Plasma metabolomics identifies markers of impaired renal function: A meta-analysis of 3,089 persons with type 2 diabetes

Nete Tofte[1], Nicole Vogelzangs[2,3], Dennis Mook-Kanamori[4], Adela Brahimaj[5,6], Jana Nano[5,7,8], Fariba Ahmadizar[5], Ko Willems van Dijk[9], Marie Frimodt-Møller[1], Ilja Arts[2,3], Joline W.J. Beulens[10], Femke Rutters[10], Amber A. van der Heijden[11], Maryam Kavousi[5], Coen D.A. Stehouwer[3,12], Giel Nijpels[11], Marleen M.J. van Greevenbroek[3,12], Carla J.H. van der Kallen[3,12], Peter Rossing[1,13], Tarunveer S. Ahluwalia[1*], Leen M. 't Hart[10,14*]

[*]*equal author contributions*

[1]Steno Diabetes Center Copenhagen, [2]Department of Epidemiology & Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, [3]Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, [4]Departments of Clinical Epidemiology and Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands, [5]Department of Epidemiology, Erasmus University Medical Center, [6]Department of General Practice, Erasmus MC, [7]Institute of Epidemiology, Helmholtz Zentrum Munich, [8]German Center for Diabetes Research Munich, [9]Departments of Human Genetics and Internal Medicine/Endocrinology, Leiden University Medical Center, Leiden, the Netherlands, [10]Department of Epidemiology and Biostatistics, Amsterdam University Medical Center, Location VUmc, [11]Department of General Practice and Elderly Care Medicine, Amsterdam University Medical Center, Location VUmc, [12]Department of Internal Medicine, Maastricht University Medical

jc.2019-40627. See endocrine.org/publications for Accepted Manuscript disclaimer and additional information.

Centre, [13]University of Copenhagen, [14]Department of Cell and Chemical Biology & Department of Biomedical Data Sciences, section Molecular Epidemiology, Leiden University Medical Center.

Keywords: Metabolomics, NMR, renal function, albuminuria, meta-analysis

Corresponding authors and requests for reprint:

Nete Tofte, e-mail: nete.tofte@regionh.dk and Tarunveer Singh Ahluwalia, e-mail: tarun.veer.singh.ahluwalia@regionh.dk, Steno Diabetes Center Copenhagen, Niels Steensens Vej 2, 2820, Gentofte, Denmark

Grant support:

This work was performed within the framework of the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI) Metabolomics Consortium funded by BBMRINL, a research infrastructure financed by the Dutch government (NWO, grant no 184.021.007 and 184033111). It was furthermore funded by ZonMW Priority Medicines Elderly (grant 113102006) and the Parelsnoer Initiative (PSI). PSI is part of and is funded by the Dutch Federation of University Medical Centres and, from 2007 to 2011 received initial funding from the Dutch Government. CODAM was supported by grants from the Netherlands Organization for Scientific Research (940–35–034), the Dutch Diabetes Research Foundation (98.901), the Parelsnoer Initiative (PSI). The work of NV was supported through a grant from the Maastricht University Medical Center+. DM-K is supported by the Dutch Science Organization (ZonMW VENI Grant 916.14.023). The Netherlands Cardiovascular Research Initiative funded the metabolomics measurements in the NEO study: an initiative with support of the Dutch Heart Foundation (CVON2014-02 ENERGISE). The European Regional Development Fund supported the Maastricht

Study via OP-Zuid, the Province of Limburg, the Dutch Ministry of Economic Affairs (grant 31O.041), Stichting De Weijerhorst (Maastricht, the Netherlands), the Pearl String Initiative Diabetes (Amsterdam, the Netherlands), CARIM School for Cardiovascular Diseases (Maastricht, the Netherlands), Stichting Annadal (Maastricht, the Netherlands), Health Foundation Limburg (Maastricht, the Netherlands) and by unrestricted grants from Janssen-Cilag B.V. (Tilburg, the Netherlands), Novo Nordisk Farma B.V. (Alphen aan den Rijn, the Netherlands) and Sanofi-Aventis Netherlands B.V. (Gouda, the Netherlands). TSA is supported by the Novo Nordisk Foundation Grant NNF18OC0052457 and by Steno Diabetes Center Copenhagen, Denmark. The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Disclosure summary

Dennis Mook-Kanamori works as a part-time clinical research consultant for Metabolon, Inc.

All other authors declare no conflicts of interest that could be perceived as prejudicing the impartiality of the research reported.

Abstract

Context: There is a need for novel biomarkers and better understanding of the pathophysiology of diabetic kidney disease.

Objective: To investigate associations between plasma metabolites and kidney function in people with type 2 diabetes (T2D).

Design: 3,089 samples from individuals with T2D, collected between 1999 and 2015, from five independent Dutch cohort studies were included. Up to 7 years follow-up was available in 1,100 individuals from two of the cohorts.

Main outcome measures: Plasma metabolites (*n=149)* were measured by nuclear magnetic resonance spectroscopy. Associations between metabolites and estimated glomerular filtration rate (eGFR), urinary albumin-to-creatinine ratio (UACR) and eGFR slopes were investigated in each study followed by random effect meta-analysis. Adjustments included traditional cardiovascular risk factors and correction for multiple testing.

*Results***:** In total, 125 metabolites were significantly associated ($P_{FDR} = 1.5\times10^{-32}$ - 0.046; β = -11.98-2.17) with eGFR. Inverse associations with eGFR were demonstrated for branched-chain and aromatic amino acids, glycoprotein acetyls, triglycerides, lipids in VLDL subclasses, and fatty acids ($P_{FDR}<0.03$). We observed positive associations with cholesterol and phospholipids in HDL and Apo A1 ($P_{FDR}<0.05$). Albeit some metabolites were associated with UACR levels ($P<0.05$), significance was lost after correction for multiple testing. Tyrosine and HDL-related metabolites were positively associated with eGFR slopes before adjustment for multiple testing ($P_{Tyr}=0.003$; $P_{HDLrelated}<0.05$), but not after.

Conclusions: This study identified metabolites associated with impaired kidney function in T2D, implying involvement of lipid and amino acid metabolism in the pathogenesis. Whether these processes precede or are consequences of renal impairment needs further investigation.

Précis

Plasma metabolomics of 3,089 individuals with type 2 diabetes identified associations of metabolites, mainly from the lipid and amino acid metabolism, with impaired kidney function in a meta-analysis.

Introduction

Diabetic kidney disease (DKD) is a frequent complication of diabetes. DKD may lead to endstage renal disease and is independently associated with a higher risk of all-cause and cardiovascular mortality [1]. DKD is often asymptomatic until the very late stages. Therefore, yearly screening of individuals with type 2 diabetes (T2D) with measurement of kidney functions through estimated glomerular filtration rate (eGFR) and urinary albumin to creatinine ratio (UACR), is recommended in clinical practice. eGFR and UACR are surrogate markers of DKD [2]; however, UACR may not be affected in all individuals with DKD and the decline in eGFR, albeit gradual, is majorly detectable in the later stages of DKD (CKD 3 onwards). Moreover, targeted treatment options for DKD are missing and, thus, are currently limited to control of traditional cardiovascular risk factors such as levels of blood pressure, blood lipids, and blood glucose [3]. There have been several genome-wide association studies (GWAS) for DKD and kidney function [4-7] suggesting a genetic component. There is an urgent need to identify lifestyle-associated biomarkers for early detection of individuals at a risk for DKD and related metabolic functions [8]. Using surrogate quantitative measures for DKD (kidney function decline, and albuminuria) may offer greater statistical power and a better understanding of DKD pathophysiology, further leading to discovery of novel treatment targets.

Advances in metabolomics technologies have allowed for a more in-depth characterization of circulating metabolites, thereby adding information simultaneously on multiple metabolic pathways and allowing for a better understanding of the underlying metabolic processes in DKD [9]. Additionally, this also adds to the missing lifestyle information required to uncover novel disease mechanisms. Nuclear magnetic resonance (NMR) spectrometry provides a platform for the targeted measurement of both amino acids, lipoprotein subclasses, and other metabolites [10 11].

