Clinica Chimica Acta 489 (2019) 242–248

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Metabolites related to eGFR: Evaluation of candidate molecules for GFR estimation using untargeted metabolomics

Titan S.M.[a,*], Venturini G.[b], Padilha K.[b], Tavares G.[a], Zatz R.[a], Bensenor I.[c], Lotufo P.A.[c], Rhee E.P.[d,e], Thadhani R.I.[d,f], Pereira A.C.[b]

[a] Nephrology Division, University of Sao Paulo Medical School, Sao Paulo, SP, Brazil
[b] Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo Medical School, Sao Paulo, SP, Brazil
[c] Epidemiological and Clinical Research Center, Universitary Hospital, Sao Paulo Univ, Sao Paulo, Brazil
[d] Division of Nephrology, Department of Medicine, Massachusetts General Hospital, Boston, MA, United States
[e] Division of Endocrinology, Department of Medicine, Massachusetts General Hospital, Boston, MA, United States
[f] Vice Dean of Research, Cedars-Sinai Medical Center, Los Angeles, CA, United States

ARTICLE INFO

Key-words:
Metabolomics
GFR
CKD

ABSTRACT

Background: Metabolomics can be used to identify novel metabolites related to renal function and that could therefore be used for estimating GFR. We evaluated metabolites replicated and related to eGFR in 3 studies (CKD and general population).

Methods: Metabolomics was performed by GC–MS. The Progredir Cohort (n = 454, class 3 and 4 CKD) was used as the derivation study and adjusted linear regression models on eGFR-CKDEPI were built. Bonferroni correction was applied for selecting metabolites to be independently validated in the Diabetic Nephropathy Study (n = 56, macroalbuminuric DN) and in the Baependi Heart Study (BHS, n = 1145, general population).

Results: In the Progredir Cohort, 72 metabolites where associated with eGFR. Of those, 11 were also significantly associated to eGFR in the DN Study and 8 in the BHS. Four metabolites were replicated and significantly associated to eGFR in all 3 studies: D-threitol, myo-inositol, 4-deoxierythronic acid and galacturonic acid. In addition, pseudouridine was strongly correlated to eGFR only in the 2 CKD populations.

Conclusions: Our results demonstrate metabolites that are potential biomarkers of renal function: D-threitol, myo-inositol, 4-deoxierythronic acid, galacturonic acid and pseudouridine. Further investigation is needed to determine their performance against otherwise gold-standard methods, most notably among those with normal eGFR.

1. Introduction

The kidney is involved in several biological functions, such as filtration, excretion, and metabolism of substances, in addition to blood pressure regulation, modulation of hemoglobin production, among others. Thus, assessment of renal function can be done by several approaches, involving one or more of the above mentioned kidney functions. Estimation of the glomerular filtration rate is the most widely used measure for diagnosis and evaluation of kidney disease, followed by urinary protein and/or albumin excretion. However, despite intense efforts in the last 40 years, our ability to accurately determine the GFR is still limited. While the use of formulas for estimating GFR based on creatinine and more recently using cystatin C has improved the accuracy and earlier diagnosis of CKD, these formulas present limitations [1–6]. Creatinine is largely influenced by muscle mass and cystatin C can also be influenced by other factors (inflammation, smoking, drugs, adiposity), besides being still of limited use due to availability and cost. In this sense, identification and understanding of novel markers of kidney function (which may be related to filtration and/or other kidney functions) might provide new insights on the biology of kidney disease and eventually generate new biomarkers of clinical outcomes.

Metabolomics is a tool that could be leveraged in this scenario. The untargeted approach (in which no particular metabolite is being specifically targeted for measurement) offers the opportunity of finding new molecules that might be related to a phenotype or outcome. Particularly in the setting of renal function measurement, metabolomics can be used to evaluate novel molecules that could potentially improve GFR estimation. Few studies have evaluated the association of metabolomics-derived metabolites to eGFR. Sekula et al. [7] showed serum metabolites associated with eGFR in the Caucasian population. In this

* Corresponding author.
E-mail address: smotitan@gmail.com (S.M. Titan).

https://doi.org/10.1016/j.cca.2018.08.037
Received 26 April 2018; Received in revised form 21 August 2018; Accepted 22 August 2018
Available online 25 August 2018
0009-8981/

important paper, the authors selected from KORA F4 Study and replicated in TwinsUK Registry (both with a general population profile and with approximately 5% of CKD participants) metabolites related to eGFR assessed by creatinine and cystatin C-based formulas. Moderate associations were found for 6 metabolites: C-mannosyltryptophan, pseudouridine, *N*-acetylalanine, erythronate, myo-inositol, and *N*-acetylcarnosine, with C-mannosyltryptophan and pseudouridine showing the highest correlations. Using the same metabolomics platform but in an African American population free of CKD, Yu et al. [8] reported 74 metabolites related to eGFR, with 3-indoxyl sulfate, *N*-acetylalanine, phenylacetylglutamine, betaine, carnitine and erythritol among the top associations. While bringing important insights to the field, these results also show how results can vary depending on factors such as race and eGFR range. More studies are needed and associations should be tested in other population profiles, including CKD patients and not restricted to Caucasian and African-American populations.

