Serum Orotidine: A Novel Biomarker of Increased CVD Risk in Type 2 Diabetes Discovered Through Metabolomics Studies

Hetal S. Shah,[1,2] Lorena Ortega Moreno,[3,4] Mario Luca Morieri,[5] Yaling Tang,[1,2] Christine Mendonca,[1] Jenny Marie Jobe,[1] Jonathan B. Thacker,[6] Joanna Mitri,[2] Stefano Monti,[7] Monika A. Niewczas,[1,2] Subramaniam Pennathur,[6] and Alessandro Doria[1,2]

Diabetes Care 2022;45:1882–1892 | https://doi.org/10.2337/dc21-1789

OBJECTIVE

To identify novel biomarkers of cardiovascular disease (CVD) risk in type 2 diabetes (T2D) via a hypothesis-free global metabolomics study, while taking into account renal function, an important confounder often overlooked in previous metabolomics studies of CVD.

RESEARCH DESIGN AND METHODS

We conducted a global serum metabolomics analysis using the Metabolon platform in a discovery set from the Joslin Kidney Study having a nested case-control design comprising 409 individuals with T2D. Logistic regression was applied to evaluate the association between incident CVD events and each of the 671 metabolites detected by the Metabolon platform, before and after adjustment for renal function and other CVD risk factors. Significant metabolites were followed up with absolute quantification assays in a validation set from the Joslin Heart Study including 599 individuals with T2D with and without clinical evidence of significant coronary heart disease (CHD).

RESULTS

In the discovery set, serum orotidine and 2-piperidinone were significantly associated with increased odds of incident CVD after adjustment for glomerular filtration rate (GFR) (odds ratio [OR] per SD increment 1.94 [95% CI 1.39–2.72], $P = 0.0001$, and 1.62 [1.26–2.08], $P = 0.0001$, respectively). Orotidine was also associated with increased odds of CHD in the validation set (OR 1.39 [1.11–1.75]), while 2-piperidinone did not replicate. Furthermore, orotidine, being inversely associated with GFR, mediated 60% of the effects of declining renal function on CVD risk. Addition of orotidine to established clinical predictors improved ($P < 0.05$) C statistics and discrimination indices for CVD risk (ΔAUC 0.053, rIDI 0.48, NRI 0.42) compared with the clinical predictors alone.

CONCLUSIONS

Through a robust metabolomics approach, with independent validation, we have discovered serum orotidine as a novel biomarker of increased odds of CVD in T2D, independent of renal function. Additionally, orotidine may be a biological mediator of the increased CVD risk associated with poor kidney function and may help improve CVD risk prediction in T2D.

Individuals with type 2 diabetes (T2D) have two- to four-fold higher risk of CVD than the general population. Such increased cardiovascular (CV) risk is a major

[1]Section on Genetics and Epidemiology, Research Division, Joslin Diabetes Center, Boston, MA
[2]Department of Medicine, Harvard Medical School, Boston, MA
[3]Department of Basic Health Sciences, Universidad Rey Juan Carlos, Alcorcón, Spain
[4]High Performance Research Group in Physiopathology and Pharmacology of the Digestive System (NeuGut), Universidad Rey Juan Carlos, Alcorcón, Spain
[5]Department of Medicine, University of Padova, Padova, Italy
[6]Division of Nephrology, Department of Internal Medicine, University of Michigan, Ann Arbor, MI
[7]Computational Biomedicine, Department of Medicine, Boston University, Boston, MA

Corresponding authors: Hetal S. Shah, hetal.shah@joslin.harvard.edu, and Alessandro Doria, alessandro.doria@joslin.harvard.edu

Received 24 August 2021 and accepted 26 April 2022

This article contains supplementary material online at https://doi.org/10.2337/figshare.19747399.

contributor to the high morbidity, mortality, and socioeconomic burdens of T2D (1–3). Currently known risk factors such as age, sex, BMI, HbA_{1c}, diabetes duration, blood pressure, and lipids do not fully explain the mechanism of increased CV risk in T2D (4). Global metabolomics studies, which examine molecules that are more proximal to the disease phenotypes than those investigated by other -omics, may potentially enrich our systemic understanding of the underlying etiopathogenesis of CVD in patients with T2D and guide clinical interventions for better prevention and management of this complication of diabetes (5).

Several studies have investigated the association between metabolomics profiles and CV outcomes in the general population. However, few such studies have been conducted in people with T2D, in whom the glycemic milieu could present with further altered metabolic profiles. Furthermore, for the metabolites thus far identified to be associated with CV in the general population, including acetylcarnitines (6,7), amino acids (8,9), monosaccharides (10), biogenic amines (8), glycerophospholipids (10), and sphingolipids (11), few studies have explored how these effects may be confounded by, or the expression of, reduced glomerular filtration rate (GFR)—a well-known powerful CV risk factor (12,13) increasingly shown to have a potentially large impact on the small molecules comprising the metabolome, even within normal ranges (14). T2D-specific studies (15,16) have also overlooked this essential CV risk factor. Additionally, many of the studies conducted to date have been limited by small sample sizes; cross-sectional study designs; targeted or narrow-coverage metabolomics panels; older, less sensitive and specific assay technologies; and lack of validation or replication of findings.

We aimed to identify novel metabolic pathways of CVD risk in T2D, taking into account their relationship with kidney function, through global serum metabolomics in a well-powered discovery set consisting of a nested case-control study of individuals with T2D with and without incident CVD events selected from a cohort with sufficient follow-up. Top findings in the discovery set were evaluated in a validation set using targeted absolute quantification assays.

RESEARCH DESIGN AND METHODS

Study Populations

Discovery Set: The Joslin Kidney Study

The Joslin Kidney Study (JKS) is a longitudinal observational study of the natural history of declining renal function in type 1 diabetes and T2D. Participants in the T2D portion of the JKS (n = 1,476) were recruited between 2003 and 2009 from among patients attending the Joslin Clinic aged 35–64 years old, with T2D diagnosed after 30 years of age according to standard clinical criteria (17,18). By design, 50% of participants recruited into the study had normoalbuminuria and 50% had albuminuria. Study protocol and informed consent procedures were approved by the Joslin Diabetes Center Committee on Human Studies (JDC-CHS).

Outcome Assessment and Study Design. Between 2012 and 2013, 1,178 participants were interviewed for assessment of CV history at enrollment and nonfatal CV events that occurred during follow-up, which were verified through careful review of each individual's medical record. Self-reported CV outcomes have previously been shown to have strong agreement with actual events reported in medical data (19–21). Additionally, the National Death Index database (22) was queried to identify individuals who had died as of the end of 2015. CV deaths were defined on the basis of ICD-9 codes 401–448.9 or ICD-10 codes I10–I74.9 or if CVD was listed as the secondary cause of death and diabetes or renal failure listed as the primary cause. Those participants who had baseline Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) serum creatinine–derived estimated GFR (SCre-eGFR) values between 30 and 120 mL/min/1.73 m^2 served as the discovery set for the current nested case-control study (n = 409). Case subjects (n = 115) included participants who experienced fatal or nonfatal CV events during follow-up, including CV death, myocardial infarction (MI), percutaneous transluminal coronary angioplasty (PTCA), coronary artery bypass grafting (CABG), or stroke. Control subjects (n = 294) were a random sample of those participants who did not experience CV events during follow-up, group matched to case subjects by sex, baseline SCre-eGFR ($>$60 vs. $\leq$60 mL/min/1.73 m^2), length of blood sample storage ($<$10 vs. $\geq$10 years), and HbA_{1c} ($<$8 vs. $\geq$8%). Baseline CVD was defined based on history of any nonfatal CV events (MI, PTCA, CABG, or stroke) before study entry.