Previous studies involving NMR metabolomics and DKD have mainly been performed in individuals with type 1 diabetes [12-16]. Two recent European studies, using NMR metabolomics in T2D, demonstrated that tyrosine is a marker of microvascular complications, but DKD was not investigated separately [17]. Second, Barrios et al. demonstrated associations of several metabolites, mainly lipids, amino acids, and energy metabolites, with measures of kidney function and incident DKD in 926 persons with T2D and 4838 persons without diabetes [18], although only taking into account a limited number of relevant confounders.

The aim was to investigate associations between plasma metabolites and kidney function in five independent Dutch cohorts of individuals all diagnosed with T2D. This study is hypothesis-generating and may identify more metabolic traits of both amino acids and lipids in DKD.

Materials and methods

Participants

In total, 3,089 persons with T2D from five independent Dutch cohort studies, the Hoorn Diabetes Care System *(DCS)* West-Friesland [19], the Maastricht Study [20], the Rotterdam Study *(RS)* [21], the Netherlands Epidemiology of Obesity Study *(NEO)* [22] and the Cohort of Diabetes and Atherosclerosis Maastricht study *(CODAM)* [23] were included. The selection processes of the independent studies have previously been described in detail; a brief description of the selection from each cohort for the present study, is provided below. The studies were all conducted following the Declaration of Helsinki, the local ethics committees approved the original protocols, and all participants gave written informed consent.

The Hoorn Diabetes Care System West-Friesland

The DCS provides diabetes care to people with T2D living in the West-Friesland region, who yearly visit the DCS research centre [19]. At the yearly visits, a medical exam is performed, and blood is drawn for biochemistry. Individuals are advised on health and treatment and have been invited to participate in the DCS research and biobank *(n=5,000+)*. For the present study, a random sample of individuals from the DCS biobank *(n=750)* as well as a selected group of individuals *(n=245)* was included, all with available plasma samples collected in 2008-2009. The selected group consisted of individuals with known diabetes complications and individuals who were unable to reach the treatment target of HbA_{1c} < 53 mmol/mol. Annual measurements of eGFR were available for calculation of eGFR slopes (median four years, interquartile range 2-6 years) in all participants.

The Maastricht Study

The Maastricht Study is a prospective population-based cohort study of individuals aged between 40 and 75 years in the southern part of the Netherlands. Inclusion began in 2010 and is ongoing [20]. The cohort is enriched with people with T2D. The study is an in-depth phenotyping study focusing on aetiology, complications and comorbidities of T2D. For the present study, all participants with T2D and available plasma samples *(n=848)* were included.

The Rotterdam Study

The Rotterdam Study is a prospective population-based cohort study in the Ommoord district in Rotterdam [21]. All inhabitants in the district, aged above 55 years, were invited to participate in this

study since 1989, with visits being performed every 3-4 years, the plasma samples analysed for the present study were from RS 1-4 and RS 2-2 cohorts (2002-2005) including people with T2D (*n=426*).

The Netherlands Epidemiology of Obesity Study

The NEO study is a prospective population-based cohort study, including individuals aged between 45 and 65 years *(n=6,671)* from 2008-2012, designed for deep phenotyping of pathways leading to obesity-related diseases [22]. In the present study, all individuals with T2D at baseline (*n=675*), were included.

The Cohort of Diabetes and Atherosclerosis Maastricht study

The CODAM study is a prospective observational cohort study of individuals at increased risk of T2D and cardiovascular disease aged above 40 *(n=574)* aiming to investigate the effects of glucose metabolism, lipids, lifestyle, and genetics on (development of) T2D and cardiovascular complications [23]. Baseline samples were collected from 1999-2002, and eGFR measurements were available from a follow-up visit seven years after baseline. All individuals with T2D and available plasma samples (*n=145)* were included in the present study.

Outcome

The eGFR was calculated from serum creatinine measured locally by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation [24]. eGFR slope was calculated based on measurements from annual visits in DCS, in participants with at least two measurements and a minimum follow-up of three years and from the 7-year follow-up visit in CODAM.

Urinary albumin to creatinine ratio (UACR) was measured in first-morning void spot urine in NEO, CODAM, and DCS. In the Maastricht study, urine albumin excretion (UAE) was based on the average

of two 24-hour urine collections. Urine albumin excretion was not measured in the Rotterdam study. People were stratified as having microalbuminuria if UACR was ≥ 2.5 mg/mmol for men and 3.5 mg/mmol for women. In the Maastricht study, microalbuminuria was present if the UAE was ≥ 30 mg/d. Macroalbuminuria was defined as UACR ≥ 25 mg/mmol (men) and 35 mg/mmol (women), in the Maastricht study as UAE ≥ 300 mg/d.

Standardized methods measured levels of Haemoglobin A1c (HbA1c), serum/plasma cholesterol, and triglyceride. Brachial blood pressure was measured after at least 5 min rest with an automatic device and an appropriately sized cuff. Body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared. Smoking status was defined as yes or no current smoker. Diabetes duration was obtained from medical records or self-reported. Medication use was registered according to Anatomical Therapeutic Chemical (ATC) classification coding: statins (C10AA, C10BA, C10BX), other lipid-modifying agents (C10AB, C10AC, C10AD, C10AX, C10BA), renin-angiotensin system (RAS)-blocking agents (angiotensin converting enzyme inhibitors and angiotensin II receptor blockers, C09), all other antihypertensives (C02, C03, C07, C08), oral glucose lowering drugs (A10B, mainly metformin and sulfonylurea) and insulins (A10).

Metabolic Biomarker Profiling

The fasted EDTA plasma samples were stored at -80°C until analysis. The sample storage time varied from 1-15 years. Metabolic biomarkers (*n=149*) were quantified from plasma samples of 3,089 individuals using high-throughput proton NMR metabolomics (Nightingale Health Ltd, Helsinki, Finland). The method provides concurrent quantification of lipids, 14 lipoprotein subclasses, fatty acid composition, and various low-molecular metabolites, including amino acids, ketone bodies, and gluconeogenesis-related metabolites in molar concentration units [25]. The samples do not undergo any

extraction steps and the serum samples are never in contact with the NMR detector; thus, there is no significant batch effect in the NMR-based metabolite quantification. Since the pre-analytical conditions may vary slightly between different studies, it is though recommended to meta-analyse the data as has been done in the present study. Details of the experimentation and applications of the NMR metabolomics platform have been described previously [10 11]. Metabolites with equal to or less than 20% missing values were included. After excluding missing data (on metabolomics measures), we performed quantile normalization (using the R functions, "scale" and "quantile") where we added the 10^{th} percentile to the normalized values, on natural log-transformed data (normally distributed).

In order to check the relatedness between individual metabolite levels found associated with renal function, we performed a sensitivity analysis (pairwise correlation) within the DCS cohort (n=995).

Statistical analyses

Continuous variables were reported as means ± standard deviation (SD) for normally distributed data, skewed data were reported as median [interquartile range (IQR)]. Categorical variables were presented as total numbers with corresponding percentages. The combined summary for all variables was performed using a weighted arithmetic mean method.

Cross-sectional analyses of each cohort using linear regression were performed to assess associations of single plasma metabolites with two continuous outcome variables: 1) eGFR and 2) UACR/UAE. Cross-sectional analyses of each cohort using logistic regression were performed to test associations between single plasma metabolites and the following two categorical outcomes related to deteriorating kidney function: 1) eGFR < 60 ml/min/1.73m^2 and 2) micro- or macroalbuminuria. Longitudinal eGFR measurements were used to calculate annual slopes for two of the cohorts (DCS and CODAM). Linear regression analysis was performed to assess associations between metabolites and eGFR slopes.

Adjustment of potential confounders included the following: age, sex, use of statins, other lipid-modifying agents, oral glucose-lowering medications, insulins, RAS-blocking agents and other antihypertensives, systolic blood pressure, body mass index, smoking, diabetes duration, HbA_{1c}, and baseline UACR/UAE or eGFR, where appropriate. In RS, no urine albumin or HbA_{1c} assessment was performed, and no data on diabetes duration was available. This cohort was, therefore, only included in analyses of eGFR and not adjusted for UAER, diabetes duration, or HbA_{1c}. Results from NEO were not adjusted for diabetes duration because of many missing observations. Individual small-sized cohorts (n < 200) having a low number of cases (n_{cases_eGFR} and $n_{cases\_UACR} \leq 10\%$) while running the logistic regression models were excluded from the meta-analysis [25]. Also, individuals with missing covariate data were excluded from each analysis.

