Identification of Novel Biomarkers and Pathways for Coronary Artery Calcification in Non-diabetic Patients on Hemodialysis Using Metabolomic Profiling

Wei Chen, Jessica Fitzpatrick, Stephen M. Sozio, Bernard G. Jaar, Michelle M. Estrella, Dario F. Riascos-Bernal, Tong Tong Wu, Yunping Qiu, Irwin J. Kurland, Ruth F. Dubin, Yabing Chen, Rulan S. Parekh, David A. Bushinsky, Nicholas E.S. Sibinga

1Department of Medicine, Albert Einstein College of Medicine, Bronx, New York, USA; 2Department of Medicine, University of Rochester School of Medicine and Dentistry, Rochester, New York, USA; 3Department of Medicine and Pediatrics, Hospital for Sick Children and University Health Network, University of Toronto, Toronto, Ontario, Canada; 4Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; 5Department of Epidemiology, Bloomberg School of Public Health, Baltimore, Maryland, USA; 6Nephrology Center of Maryland, Maryland, USA; 7Kidney Health Research Collaborative, Department of Medicine, University of California, San Francisco, San Francisco, California, USA; 8San Francisco VA Health Care System, San Francisco, California, USA; 9Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, New York, USA; 10Department of Biostatistics and Computational Biology, University of Rochester School of Medicine and Dentistry, Rochester, New York, USA; 11Einstein-Mount Sinai Diabetes Research Center, Albert Einstein College of Medicine, Bronx, New York, USA; 12Department of Medicine, University of California San Francisco, San Francisco, California, USA; 13Department of Pathology, University of Alabama at Birmingham and Birmingham Veterans Affairs Medical Center, Birmingham, Alabama, USA

CORRESPONDING AUTHOR:
Wei Chen, MD, MS
Assistant Professor of Medicine
Albert Einstein College of Medicine
1300 Morris Park Avenue, Ullmann 615
Bronx, NY 10461
Tel: 718-430-3825
Email: weichen@montefiore.org

Copyright 2020 by American Society of Nephrology.
ABSTRACT

Background: A better understanding of pathophysiology involving coronary artery calcification (CAC) in hemodialysis (HD) patients will help to develop new therapies. We sought to identify the differences in metabolomics profiles between HD patients with and without CAC.

Methods: In this case-control study nested within a cohort of 568 incident HD patients, cases were non-diabetics with a CAC score >100 (n=51), and controls were non-diabetics with a CAC score of 0 (n=48). We measured 452 serum metabolites in each participant. Metabolites and pathway scores were compared using Mann–Whitney U tests, partial least squares-discriminant analyses, and pathway enrichment analyses.

Results: Compared to controls, cases were older (64±13 vs. 42±12 years) and were less likely to be African American (51% vs. 94%). We identified three metabolites in bile acid synthesis (chenodeoxycholic, deoxycholic, and glycolithocholic acids) and one pathway (arginine/proline metabolism). After adjusting for demographics, higher levels of chenodeoxycholic, deoxycholic, and glycolithocholic acids were associated with higher odds of having CAC: comparing the third with the first tertile of each bile acid, the OR (95% CI) was 6.34 (1.12-36.06), 6.73 (1.20-37.82), and 8.53 (1.50-48.49), respectively. These associations were no longer significant after further adjustment for coronary artery disease and medication use. Per 1 unit higher in the first principal component score, arginine/proline metabolism was associated with CAC after adjusting for demographics (OR: 1.83 (95% CI: 1.06-3.15)), and the association remained significant with additional adjustments for statin use (OR: 1.84 (95% CI: 1.04-3.27)).

Conclusions: Among HD patients without diabetes mellitus, chenodeoxycholic, deoxycholic, and glycolithocholic acids may be potential biomarkers for CAC, and arginine/proline metabolism is a plausible mechanism to study for CAC. These findings need to be confirmed in future studies.
INTRODUCTION

In patients with end stage renal disease (ESRD) on hemodialysis (HD), coronary artery calcification (CAC) is prevalent and independently predicts the risk of cardiovascular disease, which is the leading cause of death in this population. The mechanisms of CAC in HD patients beyond the passive process of calcium phosphate precipitation are not well understood. In this hypothesis-generating study, we used metabolomic profiling to identify new biomarkers and pathways for CAC among HD patients. Metabolomic profiling, which studies small molecules such as amino acids in biological specimens, may provide more insight into the factors involved in cardiovascular calcification.

In general, the metabolome is downstream of transcriptional and translational processes and incorporates inputs from diet, environment and microbiome — thus metabolomics reflects a summative assessment of genes, proteins, and exogenous factors. In addition, metabolomics has the potential for rapid translational application, as some metabolites may be amenable to therapeutic targeting.

Previous studies have examined serum metabolites and cardiovascular mortality in HD patients, but none has studied the relationship between serum metabolites and cardiovascular calcification. We performed a case-control study nested in an existing HD cohort from the Predictors of Arrhythmic and Cardiovascular Risk in ESRD (PACE) study, and compared serum metabolites in non-diabetic HD patients with and without CAC. For metabolites and pathways identified through this analysis, we further explored their relationships with mineral metabolism, calcification inhibitors, and inflammation.

