

Metabolites Associated With the Risk of Incident Venous Thromboembolism: A Metabolomic Analysis

Xia Jiang, PhD; Oana A. Zeleznik, PhD; Sara Lindström, PhD; Jessica Lasky-Su, ScD; Kaitlin Hagan, PhD; Clary B. Clish, PhD; A. Heather Eliassen, ScD; Peter Kraft, PhD; Christopher Kabrhel, MD, MPH

Background—Venous thromboembolism (VTE) is a complex thrombotic disorder that constitutes a major source of mortality and morbidity. To improve understanding of the cause of VTE, we conducted a metabolomic analysis in a case-control study including 240 incident VTE cases and 6963 controls nested within 3 large prospective population-based cohorts, the Nurses' Health Study, the Nurses' Health Study II, and the Health Professionals Follow-Up Study.

Methods and Results—For each individual, we measured 211 metabolites and collected detailed information on lifestyle factors. We performed logistic regression and enrichment analysis to identify metabolites and biological categories associated with incident VTE risk, accounting for key confounders, such as age, sex, smoking, alcohol consumption, body mass index, and comorbid diseases (eg, cancers). We performed analyses of all VTEs and separate analyses of pulmonary embolism. Using the basic model controlling for age, sex, and primary disease, we identified 60 nominally significant VTE- or pulmonary embolism-associated metabolites ($P < 0.05$). These metabolites were enriched for diacylglycerols ($P_{\text{permutation}} < 0.05$). However, after controlling for multiple testing, only 1 metabolite (C5 carnitine; odds ratio, 1.25; 95% confidence interval, 1.10–1.41; $P_{\text{corrected}} = 0.03$) remained significantly associated with VTE. After further adjustment for body mass index, no metabolites were significantly associated with disease after accounting for multiple testing, and no metabolite classes were enriched for nominally significant associations.

Conclusions—Although our findings suggest that circulating metabolites may influence the risk of incident VTE, the associations we observed were confounded by body mass index. Larger studies involving additional individuals and with broader metabolomics coverage are needed to confirm our findings. (*J Am Heart Assoc.* 2018;7:e010317. DOI: 10.1161/JAHA.118.010317)

Key Words: carnitine • incidence • lipids • metabolite • metabolome • pulmonary embolism • venous thromboembolism

Venous thromboembolism (VTE), a spectrum of disease that comprises deep venous thrombosis (DVT) and pulmonary embolism (PE), is a common and potentially lethal medical condition. VTE accounts for 100 000 annual deaths in the United States¹ and is a major source of morbidity. With its high early mortality, sustained morbidity, and high incidence (1–2/1000 Americans per year),² the primary prevention of VTE remains an issue of major public health importance; and an in-depth understanding of disease

physiological characteristics is required to improve primary prevention.

VTE is a multifactorial disorder involving complex genetic, metabolic, and environmental interactions. Despite the many VTE-associated factors that have been identified, the biological mechanisms by which most of them affect disease risk remain unexplained. For example, although VTE susceptibility genes discovered through earlier studies largely involve direct alterations in the clotting system,^{3,4} more recent genome-

From the Program in Genetic Epidemiology and Statistical Genetics (X.J., P.K.) and Department of Epidemiology (A.H.E.), Harvard T.H. Chan School of Public Health, Boston, MA; Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden (X.J.); Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA (O.A.Z., J.L.-S., K.H., C.K.); Departments of Medicine (O.A.Z., K.H.) and Emergency Medicine (C.K.), Harvard Medical School, Boston, MA; Epidemiology, University of Washington, Seattle, WA (S.L.); Broad Institute, Cambridge, MA (C.B.C.); Department of Emergency Medicine, Center for Vascular Emergencies, Massachusetts General Hospital, Boston, MA (C.K.); and Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA (S.L.).

Accompanying Tables S1 through S3 and Figure S1 are available at <https://www.ahajournals.org/doi/suppl/10.1161/JAHA.118.010317>

Correspondence to: Christopher Kabrhel, MD, MPH, Department of Emergency Medicine, Center for Vascular Emergencies, Massachusetts General Hospital, Zero Emerson Place, Suite 3B, Boston, MA 02114. E-mail: ckabrhel@partners.org

Received July 10, 2018; accepted October 16, 2018.

© 2018 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Clinical Perspective

What Is New?

- We performed a population-based case-control study to systematically investigate the relationship between circulating metabolites and risk of incident venous thromboembolism (VTE).
- After controlling for common confounders, such as age, sex, and provoking diseases, we found that C5 carnitine was significantly associated with incident VTE and that the metabolomics category diacylglycerols were enriched in both VTE and pulmonary embolism.

What Are the Clinical Implications?

- By demonstrating nominal associations in prospectively collected data, our findings provide preliminary evidence in support of an altered physiological state before the VTE event and may specifically inform future research into metabolic changes and pathophysiological characteristics of VTE.

wide association studies have discovered loci (*TSPAN15* and *SLC44A2*) outside of the coagulation cascade.^{5–7} Similarly, the mechanism by which most environmental factors trigger VTE remains unclear.^{8–14} Thus, although identifying risk factors for VTE represents important progress, there are still major gaps in our understanding of VTE pathophysiological characteristics.

Metabolites are the final downstream products of the genome and are shaped by environmental factors; thus, they are well suited to reveal the physiologic state before disease onset as well as the complex interplay between genes, environmental risk factors, and VTE. However, although high-throughput metabolomic methods have identified metabolites contributing to cardiovascular disease risk,^{15–19} this approach has not been widely applied to VTE. To date, only one retrospective pilot study of 40 male VTE cases and 40 controls, with blood samples collected after VTE diagnosis, has been published.²⁰

The aim of this study was to analyze 240 incident VTE cases and 6963 controls with available prediagnostic metabolomics data, nested within 3 well-established large-scale population cohorts, to identify metabolites associated with VTE and to improve our understanding of VTE pathophysiological characteristics.

Methods

The data, analytic methods, and study materials will not be made available to other researchers for purposes of reproducing the results or replicating the procedure. All codes are

available and can be provided by the authors on request (xiajiang@hsph.harvard.edu).

Study Population

We performed a large-scale case-control study consisting of 7203 individuals nested within 3 prospective cohorts, the NHS (Nurses' Health Study), the NHS II, and the HPFS (Health Professionals Follow-Up Study), to identify associations of specific metabolites and metabolite profiles with the risk of incident VTE. Details on characteristics of each cohort have been described elsewhere.^{21–23} Briefly, the NHS was established in 1976, when 121 700 female registered nurses, 30 to 55 years of age, completed an initial questionnaire. During 1989 to 1990, blood samples were collected from 32 826 NHS participants.²⁴ The NHS II was established in 1989, when 116 429 female registered nurses, 25 to 42 years of age, completed an initial questionnaire. During 1996 to 1999, blood samples were collected from 29 611 NHS II participants.²⁵ The HPFS was established in 1986, when 51 529 male dentists, optometrists, osteopaths, podiatrists, pharmacists, and veterinarians in the United States, 40 to 75 years of age, completed an initial questionnaire. During 1993 to 1995, blood samples were collected from 18 225 HPFS participants.²⁶ Participants in the 3 cohorts have been followed up by mailed questionnaires every 2 years to update exposure information and ascertain nonfatal incident diseases. The follow-up rates for each biennial cycle have consistently been >90%.

We identified self-reported incident VTE cases (first event of either PE or DVT) among individuals who were free of cardiovascular disease at the time of blood draw. Specifically, participants were asked every 2 years whether a physician had diagnosed them with PE since the prior questionnaire. Moreover, in NHS, women reported DVT events by writing in DVT as an "other major illness" occurring since the prior questionnaire. In NHS II and HPFS, participants were asked whether a physician had diagnosed them with DVT since the prior questionnaire.

In the current study, we included 240 incident VTE cases (of which, 125 were incident PE cases) and 6963 controls, for whom metabolomics data had been generated through previous case-control studies of other phenotypes, including amyotrophic lateral sclerosis, rheumatoid arthritis, Parkinson disease, diabetes mellitus, chronic stress, and ovarian, pancreatic, and prostate cancer. For all subjects, blood samples were collected before VTE diagnosis. All subjects were of European ancestry. Study protocols were approved by the Institutional Review Boards of Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health. This study and the consent process for the use of blood samples were approved by the Human Research Committee of

Partners Healthcare and the Institutional Review Board of the T.H. Chan Harvard School of Public Health (protocols 2013P000713 [Genetic Risk Factors for Venous Thromboembolism] and 2006P001122 [Pulmonary Embolism in the NHS, NHS II, HPFS, and UK Biobank]). Cohort subjects received a letter informing them that blood samples would be used for genetic and other -omic research and were given the opportunity to opt out of such research. Any participants who opted out were not included in this analysis.

Metabolomics Assays

The blood samples were collected in heparin tubes from women in NHS and in EDTA tubes from men in HPFS. Participants arranged to have their blood collected and shipped the sample via overnight mail with an icepack to a central laboratory. On arrival, the samples were processed, and aliquots of plasma, white blood cells, and red blood cells were stored in a liquid nitrogen freezer. More than 95% of samples arrived within 24 hours of collection.²⁷ As mentioned, plasma metabolomics data were combined from a series of smaller studies of NHS, NHS II, and HPFS samples conducted over a period of several years. Data from each substudy were log transformed and converted to Z scores before merging the data sets. Metabolomics data were acquired using several liquid chromatography–tandem mass spectrometry methods designed to measure polar metabolites and lipids, as described in detail previously.^{28–30} Briefly, high-resolution, accurate mass profiling data were acquired using liquid chromatography–tandem mass spectrometry systems composed of Nexera X2 U-HPLC systems (Shimadzu Corp, Marlborough, MA) coupled to Q Exactive and Exactive Plus orbitrap mass spectrometers (Thermo Fisher Scientific, Waltham, MA). Polar metabolites were separated using hydrophilic interaction liquid chromatography with positive ion mode mass spectrometry detection, whereas C18 chromatography with negative ion mode detection and C8 chromatography with positive ion mode detection were used to profile metabolites of intermediate polarity and lipids, respectively. Raw data were processed using TraceFinder software (Thermo Fisher Scientific) and Progenesis Q1 (Non-linear Dynamics, Newcastle upon Tyne, UK). In addition, pooled plasma samples were analyzed after intervals of ≈ 20 participant samples to enable standardizing temporal drift in instrument response over time and between batches. For each method, metabolite identities were confirmed using authentic reference standards or reference samples.

