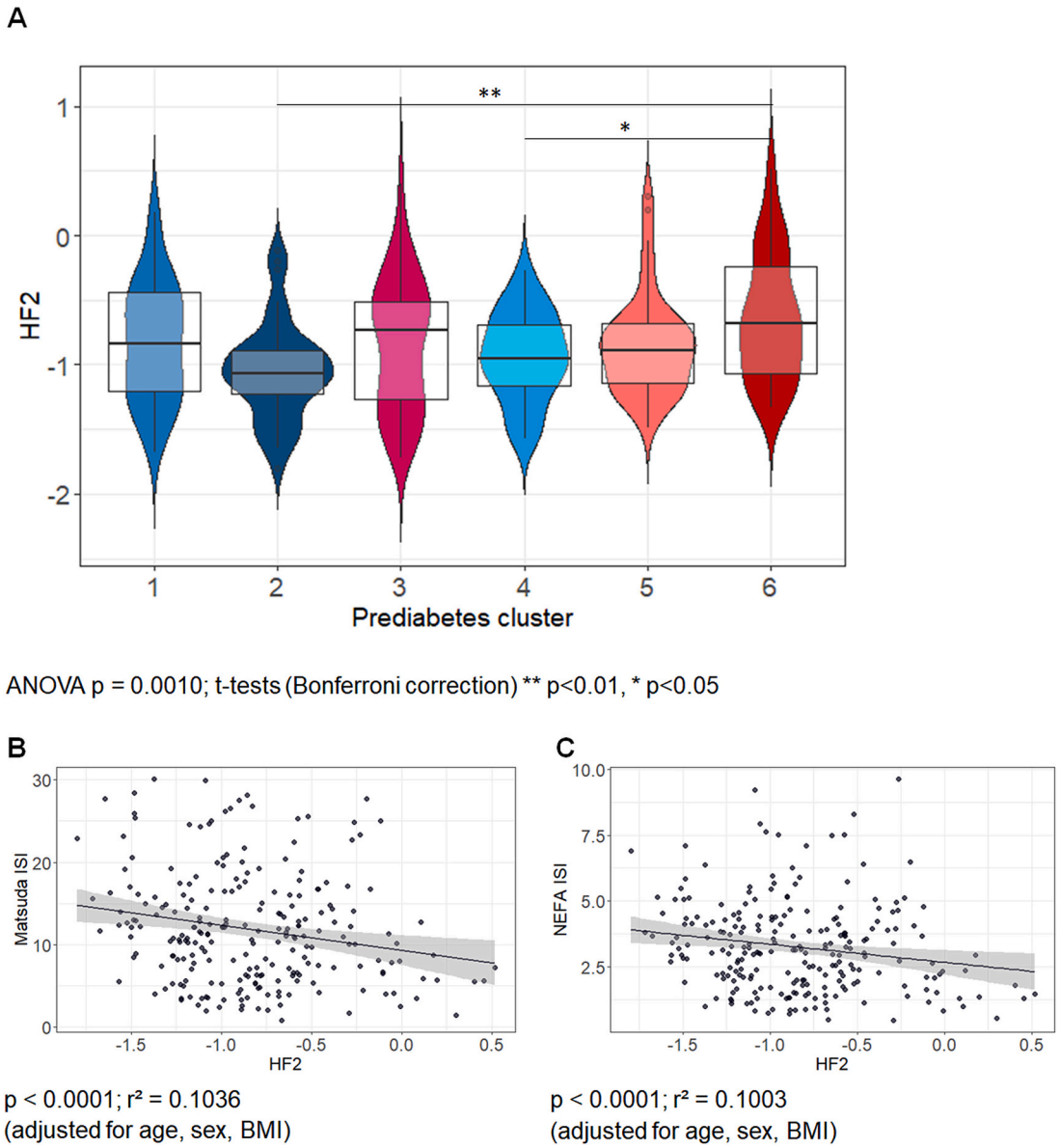


Fig. 3. Urinary peptidome classifier HF2 in prediabetes clusters (A) and correlation with insulin sensitivity indices in prediabetes persons (B, Matsuda-ISI and C, NEFA-ISI).
Abbreviations: HF, heart failure; Matsuda-ISI, Matsuda insulin sensitivity index; NEFA-ISI, insulin sensitivity index based on insulin and non-esterified fatty acids.

and NEFA-ISI ($p = 0.0143$; $r^2 = 0.0360$; Fig. 2 B and C; adjusted for age, sex, BMI).