In the present study, we aimed to (1) identify metabolites related to baseline eGFR in a CKD cohort from Brazil (a country with highly admixed population), the Progredir Cohort; (2) test the replication of these findings in two other Brazilian cohorts, namely in another CKD population (Diabetic Nephropathy Study) and a sample from the general population (Baependi Heart Study).

2. Methods

2.1. Derivation study

The Progredir Cohort is an ongoing CKD cohort in Sao Paulo, Brazil. Details of methods and population recruitment have been published elsewhere [9]. Briefly, patients from the Hospital das Clínicas Outpatient Clinics, Sao Paulo, were invited to participate if age $>$ 30 y and at least 2 measurements of creatinine (with a minimum interval of 3 months) were $>$ 1.6 mg/dL for men and $>$ 1.4 mg/dL for women. Patients attending oncology, psychiatry, urology, HIV/AIDS, viral hepatitis and glomerulonephritis services were excluded. The remaining candidates were contacted by phone and invited to participate if no exclusion criteria (hospitalization or acute myocardial infarction in the last 6 months, autoimmune diseases, pregnancy, psychiatric diseases, ongoing chemo or immunosuppressive therapy, ongoing renal replacement therapy, glomerulonephritis, HIV/AIDS infection, hepatitis C and B, and any previous organ transplantation) were met. Recruitment took place from March 2012 to December 2013, and 454 participants were enrolled.

Each participant visited the research center for interviews and clinical exams according to standard protocols. Interview and clinical examination were performed by trained personal under strict quality control. Overnight fasting blood samples, 24-h and spot urine were collected. Urine and serum aliquots were prepared and stored in liquid nitrogen. Baseline fasting serum samples were used for metabolomics measurements. Glomerular filtration rate was estimated by the CKD-EPI Eq. [10] and albumin-to-creatinine ratio (ACR) was performed using morning spot urine. Diabetes was defined by any of the following criteria: previous medical history of diabetes, use of medication to treat diabetes, fasting plasma glucose $\geq$126 mg/dL, HbA1C $\geq$6.5%, or a 2-h plasma glucose $\geq$200 mg/dL. Hypertension was defined as previous medical history or use of anti-hypertensive medication. Laboratory measurements were determined using conventional techniques [9].

2.2. Replication studies

Two studies were used for replication. The Diabetic Nephropathy Study (DN Study) comprises 56 participants from a clinical trial designed to compare the association of dual therapy with ACE inhibitor plus angiotensin II receptor blocker versus monotherapy with ACE inhibitor on proteinuria progression [11]. Recruitment took place from 2005 to 2008. For this analysis, only baseline data (prior to intervention) of the 56 participants with macroalbuminuric diabetic nephropathy was used. Fasting plasma samples were stored at -20 °C and used for metabolomics.

The Baependi Heart Study (BHS) is a general population cohort set in a rural area in the state of Minas Gerais, Brazil. Details of population recruitment (2005–06) and baseline profile have been published [12]. Metabolomics was performed on baseline serum samples of 1145 participants from the 5-year follow-up visit (stored at -80 °C). Diabetes was defined as previous medical history of diabetes, use of medication to treat diabetes or fasting plasma glucose $\geq$126 mg/dL. Glomerular filtration rate was also estimated by the CKD-EPI equation in the two replication studies [10].

2.3. Ethics

The Progredir study was approved by two local Ethics Committees, the Universitary Hospital Ethics Committee (CEP-HU, Sao Paulo University) and the Ethic Committee for Analysis of Research Projects (CAPPesq, Hospital das Clínicas, Sao Paulo University). The DN study and the BHS were approved by the Ethic Committee for Analysis of Research Projects (CAPPesq, Hospital das Clínicas, Sao Paulo University). Written informed consent was obtained from all participants and all research was performed in accordance with the 2013 Helsinki Declaration principles.

2.4. Metabolomics

Metabolomics analyses were performed according to Oliver Fiehn [13], with few modifications. Samples (serum in the Progredir Cohort and in the BHS, plasma in the DN Study) were thawed on ice at 4 °C for 30–60 min. Metabolites from each aliquot of serum/plasma (70 μL for Progredir and DN Study and 100 μL for BHS) were extracted with a 300 μL solvent mixture of acetonitrile (methanol in BHS), isopropanol and deionized water (3:3:2 v/v) and spiked with a 5 μL internal standard solution (RTL). After vortexing for 15 s, the mixture was centrifuged for 15 min at 15,800 $\times$ g at 4 °C. Supernatant (320 μL) was transferred to a new microcentrifuge tube, followed by lyophilization in a Speedvac concentrator for 18 h. Subsequently, the residue was suspended in 50 μL methoxyamine in pyridine (Sigma-Aldrich) solution (40 mg/ml), 3 μL of FAME (Fatty acid methyl ester – Sigma-Aldrich) was added, and the mixture was vortexed for 3 min. This methoximation reaction was performed at room temperature for 16 h, followed by trimethylsilylation for 1 h adding 100 μL MSTFA (*N*-methyl-*N*-trimethylsilyltrifluoroacetamine) with 1% TMCS (trimethylchlorosilane) (Sigma-Aldrich). After derivatization, 1 μL of this derivative was used for Gas Chromatography Mass Spectrometry (GC–MS) analysis.