A random-effects meta-analysis of the respective study sets was performed using the R meta-package (Meta v4.8-4) for cross-sectional and longitudinal data (eGFR slopes). We compared the results from the meta-analyses of four cohorts in the logistic models and of five cohorts in the linear models with the results from the meta-analyses including only the three cohorts with all covariates available (DCS, MS and CODAM). For this analyses, a meta regression model in the R package called *metaphor* [26] was applied. We used a fixed-effects model because the (residual) heterogeneity within each subset has already been accounted for by fitting random-effects models [27]. The fixed-effects model did not substantially change the results ($p > 0.05$)[25] and therefore data from the meta-analyses including data from all five cohorts are presented, unless stated otherwise. Further, sensitivity analyses were performed in the largest cohort (the DCS cohort): 1) excluding persons aged above 75 years and 2) including only individuals with eGFR $\geq$ 60 ml/min/1.73m^2.

Correction for multiple testing was performed by the false-discovery rate (FDR) method [28]. A two-tailed FDR-adjusted p-value (P_{FDR}) < 0.05 was considered statistically significant. Data analysis was performed with R-Studio v1.0.143.

For sensitivity analysis pairwise correlation coefficients (r2) using Pearson's method were estimated for the scaled, non-missing metabolite levels associating with eGFR (n=125). These measures (r2) were plotted as a heatmap using the 'heatmap.2' function in the statistical R package 'gplots'. The dendrogram (hierarchical clustering) based on correlation as a distance measure and the metabolite groupings (n=14) were also added to the correlation heatmap.

Results

In the combined populations, 59% of the individuals were men, the mean ± SD age was 64 ± 8 years, and the mean eGFR was 82 ± 16 ml/min/1.73m^2. Clinical characteristics for each participating cohort and their combined summary are presented in table 1. The largest differences between cohorts were observed in means of age, systolic blood pressure, and diabetes duration, and use of different classes of medications. Heterogeneity between cohorts on baseline characteristics was accounted for using random-effects meta-analysis. Overall, at baseline, 332 individuals had microalbuminuria, 40 had macroalbuminuria, 293 had eGFR 30-60 ml/min/1.73 m^2 and 7 had eGFR ≤ 30 ml/min/1.73 m^2. The eGFR slopes were based on a median of 4 measurements in 994 participants in DCS and 2 measurements in 106 participants in CODAM, respectively. The mean percentage of missing metabolite values in each of the cohorts was between 0.1 % and 0.6 %, ranging from minimum 0 % to maximum 19%.

Cross-sectional associations between metabolites and eGFR

As a continuous measure, eGFR was significantly associated with 125 metabolites (Figure 1, Table 2) [25]. The aromatic amino acids (AAAs) phenylalanine (β = -3.05, $P_{FDR} = 1.5 \times 10^{-32}$, P_{het}= 0.99), and histidine (β = -1.11, $P_{FDR} = 3.6 \times 10^{-5}$, P_{het}= 0.5), the branched-chain amino acid (BCAAs) isoleucine (β = -1.92, $P_{FDR} = 3.8 \times 10^{-10}$, P_{het}= 0.49) and leucine (β = -1.16, $P_{FDR} = 1.4 \times 10^{-3}$, P_{het}= 0.22), and the non-essential amino acid glutamine (β = -1.54, $P_{FDR} = 8.2 \times 10^{-10}$, P_{het}= 0.60) were strongly and inversely associated with eGFR, while the AAA tyrosine was positively associated with eGFR. The glycolysis related metabolites glucose and lactate were also positively associated with eGFR, while citrate and glycoprotein acetyls were inversely associated with eGFR (P_{FDR}<0.001). Inverse associations with eGFR were observed for acetoacetate, measures of fatty acids, triglycerides (TGs) in all lipoprotein subclasses, cholesterols, phospholipids in and particle concentrations of very low-density lipoproteins (VLDL), intermediate density lipoprotein (IDL), low-density lipoproteins (LDL), sphingomyelins and apolipoprotein B (P_{FDR}<0.05). Cholesterol and phospholipids in and particle concentrations of high-density lipoprotein (HDL) subclasses, HDL particle size, and apolipoprotein A1 were positively associated with eGFR (P_{FDR}<0.05).

In the logistic regression analyses of having a low eGFR < 60 ml/min/1.73 m^2, significant results were demonstrated for 106 metabolite measures, of which 104 overlapped with those significant in the linear regression analyses [25]. Phenylalanine was the strongest signal associated with the maximum likelihood of having low eGFR in the logistic regression model (OR = 1.67, $P_{FDR} = 4.1 \times 10^{-13}$, P_{het}= 0.38).

Cross-sectional associations between metabolites and albuminuria

UAER/UAE level was associated with 11 metabolites (p<0.05, Figure 2) of which nine were also associated with eGFR. Positive associations with glucose, glycoprotein acetyls, phosphatidylcholine, and

two lipid measures in VLDL subclasses were demonstrated. Citrate, glutamine, and (free) cholesterol and phospholipids in very large HDL were negatively associated with UACR. Significance for all tested associations was lost after correction for multiple testing [25]. In the logistic regression analyses, albuminuria (micro- or macroalbuminuria) was associated with 22 metabolite measures ($p<0.05$) of which 18 were also associated with eGFR. Tyrosine was inversely associated with albuminuria while glucose, glycoprotein acetyls, phosphoglycerides, phosphatidylcholine, apolipoprotein B, content of TGs in lipoproteins, total, free, and VLDL cholesterol, some lipids in (very) small VLDL, and several fatty acids were positively associated with albuminuria. However, after adjustment for multiple testing, the results were no longer significant [25]. In the Maastricht Study, where albuminuria was measured in 24-hour urine samples, we observed minor changes to betas and p-values when adjusting for eGFR. Moreover, the heterogeneity between studies was small.

Associations between metabolites and eGFR slope

Eleven metabolites were associated with eGFR slopes ($p<0.05$) before adjustment for multiple testing. Tyrosine and HDL related metabolites were positively associated with eGFR slopes and thereby demonstrated the same directionality as in the cross-sectional analyses. However, after adjustment for multiple testing the results were no longer significant [25].

Sensitivity analyses

Sensitivity analyses, including only the DCS cohort and excluding individuals aged above 75 years (n=163), did not substantially affect the results. This result suggests that the observed associations were not driven by age. We also did not observe significant differences in the metabolites most significantly associated with eGFR after omission of people with eGFR < 60 ml/min/m^2 (n=237).

It is known that eGFR tends to increase initially and then decrease during DKD progression. Therefore we performed another sensitivity analysis within the DCS cohort where we compared metabolite associations with eGFR tertiles. We observed 74% consistency in the directionality of effects across tertiles (2 out of 3). 100% consistency could not be achieved due to a lack of statistical power for this nested analysis in the DCS cohort (data not shown).

Thirdly, to account for the relatedness between different metabolite levels associated with eGFR/UACR, a heatmap depicting correlation between individual metabolites, and metabolite groups has been illustrated within the DCS cohort (n=995). The lipoprotein groups indicate that the metabolites are not entirely independent. A general trend shows highly correlated metabolites within a specific lipoprotein group (VLDL lipoproteins and HDL lipoproteins) groups while also highlighting inter group differences (positive (blue) vs. negative (red) correlations; Figure 3).

Discussion

In the present study, we investigated associations between plasma metabolites and kidney function (cross-sectionally and longitudinally) in individuals with T2D. The key findings include several novel associations between eGFR and circulating amino acids, triglycerides, lipids in VLDL subclasses, free fatty acids, and lipids in HDL subclasses cross-sectionally. eGFR slopes associated with tyrosine and subclasses of HDL lipoproteins (before correction for multiple testing). None of the metabolites measured were significantly associated with urinary albumin excretion. These findings are in line with results from a recent study in type 1 diabetes where metabolites were mainly associated with eGFR and not albuminuria [29].