Validation Set: The Joslin Heart Study

The Joslin Heart Study (JHS) comprises non-Hispanic Whites who attended the Joslin Clinic and/or the Beth Israel Deaconess Medical Center (BIDMC) between 2001 and 2014. All participants had T2D, defined as diabetes diagnosed after age 30 years according to ADA criteria and not requiring insulin treatment for at least 2 years after diagnosis (23,24). The JDC-CHS and the BIDMC Committee on Clinical Investigations approved the study protocol, and all subjects gave written informed consent. Detailed inclusion criteria and the clinical characteristics of participants recruited up to 2006 have previously been described (23,24).

Outcomes Assessment and Study Design. The JHS was a cross-sectional, unmatched case-control study. Case subjects were a random sample of individuals with T2D and coronary heart disease, defined as angiographically documented stenosis $>$50% in a major coronary artery that was documented by cardiac catheterization at BIDMC between 2001 and 2014. Control subjects were randomly selected from among Joslin patients between 2001 and 2014 as fulfilling the following additional criteria: *1*) age between 55 and 74 years, *2*) T2D for $\geq$5 years, *3*) negative CVD history (i.e., normal resting electrocardiogram, absence of cardiac symptoms, and no hospitalization for CV events), and *4*) normal response to an exercise treadmill screening test performed within 6 months prior to study enrollment. A total of 740 case and 766 control subjects are currently included in the JHS; among these, 300 CHD case and 299 control subjects were randomly selected as the validation set for the current study.

Biochemical Analyses

Serum and urine samples were obtained at baseline during the participants' clinic visits, without the requirement of being in a fasting state, and were subsequently stored at -80°C.

Assessment of Renal Function

Serum Creatinine and Cystatin C. For the discovery set, baseline serum creatinine concentrations were measured at the Advanced Research and Diagnostic Laboratory,

University of Minnesota, MN, with Roche assays, and for the validation set, they were assayed at the Joslin Diabetes Center clinical laboratory with use of modified Jaffe picrate methods (Supplementary Material). Baseline serum cystatin C concentrations were measured in the discovery and validation sets at the Advanced Research and Diagnostic Laboratory in 2017–2018 and 2019–2020, respectively, with Gentian cystatin C immunoassay (Gentian AS, Moss, Norway) on the Roche analyzer (Roche Diagnostics, Indianapolis, IN) (25). The CKD-EPI formulas (26) were used to estimate SCre-eGFR and serum cystatin C–derived estimated GFR (SCys-eGFR) for both study populations.

accuGFR. In a randomly selected subset of 225 individuals (112 control and 113 case subjects) of the discovery set, accuGFR (Metabolon, Durham, NC) was estimated from an equation based on a panel of four novel metabolites, including pseudouridine, acetylthreonine, phenylacethylglutamine, and tryptophan, which were quantitatively measured in baseline serum samples with ultrahigh-performance liquid chromatography–tandem mass spectrometry methods as previously described (27). This accuGFR algorithm does not include creatinine, cystatin C, or demographic variables and has been shown to have a great degree of accuracy in assessing glomerular filtration as demonstrated by a high concordance with measured GFR (27).

Urinary Albumin-to-Creatinine Ratio. In both the discovery and validation sets, baseline albumin-to-creatinine ratios (ACRs) were assessed in urine as previously described (28,29). Albumin concentrations were measured via immunonephelometry at the Joslin clinical laboratory, and creatinine was measured with the modified Jaffe picrate method (Supplementary Material). Normoalbuminuria (ACR$<$30 mg/g), microalbuminuria (ACR 30–300 mg/g), and macroalbuminuria (ACR $>$300 mg/g) were determined with the geometric mean ACR for the preceding 2-year interval of the baseline exam for each patient.

Global Metabolomics Profiling

Global biochemical profiles were determined by Metabolon in 409 human serum samples of the discovery set with use of their proprietary platform, which uses gas and liquid chromatography–mass spectrometry in positive and negative modes (30). Briefly, samples were maintained at -80°C until processed and then run with quality control samples as per Metabolon standards. Instrument (4%) and process (8%) variability met Metabolon's acceptance criteria. Raw data extraction, peak identification, and quality control processing then took place. A total of 903 metabolic compounds were identified. Values were normalized in terms of raw area counts and then volume normalized. Each biochemical was then rescaled to set the median equal to 1, and missing values were imputed with the minimum.

Targeted Assays

For validation of methods and replication of findings from the global metabolomics study of the discovery set (JKS), targeted quantitative assays for the top metabolites, orotidine and 2-piperidinone, were developed at the Molecular Phenotyping Core at the University of Michigan. Serum samples from the validation set (JHS) (n = 599) as well as 20 replicates from the discovery set (JKS) were assayed for these two metabolites via UHPLC-MS/MS experiments performed on an Agilent system (Agilent Technologies, Santa Clara, CA) (Supplementary Material).

Statistical Methods

All statistical analyses were performed in SAS, version 9.4 (SAS Institute, Cary, NC), or R, version 3.6.1 (R Foundation, Vienna, Austria). Clinical characteristics of case and control subjects in both the discovery (n = 409) and validation (n = 599) sets are presented with means ± SD or median (interquartile range) for continuous variables and frequencies (%) for categorical variables, and differences were tested with two-sample independent t tests or χ^2 tests as appropriate. Nonparametric variables were assessed with Kruskal-Wallis tests. Correlations between accuGFR and SCre-eGFR/SCys-eGFR in case and control subjects were examined by generalized linear regression models.

Global Metabolomics Analysis

In the discovery set (n = 409), 903 serum metabolites were measured in the Metabolon platform panel. For the 232 of these metabolites with $>$20% sample missingness, differences in nondetectability rates between case and control subjects were examined with χ^2 tests. The remaining 671 that were detectable in at least 80% of the samples were used for subsequent metabolomics analyses. Volume-normalized median-scaled metabolites were transformed into ranked z scores. In a global metabolomics analysis, these standardized metabolites were independently evaluated as predictors of CVD outcomes in multivariable conditional logistic regression models, initially only accounting for the matching strata (sex, baseline SCre-eGFR ($>$60 vs $\leq$60 mL/min/1.73 m^2), length of blood sample storage ($<$10 vs. $\geq$10 years), and HbA_{1c} ($<$8 vs. $\geq$8%)) only (model 1) and then with further adjustments by baseline SCre-eGFR (model 2), followed by repeating the analysis with inclusion of baseline SCys-eGFR instead of SCre-eGFR as a confounder (model 3). The association of each metabolite with CVD was expressed as the odds ratio (OR) per 1 SD difference. Benjamin-Hochberg false discovery rate (31) corrections for multiple testing were applied and q values $<$ 0.05 considered significant.

Metabolites with q values $<$0.05 in the SCys-eGFR–adjusted analysis were further tested in models with adjustment for metabolic parameters (BMI, triglycerides, HDL cholesterol [HDL-C], HbA_{1c}, and ACR), medication use (diuretics, β-blockers, calcium channel antagonists, renin-angiotensin system blockers [RASBs], statins, nitrates), history of prior CVD, race, and smoking status (model 4). For model 4, a Bonferroni cutoff of $P < 0.01$ was used to determine significance, accounting for the four analytes for which the full adjustments were applied. While q value cutoff points were used for metabolite selection from models 1–3, and Bonferroni corrections used for model 4, statistical significance in other analyses was interpreted according to a common cutoff point of 0.05 for P values.

