Circulating metabolic biomarkers of renal function in diabetic and non-diabetic populations

Clara Barrios1,2, Jonas Zierer1,3,22, Peter Würtz4,5, Toomas Haller4, Andres Metspalu6, Christian Gieger7,8, Barbara Thorand7, Christa Meisinger7,9, Melanie Waldenberger7,8, Olli Raitakari10,11, Terho Lehtimäki12, Sol Otero2,13, Eva Rodríguez1, Juan Pedro-Botet14, Mika Kähönen15, Mika Ala-Korpela16,17,18,19,20,21, Gabi Kastenmüller1, Tim D. Spector1, Julio Pascual1 & Cristina Menni1

Using targeted NMR spectroscopy of 227 fasting serum metabolic traits, we searched for novel metabolic signatures of renal function in 926 type 2 diabetes (T2D) and 4838 non-diabetic individuals from four independent cohorts. We furthermore investigated longitudinal changes of metabolic measures and renal function and associations with other T2D microvascular complications. 142 traits correlated with glomerular filtration rate (eGFR) after adjusting for confounders and multiple testing: 59 in diabetics, 109 in non-diabetics with 26 overlapping. The amino acids glycine and phenylalanine and the energy metabolites citrate and glycerol were negatively associated with eGFR in all the cohorts, while alanine, valine and pyruvate depicted opposite association in diabetics (positive) and non-diabetics (negative). Moreover, in all cohorts, the triglyceride content of different lipoprotein subclasses showed a negative association with eGFR, while cholesterol, cholesterol esters (CE), and phospholipids in HDL were associated with better renal function. In contrast, phospholipids and CEs in LDL showed positive associations with eGFR only in T2D, while phospholipid content in HDL was positively associated with eGFR both cross-sectionally and longitudinally only in non-diabetics. In conclusion, we provide a wide list of kidney function–associated metabolic traits and identified novel metabolic differences between diabetic and non-diabetic kidney disease.
Chronic kidney disease (CKD) is a major public health problem affecting more than 10% of the population in Western countries\(^1\), leading to increased cardiovascular (CV) morbidity and mortality\(^2\). The renal microvascular complication of diabetes (DKD) is the leading cause of end-stage renal disease (ESRD). Despite the efforts in early diagnosis and therapeutic interventions in diabetes control, the rate of ESRD caused by DKD decreases less than the rates of all other diabetes complications\(^3\).

In recent years, several studies investigated metabolic profiles associated with renal function\(^4,5\) in the general population\(^6\) and in type 1 diabetic (T1D) patients\(^7–9\) to identify biomarkers for disease progression\(^8\) and mortality\(^10\). Other studies have looked for metabolic markers of type 2 diabetic (T2D) kidney disease, but sample sizes were small, they lacked independent replication\(^11\) or were performed in experimental animal models\(^12\).

Here, we used targeted nuclear magnetic resonance (NMR) spectroscopy to investigate metabolic signatures of renal function in T2D and non-diabetic individuals, combining four European cohorts. Additionally, to gain insights in potential mechanisms of the cross-sectional associations, we investigated longitudinal changes of metabolite levels and renal function and associations with other microvascular complications of T2D.

Results
Levels of 227 fasting serum metabolic traits including small molecules, lipids, lipoprotein subclasses, their lipids component and fatty acids (Supplementary Table 1), were obtained for 5764 individuals from four independent European cohorts, including 926 T2D patients (Fig. 1). The demographic characteristics of all cohorts are presented in Table 1. We calculated associations of all 227 metabolic traits with renal function in each cohort individually and meta-analyzed results for diabetic and non-diabetic cohorts (Supplementary Table 2). To assess the confounding effect of drug usage, we ran the same models in 1054 individuals from TwinsUK additionally adjusting for statin and hormone replacement therapy (HRT), and in 655 individuals from GenodiabMar adjusting for statin usage. Results remain consistent (Supplementary Table 3).