The aromatic amino acid phenylalanine was strongest inversely associated with eGFR levels, followed by the branched-chain amino acid isoleucine and the polar amino acid glutamine. This result replicates

findings from a recent study by Barrios et al., where a strong inverse cross-sectional association between phenylalanine and eGFR was demonstrated both in 926 individuals with T2D and 4,838 individuals without diabetes, after adjusting for age, sex, BMI, statin use and hormone replacement therapy [18]. They did not find an association between phenylalanine and longitudinal changes in eGFR, similar to the findings in the present study. Higher phenylalanine and lower tyrosine levels due to impaired renal conversion of phenylalanine, have previously been reported in renal disease [30 31]. In our study, a positive association between tyrosine and eGFR was also observed cross-sectionally and longitudinally. Recently, the results of ADVANCE trial in 3587 individuals with T2D demonstrated a positive association of phenylalanine with macrovascular diseases and all-cause mortality; however, this was attenuated after adjustment for cardiovascular risk factors. Moreover, this study showed an inverse association between tyrosine and microvascular complications [17]. This trial demonstrated that higher levels of aromatic (histidine) and branched-chain (leucine) amino acids, were associated with a lower risk of all-cause mortality. This is in line with our findings in which histidine and leucine were inversely associated with eGFR. In a study by Niewczas et al. comparing T2D progressors (n=40) to end-stage renal disease (ESRD) and non-progressors (n=40), higher levels of phenylalanine, tyrosine and leucine were associated with a lower risk of progression to ESRD during 12 years follow-up, albeit not statistically significant after adjustment for HbA_{1c}, albumin excretion, eGFR and multiple testing [32]. The population had a longer diabetes duration compared to the present and one may speculate that a different metabolite profile could be observed in the early courses of T2D and/or DKD. Unfortunately, we were limited to examining the longitudinal associations with the eGFR slope and did not have data on ESRD. On the other hand, our study is statistically well-powered compared to the study by Niewczas et al. Moreover, methodological differences between the current (NMR based) and previous studies referred to above (mass-spectrometry (MS) based) might explain some of the discrepancies.

In the present study, isoleucine and leucine were inversely linearly associated with eGFR, and in line with this, higher isoleucine levels were also associated with an increased likelihood of decreased kidney function in the logistic model. In a previous paper from the same five cohorts, higher levels of isoleucine and leucine were associated with a higher risk of having an HbA_{1c} > 53 mmol/l as a measure of dysregulated diabetes [33]. A link between BCAAs and the development of diabetes [34] as well as insulin resistance [35 36], has previously been demonstrated. This suggests that BCAAs (isoleucine and leucine) levels may be indicative of not only dysregulated diabetes and insulin resistance but also progression to DKD. Indeed, insulin resistance has previously been hypothesized to play an essential role in DKD [37], especially in individuals with T2D. Several mechanisms could potentially explain the increased levels of BCAAs in individuals with insulin resistance. For instance, in individuals without diabetes, it has been proposed that high BCAA levels originate from the gut microbiome [38]. Exploring this further in individuals with T2D could shed further light on the pathophysiology and potentially reveal treatment targets.

We demonstrated in the current study, an inverse association between eGFR and glutamine in linear regression analyses. This inverse relationship is in line with the previous study by Niewczas et al., where higher glutamine was associated with a higher risk of ESRD, although not statistically significant after adjustment for HbA_{1c}, albumin excretion, eGFR and multiple testing [32]. In contrast, we previously reported that higher glutamine was most significantly associated with having an HbA_{1c} < 53 mmol/l [33].

In the present study, eGFR was inversely associated with TGs, lipid measures in VLDL subclasses, sphingomyelin and fatty acids like omega-6 and linoleic acid but positively associated with lipid measures, except triglycerides, in HDL subclasses. This inverse relationship is in line with results in both individuals with and without diabetes, from the previously mentioned study by Barrios et al [18]. The associations between these lipid measures and longitudinal endpoints of ESRD, macro- or microvascular

complications have to our knowledge not been tested in any study in people with T2D. We identified the lipids in HDL to be associated positively with longitudinal eGFR slopes among T2D, which became non-significant after FDR correction. These could be correct signals reflecting cross-sectional results on eGFR in the current study. The longitudinal results in the current study may have missed the multiple testing threshold due to lower statistical power (n=1100) compared to cross-sectional meta data where we had 3089 individuals. On the other hand, we cannot undermine the possibility of inter-metabolite relatedness as suggested by the sensitivity analysis (correlation heatmap) that may reduce the number of tests. However, in a longitudinal study by the FinnDiane Study Group including 3544 individuals with type 1 diabetes, the triglyceride-cholesterol imbalance was associated with progression in albuminuria and all-cause mortality, as was a large HDL cholesterol [14]. Previous studies in large cohorts of dyslipidaemia and DKD have demonstrated associations between routinely measured higher TG and lower HDL cholesterol levels to the progression of DKD in individuals with T2D [39 40]. Higher TG and lower HDL cholesterol are clinical components of the metabolic syndrome and may also be consequences of the underlying insulin resistance [41]. Targeting HDL cholesterol or phospholipids may be a future therapeutic approach, although previous clinical studies with HDL cholesterol increasing agents, for example torcetrapib [42] in high-risk individuals of cardiovascular disease, have demonstrated increased risk of mortality and morbidity of unknown mechanism. In a small study by Drew et al (n=13), intravenous reconstituted HDL infusion in patients with T2D increased plasma insulin. It activated AMP-activated protein kinase in skeletal muscle when compared to placebo [43].

The major strengths of this study are a the large number of individuals with T2D from five independent cohorts as well as adjustment for an extensive set of relevant clinical covariates and multiple testing, which diminished the risk of false-positive results. Moreover, the use of the NMR platform provided standardized measures of metabolites, allowing exploration of measures beyond routinely measured

biomarkers. This method was faster, cheaper, and thereby more accessible than the more extensive MS-based metabolomics methods and could potentially be easier to adapt to a clinical setting, although it did not provide as in-depth a characterisation as for example MS. The use of medications agents may affect the metabolome substantially. In a recently published paper, also from the BBMRI consortium, significant associations between cardiometabolic agents and several of the measured metabolites were demonstrated [44]. In the present study, results were adjusted for several medications commonly used in diabetes. We further recognize some limitations of the study. Metabolites were measured at one timepoint and therefore did not provide information regarding within-subject variation in metabolite levels. In some of the original studies, information regarding albuminuria measurements and history of micro- or macroalbuminuria was very limited. Further, we were limited by a low number of cases in the CODAM study, no diabetes duration in the RS study, and only having follow-up measurements from two cohorts. Information regarding diet, which may affect the measured metabolites, was not available in all cohorts and not uniformly captured. Therefore, we could not control for diet in this study. Information regarding other renal markers, such as for example vitamin D, parathyroid hormone and calcium, was also not available. Since the samples originate from different studies, the storage time before metabolite measurements differed. Though all samples have been stored at -80°C as recommended [45] and we did find consistent results across studies, regardless of the year of sampling. This finding was also demonstrated by low heterogeneity between studies, suggesting no such influence.

In conclusion, the current largest-to-date study identifies metabolites associated with an impaired eGFR among individuals with T2D while none of the metabolites measured were significantly associated with urinary albumin excretion. These results suggest that cholesterol and phosphoglycerides in HDL subclasses may be associated with a better kidney function while high levels of several amino acids and

fatty acids, lipids in VLDL subclasses, and TGs in all lipoproteins are associated with an impaired kidney function. The findings suggest alterations in the metabolome associated with renal impairment in T2D of primary importance in non-albuminuric DKD or the early stages of disease before development of albuminuria. Further longitudinal studies are needed to clarify whether alterations in metabolite levels precede or are consequences of renal impairment and whether a biomarker panel of both amino acid and lipid measures could potentially lead to improved prediction in the development of DKD.