A flowchart for the metabolomics analysis is provided in Supplementary Fig. 1.

Power

Given our sample size of 409 individuals, with 28% CVD events, we had 80% power to detect ORs of 1.39, 1.64, and 1.76 for the effects on CVD per SD increase in metabolite, at α-errors of 0.05, 0.001, and 0.0001, respectively.

Validation Analyses of Orotidine and 2-Piperidinone

Orotidine and 2-piperidinone levels quantified via targeted assays (developed at the University of Michigan) were transformed into ranked *z* scores, and these standardized metabolites were used for further analyses. In the discovery set replicates (n = 20), correlations of measurements obtained from these assays were examined with their counterparts in the Metabolon platform panel. In the validation set (n = 599), these assayed metabolites were examined for associations with CHD using multivariable logistic regression models with the analytical approach described above.

Mediation Analysis

A mediation analysis was performed to evaluate the potential of orotidine as a mediator of the effects of low GFR on increased odds of CVD in the discovery set or CHD in the validation set. The first step in this analysis was to obtain the effect estimate (observed β [βobs]) from the conditional logistic regression model testing baseline SCys-eGFR as a predictor of incident CVD events in the discovery set, accounting for matching strata (sex, sample storage time, SCre-eGFR, HbA_{1c}) and with adjustment for the full model covariates including metabolic markers (BMI, lipids, ACR), race, smoking status, and medications. The second step was to obtain the effect estimate (β1) of SCys-eGFR as a predictor of orotidine in a linear regression model with adjustment for sex, storage time, full model covariates, and case-control status. In the third step, the effect (β2) of orotidine on CVD (or CHD) was obtained from a conditional logistic regression model, accounting for strata derived from group matching as mentioned above, as well as with adjustment for full model covariates and SCys-eGFR. The expected β (βexp) was derived from the product of β1 and β2. Finally, % mediation was calculated by dividing the expected β by the observed (βexp / βobs).

Prediction of Incident CVD

Improvement in risk prediction of incident CVD in the discovery set was assessed by logistic regression models with methods developed by Kennedy and Pencina (32). The performance of the following three models were compared: model A, including orotidine alone; model B, of standard clinical predictors of CVD, including the American College of Cardiology/American Heart Association atherosclerotic CVD (ASCVD) risk estimator, prior history of CVD, BMI, and SCys-eGFR; and model C, comprising orotidine plus clinical predictors in model B. By plotting of sensitivity against (1 − specificity) for all possible thresholds, receiver operating characteristic (ROC) curves were formed for the three models, and area under the ROC curve (AUC), or C statistic, was computed. The relative integrated discrimination improvement (rIDI) (33) and the category-free net reclassification index (NRI) (33) were also estimated.

RESULTS

Clinical Characteristics of the Discovery Cohort

We conducted a nested case-control study from a cohort of Joslin Diabetes Center patients with T2D who were followed for a median of 7.4 years (range 4.9–9.0). Included were 115 incident case subjects with a major CV event during follow-up and 294 control subjects without such an event. Clinical characteristics at baseline are shown in Table 1. Mean ± SD age was 57.7 ± 5.4 years for case and 58.0 ± 5.6 years for control subjects, with durations of diabetes of 15.2 ± 8.2 and 14.4 ± 7.8 years, respectively. Women comprised 30% of case and 35% of control subjects. Case subjects had significantly higher BMI and triglycerides and lower HDL-C and more frequently had a history of CVD compared with control subjects, whereas HbA_{1c}, systolic and diastolic blood pressures, and smoking history were similar in the two groups. SCre-eGFR did not significantly differ between case and control subjects (71.0 ± 21.4 vs. 74.5 ± 20.0 mL/min/1.73 m^2, respectively), consistent with the matching of the two study groups by this variable. However, SCys-eGFR was significantly lower in case than in control subjects (65.8 ± 28.3 vs. 77.4 ± 25.4 mL/min/1.73 m^2, $P < 0.0001$). A similar difference in baseline GFR was also observed with GFR estimates obtained with a third method (accuGFR) in a subset of study subjects (Supplementary Table 1 and Supplementary Fig. 2). Case subjects also had a significantly higher prevalence of microalbuminuria (48.7 vs. 34.7%) and macroalbuminuria (19.1 vs. 7.5%) and higher baseline use of antihypertensive medications, including RASBs, calcium channel blockers, β-blockers, and diuretics, as well as aspirin, allopurinol, nitrates, and statins (Table 1).

Global Metabolomics Analysis of CVD in T2D

Of the 903 metabolites detected in at least one sample in the discovery cohort, 232 were undetectable in >20% of samples. Of these, 28 had significantly different ($P < 0.05$) nondetectability rates between case and control subjects, with the majority more frequently undetectable in control subjects (Supplementary Table 2). These were mostly metabolites related to drugs such as atorvastatin, metoprolol, and diuretics. Only five metabolites were more frequently undetectable in case than in control subjects, including α-CEHC (carboxyethylhydroxychroman) sulfate (tocopherol metabolism), guanosine (purine metabolism), paroxetine (selective serotonin reuptake inhibitor antidepressant), glycohyocholate, and tauro-β-muricholate (bile acid metabolism) (Supplementary Table 2).

The 671 metabolites that were detectable in ≥80% of samples were included in the global metabolomics analysis. Of these, 50% were lipid analytes, 23% were amino acids, and 11% were xenobiotics; the remaining were peptides, nucleotides, energy metabolites, carbohydrates, and cofactors and vitamins (Supplementary Fig. 3*A*). In an initial conditional logistic regression analysis accounting only for the matching strata (model 1), 72 metabolites attained a q value <0.05 for association with CV events (Fig. 1*A*). Of these, 25% were lipids, 39% amino acids and their modified derivatives, 14% nucleotides, and 11% xenobiotics (Supplementary Fig. 3*B*). Many of these metabolites exhibited a high degree of mutual correlation (Supplementary Fig. 4). A majority of the modified derivatives of amino acids associated with CVD were uremic solutes (34,35), suggesting a possible confounding effect of differences in kidney function between case and control subjects (Supplementary Table 3). After further adjustment of the primary model for baseline SCre-eGFR (model 2), 48 metabolites remained significant (q value <0.05) (Fig. 1*B*). When this analysis was repeated with adjustment for SCys-eGFR rather than SCre-eGFR (model 3), four

Table 1—Baseline clinical characteristics of case and control subjects in the discovery set