METHODS

Study design and population

We conducted a nested case-control study within PACE, a prospective cohort designed to determine cardiovascular and dialysis-related risk factors in patients on HD. PACE was approved by the Johns Hopkins School of Medicine and MedStar Institutional Review Boards; 568 incident HD
patients (i.e. on HD for <6 months) were recruited from 25 free-standing outpatient HD units and 2 hospital-based outpatient units in Baltimore, Maryland and its surrounding area from 2008–2012. CAC was measured at the baseline visit by computed tomography and quantified by Agatston score.\textsuperscript{11} All PACE participants were evaluated for potential inclusion into our metabolomics study. A total of 130 non-diabetic participants had serum stored and underwent CAC measurement (Figure 1 and Supplemental Methods; n was used to denote the number of participants and m for the number of metabolites). We limited our study to participants without diabetes mellitus to avoid potential confounding effects of diabetes-associated metabolic dysregulation.\textsuperscript{12} Controls (n=48) were participants with a CAC score of 0. Cases (n=51) were defined as having a CAC score >100. The remaining 31 participants with CAC scores between 0 and 100 were not included in the study. CT is a sensitive method to detect coronary calcium with a sensitivity as high as 99% using a cutoff score of 0; however, its specificity is lower, approximately 65%.\textsuperscript{13,14} By using a cut off of 100 to define cases, we increased the specificity to 77%.\textsuperscript{14} We compared the baseline characteristics of study population with unselected participants in Table S1. Cases and controls were not matched.

**Measurement of serum metabolites**

Serum was collected on a non-HD day at the baseline visit after ~8 hours of fasting and stored at -80°C. Using triple quadrupole mass spectrometry (AB Sciex 6500 + QTRAP; Concord, Ontario, Canada) coupled with a Waters Ultra Performance Liquid Chromatography system (Milford, Massachusetts, USA) as previously described,\textsuperscript{15} we performed widely targeted metabolomics (452 metabolites) in stored serum. A pooled quality control sample was injected 6 times and used to calculate the coefficients of variation (CVs). A distribution of CVs is shown in Figure S1a. Metabolites with a CV <30% (n=247) were included in the analyses, and their pathways are shown in Figure S1b. Missing metabolite values (1.3%) were replaced by half of the minimum positive value in the original dataset.
Measurement of covariates

Potential confounders included demographics, education level, smoking history, comorbidities, medication use, dialysis access, and dialysis clearance. Participants’ demographic factors (age, sex, and race), education level, smoking history, and medication use were self-reported. Comorbidities including diabetes mellitus, coronary artery disease (CAD), and hypertension were adjudicated by a committee of physicians. Dialysis clearance was assessed by single-pool Kt/V (spKt/V). We also measured serum markers of mineral metabolism (calcium, phosphorous, intact parathyroid hormone (PTH), C-terminal fibroblast growth factor (FGF23), and soluble klotho), circulating inhibitors of cardiovascular calcification (osteoprotegerin, dephosphorylated and uncarboxylated matrix glutamate (Gla) protein (dp-ucMGP), and fetuin-A), secondary calciprotein particle (CPP2) size and the time of transformation from primary to secondary calciprotein particle (T_{50}), and high sensitive C-reactive protein. Details of these measurements are available in Supplemental Methods.

Statistical and pathway analyses

Baseline participant characteristics were examined by CAC status using two-sample t-test, Mann-Whitney U test, or chi-squared test, as appropriate. Metabolomic data were normalized using auto scaling (mean-centered and divided by the standard deviation of each metabolite). Figure 1 includes an analysis flowchart to identify significant metabolites and pathways.

First, we compared each metabolite (non-normally distributed) by the CAC status using Mann-Whitney U tests and partial least squares-discriminant analyses (PLS-DA). For Mann-Whitney U tests, we used a p-value threshold of <0.005 and a fold difference threshold of 1.0 to obtain approximately 10 metabolites that differentiated cases from controls. The results were presented in a volcano plot. PLS-DA is commonly used to analyze metabolomics datasets because of its ability to analyze data with high collinearity. For PLS-DA, we identified the top 15 metabolites based on the variable importance in project (VIP) scores. For the top metabolites (m=15) identified by Mann-Whitney U tests or PLS-DA, we
examined the association of each metabolite with CAC using multiple logistic regression models adjusting for demographics (age, sex and race). For the metabolites that had a p value <0.05 after adjusting for demographics (m=3; chenodeoxycholic, deoxycholic and glycolithocholic acids), we generated fully adjusted models using backward elimination. In addition to individual metabolite (in continuous scale) and demographics, backward elimination selected from the following 9 covariates: body mass index, smoking history, education level, spkt/V, CAD, and medication use including vitamin D therapy, calcium-based phosphate binder, renin-angiotensin-aldosterone system (RAAS) blockade and statin. Significance threshold to remove a covariate from the models was 0.1. The backward elimination approach generated the same set of covariates for chenodeoxycholic, deoxycholic and glycolithocholic acids, and the final full models were adjusted for demographics, CAD, and use of calcium-base phosphate binder, RAAS blockade and statin. Additionally, we re-fitted logistic models using tertiles of the 3 bile acids, and log-transformed the bile acids to examine their correlations with each other using Pearson Correlation.

Second, we performed pathway enrichment analyses using the Small Molecule Pathway Database. To produce robust results, we used the metabolites that met the following 3 criteria: a VIP score >1, p-value <0.05, and available human metabolome database identification (m=34). Top 3 pathways were identified based on the combination of pathway significance and impact. Pathway significance was assessed using global tests with a false discovery rate adjusted p-value threshold of 0.05, and pathway impact was assessed using topology analysis with relative-betweenness centrality.

