Journal of the American Heart Association

ORIGINAL RESEARCH

Large-Scale Metabolomics and the Incidence of Cardiovascular Disease

Lars Lind, MD, PhD; Tove Fall, PhD; Johan Ärnlöv, MD, PhD; Sölve Elmståhl, MD, PhD; Johan Sundström, MD, PhD

BACKGROUND: The study aimed to show the relationship between a large number of circulating metabolites and subsequent cardiovascular disease (CVD) and subclinical markers of CVD in the general population.

METHODS AND RESULTS: In 2278 individuals free from CVD in the EpiHealth study (aged 45–75 years, mean age 61 years, 50% women), 790 annotated nonxenobiotic metabolites were measured by mass spectroscopy (Metabolon). The same metabolites were measured in the PIVUS (Prospective Investigation of Vasculature in Uppsala Seniors) study (n=603, all aged 80 years, 50% women), in which cardiac and carotid artery pathologies were evaluated by ultrasound. During a median follow-up of 8.6 years, 107 individuals experienced a CVD (fatal or nonfatal myocardial infarction, stroke, or heart failure) in EpiHealth. Using a false discovery rate of 0.05 for age- and sex-adjusted analyses and $P<0.05$ for adjustment for traditional CVD risk factors, 37 metabolites were significantly related to incident CVD. These metabolites belonged to multiple biochemical classes, such as amino acids, lipids, and nucleotides. Top findings were dimethylglycine and N-acetylmethionine. A lasso selection of 5 metabolites improved discrimination when added on top of traditional CVD risk factors (+4.0%, $P=0.0054$). Thirty-five of the 37 metabolites were related to subclinical markers of CVD evaluated in the PIVUS study. The metabolite 1-carboxyethyltyrosine was associated with left atrial diameter as well as inversely related to both ejection fraction and the echogenicity of the carotid artery.

CONCLUSIONS: Several metabolites were discovered to be associated with future CVD, as well as with subclinical markers of CVD. A selection of metabolites improved discrimination when added on top of CVD risk factors.

Key Words: amino acids ■ cardiovascular disease ■ epidemiology ■ mass spectroscopy ■ metabolomics

Omics technologies could be used in epidemiology either to search for novel insights into the pathogenesis of diseases or to identify clinically relevant risk markers that could add to the predictive power of established risk factors. One such example is the use of circulating proteomics to predict cardiovascular disease (CVD), where studies have identified both already well-known biomarkers, such as NT-proBNP (N-terminal pro-B-type natriuretic peptide), but also some new proteins linked to incident CVD.[1,2]

Another omics modality is metabolomics, the study of small molecules, measured either by mass spectroscopy (MS) or by nuclear magnetic resonance techniques.

In a systematic review from 2017,[3] the authors found 12 studies in which metabolomic markers were associated with incident CVD. Metabolites from several metabolite classes, such as carnitines, amino acids, and lipids, were related to future CVD. However, the authors concluded that a quantitative synthesizing of the previous findings was difficult, given the use of different analytical platforms and statistical methods. Since 2017, a few additional large studies have been published, using MS or nuclear magnetic resonance.[4–8] Also, in this case, the results are not easily comparable, because the number and nature of annotated metabolites varied greatly between studies (see overview in Table S1).

Correspondence to: Lars Lind, MD, PhD, Department of Medical Sciences, Uppsala University, 751 85 Uppsala, Sweden. Email: lars.lind@medsci.uu.se

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.122.026885

For Sources of Funding and Disclosures, see page 11.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- In a population-based sample, we found 37 metabolites to be related to incident cardiovascular disease.
- These metabolites belonged to multiple biochemical classes, such as amino acids, lipids, and nucleotides.
- Most metabolites were also associated with subclinical markers of cardiovascular disease, such as enlarged dimeter of the left atrium and a reduced ejection fraction.

What Are the Clinical Implications?

- A selection of 5 metabolites improved discrimination of cardiovascular disease by 4% (C statistic) when added to traditional cardiovascular disease risk factors.
- If reproduced by others, metabolites might be included in risk prediction scores in the future.

Nonstandard Abbreviations and Acronyms

MS	mass spectroscopy
PIVUS	Prospective Investigation of Vasculature in Uppsala Seniors

Biomarker discovery could be used for several purposes, including to identify new pathophysiological pathways that potentially could be used as targets for drug development, identify new biomarkers that could be used as diagnostics at the individual level, and identify new biomarkers that could be used for improved risk stratification.

The primary aim of the present study was to evaluate if we could identify new pathophysiological pathways for CVD by large-scale metabolomics. If so, we could, as a secondary aim, evaluate if those metabolites could improve risk stratification in relation to traditional CVD risk factors. For the primary aim, we investigated the metabolomic profile associated with incident CVD (myocardial infarction, stroke, or heart failure) using MS-based data on almost 800 nonxenobiotic annotated metabolites in the EpiHealth study. This is the study with the greatest number of metabolites evaluated on incident CVD performed to date. To support our findings in the EpiHealth study, we investigated if associations between the metabolites of interest were also related to indices of subclinical CVD in an independent sample, the PIVUS (Prospective Investigation of Vasculature in Uppsala Seniors) study. The 5 indices of subclinical CVD evaluated (left atrial diameter, left ventricular ejection fraction, left ventricular mass, intima-media thickness, and the echogenicity of the carotid artery wall) have all previously been shown to be associated with incident CVD.[9–14]

METHODS

Samples

EpiHealth is a population-based study conducted with the same protocol in 2 Swedish cities, Uppsala and Malmö. Approximately 25 000 individuals participated from 2011 to 2018.

The age range was 45 to 75 years, and 50% were women. Traditional CVD risk factors were measured. In the first 2342 subjects included in Uppsala, metabolomics was analyzed on frozen plasma samples. The cohort has been followed for almost 10 years on incident CVD. Details on the examination have been provided previously.[15]

The PIVUS study is a population-based study conducted in Uppsala, Sweden. At age 70 years, 1016 subjects were investigated from 2001 to 2004 (50% were women). Ten years later, all subjects still alive were invited to a new examination, in which 604 participated. At that time point, traditional CVD risk factors were measured, and an echocardiogram and a carotid ultrasound were obtained. Metabolomics was analyzed on frozen plasma samples in all but 1 individual (because of a lost vial). Thus, the present study used only data from those aged 80 years.

Details on the examinations have been given previously.[16] Eleven percent had a history of myocardial infarction, 10% had suffered from a stroke, and 9% reported a heart failure diagnosis at the time of the investigation.

Both studies have been approved by the ethics committee at Uppsala University, and all study participants have given their informed consent to participate. The data that support the finding of this study are available from the corresponding author upon reasonable request.

CVD Risk Factors

Blood pressure was measured in the supine position in the PIVUS study and in the sitting position in EpiHealth (M10-IT; Omron, Kyoto, Japan). Blood was drawn after an overnight fast in the PIVUS study and after a minimum of 6 hours in EpiHealth. Glucose and low-density lipoprotein and high-density lipoprotein cholesterol were measured by standard techniques at the clinical chemistry laboratory at the university hospital. Plasma was frozen at −80 °C for later metabolomics analysis. Smoking and medications were assessed by a questionnaire. Body mass index (BMI) was calculated by measured height and weight. Diabetes was defined as glucose ≥7.0 mmol/L or use of antidiabetic medication.

Metabolomics

In both samples, the same nontargeted metabolomics (Metabolon) was performed on plasma samples. Samples were prepared using the automated MicroLab STAR system (Hamilton). Several internal standards were added before the first step in the extraction process for quality control purposes. To remove protein, we dissociated small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 minutes (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into 5 fractions: 2 for analysis by 2 separate reverse phases/ultraperformance liquid chromatography–MS/MS methods with positive ion mode electrospray ionization, 1 for analysis by reverse phases/ultraperformance liquid chromatography–MS/MS with negative ion mode electrospray ionization, 1 for analysis by hydrophilic interaction/ultraperformance liquid chromatography–MS/MS with negative ion mode electrospray ionization, and 1 sample was reserved for backup. Only annotated, nonxenobiotic metabolites with a detection rate >75% in all samples were used in the analyses (n=790). The values were normalized and given in arbitrary units.

CVD End Point

Data from the Swedish registries of mortality and in-hospital care were used to follow up incident cases of CVD in the EpiHealth study. In this study we used a combined CVD end point consisting of myocardial infarction (*International Classification of Diseases, Tenth Revision* [*ICD-10*] code I21), ischemic stroke (*ICD-10* code I63), and heart failure (*ICD-10* codes I50 and I11.0). Both fatal and nonfatal events were included in the combined end point.