3.3. HF2 in prediabetes clusters

The urinary peptide-based classifier HF2 showed significant differences in participants from different prediabetes clusters and with diabetes (ANOVA $p = 0.008$, Table 1). HF2 was highest in cluster 6 with significant differences to clusters 2 and 4 ($p = 0.001$ and $p = 0.026$, Fig. 3 A). HF2 inversely associated with Matsuda-ISI ($p < 0.0001$; $r^2 = 0.1036$) and NEFA-ISI ($p < 0.0001$; $r^2 = 0.1003$; Fig. 3 B and C; adjusted for age, sex, BMI).

3.4. CAD238 in prediabetes clusters

The urinary peptide-based classifier CAD238 showed significant differences in participants of different prediabetes clusters and with diabetes (ANOVA $p = 0.007$, Table 1). CAD238 was highest in clusters 5 and 6 (Table 1) with highest values in cluster 6 and significant or trending significant differences between cluster 6 vs. cluster 2 and 3, respectively ($p = 0.039$ and $p = 0.081$, Fig. 4 A). CAD238 inversely associated with Matsuda ISI ($p < 0.0001$; $r^2 = 0.0999$) and NEFA-ISI (p

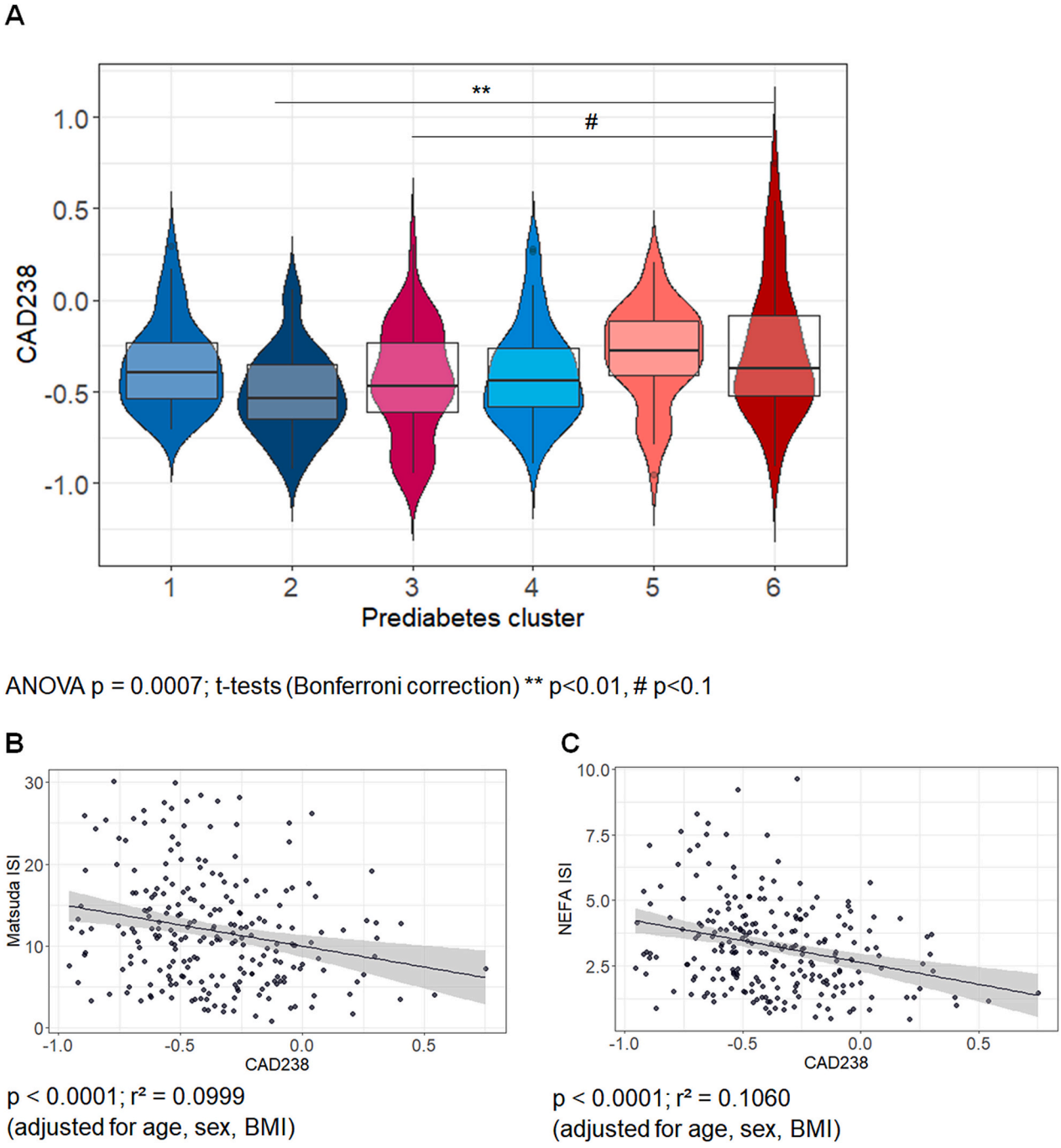


Fig. 4. Urinary peptidome classifier CAD238 in prediabetes clusters (A) and correlation with insulin sensitivity indices in prediabetes persons (B, Matsuda-ISI and C, NEFA-ISI).

Abbreviations: CAD, coronary artery disease; Matsuda-ISI, Matsuda insulin sensitivity index; NEFA-ISI, insulin sensitivity index based on insulin and non-esterified fatty acids.

< 0.0001; $r^2 = 0.1060$; Fig. 4 B and C; adjusted for age, sex, BMI).

3.5. Lasso regression of urinary peptides differentiating low-risk vs. high-risk clusters

A total of 1621 sequenced urinary peptides were detected by CE-MS with a frequency threshold of 30 %. To identify a combination of urinary peptides differentiating high-risk from low-risk clusters, machine learning analysis was performed with lasso regression, as described in the methods section, taking accuracy as selection criterion for the optimal model. The optimal combination comprised 112 urinary peptides (Fig. 5 A, Suppl. Table 2). Th feature contribution of the lasso model peptides is shown in Fig. 2 B. The model predicted low-risk vs. high-risk clusters in the test set with an area under the receiver operating characteristic curve (AUC-ROC) of 0.868 (95 % CI 0.755–0.981; Fig. 5 C), a sensitivity of 72 % and a specificity of 83 % (Fig. 5 D). The Screening Efficiency Ratio (SER), defined as true positive [TP] + false positive [FP] + true negative [TN] + false negative [FN]) / TP was 18 + 4 + 20 + 7 / 18 = 2,72. The