In the current study, we excluded unknown metabolites and metabolites unlikely to be informative in analysis with an intraclass correlation (ICC) < 0.6 or with no variance; we further excluded nonparticipant samples, such as drift pool samples. Finally, we included 211 named metabolites that

were measured in all 3 cohorts. Most of these metabolites (81%) were lipids or lipid intermediates.

Statistical Analyses

We conducted unconditional logistic regression to identify metabolites that were associated with the risk of VTE and the PE subtype. The relative concentration of each metabolite was initially transformed using natural logarithm, and then scaled to a mean value=0 and an SD=1. We quantified the odds ratios (ORs) and 95% confidence intervals for the relationship between each metabolite (one metabolite at a time), modeled as a continuous variable, and disease status (VTE or PE), modeled as a binary variable, while controlling for potential confounders. We first performed the association analysis using all subjects in a “basic” model adjusted for age, sex, and provoking disease (whether the participant was a case or control in case-control studies of other phenotypes). We then fitted a “full” model incorporating additional covariates, including alcohol consumption, smoking status, body mass index (BMI), and fasting status at blood draw, all measured at the time of blood draw (before VTE diagnosis).

We further grouped metabolites into categories on the basis of their biological class and identified categories that were enriched (or depleted) for nominal associations with VTE (or PE). We compared the observed hypergeometric *P* values by testing the null hypothesis that the probability of a nominally significant association did differ across classes to those calculated in 5000 simulated null data sets in which case status had been randomly permuted within sex strata. The permutation *P* value for each class was calculated as $k/5001$, where *k* is the number of permutations, where $P_{\text{permuted}} \leq P_{\text{observed}}$.

Principal component (PC) analysis redefines a data set into a new set of variables (PCs) that are mutually orthogonal linear combinations of the original variables. We next applied this PC analysis approach to convert the 211 correlated metabolites into a set of linearly uncorrelated variables (PCs). We used a method named “singular value decomposition” to examine the covariance/correlations between individuals and the built-in R function “prcomp” to perform the analysis. We extracted the first several PCs (*N*=19 PCs) that explained at least 80% of the total variance and repeated the association analyses (the same crude and full models as previously mentioned) using PCs as exposures instead of the metabolites.

Finally, we performed 2 sensitivity analyses. In our first sensitivity analysis, we used multiple imputation to account for missing information for some covariates and metabolites, using “fully conditional specification” models through the R package “mice” (multivariate imputation by chained equations).³¹ This procedure predicts the missing value of a variable using existing values from other variables, thereby allowing individuals who previously lacked data back into the analysis, with covariate values similar to those actually observed among individuals with

Table 1. Characteristics of Study Subjects by Disease Status

| Characteristics | VTE | PE | Controls | P Values | |
|-----------------------------------|-------------|-------------|-------------|------------------------|------------------------|
| | (N=240) | (N=125) | (N=6963) | VTE vs Controls | PE vs Controls |
| Age, mean±SD, y | 56.2 (9.4) | 55.4 (8.9) | 52.7 (9.8) | 5.2×10^{-8} * | 0.0011* |
| Sex | | | | | |
| Women | 170 (70.8) | 100 (80.0) | 5744 (82.5) | 0.0005* | 0.48 |
| Men | 70 (29.2) | 25 (20.0) | 1219 (17.5) | | |
| Smoking status | | | | | |
| Current | 18 (7.6) | 5 (4.0) | 654 (9.4) | 0.53 | 0.12 |
| Past | 89 (37.7) | 45 (36.3) | 2427 (35.0) | | |
| Never | 129 (54.7) | 74 (59.7) | 3854 (55.6) | | |
| Pack-years, mean±SD | 25.4 (20.5) | 25.4 (21.0) | 19.4 (17.0) | 0.0039* | 0.06 |
| Alcohol consumption, mean±SD, g/d | 9.7 (11.0) | 7.5 (9.0) | 9.5 (12.0) | 0.85 | 0.09 |
| Fasting status at blood draw | | | | | |
| Yes | 175 (72.9) | 95 (76.0) | 5154 (74.0) | 0.71 | 0.69 |
| No | 65 (27.1) | 30 (24.0) | 1808 (26.0) | | |
| BMI, kg/m ² | | | | | |
| <25 | 73 (30.8) | 32 (25.8) | 3526 (51.5) | 0.0005* | 0.0005* |
| 25–29 | 105 (44.3) | 57 (46.0) | 2193 (32.1) | | |
| ≥30 | 59 (24.9) | 35 (28.2) | 1121 (16.4) | | |
| Mean±SD | 27.6 (5.4) | 28.3 (5.1) | 25.8 (5.0) | 5.1×10^{-7} * | 2.6×10^{-7} * |
| With an end point [†] | | | | | |
| Yes | 133 (56.8) | 81 (65.9) | 3159 (48.2) | 0.013* | 0.0005* |
| No | 101 (43.2) | 42 (34.1) | 3396 (51.8) | | |

Data are given as number (percentage) unless otherwise indicated. BMI indicates body mass index; PE, pulmonary embolism; VTE, venous thromboembolism.

* $P < 0.05$. PE is a subset of VTE.

[†]With an endpoint was a case in one of the previous case-control studies of other phenotypes (amyotrophic lateral sclerosis, rheumatoid arthritis, Parkinson disease, diabetes mellitus, chronic stress, and ovarian, pancreatic, and prostate cancer) from which our study population was derived.

the same pattern of other covariates. We performed 5 rounds of multiple imputations. Predictive mean matching was used for numeric variables, such as metabolites; logistic regression was used for binary variables with 2 levels; and bayesian polytomous regression was used for factor variables with >2 levels. We combined and averaged the estimates from 5 repeated complete data analyses into a single set of estimates and standard errors by Rubin's rules. In our second sensitivity analysis, we limited the analysis to subjects who were controls in the studies from which they were derived, to reduce the impact of confounding by comorbid illnesses, especially cancers.

Results

Characteristics of Study Subjects

Metabolomics assays of serum samples from 240 incident VTE cases (125 incident PE cases) and 6963 controls resulted in 211 named metabolites measured in all 3 cohorts (the study flow is shown in Figure S1). Those metabolites cover

mainly lipids and lipid intermediates (ie, 12 diacylglycerols, 42 triacylglycerols, 23 phosphatidylethanolamines, 36 phosphatidylcholines, 12 cholesterol esters, 7 sphingomyelins, 9 lysophosphatidylcholines, 6 lysophosphatidylethanolamines, and 22 carnitines), as well as other small molecules, such as amino acids and nucleotides (Tables S1 and S2). As shown in Table 1, compared with the controls and consistent with the epidemiological characteristics of VTE, we found that VTE (and PE) cases were significantly older in age, were more likely to be men, had more prevalent comorbidities (ie, cancers), and were more likely to be overweight. Despite the comparable proportions of current and past smokers observed among cases and controls, the cases smoked more heavily.

Metabolites Associated With VTE or PE in the Basic Model

Using multivariable logistic regression models adjusted for age, sex, and original study end point (basic model), we identified 50 metabolites that were nominally associated with VTE ($P < 0.05$).

Table 2. Metabolites Significantly Associated With VTE or PE, Models Adjusted for Age, Sex, and End Points