	Control subjects (*N* = 294)	Case subjects (*N* = 115)	*P*
Female*	104 (35.4)	35 (30.4)	0.34
Age (years)	58.0 ± 5.6	57.7 ± 5.4	0.64
Age at diagnosis (years)	43.6 ± 8.5	42.4 ± 8.8	0.20
Diabetes duration (years)	14.4 ± 7.8	15.2 ± 8.2	0.38
Sample storage time (years)	10.7 ± 2.2	11.2 ± 2.2	0.05
BMI (kg/m^2)	31.2 ± 6.8	34.0 ± 7.4	0.0003
HbA_{1c} (%)	8.1 ± 1.4	8.2 ± 1.8	0.48
HbA_{1c} (mmol/mol)	65 ± 15	66 ± 20	0.48
Systolic BP (mmHg)	133.6 ± 17.3	135.1 ± 19.9	0.47
Diastolic BP (mmHg)	75.8 ± 10.8	74.9 ± 10.1	0.46
Cholesterol (mg/dL)	175.7 ± 38.9	180.0 ± 42.2	0.33
Triglycerides (mg/dL)#	127.5 (82.5–210.5)	160.0 (105.0–253.0)	0.007
HDL-C (mg/dL)	48.1 ± 16.9	42.4 ± 11.8	0.001
SCre-eGFR (mL/min/1.73 m^2)	74.5 ± 20.0	71.0 ± 21.4	0.12
SCys-eGFR (mL/min/1.73 m^2)	77.4 ± 25.4	65.8 ± 28.3	<0.0001
Prior CVD*	46 (16.1)	30 (38.5)	<0.0001
Urinary ACR (mg/g)#	14.6 (7.9–48.6)	60.0 (14.7–165.0)	<0.0001
Urinary ACR category			<0.0001
Normoalbuminuria	170 (57.8)	37 (32.2)	
Microalbuminuria	102 (34.7)	56 (48.7)	
Macroalbuminuria	22 (7.5)	22 (19.1)	
Diabetes treatment*			0.39
Diet alone	10 (3.5)	1 (0.9)	
Oral medication	78 (27.5)	34 (31.2)	
Oral medication + insulin	74 (26.1)	32 (29.4)	
Insulin	122 (42.9)	42 (38.5)	
Smoking*			0.43
Present	23 (8.2)	10 (9.3)	
Former	134 (47.5)	58 (53.7)	
Never	125 (44.3)	40 (37.0)	
RASB*	195 (69.2)	90 (82.6)	0.007
Calcium channel blockers	40 (13.6)	34 (29.6)	0.0002
β-Blockers	78 (26.5)	48 (41.7)	0.003
α-Blockers*	11 (3.7)	6 (5.2)	0.50
Diuretics*	18 (6.4)	18 (16.5)	0.002
Statins*	180 (61.2)	82 (71.3)	0.06
Thiazolidinediones*	44 (15.0)	19 (16.5)	0.70
Nitrates*	13 (4.4)	12 (10.4)	0.03
Aspirin*	183 (62.2)	82 (71.3)	0.08
Allopurinol*	3 (1.0)	9 (7.8)	0.0002
Race*			0.71
White	235 (83.3)	86 (79.6)	
Black	24 (8.5)	13 (12.0)	
Asian	10 (3.5)	3 (2.8)	
Native American	1 (0.4)	1 (0.9)	
Latino	10 (3.6)	3 (2.8)	
Other	2 (0.7)	2 (1.9)	

Data are means ± SD or median (interquartile range) for continuous variables and *n* (%) for categorical variables (∗). Two-sample independent *t* test *P* values for difference in means between case and control subjects are presented for continuous variables and χ^2 probabilities for categorical variables. BP, blood pressure. #Kruskal-Wallis *P* values for nonparametric variables.

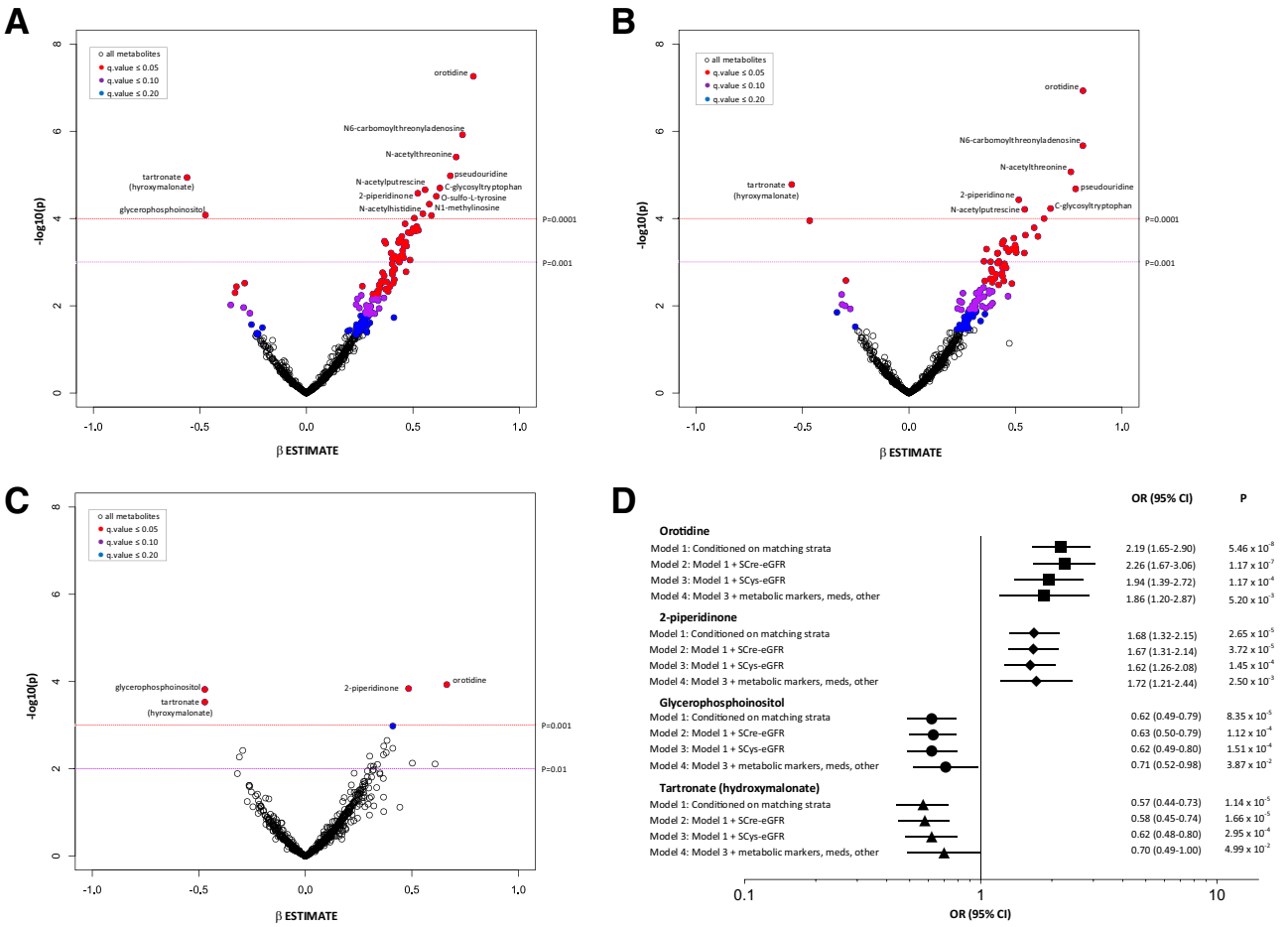

Figure 1—Global metabolomics screen for CVD risk in T2D in the discovery set. Volcano plots showing results of global metabolomics study using conditional logistic regression models to test effects of 1-SD increases in each metabolite on incident CVD risk. *A*: Model conditioned on the matching strata (sex, baseline SCre-eGFR, sample storage time, and HbA_{1c}) (model 1). *B*: Model 1 plus SCre-eGFR adjustments (model 2). *C*: Model 1 plus SCys-eGFR adjustments (model 3). *D*: Forest plot showing effects of top four metabolites on CVD. ORs and 95% CIs presented for effects of each of the four metabolites on CVD risk in various models. In model 4, the metabolic markers include baseline BMI, HbA_{1c}, ACR, HDL-c, and triglycerides. Medications include diuretics, nitrates, β-blockers, calcium channel antagonists, statins, ACE inhibitors, and angiotensin receptor blockers. Other covariates include baseline CVD, race, and smoking status. A Bonferroni cutoff of $P < 0.01$ for four tests was used to determine significance for model 4. meds, medications.