Markers of renal function common for diabetics and non-diabetics. After adjusting for age, gender, BMI and multiple testing, 26 metabolic traits where consistently associated with renal function across diabetic and non-diabetic cohorts (Table 2). The strongest cross-sectional associations with eGFR were observed for glycine and phenylalanine (P < 0.001) with association magnitudes of \(-8.37 [-9.73: -7.02]\) and \(-7.92 [-9.27: -6.57]\), respectively, for the diabetic group and \(-1.29 [-1.66: -0.92]\) and \(-1.69 [-2.07: -1.32]\) for the non-diabetic group (Fig. 2).

Levels of triglycerides in different sizes of intermediate- and low-density lipoprotein (IDL and LDL, respectively) particles were consistently inversely associated with the eGFR, while several high-density lipoprotein (HDL) subclasses of different sizes rich in lipids, cholesterol, cholesteryl esters, phospholipids and Apolipoprotein-A1 (Apo-A1) were consistently positively associated with eGFR (Fig. 3).

Citrate and glycerol were consistently negatively correlated with eGFR in all cohorts and also correlated with longitudinal change of eGFR (\(\beta [95\% CI] = -0.04 [-0.06: -0.02]\), \(p = 1.5 \times 10^{-6}\) and \(-0.02 [-0.04: -0.01]\),...
Table 1. General characteristics of the study populations. MZ = monozygotic, DZ = dizygotic. BMI = Body mass index. eGFR = estimated glomerular filtration rate (CKD-EPI equation). *Grades of Chronic Kidney Disease are stratified per KDIGO recommendation as: G2 60–89, G3 59–30, G4 15–29 and G5 < 15 mL/min/1.73 m². Values for categorical variables are given as n (percentage); values for continuous variable as mean ±SD.

|                        | Diabetic cohorts | Non-diabetic cohorts |
|------------------------|------------------|----------------------|
|                        | GenodiabMar | TwinsUK (diabetics) | Kora (diabetics) | TwinsUK | Kora (non-diabetic) | YoungFinns |
| n                      | 655          | 111                  | 160              | 1168     | 1624                | 2046       |
| Age (years)            | 69.70 ± 9.32 | 68.64 ± 8.38         | 66.74 ± 7.44     | 64.83 ± 7.91 | 60.30 ± 8.83      | 41.88 ± 5.00 |
| Gender                 | 256 (39.1%)  | 105 (94.6%)          | 71 (44.4%)       | 1118 (95.7%) | 845 (52.0%)        | 1115 (54.5%) |
| BMI (kg/m²)            | 30.32 ± 5.05 | 29.33 ± 5.55         | 31.48 ± 5.53     | 26.05 ± 4.61 | 27.82 ± 4.58      | 26.54 ± 5.05 |
| eGFR (mL/min/1.73 m²)  | 58.64 ± 28.83  | 75.80 ± 18.64        | 76.59 ± 18.15    | 79.87 ± 14.53 | 87.80 ± 15.38    | 94.75 ± 12.53 |
| creatinine (mg/dL)     | 1.26 ± 0.62   | 0.84 ± 0.24          | 0.96 ± 0.48      | 0.80 ± 0.16   | 0.83 ± 0.23       | 0.87 ± 0.15   |
| CKD* (grades)          | 116/213/202/72/52 | 20/76/12/3/0     | 39/95/24/1/1    | 32/22/76/118/20 | 795/752/73/2/2 | 1339/699/77/1/0 |
| N and (%)              | (17.7/32.5/30.8/10.9/7.9) | (18/68.4/10/8/2.7/0) | (24.3/59/3/15/0/0.6/0) | (27.5/62/1/10/1/0) | (48.9/66/4/4/0/1.1) | (65/4/4/1/0/3/0.04/0) |

p = 2.6 × 10\(^{-3}\) respectively, in TwinsUK and YoungFinns. However, their circulating levels at baseline did not predict eGFR at follow-up in GenodiabMar and diabetic-KORA cohorts (Supplementary Tables 4 and 5).