Ultrasound

In the PIVUS study, a 2-dimensional echocardiography examination was performed with an Acuson XP124 cardiac ultrasound unit (Acuson).

A 2.5-MHz transducer was used for the majority of the examinations. Left ventricular dimensions were measured with M-mode online from the parasternal projection, using a leading-edge-to-leading-edge convention. Measurements included left atrial diameter, interventricular septal thickness, posterior wall thickness, and left ventricular diameter in end-diastole and end-systole. Left ventricular mass was determined from the Penn conversion. Left ventricular mass was then indexed for height ($g/m^{2.7}$) to obtain left ventricular mass index.

Left ventricular volumes were calculated according to the Teichholz formula (7*D3/[2.4+D]), and from that the ejection fraction was calculated.

Carotid artery ultrasound was performed using a 10-MHz probe investigating both arteries. Intima-media thickness was measured over a 10-mm distance proximal to the bifurcation in the far wall by semiautomated software. In the same segment, the gray scale of the intima-media complex was determined (echogenicity of the intima-media complex). The mean value of the 2 arteries was used for both indices. Details on the carotid artery ultrasound measurements have been reported previously.[17]

Statistical Analysis

The metabolomic variables were subjected to inverse rank normalization to obtain normally distributed variables on an SD scale, in both EpiHealth and the PIVUS study.

In EpiHealth, we investigated the association between metabolites and incident CVD using separate Cox proportional hazard analysis for each metabolite. Two levels of adjustment were performed. First, adjustment for age and sex, and second, additional adjustment for traditional CVD risk factors (systolic blood pressure, diabetes, low-density lipoprotein and high-density lipoprotein cholesterol, BMI, and current smoking).

For these analyses, a false discovery rate <0.05 for the age- and sex-adjusted analysis and $P<0.05$ for adjustment for traditional CVD risk factors were regarded as significant.

A lasso logistic regression analysis using the 37 significant metabolites from the previous step as independent variables together with age and sex was conducted to sort out the most influential independent metabolites based on their β coefficients. The sample was split in 2 parts, and we used lasso with 10 cross-validations. This procedure selects λ* to be the λ that gives the minimum of the cardiovascular function. Because we knew that we had 107 incident CVD cases during the follow-up period and we had 6 independent significant traditional CVD risk factors (including age and sex), we decided a priori that we should only add 5 metabolites, at most, to the model, including the risk factors, so as not to induce any overfitting problems (based on a rule of thumb that there should be around 10 cases per covariate in the model).

First, the traditional CVD risk factors (including age and sex) entered a logistic regression model, and the C index was calculated (after excluding BMI and diabetes, which were far from significant in the model and did not influence the C index).

Second, the top 5 metabolites in the lasso regression model (all showing $P<0.05$) entered a logistic regression model together with age and sex to calculate the C statistics.

In the third step, we entered the 5 metabolites together with age, sex, and traditional risk factors and calculated the C index. In a fourth step, the model with the traditional risk factors only were compared with the model including both traditional risk factors and the 5 metabolites.

In the PIVUS study, the 37 metabolites identified to be of interest in EpiHealth were one-by-one related to each of the 5 markers of subclinical CVD by linear regression analysis. Similar to the previous analyses in EpiHealth, 2 levels of adjustment were performed: (1) adjustment for sex (the age was the same in all subjects) and (2) additional adjustment for traditional CVD risk factors. In this exploratory validation, $P<0.05$ was considered statistically significant.

Stata 16.1 was used for the calculations (StataCorp, College Station, TX).

RESULTS

Basic characteristics of the 2 cohorts are provided in Table 1. After excluding 64 participants with prevalent CVD at baseline, 2278 persons were at risk during a median follow-up of 8.6 years (maximum 9.6 years, 18 852 person-years at risk). During that period, 107 individuals suffered from CVD.

Metabolites Versus Incident CVD

Of the 790 evaluated metabolites, 37 were associated with incident CVD (overview in Figure 1 and details in Table 2 and Table S2). These metabolites belonged to multiple biochemical classes, such as amino acids, lipids, and nucleotides. Also, within the lipid group, different functional classes were represented, such as steroids and phosphatidylethanolamines. The top findings were 2 amino acid derivates, dimethylglycine and N-acetylmethionine. As seen in Table 2, the relationships were generally positive, but a few of the steroid hormones were (inversely) associated with incident CVD.

Metabolites and Prediction of CVD

A model with age, sex, and traditional cardiovascular risk factors (systolic blood pressure, low-density lipoprotein and high-density lipoprotein cholesterol, and current smoking) was developed, with a C index of 0.74 (95% CI, 0.70–0.78) for discrimination of subsequent CVD.

A lasso regression identified the 5 metabolites of top importance for incident CVD (5-methylthioadenosine, dimethylglycine, pregnenediol sulfate, imidazole propionate, and 1-carboxyethyltyrosine). When those were used in a model together with age and sex, the model resulted in a C index of 0.77.

When the 5 metabolites were added to the established risk factors, the C index increased by 4.0% (0.78 [95% CI, 0.74–0.81] versus 0.74 [95% CI, 0.70–0.78]) when compared with the model with traditional risk factors only ($P=0.005$; Figure 2).

When we performed a bootstrap analysis with 10 000 repetitions of the difference in C statistics between the traditional risk factors and the model that included 5 metabolites together with the traditional risk factors, the mean difference across the subsamples was close to the calculation in the total sample, with a 95% CI being highly significant (bootstrapped mean C statistic difference between models 0.0391 [95% CI, 0.0134–0.0649]; $P=0.003$).

Metabolites Versus Subclinical Markers of CVD

As can be seen in the overview in Figure 3, 35 of the 37 metabolites identified in EpiHealth were related to

Table 1. Basic Characteristics of the 2 Samples

	EpiHealth, n=2278	PIVUS, n=603
Age, y	61.1 (8.4)	80.1 (0.2)
Female sex, %	50.2	50.4
Systolic blood pressure, mm Hg	134.1 (16.8)	146.4 (19.2)
LDL cholesterol, mmol/L	3.9 (0.9)	3.4 (0.8)
HDL cholesterol, mmol/L	1.5 (0.3)	1.4 (0.4)
Diabetes (%)	8.0	11.2
Smoking	6.7 (8.9) y smoked	3.1% current smokers
Body mass index, kg/m^2	26.4 (3.8)	26.8 (4.3)
Left atrial diameter, mm	NA	42.2 (6.6)
Ejection fraction, %	NA	64.8 (10)
Left ventricular mass index, g/m$^{2.7}$	NA	45.1 (12)
Intima-media thickness, mm	NA	0.95 (0.16)
Echogenicity of the carotid artery wall	NA	59.8 (15.4)

Mean and SD or proportions are given. HDL indicates high-density lipoprotein; LDL, low-density lipoprotein; NA, not assessed; and PIVUS, Prospective Investigation of Vasculature in Uppsala Seniors.