False Negative Burden Ratio (FNBR), defined as TP + FP + TN + FN / FN was 18 + 4 + 20 + 7 / 7 = 7. Comparison of the lasso regression model with the predefined urinary peptide-based classifiers revealed only one common peptides with CKD273, and no overlap with CAD283 and HF2 (Suppl. Table 2). As an alternative approach to identify an optimum set of proteins explaining the difference between the low risk and high-risk clusters, an iterative model fitting approach was used, as described in the methods section (Suppl. Fig. 1). An optimum model was not identified due to the lack of significant differences between iterations. An AUC-ROC of 0.775 was reached after 5 steps. With $n = 74$ peptide variables, the AUC-ROC of the model reached 0.853, approaching the AUC of the lasso model. The $n = 5$ peptides with the highest importance are printed bold in Suppl. Table 2.

4. Discussion

The prediabetes clusters delineate groups of individuals with similar metabolic status before the onset of diabetes and stratify populations for the risk of developing diabetes and complications of diabetes, such as CKD [4,24]. Prediabetes clusters 3 and 5 are associated with an elevated risk of developing diabetes [4]. Individuals in cluster 6 have an elevated risk of CKD and all-cause mortality, but only a moderate risk of developing overt diabetes mellitus [4]. Consistent with this, a slightly increased urinary albumin excretion was already detectable in cluster 6. In the current urinary peptidome analysis, an association between the investigated urinary peptidome classifiers and insulin sensitivity was observed. Insulin sensitivity plays a major role in renal hemodynamics, as well as podocyte viability and tubular function, such that insulin resistance has long been discussed as a driver of diabetic kidney disease [25]. The CKD-related classifier CKD273 in clusters 3 and 6 was as high as in individuals with diabetes mellitus (Table 1). This aligns with the increased risk of kidney disease in cluster 6 and suggests that the biomarker CKD273 captures the elevated CKD risk in cluster 6. Indeed, in patients with diabetes, CKD273 has been shown to be associated with an increased risk of development of microalbuminuria over a median follow-up of 2.5 years [8]. Notably, urine peptidome analysis requires only one urine sample, and does not need a functional test, such as an OGTT, for cluster assignment. CKD273 may therefore be a readily implementable tool to detect prediabetes persons at risk for CKD.