Two types of controls were done every day in GC–MS analysis, 3 times per day. We used a blank and a quality control (QC) samples to verify the existence of impurities in the reagents and equipment contamination, and to check the sensitivity and compliance of the injector system, respectively. Quality control samples were prepared using a pool of serum samples and blank samples were prepared following the same procedures as those performed for study samples, replacing the serum by deionized water.

One microliter of each sample was injected into an Agilent 7890B GC system operated in splitless mode. A DB5-MS + 10 m Duraguard capillary column (Agilent 122-5532G) where helium carrier gas flowed at a rate of 0.736 (Progredir and DN Study) and of 1.1 mL/min (BHS) as used for metabolite separation. The injector temperature was set at 250 °C. The column temperature was held at 60 °C for 1 min, and then increased to 310 °C at a rate of 10 °C/min during 37 min. The column effluent was introduced into the ion source of an Agilent 5977A mass selective detector. The detector operated in the electron impact ionization mode (70 eV) and mass spectra were recorded after a solvent delay of 6.5 min with 3 scans per second. The MS quadrupole temperature was set at 180 °C and the ion source temperature was set at

Table 1
Descriptive baseline characteristics of participants in the three studies.

	Progredir Cohort	DN Study	BHS
	n = 454	n = 56	n = 1145
Age (years; mean/std)	67.5 (11.9)	58.1 (10.2)	48 (17)
Sex (men; n/%)	287 (63.2%)	33 (61.1%)	481 (42%)
Race (white; n/%)	300 (66.1%)	23 (41.1%)	893 (78%)
Hypertension (n/%)	409 (90.1%)	52 (92.9%)	335 (31%)
Diabetes (n/%)	257 (56.6%)	56 (100%)	82 (7%)
Previous myocardial infarction (n/%)	147 (32.4%)	16 (29.6%)	18 (1.5%)
Previous stroke (n/%)	73 (16.1%)	7 (12.5%)	10 (1%)
Smoking (current or previous; n/%)	269 (59.3%)	28 (50%)	388 (34%)
SBP (mmHg; mean/std)	140 (24)	149 (23)	125 (16)
DBP (mmHg; mean/std)	76 (13)	81 (14)	76 (10)
Body-mass index (mean/std)	29.4 (5.4)	24.1 (3.4)	25.8 (5.2)
eGFR-CKDEPI (mL/min/1.73m^2; mean/std)	38.4 (14.6)	46.3 (21.5)	94.8 (19.8)
Urea (mg/dL; median/std)	75 (29)	61 (25)	27 (9)
Albuminuria (mg/dL; median/IQR)	80 (16–640)	–	–
24 h proteinuria (g/d/1.73m^2; median/IQR)	0.17 (0.08–0.61)	2.92 (1.61–5.12)	–
Glycemia (mg/dL; mean/std)	119 (45)	151 (66)	93 (21)
Glycated hemoglobin (%; mean/std)	6.7 (1.4)	8.3 (1.9)	5.7 (0.8)
Total cholesterol (mg/dL mean/std)	169 (40)	190 (58)	195 (40)
LDL-cholesterol (mg/dL; mean/std)	91 (32)	106 (52)	122 (34)
HDL-cholesterol (mg/dL; mean/std)	46 (14)	50 (14)	48 (12)
Triglycerides (mg/dL; median/IQR)	142 (99–192)	162 (103–270)	132 (80)
Bicarbonate (mmol/L; mean/std	25.6 (2.9)	24.9 (3.1)	114 (86–156)
Hemoglobin (g/dL; mean/std)	13.1 (1.9)	12.8 (1.9)	–
Hematocrit (%; mean/std)	39.5 (6.0)	38.7 (5.3)	–
Albumin (mg/dL; mean/std)	4.3 (0.3)	3.3 (0.5)	–
Phosphorus (mg/dL; mean/std)	3.6 (0.6)	4.1 (0.7)	–
Calcium (mg/dL; mean/std)	9.6 (0.6)	9.3 (0.8)	–
Parathormone (pg/mL; median/IQR)	93 (64–143)	60 (37–95)	–
Diabetic retinopathy (yes among diabetics; n/%)	85 (33.1%)	44 (81.5%)	–

DN Study Diabetic Nephropathy Study; BHS, Baependi Heart Study.

280 °C. Each sample was analyzed in 3 technical replicates in the Progredir Cohort and DN Study. Only one measurement was performed in the BHS.

To achieve higher specificity in the identification of metabolites acquired by the GC–MS technique, we used both the measurement of the retention time and EI spectra. Absolute retention times were locked to the internal standard d27-myristic acid 3 mg/mL (Product # 366889; Sigma- Aldrich; RT of the locking standard is 16.752 min) using the RTL system provided in Agilent's MassHunter software. Retention time locking reduces the retention time variation and is employed within the Agilent Fiehn GC–MS Metabolomics RTL Library for identification of approximately 1447 common metabolites (Metabolomics GCMS # G166766A). The metabolites that were not identified in Agilent Fiehn GC–MS Library were identified using the NIST Library (2014).

2.5. Chemicals and reagents

Acetonitrile (LC-MS Chromasolv), 2-propanol (LC-MS Ultra Chromasolv), Pyridine (LC-MS Chromasolv Plus, Methoxyamine hydrochloride, FAME (fatty acid methyl ester), MSTFA (*N*-methyl-*N*-trimethylsilyltrifluoroacetamine), TMCS (trimethylchlorosilane), Memhanol (UHPLC), and myristic acid D27 (all from Sigma-Aldrich).