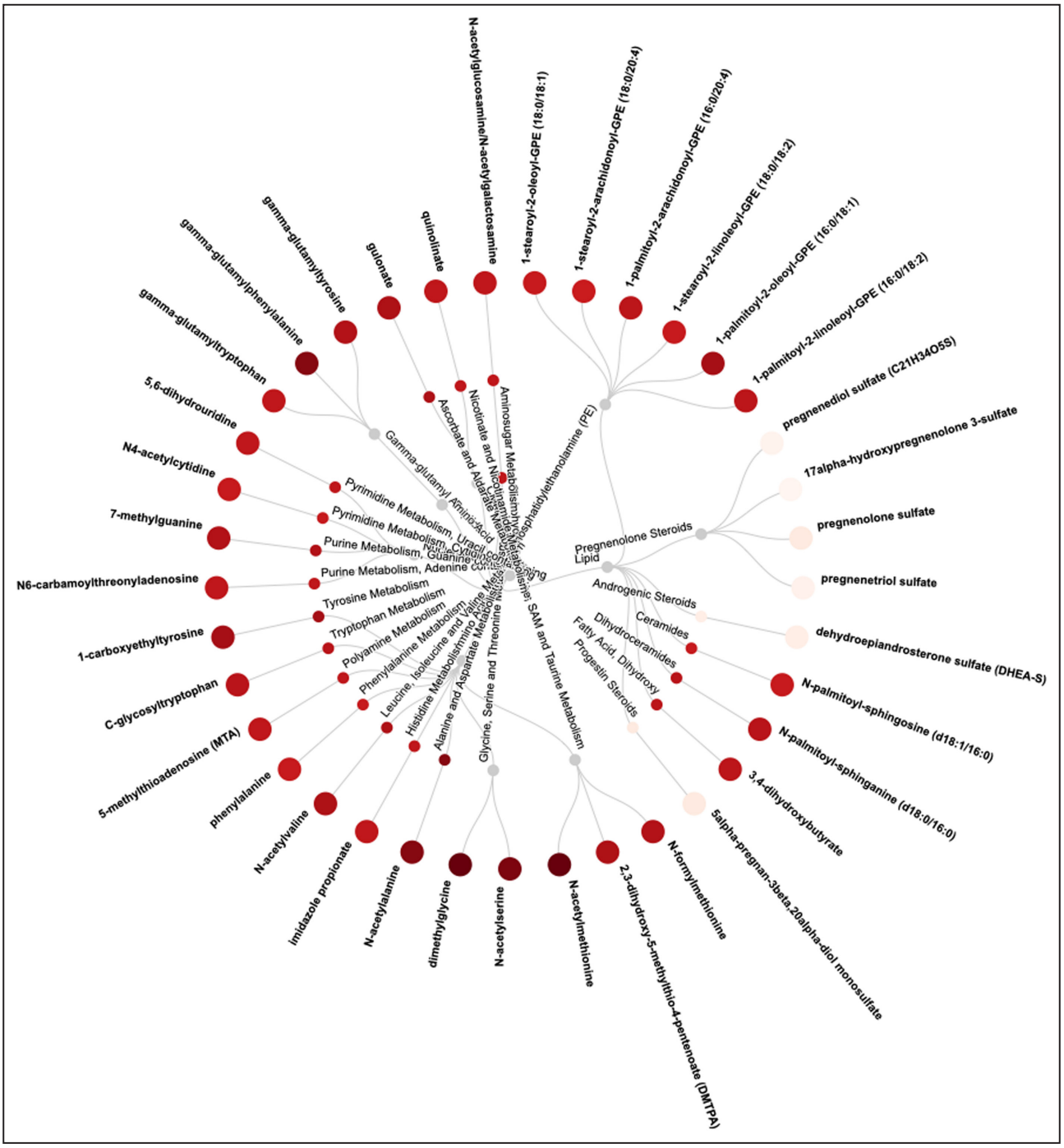

Figure 1. Overview of the 37 significant metabolites related to incident cardiovascular disease in the EpiHealth sample. The intensity of the red color is proportional to the estimate (hazard ratio) so that the pale dots correspond to hazard ratios <1.0. Subpathway and metabolite names for these 37 metabolites, with the start from 1-stearoyl-2-oleoyl-GPE (18:0/18:1) at the top of the figure and then in clockwise order, are given in Table S2. GPE indicates glycero-3-phosphoethanolamine; and SAM, S-adenosylmethionine.

subclinical markers of CVD evaluated in the PIVUS study. Eleven of the 37 metabolites showed $P<0.05$ in relation to left ventricular mass index in the age- and sex-adjusted analysis, but none of these metabolites were significantly associated with left ventricular mass index following adjustment for established cardiovascular risk factors (see Table S3 for details). Thirty-four of the metabolites were related to ejection fraction. Most of these relationships were inverse. Twenty-nine of those associations with ejection fraction were statistically significant following multivariable adjustment. Twenty-nine of the metabolites were related to left atrial diameter.

Table 2. Relationships Between Metabolites and Incident Cardiovascular Disease

Super pathway	Subpathway	Chemical name	Age and sex adjusted				Multiple adjusted			
			HR	Low 95% CI	High 95% CI	*P* value	HR	Low 95% CI	High 95% CI	*P* value
Amino acid	Alanine and aspartate metabolism	N-acetylalanine	1.52	1.25	1.86	0.000036	1.51	1.22	1.88	0.00017
Amino acid	Glycine, serine, and threonine metabolism	N-acetylserine	1.54	1.26	1.88	0.000030	1.54	1.25	1.9	0.000064
Amino acid	Glycine, serine, and threonine metabolism	Dimethylglycine	1.58	1.3	1.95	0.000010	1.63	1.32	2.01	5.9e-06
Amino acid	Histidine metabolism	Imidazole propionate	1.38	1.13	1.7	0.0019	1.36	1.11	1.68	0.0045
Amino acid	Leucine, isoleucine, and valine metabolism	N-acetylvaline	1.43	1.17	1.75	0.00031	1.4	1.15	1.73	0.00095
Amino acid	Methionine, cysteine, SAM, and taurine metabolism	N-formylmethionine	1.42	1.17	1.73	0.00040	1.4	1.15	1.73	0.00093
Amino acid	Methionine, cysteine, SAM, and taurine metabolism	2,3-dihydroxy-5-methylthio-4-pentenoate	1.43	1.16	1.79	0.00074	1.42	1.13	1.77	0.0028
Amino acid	Methionine, cysteine, SAM, and taurine metabolism	N-acetylmethionine	1.58	1.3	1.93	6.8e-06	1.55	1.27	1.92	0.000027
Amino acid	Phenylalanine metabolism	Phenylalanine	1.36	1.13	1.67	0.0016	1.35	1.11	1.65	0.0038
Amino acid	Polyamine metabolism	5-methylthioadenosine	1.38	1.14	1.68	0.0010	1.36	1.11	1.68	0.0037
Amino acid	Tryptophan metabolism	C-glycosyltryptophan	1.4	1.13	1.73	0.0021	1.38	1.09	1.72	0.0053
Amino acid	Tyrosine metabolism	1-carboxyethyltyrosine	1.45	1.17	1.79	0.00047	1.43	1.15	1.8	0.0017
Carbohydrate	Aminosugar metabolism	N-acetylglucosamine/N-acetylgalactosamine	1.38	1.13	1.68	0.0021	1.32	1.06	1.65	0.012
Cofactors and vitamins	Ascorbate and aldarate metabolism	Gulonate	1.43	1.17	1.75	0.00048	1.4	1.14	1.72	0.0011
Cofactors and vitamins	Nicotinate and nicotinamide metabolism	Quinolinate	1.38	1.13	1.68	0.0021	1.35	1.07	1.68	0.0091
Lipid	Androgenic steroids	Dehydroepiandrosterone sulfate	0.68	0.54	0.84	0.00035	0.67	0.54	0.84	0.00029
Lipid	Ceramides	N-palmitoyl-sphingosine (d18:1/16:0)	1.36	1.12	1.65	0.0023	1.43	1.13	1.84	0.0037
Lipid	Dihydroceramides	N-palmitoyl-sphinganine (d18:0/16:0)	1.4	1.15	1.7	0.00063	1.36	1.09	1.7	0.0063
Lipid	Fatty acid, dihydroxy	3,4-dihydroxybutyrate	1.39	1.14	1.7	0.0012	1.35	1.09	1.67	0.0052
Lipid	Phosphatidylethanolamine	1-stearoyl-2-oleoyl-GPE (18:0/18:1)	1.36	1.12	1.65	0.0019	1.27	1.03	1.57	0.022
Lipid	Phosphatidylethanolamine	1-stearoyl-2-arachidonoyl-GPE (18:0/20:4)	1.35	1.12	1.63	0.0017	1.25	1.02	1.52	0.029
Lipid	Phosphatidylethanolamine	1-palmitoyl-2-arachidonoyl-GPE (16:0/20:4)	1.39	1.15	1.7	0.00059	1.31	1.08	1.6	0.0067

(*Continued*)

Table 2. Continued

Super pathway	Subpathway	Chemical name	Age and sex adjusted				Multiple adjusted			
			HR	Low 95% CI	High 95% CI	*P* value	HR	Low 95% CI	High 95% CI	*P* value
Lipid	Phosphatidylethanolamine	1-stearoyl-2-linoleoyl-GPE (18:0/18:2)	1.36	1.13	1.65	0.0016	1.28	1.05	1.57	0.015
Lipid	Phosphatidylethanolamine	1-palmitoyl-2-oleoyl-GPE (16:0/18:1)	1.45	1.19	1.75	0.00021	1.34	1.09	1.65	0.0051
Lipid	Phosphatidylethanolamine	1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	1.39	1.15	1.68	0.00080	1.32	1.09	1.62	0.0044
Lipid	Pregnenolone steroids	Pregnenediol sulfate (C21H34O5S)	0.65	0.51	0.82	0.00026	0.64	0.51	0.81	0.00024
Lipid	Pregnenolone steroids	17α-hydroxypregnenolone 3-sulfate	0.64	0.49	0.84	0.00090	0.6	0.46	0.79	0.00020
Lipid	Pregnenolone steroids	Pregnenolone sulfate	0.7	0.56	0.87	0.0011	0.68	0.54	0.84	0.00063
Lipid	Pregnenolone steroids	Pregnenetriol sulfate	0.66	0.52	0.84	0.00093	0.64	0.5	0.82	0.00036
Lipid	Progestin steroids	5α-pregnan-3β, 20α-diol monosulfate	0.7	0.57	0.85	0.00053	0.71	0.58	0.88	0.0015
Nucleotide	Purine metabolism, adenine containing	N6-carbamoylthreonyladenosine	1.38	1.13	1.68	0.0021	1.34	1.08	1.65	0.0076
Nucleotide	Purine metabolism, guanine containing	7-methylguanine	1.42	1.16	1.72	0.00053	1.34	1.09	1.63	0.0046
Nucleotide	Pyrimidine metabolism, cytidine containing	N4-acetylcytidine	1.36	1.12	1.65	0.0020	1.34	1.08	1.63	0.0054
Nucleotide	Pyrimidine metabolism, uracil containing	5,6-dihydrouridine	1.38	1.13	1.7	0.0019	1.38	1.12	1.7	0.0027
Peptide	γ-glutamyl amino acid	γ-glutamyltryptophan	1.38	1.14	1.68	0.0012	1.35	1.11	1.65	0.0031
Peptide	γ-glutamyl amino acid	γ-glutamylphenylalanine	1.52	1.23	1.86	0.000056	1.51	1.21	1.86	0.00026
Peptide	γ-glutamyl amino acid	γ-glutamyltyrosine	1.42	1.16	1.73	0.00050	1.38	1.12	1.68	0.0030