The cause of increased mortality in cluster 6 has so far been elusive [4]. In the current analysis, the cardiovascular disease-related classifiers HF2 and CAD238 were elevated in prediabetes cluster 6. These results indicate an increased risk of heart disease, such as CAD and HF, in persons of prediabetes cluster 6, which may underlie the previously unexplained increased mortality. This is further supported by recent work showing higher mortality in high-risk prediabetes clusters in a cohort undergoing coronary angiography [24].

The current study highlights the potential of urinary peptidome analysis in identifying persons of high-risk prediabetes clusters and facilitating early risk stratification before the onset of diabetes. Lifestyle intervention is the recommended approach for managing prediabetes, according to current guidelines. However, with prediabetes affecting over 25 % of the population in the US and Europe, and the limited scalability of lifestyle interventions, effective implementation poses significant challenges [2]. Stratifying prediabetes into biologically similar groups facilitates identification of subpopulations that respond differently to lifestyle interventions [26,27]. Regarding pharmacological approaches, SGLT2 inhibitor treatment did not improve whole-body insulin sensitivity or insulin secretion in an unstratified prediabetes population [28,29]. Identifying individuals with high-risk prediabetes probably improves selection of those who would most benefit from a medical treatment, including time-intensive lifestyle intervention or pharmacological approaches [30].

In the current study, a set of urinary peptides was able to differentiate high- and low-risk prediabetes clusters. The urinary classifier CKD273 has demonstrated effectiveness in identifying high-risk patients with type 2 diabetes [31] and monitoring therapeutic success in patients with diabetes [8,32]. Urinary peptide classifiers have also been able to predict in silico treatment efficacy in cardiovascular disease and CKD [33]. The combined evidence suggests that urinary peptidome measurements are a promising tool for monitoring the success of intervention in prediabetes or specific clusters of prediabetes. Prediabetes with fluctuating hyperglycemia is also common and needs to be stratified in vulnerable cohorts with potential altered physiology, such as kidney or liver transplant recipients [34,35]. Urinary peptidome analysis, requiring a single urine sample and potentially repeated measurement over time, has the potential to improve the selection of individuals who would benefit most from the time- and cost-intensive interventions.

The present study is limited by the available sample size and retrospective single-measurement nature of the data. Another limitation of the study is the lack of external validation of the peptide combination distinguishing low- and high-risk prediabetes. However, it is the first study to investigate use of urinary peptidome analysis for risk stratification in prediabetes.

5. Conclusion

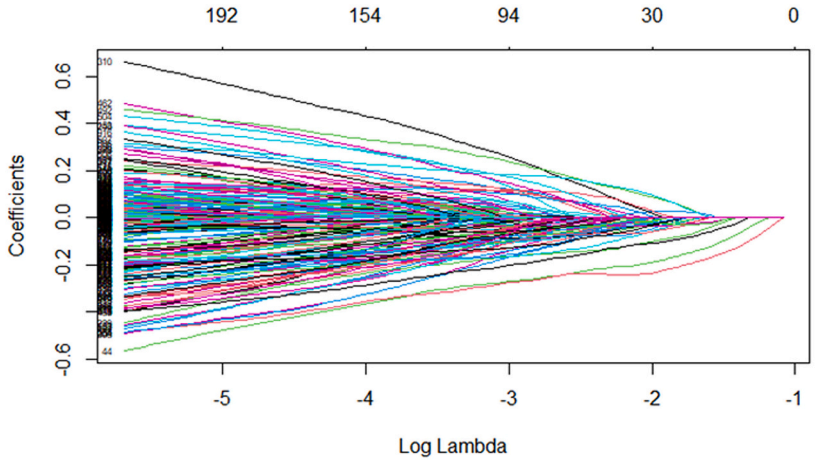
In conclusion, urinary peptidome classifiers corroborate the increased risk of CKD and suggest an elevated risk of heart failure and coronary artery disease in individuals from prediabetes cluster 6, providing an explanation for the increased mortality independent of diabetes development in this high-risk prediabetes cluster. Urinary peptidome analysis and investigation of urinary peptide-based classifiers facilitate the selection of individuals with prediabetes who are at high risk for complications and are most likely to benefit from therapeutic intervention.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.metabol.2025.156174.