2.6. Data processing and statistical analysis

Identification of compounds was made comparing the mass spectra and retention time (RT) of all detected compounds with the Agilent Fiehn GC–MS Metabolomics RTL Library (ver A.02.02) and the National Institute of Standards and Technology (NIST) library 11 (2014) using Unknowns - Agilent MassHunter Workstation Quantitative Analysis (ver B.06.00). Missing values were considered as non-available. Mean values among technical triplicates were calculated (except for the BHS, which had only one measurement per sample). We excluded metabolites that were not represented in at least 50% of the population in all three studies, leaving 293 metabolites in the Progredir Study, 228 in the DN Study and 319 in the BHS.

Initially, we evaluated the association of metabolites (log2) with eGFR by linear regression in models adjusted only for batch, sex, age and adjusted for batch, sex, age, diabetes, systolic blood pressure (SBP), and smoking in the 3 studies. We then selected metabolites that reached significance at Bonferroni correction (0.05/293) in the Progredir Study and that were replicated and presented nominal significant associations to eGFR in the validation studies ($p = .05$). Spearman's correlation coefficients were calculated for eGFR and metabolites (log2) and scatter plots were built. Statistical analyses were performed using SPSS® (ver 20.0) and R (https://www.r-project.org/).

3. Results

In Table 1, we show the descriptive data of the three studies. In the Progredir Cohort, mean age was 67 years, mean eGFR was 38.4 (14.6) mL/min/1.73m^2 and approximately 57% of participants were diabetic and 90% hypertensive. In the DN Study, mean age was 58 y, mean eGFR was 46.3 (21.3) mL/min/1.73m^2, and most participants had macroalbuminuria and hypertension. The population in the BHS was younger with mean age of 48 years, presenting 7% of diabetics and 30% of hypertensive individuals. As expected considering its general population profile, mean eGFR was 94.8 (19.8) mL/min/1.73m^2.

First, we selected metabolites that were significantly associated with renal function in the Progredir Study applying a Bonferroni correction, after adjustment for sex, age and batch (model 1) and sex, age, diabetes, SBP, smoking and batch (model 2). These metabolites are shown in Table 2, with very similar results for models 1 and 2. We next tested metabolites related to eGFR applying the same adjustments in the 2 replication studies (Supplementary Tables 1 and 2). In Table 3, we show metabolites associated to eGFR in the Progredir Study that were

Table 2
Metabolites related eGFR-CKDEPI among 454 CKD participants in the Progredir Study (adjusted for age, sex, SBP, DM, smoking and batch).