Third, we tested the associations between the top 3 pathways and CAC status. In order to do that, we used principal component analysis and metabolites with CV<30% (m=247) to score the top 3 pathways identified in step 2. The first principal component (PC1) score of each pathway was then used to represent the pathway. We examined the association between the PC1 of each pathway with CAC using multiple logistic regression. Models were first adjusted for demographics, then similar to step 1, we generated fully adjusted models using backward elimination, which selected covariates in
addition to individual PC1 of each pathway and demographics. The fully adjusted models were adjusted for demographics and statin use for arginine/proline metabolism and urea cycle pathways, and adjusted for demographics, CAD, and use of calcium-based phosphate binder, RAAS blockade and statin for bile acid synthesis pathway. For secondary analyses, we used Tobit regression to simultaneously model the presence and severity of CAC and to examine their associations with key metabolites and pathways identified in Step 1 and 2. Because CAC score has a right skewed distribution and zero scores, we log-transformed CAC score plus 1, and used left censoring at 0 and bootstrap techniques with 999 repetitions.

Lastly, using Spearman’s rank correlation, we examined the correlations of key metabolites and/or pathways with the markers of mineral metabolism, inflammation, circulating inhibitors of cardiovascular calcification, and properties of calciprotein particle transformation. MetaboAnalyst 4.0 (McGill University, Montreal, Quebec, CA) and STATA 14.1 (StataCorp, College Station, TX, USA) were used for statistical and pathway analyses. A two-sided p-value <0.05 was considered statistically significant in the multiple logistic and Tobit regression models.

RESULTS

Participant characteristics

As shown in Table 1, the mean age of participants was 53 ± 16 years, 38% were women, and 72% were African American. All participants had hypertension and 27% had CAD. Among cases, the median CAC score was 466 [interquartile range (IQR) 246-981]. Compared to controls, cases were older (64±13 vs. 42±12 years), less likely to be African American (51% vs. 94%) and more likely to have CAD (45% vs. 8%). Cases were also less likely to be taking medications that blocked RAAS (36% vs. 58%) and more likely to be on statin therapy (42% vs. 13%).
Identification of significant metabolites

Using the Mann-Whitney U test, and a set threshold of fold difference (1.0) and raw p-value (0.005), we identified 9 metabolites that differed between cases and controls, as shown in the volcano plot (Figure S2, Table S2). Compared to the controls, bile acid levels in cases were more than 2-fold higher: chenodeoxycholic acid, deoxycholic acid, and glycolithocholic acid (Figure 2). Using PLS-DA, we identified the top 15 metabolites based on VIP scores (Figure S3 & S4). The metabolites that were identified by the volcano plot and PLS-DA are summarized in Table S2. In addition to the 9 metabolites shown in the volcano plot, PLS-DA identified 3 metabolites involved in arginine/proline metabolism: arginine, ornithine, and succinic acid (Figure 2).

After adjusting for demographics, only bile acids remained significantly associated with CAC status (Table S2). Higher levels of chenodeoxycholic, deoxycholic and glycolithocholic acids were associated with higher odds of CAC. After adjusting for demographics, compared to the participants with bile acid levels in the first tertile, those in the third tertile had higher odds of CAC with an odds ratio (OR) of 6.34 (95% CI: 1.12-36.06, p=0.04), 6.73 (95% CI: 1.20-37.82, p=0.03), and 8.53 (95% CI: 1.50-48.49, p=0.02) for chenodeoxycholic, deoxycholic, and glycolithocholic acid, respectively (Table 2). However, the associations were no longer significant after additional adjustment for CAD, and use of calcium-based phosphate binder, RAAS blockade and statin.

Chenodeoxycholic, deoxycholic and glycolithocholic acids were correlated with each other (Figure S5). For secondary analyses, Tobit regression revealed overall similar observations as the logistic regression (Table S3). After adjusting for demographics, the difference in the CAC score of those with CAC score >0 (weighted by the probability of having a CAC score>0) and in the probability of having CAC score >0 (weighted by the expected value of CAC score if >0) was 0.15 per 1 unit higher in relative intensity of chenodeoxycholic acid (p=0.02). The association remained significant after further adjustment for CAD and medication use (p=0.03). Deoxycholic and glycolithocholic acids were
associated with CAC using Tobit regression after adjusting for demographics, but not in the fully adjusted models.

**Identification of significant pathways**

Pathway enrichment analyses revealed 3 top pathways: arginine/proline metabolism, urea cycle, and bile acid synthesis (Table 3). Metabolites analyzed in these 3 pathways are shown in Table S4. When comparing cases with controls, 5 metabolites in arginine/proline metabolism, 4 in urea cycle and 3 in bile acid synthesis had a p-value <0.05 (Figure 2). L-arginine, ornithine, and citrulline were present in both arginine/proline metabolism and urea cycle pathways.

We used principal component analysis to generate scores for the 3 pathways. The PC1 score explained 20.3% of variance in the arginine/proline metabolism, 23.9% in urea cycle, and 37.2% in bile acid synthesis (Figure S6). PC1s of arginine/proline metabolism and urea cycle were associated with CAC status, but only PC1 of arginine/proline metabolism remained significant after adjusting for demographics and statin use, with an adjusted OR of 1.84 (95% CI: 1.04-3.27) per 1 unit higher of PC1 score (p=0.04; Table 4). In secondary analyses using Tobit regression, we found that the presence and severity of CAC were associated with the PC1 of both arginine/proline metabolism and urea cycle pathways in the fully adjusted model (p=0.02 and 0.04, respectively), but not with PC1 of bile acid synthesis (Table S3).