Author contributions

NT, NV, MvG, CvdK, GN, PR, TSA, and LMtH conceived and designed the research. All authors contributed to the acquisition and/or interpretation of the data; NV, DM-K, AB, JN, FA, and TSA performed the statistical analyses; TSA also performed the overall meta-analyses; NT drafted the manuscript. All authors critically revised the manuscript for key intellectual content, suggested revisions, and approved the final version of the manuscript. The authors declare that there are no conflicts of interest that could be perceived as prejudicing the impartiality of the research reported.

Acknowledgements

The authors would like to thank all people in the studies for their cooperation.

Data Availability

The datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding authors on reasonable request.

References

1. Bjerg L, Hulman A, Carstensen B, Charles M, Witte DR, Jorgensen ME. Effect of duration and burden of microvascular complications on mortality rate in type 1 diabetes: an observational clinical cohort study. Diabetologia 2019 doi: 10.1007/s00125-019-4812-6[published Online First: Epub Date]|.
2. Fox CS, Matsushita K, Woodward M, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. Lancet 2012;**380**(9854):1662-73 doi: 10.1016/S0140-6736(12)61350-6[published Online First: Epub Date]|.
3. National Kidney F. KDOQI Clinical Practice Guideline for Diabetes and CKD: 2012 Update. Am J Kidney Dis 2012;**60**(5):850-86 doi: 10.1053/j.ajkd.2012.07.005[published Online First: Epub Date]|.
4. Wuttke M, Li Y, Li M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet 2019;**51**(6):957-72 doi: 10.1038/s41588-019-0407-x[published Online First: Epub Date]|.
5. 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**(7):1414-27 doi: 10.2337/db17-0914[published Online First: Epub Date]|.
6. Salem RM, Todd JN, Sandholm N, et al. Genome-Wide Association Study of Diabetic Kidney Disease Highlights Biology Involved in Glomerular Basement Membrane Collagen. Journal of the American Society of Nephrology : JASN 2019;**30**(10):2000-16 doi: 10.1681/ASN.2019030218[published Online First: Epub Date]|.
7. Ahluwalia TS, Schulz CA, Waage J, et al. A novel rare CUBN variant and three additional genes identified in Europeans with and without diabetes: results from an exome-wide association study of albuminuria. Diabetologia 2019;**62**(2):292-305 doi: 10.1007/s00125-018-4783-z[published Online First: Epub Date]|.
8. Ahmad S, Ahluwalia TS. Editorial: The Role of Genetic and Lifestyle Factors in Metabolic Diseases. Front Endocrinol (Lausanne) 2019;**10**:475 doi: 10.3389/fendo.2019.00475[published Online First: Epub Date]|.
9. Ahluwalia TS, Kilpelainen TO, Singh S, Rossing P. Editorial: Novel Biomarkers for Type 2 Diabetes. Front Endocrinol (Lausanne) 2019;**10**:649 doi: 10.3389/fendo.2019.00649[published Online First: Epub Date]|.
10. Soininen P, Kangas AJ, Wurtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circ Cardiovasc Genet 2015;**8**(1):192-206 doi: 10.1161/CIRCGENETICS.114.000216[published Online First: Epub Date]|.
11. Soininen P, Kangas AJ, Wurtz P, et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 2009;**134**(9):1781-5 doi: 10.1039/b910205a[published Online First: Epub Date]|.
12. Makinen VP, Soininen P, Forsblom C, et al. Diagnosing diabetic nephropathy by 1H NMR metabonomics of serum. MAGMA 2006;**19**(6):281-96 doi: 10.1007/s10334-006-0054-y[published Online First: Epub Date]|.
13. Makinen VP, Soininen P, Forsblom C, et al. 1H NMR metabonomics approach to the disease continuum of diabetic complications and premature death. Mol Syst Biol 2008;**4**:167 doi: 10.1038/msb4100205[published Online First: Epub Date]|.
14. Makinen VP, Soininen P, Kangas AJ, et al. Triglyceride-cholesterol imbalance across lipoprotein subclasses predicts diabetic kidney disease and mortality in type 1 diabetes: the FinnDiane Study. J Intern Med 2013;**273**(4):383-95 doi: 10.1111/joim.12026[published Online First: Epub Date]|.
15. Makinen VP, Tynkkynen T, Soininen P, et al. Metabolic diversity of progressive kidney disease in 325 patients with type 1 diabetes (the FinnDiane Study). Journal of proteome research 2012;**11**(3):1782-90 doi: 10.1021/pr201036j[published Online First: Epub Date]|.

16. Makinen VP, Kangas AJ, Soininen P, Wurtz P, Groop PH, Ala-Korpela M. Metabolic phenotyping of diabetic nephropathy. Clin Pharmacol Ther 2013;**94**(5):566-9 doi: 10.1038/clpt.2013.158[published Online First: Epub Date]|.
17. Welsh P, Rankin N, Li Q, et al. Circulating amino acids and the risk of macrovascular, microvascular and mortality outcomes in individuals with type 2 diabetes: results from the ADVANCE trial. Diabetologia 2018;**61**(7):1581-91 doi: 10.1007/s00125-018-4619-x[published Online First: Epub Date]|.
18. Barrios C, Zierer J, Wurtz P, et al. Circulating metabolic biomarkers of renal function in diabetic and non-diabetic populations. Sci Rep 2018;**8**(1):15249 doi: 10.1038/s41598-018-33507-7[published Online First: Epub Date]|.
19. van der Heijden AA, Rauh SP, Dekker JM, et al. The Hoorn Diabetes Care System (DCS) cohort. A prospective cohort of persons with type 2 diabetes treated in primary care in the Netherlands. BMJ Open 2017;**7**(5):e015599 doi: 10.1136/bmjopen-2016-015599[published Online First: Epub Date]|.
20. Schram MT, Sep SJ, van der Kallen CJ, et al. The Maastricht Study: an extensive phenotyping study on determinants of type 2 diabetes, its complications and its comorbidities. Eur J Epidemiol 2014;**29**(6):439-51 doi: 10.1007/s10654-014-9889-0[published Online First: Epub Date]|.
21. Ikram MA, Brusselle GGO, Murad SD, et al. The Rotterdam Study: 2018 update on objectives, design and main results. Eur J Epidemiol 2017;**32**(9):807-50 doi: 10.1007/s10654-017-0321-4[published Online First: Epub Date]|.
22. de Mutsert R, den Heijer M, Rabelink TJ, et al. The Netherlands Epidemiology of Obesity (NEO) study: study design and data collection. Eur J Epidemiol 2013;**28**(6):513-23 doi: 10.1007/s10654-013-9801-3[published Online First: Epub Date]|.
23. Jacobs M, van Greevenbroek MM, van der Kallen CJ, et al. Low-grade inflammation can partly explain the association between the metabolic syndrome and either coronary artery disease or severity of peripheral arterial disease: the CODAM study. Eur J Clin Invest 2009;**39**(6):437-44 doi: 10.1111/j.1365-2362.2009.02129.x[published Online First: Epub Date]|.
24. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;**150**(9):604-12
25. Nete T, Nicole V, Dennis M-K, et al. *Data from: Plasma metabolomics identifies markers of impaired renal function: A meta-analysis of 3,089 persons with type 2 diabetes. Deposited to Figshare.com 20 March 2020.* *http://doi.org/10.6084/m9.figshare.10050068.v1*, 2020.
26. Dokkedahl S, Kok RN, Murphy S, Kristensen TR, Bech-Hansen D, Elklit A. The psychological subtype of intimate partner violence and its effect on mental health: protocol for a systematic review and meta-analysis. Syst Rev 2019;**8**(1):198 doi: 10.1186/s13643-019-1118-1[published Online First: Epub Date]|.
27. Rubio-Aparicio M, López-López JA, Viechtbauer W, Marín-Martínez F, Botella J, Sánchez-Meca J. Testing Categorical Moderators in Mixed-Effects Meta-analysis in the Presence of Heteroscedasticity. The Journal of Experimental Education 2019:1-23 doi: 10.1080/00220973.2018.1561404[published Online First: Epub Date]|.
28. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. J Clin Epidemiol 2014;**67**(8):850-7 doi: 10.1016/j.jclinepi.2014.03.012[published Online First: Epub Date]|.
29. Tofte N, Suvitaival T, Trost K, et al. Metabolomic Assessment Reveals Alteration in Polyols and Branched Chain Amino Acids Associated With Present and Future Renal Impairment in a Discovery Cohort of 637 Persons With Type 1 Diabetes. Front Endocrinol (Lausanne) 2019;**10**:818 doi: 10.3389/fendo.2019.00818[published Online First: Epub Date]|.
30. Kopple JD. Phenylalanine and tyrosine metabolism in chronic kidney failure. J Nutr 2007;**137**(6 Suppl 1):1586S-90S; discussion 97S-98S doi: 10.1093/jn/137.6.1586S[published Online First: Epub Date]|.