Metabolite (log base 2)	**Model 1** Beta	Bonferroni p	**Model 2** Beta	Bonferroni p
Pseudouridine	-12.80	1.11E-62	-12.81	2.62E-62
D-threitol	-13.58	1.82E-60	-13.55	1.63E-59
Unidentified 1 (m/z 405)	-12.32	1.60E-56	-12.30	9.88E-56
Myo-inositol	-14.64	4.33E-53	-14.63	1.17E-52
Butanoic acid	-9.45	1.51E-44	-9.47	2.32E-44
5-hydroxyindol	-6.98	7.18E-32	-6.97	2.92E-31
Galactonic acid	-7.30	1.27E-30	-7.28	7.21E-30
Xylitol	-7.35	2.12E-28	-7.31	1.04E-27
2-O-Glycerol-α-d-galactopyranoside	-7.28	4.73E-24	-7.28	1.64E-23
Lactose	-5.06	5.36E-23	-5.15	5.87E-23
Trans-aconitic acid	-10.60	8.96E-21	-10.61	1.57E-20
4-Deoxyerythronic acid	-6.42	1.52E-19	-6.42	2.23E-19
3-indolelactic acid	-8.85	1.11E-18	-8.89	1.30E-18
Ribonic acid	-7.90	1.83E-17	-7.85	4.67E-17
Threonic acid	-9.33	1.19E-16	-9.46	4.56E-17
(S)-3,4-Dihydroxybutyric acid	-9.94	1.92E-16	-9.82	9.51E-16
Quinic acid	-5.30	5.04E-16	-5.48	4.65E-17
Arabinose	-7.95	8.83E-16	-7.83	5.24E-15
Phenol	-4.67	1.38E-15	-4.67	1.57E-15
Creatinine	-7.23	1.67E-15	-7.38	1.76E-15
D-mannitol	-3.41	5.26E-14	-3.47	3.80E-14
4-Deoxyerythronic acid	-6.63	5.52E-14	-6.81	8.89E-15
Tyrosine 1	6.04	4.50E-13	6.05	1.25E-12
Glycerol	6.86	1.97E-11	6.86	2.11E-11
Galacturonic acid 1	-6.13	4.20E-11	-6.18	5.62E-11
Gluconic acid 2	-6.60	4.54E-11	-6.71	2.20E-11
L-cystine 2	-5.97	3.25E-10	-5.96	5.66E-10
L-cystine 3	-3.51	2.13E-09	-3.45	6.18E-09
2-tert-Butyl-4-methoxyphenol	-7.54	4.56E-09	-7.72	1.48E-09
2-Propenoic acid	-12.38	5.66E-09	-12.83	1.51E-09
Xanthine	-3.37	2.45E-08	-3.40	1.76E-08
Ribose	-4.24	3.89E-08	-4.23	5.52E-08
Citrulline 2	-5.32	4.21E-08	-5.40	2.88E-08
Gluconic acid	-7.57	9.80E-08	-7.53	1.61E-07
L-glutamine 1	-5.77	1.11E-07	-6.05	4.13E-08
Acetamide	-7.34	1.20E-07	-7.42	8.06E-08
Glycine	-4.26	1.62E-07	-4.35	8.23E-08
Diethanolamine	-3.80	2.65E-07	-3.80	3.09E-07
O-phosphocolamine	-6.66	3.83E-07	-6.57	8.18E-07
Unidentified 2 (m/z 302)	-5.20	1.01E-06	-5.58	1.08E-07
Acetohydroxamic acid	-6.36	1.65E-06	-6.31	2.07E-06
L-glutamine 2	-4.50	1.82E-06	-4.65	6.71E-07
Galactitol	-4.69	1.87E-06	-4.51	9.21E-06
4-hydroxyphenylacetic acid	-3.35	2.55E-06	-3.26	7.98E-06
L-lysine 2	-2.92	4.23E-06	-2.95	3.26E-06
Hypoxanthine	-4.34	8.93E-06	-4.38	7.89E-06
p-Cresol glucuronide	-3.14	2.37E-05	-2.99	1.86E-04
Unidentified 3 (m/z 273)	-9.82	2.89E-05	-9.70	4.46E-05
p-cresol	-3.95	3.29E-05	-3.87	1.10E-04
Isotridecanol	-4.30	1.48E-04	-4.22	2.89E-04
Pyroglutamic acid	-6.19	1.65E-04	-6.24	1.75E-04
L-tyrosine 2	-3.74	4.39E-04	-3.65	8.36E-04
Uridine	-6.68	6.56E-04	-6.63	6.77E-04
4-Nitropyrazole	-6.01	6.74E-04	-5.93	1.05E-03
2-Pyrrolidinone	-6.54	9.09E-04	-6.44	1.43E-03
Dihydrouracil	-6.67	1.07E-03	-6.36	3.13E-03
2-hydroxybutyric acid	-5.67	1.39E-03	-5.78	9.83E-04
Tyrosine 2	-2.98	2.38E-03	-3.65	8.36E-04
2-Ethylhydracrylic acid	-4.44	2.65E-03	-4.25	7.97E-03
Glycine	-4.31	2.67E-03	-4.21	4.21E-03
Eicosapentaenoic acid	-2.78	3.08E-03	-2.93	9.23E-04
trans-4-hydroxy-L-proline 1	-2.97	3.42E-03	-2.94	4.72E-03
Unidentified 4 m/z 296	-4.88	4.23E-03	-4.57	1.74E-02
Arachidic acid	-4.34	6.93E-03	-4.35	7.37E-03
L-glutamic acid 2	3.46	8.68E-03	3.45	9.06E-03
4-Bromo-1-butanol	-5.36	1.13E-02	-5.34	1.24E-02
Unidentified m/z 334	-6.44	1.15E-02	-6.35	1.79E-02
Glycolic acid	-5.70	2.27E-02	-5.66	2.65E-02
Citric acid	-5.24	2.39E-02	-5.30	2.11E-02
(Z)13-Docosenoic acid	-7.48	2.75E-02	-7.74	1.73E-02
5alpha-Cholestanol	-4.61	3.07E-02	-4.59	3.46E-02
3-indoleacetic acid	-3.17	5.25E-02	-3.21	4.12E-02

Model 1 is adjusted for sex, age, and batch.
Model 2 is adjusted for sex, age, DM, SBP, smoking and batch.
Color scale ranging from green (positive associations) to red (negative associations).

Table 3
Metabolites replicated in the two validation studies and with significant associations to eGFR (adjusted for age, sex, SBP, DM, smoking and batch).

Replicated and related to eGFR in the DN Study			Replicated and related to eGFR in the Baependi Heart Study			Replicated and related to eGFR in the 3 studies
Metabolite	Beta	p	Metabolite	Beta	p	
Pseudouridine	−25.9	7.89E-06	D-threitol	−2.95	9.59E-08	D-threitol
D-threitol	−21.8	8.30E-05	Glycine	−1.95	0.001	Myo-inositol
Myo-inositol	−16.2	0.0001	4-Deoxyerythronic acid	−1.69	0.001	4-Deoxyerythronic acid
L-glutamic acid	22.3	0.01	L-Lysine	−2.47	0.002	Galacturonic acid
Galactonic acid	−12.9	0.02	L-glutamine	−1.69	0.002	
Butanoic acid	−22.0	0.02	Galacturonic acid	−1.93	0.003	
4-Deoxyerythronic acid	−12.9	0.02	Myo-inositol	−1.24	0.02	
Unidentified m/z 334	−21.8	0.04	Citric acid	−1.26	0.04	
(Z)13-Docosenoic acid	26.2	0.04				
Galacturonic acid	−10.3	0.04				
5alpha-Cholestanol	−24.5	0.04				