| Metabolites | VTE | | | | | PE | | | | |
|--------------------------------|------------------|---------------|-------------------|-------|------|------------------|---------------|-------------------|------|------|
| | OR (95% CI) | Crude P Value | Corrected P Value | | | OR (95% CI) | Crude P Value | Corrected P Value | | |
| | | | TPFP | FWR | FDR | | | TPFP | FWR | FDR |
| C5 carnitine | 1.25 (1.10–1.41) | 0.0005* | 0.07 | 0.03* | 0.07 | 1.24 (1.04–1.47) | 0.015* | 0.70 | 0.73 | 1.00 |
| C34:2 diacylglycerol | 1.30 (1.10–1.55) | 0.003* | 0.33 | 0.11 | 0.67 | 1.46 (1.13–1.88) | 0.004* | 0.36 | 0.35 | 0.59 |
| C34:1 diacylglycerol | 1.29 (1.09–1.53) | 0.003* | 0.38 | 0.16 | 0.59 | 1.36 (1.06–1.74) | 0.015* | 0.66 | 0.76 | 1.00 |
| C36:1 phosphatidylethanolamine | 1.31 (1.09–1.57) | 0.003* | 0.07 | 0.09 | 0.36 | 1.39 (1.07–1.82) | 0.015* | 0.30 | 0.46 | 0.59 |
| C36:1 diacylglycerol | 1.28 (1.08–1.51) | 0.004* | 0.38 | 0.22 | 0.68 | 1.33 (1.05–1.70) | 0.020* | 0.64 | 0.72 | 1.00 |
| C52:2 triacylglycerol | 1.31 (1.09–1.58) | 0.004* | 0.41 | 0.33 | 0.65 | 1.41 (1.07–1.86) | 0.014* | 0.66 | 0.74 | 1.00 |
| C50:3 triacylglycerol | 1.31 (1.09–1.57) | 0.004* | 0.41 | 0.25 | 0.70 | 1.54 (1.17–2.04) | 0.002* | 0.28 | 0.21 | 0.64 |
| C36:2 diacylglycerol | 1.28 (1.08–1.51) | 0.005* | 0.38 | 0.27 | 0.68 | 1.39 (1.08–1.78) | 0.009* | 0.54 | 0.58 | 0.69 |
| C34:3 diacylglycerol | 1.30 (1.08–1.56) | 0.005* | 0.38 | 0.25 | 0.68 | 1.45 (1.10–1.90) | 0.007* | 0.57 | 0.58 | 0.73 |
| C32:1 diacylglycerol | 1.29 (1.08–1.53) | 0.005* | 0.38 | 0.18 | 0.59 | 1.38 (1.06–1.79) | 0.016* | 0.58 | 0.66 | 1.00 |
| Valine | 1.20 (1.05–1.36) | 0.006* | 0.56 | 0.48 | 0.80 | 1.32 (1.11–1.57) | 0.002* | 0.24 | 0.21 | 0.38 |
| Isoleucine | 1.19 (1.05–1.35) | 0.006* | 0.41 | 0.48 | 0.76 | 1.22 (1.03–1.45) | 0.022* | 0.75 | 0.82 | 1.00 |
| C14:0 lysophosphatidylcholine | 1.23 (1.06–1.43) | 0.006* | 0.43 | 0.45 | 0.73 | 1.24 (1.01–1.53) | 0.038* | 0.87 | 0.92 | 1.00 |
| C50:2 triacylglycerol | 1.27 (1.07–1.51) | 0.007* | 0.63 | 0.48 | 0.82 | 1.37 (1.06–1.78) | 0.017* | 0.61 | 0.76 | 0.92 |
| C48:3 triacylglycerol | 1.27 (1.06–1.51) | 0.009* | 0.58 | 0.36 | 0.76 | 1.41 (1.08–1.85) | 0.011* | 0.39 | 0.42 | 0.89 |
| C48:2 triacylglycerol | 1.27 (1.06–1.52) | 0.009* | 0.58 | 0.43 | 0.73 | 1.36 (1.04–1.78) | 0.026* | 0.73 | 0.71 | 1.00 |
| C52:1 triacylglycerol | 1.26 (1.06–1.50) | 0.010* | 0.56 | 0.45 | 0.81 | 1.33 (1.02–1.73) | 0.032* | 0.75 | 0.76 | 1.00 |
| C32:2 diacylglycerol | 1.27 (1.06–1.52) | 0.011* | 0.63 | 0.54 | 0.84 | 1.31 (1.00–1.73) | 0.05 | 0.94 | 0.97 | 1.00 |
| C54:1 triacylglycerol | 1.24 (1.05–1.47) | 0.011* | 0.60 | 0.48 | 0.83 | 1.25 (0.98–1.60) | 0.08 | 0.92 | 0.96 | 1.00 |
| C50:4 triacylglycerol | 1.25 (1.05–1.49) | 0.013* | 0.73 | 0.66 | 0.95 | 1.49 (1.14–1.95) | 0.004* | 0.39 | 0.44 | 0.69 |
| C50:1 triacylglycerol | 1.24 (1.04–1.48) | 0.014* | 0.64 | 0.64 | 0.87 | 1.32 (1.01–1.71) | 0.038* | 0.77 | 0.88 | 1.00 |
| Trimethylamine-N-oxide | 1.17 (1.03–1.32) | 0.014* | 0.79 | 0.74 | 0.92 | 1.14 (0.96–1.36) | 0.12 | 1.00 | 1.00 | 1.00 |
| C54:2 triacylglycerol | 1.23 (1.04–1.46) | 0.017* | 0.77 | 0.69 | 0.89 | 1.30 (1.01–1.67) | 0.043* | 0.85 | 0.89 | 1.00 |
| C38:5 diacylglycerol | 1.23 (1.04–1.45) | 0.017* | 0.76 | 0.80 | 0.97 | 1.25 (0.97–1.59) | 0.08 | 0.99 | 0.99 | 1.00 |
| C46:2 triacylglycerol | 1.23 (1.03–1.47) | 0.019* | 0.79 | 0.60 | 0.87 | 1.30 (1.00–1.68) | 0.049* | 0.87 | 0.87 | 1.00 |
| C3 carnitine | 1.16 (1.02–1.32) | 0.021* | 0.77 | 0.82 | 0.97 | 1.28 (1.08–1.52) | 0.005* | 0.31 | 0.44 | 0.69 |
| C46:1 triacylglycerol | 1.23 (1.03–1.47) | 0.021* | 0.75 | 0.62 | 0.89 | 1.27 (0.98–1.64) | 0.07 | 0.94 | 0.94 | 1.00 |
| C36:2 phosphatidylethanolamine | 1.22 (1.03–1.45) | 0.021* | 0.77 | 0.72 | 0.93 | 1.39 (1.07–1.79) | 0.012* | 0.28 | 0.44 | 0.64 |
| C20:5 lysophosphatidylcholine | 0.82 (0.69–0.97) | 0.022* | 0.79 | 0.82 | 0.97 | 0.82 (0.64–1.06) | 0.13 | 1.00 | 1.00 | 1.00 |
| C38:4 phosphatidylethanolamine | 1.21 (1.02–1.44) | 0.026* | 0.79 | 0.83 | 1.00 | 1.25 (0.97–1.60) | 0.09 | 0.97 | 1.00 | 1.00 |
| C36:3 diacylglycerol | 1.21 (1.02–1.44) | 0.026* | 0.85 | 0.85 | 1.00 | 1.39 (1.08–1.78) | 0.010* | 0.59 | 0.61 | 0.83 |
| C34:2 phosphatidylethanolamine | 1.23 (1.02–1.48) | 0.028* | 0.65 | 0.78 | 0.93 | 1.47 (1.10–1.95) | 0.008* | 0.15 | 0.23 | 0.13 |
| Leucine | 1.15 (1.02–1.31) | 0.029* | 0.86 | 0.88 | 1.00 | 1.25 (1.05–1.49) | 0.012* | 0.59 | 0.69 | 0.89 |
| Kynurenic acid | 1.16 (1.01–1.32) | 0.030* | 0.79 | 0.84 | 0.97 | 1.12 (0.94–1.34) | 0.20 | 1.00 | 1.00 | 1.00 |

Continued

Table 2. Continued

| Metabolites | VTE | | | | | PE | | | | |
|---------------------------------------|------------------|----------------------|--------------------------|------|------|------------------|----------------------|--------------------------|------|------|
| | OR (95% CI) | Crude <i>P</i> Value | Corrected <i>P</i> Value | | | OR (95% CI) | Crude <i>P</i> Value | Corrected <i>P</i> Value | | |
| | | | TPFP | FWR | FDR | | | TPFP | FWR | FDR |
| C48:1 triacylglycerol | 1.21 (1.02–1.45) | 0.030* | 0.87 | 0.82 | 0.97 | 1.27 (0.98–1.64) | 0.08 | 0.96 | 0.98 | 1.00 |
| C32:0 diacylglycerol | 1.20 (1.02–1.42) | 0.030* | 0.89 | 0.89 | 1.00 | 1.23 (0.96–1.56) | 0.10 | 0.99 | 1.00 | 1.00 |
| C36:1 phosphatidylcholine plasmalogen | 0.83 (0.70–0.98) | 0.031* | 0.86 | 0.88 | 1.00 | 0.81 (0.64–1.04) | 0.10 | 0.99 | 1.00 | 1.00 |
| 1-Methyladenosine | 1.16 (1.01–1.33) | 0.031* | 0.90 | 0.88 | 1.00 | 1.05 (0.87–1.25) | 0.63 | 1.00 | 1.00 | 1.00 |
| C52:3 triacylglycerol | 1.22 (1.02–1.46) | 0.032* | 0.87 | 0.91 | 1.00 | 1.45 (1.09–1.93) | 0.010* | 0.55 | 0.56 | 0.89 |
| Glutamate | 1.15 (1.01–1.30) | 0.033* | 0.90 | 0.92 | 1.00 | 1.13 (0.94–1.34) | 0.19 | 1.00 | 1.00 | 1.00 |
| Alanine | 1.15 (1.01–1.31) | 0.033* | 0.89 | 0.91 | 1.00 | 1.13 (0.95–1.35) | 0.17 | 1.00 | 1.00 | 1.00 |
| C36:4 phosphatidylethanolamine | 1.21 (1.01–1.45) | 0.035* | 0.90 | 0.92 | 1.00 | 1.25 (0.96–1.62) | 0.10 | 0.99 | 1.00 | 1.00 |
| C4-OH carnitine | 1.16 (1.01–1.33) | 0.037* | 0.60 | 0.69 | 0.95 | 1.22 (1.01–1.47) | 0.043* | 0.74 | 0.77 | 1.00 |
| C38:2 phosphatidylcholine | 1.20 (1.01–1.43) | 0.038* | 0.91 | 0.96 | 1.00 | 1.34 (1.04–1.73) | 0.024* | 0.78 | 0.86 | 1.00 |
| Tryptophan | 0.88 (0.77–0.99) | 0.041* | 0.95 | 0.97 | 1.00 | 0.79 (0.67–0.93) | 0.004* | 0.32 | 0.36 | 0.59 |
| Hydroxyproline | 1.14 (1.00–1.29) | 0.045* | 0.95 | 0.94 | 1.00 | 1.08 (0.91–1.29) | 0.39 | 1.00 | 1.00 | 1.00 |
| C36:3 phosphatidylcholine plasmalogen | 0.84 (0.71–1.00) | 0.046* | 0.91 | 0.92 | 1.00 | 0.89 (0.69–1.14) | 0.37 | 1.00 | 1.00 | 1.00 |
| C50:5 triacylglycerol | 1.19 (1.00–1.42) | 0.048* | 0.94 | 0.95 | 1.00 | 1.37 (1.05–1.78) | 0.018* | 0.64 | 0.66 | 0.95 |
| C30:0 diacylglycerol | 1.20 (1.00–1.43) | 0.048* | 0.92 | 0.88 | 1.00 | 1.21 (0.93–1.57) | 0.16 | 1.00 | 1.00 | 1.00 |
| C52:0 triacylglycerol | 1.18 (1.00–1.40) | 0.049* | 0.91 | 0.90 | 1.00 | 1.19 (0.93–1.52) | 0.16 | 1.00 | 1.00 | 1.00 |
| C38:3 phosphatidylcholine | 1.19 (1.00–1.41) | 0.053 | 0.95 | 0.98 | 1.00 | 1.30 (1.01–1.68) | 0.043* | 0.87 | 0.85 | 1.00 |
| C22:5 CE | 0.86 (0.74–1.00) | 0.06 | 0.95 | 0.98 | 1.00 | 0.79 (0.65–0.96) | 0.016* | 0.71 | 0.66 | 1.00 |
| C52:4 triacylglycerol | 1.18 (0.99–1.41) | 0.07 | 0.95 | 0.99 | 1.00 | 1.42 (1.08–1.87) | 0.013* | 0.63 | 0.63 | 0.89 |
| C18 carnitine | 0.89 (0.78–1.01) | 0.07 | 0.99 | 0.99 | 1.00 | 0.82 (0.70–0.98) | 0.026* | 0.85 | 0.87 | 1.00 |
| C36:4 diacylglycerol | 1.15 (0.97–1.37) | 0.12 | 0.99 | 1.00 | 1.00 | 1.35 (1.04–1.75) | 0.024* | 0.77 | 0.81 | 1.00 |
| C52:6 triacylglycerol | 1.15 (0.96–1.36) | 0.12 | 0.99 | 1.00 | 1.00 | 1.31 (1.01–1.70) | 0.040* | 0.77 | 0.84 | 1.00 |
| C54:5 triacylglycerol | 1.14 (0.96–1.36) | 0.12 | 1.00 | 1.00 | 1.00 | 1.30 (1.00–1.67) | 0.047* | 0.88 | 0.96 | 1.00 |
| C36:3 phosphatidylethanolamine | 1.14 (0.96–1.36) | 0.13 | 0.99 | 1.00 | 1.00 | 1.32 (1.02–1.71) | 0.036* | 0.64 | 0.78 | 0.92 |
| C20:5 CE | 0.89 (0.75–1.05) | 0.17 | 1.00 | 1.00 | 1.00 | 0.75 (0.59–0.97) | 0.027* | 0.74 | 0.78 | 1.00 |
| Glycine | 0.93 (0.82–1.06) | 0.31 | 1.00 | 1.00 | 1.00 | 0.83 (0.69–1.00) | 0.049* | 0.96 | 0.97 | 1.00 |