Only metabolites with false discovery rate <0.05 for the age- and sex-adjusted analysis and *P*<0.05 for the multiple-adjusted analysis are given. GPE indicates glycero-3-phosphoethanolamine; HR, hazard ratio; and SAM, S-adenosylmethionine.

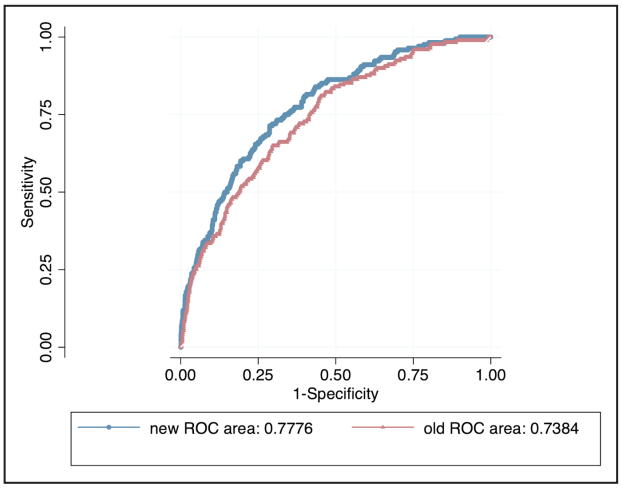

Figure 2. Area under the curve for traditional risk factors vs traditional risk factors plus metabolites on the outcome incident cardiovascular disease in the EpiHealth study.
The ROC curve (old ROC area) for the traditional risk factors of importance (systolic blood pressure, low-density lipoprotein and high-density lipoprotein cholesterol, and current smoking) is given in red (denoted old ROC curve), whereas the curve in black (denoted new ROC curve) also includes the 5 metabolites 5-methylthioadenosine, dimethylglycine, pregnenediol sulfate, imidazole propionate, and 1-carboxyethyltyrosine. P=0.0054 for the difference between curves. ROC indicates receiver operating characteristic.

Only 9 of those relationships versus left atrial diameter still showed P<0.05 following multivariable adjustment.

Four metabolites were related to intima-media thickness, and 2 of those showed P<0.05 following multivariable adjustment (both inverse). Nine metabolites were related to echogenicity of the intima-media complex, and 2 of those showed P<0.05 following multivariable adjustment.

Of particular interest was 1-carboxyethyltyrosine, being related to left atrial diameter, as well as inversely related to both ejection fraction and the echogenicity of the carotid artery (echogenicity of the intima-media complex) also after adjustment for traditional cardiovascular risk factors.

DISCUSSION

The present study, which investigated almost 800 different nonxenobiotic metabolites, showed that 37 metabolites were related to an increased risk of subsequent CVD. As expected from previous studies in this field, these metabolites represent several different biochemical classes and biological pathways. The top 5 metabolites improved the discrimination of incident CVD when added to established cardiovascular risk factors. Most of these 37 metabolites were related to different markers of subclinical CVD in an independent sample. Of particular interest was 1-carboxyethyltyrosine, being related to both cardiac performance and structural changes in the carotid artery wall.

Comparison With the Literature

As summarized in Table 1, a great number of metabolites have previously been linked to incident CVD. The most consistent findings have been for certain amino acids, such as branch-chained and aromatic amino acids, acetylcarnitines, and certain lipid classes, such as lysophophatidylcholines and sphingomyelins. The present study highlighted the previously mentioned amino acids and phosphatidylethanolamines, but also found steroids, mainly from the pregnenolone pathway, and metabolites, from the nucleotide purine metabolism, to be linked to future CVD.

Although this investigation and previous studies have found relationships between CVD and many different metabolites, the biological significance remains to be established. First, associations are not always causal. One way to evaluate causation is by Mendelian randomization using genetic information. However, although information on the genetic associations exists for some metabolites,[18] the genetic basis for the majority of the 790 metabolites evaluated in the present study does not exist. Another fact that hampers the use of Mendelian randomization in metabolomics research is that a certain genetic locus is often related to several metabolites in the same pathway or chemical class. Such pleiotropy violates the fundamental assumptions of the Mendelian randomization analysis. Thus, at present, it is hard to evaluate if the majority of the reported associations are causally related to CVD or not, but we hope that the identification of these metabolites and metabolic pathways could serve as inspiration to evaluate such associations in experimental models in which interventions could be made to evaluate causality.

In a pathway enrichment analysis (using MetaboAnalyst; www.metaboanalyst.cd) of the 37 metabolites of interest, no pathway was significantly enriched and, in most cases, only 1 metabolite was identified per pathway, suggesting that no physiological pathway was of major importance compared with the others. Thus, incident CVD is linked to many pathophysiological pathways, and the present metabolomics analysis could not identify a pathway of superior interest.

Steroids from the pregnenolone pathway identified in the present study were generally inversely related to future CVD. This was somewhat surprising, given that pregnenolone metabolites have been reported to be elevated in patients with myocardial infarction.[19] One reason for increased levels of pregnenolone metabolites following an acute myocardial infarction might be