Metabolic profiles associated to renal function in diabetics. In the three diabetic cohorts, 59 metabolic measures were consistently associated with eGFR after meta-analysis at P < 0.001. Of those, 33 traits were associated with eGFR only in diabetics but not in non-diabetics (Supplementary Table 2). 6 of these traits were concentrations of cholesterol esters in LDL and IDL subclasses, 4 phospholipids in LDL and IDL, and 14 cholesterol and lipid concentrations in LDL that followed a positive association with eGFR in diabetics (Fig. 4 and Supplementary Fig. 1). Also, esterified cholesterol (EC) (β = 4.35 [2.96: 5.74], p = 9.3 × 10\(^{-10}\)) and total cholesterol (β = 3.68 [2.29: 5.07], p = 2.0 × 10\(^{-7}\)) were positively associated with eGFR in diabetics. However, none of this lipoprotein subclasses predicted the change of renal function and only triglycerides to total lipids ratio in large VLDL (β = 0.13 [0.01: 0.25]) and total cholesterol to total lipids ratio in small VLDL (β = −0.12 [−0.23: 0.00]) were associated to eGFR in the longitudinal analysis (Supplementary Table 5).

To further explore the relationship of metabolic profiles with other microvascular complications of diabetes, we calculated cross-sectional associations of metabolite profiles with proteinuria (independently of the eGFR), as well as cross-sectional odds-ratios for diabetic nephropathy (DN) and diabetic retinopathy (DR) in GenodiabMar. Glycine and phenylalanine were common risk factors not only for DN but also for both DR and proteinuria (Fig. 5 and Supplementary Table 6). Similarly, pyruvate, which was associated with better renal function, showed an inverse association with proteinuria. Glyceraldehyde, citrate, and pyruvate showed concordant albeit non-significant association with the retinal microvascular damage. Moreover, triglyceride contents in LDL, large and medium LDL, and small VLDL were consistently associated with decreased eGFR as well as higher risk of DN and DR, though to a lesser extent. In contrast, many other lipoprotein subclasses, including most HDLs, did not appear to be associated with DR (P > 0.05). As expected serum albumin was strongly associated with better renal function only in GenodiabMar cohort, due to a higher prevalence and a more severe diabetic nephropathy in this population.

Metabolic profiles associated to renal function in non-diabetics. In the three non-diabetic cohorts 109 metabolic measures were consistently associated with eGFR (P < 0.001) (Supplementary Table 2). 83 of these were associated with renal function only in non-diabetics but not in the diabetic group. Cholesterol and triglyceride levels in LDL particles of all sizes were negatively associated with eGFR. In contrast, phospholipid in large HDL were positively associated with renal function in non-diabetic populations cross-sectionally (phospholipids to total lipids ratio in very large HDL: β = 0.69 [0.32:1.06], p = 2.8 × 10\(^{-3}\)) and longitudinally (phospholipids in very large HDL: 0.04 [0.02:0.05], p = 1.3 × 10\(^{-3}\)) in TwinsUK (Supplementary Tables 2 and 4). Also, this lipid ratio predicted longitudinal change of eGFR independently of baseline eGFR in KORA (0.04 [0.01:0.07], p = 3.7 × 10\(^{-3}\)) (Supplementary Table 5).

Discordant metabolic measures between diabetics and non-diabetics. Four metabolites were positively associated with eGFR in diabetics, while they were negatively associated in the non-diabetic individuals at P < 0.001. However, the effect directions of these metabolites were not consistent throughout all cohorts. These include the amino acids alanine (T2D: β = 3.32 [1.92:4.72], non-T2D: −1.39 [−1.75: −1.03]), valine (T2D: 4.48 [3.08: 5.88], non-T2D: −0.88 [−1.27: −0.48]), the glycolysis product pyruvic acid (T2D: 3.37 [1.93:4.82], non-T2D: −0.70 [−1.06: −0.34]), and albumin (T2D: 2.55 [1.16: 3.93], non-T2D: −0.70 [−1.06: −0.35]) (Supplementary Table 2). Albumin also predicted the longitudinal change of eGFR (0.18 [0.06: 0.30] p = 3.8 × 10\(^{-3}\)) in the GenodiabMar and KORA cohorts (Supplementary Table 5). Also, there was a trend towards an increase of small, medium, and large LDL particles rich in phospholipids and cholesterol with better eGFR in T2D, that followed an opposite albeit non-significant association in non-diabetic populations (Fig. 4).