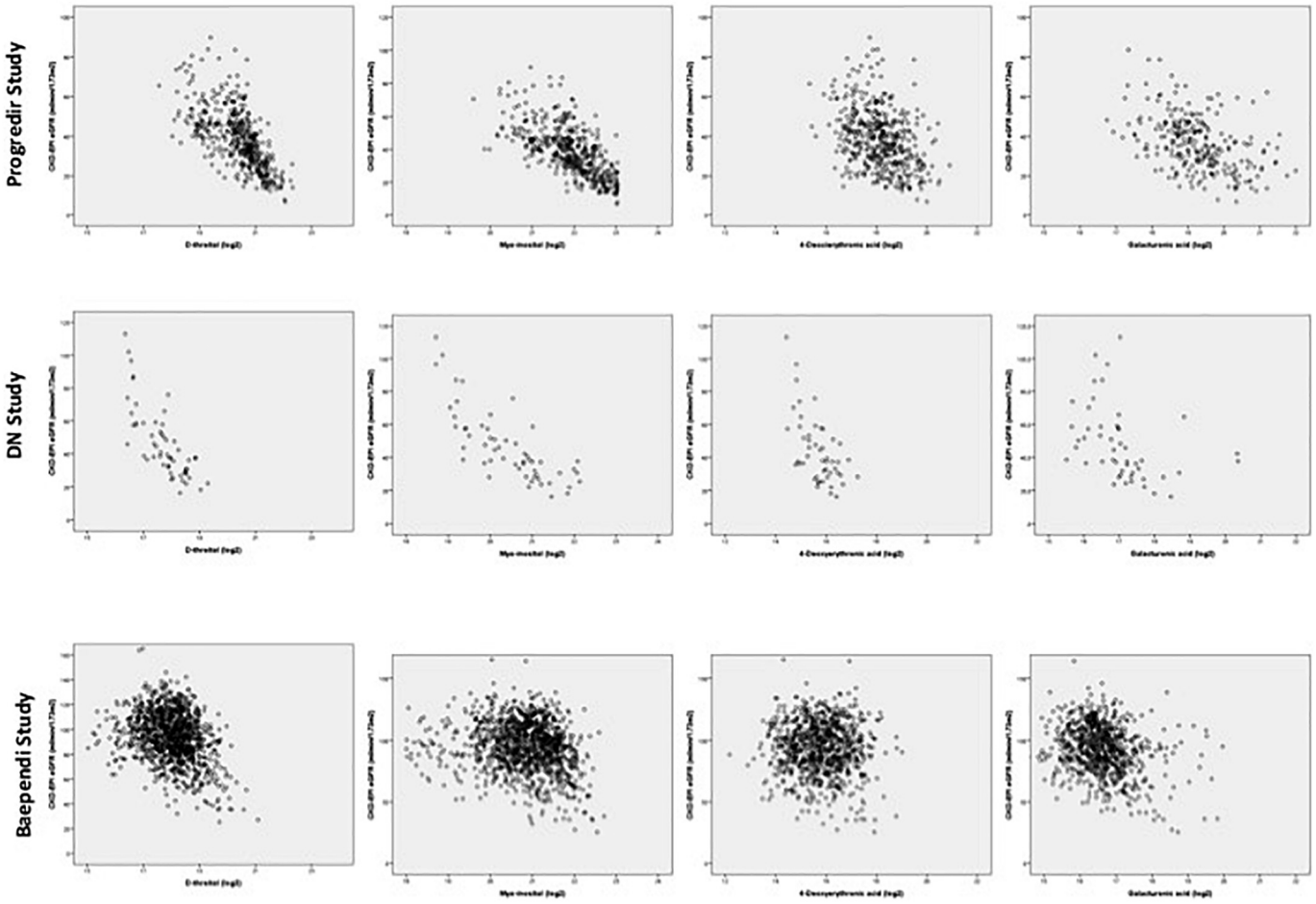

Fig. 1. Scatter plots of selected metabolites (D-threitol, myo-inositol, 4-deoxyerithronic acid and galacturonic acid; log2) and eGFR (mL/min/1.73m^2) in the 3 studies (Progredir Study, Diabetic Nephropathy Study and Baependi Heart Study).

replicated and significantly related to eGFR in the DN Study (11 metabolites) and in the Baependi Heart Study (8 metabolites). In addition, only 4 metabolites could be replicated and were significantly related to eGFR in the three studies: D-threitol, myo-inositol, 4-deoxyerithronic acid and galacturonic acid.

In Fig. 1 and Table 4, we show the scatter plots and the correlation coefficients between these 4 metabolites (log2) and eGFR in the 3 studies, respectively. It is interesting to note that while the metabolites show moderate to high correlation coefficients to eGFR in the CKD populations (Progredir Cohort and DN Study), the correlation is significantly attenuated in the general population. D-threitol and myo-inositol were the metabolites that showed the highest correlation to eGFR among all. In addition, pseudouridine (Table 4) also presented a strong association to eGFR in the two CKD studies (Progredir and DN Study) but was not correlated to eGFR in the BHS (general population).

Table 4
Spearman correlation between candidate metabolites and eGFR in the three studies.

	Progredir study		DN study		BHS	
	n = 454		n = 56		n = 1145	
	R	*p*	*R*	*p*	*R*	*p*
D-threitol	−0.71	8.70E-70	−0.80	9.62E-14	−0.28	1.83E-21
Myo-inositol	−0.70	4.92E-66	−0.81	5.48E-14	−0.20	2.00E-11
4-Deoxyerythronic acid	−0.38	7.87E-17	−0.56	2.49E-05	−0.023	0.50
Galacturonic acid	−0.47	6.10E-15	−0.58	9.30E-06	−0.23	9.89E-11
Pseudouridine	−0.71	1.26E-65	−0.86	3.76E-17	–	NS

DN Study Diabetic Nephropathy Study; BHS, Baependi Heart Study.