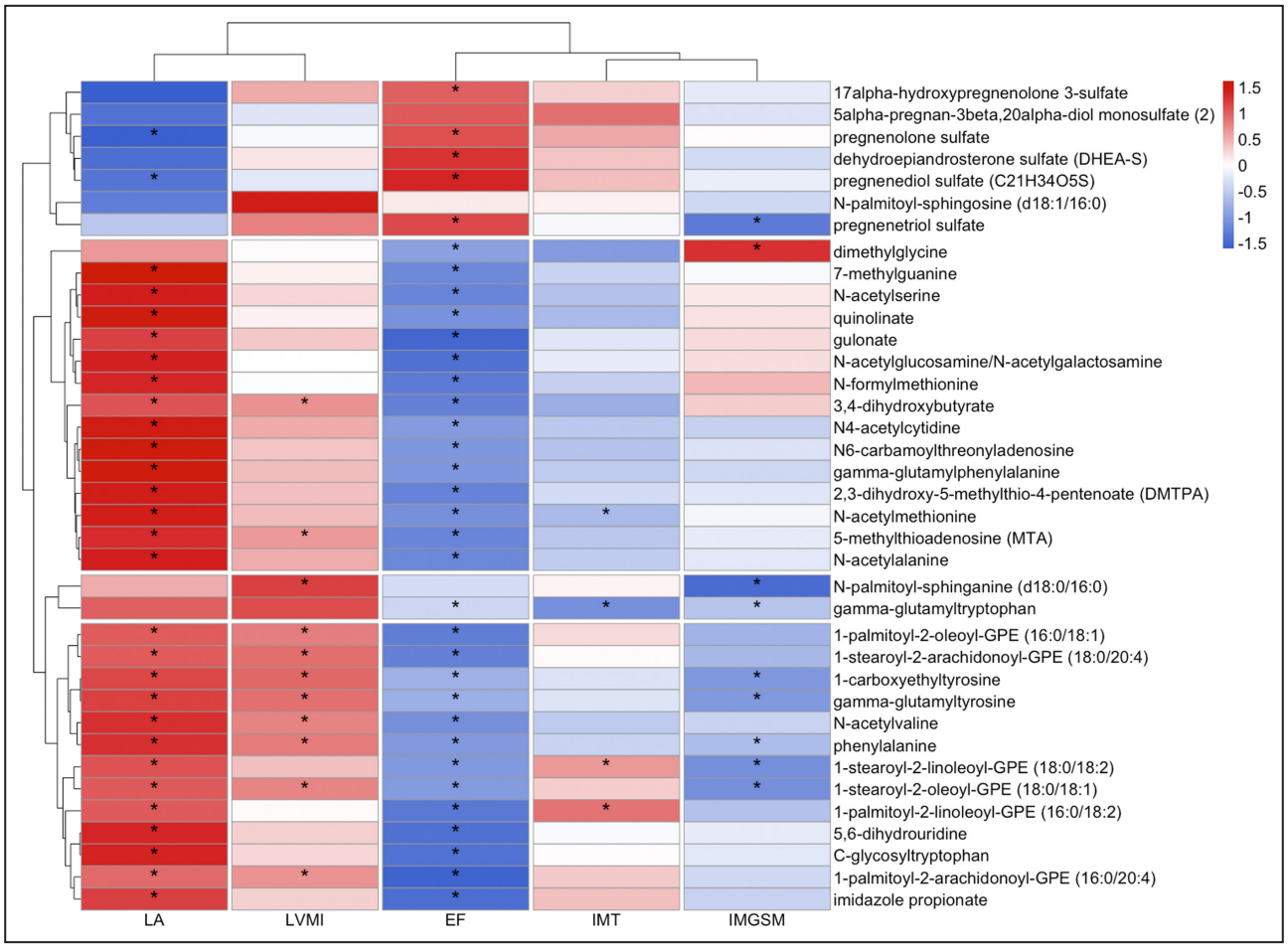

Figure 3. Hierarchically clustered heat map of the relationships (given as regression coefficients) between the 37 metabolites found to be related to incident CVD and markers of subclinical CVD in the PIVUS study following adjustment for age and sex.

Red filling indicates positive relationships, whereas blue filling indicates inverse relationships. A star denotes that the relationships showed $P<0.05$. The heat map was created in R 4.2 using the package pheatmap. CVD indicates cardiovascular disease; EF, ejection fraction; GPE, glycero-3-phosphoethanolamine; IMGSM, echogenicity of the intima-media complex; IMT, carotid artery intima-media thickness; LA, left atrial diameter; LVMI, left ventricular mass index; and PIVUS, Prospective Investigation of Vasculature in Uppsala Seniors.

activation of the hypothalamic–pituitary–adrenal axis caused by the stressful event.[20] However, in this case, the pregnenolone metabolites were evaluated years before the event, and it has been published that low levels of downstream metabolites, such as testosterone[21] and cortisol,[22,23] are associated with future CVD or a high CVD risk. A recent experimental study showed that testosterone increased endothelial nitric oxide synthesis and activity of superoxide dismutase and reduced catalase activity, with potential vascular protective effects,[24] suggesting potential pathophysiological events at the molecular level by low levels of this androgen.

It is well established that urate levels, being a breakdown product of purine metabolism, is a risk factor for CVD (see meta-analysis in Li et al[25]) A recent Mendelian randomization study has suggested this association to be causal.[26] There are several possible explanations whereby high urate levels could induce CVD. Urate could induce oxidative stress by upregulating xanthine oxidoreductase.[27] Urate could also increase proinflammatory cytokines and active platelets.[28] Whether or not these pathophysiological mechanisms are also valid for the purine metabolites identified in the present study is not known. However, the upstream nucleotide metabolism is not well studied in the context of CVD. Here, 4 of the 37 metabolites of interest belonged to purine or pyramidine metabolism. This is a novel and potentially important finding that merits further studies.

Glucose reacts nonenzymatically with other metabolites, such as proteins, to form advanced glycated end products. This is seen with aging, but is also accelerated in diabetes. The metabolite 1-carboxyethyltyrosine, being related to both cardiac performance and composition of the carotid artery wall, is such an example. Advanced glycated end products have mainly been

linked to diabetes and its complications.[29] As reviewed in Barlovic et al[30] and Del Turco et al,[31] advanced glycated end products could induce inflammation and oxidative stress by binding to a receptor, receptor of advanced glycated end-products (RAGE), and could also affect extracellular matrix composition. Elevated levels of RAGE have been linked to atherosclerosis in humans.[32] Whether 1-carboxyethyltyrosine could influence myocardial function and the composition of the arterial wall by an advanced glycated end product/RAGE-dependent mechanism remains to be established. Other aromatic amino acids, such as phenylananine and tryptophan, have previously been related to CVD,[33,34] as found in the present study, but reports on tyrosine and CVD are scarce. In a nested case–control study, tyrosine levels were not different between subjects with future CVD and controls.[33]

We demanded a metabolite to show $P<0.05$ also after adjustment for traditional risk factors, including BMI, to qualify as significantly related to incident CVD. The inclusion of BMI as a confounder is of great importance in this setting, because a large number of metabolites are related to obesity, as reviewed in Rangel-Huerta et al.[35] However, we were not interested in only identifying a great number of BMI-associated metabolites, and therefore we adjusted for BMI. This could be the reason why we were not able to replicate some of the findings in previous studies, such as lysophophatidylcholines and sphingomyelins.

We are not aware of any other sample with a similarly high number of metabolites measured at the same platform and with sufficient cases to have an appropriate power that could be used for validation of the findings in the EpiHealth study, so our 37 candidate metabolites have to be reproduced by others in the future to be regarded as validated. However, we would expect that metabolites being related to future CVD to be related to some marker of subclinical CVD, because impairments in those markers usually precede an overt CVD event. We could see that almost all of our 37 metabolites of interest were related to at least 1 or 2 of the 5 markers of subclinical CVD in an independent sample. Although not a formal validation of the 37 metabolites, it is reassuring to see that the vast majority of the metabolites of interest are related markers of subclinical CVD in our supportive analysis in the PIVUS study, and therefore most likely are not mere chance findings.

In the present study we conducted an analysis to evaluate if some metabolites could improve the discrimination of CVD on top of traditional CVD risk factors and found the top 5 metabolites to increase the C statistics by 3.9%. We are fully aware that this is most likely to be an overestimation of the effect of the metabolites, because the selection of metabolites and the C statistic test were performed in the same sample, but we see this test as an inspiration for future studies to improve risk prediction by metabolites. An alternative to validation in an independent sample is to split the present sample in 2 parts and then generate a list of metabolites of interest in 1 subsample and perform the C statistics test in the other subsample. However, because of the limited number of cases, we do not have the power to do so in a meaningful way. As an alternative, we performed a bootstrap analysis of the C statistic test in the total sample and found that the increase in the C statistics found in the traditional analysis was stable across 10 000 repetitive subsamples of the cohort, which would serve as an internal validation. In this study, we limited the number of metabolites to be used in the discrimination analysis to 5 because of the restricted number of CVD cases, although the lasso analysis disclosed 15 metabolites of interest. Thus, in future samples with a higher number of incident cases, it is likely that >5 metabolites could be evaluated to improve discrimination further.

The major strength of the present study is the use of a metabolomic platform that made it possible to evaluate a great number of nonxenobiotic metabolites in the same sample. It is also a strength that we used the same metabolomic platform in another study in which we had measured several markers of subclinical CVD, although that study (PIVUS) did not have sufficient follow-up time to generate the number of cases needed to evaluate incident CVD in a meaningful way.

Some limitations have already been discussed above, and we should also acknowledge that we have been studying almost exclusively individuals with European ancestry, so future studies in other ethnic groups are warranted. The way to calculate ejection fraction from M-mode in PIVUS is rather old fashioned but has been used since the first investigations in PIVUS back in 2001 to be comparable throughout this longitudinal study. If anything, an old-fashioned technique would drive associations toward the null hypothesis and is not likely to produce false-positive results.

In the present study, we used a false discovery rate of <0.05 for the association adjusted for age and sex and $P<0.05$ for the following multiple-adjusted analysis as the criteria of significance for a metabolite. This is a commonly used approach, but it must be emphasized that a couple of the 37 metabolites identified to be of interest could be false-positive findings on multiple-adjusted analysis, so these results should be taken with caution until reproduced by others.

The plasma metabolome is the summary of metabolic processes in a great variety of tissues. If a metabolite is associated (either up- or downregulated) with a certain disease, it cannot convey whether the altered concentration is caused by an increased production or an increased extraction and which tissues are involved. Thus, findings of an altered plasma metabolome in a certain disease should mainly be taken as an

indication that metabolic processes could be altered, but the more detailed knowledge of these processes must be verified in experimental studies or in more sophisticated studies in humans.