31. Druml W, Roth E, Lenz K, Lochs H, Kopsa H. Phenylalanine and tyrosine metabolism in renal failure: dipeptides as tyrosine source. Kidney Int Suppl 1989;**27**:S282-6
32. Niewczas MA, Sirich TL, Mathew AV, et al. Uremic solutes and risk of end-stage renal disease in type 2 diabetes: metabolomic study. Kidney international 2014;**85**(5):1214-24 doi: 10.1038/ki.2013.497[published Online First: Epub Date]|.
33. 't Hart LM, Vogelzangs N, Mook-Kanamori DO, et al. Blood metabolomic measures associate with present and future glycemic control in type 2 diabetes. J Clin Endocrinol Metab 2018 doi: 10.1210/jc.2018-01165[published Online First: Epub Date]|.
34. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;**17**(4):448-53 doi: 10.1038/nm.2307[published Online First: Epub Date]|.
35. Lee CC, Watkins SM, Lorenzo C, et al. Branched-Chain Amino Acids and Insulin Metabolism: The Insulin Resistance Atherosclerosis Study (IRAS). Diabetes Care 2016;**39**(4):582-8 doi: 10.2337/dc15-2284[published Online First: Epub Date]|.
36. Lynch CJ, Adams SH. Branched-chain amino acids in metabolic signalling and insulin resistance. Nat Rev Endocrinol 2014;**10**(12):723-36 doi: 10.1038/nrendo.2014.171[published Online First: Epub Date]|.
37. Karalliedde J, Gnudi L. Diabetes mellitus, a complex and heterogeneous disease, and the role of insulin resistance as a determinant of diabetic kidney disease. Nephrol Dial Transplant 2016;**31**(2):206-13 doi: 10.1093/ndt/gfu405[published Online First: Epub Date]|.
38. Pedersen HK, Gudmundsdottir V, Nielsen HB, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature 2016;**535**(7612):376-81 doi: 10.1038/nature18646[published Online First: Epub Date]|.
39. Russo GT, De Cosmo S, Viazzi F, et al. Plasma Triglycerides and HDL-C Levels Predict the Development of Diabetic Kidney Disease in Subjects With Type 2 Diabetes: The AMD Annals Initiative. Diabetes Care 2016;**39**(12):2278-87 doi: 10.2337/dc16-1246[published Online First: Epub Date]|.
40. Sacks FM, Hermans MP, Fioretto P, et al. Association between plasma triglycerides and high-density lipoprotein cholesterol and microvascular kidney disease and retinopathy in type 2 diabetes mellitus: a global case-control study in 13 countries. Circulation 2014;**129**(9):999-1008 doi: 10.1161/CIRCULATIONAHA.113.002529[published Online First: Epub Date]|.
41. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. Lancet 2005;**365**(9467):1333-46 doi: 10.1016/S0140-6736(05)61032-X[published Online First: Epub Date]|.
42. Barter PJ, Caulfield M, Eriksson M, et al. Effects of torcetrapib in patients at high risk for coronary events. The New England journal of medicine 2007;**357**(21):2109-22 doi: 10.1056/NEJMoa0706628[published Online First: Epub Date]|.
43. Drew BG, Duffy SJ, Formosa MF, et al. High-density lipoprotein modulates glucose metabolism in patients with type 2 diabetes mellitus. Circulation 2009;**119**(15):2103-11 doi: 10.1161/CIRCULATIONAHA.108.843219[published Online First: Epub Date]|.
44. Liu J, Lahousse L, Nivard MG, et al. Integration of epidemiologic, pharmacologic, genetic and gut microbiome data in a drug-metabolite atlas. Nat Med 2020;**26**(1):110-17 doi: 10.1038/s41591-019-0722-x[published Online First: Epub Date]|.
45. Kirwan JA, Brennan L, Broadhurst D, et al. Preanalytical Processing and Biobanking Procedures of Biological Samples for Metabolomics Research: A White Paper, Community Perspective (for "Precision Medicine and Pharmacometabolomics Task Group"-The Metabolomics Society Initiative). Clin Chem 2018;**64**(8):1158-82 doi: 10.1373/clinchem.2018.287045[published Online First: Epub Date]|.

Figure 1 Volcano plot of eGFR and associated metabolites

The Y axis represents $-\log_{10}$ (p_{FDR} value) for metabolite eGFR association. Blue line represents p value = 0.05, and red line FDR-adjusted p value = 0.05. Metabolites are mmol/L, log transformed. P_{FDR} values $< 1.0 \times 10^{-10}$ are depicted as 1.0×10^{-10}, beta estimates $\geq$ -3.0 are depicted as -3.0 to fit into the figure. Top ten most significant metabolites are named. Each dot signifies one metabolite and the size of the dot relates to the observed estimates from the random effects meta-analysis. Colour of the dot determines the metabolite group listed under the colour key.

Figure 2 Volcano plot of UAER/UAE and associated metabolites

The Y axis represents $-\log_{10}$ (p value) for metabolite albuminuria (UAER/UAE) association. Blue line represents p value = 0.05, and red line FDR-adjusted p value = 0.05. Metabolites are mmol/L, log transformed. P_{FDR} values $< 1.0 \times 10^{-10}$ are depicted as 1.0×10^{-10}Beta estimates ≥ 0.15 are depicted as 0.15. Top ten most significant metabolites have been named. Each dot signifies one metabolite and the size of the dot relates to the observed estimates from the random effects meta-analysis. Colour of the dot determines the metabolite group listed under the colour key.

Figure 3 Correlation heatmap of individual metabolites, metabolite groups and hierarchical clustering

125 eGFR associated metabolite levels within the DCS cohort are plotted (quantile normalized). Each small square on the X and Y axis depict individual metabolites. Higher colour intensity (blue/red) indicates high values of positive/negative pairwise correlation measure between metabolites. Dendrogram depict clustering based on correlation distance measures. Horizontal coloured bar on top of the heatmap depicts the metabolite groups.

Table 1 Baseline characteristics of the five study cohorts of people with T2D n = 3,089