4. Discussion

In this analysis, we explored metabolites related to renal function, assessed here by the creatinine-based eGFR CKD-EPI formula. Eleven metabolites were replicated in the 2 CKD studies, showing moderate to high correlation coefficients to eGFR. However, only 8 metabolites from the derivation study could be replicated in the Baependi Heart Study, a population characterized by younger participants mostly with normal eGFR. Interestingly, the associations to eGFR in the Baependi Heart Study, although significant, tended to be much weaker than those observed in the CKD studies. In addition, metabolites such as pseudouridine, that presented a very high correlation to eGFR in the 2 CKD studies (−0.71 in the PS and − 0.80 in the DNS) showed no correlation at all to eGFR in the general population. These results suggest that estimating renal function using metabolic biomarkers in individuals with normal eGFR may be more troublesome. In our approach, we derived candidate molecules from a CKD population and this could be the reason why metabolites showed lower correlations to eGFR in the general population. However, among all metabolites related to eGFR in the Baependi Heart Study, none showed a high correlation coefficient to eGFR, suggesting that variability tends to be higher in the normal population, a condition where possibly other factors predominate over eGFR in the determination of metabolite interindividual variation. Approaches to overcome this problem could be the use of a panel of biomarkers and creating adjusting factors (as is done in MDRD and CKD-EPI formulas). A better understanding of how metabolites relate to eGFR in different strata of renal function might yield new strategies to predict renal function in a broad range of values, from severe decrease in eGFR to hyperfiltration.

From our analysis, 4 molecules emerged as potential candidates for estimating eGFR, by using an analytical approach of reproducibility in different cohort profiles: D-threitol, myo-inositol, 4-deoxyerithronic acid and galacturonic acid. In addition, pseudouridine was highly correlated to eGFR only in the CKD studies and not in the general population. Among these 5 metabolites, D-threitol has also been described as associated to eGFR in a previous publication [8], as well as myo-inositol and pseudouridine [7].

D-threitol is a polyol and end-product of D-xylose, a pentose. Myo-inositol is a sugar alcohol that can be obtained from diet or derived endogenously from glucose. Interestingly, it has been shown that myo-inositol oxygenase, the enzyme responsible for myo-inositol metabolism to glucoronate [14], is a renal-specific proximal tubule protein, shown to be increased in serum of animals and plasma of critically ill patients with acute kidney injury [15]. 4-deoxyerithronic acid is an organic acid derived from threonine metabolism [16] and it has already been shown to be increased in uremia [17]. Galacturonic acid is an oxidized form of D-galactose and can either be obtained from diet or produced endogenously. Some speculative reasoning on why these molecules are related to eGFR can be drawn. They are all small molecules and may present filtration properties (free passage by glomerular filtration barrier and little tubular secretion and absorption) that allows their serum/plasma concentration to reflect GFR. Alternatively, their metabolism may be related to a specific kidney function and thus serum/plasma concentration may be used to predict renal function, as may be the case for myo-inositol. More studies would be necessary to address these questions since the available literature about these metabolites in humans is still very scarce.

Our study presents some limitations. First, we used GC–MS to measure metabolites. This may limit our ability to see some classes of metabolites, particularly polar metabolites, which are better evaluated with LC-MS. The metabolomics platform used may also help explain differences between our results and studies in the literature that have reported other metabolites importantly related to renal function, such as C-mannosyltryptophan, *N*-acetylalanine, erythronate, and *N*-acetylcarnosine [7], metabolites that did not appear in our analyses. Secondly, we could not test our metabolites against a gold-standard measure of GFR, something that would be necessary for evaluating the true performance of metabolites in estimating GFR. Previously, only pseudouridine and C-mannosyltryptophan have been tested and validated against GFR measured as urinary clearance of ^{125}I-iothalamate in 200 participants with CKD [7], but not in populations with renal function in the normal range.

In conclusion, our results show 4 metabolites that were reproduced in different cohorts and consistently related to eGFR, thereby being candidate molecules for estimating eGFR in a broad spectrum of renal function range: D-threitol, myo-inositol, 4-deoxyerithronic acid and galacturonic acid. In addition, pseudouridine had a very good performance in predicting eGFR in the CKD population, but not in a population with normal or near-normal eGFR. Further studies are necessary to validate these findings against a gold-standard measure of glomerular filtration.

Conflicts of interest

Dr. Lotufo received horonaria from Abbot-Brazil, AbbVie-Brazil and Amgen for lectures.

Dr. Thadhani is a consultant to Fresenius Medical Care North America.

Corresponding author: Titan SM, Nephrology Division, Hospital das Clínicas - FMUSP.

Av Dr. Enéas de Carvalho Aguiar, 255, Zip 05403–000, São Paulo, SP, Brazil. *E*-mail address: smotitan@gmail.com

Acknowledgements

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Sao Paulo, Brazil; Brazilian Ministry of Health (Science and Technology Department), Brazil; FINEP, Brazil; CNPq, Brazil.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cca.2018.08.037.