DOI:10.1038/s41598-018-33507-7
Table 2. Metabolic traits consistently associated with renal function in diabetic and non-diabetic cohorts. 26 metabolic traits from 14 metabolic classes, listed here, were associated with eGFR consistently across both diabetic and non-diabetic cohorts. Signs represent the directions of the regression coefficients in each diabetic (GenodiabMar, diabetics-TwinsUK, diabetics-KORA) and non-diabetic (TwinsUK, KORA, YoungFins) cohort. Results were meta-analyzed for diabetic and non-diabetic cohorts separately. Association magnitudes are eGFR per 1 SD (log-transformed) concentration. For detailed list of associations of all 227 analyzed metabolic traits see Supplementary T2.

Discussion

In the largest study of its kind, including 926 T2D diabetics and 4838 non-diabetics from four independent European cohorts, we identified 142 metabolic traits consistently associated with renal function at P < 0.001 and with concordant effects across cohorts: 59 in diabetics, 109 in non-diabetics, with an overlap of 26 traits. When comparing the effect directions, associations were largely concordant between diabetic and non-diabetic cohorts (R^2 = 0.60, Fig. 2). However, there were some notable exceptions. For instance, phospholipids and CE in IDL and LDL were positively correlated with eGFR only in diabetic individuals, while phospholipid content in HDL was positively associated with eGFR both cross-sectionally and longitudinally only in non-diabetics. We additionally identified four traits, valine, alanine, pyruvate, and albumin that were negatively associated with eGFR in non-diabetics, and positively associated in diabetics, though the positive associations were not consistent across all cohorts but driven by the GenodiabMar cohort that has a wider range of renal function impairment.

The metabolic measures identified fall into three categories: amino acids, energy-related metabolites, and lipoprotein subclasses particles and their lipids composition.
Amino acids. Phenylalanine, serves as precursor for tyrosine in the liver and kidneys. It has been previously associated with insulin resistance, increased risk of T2D and is a predictor of CV events. Moreover, reduced rates of conversion of phenylalanine to tyrosine were observed in CKD, leading to decreased circulating levels of tyrosine and increased levels of phenylalanine. While previous studies found the negative association of eGFR with tyrosine stronger than its positive associations with phenylalanine, we found tyrosine levels decreased only in diabetic patients (P < 0.001). Also, the log-fold change of phenylalanine over tyrosine correlated stronger
with eGFR in GenodiabMar than either individually ($\beta = 13.13$ [11.38:14.88], $p = 5.2 \times 10^{-42}$). While increased concentration of phenylalanine was associated with worse renal function in both diabetics and non-diabetics, it was not predictive for disease progression in this study. This suggests that phenylalanine might not be a predictor of renal decline but rather a consequence of renal dysfunction and a marker of vascular damage\(^\text{18}\). Indeed, phenylalanine was also associated at $p < 0.05$ with DR and albuminuria suggesting an association with endothelial microvascular damage.

Similarly, glycine is converted to serine in the kidneys\(^\text{22,23}\). Thus, impairment of renal function leads to accumulation of glycine, which was consistently observed in both diabetic and non-diabetics. Glycine also correlated with albuminuria and DR at $p < 0.001$ but did not predict longitudinal change of renal function. Previous studies report glycine to be negatively associated with CV risk factors and T2D\(^\text{17}\), but renal function was not included as a covariate. Also, eGFR usually increases in early stages of diabetes, before renal function declines. Thus, the observed associations with risk for T2D might be confounded by renal function. Our results highlight the importance of including renal function as cofactor when studying diabetes, and are in line with T2D experimental animal model studies, that revealed a lower urine excretion of glycine and accumulation of this metabolite in diabetic kidney tissues\(^\text{12}\).

### Energy-related metabolites.

Alanine is a major precursor of hepatic and renal gluconeogenesis and glycolysis via pyruvate pathways. Together with glycerol, which was also negatively associated with eGFR, and glutamine, they constitute 90% of the substrates of gluconeogenesis. Metabolic acidosis induced by CKD leads to increased abundance of circulating alanine, glutamine, and glutamate\(^\text{23}\). However, in the diabetic milieu, glucose metabolism is heavily disturbed with an increased rate of gluconeogenesis\(^\text{24}\). Consequently, the decline of renal function has a different impact on gluconeogenesis in diabetics, as evidenced by the different directions of associations for alanine and pyruvate in this study.