metabolites reached a *q* value of 0.05. (Fig. 1*C*). For two of these (orotidine and 2-piperidinone), the odds of having CVD events increased for each 1-SD increment by 1.94-fold (95% CI 1.39–2.72, P = 0.0001) and 1.62-fold (1.26–2.08, P = 0.0001), respectively (Fig. 1*D*). The other two metabolites that attained a *q* value <0.05 in the SCys-eGFR–adjusted analysis (glycerophosphoinositol and tartronate) were both associated with 38% decreased odds of CVD (Fig. 1*D*). Orotidine and 2-piperidinone remained significant ($P < 0.01$, with Bonferroni threshold for four tests) after further adjustments for baseline metabolic parameters (BMI, triglycerides, HDL-C, HbA_{1c}, and ACR), medication use (diuretics, β-blockers, calcium channel antagonists, RASBs, statins, nitrates), history of prior CVD, race, and smoking status (model 4) (Fig. 1*D*). By contrast, the *P* values for glycerophosphoinositol and tartronate in the fully adjusted model (model 4) were >0.01. In the subset (n = 225) having accuGFR measurements, all four metabolites remained robust to adjustments with accuGFR (data not shown). Orotidine was also further examined for association with components of the composite CVD end point. It was significantly ($P < 0.05$) associated with CV death, nonfatal MI, and PTCA but not with CABG (P = 0.1) or stroke (P = 0.2) after SCys-eGFR adjustment (Supplementary Table 4). Additionally, in a sensitivity analyses with exclusion of patients with CVD at baseline, orotidine was still significantly associated with CVD events (Supplementary Table 5). The association seemed to be weaker among participants with CVD at baseline. However, the *P* value for CVD history × orotidine interaction was not significant for models 1–3 and only marginally significant for model 4 (Supplementary Table 5).

Included in the 671 metabolites that were detectable in 80% of the samples were 38 metabolites that had previously been found to be associated with CV risk in the general population (6–11,15,16). Five of these were associated with increased CV odds in our discovery set at $q < 0.05$. However, none attained $q < 0.05$ after adjustment for SCys-eGFR and only three—stearoylcarnitine acetylcarnitine and palmitoylcarnitine—remained nominally significant ($P < 0.05$) (Supplementary

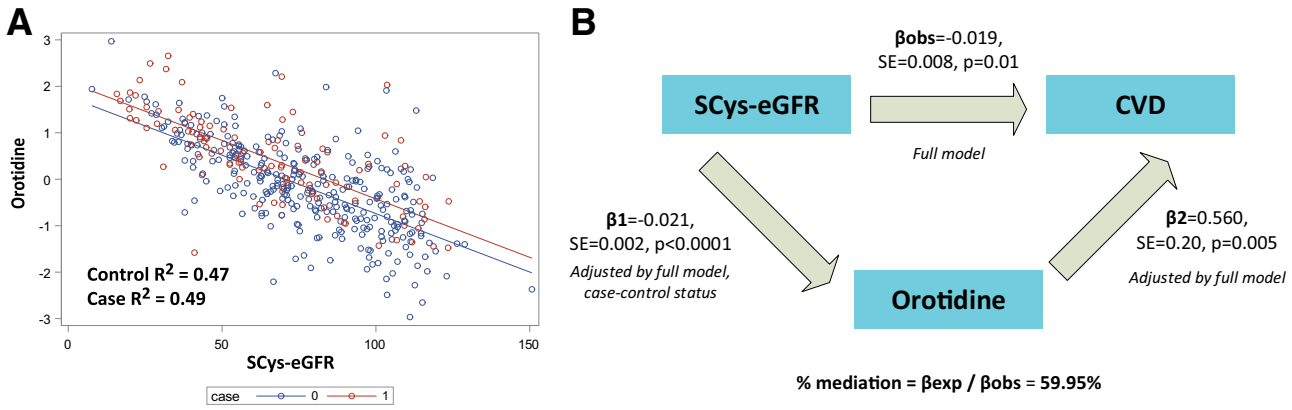

Figure 2—*A*: Correlation between orotidine and SCys-eGFR in the discovery set. Scatter plots and R^2 coefficient of variation presented for orotidine and SCys-eGFR correlations among CVD case (red) and control (blue) subjects. *B*: Mediation analysis for estimating how much of the effect of SCys-eGFR on CVD is mediated by orotidine in the discovery set.

Table 6). In this latter analysis, for each 1-SD increase in stearoylcarnitine, acetylcarnitine, and palmitoylcarnitine levels, the odds of having CVD events increased by 1.44-fold (95% CI 1.13–1.84, P = 0.003), 1.35-fold (1.03–1.75, P = 0.03), and 1.38-fold (1.08–1.76, P = 0.01), respectively.

Validation of Top Metabolites in the JHS

Quantitative measurements of orotidine and 2-piperidinone developed specifically for this study were highly correlated with those from the Metabolon platform (R^2 = 0.84 for orotidine and R^2 = 0.72 for 2-piperidinone in 20 replicates from the discovery cohort [Supplementary Fig. 5]). These assays were used to evaluate the association between these two metabolites and CHD in a validation set from the JHS, including 300 CHD-positive case subjects (defined on the basis of angiographic evidence of significant coronary stenosis) and 299 CHD-negative control subjects (defined on the basis of a negative CVD history and a normal exercise treadmill test), all with T2D (Supplementary Table 7). Among these subjects, orotidine showed a significant association with CHD, with OR of 1.54 (95% CI 1.28–1.86, $P < 0.0001$) for model 1 and ORs of 1.53 (1.23–1.89, P = 0.0001) and 1.39 (1.11–1.75, P = 0.005) in the SCre-eGFR–adjusted (model 2) and SCys-eGFR–adjusted (model 3) analyses, respectively (Supplementary Table 8). Orotidine was still associated with a 1.32-fold increase (1.02–1.71) in the odds of CHD (P = 0.036) on adjustment for blood pressure, age, triglycerides, HDL-C, smoking status, and diabetes treatment history on top of sex, SCys-eGFR, and storage time. By contrast, 2-piperidinone was not significantly associated with CHD in the validation set regardless of adjustment for estimated GFR (Supplementary Table 8). However, there was a significant trend observed (P = 0.027) per quartile increase in 2-piperidinone levels in sex- and storage time–adjusted models and nominally significant trends (P = 0.054) observed after SCys-eGFR adjustments.

Orotidine as a Mediator of the Increased CVD Associated With Lower GFR

Consistent with the well-known inverse association between kidney function and CV risk (12,13), in full model–adjusted conditional logistic regression analyses, lower baseline SCys-eGFR values were associated with higher incident CVD in the discovery set (βobs = −0.0194 ± 0.008, P = 0.01, equivalent to a 1.9% increase [95% CI 0.4–3.5] in CVD odds per 1 mL/min/1.73 m^2 decrease in SCys-GFR). At the same time, lower SCys-eGFR levels were associated with increased orotidine levels (β_1 = −0.0208 ± 0.0017, $P < 0.0001$ [Fig. 2*A*]), and, as described above, for any given value of GFR, increased orotidine levels were associated with increased CVD odds in the fully adjusted model (β_2 = 0.5603 ± 0.20, P = 0.05). Thus, a mediation analysis was performed to examine whether and to what extent the increase in serum orotidine levels associated with reduced kidney function mediated the association between low SCys-GFR and increased CVD odds (Fig. 2*B*). The expected β ($\beta_1 \times \beta_2$) for the association between SCys-GFR and CVD, if this was entirely mediated through orotidine, was estimated as −0.0116. Thus, as shown in Fig. 2, orotidine could be estimated to mediate 59.95% (−0.0116/−0.0194 × 100) of the effect of SCys-eGFR on CVD odds. The corresponding calculation in the validation set using fully adjusted models yielded an estimate of 66.7% of the effect of SCys-eGFR on CHD odds (βexp/βobs = [(−0.0241 × 0.2727)/−0.0099 × 100]) being mediated by the increase in orotidine associated with reduced kidney function. Consistent with these results, in both the discovery and validation sets, the βobs effects of SCys-GFR on CVD/CHD odds became nonsignificant with adjustment for orotidine levels.