The logistic regression analysis was performed for each metabolite using the basic model adjusted for age, sex, and end point. CE indicates cholesterol ester; CI, confidence interval; FDR, false discovery rate; FWR, family-wise error rate; OR, odds ratio; PE, pulmonary embolism; TPFP, tail probability of the proportion of false positives; VTE, venous thromboembolism. * $P < 0.05$.

However, after multiple correction, only 1 metabolite remained significant (C5 carnitine; OR, 1.25; 95% confidence interval, 1.10–1.41; $P_{\text{corrected}}=0.03$) (Table 2). Likewise, we identified 41 metabolites that showed nominal associations with PE ($P < 0.05$), yet none survived multiple testing correction (Table 2). Most of these PE-associated metabolites (at $P < 0.05$) were also associated with incident VTE. The total 60 nominally significant metabolites we identified were highly correlated with each other (median values for Pearson correlation, 0.42 and -0.13 ; maximum values for Pearson correlation, 0.97 and -0.53 ,

respectively) and clustered into subgroups, as presented in a heat map (Figure 1).

Enrichment Analysis

We further identified specific metabolite categories that were enriched for nominally significant associations with VTE or PE cases using a permutation test to account for correlation within metabolite classes. We found diacylglycerols to show more associations with VTE and PE than non-diacylglycerol

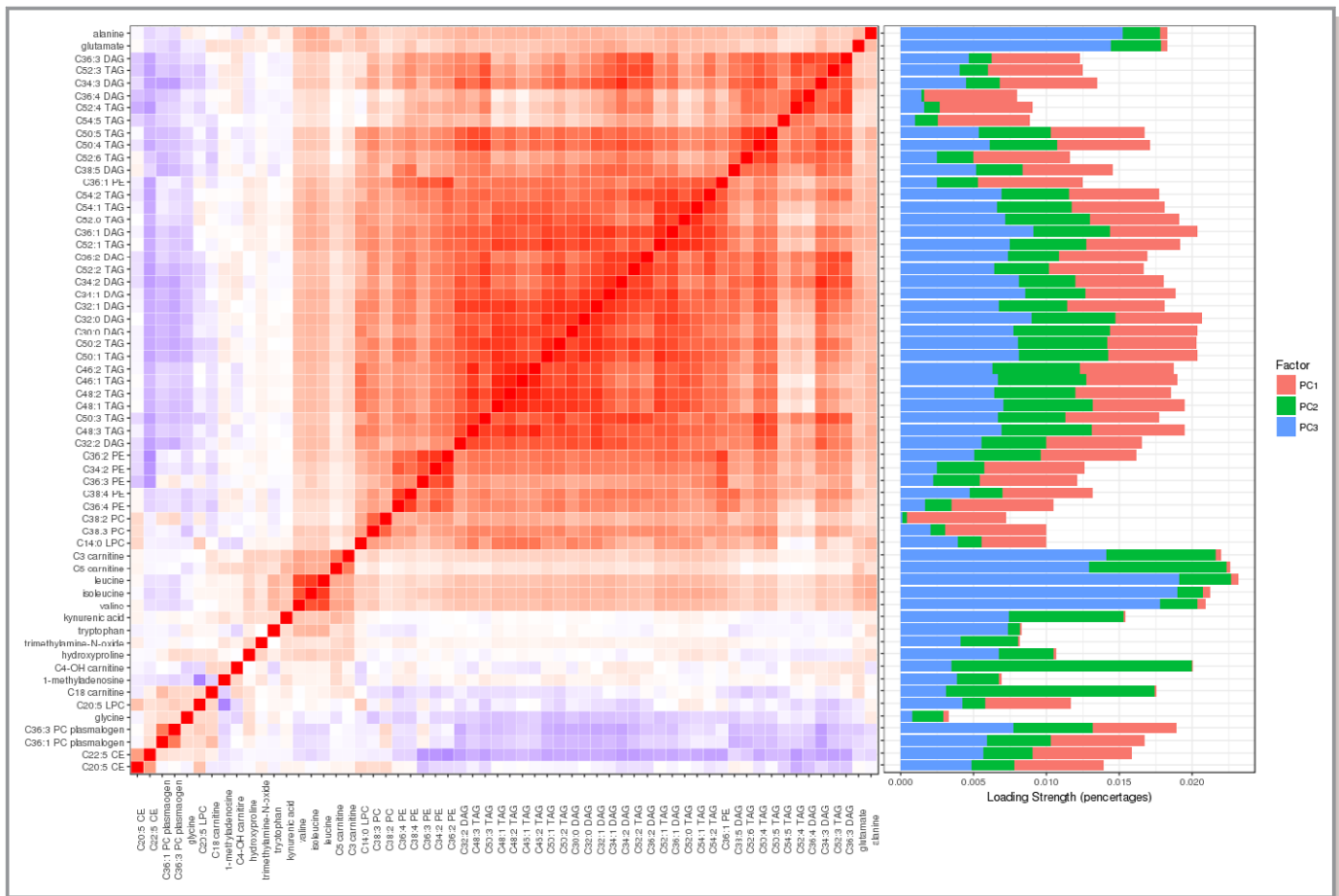


Figure 1. Left panel: correlation matrix of the 60 significant ($P < 0.05$) venous thromboembolism-, pulmonary embolism-associated metabolites identified from the basic model. We calculated Pearson correlation coefficients for each metabolite with all other metabolites and mapped these pairwise coefficients into a heat map. The color of each checker represents the directions and magnitudes of correlation. Red reflects positive correlations, blue reflects negative correlations, and white reflects the midpoint (0). A darker color represents a stronger correlation. Metabolites with similar magnitude of correlation are clustered into blocks. Right panel: stacked bar plot of loadings for each metabolite. The length of the bar represents the strength of loading (the weights calculated from principal component [PC] analysis). A longer bar means a higher proportion of that metabolite tagged by the particular PC. The percentage of variance of each of the 3 PCs is as follows: PC1, 48%; PC2, 6%; PC3, 5%. CE indicates cholesterol ester.

metabolites (enrichment $P_{\text{permutation}} = 0.004$ for VTE and $P_{\text{permutation}} = 0.03$ for PE) (Table 3).

Metabolites Associated With VTE or PE in the Full Model

Obesity, a well-established primary risk factor for VTE, has been associated with a broad range of metabolic alterations.³² We, therefore, performed additional analyses incorporating BMI and other potential confounders (smoking status, alcohol consumption, and fasting status) into the regression model (full model). We identified 3 metabolites associated with VTE and 3 metabolites associated with PE with nominal significance ($P < 0.05$). However, none survived multiple testing correction (Table 4).

We next performed analysis to explore the possibility that metabolites may mediate the risk effect of BMI in VTE or PE. Consistent with established knowledge, we observed an association between BMI and disease risk in our data, in which each unit increase in BMI was associated with a 7% increase in VTE risk (OR, 1.07; 95% confidence interval, 1.05–1.10) and a 9% increase in PE risk (OR, 1.09; 95% confidence interval, 1.06–1.12), adjusting for age, sex, and original study end point. We found almost no alterations in the effect of BMI when sequentially adjusting for the 60 nominally significant VTE- or PE-associated metabolites (Figures 2 and 3). We also performed a formal mediation analysis using the R package “mediation.” We found a significant direct effect between BMI and VTE/PE, but no significant mediation effects through metabolites (Table S3). These results indicate that the effect

Table 3. Enrichment Analysis for Significant Metabolites

| Category | No. of Significant Metabolites in the Category | No. of Total Metabolites in the Category | Disease | Permutation <i>P</i> Value |
|--------------------------|--|--|---------|----------------------------|
| Carnitine | 3 | 24 | VTE | 0.41 |
| Diacylglycerols | 11 | 12 | VTE | 0.004* |
| Lysophosphatidylcholine | 2 | 9 | VTE | 0.23 |
| Phosphatidylcholine | 3 | 36 | VTE | 0.54 |
| Phosphatidylethanolamine | 5 | 23 | VTE | 0.32 |
| Triacylglycerols | 16 | 42 | VTE | 0.14 |
| Other | 10 | 40 | VTE | 0.39 |
| Carnitine | 4 | 24 | PE | 0.35 |
| Cholesteryl esters | 2 | 12 | PE | 0.28 |
| Diacylglycerols | 8 | 12 | PE | 0.03* |
| Lysophosphatidylcholine | 1 | 9 | PE | 0.25 |
| Phosphatidylcholine | 2 | 36 | PE | 0.53 |
| Phosphatidylethanolamine | 4 | 23 | PE | 0.32 |
| Triacylglycerols | 15 | 42 | PE | 0.11 |
| Other | 5 | 40 | PE | 0.59 |

Only metabolite categories with nominal significant associations (from Table 2) were shown in the table. PE indicates pulmonary embolism; VTE, venous thromboembolism. **P*<0.05.

of BMI on VTE or PE is most likely driven through other pathways.