**Correlations with significant metabolites and pathways**

We then examined the correlations of bile acid synthesis and arginine/proline metabolism pathways and their key metabolites with serum markers of mineral metabolism and inflammation, circulating inhibitors of cardiovascular calcification, and properties of calciprotein particles. Results are shown in a heat map (Figure 3) and in Table S5. The PC1 score of bile acid synthesis was positively correlated with FGF-23 and osteoprotegerin levels and negatively correlated with T_{50}. Chenodeoxycholic, deoxycholic, and glycolithocholic acids correlated negatively with C-reactive protein,
but were not correlated with serum markers of mineral metabolism or circulating inhibitors of calcification. The PC1 score of arginine/proline metabolism correlated positively with levels of serum phosphorous, FGF-23, osteoprotegerin, and dp-ucMGP.

DISCUSSION

In patients with ESRD on HD, CAC is a prevalent and independent predictor of cardiovascular morbidity and mortality. In this nested case-control study, we compared metabolites and pathway scores in 99 non-diabetic, incident HD patients with a CAC score >100 and a CAC score of 0. Compared to those with a CAC score of 0, patients with a CAC score >100 have more than 3 times the risk of coronary events. We identified three metabolites in bile acid synthesis (chenodeoxycholic, deoxycholic, and glycolithocholic acids), and one metabolic pathway (arginine/proline metabolism) that were associated with CAC.

We found that higher levels of chenodeoxycholic, deoxycholic, and glycolithocholic acids were associated with CAC after adjusting for demographics. Similarly, in a study of patients with pre-dialysis chronic kidney disease (CKD), higher serum deoxycholic acid levels were associated with greater CAC after adjusting for demographics, comorbidities and kidney function. However, after we further adjusted for comorbidities and medication use, the association between bile acids and CAC status was no longer significant. The lack of association in the fully-adjusted logistic models could be due to small sample size. In the sensitivity analyses using Tobit regression, chenodeoxycholic acid remained significantly after adjusting for demographics, CAD, and medication use. Compared to logistic regression, Tobit regression is more likely to identify predictors for cardiovascular calcification.

Elevated bile acid levels may reflect disruption of intestinal barrier function or alteration in the intestinal microbiome, making them potential candidates as biomarkers for CAC in HD patients. Bile acids are important for intestinal nutrient absorption and biliary secretion of lipid and toxic metabolites (Figure S7a). Chenodeoxycholic acid is a primary bile acid, as it is directly synthesized from
cholesterol in liver. After synthesis, chenodeoxycholic acid is secreted into intestinal tract, where it is converted by intestinal bacteria to lithocholic acid, a secondary bile acid. Conjugation of lithocholic acid with glycine forms glycolithocholic acid, increasing solubility and decreasing toxicity. Cholic acid, another primary bile acid, is also converted in the intestinal tract to deoxycholic acid, a secondary bile acid. After bile acids are secreted into the intestinal tract, they are reabsorbed in the ileum or by portal circulation back to the liver. This enterohepatic circulation of bile acids is generally highly efficient. Patients on HD may have a disrupted intestinal barrier, which results in translocation of bile acids, endotoxins and bacterial metabolites in the systemic circulation. These translocated metabolites and toxins may then contribute to chronic inflammation and ultimately to cardiovascular disease.

Animal studies have shown that bile acids may have direct effects on cardiovascular calcification via farnesoid X receptor, a bile acid nuclear receptor. In our study, bile acid synthesis was one of the top 3 pathways identified by the pathway enrichment analyses, but its pathway impact was low compared to that of arginine/proline metabolism and urea cycle (0.01 versus 0.21 and 0.23), and the association between the bile acid synthesis pathway and CAC status was not significant. We also did not observe any significant correlation of bile acids with serum markers of mineral metabolism or circulating inhibitors of calcification. Our findings do not support a causal relationship between bile acid synthesis and the development of CAC.

We found that arginine/proline metabolism was associated with CAC in both fully adjusted logistic and Tobit regression models. Urea cycle pathway has three common metabolites with arginine/proline metabolism (i.e. L-arginine, ornithine, and citrulline), and was associated with CAC only in Tobit regression. Figure S7b illustrates the intracellular arginine/proline metabolism, one of the central pathways for the biosynthesis of arginine and proline from glutamine. Compared to controls, cases had higher levels of L-arginine, ornithine, citrulline, and succinic acid and lower levels of 4-hydroxy-L-proline. None of the amino acids were associated with CAC after adjusting for
demographics, but their relative-betweenness centrality resulted in a high pathway impact. Assuming intracellular metabolism of these amino acids mirrors their extracellular metabolism and serum levels, their directionality seems to suggest that there was an increased synthesis of L-arginine and decreased metabolism of proline in participants with CAC.