Orotidine Improves Risk Prediction of Incident CVD

In prediction models of incident CVD in the discovery set, the AUC for model A (orotidine only) was 0.71, as compared with 0.70 for model B (traditional clinical predictors). In model C, addition of orotidine to the clinical predictors (model B) significantly improved the AUC (ΔAUC +0.053 [SE 0.026], P = 0.04) (Fig. 3). There was also a significant improvement in the rIDI (+0.481 [SE 0.016], P = 0.03) and NRI (0.424 [SE 0.136], P = 0.02), with 15% of events and 27% of nonevents correctly reclassified with model C compared with model B. All models were adequately calibrated as per the Hosmer-Lemeshow goodness-of-fit test.

CONCLUSIONS

As the T2D epidemic expands globally, and its associated CVD complications

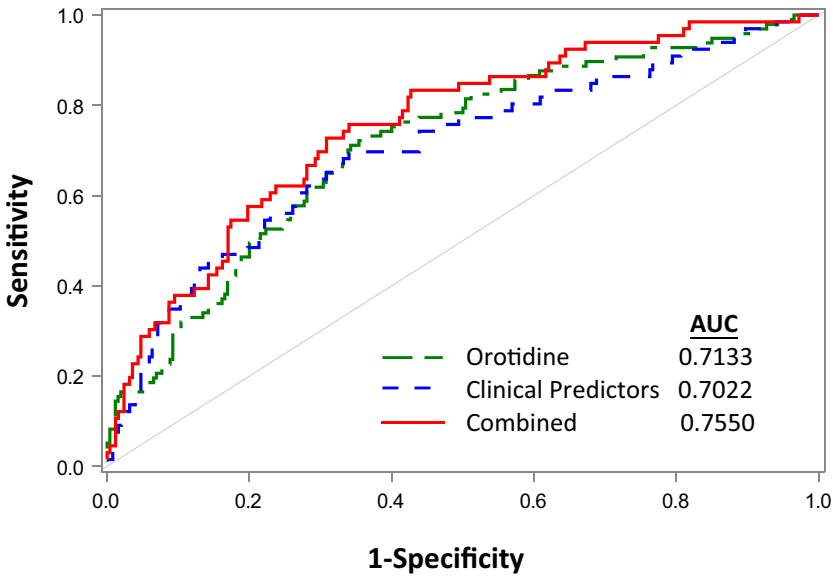

Figure 3—ROC curves for prediction of incident CVD risk in discovery set. Models tested include orotidine only, clinical predictors only, and both combined (red). Clinical predictors include the American College of Cardiology/American Heart Association atherosclerotic CVD risk score, BMI, SCys-eGFR, and prior CVD history. The AUC (C statistics) are presented.

contribute to its large health care and socioeconomic burdens, there is an urgent need to gain deeper insights into the pathophysiology of CVD in T2D to better predict and prevent these complications. To this end, we conducted a global metabolomics analysis of incident CVD risk in 409 individuals with T2D, using a nested case-control study design, and validated top signals in an independent data set using absolute quantification methods. Through this approach, we have identified serum orotidine as a novel, validated biomarker of increased CV risk in T2D.

One striking finding of our study was the profound confounding effect of kidney function. Indeed, the majority of the 72 significant metabolites in the primary eGFR-unadjusted metabolomics analysis in the discovery set were uremic solutes (34,35) (Supplementary Table 3), and most of these remained significant even after adjustment for SCre-eGFR, suggesting existence of residual confounding of renal function. Most of these associations disappeared after adjustment for SCys-eGFR—a more precise estimate of GFR in this data set as demonstrated by its stronger correlations with accuGFR in a subset of participants. As seen in Table 1 for the discovery set there was no significant difference in SCre-eGFR levels among case and control subjects, due to matching of control subjects to case subjects by this variable, while SCys-eGFR was significantly lower in case than in control subjects, as one would expect given the known association between decreased kidney function and CV risk. Nonetheless, after SCys-eGFR adjustment, four metabolites—orotidine, 2-piperidinone, glycerophosphoinositol, and tartronate—remained associated with increased CVD odds in the discovery metabolomics analysis (Fig. 1*C*). Of these, orotidine and 2-piperidinone remained significantly associated with CVD in a fully adjusted model (Fig. 1*D*), and of these two, orotidine was further associated with CHD in an independent validation set with use of absolute quantification assays.

Orotidine—a nucleoside consisting of orotic acid attached to a ribose ring—is derived from the dephosphorylation of orotidine-5-phosphate (OMP), an intermediate in the synthesis of pyrimidine nucleotides (36). Serum and urine levels of orotidine are increased in hereditary orotic aciduria (37,38), a rare autosomal recessive disorder resulting from loss of activity of uridine monophosphate (UMP) synthase (the enzyme catalyzing the last two steps of the pyrimidine nucleotide de novo synthetic pathway) and characterized by failure to thrive, developmental delay, and megaloblastic anemia. More commonly, increased serum levels of orotidine are found in people with reduced kidney function (39,40), orotidine being one of the uremic solutes that are retained in the circulation when glomerular filtration is impaired. Several studies have implicated uremic solutes in the link between chronic kidney disease and increased CVD risk (41). However, whether these metabolites play a causal role in this link or are just an epiphenomenon of impaired kidney function has been challenging to address in epidemiological studies due to the high collinearity between their serum concentrations and GFR. Our finding of serum orotidine being the only uremic solute that remained associated with CVD after adjustment for GFR points to this molecule as a prime candidate for being a causal mediator of the link between CKD and CVD/CHD among people with T2D. Indeed, mediation analyses indicated that orotidine mediated almost 60% of the effect of SCys-eGFR on incident CVD events in the discovery set and 67% of the association between SCys-eGFR and prevalent CHD in the validation set.

The biological mechanisms linking orotidine to CV risk are unclear at this time. A study on the impact of uremic toxins on spermatozoa mobility included orotidine, but this was not among the molecules having the strongest impact (42). Studies have shown that accumulation of orotate (orotic acid) may induce intracellular lipid accumulation by impairing hepatic mitochondrial respiration via inhibition of DHODH activity (43,44). However, whether orotidine has the same effect is unknown. Another possibility is that increased serum orotidine is a marker of increased orotic acid and that the association between orotidine and CVD is due to its correlation with this other molecule. In the Metabolon platform, orotate significantly correlated with orotidine ($\beta = 0.42$, SE = 0.05, $P < 0.0001$) and was associated with CVD in the eGFR-unadjusted analysis (OR 1.30 [95% CI 1.04–1.62], $P = 0.02$) but not in the SCys-eGFR–adjusted model (1.20 [0.96–1.52], $P = 0.1$). However, this could still be a mechanism underlying the association between orotidine and CVD if orotidine is a better marker of intracellular orotate levels than serum orotate.