---

**Figure 4.** Lipoprotein classes associated with eGFR in diabetic and non-diabetic cohorts. Associations of lipoprotein subclasses with eGFR were calculated in three type 2 diabetic (T2D) and three non-diabetic (non-T2D) cohorts and results were meta-analyzed (black). Here we report regression coefficients and their respective 95% confidence interval per 1-SD (log-transformed) metabolite concentration for each cohort and the meta-analyses. For detailed list of results, including heterogeneity of effect estimates, and full metabolites names see Supplementary Table 2.
Citrate is an important metabolic substrate in the kidney accounting for up to 10% of the energy production that counteract metabolic acidosis. In agreement with our findings, different studies reported increased concentration with the decline of eGFR.

Lipoprotein subclasses and their lipids component. Lipoproteins abnormalities are not always detected in CKD subjects when using standard clinical measures. Particularly total and LDL cholesterol are usually normal and even low in advanced CKD. The CKD-induced lipid profile has specific characteristics distinct from the general population. Besides quantitative changes, renal patients have several qualitative lipid alterations that cannot be detected by routine determinations and some alterations of the lipoprotein composition and size may contribute to the CV complications observed in CKD patients. Interestingly, in the present study non-classical lipid profiles showed association with renal function and remained associated after adjustment for statin usage (Supplementary Table 2). Some epidemiological studies revealed controversial results regarding lipid-lowering therapy and reduction of cardiovascular mortality in CKD and emphasize the need for further studies such as the present analysis. In our study, the lipid content in the different lipoprotein particles showed considerable differences, which highlights the potential importance of performing a more detailed lipidomic analysis that may reveal different risk patterns that would otherwise be missed.

Some of the largest differences found with renal function between diabetics and non-diabetics were the negative associations of small to large VLDL and LDL subclasses and their respective cholesterol and triglyceride content observed only in non-diabetics, as well as the positive associations of small to large IDL and LDL subclasses and their cholesterol, EC and phospholipid content observed only in diabetics.

The positive association of the pro-atherogenic LDL and IDL with eGFR observed in this study, are likely not reflecting a positive effect of these lipoproteins on renal function but rather a better nutritional status in subjects with better renal function. Higher prevalence of individuals with worse renal function in the T2D cohorts is the likely cause of these counter-intuitive associations. Of note, the phospholipid and CE content of these lipoproteins were not significantly different between diabetics and non-diabetics.
lipoprotein particles may be related to increased lipid transfer proteins (LTP) activity (CETP and PLTP) present in diabetic subjects. On the contrary, lower activity of LTPs associates with lower CV risk, which might be related to the negative associations of LDL subclasses with renal function in non-diabetic subjects. Interestingly, triglyceride ratios in LDL and IDL were negatively associated with renal function consistently between diabetics and non-diabetics. Also, triglyceride to total lipid ratios showed stronger association in diabetes compared to non-diabetes (Supplementary Table 2). To the best of our knowledge, no studies have investigated the activity of LTP and their association with renal damage and whether pharmacological targeting of this proteins might influence in renal function.

Longitudinal analysis revealed a positive association of several HDL particle and Apo-A1 with renal function over time (Supplementary Table 4). However, their circulating levels at baseline were not associated with a better renal function in T2D at follow-up. Diabetic dyslipidemia presents particularities regarding quantitative lipoprotein abnormalities and also qualitative and kinetic abnormalities that results in a more atherogenic lipid profile. Changes in HDL composition in T2D have been shown to affect cholesterol efflux. Moreover, higher proteinuria may increases the loss of particles derived from HDL catabolism. In our study, the ratio of phospholipids to total lipids in very-large HDL was associated with longitudinal change of eGFR and predicted future eGFR only in non-diabetics. Phospholipids in HDL enhance its cholesterol efflux capacity, which is impaired in diabetics and may explain the observed differences.

Although many of our findings are shared with previous studies on T1D, we found some differences. For example, we did not find any association of eGFR with sphingomyelin or total fatty acids that were markers of kidney injury and mortality in T1D. This may suggest differences between T1D and T2D metabolic profiles and the importance of analyze both conditions individually.