HD patients with CAC might have increased synthesis of L-arginine and decreased metabolism of proline to counteract the calcifying milieu. There are few studies that examined the effect of arginine on cardiovascular calcification. In an adenine-induced model of renal failure and arterial calcification, dietary L-arginine supplementation suppressed arterial calcification in 7 out of 10 rats. The exact mechanism of how arginine may attenuate arterial calcification is unclear. In vitro, L-arginine attenuated precipitation of calcium and phosphate. Arginine is a precursor for nitric oxide (NO; Figure S7b), and its effect on arterial calcification may be mediated by NO. To the best of our knowledge, no animal study has examined the effect of L-proline on arterial calcification. In cell culture, L-proline inhibited the apoptosis of human vascular smooth muscle cells induced by calcium and phosphate. Patients with CKD can develop both intimal calcification, which is an indicator of atherosclerosis, and medial calcification, which is characterized by diffuse calcium and phosphate deposition. Unfortunately, CT scan cannot differentiate between intimal and medial calcification. Altered arginine/proline metabolism may reflect atherosclerosis rather than mineralization of calcium and phosphate. In rabbits fed with high cholesterol diet, arginine supplementation limited the development of atherosclerosis. In our exploratory analyses, we found that arginine/proline metabolism was positively correlated with serum phosphorous, FGF-23, osteoprotegerin, and dp-ucMGP. Arginine/proline metabolism may serve as a scaffold on which a variety of regulatory mechanisms for arterial calcification are integrated. Although controversial, arginine supplementation has been used in a wide variety of conditions such as hypertension and erectile dysfunction. As arginine/proline metabolic pathway could be a potential therapeutic target, the role of arginine in arterial calcification merits further investigation.
Our study has several limitations. Because metabolites and CAC were both measured at baseline, the temporal relationship between serum metabolites and CAC could not be studied, thus limiting the inference of potential causal relationships. We were not able to address all potential or unmeasured confounders. Diabetes and hypertension are major potential confounders because both are associated with arterial calcification and metabolic dysregulation.\textsuperscript{12,42,43} All participants in our study were incident to HD, non-diabetic, and diagnosed with hypertension. Liver and residual renal function may influence the association of serum metabolites with CAC. Unfortunately, liver function tests were not available in PACE and only approximately a third of study population had available data on residual renal function. Due to the limitation in sample size, we did not include residual renal function in our analyses. Last, the study cohort is not representative of the ESRD population in the United States and may need validation in a larger and more diverse ESRD population.

Our study has several strengths. First, we studied the key metabolites and pathways in detail in regard to their relationships with serum markers of mineral metabolism, inflammation, and circulating inhibitors of calcification. Our findings provide an important framework for future studies that investigate the roles of these novel metabolites and pathway on CAC. Second, all samples were collected after 8 hours of fasting, thus minimizing the effects of diet on metabolites.\textsuperscript{44} Third, by focusing on incident HD patients who had similar characteristics including the status of diabetes and hypertension, we minimized potential residual confounding.

In this nested case-control study of non-diabetic, incident HD patients, we used a relatively unbiased approach of metabolomics profiling, and identified three novel metabolites (chenodeoxycholic, deoxycholic, and glycolithocholic acids) and one pathway (arginine/proline metabolism) that were associated with CAC. Chenodeoxycholic, deoxycholic, and glycolithocholic acids may be potential serum biomarkers for CAC, while arginine and proline metabolism may emerge as a new pathway in the pathogenesis of CAC and could be a potential treatment target. Our findings provide new insight
into pathophysiology of CAC in patients on HD and lay an important framework for future studies that investigate the roles of bile acids and arginine/proline metabolism on CAC.

DISCLOSURES

DAB is a consultant for Relyspa/Vifor/Fresenius, Amgen, Sanofi/Genzyme, and Tricida and has an equity interest in Amgen and Tricida. All remaining authors have nothing to disclose.

FUNDING

The research described was supported by NIH/National Center for Advancing Translational Science (NCATS) Einstein-Montefiore CTSA Grant Number UL1 TR002556. The PACE Study was funded by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK, R01 DK072367), National Kidney Foundation of Maryland, and by NIDDK (DK090070). WC is supported by the American Society of Nephrology Carl W. Gottschalk Research Grant and NIDDK (K23 DK114476). DAB is supported by NIDDK (R01 DK075462). RSP is supported by the Canada Research Chair in chronic kidney disease epidemiology. DFR-B is supported by the American Heart Association Career Development Award 19CDA34660217. NESS is supported by grants from NIH/NHLBI (R01 HL133861 and R01 HL149921), American Heart Association (19TPA34890070), and the Irma T. Hirschl-Monique Weill-Caulier Charitable Trust. Stable Isotope and Metabolomics Core Facility of the Diabetes Research and Training Center of the Albert Einstein College of Medicine is supported by NIH/NCI grant P60DK020541.

ACKNOWLEDGMENTS

We thank participants, nephrologists, and staff of the DaVita and MedStar dialysis units in the Baltimore region who contributed to the PACE study.

AUTHOR CONTRIBUTIONS
W Chen: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Validation; Visualization; Writing - original draft; Writing - review and editing

J Fitzpatrick: Data curation; Methodology; Writing - review and editing

S Sozio: Conceptualization; Data curation; Funding acquisition; Writing - review and editing

B Jaar: Data curation; Writing - review and editing

M Estrella: Data curation

D Riascos-Bernal: Conceptualization; Supervision; Writing - review and editing

T Tong Wu: Formal analysis; Validation; Writing - review and editing

Y Qiu: Investigation; Methodology; Project administration; Writing - review and editing

I Kurland: Investigation; Methodology; Project administration

R Dubin: Investigation; Methodology; Writing - review and editing

Y Chen: Resources; Supervision; Writing - review and editing

R Parekh: Conceptualization; Data curation; Investigation; Methodology; Resources; Supervision; Writing - review and editing