The PIVUS sample has a mean age 10 years older than the EpiHealth sample. Ideally, the supportive information gained in the PIVUS study should be derived from a sample with the same age distribution, but we still find it reassuring that most metabolites found to be linked to incident CVD in EpiHealth are linked to deviations in subclinical markers of CVD, even if PIVUS is an older sample.

In conclusion, several metabolites were discovered to be associated with subsequent CVD as well as with subclinical markers of CVD. A selection of those improved discrimination for the prediction of incident cardiovascular events when added to established cardiovascular risk factors.

ARTICLE INFORMATION

Received May 20, 2022; accepted December 2, 2022.

Affiliations

Department of Medical Sciences, Uppsala University, Uppsala, Sweden (L.L., T.F., J.S.); Division of Family Medicine and Primary Care, Department of Neurobiology, Care Science and Society, Karolinska Institutet, Huddinge, Sweden (J.Ä.); School of Health and Social Studies, Dalarna University, Falun, Sweden (J.Ä.); and Division of Geriatric Medicine, Department of Clinical Sciences in Malmö, Lund University, Malmö, Sweden (S.E.).

Sources of Funding

The EpiHealth study is funded as a strategic research area by the Swedish government. PIVUS was funded by Uppsala University Hospital.

Disclosures

None.

Supplemental Material

Tables S1–S3

REFERENCES

1. Lind L, Zanetti D, Ingelsson M, Gustafsson S, Ärnlöv J, Assimes TL. Large-scale plasma protein profiling of incident myocardial infarction, ischemic stroke, and heart failure. *J Am Heart Assoc.* 2021;10:e023330. doi: 10.1161/JAHA.121.023330
2. Lind L, Ärnlöv J, Sundström J. Plasma protein profile of incident myocardial infarction, ischemic stroke, and heart failure in 2 cohorts. *J Am Heart Assoc.* 2021;10:e017900. doi: 10.1161/JAHA.120.017900
3. Ruiz-Canela M, Hruby A, Clish CB, Liang L, Martínez-González MA, Hu FB. Comprehensive metabolomic profiling and incident cardiovascular disease: a systematic review. *J Am Heart Assoc.* 2017;6:e005705. doi: 10.1161/JAHA.117.005705
4. Cruz DE, Tahir UA, Hu J, Ngo D, Chen ZZ, Robbins JM, Katz D, Balasubramanian R, Peterson B, Deng S, et al. Metabolomic analysis of coronary heart disease in an African American cohort from the Jackson heart study. *JAMA Cardiol.* 2022;7:184–194. doi: 10.1001/jamacardio.2021.4925
5. Holmes MV, Millwood IY, Kartsonaki C, Hill MR, Bennett DA, Boxall R, Guo Y, Xu X, Bian Z, Hu R, et al. Lipids, lipoproteins, and metabolites and risk of myocardial infarction and stroke. *J Am Coll Cardiol.* 2018;71:620–632. doi: 10.1016/j.jacc.2017.12.006
6. Floegel A, Kuhn T, Sookthai D, Johnson T, Prehn C, Rolle-Kampczyk U, Otto W, Weikert C, Illig T, von Bergen M, 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. doi: 10.1007/s10654-017-0333-0
7. Seah JYH, Chew WS, Torta F, Khoo CM, Wenk MR, Herr DR, Choi H, Tai ES, van Dam RM. Plasma sphingolipids and risk of cardiovascular diseases: a large-scale lipidomic analysis. *Metabolomics.* 2020;16:89. doi: 10.1007/s11306-020-01709-8
8. Lind L, Salihovic S, Ganna A, Sundström J, Broeckling CD, Magnusson PK, Pedersen NL, Siegbahn A, Prenni J, Fall T, et al. A multi-cohort metabolomics analysis discloses sphingomyelin (32:1) levels to be inversely related to incident ischemic stroke. *J Stroke Cerebrovasc Dis.* 2020;29:104476. doi: 10.1016/j.jstrokecerebrovasdis.2019.104476
9. Levy D, Garrison RJ, Savage DD, Kannel WB, Castelli WP. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham heart study. *N Engl J Med.* 1990;322:1561–1566. doi: 10.1056/NEJM199005313222203
10. Lind L, Sundström J. Change in left ventricular geometry over 10 years in the elderly and risk of incident cardiovascular disease. *J Hypertens.* 2019;37:325–330. doi: 10.1097/HJH.0000000000001897
11. Aurigemma GP, Gottdiener JS, Shemanski L, Gardin J, Kitzman D. Predictive value of systolic and diastolic function for incident congestive heart failure in the elderly: the cardiovascular health study. *J Am Coll Cardiol.* 2001;37:1042–1048. doi: 10.1016/S0735-1097(01)01110-X
12. Gottdiener JS, Kitzman DW, Aurigemma GP, Arnold AM, Manolio TA. Left atrial volume, geometry, and function in systolic and diastolic heart failure of persons > or =65 years of age (the cardiovascular health study). *Am J Cardiol.* 2006;97:83–89. doi: 10.1016/j.amjcard.2005.07.126
13. Wohlin M, Sundström J, Andrén B, Larsson A, Lind L. An echolucent carotid artery intima-media complex is a new and independent predictor of mortality in an elderly male cohort. *Atherosclerosis.* 2009;205:486–491. doi: 10.1016/j.atherosclerosis.2009.01.032
14. Lorenz MW, Polak JF, Kavousi M, Mathiesen EB, Völzke H, Tuomainen TP, Sander D, Plichart M, Catapano AL, Robertson CM, et al. Carotid intima-media thickness progression to predict cardiovascular events in the general population (the PROG-IMT collaborative project): a meta-analysis of individual participant data. *Lancet.* 2012;379:2053–2062. doi: 10.1016/S0140-6736(12)60441-3
15. Lind L, Elmståhl S, Bergman E, Englund M, Lindberg E, Michaëlsson K, Nilsson PM, Sundström J. EpiHealth: a large population-based cohort study for investigation of gene-lifestyle interactions in the pathogenesis of common diseases. *Eur J Epidemiol.* 2013;28:189–197. doi: 10.1007/s10654-013-9787-x
16. Lind L, Fors N, Hall J, Marttala K, Stenborg A. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the prospective investigation of the vasculature in Uppsala seniors (PIVUS) study. *Arterioscler Thromb Vasc Biol.* 2005;25:2368–2375. doi: 10.1161/01.ATV.0000184769.22061.da
17. Lind L, Andersson J, Rönn M, Gustavsson T. The echogenecity of the intima-media complex in the common carotid artery is closely related to the echogenecity in plaques. *Atherosclerosis.* 2007;195:411–414. doi: 10.1016/j.atherosclerosis.2007.03.029
18. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet.* 2014;46:543–550. doi: 10.1038/ng.2982
19. Trainor PJ, Hill BG, Carlisle SM, Rouchka EC, Rai SN, Bhatnagar A, DeFilippis AP. Systems characterization of differential plasma metabolome perturbations following thrombotic and non-thrombotic myocardial infarction. *J Proteomics.* 2017;160:38–46. doi: 10.1016/j.jprot.2017.03.014
20. Paganelli F, Frachebois C, Velut JG, Boullu S, Sauze N, Rosso JP, Barnay P, Sbragia P, Gelisse R, Grino M, et al. Hypothalamo-pituitary-adrenal axis in acute myocardial infarction treated by percutaneous transluminal coronary angioplasty: effect of time of presentation. *J Endocrinol Invest.* 2003;26:407–413. doi: 10.1007/BF03345195
21. Corona G, Rastrelli G, Monami M, Guay A, Buvat J, Sforza A, Forti G, Mannucci E, Maggi M. Hypogonadism as a risk factor for cardiovascular mortality in men: a meta-analytic study. *Eur J Endocrinol.* 2011;165:687–701. doi: 10.1530/EJE-11-0447
22. Dai W, Wagh SA, Chettiar S, Zhou GD, Roy R, Qiao X, Visich PS, Hoffman EP. Blunted circadian cortisol in children is associated with poor cardiovascular health and may reflect circadian misalignment. *Psychoneuroendocrinology.* 2021;129:105252. doi: 10.1016/j.psyneuen.2021.105252