In addition to the novel findings with orotidine, we found evidence of association with CVD ($q < 0.05$) in the discovery set for five metabolites that had previously been found to be associated with CV risk in the general population

(6–11,15,16). Interestingly, a confounding effect of GFR was also observed for most of these metabolites, suggesting that many of these previous findings may have been related to the lack of adjustment for differences in kidney function between CVD case and control subjects. The only exceptions were the three long-chain acylcarnitines stearoylcarnitine, acetylcarnitine, and palmitoylcarnitine, for which the association with CVD remained nominally significant after adjustment for SCys-eGFR. These findings confirm the importance of alterations of energy and fatty acid metabolism in the etiology of CVD in general (45) and in the specific case of T2D (46). Branched-chain amino acids that were in our panel, including leucine and isoleucine, though shown to be associated previously with incident CVD (47), and with T2D risk in the general population, were not associated with CVD in our data set. In a previous study in the PREvención con DIeta MEDiterránea (PREDIMED) cohort, an association was found between ceramides and incident CVD (48), but these were not detected in our Metabolon platform panel, and the four other ceramides on our panel were not significant (Supplementary Table 9). In another PREDIMED lipidomic study, investigators reported a group of polyunsaturated phosphatidylcholine (PC), lysoPC, PC plasmalogens, and cholesterol esters to be associated with CVD protection and a group of monoacylglycerols/diacylglycerols and short triacylglycerols associated with CVD risk (49). The CV-protective metabolites were not in our panel. Of the risk metabolites, while there were no triacylglycerols, 7 of the 15 monoacylglycerols and 4 of the 16 diacylglycerols in our panel remained nominally significant ($P < 0.05$) in the SCys-eGFR–adjusted analysis (Supplementary Table 9). A possible explanation for these discrepant results aside from the different methodology is that confounding by T2D or renal function was not examined in the PREDIMED study.

Aside from the potential biological significance of this novel biomarker, orotidine also significantly improved the performance of risk prediction models containing conventional clinical risk factors of CVD (model C compared with model B in Fig. 3), suggesting its potential as a tool for better identification of individuals with T2D at high risk of CVD events. This may in turn enhance cost-effectiveness via improved targeted prevention strategies against CVD in T2D.

Our study had several strengths including the specific focus on individuals with T2D, the prospective design of the discovery set, use of a well-validated metabolomics platform, attention to the confounding effects of kidney function and other variables, and external validation of findings in another data set with a different design. Also, that discovery cases were defined on the basis of self-reported CV events, whereas for the validation cases there was angiographic evidence of coronary artery disease, and that, despite these differences, orotidine was associated with the CV end points in both sets, adds to the strength of our findings. Nonetheless, some potential limitations should be acknowledged. First, while our study was larger than many of the previous ones, its power was constrained by the number of events in the cohort under investigation. Thus, weaker but still biologically relevant associations between other metabolites and CVD might have gone undetected. Second, while we used two different methods to estimate kidney function in case and control subjects, residual confounding by differences in GFR between the two study groups is still possible. This concern, however, is alleviated by the fact that the association between orotidine and CVD remained significant after adjustment with a third GFR estimation method (accuGFR) in a subset of the discovery set for which this measure was available. Third, the CVD outcome in the discovery set was heterogenous, comprising not only hard end points (MI, stroke, and CV mortality) but also PTCA and CABG. However, in a sensitivity analysis, we observed that the associations with orotidine remained when considering each of the end points separately (Supplementary Table 4). Fourth, our analysis included all individuals with or without baseline CVD, since a metabolite could be a risk factor for new-onset CVD events as well, and in a sensitivity analysis (Supplementary Table 5), while there was no significant interaction between orotidine and CVD history, we could not entirely rule out a possible negative interaction (i.e., orotidine may be a stronger risk factor for CVD in people who have negative CVD history) due to the small sample size of the group with positive CVD history. Fifth, the cross-sectional nature of the association between estimated GFR and orotidine cannot exclude reverse causation and should be addressed in future prospective studies. Sixth, we were unable to assess the impact of race/ethnic heterogeneity on our findings, as 83% of subjects in the discovery set and all subjects in the validation set were non-Hispanic White. In the discovery set, adjusting for race did not impact the results. Furthermore, this study did not address multivariate analysis incorporating mutual correlations between metabolites, as it was focused on individual metabolites as predictors of CVD in diabetes. Additionally, aside from smoking, we could not address confounding by other lifestyle factors such as diet or physical activity as these data were not obtained in either study set. Lastly, serum samples for the metabolomics study were obtained without regard to participants' fasting status. While this could have added noise to our findings and possibly result in some false negatives, it is highly unlikely that our positive results were confounded by variability in fasting status, since for this to happen the fasting status of participants at the time of baseline sampling would have had to be associated with future caseness. If anything, the noise added by fasting variability might have biased results toward the null, making our positive findings even more remarkable.

In summary, through an untargeted metabolomics analysis, we have discovered orotidine as a novel biomarker of increased CVD risk in T2D, having robust effects independent of estimated GFR. Importantly, we have validated this biomarker via absolute quantification methods in an independent data set. In future studies investigators should explore the causal pathways of the CV effects of orotidine, for example, through Mendelian randomization methods and/or cellular and animal studies, to understand whether this is a viable target for novel interventions aimed at decreasing the burden of CVD among patients with T2D.

Acknowledgments. The authors acknowledge study coordinators, clinical research nursing staff, and participants of the JSK and JHS. The authors acknowledge Dr. Andrzej Krolewski at

the Joslin Diabetes Center for giving the authors access to the JSK samples and database.

Funding. This research was supported by National Institutes of Health (NIH) grants HL132254 (to A.D.), DK082841 (to S.P.), DK036836 (Molecular Phenotyping and Genotyping Core of the Joslin Diabetes Research Center), DK081943 (to the University of Michigan George M. O'Brien Michigan Kidney Translational Core Center), and DK089503 (to the Michigan Nutrition Obesity Research Center). L.O.M. was supported by grants Ayudas Atracción de Talento modalidad and Comunidad de Madrid (BMD-5800). M.A.N was supported by the JDRF (CDA-2015-89-A-B).