PC Analysis

We further reduced the dimension of analysis by only examining the first 19 PCs that explained accumulatively 80% of the variance in metabolites. Consistent with results shown from our main analysis, we found PC3, which captured a higher proportion of loading strength in C5 carnitine (Figure 1), to be significantly

associated with both VTE and PE in the basic model (*P*=0.001 and *P*=0.002, respectively). These results withstood Bonferroni correction for the numbers of PCs tested ($P_{\text{threshold}}=0.05/19=0.003$). However, none remained significant after adjusting for BMI (Table 5).

To ensure that the adjustment for potential confounding factors has been used effectively, we compared the distribution of our potential confounding factors as well as VTE/PE status across the quantiles of the exposure, PC3. As shown in Table 6, age, sex, smoking status and intensity, fasting status, BMI, comorbid illness, and VTE/PE status all differed significantly across different categories of exposure. We thus believe it is appropriate to include these potential confounding factors in our regression models.

Multiple Imputation and Sensitivity Analysis

To improve statistical power, we used multiple imputation to estimate missing values in the data set. The distribution of metabolites and covariates remained unchanged before and after imputation, indicating plausibility of imputation values (data not shown). We found 42 nominally significant metabolites (37 VTE associated and 24 PE associated, with an overlap of 19) (Table 7), most of which (41 of 42 [98%]) were also identified by the original basic model, in which only subjects with complete information on the covariates were included and analyzed.

Finally, to reduce potential confounding, we performed another sensitivity analysis removing all individuals who were cases in the original metabolomics studies from which our data set was compiled. When we limited our analysis to subjects who were controls in the original studies, we found 37 nominally significant metabolites (34 VTE associated and 11 PE associated, with an overlap of 8) (Table 8); 33 of these (89%) were also discovered in the original basic model. In both sensitivity analyses, we observed highly consistent estimates (ORs) compared with the results from the original basic model.

Table 4. Metabolites Significantly Associated With VTE or PE, Models Adjusted for BMI

| Metabolites | VTE | | | | | PE | | | | |
|--------------------------------|------------------|----------------------|--------------------------|------|------|------------------|----------------------|--------------------------|------|------|
| | OR (95% CI) | Crude <i>P</i> Value | Corrected <i>P</i> Value | | | OR (95% CI) | Crude <i>P</i> Value | Corrected <i>P</i> Value | | |
| | | | TPFP | FWR | FDR | | | TPFP | FWR | FDR |
| C14:0 lysophosphatidylcholine | 1.22 (1.04–1.44) | 0.02* | 0.83 | 0.88 | 1.00 | 1.25 (1.00–1.57) | 0.06 | 1.00 | 1.00 | 1.00 |
| C5 carnitine | 1.17 (1.02–1.34) | 0.03* | 0.83 | 0.89 | 1.00 | 1.18 (0.98–1.43) | 0.09 | 1.00 | 1.00 | 1.00 |
| Tryptophan | 0.87 (0.76–0.99) | 0.04* | 0.96 | 0.95 | 1.00 | 0.77 (0.65–0.92) | 0.003* | 0.28 | 0.30 | 0.64 |
| C50:3 triacylglycerol | 1.15 (0.94–1.40) | 0.18 | 1.00 | 1.00 | 1.00 | 1.36 (1.00–1.85) | 0.048* | 0.97 | 0.96 | 1.00 |
| C34:2 phosphatidylethanolamine | 1.14 (0.94–1.38) | 0.19 | 1.00 | 1.00 | 1.00 | 1.36 (1.01–1.82) | 0.04* | 0.68 | 0.77 | 1.00 |

The logistic regression analysis was performed for each metabolite using the full model adjusted for age, sex, end point, BMI, alcohol consumption, smoking status, and fasting status. BMI indicates body mass index; CI, confidence interval; FDR, false discovery rate; FWR, family-wise error rate; OR, odds ratio; PE, pulmonary embolism; TPFP, tail probability of the proportion of false positives; VTE, venous thromboembolism.

**P*<0.05.

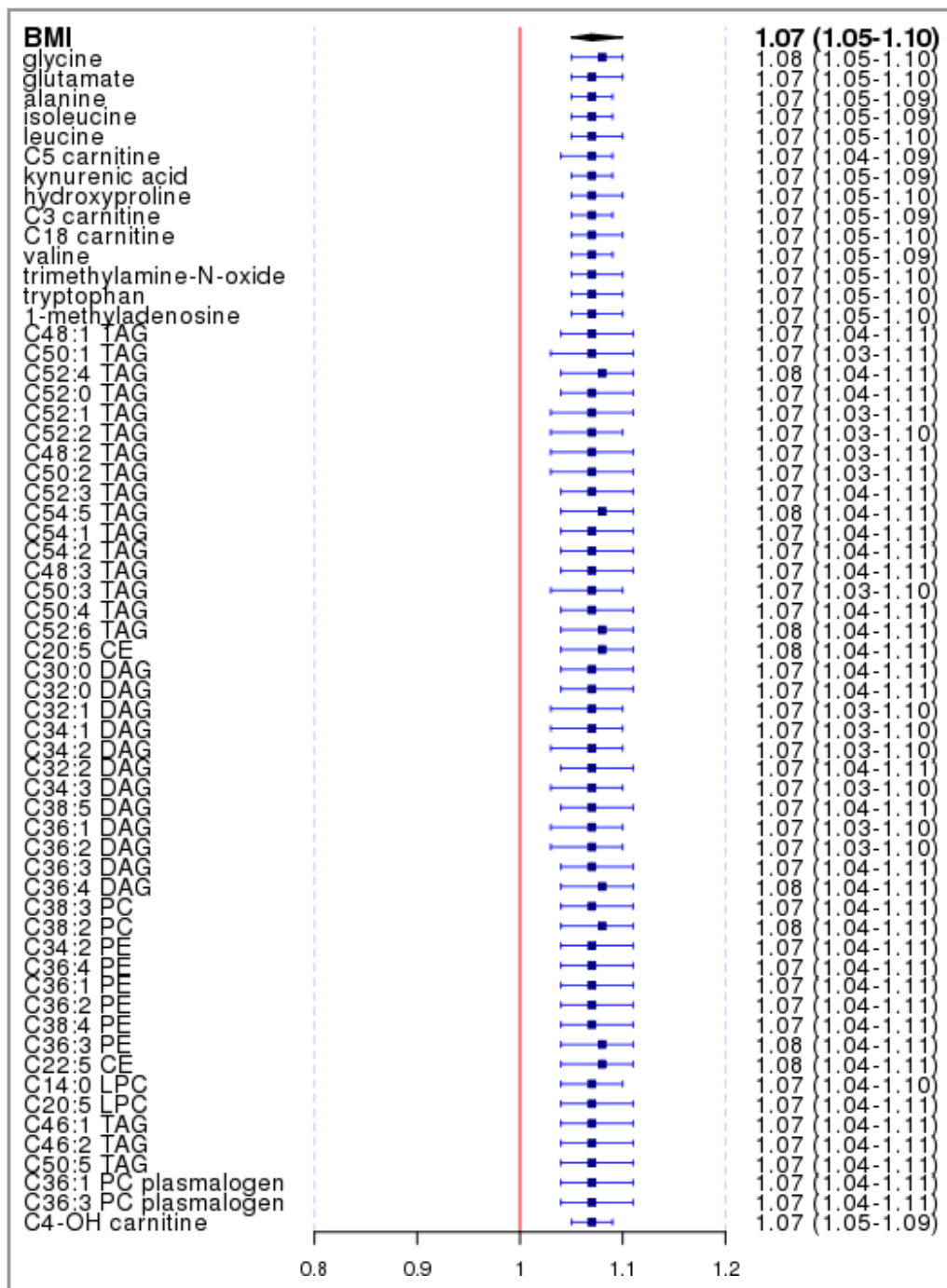


Figure 2. The effect size of body mass index (BMI) on venous thromboembolism (VTE), sequentially adjusted for each of the 60 significant VTE- or pulmonary embolism-associated metabolites. Blue squares and horizontal bars represent the odds ratios and confidence intervals of BMI adjusted for that particular metabolite. The first row (diamond) represents the “crude” effect of BMI without adjustment of any metabolite. CE indicates cholesterol ester.

Discussion

We performed a population-based case-control study to systematically investigate the relationship between circulating metabolites and risk of incident VTE. After controlling for common confounders, such as age, sex, and provoking

diseases, we found that C5 carnitine was significantly associated with incident VTE and that the metabolomics category DAGs were enriched in both VTE and PE. After further adjustment for BMI, no metabolites were significantly associated with disease accounting for multiple testing.