23. Rosmond R, Björntorp P. The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. *J Intern Med.* 2000;247:188–197. doi: 10.1046/j.1365-2796.2000.00603.x
24. Koukoulis GN, Filiponi M, Gougoura S, Befani C, Liakos P, Bargiota A. Testosterone and dihydrotestosterone modulate the redox homeostasis of endothelium. *Cell Biol Int.* 2022;46:660–670. doi: 10.1002/cbin.11768
25. Li M, Hu X, Fan Y, Li K, Zhang X, Hou W, Tang Z. Hyperuricemia and the risk for coronary heart disease morbidity and mortality: a systematic review and dose-response meta-analysis. *Sci Rep.* 2016;6:19520. doi: 10.1038/srep19520
26. Yang F, Lu Y, Chen S, Wang K, Hu T, Cui H. Sex-specific effect of serum urate levels on coronary heart disease and myocardial infarction prevention: a Mendelian randomization study. *Nutr Metab Cardiovasc Dis.* 2022;32:1266–1274. doi: 10.1016/j.numecd.2022.01.022
27. Battelli MG, Bolognesi A, Polito L. Pathophysiology of circulating xanthine oxidoreductase: new emerging roles for a multi-tasking enzyme. *Biochim Biophys Acta.* 2014;1842:1502–1517. doi: 10.1016/j.bbadis.2014.05.022
28. Inaba S, Sautin Y, Garcia GE, Johnson RJ. What can asymptomatic hyperuricaemia and systemic inflammation in the absence of gout tell us? *Rheumatology (Oxford).* 2013;52:963–965. doi: 10.1093/rheumatology/ket001
29. Goldin A, Beckman JA, Schmidt AM, Creager MA. Advanced glycation end products: sparking the development of diabetic vascular injury. *Circulation.* 2006;114:597–605. doi: 10.1161/CIRCULATIONAHA.106.621854
30. Barlovic DP, Soro-Paavonen A, Jandeleit-Dahm KA. RAGE biology, atherosclerosis and diabetes. *Clin Sci (Lond).* 2011;121:43–55. doi: 10.1042/CS20100501
31. Del Turco S, Basta G. An update on advanced glycation endproducts and atherosclerosis. *Biofactors.* 2012;38:266–274. doi: 10.1002/biof.1018
32. Sandstedt J, Vargmar K, Björkman K, Ruetschi U, Bergström G, Hultén LM, Skiöldebrand E. COMP (cartilage Oligomeric matrix protein) Neoepitope: a novel biomarker to identify symptomatic carotid stenosis. *Arterioscler Thromb Vasc Biol.* 2021;41:1218–1228. doi: 10.1161/ATVBAHA.120.314720
33. Kume S, Araki S, Ono N, Shinhara A, Muramatsu T, Araki H, Isshiki K, Nakamura K, Miyano H, Koya D, et al. Predictive properties of plasma amino acid profile for cardiovascular disease in patients with type 2 diabetes. *PLoS One.* 2014;9:e101219. doi: 10.1371/journal.pone.0101219
34. Würtz P, Havulinna AS, Soininen P, Tynkkynen T, Prieto-Merino D, Tillin T, Ghorbani A, Artati A, Wang Q, Tiainen M, et al. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation.* 2015;131:774–785. doi: 10.1161/CIRCULATIONAHA.114.013116
35. Rangel-Huerta OD, Pastor-Villaescusa B, Gil A. Are we close to defining a metabolomic signature of human obesity? A systematic review of metabolomics studies. *Metabolomics.* 2019;15:93. doi: 10.1007/s11306-019-1553-y

SUPPLEMENTAL MATERIAL

Table S1. Overview of published studies relating metabolomic data to incident CVD. PMID= PubMed identifier.

Publication	PMID	No. of metabolites evaluated	Major findings
Shah 2010	20173117	60	Short- and medium-chained acetylcarnitines
Wang 2011	21475195	Untargeted	Choline, Betaine, TMAO
Shah 2012	22607863	60	Dicarboxylacylcarnitines, fatty acids, proline
Kalim 2013	24308938	165	Carnitines, TMAO
Rizza 2014	24468136	48	Medium-long-chain acylcarnitines, Alanine
Vaarhorst 2014	24952859	Untargeted	creatinine, serine, glucose, 1,5-anhydrosorbitol, TMAO, ornithine, citrate, glutamate, glycoproteins, valine
Stegman 2014	24622385	135	Triacylglycerol 54:2, Cholesterol ester 16:1, Phosphatidylethanolamine 36:5
Ganna 2014	25502724	Untargeted	Lysophosphatidylcholine 18:2, Lysophosphatidylcholine 18:1, Monoglyceride 18:2, Sphingomyelin 28:1
Kume 2014	24971671	31	ethanolamine, hydroxyproline, glutamic acid, 3-methylhistidine, tryptophan
Zeng 2014	24760976	356	γ-Glutamyl dipeptide pathway, Lysophosphatidylcholines, 2-Hydroxybutyrates
Wurtz 2015	25573147	68	Phenylalanine, Monounsaturated fatty acids, ω-6 fatty acids, Docosahexaenoic acid
Alsherhy 2016	27756783	310	Hexosylceramides, Phosphatidylcholines, Cholesteryl Ester (16:0), Triacylglycerol (56:6)

Floegel 2018	29181692	105	sphingomyelins, diacyl-phosphatidylcholines, acyl-alkyl-phosphatidylcholines
Holmes 2018	29420958	225	Lipoproteins, Glycoprotein acetyls, ketone bodies, glucose, docosahexaenoic acid
Stenemo 2019	31148414	206	Urobilin, sphingomyelin (30:1)
Lind 2020	31806450	204	sphingomyelin (32:1)
Seah 2020	32816082	79	Monohexosylceramides, 18:1 sphingolipids
Tahir 2021	33464957	312	Homoarginine, 2-aminoadipic acid, uridine, methylhistidine, 4-acetamidobutanoate, choline, ectoine, *N*-acetylspermidine, dimethylguanidino valeric acid (DMGV), *N*-acetyl-l-alanine, and *N2-N2* dimethylguanosine
Cruz 2022	34851361	303	N-acylamides, leucine, lipid species

Table S2. Subpathway and metabolite names for the 37 metabolites being related to incident CVD given in Figure 1 with the start from 1-stearoyl-2-oleoyl-GPE (18:0/18:1) at the top of the figure and then given in clockwise order.

Subpathway	Metabolite
Phosphatidylethanolamine (PE)	1-stearoyl-2-oleoyl-GPE (18:0/18:1)
Phosphatidylethanolamine (PE)	1-stearoyl-2-arachidonoyl-GPE (18:0/20:4)
Phosphatidylethanolamine (PE)	1-palmitoyl-2-arachidonoyl-GPE (16:0/20:4)
Phosphatidylethanolamine (PE)	1-stearoyl-2-linoleoyl-GPE (18:0/18:2)
Phosphatidylethanolamine (PE)	1-palmitoyl-2-oleoyl-GPE (16:0/18:1)
Phosphatidylethanolamine (PE)	1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)
Pregnenolone Steroids	pregnenediol sulfate (C21H34O5S)
Pregnenolone Steroids	17alpha-hydroxypregnenolone 3-sulfate
Pregnenolone Steroids	pregnenolone sulfate
Pregnenolone Steroids	pregnenetriol sulfate
Androgenic Steroids	dehydroepiandrosterone sulfate (DHEA-S)
Ceramides	N-palmitoyl-sphingosine (d18:1/16:0)
Dihydroceramides	N-palmitoyl-sphinganine (d18:0/16:0)
Fatty Acid, Dihydroxy	3,4-dihydroxybutyrate
Progestin Steroids	5alpha-pregnan-3beta,20alpha-diol monosulfate
Methionine, Cysteine, SAM and Taurine Metabolism	N-formylmethionine
Methionine, Cysteine, SAM and Taurine Metabolism	2,3-dihydroxy-5-methylthio-4-pentenoate (DMTPA)
Methionine, Cysteine, SAM and Taurine Metabolism	N-acetylmethionine
Glycine, Serine and Threonine Metabolism	N-acetylserine
Glycine, Serine and Threonine Metabolism	dimethylglycine
Alanine and Aspartate Metabolism	N-acetylalanine
Histidine Metabolism	imidazole propionate
Leucine, Isoleucine and Valine Metabolism	N-acetylvaline
Phenylalanine Metabolism	phenylalanine
Polyamine Metabolism	5-methylthioadenosine (MTA)
Tryptophan Metabolism	C-glycosyltryptophan
Tyrosine Metabolism	1-carboxyethyltyrosine
Purine Metabolism, Adenine containing	N6-carbamoylthreonyladenosine

Purine Metabolism, Guanine containing	7-methylguanine
Pyrimidine Metabolism, Cytidine containing	N4-acetylcytidine
Pyrimidine Metabolism, Uracil containing	5,6-dihydrouridine
Gamma-glutamyl Amino Acid	gamma-glutamyltryptophan
Gamma-glutamyl Amino Acid	gamma-glutamylphenylalanine
Gamma-glutamyl Amino Acid	gamma-glutamyltyrosine
Ascorbate and Aldarate Metabolism	gulonate
Nicotinate and Nicotinamide Metabolism	quinolinate
Aminosugar Metabolism	N-acetylglucosamine/N-acetylgalactosamine

Table S3. Relationships between metabolites being related to incident CVD in Table 2 and markers of subclinical CVD in the PIVUS study.