	DCS *(n=995)*	**Maastricht** *(n=848)*	**RS** (*n=426)*	**NEO** (*n=675)*	**CODAM** (*n=145)*	**Combined** (*n=3089*)
Age, years	63 (10)	63 (8)	76 (6)	58 (5)	61 (6)	64 (8)
Sex: Men, n (%)	569 (57)	580 (68)	204 (48)	376 (56)	95 (66)	1824 (59)
Diabetes duration, years	5.9 [2.9-9.1]	6.0 [2.0-10.0]	NA	4.6 [0-7]	0.0 [0.0-5.8]	5.2
Body mass index, kg/ m^2	30.6 (5.5)	29.9 (4.9)	29.1 (4.5)	33.0 (5.3)	30.3 (4.6)	30.7 (5.1)
SBP, mmHg	143 (19)	142 (18)	156 (21)	137 (17)	148 (19)	143 (19)
HbA_{1c}, mmol/mol	48 [43-55]	50 [45-56]	NA	49 [40-55]	50 [44-57]	48.9
HbA_{1c}, %	6.5 [6.1-7.2]	6.7 [6.3-7.3]	NA	6.6 [5.8-7.0]	6.7 [6.1-7.3]	6.6
eGFR, ml/min/1.73m^2	81 (18)	83 (16)	73 (14)	88 (14)	92 (14)	82 (16)
UACR, mg/mmol	0.6 [0.4-1.3]	NA	NA	0.6 [0.3-1.0]	0.5 [0.3-1.1]	0.6
UAE, mg/d	NA	10 [5-22]	NA	NA	NA	-
Total cholesterol, mmol/ L	4.7 (1.9)	4.4 (1.0)	5.3 (1.0)	5.0 (1.2)	5.2 (1.2)	4.7 (1.3)
LDL cholesterol, mmol/ L	2.6 (0.9)	2.4 (0.9)	3.6 (0.9)	3.0 (1.1)	3.2 (0.9)	2.7 (0.9)
HDL cholesterol, mmol/ L	1.2 (0.4)	1.3 (0.4)	1.3 (0.4)	1.2 (0.3)	1.1 (0.3)	1.2 (0.3)
Triglycerides, mmol/ L	1.6 [1.2-2.2]	1.5 [1.1-2.1]	1.5 [1.2-2.1]	1.6 [1.2-2.2]	1.8 [1.2-2.4]	1.56
Smoking, n (%)	167 (17)	128 (15)	59 (14)	120 (18)	26 (18)	500 (16)
History of CVD*, n (%)	165 (18)	219 (26)	100 (24)	78 (12)	25 (17)	587 (19)
Medication						
Statins, n (%)	685 (69)	627 (74)	123 (29)	355 (53)	35 (24)	1825 (59)
Other lipid-modifying agents, n (%)	31 (3)	41 (5)	8 (2)	4 (1)	3 (2)	87 (3)
RAS-blocking agents, n (%)	498 (50)	499 (59)	152 (36)	313 (46)	42 (29)	1504 (49)
Other antihypertensives, n (%)	568 (57)	428 (50)	267 (63)	424 (63)	69 (48)	1756 (57)
Oral glucose lowering drugs, n (%)	824 (83)	623 (73)	195 (46)	330 (49)	64 (44)	2036 (66)
Insulin use, n (%)	263 (26)	175 (21)	47 (11)	87 (13)	13 (9)	585 (19)
Categorical outcomes						
eGFR 30-60 ml/min/1.73m^2, n (%)	120 (12)	69 (8)	76 (18)	22 (3)	6 (4)	293 (9)
eGFR ≤ 30 ml/min/1.73m^2, n (%)	4 (<1)	1 (<1)	1 (<1)	1 (<1)	0	7 (0.5)
Microalbuminuria, n (%)	118 (12)	149 (18)	NA	53 (8)	12 (8)	332 (11)
Macroalbuminuria, n (%)	18 (2)	10 (1)	NA	10 (2)	2 (1)	40 (1)

Mean (SD) or median [IQR] for continuous variables, n (%, rounded) for categorical variables. RAS denotes renin-angiotensin system, eGFR estimated glomerular filtration rate, UACR urinary albumin to creatinine ratio, UAE urinary albumin excretion, CVD cardiovascular disease, SBP systolic blood pressure. DCS The Hoorn Diabetes Care System, NEO The Netherlands Epidemiology of Obesity Study, CODAM The Cohort of Diabetes and Atherosclerosis Maastricht study, RS Rotterdam Study. *Self-reported.

Table 2 Metabolites associated with eGFR (n=3,079) – divided by metabolite subgroups

Metabolites	Beta	SE	P	P_{FDR}	Phet
Amino acids					
Phenylalanine	-3.05	0.25	2.0×10^{-34}	1.5×10^{-32}	0.99
Isoleucine	-1.92	0.31	3.9×10^{-10}	1.5×10^{-8}	0.30
Glutamine	-1.54	0.25	8.2×10^{-10}	2.5×10^{-8}	0.49
Histidine	-1.11	0.25	1.0×10^{-5}	3.6×10^{-5}	0.50
Leucine	-1.16	0.35	8.4×10^{-4}	1.4×10^{-3}	0.22
Tyrosine	0.62	0.26	1.7×10^{-2}	2.1×10^{-2}	0.68
Inflammation					
Glycoprotein acetyls, mainly a1-acid glycoprotein	-1.68	0.38	8.9×10^{-6}	3.5×10^{-5}	0.10
Glycolysis related metabolites					
Citrate	-2.27	0.47	1.4×10^{-6}	8.9×10^{-6}	0.02
Lactate	1.04	0.25	2.6×10^{-5}	8.1×10^{-5}	0.47
Glucose	1.12	0.38	3.0×10^{-3}	5.0×10^{-3}	0.99
Ketone bodies					
Acetoacetate	-0.77	0.30	1.0×10^{-2}	1.0×10^{-2}	0.99
Fatty acids					
18:2 Linoleic acid	-1.64	0.29	1.3×10^{-8}	2.2×10^{-7}	0.37
Omega-6 fatty acids	-1.28	0.30	1.6×10^{-5}	5.4×10^{-5}	0.33
Polyunsaturated fatty acids	-1.17	0.31	1.6×10^{-4}	3.4×10^{-4}	0.29
Monounsaturated fatty acids; 16:1, 18:1	-1.10	0.32	6.4×10^{-4}	1.1×10^{-3}	0.21
Total fatty acids	-1.06	0.34	2.0×10^{-3}	3.3×10^{-3}	0.19
Saturated fatty acids	-0.73	0.35	3.9×10^{-2}	4.6×10^{-2}	0.16
Lipoprotein particle sizes					
HDL_D	1.74	0.38	5.7×10^{-6}	2.4×10^{-5}	0.11
VLDL_D	-1.07	0.30	3.0×10^{-4}	5.0×10^{-4}	0.29
Cholesterol					
Total cholesterol in HDL	2.05	0.31	3.2×10^{-11}	1.6×10^{-9}	0.29
Total cholesterol in HDL2	2.17	0.37	5.7×10^{-9}	1.3×10^{-7}	0.13
Total cholesterol in VLDL	-1.81	0.38	2.1×10^{-6}	1.1×10^{-5}	0.10
Remnant cholesterol (non-HDL, non-LDL -cholesterol)	-1.89	0.46	3.9×10^{-5}	1.1×10^{-4}	0.19
Free cholesterol	-1.12	0.29	1.4×10^{-4}	3.1×10^{-4}	0.78
Total cholesterol in IDL	-0.88	0.29	2.2×10^{-3}	3.7×10^{-3}	0.52
Total cholesterol in LDL	-0.80	0.29	5.3×10^{-3}	7.8×10^{-3}	0.43
Serum total cholesterol	-0.63	0.30	3.4×10^{-2}	4.1×10^{-2}	0.59
Apolipoproteins					
Apolipoprotein A-1	1.52	0.28	7.6×10^{-8}	1.1×10^{-6}	0.87
Apolipoprotein B	-1.53	0.42	2.3×10^{-4}	4.7×10^{-4}	0.07
Glycerides & phospholipids					
Triglycerides in HDL	-1.10	0.26	2.3×10^{-5}	7.4×10^{-5}	0.54
Serum total triglycerides	-1.41	0.35	5.1×10^{-5}	1.4×10^{-4}	0.15
Triglycerides in VLDL	-1.44	0.36	5.1×10^{-5}	1.4×10^{-4}	0.13
Diacylglycerol	-0.93	0.36	8.8×10^{-3}	1.2×10^{-2}	0.15
Sphingomyelins	-0.63	0.28	2.7×10^{-2}	3.3×10^{-2}	0.55
Triglycerides in LDL	-0.61	0.29	3.2×10^{-2}	4.0×10^{-2}	0.34
Lipoprotein subclasses of VLDL*					
Cholesterol esters in medium VLDL	-1.72	0.30	1.3×10^{-8}	2.2×10^{-7}	0.27
Cholesterol esters in small VLDL	-2.12	0.38	2.5×10^{-8}	3.9×10^{-7}	0.12
Total cholesterol in medium VLDL	-1.63	0.32	3.0×10^{-7}	3.5×10^{-6}	0.22
Lipoprotein subclasses of HDL*					
Phospholipids in small HDL	1.55	0.26	2.0×10^{-9}	5.0×10^{-8}	0.61
Free cholesterol in small HDL	1.35	0.26	2.0×10^{-7}	2.6×10^{-6}	0.90
Total lipids in small HDL	1.30	0.26	6.3×10^{-7}	5.6×10^{-6}	0.82