D Bushinsky: Conceptualization; Methodology; Supervision; Writing - review and editing

N Sibinga: Conceptualization; Methodology; Supervision; Writing - original draft; Writing - review and editing

Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.
SUPPLEMENTAL MATERIAL

Table of Contents

Supplemental Methods

Table S1. Comparing baseline characteristics of study population with unselected participants
Table S2. Summary of metabolites that differed between cases and controls
Table S3. Tobit regression of log-transformed CAC score
Table S4. Metabolites analyzed in arginine/proline metabolism, urea cycle and bile acid synthesis
Table S5. Correlations of bile acid synthesis and arginine/proline metabolism pathways and their key metabolites with serum markers of mineral metabolism (Ca, Phos, Mg, PTH, FGF23, klotho), circulating inhibitors of calcification (osteoprotegerin, dp-ucMGP, fetuin-A, CPP2, T_{50}) and inflammation (CRP).

Figure S1. a) Distribution of coefficients of variation (CVs) for all metabolites measured (m=452). b) Pathways and number of metabolites (m=247) analyzed.

Figure S2. Volcano plot identified metabolites that differed between cases and controls.

Figure S3. Synchronized 3-dimensional plot of partial least squares-discriminant analyses. Cases are in green and controls are in red.

Figure S4. Top 15 serum metabolites listed by their variable importance in projection (VIP) scores using partial least squares discriminant analyses.

Figure S5. Correlation among 3 bile acids.

Figure S6. Scree plots of arginine/proline metabolism (m=15), urea cycle (m=10) and bile acid synthesis (m=12).

Figure S7. Bile acid synthesis and intracellular arginine and proline metabolism, simplified version.

References
REFERENCES


## Table 1. Baseline participant characteristics by CAC status

<table>
<thead>
<tr>
<th></th>
<th>Total (n=99)</th>
<th>Control (CAC=0, n=48)</th>
<th>Case (CAC&gt;100, n=51)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, year</td>
<td>53 ± 16</td>
<td>42 ± 12</td>
<td>64 ± 13</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>38 (38)</td>
<td>19 (40)</td>
<td>19 (37)</td>
<td>0.81</td>
</tr>
<tr>
<td>African American, n (%)</td>
<td>71 (72)</td>
<td>45 (94)</td>
<td>26 (51)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Body Mass Index, kg/m²</td>
<td>27 ± 6</td>
<td>28 ± 7</td>
<td>27 ± 5</td>
<td>0.66</td>
</tr>
<tr>
<td>History of smoking, n (%)</td>
<td>67 (68)</td>
<td>30 (63)</td>
<td>37 (73)</td>
<td>0.29</td>
</tr>
<tr>
<td>Education less than grade 11, n (%)</td>
<td>44 (44)</td>
<td>24 (50)</td>
<td>20 (39)</td>
<td>0.50</td>
</tr>
<tr>
<td>Coronary artery disease, n (%)</td>
<td>27 (27)</td>
<td>4 (8)</td>
<td>23 (45)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Medication use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAAS blockade, n (%)</td>
<td>41 (46)</td>
<td>23 (58)</td>
<td>18 (36)</td>
<td>0.04</td>
</tr>
<tr>
<td>Calcium-based phosphate binder, n (%)</td>
<td>37 (37)</td>
<td>15 (31)</td>
<td>22 (43)</td>
<td>0.22</td>
</tr>
<tr>
<td>Vitamin D therapy, n (%)</td>
<td>69 (70)</td>
<td>34 (71)</td>
<td>35 (69)</td>
<td>0.81</td>
</tr>
<tr>
<td>Statin use, n (%)</td>
<td>27 (27)</td>
<td>6 (13)</td>
<td>21 (42)</td>
<td>0.001</td>
</tr>
<tr>
<td>Single-pool Kt/V</td>
<td>1.8 ± 0.3</td>
<td>1.8 ± 0.3</td>
<td>1.9 ±0.3</td>
<td>0.08</td>
</tr>
<tr>
<td>Dialysis Access, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.31</td>
</tr>
<tr>
<td>AVF</td>
<td>32 (32)</td>
<td>12 (25)</td>
<td>20 (39)</td>
<td></td>
</tr>
<tr>
<td>AVG</td>
<td>5 (5)</td>
<td>3 (6)</td>
<td>2 (4)</td>
<td></td>
</tr>
<tr>
<td>Catheter</td>
<td>62 (63)</td>
<td>33 (69)</td>
<td>29 (57)</td>
<td></td>
</tr>
<tr>
<td>Markers of mineral metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum calcium, mg/dL</td>
<td>8.7 ± 0.7</td>
<td>8.7 ± 0.7</td>
<td>8.7 ± 0.7</td>
<td>0.82</td>
</tr>
<tr>
<td>Serum phosphorous, mg/dL</td>
<td>5.3 ± 1.2</td>
<td>5.4 ± 1.4</td>
<td>5.2 ± 1.0</td>
<td>0.37</td>
</tr>
<tr>
<td>Serum magnesium, mEq/L</td>
<td>1.8 ± 0.2</td>
<td>1.8 ± 0.2</td>
<td>1.7 ± 0.2</td>
<td>0.36</td>
</tr>
<tr>
<td>Intact parathyroid hormone, pg/mL</td>
<td>400 (270-581)</td>
<td>402 (293-583)</td>
<td>387 (252-580)</td>
<td>0.49</td>
</tr>
<tr>
<td>FGF23, RU/ml</td>
<td>777 (222-1310)</td>
<td>819 (171-1284)</td>
<td>754 (316-1500)</td>
<td>0.53</td>
</tr>
<tr>
<td>Soluble klotho, pg/mL</td>
<td>356 (268-489)</td>
<td>397 (285-524)</td>
<td>306 (263-453)</td>
<td>0.05</td>
</tr>
<tr>
<td>Markers of cardiovascular calcification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fetuin-A, mg/L</td>
<td>533 ± 193</td>
<td>547 ± 192</td>
<td>519 ± 195</td>
<td>0.48</td>
</tr>
<tr>
<td>Dp-ucMGP, ppm</td>
<td>1441 (967-2042)</td>
<td>1258 (783-1800)</td>
<td>1734 (1101-2437)</td>
<td>0.02</td>
</tr>
<tr>
<td>Osteoprotegerin, pmol/L</td>
<td>10.1 ± 4.8</td>
<td>8.2 ± 3.4</td>
<td>11.8 ± 5.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CPP2 size, nm</td>
<td>289 (192-394)</td>
<td>289 (187-394)</td>
<td>289 (207-394)</td>
<td>0.97</td>
</tr>
<tr>
<td>T₅₀, min</td>
<td>307 (227-383)</td>
<td>306 (252-378)</td>
<td>315 (213-398)</td>
<td>0.71</td>
</tr>
<tr>
<td>C-reactive protein, µg/mL</td>
<td>4.5 (1.7-11.4)</td>
<td>3.4 (1.5-8.7)</td>
<td>5.3 (2.1-17.1)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Abbreviations: RAAS, renin-angiotensin-aldosterone system; AVF, arteriovenous fistula; AVG, arteriovenous graft; dp-ucMGP, Dephosphorylated and uncarboxylated matrix Gla protein; FGF23, fibroblast growth factor-23; CPP2, secondary calciprotein particle; T₅₀, transformation time from primary to secondary calciprotein particle. Note: If normally distributed, values for continuous variables
with normal distribution are provided as mean ± standard deviation and were tested with two-sample t-test. Otherwise, they are provided as median (interquartile range) and were tested using Mann-Whitney test. Categorical variables are presented as absolute number with percentage and were tested with chi-square test.
Table 2. Logistic regression of CAC status with chenodeoxycholic, deoxycholic, and glycolithocholic acids