CRediT authorship contribution statement

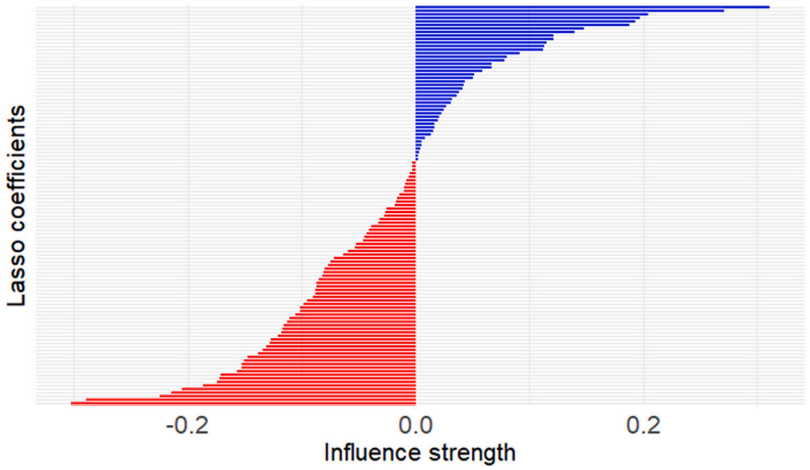
Anja Schork: Writing – original draft, Visualization, Formal analysis. **Andreas Fritsche:** Writing – review & editing, Supervision, Conceptualization. **Erwin D. Schleicher:** Writing – review & editing, Validation, Conceptualization. **Andreas Peter:** Writing – review & editing, Resources, Investigation. **Martin Heni:** Writing – review & editing, Validation. **Norbert Stefan:** Writing – review & editing, Validation. **Reiner Jumpertz von Schwartzenberg:** Writing – review & editing, Validation. **Martina Guthoff:** Writing – review & editing, Validation, Supervision, Investigation. **Harald Mischak:** Writing – review & editing, Validation, Software, Resources, Investigation, Conceptualization. **Justyna Siwy:** Writing – review & editing, Validation, Software, Investigation. **Andreas L. Birkenfeld:** Writing – review

A

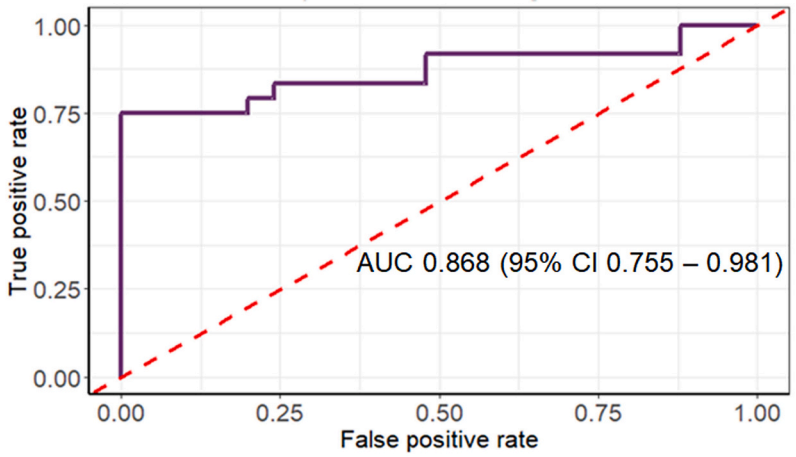


The final values used for the model were alpha = 0.55 and lambda = 0.0372.
Number of non-zero coefficients: 113 (112 peptides + intercept).

B



C



D

	Predicted	
Actual	high-risk	low-risk
high-risk	18	7
low-risk	4	20

Sensitivity: 0.720. Specificity: 0.833.
Precision: 0.741. Recall: 0.833.
F1-Score: 0.784.