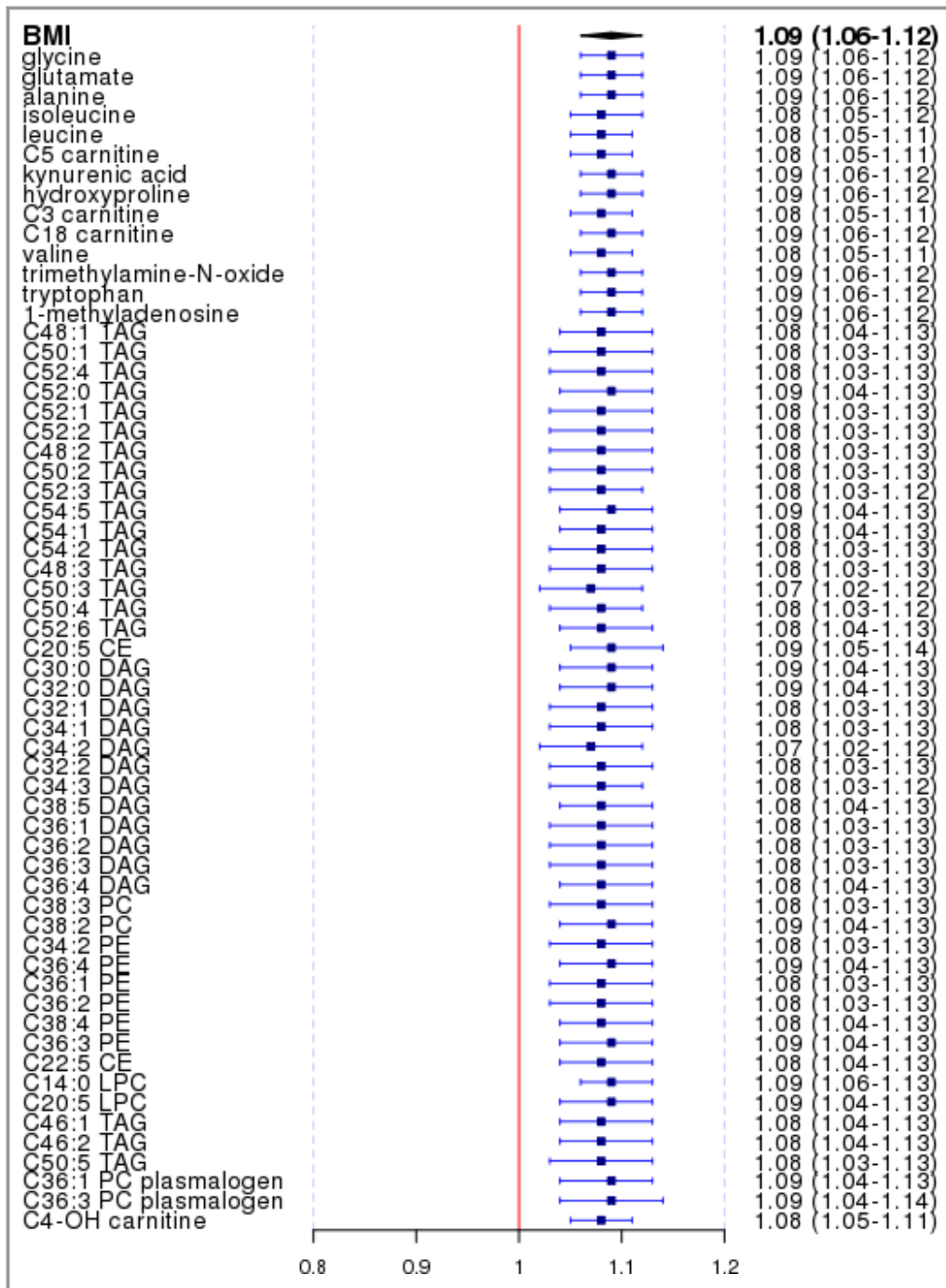


Figure 3. The effect size of body mass index (BMI) on pulmonary embolism (PE), sequentially adjusting for each of the 60 significant venous thromboembolism- or PE-associated metabolites. Blue squares and horizontal bars represent the odds ratios and confidence intervals of BMI adjusted for that particular metabolite. The first row (diamond) represents the “crude” effect of BMI without adjustment of any metabolite. CE indicates cholesterol ester.

Scant research has been performed to date using modern metabolomic techniques to identify risk factors for VTE. Only one earlier study has examined the relationship between metabolites and VTE risk.²⁰ This study was small, involving 80 subjects, used a case-control design with blood sampling after VTE diagnosis, was limited to men <55 years of age, and

lacked adjustment for confounders. A second study identified metabolites associated with high- versus low-risk PE among cases but did not examine associations with incident VTE.³³ More studies are available analyzing associations between metabolites and incident cardiovascular disease. In the Women’s Health Initiative, our colleagues identified a

Table 5. Association Between 19 PCs and the Disease

| PCs | Cumulative Proportion | VTE | | | | PE | | | |
|-------------|-----------------------|----------|-------|---------|---------|----------|-------|---------|---------|
| | | Estimate | SD | Z Score | P Value | Estimate | SD | Z Score | P Value |
| Crude model | | | | | | | | | |
| 1 | 0.48 | -0.067 | 0.137 | -0.490 | 0.624 | -0.115 | 0.182 | -0.631 | 0.528 |
| 2 | 0.54 | 0.017 | 0.065 | 0.256 | 0.798 | 0.038 | 0.090 | 0.423 | 0.672 |
| 3 | 0.59 | -0.202 | 0.063 | -3.206 | 0.001* | 0.281 | 0.090 | 3.124 | 0.002* |
| 4 | 0.62 | -0.084 | 0.064 | -1.315 | 0.189 | -0.125 | 0.091 | -1.382 | 0.167 |
| 5 | 0.65 | 0.031 | 0.067 | 0.469 | 0.639 | -0.006 | 0.097 | -0.063 | 0.950 |
| 6 | 0.67 | 0.019 | 0.064 | 0.290 | 0.772 | 0.023 | 0.094 | 0.250 | 0.803 |
| 7 | 0.69 | 0.048 | 0.069 | 0.693 | 0.488 | 0.042 | 0.099 | 0.428 | 0.669 |
| 8 | 0.71 | -0.012 | 0.065 | -0.185 | 0.853 | -0.079 | 0.092 | -0.858 | 0.391 |
| 9 | 0.72 | 0.086 | 0.065 | 1.321 | 0.186 | 0.212 | 0.091 | 2.327 | 0.020 |
| 10 | 0.73 | -0.042 | 0.062 | -0.678 | 0.497 | -0.166 | 0.089 | -1.869 | 0.062 |
| 11 | 0.74 | 0.042 | 0.065 | 0.647 | 0.518 | -0.061 | 0.091 | -0.669 | 0.503 |
| 12 | 0.75 | 0.065 | 0.064 | 1.010 | 0.312 | -0.053 | 0.091 | -0.578 | 0.563 |
| 13 | 0.76 | -0.059 | 0.065 | -0.915 | 0.360 | -0.045 | 0.091 | -0.491 | 0.623 |
| 14 | 0.77 | 0.096 | 0.064 | 1.504 | 0.133 | 0.092 | 0.091 | 1.009 | 0.313 |
| 15 | 0.77 | 0.002 | 0.062 | 0.032 | 0.975 | -0.059 | 0.088 | -0.675 | 0.500 |
| 16 | 0.78 | 0.142 | 0.063 | 2.249 | 0.025 | 0.153 | 0.089 | 1.720 | 0.085 |
| 17 | 0.79 | -0.123 | 0.066 | -1.857 | 0.063 | 0.056 | 0.088 | 0.633 | 0.526 |
| 18 | 0.79 | -0.035 | 0.063 | -0.560 | 0.576 | -0.073 | 0.088 | -0.833 | 0.405 |
| 19 | 0.80 | -0.069 | 0.065 | -1.060 | 0.289 | -0.004 | 0.090 | -0.041 | 0.967 |
| Full model | | | | | | | | | |
| 1 | 0.48 | 0.005 | 0.147 | 0.033 | 0.974 | -0.052 | 0.194 | -0.268 | 0.789 |
| 2 | 0.54 | 0.017 | 0.072 | 0.240 | 0.810 | 0.001 | 0.101 | 0.008 | 0.994 |
| 3 | 0.59 | -0.082 | 0.071 | -1.154 | 0.249 | -0.109 | 0.100 | -1.086 | 0.278 |
| 4 | 0.62 | -0.020 | 0.069 | -0.288 | 0.774 | -0.059 | 0.097 | -0.601 | 0.548 |
| 5 | 0.65 | -0.029 | 0.072 | -0.402 | 0.688 | -0.080 | 0.105 | -0.761 | 0.447 |
| 6 | 0.67 | 0.010 | 0.069 | 0.149 | 0.882 | 0.001 | 0.100 | 0.013 | 0.989 |
| 7 | 0.69 | -0.005 | 0.075 | -0.069 | 0.945 | 0.018 | 0.107 | 0.163 | 0.871 |
| 8 | 0.71 | -0.048 | 0.070 | -0.683 | 0.495 | -0.140 | 0.100 | -1.396 | 0.163 |
| 9 | 0.72 | 0.017 | 0.070 | 0.236 | 0.814 | 0.121 | 0.099 | 1.224 | 0.221 |
| 10 | 0.73 | 0.011 | 0.067 | 0.166 | 0.868 | -0.118 | 0.095 | -1.246 | 0.213 |
| 11 | 0.74 | 0.114 | 0.071 | 1.615 | 0.106 | 0.048 | 0.099 | 0.488 | 0.626 |
| 12 | 0.75 | 0.022 | 0.068 | 0.330 | 0.741 | -0.118 | 0.097 | -1.220 | 0.222 |
| 13 | 0.76 | -0.058 | 0.069 | -0.837 | 0.403 | -0.043 | 0.097 | -0.441 | 0.659 |
| 14 | 0.77 | 0.090 | 0.068 | 1.317 | 0.188 | 0.091 | 0.097 | 0.935 | 0.350 |
| 15 | 0.77 | 0.018 | 0.066 | 0.268 | 0.789 | 0.006 | 0.095 | 0.067 | 0.946 |
| 16 | 0.78 | 0.113 | 0.067 | 1.682 | 0.093 | 0.101 | 0.096 | 1.053 | 0.293 |
| 17 | 0.79 | -0.096 | 0.070 | -1.359 | 0.174 | 0.089 | 0.094 | 0.944 | 0.345 |
| 18 | 0.79 | 0.007 | 0.067 | 0.112 | 0.911 | -0.044 | 0.094 | -0.464 | 0.643 |
| 19 | 0.80 | -0.084 | 0.070 | -1.210 | 0.226 | -0.041 | 0.097 | -0.425 | 0.671 |

PC indicates principal component; PE, pulmonary embolism; VTE, venous thromboembolism.
 *P value withstood multiple corrections (0.05/19).