<table>
<thead>
<tr>
<th></th>
<th>Chenodeoxycholic acid</th>
<th>Deoxycholic acid</th>
<th>Glycolithocholic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
<td>Demographics-adjusted</td>
<td>Fully adjusted</td>
</tr>
<tr>
<td>Case (n)</td>
<td>OR (95% CI)</td>
<td>p</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Continuous*</td>
<td>1.17 (1.04 – 1.33)</td>
<td>0.01</td>
<td>1.18 (1.01 - 1.38)</td>
</tr>
<tr>
<td>Tertile 1</td>
<td>12</td>
<td>Reference</td>
<td>reference</td>
</tr>
<tr>
<td>Tertile 2</td>
<td>13</td>
<td>1.14 (0.42 - 3.08)</td>
<td>0.80</td>
</tr>
<tr>
<td>Tertile 3</td>
<td>26</td>
<td>6.5 (2.17 - 19.43)</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>1.18 (1.04 – 1.34)</td>
<td>0.01</td>
<td>1.17 (1.00 – 1.37)</td>
</tr>
<tr>
<td>Tertile 1</td>
<td>11</td>
<td>Reference</td>
<td>reference</td>
</tr>
<tr>
<td>Tertile 2</td>
<td>14</td>
<td>1.47 (0.54 - 4.01)</td>
<td>0.45</td>
</tr>
<tr>
<td>Tertile 3</td>
<td>26</td>
<td>7.43 (2.46 - 22.42)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>1.05 (1.01 – 1.09)</td>
<td>0.01</td>
<td>1.07 (1.00 – 1.13)</td>
</tr>
<tr>
<td>Tertile 1</td>
<td>11</td>
<td>Reference</td>
<td>reference</td>
</tr>
<tr>
<td>Tertile 2</td>
<td>16</td>
<td>1.88 (0.70 – 5.09)</td>
<td>0.21</td>
</tr>
<tr>
<td>Tertile 3</td>
<td>24</td>
<td>5.33 (1.86 – 15.30)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

*Units for continuous variables were 1 unit in relative intensity for chenodeoxycholic acid and deoxycholic acid and 0.01 unit in relative intensity for glycolithocholic acid.

n=33 in each tertile

Demographics included age, sex, and race.

Full models were adjusted for demographics, coronary artery disease, use of calcium-based phosphate binder, RAAS blockage and statin for all 3 bile acids.
Table 3. Impact and significance of arginine/proline metabolism, urea cycle and bile acid synthesis

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Metabolites in SMPD pathway (m)</th>
<th>Metabolites analyzed (m)</th>
<th>Pathway impact</th>
<th>FDR adjusted p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arginine and proline metabolism</td>
<td>53</td>
<td>15</td>
<td>0.21</td>
<td>5.8 x 10^{-4}</td>
</tr>
<tr>
<td>Urea cycle</td>
<td>29</td>
<td>10</td>
<td>0.23</td>
<td>0.002</td>
</tr>
<tr>
<td>Bile acid synthesis</td>
<td>65</td>
<td>12</td>
<td>0.01</td>
<td>0.005</td>
</tr>
</tbody>
</table>