(caption on next page)

Fig. 5. Lasso regression of urinary peptides differentiating high-risk vs. low-risk prediabetes clusters: lasso coefficient paths (A), feature contribution fig. (B), ROC curve (C) and confusion matrix (D) of the lasso model.
In the plot of the lasso coefficient paths (A), each colored line represents a different peptide and its coefficient value across varying levels of regularization (log (lambda)). Numbers above the panel show the number of non-zero coefficients, below the panel the natural logarithm of lambda values. The vertical red line indicates the optimal value of lambda chosen for the lasso model. A feature contribution fig. (B) was plotted to visualize influence strength of the coefficients of the lasso model (positive coefficients are plotted blue, negative coefficients red). To test the quality of the lasso model, a receiver operating characteristic (ROC) curve (C) and confusion matrix (D) were created. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

& editing, Validation, Supervision, Resources. **Robert Wagner:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Author contribution statement

Detailed author contributions are as follows: AS analyzed and interpreted the data and drafted the manuscript. RW and AF planned the study, selected the phenotyped urine samples, helped analyzing and interpreting the data and revised the manuscript. HM and JS performed the peptidome analysis and computed predefined urinary peptide classifiers at Mosaiques diagnostics and contributed to interpretation of results. EDS, AP, MH, NS, RJvS and AL discussed interpretation of results and revised the manuscript. All authors approved the final version of the manuscript submitted.

Statement of Ethics

The study was approved by the Ethics Committee of the University of Tübingen (422/2002). All participants gave written informed consent.

Funding

The study was supported by the German Center for Diabetes Research (Deutsches Zentrum für Diabetesforschung, DZD). The DZD is funded by the German Federal Ministry for Education and Research and the states where its partner institutions are located (01GI0925). Urinary proteomics measurements were supported by a grant from the German Diabetes Association (Deutsche Diabetesgesellschaft, DDG). The funders had no role in the design, data collection, data analysis, and reporting of this study. AS was supported by the TÜFF program of the Medical Faculty of Tübingen University (2729-0-0).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: HM is the founder and co-owner of Mosaiques Diagnostics (Hannover, Germany). JS is employed by Mosaiques Diagnostics. MH reports research grants from Boehringer Ingelheim and Sanofi to the University Hospital of Tübingen, participation in advisory board for Boehringer Ingelheim, Sanofi, and Amryt, and lecture fees from Amryt, AstraZeneca, Bayer, Boehringer Ingelheim, Eli Lilly, Novartis, Novo Nordisk, and Sanofi. RW reports lecture fees from Novo Nordisk, Sanofi-Aventis, Boehringer-Ingelheim and Eli Lilly and served on the advisory board for Akcea Therapeutics, Daiichi Sankyo, Sanofi-Aventis, Eli Lilly and NovoNordisk.

Acknowledgments

We thank our participants, whose dedication made this study possible. The authors are thankful for the dedication and excellent assistance of I. Wagener, A. Eberle, E. Stehle, A. Vosseler, C. Dannecker and D. Neuscheler (all University of Tübingen). The authors acknowledge the authors of the following packages for the free statistical software R, which were used in the data analysis: tidyverse, caret, glmnet, pROC, openxlsx, readxl, psych, corrplot, ggpubr, viridis, htmlTable. We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Data availability

All requests for data and materials will be reviewed by the Data Access Steering Committee of the Institute of Diabetes and Metabolic Research, Tübingen, to verify whether the request is subject to any intellectual property or confidentiality obligations. Individual-level data may be subject to confidentiality. Any data and materials that can be shared will be released via a Material Transfer Agreement.