The present study has several strengths. First, we analyzed data from four independent cohorts, thus minimizing the risk of false positive findings. Second, we analyzed a wide range of metabolic traits beyond those commonly used in clinics. Also, we stratified for diabetes status, thus providing a direct comparison of metabolic profiles associated with diabetic and non-diabetic renal damage. We also note some study limitations. The GenodiabMar cohort was recruited from medical consultations while the other cohorts represent individuals from the general population. Thus, the presence of other medical complications as well as different grades of renal dysfunction may be confounding factors. However, by meta-analyzing results across diabetic cohorts, we controlled for population-specific effects. Also, drug use may have an important impact on metabolic profiles, although statin use did not substantially change the results in this study (Supplementary Table 3). However, further analyses specifically addressing the effects of different drugs, such as antihypertensive, other medical conditions and the potential effect of renal replacement therapies, are needed.

In conclusion, we found widespread metabolic changes associated with decline of renal function. While associations of many lipoprotein particles and their lipid composition with renal function were largely similar between diabetic and non-diabetic cohorts, several exceptions revealed metabolic differences between the conditions. Also, changes of amino acid and energy metabolism were markedly different regarding diabetes condition. Our results show alterations of lipoprotein composition in kidney disease that are currently underexploited in clinics. We also find marked metabolic differences between diabetic and non-diabetic kidney disease, suggesting that more specific markers for each condition might be able to outperform current markers of kidney disease.

**Methods**

**Study Design and Participants.** Targeted NMR metabolic profiling was conducted in 926 diabetic and 4838 non-diabetic individuals from the GenodiabMar (n = 655), TwinsUK (n = 1279, 111 with T2D), KORA (n = 1784, 160 with T2D), and Young Finns (n = 2046) cohorts. GenodiabMar is a cohort of T2D patients, recruited in a hospital, while the other cohorts were recruited from the general population. Renal function was measured as eGFR from standard creatinine using the Chronic Kidney Disease Epidemiology Collaboration equation (CKD-EPI). Longitudinal measures were available for a subset of 3644 individuals (Supplementary methods). Each local ethics committee approved the study, and subjects were included after providing informed consent. All methods were performed in accordance with the relevant guidelines and regulations.

A flowchart of the study design is depicted in Fig. 1.

**Metabolic profiling.** Metabolic profiling of 227 metabolic traits, 143 metabolite concentrations, 80 lipid ratios, 3 lipoprotein particle sizes and a semi-quantitative measure of albumin (see Supplementary Table 1 for full list), was conducted for all cohorts by Nightingale Health Ltd. (Helsinki, Finland; previously known as Brainshake Ltd) using a targeted NMR spectroscopy platform that has been extensively applied for biomarker profiling in epidemiological studies as previously described. Longitudinal measures were available for a subset of 3644 individuals (Supplementary methods).

**Statistical analysis.** All metabolic measures were log-transformed. To account for zero values a pseudo-count of 1 was added to all measurements prior to transformation. All measurements were shifted to zero mean and scaled to standard deviation (SD) of 1 (z scores) to facilitate comparisons across cohorts. The average absolute concentrations and SDs of each metabolite in each cohort are presented in Supplementary Table 1.

**Cross-sectional analysis.** We assessed the associations between metabolic profiles and renal function in each cohort individually by fitting linear regressions for all metabolic traits with eGFR as outcome, adjusting for age, gender, BMI (and family relatedness as random intercept) to account for the decline of renal function with advancing age as well as its dependency on obesity. All results were then meta-analyzed separately for T2D patients and non-diabetic cohorts using inverse variance fixed effect meta-analysis due to the expected homogeneity of effects in both subgroups. We adjusted for multiple testing using Bonferroni correction assuming 50 independent test as suggested by Li and Ji (P < 0.001) (Supplementary methods).
As metabolite profiles may be strongly affected by medication such as statin treatment or hormone replacement therapy, we tested the robustness of our results by running a sub-analysis in a subset of individuals from TwinsUK and GenodiabMar with information on treatment available, additionally adjusting for medication status.

To further investigate traits of interest, we regressed the concentration of albumin in urine against each of the metabolic traits. Finally, we calculated logistic regression models to assess the association of each metabolic trait with diabetic nephropathy and retinopathy, respectively.