Table 6. Characteristics of Study Subjects by Exposure Status (Quantiles of PC3)

| Characteristics | Quantile 1 | Quantile 2 | Quantile 3 | Quantile 4 | P Values |
|-----------------------------------|-------------|-------------|-------------|-------------|---------------------------|
| Age, mean±SD, y | 53.7 (9.77) | 52.5 (9.59) | 52.1 (9.64) | 53.0 (10.2) | 0.015* |
| Sex | | | | | |
| Women | 1455 (80.7) | 1514 (84.2) | 1541 (85.6) | 1404 (77.9) | 0.08 |
| Men | 347 (19.3) | 285 (15.8) | 259 (14.4) | 398 (22.1) | |
| Smoking status | | | | | |
| Current | 182 (10.2) | 190 (10.6) | 170 (9.5) | 130 (7.2) | 0.0002* |
| Past | 701 (39.2) | 586 (32.7) | 625 (34.9) | 604 (33.7) | |
| Never | 907 (50.7) | 1018 (56.7) | 998 (55.7) | 1060 (59.1) | |
| Pack-years, mean±SD | 21.2 (17.0) | 20.2 (17.4) | 19.0 (17.1) | 17.6 (17.1) | 9.9 × 10 ^{-9*} |
| Alcohol consumption, mean±SD, g/d | 10.1 (12.6) | 9.30 (11.9) | 9.18 (12.2) | 9.44 (11.1) | 0.77 |
| Fasting status at blood draw | | | | | |
| Yes | 1225 (68) | 1367 (76) | 1388 (77.1) | 1349 (74.9) | 2.77 × 10 ^{-6*} |
| No | 576 (32) | 432 (24) | 412 (22.9) | 453 (25.1) | |
| BMI, kg/m ² | | | | | |
| <25 | 569 (32.1) | 818 (46.2) | 1000 (56.3) | 1212 (68.9) | 1.92 × 10 ^{-55*} |
| 25–29 | 669 (37.8) | 623 (35.2) | 555 (31.2) | 451 (25.6) | |
| ≥30 | 534 (30.1) | 328 (18.5) | 222 (12.5) | 96 (5.5) | |
| Mean±SD | 28.1 (5.75) | 26.1 (4.94) | 25.1 (4.63) | 23.9 (3.63) | 9.2 × 10 ^{-152*} |
| With an end point [†] | | | | | |
| Yes | 896 (52.9) | 862 (49.7) | 783 (45.6) | 751 (45.7) | 0.044* |
| No | 797 (47.1) | 872 (50.3) | 934 (54.4) | 894 (54.3) | |
| Incident VTE | | | | | |
| No | 1725 (95.7) | 1730 (96.2) | 1753 (97.4) | 1755 (97.4) | 0.0010* |
| Yes | 77 (4.3) | 69 (3.8) | 47 (2.6) | 47 (2.6) | |
| Incident PE | | | | | |
| No | 1725 (97.7) | 1730 (97.7) | 1753 (98.5) | 1755 (98.9) | 0.0012* |
| Yes | 40 (2.3) | 40 (2.3) | 26 (1.5) | 19 (1.1) | |

Data are given as number (percentage) unless otherwise indicated. BMI indicates body mass index; PC, principal component; PE, pulmonary embolism; VTE, venous thromboembolism. * $P < 0.05$.

[†]With an endpoint was a case in one of the previous case-control studies of other phenotypes (amyotrophic lateral sclerosis, rheumatoid arthritis, Parkinson disease, diabetes mellitus, chronic stress, and ovarian, pancreatic, and prostate cancer) from which our study population was derived.

signature of 8 metabolites associated with coronary heart disease, independent of traditional risk factors.²⁸ Women with C34:2 hydroxy-phosphatidylcholine levels in the highest quartile had a 4.7-fold increase in coronary heart disease. Glutamate was also associated with coronary heart disease outcomes in this study. In addition, Lanfear et al developed a plasma metabolite profile defined by 13 amino acids, which predicts the survival of patients with heart failure incremental to conventional predictors³⁴; Olkowicz et al found 22 amino acids significantly changed in aortic stenosis³⁵; and Bjerrum et al identified a reversible proatherogenic lipid profile in patients with active inflammatory bowel disease.³⁶ These results suggest the importance of metabolite profile in the field of cardiovascular disease, but also highlight the novelty

of our results, which may specifically inform future research into metabolic changes and pathophysiological characteristics of VTE. More important, by demonstrating nominal associations in prospectively collected data, our findings provide preliminary evidence in support of an altered physiological state before the VTE event.

Our results are in line with previously reported findings. The only other available high-throughput metabolomics study, conducted by Deguchi et al, identified decreased plasma levels of 2 long-chain carnitines (>10 carbon atoms, C12:2 and C18:2) associated with VTE.²⁰ A follow-up experimental study demonstrated the addition of acyl carnitines to clotting assays inhibited factor X α -initiated clotting, with a carbon length of >14 exhibiting the highest activity. We have

Table 7. Metabolites Associated With VTE or PE Using Multiple Imputed Data

| Metabolite | VTE | | | | PE | | | |
|---------------------------------------|------------------|---------|------------------|---------|------------------|---------|------------------|---------|
| | Basic Model | | Full Model | | Basic Model | | Full Model | |
| | OR (95% CI) | P Value | OR (95% CI) | P Value | OR (95% CI) | P Value | OR (95% CI) | P Value |
| C5 carnitine | 1.25 (1.10–1.41) | 0.001* | 1.17 (1.02–1.33) | 0.02* | 1.24 (1.04–1.47) | 0.015* | 1.13 (0.95–1.36) | 0.18 |
| C36:1 phosphatidylethanolamine | 1.31 (1.09–1.57) | 0.004* | 1.26 (1.05–1.53) | 0.01* | 1.39 (1.07–1.82) | 0.015* | 1.35 (1.03–1.78) | 0.03* |
| Valine | 1.20 (1.05–1.36) | 0.006* | 1.07 (0.93–1.22) | 0.36 | 1.32 (1.11–1.57) | 0.002* | 1.14 (0.94–1.37) | 0.17 |
| Isoleucine | 1.19 (1.05–1.35) | 0.006* | 1.08 (0.94–1.24) | 0.26 | 1.22 (1.03–1.45) | 0.022* | 1.07 (0.89–1.29) | 0.47 |
| C36:1 diacylglycerol | 1.25 (1.06–1.48) | 0.008* | 1.15 (0.97–1.37) | 0.11 | 1.29 (1.01–1.64) | 0.039* | 1.17 (0.91–1.50) | 0.23 |
| C34:2 diacylglycerol | 1.26 (1.06–1.50) | 0.009* | 1.15 (0.96–1.37) | 0.14 | 1.39 (1.07–1.79) | 0.013* | 1.23 (0.95–1.61) | 0.12 |
| C36:2 diacylglycerol | 1.25 (1.05–1.48) | 0.011* | 1.14 (0.96–1.36) | 0.14 | 1.34 (1.05–1.72) | 0.021* | 1.20 (0.93–1.56) | 0.16 |
| C34:1 diacylglycerol | 1.24 (1.05–1.47) | 0.012* | 1.12 (0.94–1.34) | 0.19 | 1.28 (1.00–1.64) | 0.048* | 1.14 (0.88–1.47) | 0.33 |
| Trimethylamine-N-oxide | 1.17 (1.03–1.32) | 0.014* | 1.15 (1.01–1.30) | 0.03* | 1.14 (0.96–1.36) | 0.122 | 1.12 (0.94–1.34) | 0.19 |
| C52:2 triacylglycerol | 1.26 (1.05–1.51) | 0.015* | 1.14 (0.94–1.38) | 0.18 | 1.31 (1.00–1.72) | 0.054 | 1.16 (0.88–1.53) | 0.30 |
| C36:2 phosphatidylethanolamine | 1.24 (1.04–1.48) | 0.016* | 1.20 (1.00–1.43) | 0.05 | 1.37 (1.06–1.77) | 0.016* | 1.32 (1.01–1.71) | 0.04* |
| C14:0 lysophosphatidylcholine | 1.20 (1.03–1.38) | 0.016* | 1.16 (0.99–1.34) | 0.06 | 1.20 (0.99–1.47) | 0.070 | 1.16 (0.95–1.43) | 0.15 |
| C50:3 triacylglycerol | 1.25 (1.04–1.50) | 0.016* | 1.14 (0.95–1.37) | 0.17 | 1.41 (1.07–1.85) | 0.015* | 1.25 (0.95–1.66) | 0.12 |
| C54:1 triacylglycerol | 1.22 (1.03–1.45) | 0.020* | 1.15 (0.96–1.38) | 0.12 | 1.22 (0.96–1.55) | 0.111 | 1.14 (0.89–1.47) | 0.31 |
| C3 carnitine | 1.16 (1.02–1.32) | 0.021* | 1.10 (0.96–1.25) | 0.17 | 1.28 (1.08–1.52) | 0.005* | 1.20 (1.00–1.44) | 0.05* |
| C34:3 diacylglycerol | 1.23 (1.03–1.48) | 0.022* | 1.13 (0.94–1.36) | 0.19 | 1.32 (1.01–1.73) | 0.041* | 1.19 (0.90–1.56) | 0.22 |
| C50:2 triacylglycerol | 1.22 (1.03–1.46) | 0.024* | 1.11 (0.92–1.33) | 0.28 | 1.30 (1.00–1.68) | 0.048* | 1.15 (0.88–1.50) | 0.32 |
| C32:1 diacylglycerol | 1.22 (1.02–1.46) | 0.026* | 1.12 (0.93–1.34) | 0.24 | 1.28 (0.99–1.65) | 0.063 | 1.15 (0.88–1.50) | 0.30 |
| C52:1 triacylglycerol | 1.22 (1.02–1.46) | 0.028* | 1.12 (0.93–1.35) | 0.23 | 1.26 (0.98–1.63) | 0.076 | 1.14 (0.88–1.49) | 0.32 |
| Leucine | 1.15 (1.02–1.31) | 0.029* | 1.05 (0.92–1.20) | 0.50 | 1.25 (1.05–1.49) | 0.012* | 1.10 (0.92–1.33) | 0.29 |
| Kynurenic acid | 1.16 (1.01–1.32) | 0.030* | 1.12 (0.98–1.27) | 0.11 | 1.12 (0.94–1.34) | 0.200 | 1.07 (0.89–1.28) | 0.47 |
| 1-Methyladenosine | 1.16 (1.01–1.33) | 0.031* | 1.05 (0.91–1.21) | 0.49 | 1.05 (0.87–1.25) | 0.629 | 0.90 (0.75–1.09) | 0.29 |
| C36:1 phosphatidylcholine plasmalogen | 0.83 (0.71–0.98) | 0.031* | 0.87 (0.74–1.03) | 0.11 | 0.82 (0.65–1.03) | 0.093 | 0.87 (0.69–1.11) | 0.26 |
| Glutamate | 1.15 (1.01–1.30) | 0.033* | 1.03 (0.90–1.18) | 0.66 | 1.13 (0.94–1.34) | 0.185 | 0.98 (0.81–1.19) | 0.88 |
| Alanine | 1.15 (1.01–1.31) | 0.034* | 1.07 (0.93–1.22) | 0.36 | 1.13 (0.95–1.35) | 0.172 | 1.02 (0.85–1.23) | 0.84 |
| C38:2 phosphatidylcholine | 1.19 (1.01–1.40) | 0.036* | 1.16 (0.99–1.37) | 0.06 | 1.29 (1.03–1.62) | 0.030* | 1.25 (1.00–1.57) | 0.05* |
| C48:2 triacylglycerol | 1.21 (1.01–1.45) | 0.036* | 1.12 (0.93–1.35) | 0.23 | 1.27 (0.97–1.66) | 0.080 | 1.16 (0.88–1.53) | 0.29 |
| C50:4 triacylglycerol | 1.21 (1.01–1.44) | 0.037* | 1.12 (0.93–1.34) | 0.23 | 1.37 (1.06–1.79) | 0.018* | 1.25 (0.95–1.63) | 0.11 |
| C4-OH carnitine | 1.16 (1.01–1.33) | 0.037* | 1.07 (0.93–1.24) | 0.32 | 1.22 (1.01–1.47) | 0.043* | 1.10 (0.90–1.35) | 0.34 |
| C54:2 triacylglycerol | 1.20 (1.01–1.42) | 0.039* | 1.10 (0.92–1.32) | 0.28 | 1.24 (0.97–1.59) | 0.084 | 1.13 (0.87–1.46) | 0.37 |
| C48:3 triacylglycerol | 1.20 (1.01–1.44) | 0.041* | 1.12 (0.93–1.35) | 0.22 | 1.30 (1.00–1.69) | 0.050 | 1.20 (0.91–1.57) | 0.20 |
| Tryptophan | 0.88 (0.77–0.99) | 0.041* | 0.89 (0.78–1.01) | 0.06 | 0.79 (0.67–0.93) | 0.004* | 0.80 (0.68–0.94) | 0.01* |
| C34:2 phosphatidylethanolamine | 1.21 (1.01–1.47) | 0.042* | 1.18 (0.98–1.42) | 0.09 | 1.36 (1.03–1.79) | 0.029* | 1.32 (1.00–1.74) | 0.05 |
| C14:0 sphingomyelin | 1.17 (1.00–1.37) | 0.044* | 1.13 (0.97–1.32) | 0.12 | 1.22 (0.98–1.51) | 0.077 | 1.17 (0.94–1.46) | 0.17 |
| Hydroxyproline | 1.14 (1.00–1.29) | 0.045* | 1.12 (0.99–1.28) | 0.08 | 1.08 (0.91–1.29) | 0.391 | 1.06 (0.89–1.27) | 0.50 |
| C50:1 triacylglycerol | 1.19 (1.00–1.42) | 0.046* | 1.09 (0.91–1.30) | 0.37 | 1.22 (0.95–1.58) | 0.120 | 1.10 (0.84–1.43) | 0.50 |