References

[1] Schlesinger S, Neuenschwander M, Barbaresko J, Lang A, Maalmi H, Rathmann W, et al. Prediabetes and risk of mortality, diabetes-related complications and comorbidities: umbrella review of meta-analyses of prospective studies. Diabetologia 2022;65(2):275–85.
[2] Centers for Disease Control and Prevention. National Diabetes Statistics Report website. https://www.cdc.gov/diabetes/data/statistics-report/index.html. Accessed 2024, January 16. [.
[3] Rooney MR, Rawlings AM, Pankow JS, Echouffo Tcheugui JB, Coresh J, Sharrett AR, et al. Risk of progression to diabetes among older adults with prediabetes. JAMA Intern Med 2021;181(4):511–9.
[4] Wagner R, Heni M, Tabák AG, Machann J, Schick F, Randrianarisoa E, et al. Pathophysiology-based subphenotyping of individuals at elevated risk for type 2 diabetes. Nat Med 2021;27(1):49–57.
[5] Wagner R. Deconvoluting prediabetes: a path to understanding the origins of complications. Cell Rep Med 2023;4(3):100986.
[6] Latosinska A, Siwy J, Faguer S, Beige J, Mischak H, Schanstra JP. Value of urine peptides in assessing kidney and cardiovascular disease. Proteomics Clin Appl 2021;15(1):e2000027.
[7] Good DM, Zürbig P, Argilés A, Bauer HW, Behrens G, Coon JJ, et al. Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. Mol Cell Proteomics 2010;9(11):2424–37.
[8] Tofte N, Lindhardt M, Adamova K, Bakker SJL, Beige J, Beulens JWJ, et al. Early detection of diabetic kidney disease by urinary proteomics and subsequent intervention with spironolactone to delay progression (PRIORITY): a prospective observational study and embedded randomised placebo-controlled trial. Lancet Diabetes Endocrinol 2020;8(4):301–12.
[9] Zhang Z, Staessen JA, Thijs L, Gu Y, Liu Y, Jacobs L, et al. Left ventricular diastolic function in relation to the urinary proteome: a proof-of-concept study in a general population. Int J Cardiol 2014;176(1):158–65.
[10] Delles C, Schiffer E, von Zur Muhlen C, Peter K, Rossing P, Parving HH, et al. Urinary proteomic diagnosis of coronary artery disease: identification and clinical validation in 623 individuals. J Hypertens 2010;28(11):2316–22.
[11] Brown CE, McCarthy NS, Hughes AD, Sever P, Stalmach A, Mullen W, et al. Urinary proteomic biomarkers to predict cardiovascular events. Proteomics Clin Appl 2015;9(5–6):610–7.
[12] Schmid V, Wagner R, Sailer C, Fritsche L, Kantartzis K, Peter A, et al. Non-alcoholic fatty liver disease and impaired proinsulin conversion as newly identified predictors of the long-term non-response to a lifestyle intervention for diabetes prevention: results from the TULIP study. Diabetologia 2017;60(12):2341–51.
[13] Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22(9):1462–70.
[14] Wagner R, Fritsche L, Heni M, Fehlert E, Stefan N, Staiger H, et al. A novel insulin sensitivity index particularly suitable to measure insulin sensitivity during gestation. Acta Diabetol 2016;53(6):1037–44.
[15] Jantos-Siwy J, Schiffer E, Brand K, Schumann G, Rossing K, Delles C, et al. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. J Proteome Res 2009;8(1):268–81.
[16] Klein J, Papadopoulos T, Mischak H, Mullen W. Comparison of CE-MS/MS and LC-MS/MS sequencing demonstrates significant complementarity in natural peptide identification in human urine. Electrophoresis 2014 Apr;35(7):1060–4. https://doi.org/10.1002/elps.201300327. Epub 2013 Nov 20. PMID: 24254231.
[17] Mischak H, Vlahou A, Ioannidis JP. Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience. Clin Biochem 2013;46(6):432–43.