**Longitudinal analysis.** For TwinsUK and YoungFinns we estimated the trajectories of metabolite/eGFR change by fitting linear mixed models for each metabolite with a per-individual random effect for the time since baseline. The estimate of this random effect provides a measure of the (linear) change of metabolite concentration over time similarly to calculating the change per year for two visits. These trajectories were estimated for all metabolites individually and then compared to the change in eGFR in a separate regression model, thus assessing longitudinal correlations between metabolites and renal function.

Also, we evaluated the potential of metabolite measures as diagnostic tool by predicting the eGFR at follow-up using metabolite measures at baseline, correcting for gender and baseline eGFR, age, and BMI.

**Data Availability**
Data from the TwinsUK cohort are available upon request on the website (http://www.twinsuk.ac.uk/data-access/accessmanagement/). Data from the KORA cohort can be requested online (https://epi.helmholtz-muenchen.de/) and is subject to approval by the KORA board.

**References**
1. O'Callaghan, C., Shine, B. & Lasserson, D. S. Chronic kidney disease: a large-scale population-based study of the effects of introducing the CKD-EPI formula for eGFR reporting. *BMJ Open* 1, e000308–e000308 (2011).
2. Go, A. S., Chertow, G. M., Fan, D., McCulloch, C. E. & Hsu, C. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N. Engl. J. Med.* 351, 1296–305 (2004).
3. Gregg, E. W. et al. Changes in diabetes-related complications in the United States, 1990–2010. *N Engl J Med* 370, 1514–1523 (2014).
4. Zhang, Y., Zhang, S. & Wang, G. Metabolomic biomarkers in diabetic kidney diseases–A systematic review. *J. Diabetes Complications* 29, 1345–51 (2015).
5. Barrios, C., Spector, T. D. & Menni, C. Blood, urine and faecal metabolite profiles in the study of adult renal disease. *Arch. Biochem. Biophys.* 589, 81–92 (2016).
6. Sekula, P. et al. A Metabolome-Wide Association Study of Kidney Function and Disease in the General Population. *J. Am. Soc. Nephrol.* 27, 1175–88 (2016).
7. Mäkinen, V.-P. et al. Metabolic phenotyping of diabetic nephropathy. *Clin. Pharmacol. Ther.* 94, 566–9 (2013).
8. Mäkinen, V.-P. et al. Metabolic diversity of progressive kidney disease in 325 patients with type 1 diabetes (the FinnDiene Study). *J. Proteome Res.* 11, 1782–90 (2012).
9. Soininen, P. et al. 1H NMR metabonomics approach to the disease continuum of diabetic complications and premature death. *Mol. Syst. Biol.* 4, 167 (2008).
10. Mäkinen, V.-P. et al. Triglyceride-cholesterol imbalance across lipoprotein subclasses predicts diabetic kidney disease and mortality in type 1 diabetes: the FinnDiene Study. *J. Intern. Med.* 273, 383–95 (2013).
11. Darshi, M., Van Espen, B. & Sharma, K. Metabolomics in Diabetic Kidney Disease: Unraveling the Biochemistry of a Silent Killer. *Am. J. Nephrol.* 44, 92–103 (2016).
12. Wei, T. et al. Metabolomic analysis of potential biomarkers and drug targets involved in diabetic nephropathy mice. *Sci. Rep.* 5, 11998 (2015).
13. Möller, N., Meeke, S., Bigelow, M., Andrews, J. & Nair, K. S. The kidney is an important site for in vivo phenylalanine-to-tyrosine conversion in adult humans: A metabolic role of the kidney. *Proc. Natl. Acad. Sci. USA* 97, 1242–6 (2000).
14. Wang, T. I. et al. Metabolite profiles and the risk of developing diabetes. *Nat. Med.* 17, 448–53 (2011).
15. Floegel, A. et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolic approach. *Diabetes* 62, 639–648 (2013).
16. Stancakova, A. et al. Hyperglycaemia and a Common Variant of GCKR Are Associated With the Levels of Eight Amino Acids in 9,369 Finnish Men. *Diabetes* 61, 1895–1902 (2012).
17. Guasch-Ferré, M. et al. Metabolomics in Prediabetes and Diabetes: A Systematic Review and Meta-analysis. *Diabetes Care* 39, 833–846 (2016).
18. Wurtz, P. et al. Metabolite Profiling and Cardiovascular Event Risk: A Prospective Study of 3 Population-Based Cohorts. *Circulation* 131, 774–785 (2015).
19. Young, G. A. & Parsons, F. M. Impairment of phenylalanine hydroxylation in chronic renal insufficiency. *Clin. Sci.* 45, 89–97 (1973).
20. Boirie, Y., Albright, R., Bigelow, M. & Nair, K. S. Impairment of phenylalanine conversion to tyrosine in end-stage renal disease causing tyrosine deficiency. *Kidney Int.* 64, 591–596 (2004).
21. Kopple, J. D. Phenylalanine and tyrosine metabolism in chronic kidney failure. *J. Nutr.* 137 1586S–1590S; discussion 1597S–1598S (2007).
22. Pitts, R. F., Damian, A. C. & MacLeod, M. B. Synthesis of serine by rat kidney in vivo and in vitro. *Am. J. Physiol.* 219, 584–9 (1970).
23. Truzianello, A., De Ferrari, G., Garibotto, G., Gurrieri, G. & Robaudo, C. Renal metabolism of amino acids and ammonia in subjects with normal renal function and in patients with chronic renal insufficiency. *J. Clin. Invest.* 65, 1162–73 (1980).
24. Magnusson, I., Rothman, D. L., Katz, L. D., Shulman, R. G. & Shulman, G. I. Increased rate of gluconeogenesis in type II diabetes mellitus. A 13C nuclear magnetic resonance study. *J. Clin. Invest.* 90, 1323–7 (1992).
25. Brennan, S., Herin-Smith, K. & Hamm, L. I. Effect of pH on citrate reabsorption in the proximal convoluted tubule. *Am. J. Physiol.* 255, F301–6 (1988).
26. Kim, J.-A. et al. 1H NMR-based metabolite profiling of plasma in a rat model of chronic kidney disease. *PLoS One* 9, e85445 (2014).
27. Sharma, K. et al. Metabolomics reveals signature of mitochondrial dysfunction in diabetic kidney disease. *J. Am. Soc. Nephrol.* 24, 1901–12 (2013).
28. Vergès, B. Pathophysiology of diabetic dyslipidaemia: where are we? *Diabetologia* 58, 886–99 (2015).
29. Vergès, B. New insight into the pathophysiology of lipid abnormalities in type 2 diabetes. *Diabetes Metab.* 31, 429–39 (2005).
30. Baigent, C., Landray, M. J. & Wheeler, D. C. Misleading associations between cholesterol and vascular outcomes in dialysis patients: the need for randomized trials. *Semin. Dial.* 20, 498–503 (2007).
31. Baigent, C. & Landray, M. Which cardiovascular risk factors matter in chronic kidney disease? *Nephrol. Dial. Transplant* 22, 9–11 (2007).
Author Contributions