	Age and sex-adjusted			Multiple adjusted		
	Beta	SE	p-value	Beta	SE	p-value
Left ventricular mass index						
N-acetylvaline	.13	.043	.0031	-.0031	.041	.94
phenylalanine	.112	.044	.013	-.033	.042	.43
5-methylthioadenosine (MTA)	.122	.044	.0050	-.022	.043	.60
1-carboxyethyltyrosine	.19	.047	.000077	.017	.05	.72
N-palmitoyl-sphinganine (d18:0/16:0)	.10	.045	.025	.024	.045	.59
3,4-dihydroxybutyrate	.086	.043	.042	-.0083	.040	.84
1-stearoyl-2-oleoyl-GPE (18:0/18:1)	.11	.043	.013	.0032	.041	.93
1-palmitoyl-2-oleoyl-GPE (16:0/18:1)	.12	.043	.0060	.013	.040	.73
1-palmitoyl-2-arachidonoyl-GPE (16:0/20:4)	.10	.044	.022	.045	.039	.25
1-stearoyl-2-arachidonoyl-GPE (18:0/20:4)	.133	.044	.0027	.053	.041	.19
gamma-glutamyltyrosine	.010	.043	.023	-.017	.039	.66
Ejection fraction						
N-acetylalanine	-.15	.025	4.6e-09	-.107	.027	.000078
N-acetylserine	-.15	.025	5.9e-09	-.113	.027	.000023
dimethylglycine	-.071	.026	.0072	-.051	.026	.054
imidazole propionate	-.11	.027	.000037	-.069	.028	.013
N-acetylvaline	-.15	.025	2.9e-09	-.10	.027	.00012
N-formylmethionine	-.17	.025	1.7e-11	-.14	.026	6.5e-08
2,3-dihydroxy-5-methylthio-4-pentenoate (DMTPA)	-.18	.026	7.3e-12	-.14	.028	8.4e-07
N-acetylmethionine	-.14	.026	1.2e-07	-.11	.026	.000057
phenylalanine	-.14	.026	9.9e-08	-.10	.027	.00015

5-methylthioadenosine (MTA)	-.17	.026	5.0e-11	-.13	.028	5.4e-06
C-glycosyltryptophan	-.13	.025	1.1e-07	-.11	.026	.000049
1-carboxyethyltyrosine	-.17	.028	8.2e-10	-.13	.032	.000038
N-acetylglucosamine/N-acetylgalactosamine	-.088	.025	.00049	-.063	.026	.017
gulonate	-.13	.025	2.6e-07	-.091	.026	.00048
quinolinate	-.12	.025	3.9e-06	-.079	.027	.0032
dehydroepiandrosterone sulfate (DHEA-S)	.06	.026	.021	.039	.026	.13
3,4-dihydroxybutyrate	-.099	.026	.00013	-.052	.027	.058
1-palmitoyl-2-oleoyl-GPE (16:0/18:1)	-.11	.026	.000023	-.12	.026	.00012
1-stearoyl-2-linoleoyl-GPE (18:0/18:2)	-.086	.026	.00092	-.088	.027	.0011
1-palmitoyl-2-arachidonoyl-GPE (16:0/20:4)	-.12	.026	4.9e-06	-.11	.026	.000011
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	-.086	.026	.00078	-.092	.026	.00040
1-stearoyl-2-arachidonoyl-GPE (18:0/20:4)	-.12	.026	6.7e-06	-.11	.027	.000067
1-stearoyl-2-oleoyl-GPE (18:0/18:1)	-.092	.026	.00037	-.082	.027	.0021
pregnenolone sulfate	.076	.027	.0046	.055	.027	.043
pregnenetriol sulfate	.064	.028	.021	.058	.028	.034
pregnenediol sulfate (C21H34O5S)	.069	.028	.012	.056	.028	.044
17alpha-hydroxypregnenolone 3-sulfate	.072	.032	.026	.054	.033	.095
N6-carbamoylthreonyladenosine	-.12	.026	2.5e-06	-.096	.028	.00051
7-methylguanine	-.11	.027	.000022	-.089	.027	.00095
N4-acetylcytidine	-.13	.026	14e-07	-.095	.027	.00048
5,6-dihydrouridine	-.16	.025	1.8e-10	-.13	.026	7.9e-07
gamma-glutamylphenylalanine	-.14	.026	5.7e-08	-.106	.027	.00011
gamma-glutamyltryptophan	-.069	.026	.0067	-.048	.025	.060
gamma-glutamyltyrosine	-.095	.025	.00018	-.069	.026	.0076
Left atrium diameter						
N-acetylalanine	.18	.039	5.8e-06	.043	.039	.30

N-acetylserine	.18	.039	2.9e-06	.074	.038	.053
imidazole propionate	.12	.042	.0034	.021	.041	.60
N-acetylvaline	.20	.039	1.9e-07	.072	.039	.063
N-acetylmethionine	.15	.040	.00022	.056	.038	.14
N-formylmethionine	.17	.039	.000010	.091	.037	.013
2,3-dihydroxy-5-methylthio-4-pentenoate (DMTPA)	.20	.041	7.9e-07	.045	.04	.25
phenylalanine	.17	.040	9.9e-06	.046	.039	.23
5-methylthioadenosine (MTA)	.23	.039	3.8e-09	.092	.040	.022
C-glycosyltryptophan	.13	.039	.0010	.027	.038	.47
1-carboxyethyltyrosine	.23	.043	7.8e-08	.10	.046	.025
N-acetylglucosamine/N-acetylgalactosamine	.11	.039	.0059	.044	.036	.22
gulonate	.13	.039	.0011	.026	.037	.48
quinolinate	.16	.039	.000052	.028	.039	.47
3,4-dihydroxybutyrate	.12	.039	.0016	.023	.038	.55
1-stearoyl-2-linoleoyl-GPE (18:0/18:2)	.13	.040	.00089	.13	.039	.00077
1-stearoyl-2-oleoyl-GPE (18:0/18:1)	.14	.040	.00046	.11	.038	.0047
1-palmitoyl-2-arachidonoyl-GPE (16:0/20:4)	.13	.041	.0014	.12	.037	.0015
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	.11	.04+	.0071	.12	.037	.0012
1-palmitoyl-2-oleoyl-GPE (16:0/18:1)	.14	.039	.00030	.10	.037	.0066
1-stearoyl-2-arachidonoyl-GPE (18:0/20:4)	.15	.040	.00031	.11	.038	.0031
pregnenolone sulfate	-.14	.041	.00093	-.066	.039	.090
pregnenediol sulfate (C21H34O5S)	-.11	.043	.0097	-.062	.040	.11
N6-carbamoylthreonyladenosine	.18	.041	.000011	.057	.041	.14
7-methylguanine	.12	.041	.0024	.046	.038	.22
N4-acetylcytidine	.19	.039	1.2e-06	.055	.039	.16
5,6-dihydrouridine	.17	.039	.000023	.051	.038	.18
gamma-glutamyltyrosine	.13	.039	.00082	.026	.037	.48

gamma-glutamylphenylalanine	.198	.039	4.9e-07	.055	.039	.15
Intima-media thickness						
gamma-glutamyltryptophan	-.137	.041	.00077	-.14	.041	.00078
1-palmitoyl-2-linoleoyl-GPE (16:0/18:2)	.093	.041	.024	.065	.042	.12
N-acetylmethionine	-.089	.041	.030	-.087	.043	.041
1-stearoyl-2-linoleoyl-GPE (18:0/18:2)	.088	.042	.036	.065	.043	.12
Intima-media echogenicity (IM-GSM)						
1-carboxyethyltyrosine	-.217	.044	8.3e-07	-.10	.051	.041
N-palmitoyl-sphinganine (d18:0/16:0)	-.15	.042	.00037	-.083	.043	.050
dimethylglycine	.125	.041	.0023	.15	.042	.00011
gamma-glutamyltyrosine	-.119	.041	.0037	-.023	.041	.576
1-stearoyl-2-oleoyl-GPE (18:0/18:1)	-.105	.041	.011	-.016	.041	.69
1-stearoyl-2-linoleoyl-GPE (18:0/18:2)	-.096	.042	.021	-.028	.041	.49
phenylalanine	-.093	.042	.028	.011	.043	.79
pregnenetriol sulfate	-.096	.044	.028	-.052	.044	.25
gamma-glutamyltryptophan	-.086	.042	.039	-.014	.041	.72

Only metabolites with $p<0.05$ in the age- and sex-adjusted analysis are shown.