[18] Latosinska A, Siwy J, Mischak H, Frantzi M. Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: the past, the present, and the future. Electrophoresis 2019;40(18–19):2294–308.
[19] Mavrogeorgis E, Mischak H, Latosinska A, Siwy J, Jankowski V, Jankowski J. Reproducibility evaluation of urinary peptide detection using CE-MS. Molecules 2021;26(23).
[20] Mischak H, Kolch W, Aivaliotis M, Bouyssié D, Court M, Dihazi H, et al. Comprehensive human urine standards for comparability and standardization in clinical proteome analysis. Proteomics Clin Appl 2010;4(4):464–78.
[21] Kokkorakis M, Folkertsma P, Forte JC, Wolffenbuttel BHR, van Dam S, Mantzoros CS. GDF-15 improves the predictive capacity of Steatotic liver disease non-invasive tests for incident morbidity and mortality risk for cardio-renal-metabolic diseases and malignancies. Metabolism 2024;156047.
[22] Guo Y, You J, Zhang Y, Liu WS, Huang YY, Zhang YR, et al. Plasma proteomic profiles predict future dementia in healthy adults. Nat Aging 2024;4(2):247–60.
[23] Kuhn M. caret: Classification and Regression Training. Astrophysics Source Code Library. 2015:ascl:1505.003.
[24] Prystupa K, Delgado GE, Moissl AP, Kleber ME, Birkenfeld AL, Heni M, et al. Clusters of prediabetes and type 2 diabetes stratify all-cause mortality in a cohort of participants undergoing invasive coronary diagnostics. Cardiovasc Diabetol 2023;22(1):211.
[25] Artunc F, Schleicher E, Weigert C, Fritsche A, Stefan N, Häring HU. The impact of insulin resistance on the kidney and vasculature. Nat Rev Nephrol 2016;12(12):721–37.
[26] Sandforth A, von Schwartzenberg RJ, Arreola EV, Hanson RL, Sancar G, Katzenstein S, et al. Mechanisms of weight loss-induced remission in people with prediabetes: a post-hoc analysis of the randomised, controlled, multicentre prediabetes lifestyle intervention study (PLIS). Lancet Diabetes Endocrinol 2023;11(11):798–810.
[27] Wagner R, Heni M, Kantartzis K, Sandforth A, Machann J, Schick F, et al. Lower hepatic fat is associated with improved insulin secretion in a high-risk prediabetes subphenotype during lifestyle intervention. Diabetes 2023;72(3):362–6.
[28] Hummel J, Machann J, Dannecker C, Kullmann S, Birkenfeld AL, Häring HU, et al. Eight weeks of empagliflozin does not affect pancreatic fat content and insulin secretion in people with prediabetes. Diabetes Obes Metab 2022;24(8):1661–6.
[29] Kullmann S, Hummel J, Wagner R, Dannecker C, Vosseler A, Fritsche L, et al. Empagliflozin improves insulin sensitivity of the hypothalamus in humans with prediabetes: a randomized, double-blind, placebo-controlled, phase 2 trial. Diabetes Care 2022;45(2):398–406.
[30] Jumpertz von Schwartzenberg R, Stefan N, Wagner R, Guthoff M, Sandforth A, Icks A, et al. SGLT2 inhibition in addition to lifestyle intervention and risk for complications in subtypes of patients with prediabetes - a randomized, placebo controlled, multi-center trial (LIFETIME) – rationale, methodology and design. medRxiv. 2023:2023.11.18.23298622.
[31] Critselis E, Vlahou A, Stel VS, Morton RL. Cost-effectiveness of screening type 2 diabetes patients for chronic kidney disease progression with the CKD273 urinary peptide classifier as compared to urinary albumin excretion. Nephrol Dial Transplant 2018;33(3):441–9.
[32] Curovic VR, Eickhoff MK, Rönkkö T, Frimodt-Møller M, Hansen TW, Mischak H, et al. Dapagliflozin improves the urinary proteomic kidney-risk classifier CKD273 in type 2 diabetes with albuminuria: a randomized clinical trial. Diabetes Care 2022;45(11):2662–8.
[33] Jaimes Campos MA, Andújar I, Keller F, Mayer G, Rossing P, Staessen JA, et al. Prognosis and personalized in silico prediction of treatment efficacy in cardiovascular and chronic kidney disease: a proof-of-concept study. Pharmaceuticals (Basel) 2023;16(9).
[34] Guthoff M, Wagner R, Weichbrodt K, Nadalin S, Königsrainer A, Häring HU, et al. Dynamics of glucose metabolism after kidney transplantation. Kidney Blood Press Res 2017;42(3):598–607.
[35] Grottenthaler JM, Füger JK, Mahling M, Königsrainer A, Malek NP, Birkenfeld AL, et al. Dynamics of glucose metabolism after liver transplantation: prediabetes as a window of opportunity for patient survival and long-term kidney function. Transpl Int 2021;34(10):1959–70.
[36] https://www.dzd-ev.de/en/press/press-releases/press-releases-archive/press-releases-2021/prediabetes-subtypes-identified/index.html [Accessed 2024, November 5] [.