C.M., T.D.S., J.P., G.K. conceived and supervised the study; P.W., A.M., C.G., B.T., C.M., M.W., O.R., T.L., S.O., E.R., J.P.B., M.K., M.A.K. collected and measured data; C.B., J.Z., T.H., analyzed data; C.B., J.Z., C.M. wrote the manuscript; All authors reviewed the manuscript.

Acknowledgements

TwinUK was funded by the Wellcome Trust; European Community’s Seventh Framework Programme (FP7/2007–2013). The study also receives support from the National Institute for Health Research (NIHR) Clinical Research Facility at Guy’s & St Thomas’ NHS Foundation Trust and NIHR Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. TS is holder of an MRC Centre based at Guy’s and St Thomas’ NHS Foundation Trust and NIHR Biomedical Research Centre at King’s College London. TS is the founder of the Academy of Finland (312476 & 312477), and the Novo Nordisk Foundation (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (project numbers 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; and Diabetes Research Foundation of Finnish Diabetes Association. MAK was supported by the Sigrid Juselius Foundation, Finland. MAK works in a Unit that is supported by the University of Bristol and UK Medical Research Council (MC_UU_12013/1). PW is funded by the Academy of Finland (312476 & 312477), and the Nordoisk Foundation Competing Interests.
Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-33507-7.

Competing Interests: PW is shareholder and employee of Nightingale Health Ltd, a company offering NMR-based metabolic profiling.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018