The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or other funding sources.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. H.S.S. and A.D. contributed to study design, were responsible for integrity of data acquisition, contributed to data analysis and interpretation, and contributed to writing the manuscript and the publication process. L.O.M., M.-L.M., Y.T., S.M., and M.A.N. contributed to data analysis and interpretation and review of the manuscript. C.M., J.M.J., J.B.T., J.M., and S.P. contributed to data acquisition, analysis, and interpretation and review of manuscript. H.S.S. and A.D. are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Saeedi P, Petersohn I, Salpea P, et al.; IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2019;157:107843
2. Rawshani A, Rawshani A, Franzén S, et al. Mortality and cardiovascular disease in type 1 and type 2 diabetes. N Engl J Med 2017;376:1407–1418
3. Sarwar N, Gao P, Seshasai SR, et al.; Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet 2010;375:2215–2222
4. Orasanu G, Plutzky J. The pathologic continuum of diabetic vascular disease. J Am Coll Cardiol 2009;53(Suppl.):S35–S42
5. Kell DB, Oliver SG. The metabolome 18 years on: a concept comes of age. Metabolomics 2016;12:148
6. Shah SH, Sun JL, Stevens RD, et al. Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. Am Heart J 2012;163:844–850.e1
7. Rizza S, Copetti M, Rossi C, et al. Metabolomics signature improves the prediction of cardiovascular events in elderly subjects. Atherosclerosis 2014;232:260–264
8. Vaarhorst AA, Verhoeven A, Weller CM, et al. A metabolomic profile is associated with the risk of incident coronary heart disease. Am Heart J 2014;168:45–52.e7
9. Würtz P, Havulinna AS, Soininen P, et al. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. Circulation 2015;131:774–785
10. Ganna A, Salihovic S, Sundström J, et al. Large-scale metabolomic profiling identifies novel biomarkers for incident coronary heart disease. PLoS Genet 2014;10:e1004801
11. Floegel A, Kühn T, Sookthai D, et al. Serum metabolites and risk of myocardial infarction and ischemic stroke: a targeted metabolomic approach in two German prospective cohorts. Eur J Epidemiol 2018;33:55–66
12. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med 2004;351:1296–1305
13. Chronic Kidney Disease Prognosis Consortium; Matsushita K, van der Velde M, Astor BC, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet 2010;375:2073–2081
14. Sekula P, Goek O-N, Quaye L, et al. A metabolome-wide association study of kidney function and disease in the general population. J Am Soc Nephrol 2016;27:1175–1188
15. Stratmann B, Richter K, Wang R, et al. Metabolomic signature of coronary artery disease in type 2 diabetes mellitus. Int J Endocrinol 2017;2017:7938216
16. Chen ZZ, Gerszten RE. Metabolomics and proteomics in type 2 diabetes. Circ Res 2020;126:1613–1627
17. Ihara K, Skupien J, Kobayashi H, et al. Profibrotic circulating proteins and risk of early progressive renal decline in patients with type 2 diabetes with and without albuminuria. Diabetes Care 2020;43:2760–2767
18. Nowak N, Skupien J, Smiles AM, et al. Markers of early progressive renal decline in T2D suggest different implications for etiological studies and prognostic tests development. Kidney Int 2018;93:1198–1206
19. Bergmann MM, Byers T, Freedman DS, Mokdad A. Validity of self-reported diagnoses leading to hospitalization: a comparison of self-reports with hospital records in a prospective study of American adults. Am J Epidemiol 1998;147:969–977
20. Martin LM, Leff M, Calonge N, Garrett C, Nelson DE. Validation of self-reported chronic conditions and health services in a managed care population. Am J Prev Med 2000;18:215–218
21. Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 2004;57:1096–1103
22. National Center for Health Statistics. National Death Index User's Guide. Hyattsville, MD, National Center for Health Statistics, 2013
23. Doria A, Wojcik J, Xu R, et al. Interaction between poor glycemic control and 9p21 locus on risk of coronary artery disease in T2D. JAMA 2008;300:2389–2397
24. Qi L, Qi Q, Prudente S, et al. Association between a genetic variant related to glutamic acid metabolism and coronary heart disease in individuals with T2D. JAMA 2013;310:821–828
25. Laterza OF, Price CP, Scott MG. Cystatin C: an improved estimator of glomerular filtration rate? Clin Chem 2002;48:699–707
26. Inker LA, Schmid CH, Tighiouart H, et al.; CKD-EPI Investigators. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med 2012;367:20–29
27. Freed TA, Coresh J, Inker LA, et al. Validation of a metabolite panel for a more accurate estimation of glomerular filtration rate using quantitative LC-MS/MS. Clin Chem 2019;65:406–418
28. Warram JH, Gearin G, Laffel L, Krolewski AS. Effect of duration of type I diabetes on the prevalence of stages of diabetic nephropathy defined by urinary albumin/creatinine ratio. J Am Soc Nephrol 1996;7:930–937
29. Niewczas MA, Mathew AV, Croall S, et al. Circulating modified metabolites and a risk of ESRD in patients with type 1 diabetes and chronic kidney disease. Diabetes Care 2017;40:383–390
30. Dehaven CD, Evans AM, Dai H, Lawton KA. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. J Cheminform 2010;2:9
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 1995;57:289–300. Available from https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
32. Kennedy KF, Pencina MJ. A SAS macro to compute added predictive ability of new markers predicting a dichotomous outcome. North Carolina State University, 2010. Accessed 6 February 2021. Available from https://analytics.ncsu.edu/sesug/2010/SDA07.Kennedy.pdf
33. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–172; discussion 207–212
34. Tanaka H, Sirich TL, Plummer NS, Weaver DS, Meyer TW. An enlarged profile of uremic solutes. PLoS One 2015;10:e0135657
35. Vanholder R, De Smet R, Glorieux G, et al.; European Uremic Toxin Work Group (EUTox). Review on uremic toxins: classification, concentration, and interindividual variability. Kidney Int 2003;63:1934–1943
36. El Kouni MH, Cha S. Isolation and partial characterization of a 5′-nucleotidase specific for orotidine-5′-monophosphate. Proc Natl Acad Sci U S A 1982;79:1037–1041
37. Bailey CJ. Orotic aciduria and uridine monophosphate synthase: a reappraisal. J Inherit Metab Dis 2009;32(Suppl. 1):S227–S233
38. Suchi M, Mizuno H, Kawai Y, et al. Molecular cloning of the human UMP synthase gene and characterization of point mutations in two hereditary orotic aciduria families. Am J Hum Genet 1997;60:525–539
39. Vanholder R, Van Laecke S, Glorieux G. What is new in uremic toxicity? Pediatr Nephrol 2008;23:1211–1221
40. Liu S, Wang L, Hu C, et al. Plasma metabolomics profiling of maintenance hemodialysis

based on capillary electrophoresis - time of flight mass spectrometry. Sci Rep 2017;7:8150
41. Ravid JD, Chitalia VC. Molecular mechanisms underlying the cardiovascular toxicity of specific uremic solutes. Cells 2020;9:2024
42. Vollmer T, Ljungberg B, Jankowski V, Jankowski J, Glorieux G, Stegmayr BG. An in-vitro assay using human spermatozoa to detect toxicity of biologically active substances. Sci Rep 2019;9:14525
43. Le TT, Ziemba A, Urasaki Y, Hayes E, Brotman S, Pizzorno G. Disruption of uridine homeostasis links liver pyrimidine metabolism to lipid accumulation. J Lipid Res 2013;54:1044–1057
44. Creasey WA, Hankin L, Handschumacher RE. Fatty livers induced by orotic acid. I. Accumulation and metabolism of lipids. J Biol Chem 1961;236:2064–2070
45. Strand E, Pedersen ER, Svingen GF, et al. Serum acylcarnitines and risk of cardiovascular death and acute myocardial infarction in patients with stable angina pectoris. J Am Heart Assoc 2017;6:e003620
46. Ljubkovic M, Gressette M, Bulat C, et al. Disturbed fatty acid oxidation, endoplasmic reticulum stress, and apoptosis in left ventricle of patients with type 2 diabetes. Diabetes 2019;68:1924–1933
47. Ruiz-Canela M, Toledo E, Clish CB, et al. Plasma branched-chain amino acids and incident cardiovascular disease in the PREDIMED trial. Clin Chem 2016;62:582–592
48. Wang DD, Toledo E, Hruby A, et al. Plasma ceramides, Mediterranean diet, and incident cardiovascular disease in the PREDIMED trial (Prevención con Dieta Mediterránea). Circulation 2017;135:2028–2040
49. Razquin C, Liang L, Toledo E, et al. Plasma lipidome patterns associated with cardiovascular risk in the PREDIMED trial: a case-cohort study. Int J Cardiol 2018;253:126–132