For pathway enrichment analyses, we used Small Molecule Pathway Database (SMPD) and included 34 metabolites that had a variable importance in projection (VIP) score >1 and p-value <0.05. Global tests were used to assess the significance of pathways while topology analysis with relative-betweenness centrality was used to assess pathway impact. Abbreviation: SMPD, Small Molecule Pathway Database; FDR, false discovery rate.
Table 4. Logistic regression of CAC status with arginine/proline metabolism, urea cycle and bile acid synthesis*

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Unadjusted</th>
<th>Demographics-adjusted</th>
<th>Fully adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>p</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Arginine/proline metabolism</td>
<td>1.43 (1.06 - 1.92)</td>
<td>0.02</td>
<td>1.83 (1.06 – 3.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.84 (1.04 – 3.27)</td>
</tr>
<tr>
<td>Urea cycle</td>
<td>1.93 (1.34 - 2.80)</td>
<td>&lt;0.001</td>
<td>1.58 (0.93 – 2.69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.66 (0.94 – 2.91)</td>
</tr>
<tr>
<td>Bile acid synthesis</td>
<td>0.99 (0.82 - 1.20)</td>
<td>0.92</td>
<td>1.09 (0.84 – 1.41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.64 (0.97 – 2.75)</td>
</tr>
</tbody>
</table>

*per 1 unit higher in the first principal component score of each pathway
Demographics included age, sex, and race.
Full models were adjusted for demographics and statin use for arginine/proline metabolism and urea cycle pathways, and adjusted for demographics, coronary artery disease, use of calcium-based phosphate binder, RAAS blockage and statin for bile acid synthesis pathway.
FIGURE LEGEND

Figure 1. Overview for study population, analyses and results (n denotes number of participants and m denotes number of metabolites). We conducted a case control study in the PACE cohort. Cases were non-diabetic participants with a CAC score >100 (n=51), while controls were those with a CAC score of 0 (n=48). Metabolites (m=247) and pathway scores were compared between cases and controls. We identified 3 metabolites and 1 pathway that remained significant after adjusting for age and race.

Note: a There were 36 metabolites with variable importance in projection (VIP) score >1 and p-value <0.05. Two of which do not have human metabolome database identification (HMDB ID): lumazine (food component) and 3-hydroxybenzaldehyde (chemical), and they were not included in the pathway analysis. b Adjusted for age and race using each metabolite. c Adjusted for age and race using the first principal component of the pathways.

Figure 2. Box plots and individual data points of selected metabolites in arginine and proline metabolism, urea cycle and bile acid synthesis. Mann-Whitney U tests were used to compare metabolites between cases (CAC>100) and controls (CAC=0). Metabolites with a p-value <0.05 are presented above. L-arginine (fold difference=1.14, p=0.03), ornithine (fold difference=1.17, p=0.008) and citrulline (fold difference=1.21, p=0.01) are in both arginine and proline metabolism and urea cycle. 4-hydroxy-L-proline (fold difference=0.74, p=0.04) and succinic acid (fold difference=1.44, p=0.005) are only in arginine and proline metabolism, while L-glutamine (fold difference=1.12, p=0.02) is only in the urea cycle pathway. Chenodeoxycholic acid (fold difference=2.33, p=0.002), deoxycholic acid (fold difference=2.32, p=0.001) and glycolithocholic acid (fold difference=2.44, p=0.004) are in bile acid synthesis pathway. Note: Box plots represent median and interquartile range.

Figure 3. Correlations of bile acid synthesis and arginine/proline metabolism pathways and their key metabolites with serum markers of mineral metabolism (Ca, Phos, Mg, PTH, FGF23, klotho), circulating inhibitors of calcification (osteoprotegerin, dp-ucMGP, fetuin-A, CPP2, T50) and inflammation (CRP). The first principal component scores were used to represent the pathways. Heat map shows Spearman’s rank correlations. Correlation strength is represented by color bar (red: positive correlation; white: no correlation; blue: negative correlation). *p<0.05; **p<0.005. Abbreviations: Ca, calcium; Phos, phosphorous; Mg, magnesium; PTH, intact parathyroid hormone; FGF-23, fibroblast growth factor-23; klotho, soluble klotho; dp-ucMGP, Dephosphorylated and uncarboxylated matrix Gla protein; CPP2, secondary calciprotein particle; T50, half maximal transformation of primary to secondary calciprotein particle; CRP, C-reactive protein.
Figure 1. Overview for study population, analyses and results. Note: n denotes number of participants and m denotes number of metabolites.

There were 36 metabolites with variable importance in projection (VIP) score >1 and p-value <0.05.

Adjusted for age, sex, and race using each metabolite.

Adjusted for age, sex, and race using the first principal component of the pathways.
Figure 2. Box plots and individual data points of selected metabolites in arginine and proline metabolism, urea cycle and bile acid synthesis. Note: Box plots represent median and interquartile range.
Figure 3. Correlations of bile acid synthesis and arginine/proline metabolism pathways and their key metabolites with serum markers of mineral metabolism (Ca, Phos, Mg, PTH, FGF23, klotho), circulating inhibitors of calcification (osteoprotegerin, dp-ucMGP, fetuin-A, CPP2, T50) and inflammation (CRP). Note: *p<0.05; **p<0.005; correlation strength is represented by color bar (red: positive correlation; white: no correlation; blue: negative correlation).