References

[1] A.S. Levey, L.A. Inker, Assessment of Glomerular Filtration Rate in Health and Disease: a State of the Art Review, Clin. Pharmacol. Ther. 102 (3) (2017 Sep) 405–419.

[2] A.S. Levey, L.A. Inker, J. Coresh, GFR estimation: from physiology to public health, Am. J. Kidney Dis. 63 (5) (2014 May) 820–834.

[3] T.W. Ferguson, P. Komenda, N. Tangri, Cystatin C as a biomarker for estimating glomerular filtration rate, Curr. Opin. Nephrol. Hypertens. 24 (3) (2015 May) 295–300.

[4] L.A. Stevens, Y. Zhang, C.H. Schmid, Evaluating the performance of equations for estimating glomerular filtration rate, J. Nephrol. 21 (6) (2008 Nov-Dec) 797–807.

[5] P. Delanaye, C. Mariat, The applicability of eGFR equations to different populations, Nat. Rev. Nephrol. 9 (9) (2013 Sep) 513–522.

[6] E.J. Lamb, P.E. Stevens, Estimating and measuring glomerular filtration rate: methods of measurement and markers for estimation, Curr. Opin. Nephrol. Hypertens. 23 (3) (2014 May) 258–266.

[7] P. Sekula, O.N. Goek, L. Quaye, C. Barrios, A.S. Levey, W. Römisch-Margl, C. Menni, I. Yet, C. Gieger, L.A. Inker, J. Adamski, W. Gronwald, T. Illig, K. Dettmer, J. Krumsiek, P.J. Oefner, A.M. Valdes, C. Meisinger, J. Coresh, T.D. Spector, R.P. Mohney, K. Suhre, G. Kastenmüller, A. Köttgen, A Metabolome-Wide Association Study of Kidney Function and Disease in the General Population, J. Am. Soc. Nephrol. 27 (4) (2016 Apr) 1175–1188.

[8] B. Yu, Y. Zheng, J.A. Nettleton, D. Alexander, J. Coresh, E. Boerwinkle, Serum metabolomic profiling and incident CKD among African Americans, Clin. J. Am. Soc. Nephrol. 9 (8) (2014 Aug 7) 1410–1417.

[9] M.A.M. Domingos, A.C. Goulart, P.A. Lotufo, I.J.M. Benseñor, S.M.O. Titan, Chronic kidney disease - determinants of progression and cardiovascular risk. PROGREDIR cohort study: design and methods, Sao Paulo Med. J. 135 (2) (2017 Mar-Apr) 133–139.

[10] A.S. Levey, L.A. Stevens, C.H. Schmid, Y.L. Zhang, A.F. Castro 3rd, H.I. Feldman, J.W. Kusek, P. Eggers, F. Van Lente, T. Greene, J. Coresh, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate, Ann. Intern. Med. 150 (9) (2009 May 5) 604–612.

[11] Titan SM, M Vieira J Jr, Dominguez WV, Barros RT, Zatz R. ACEI and ARB combination therapy in patients with macroalbuminuric diabetic nephropathy and low socioeconomic level: a double-blind randomized clinical trial. Clin. Nephrol. 2011 Oct;76(4):273–83.

[12] K.J. Egan, M. von Schantz, A.B. Negrão, H.C. Santos, A.R. Horimoto, N.E. Duarte, G.C. Gonçalves, J.M. Soler, M. de Andrade, G. Lorenzi-Filho, H. Vallada, T.P. Taporoski, M. Pedrazzoli, A.P. Azambuja, C.M. de Oliveira, R.O. Alvim, J.E. Krieger, A.C. Pereira, Cohort profile: the Baependi Heart Study-a family-based, highly admixed cohort study in a rural Brazilian town, BMJ Open 6 (10) (2016 Oct 21) e011598.

[13] O. Fiehn, Metabolomics by gas chromatography-mass spectrometry: combined targeted and untargeted profiling, Curr Protoc Mol Biol. 114 (2016 Apr 1) (30.4.1-30.4.32).

[14] A.G. Thorsell, C. Persson, N. Voevodskaya, R.D. Busam, M. Hammarström, S. Gräslund, A. Gräslund, B.M. Hallberg, Structural and biophysical characterization of human myo-inositol oxygenase, J. Biol. Chem. 283 (22) (2008 May 30) 15209–15216.

[15] J.P. Gaut, D.L. Crimmins, M.F. Ohlendorf, C.M. Lockwood, T.A. Griest, N.A. Brada, M. Hoshi, B. Sato, R.S. Hotchkiss, S. Jain, J.H. Ladenson, Development of an immunoassay for the kidney-specific protein myo-inositol oxygenase, a potential biomarker of acute kidney injury, Clin. Chem. 60 (5) (2014 May) 747–757.

[16] D.B. Kassel, M. Martin, W. Schall, C.C. Sweeley, Urinary metabolites of L-threonine in type 1 diabetes determined by combined gas chromatography/chemical ionization mass spectrometry, Biomed. Environ. Mass Spectrom. 13 (10) (1986 Oct) 535–540.

[17] F.W. Bultitude, S.J. Newham, Identification of some abnormal metabolites in plasma from uremic subjects, Clin. Chem. 21 (9) (1975 Aug) 1329–1334.