Metabolites	Beta	SE	P	P_{FDR}	Phet
Lipoprotein subclasses of IDL*					
Cholesterol esters in IDL	-0.99	0.29	5.5×10^{-4}	1.0×10^{-3}	0.43
Total lipids in IDL	-0.99	0.29	5.8×10^{-4}	1.0×10^{-3}	0.48
Total cholesterol in IDL	-0.87	0.29	2.3×10^{-3}	3.7×10^{-3}	0.52
Lipoprotein subclasses of LDL*					
Triglycerides in small LDL	-0.99	0.31	1.5×10^{-3}	2.5×10^{-3}	0.25
Cholesterol esters in medium LDL	-0.90	0.29	1.6×10^{-3}	2.5×10^{-3}	0.40
Cholesterol esters in large LDL	-0.88	0.29	2.1×10^{-3}	3.3×10^{-3}	0.41

Metabolites are mmol/L except Apolipoproteins which are g/l. All metabolites are log transformed. P_{het} P-value for heterogeneity. *For the lipoprotein subclasses only the three most significant measures in each group are included in this table [25]. Adjustment of potential confounders included the following: age, sex, use of statins, other lipid-modifying agents, oral glucose lowering medications, insulins, RAS-blocking agents and other antihypertensives, SBP, BMI, smoking, diabetes duration, HbA_{1c} and baseline UACR/UAE. In RS, no urine albumin or HbA_{1c} assessment was performed and no data on diabetes duration was available, results from RS were therefore not adjusted for UAER, diabetes duration or HbA_{1c}. Results from NEO were not adjusted for diabetes duration because of many missing observations.

Figure 1

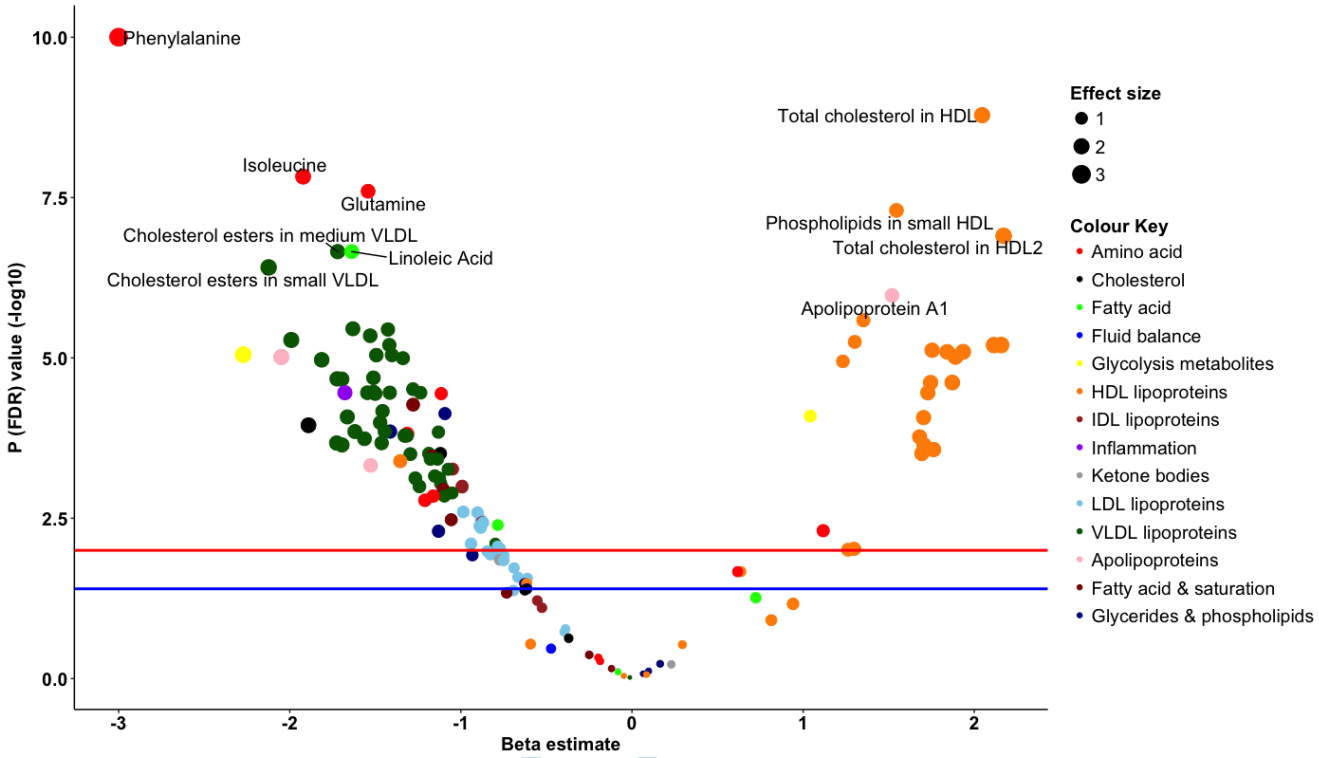

Figure 2

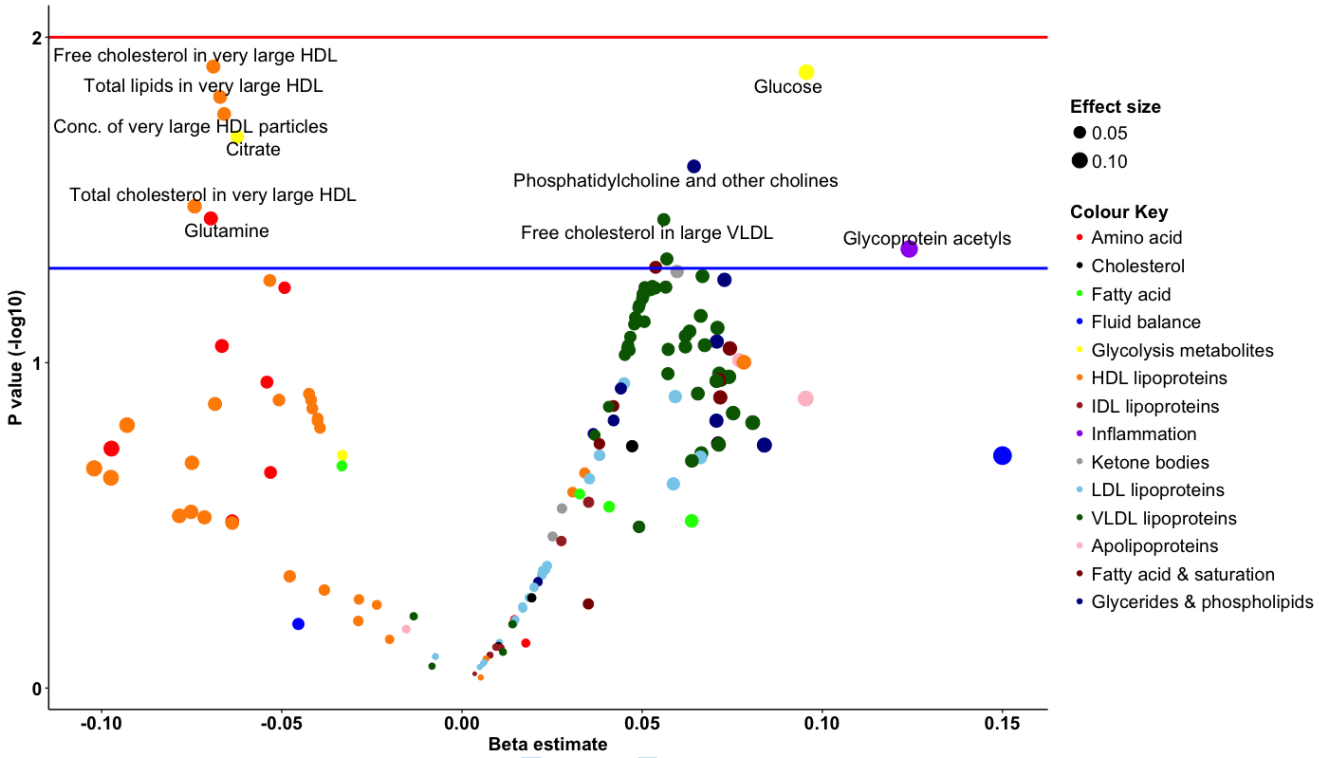

Figure 3