Continued

Table 7. Continued

| Metabolite | VTE | | | | PE | | | |
|---------------------------------------|------------------|---------|------------------|---------|------------------|---------|------------------|---------|
| | Basic Model | | Full Model | | Basic Model | | Full Model | |
| | OR (95% CI) | P Value | OR (95% CI) | P Value | OR (95% CI) | P Value | OR (95% CI) | P Value |
| C36:3 phosphatidylcholine plasmalogen | 0.84 (0.70–1.00) | 0.048* | 0.91 (0.76–1.09) | 0.32 | 0.86 (0.67–1.10) | 0.226 | 0.96 (0.75–1.24) | 0.77 |
| C36:3 diacylglycerol | 1.19 (1.00–1.41) | 0.051 | 1.11 (0.93–1.32) | 0.26 | 1.32 (1.02–1.72) | 0.034* | 1.21 (0.92–1.58) | 0.17 |
| C36:3 phosphatidylethanolamine | 1.18 (1.00–1.40) | 0.056 | 1.18 (0.99–1.40) | 0.07 | 1.30 (1.02–1.65) | 0.034* | 1.30 (1.01–1.66) | 0.04* |
| C22:5 CE | 0.87 (0.75–1.01) | 0.061 | 0.91 (0.78–1.06) | 0.23 | 0.80 (0.66–0.98) | 0.027* | 0.85 (0.69–1.04) | 0.11 |
| C18 carnitine | 0.89 (0.78–1.01) | 0.072 | 0.90 (0.80–1.03) | 0.13 | 0.82 (0.70–0.98) | 0.026* | 0.85 (0.71–1.01) | 0.06 |
| Glycine | 0.93 (0.82–1.06) | 0.308 | 1.02 (0.89–1.17) | 0.78 | 0.83 (0.69–1.00) | 0.049* | 0.93 (0.77–1.12) | 0.43 |

The basic model adjusted for age, sex, and end point; the full model adjusted for age, sex, end point, body mass index, alcohol consumption, smoking status, and fasting status. The data were imputed 5 rounds, and estimates were pooled over each round. CE indicates cholesterol ester; CI, confidence interval; OR, odds ratio; PE, pulmonary embolism; VTE, venous thromboembolism.

* $P < 0.05$.

complemented previous findings by observing the significant effect of a short-chain carnitine (C5) with VTE, suggesting that both long- and short-chain carnitines may work in concert in the complex dynamics of the hemostatic system and the pathophysiological characteristics of VTE. Further research is needed to elucidate the biological mechanisms underlying these observations.

The role of diacylglycerols in VTE pathological characteristics among humans is unknown. To the best of our knowledge, no relevant epidemiological evidence is available relating diacylglycerols and risk of incident VTE. In animals, a diacylglycerol-enriched diet has been found to inhibit arterial thrombus formation in *ApoE* and low-density lipoprotein receptor double-negative male mice,³⁷ possibly driven by the protection of vascular endothelium from injury and lowered serum low-density lipoprotein cholesterol.³⁸ Diacylglycerols constitute only a minor component of the human diet, and the amount of intake is small. It is likely that the enrichment of diacylglycerols we identified for VTE and PE reflects the intermediate effect of triacylglycerols. Well-designed epidemiological investigations are warranted to clarify this association in humans.

In addition to carnitine and diacylglycerols, we observed a few suggestive associations that are of interest, and in concordance with existing literature. Although none of these findings survived multiple correction, they may represent promising avenues for new research. We found several triacylglycerols, phosphatidylethanolamines, and amino acids (tryptophan) to be nominally linked with the risk of VTE and PE. In a study consisting of 477 postmenopausal female VTE cases and 1986 sex, age-matched controls, elevated circulating triacylglycerols (>1.05 mmol/L) have been reported to double the risk of VTE independent of age, hospitalization, malignancy,

height, and weight.³⁹ This association has been supported by several small case-control studies^{40–42} and is possibly driven by a decreased ratio of activated protein C⁴³ and elevated concentrations of coagulant factors.^{44–46} Moreover, antibodies against phosphatidylethanolamines have been detected in patients with thromboembolic events, such as PE, stroke, and antiphospholipid syndrome (a systemic thrombotic diathesis).^{47,48} In a population-based retrospective case-control study examining 234 patients with VTE and 236 matched controls, antiphosphatidylethanolamine antibodies have been found to increase thrombosis risk by 4-fold using univariate logistic regression, although the effect was attenuated after adjusting for additional covariates.⁴⁹ In addition to the lipid-related metabolites, tryptophan is an amino acid acquired from the diet with anticoagulant properties. In mice, a tryptophan derivative, TD-26, has been found to decrease death from acute PE by 90% and attenuate thrombosis weight by 60%. Downstream experiments have further demonstrated that TD-26 suppresses platelet aggregation by blocking the binding of fibrinogen to integrin and reducing protein kinase B phosphorylation in platelet phosphatidylinositol 3-kinase signaling.⁵⁰

Although our results are in line with previous findings, most of the nominally significant metabolites identified by our systemic interrogation failed to pass multiple correction or are attenuated to nonsignificant after adjusting for BMI, making it difficult to draw firm conclusions. There are several potential explanations for this. Although, to the best of our knowledge, this is the largest study of metabolites and incident VTE risk to date, our power is still limited. With the current sample size of 240 cases and >6900 controls, we had 80% power to detect an OR of 1.2 per SD increase in metabolite at a corrected α level of 0.0003, a stringent P-threshold. Second, BMI is known to influence a broad range of metabolites. Given

Table 8. Metabolites Significantly Associated With VTE or PE in the Sensitive Analysis, Excluding Individuals Recruited as a Case in Case-Control Selections of Other Outcomes

| Metabolite | VTE | | | | PE | | | |
|--------------------------------|------------------|---------|------------------|---------|------------------|---------|------------------|---------|
| | Basic Model | | Full Model | | Basic Model | | Full Model | |
| | OR (95% CI) | P Value | OR (95% CI) | P Value | OR (95% CI) | P Value | OR (95% CI) | P Value |
| C34:3 diacylglycerol | 1.53 (1.16–2.02) | 0.003* | 1.36 (0.99–1.86) | 0.06 | 1.67 (1.04–2.69) | 0.034* | 1.56 (0.92–2.64) | 0.10 |
| C34:2 diacylglycerol | 1.47 (1.14–1.91) | 0.003* | 1.30 (0.97–1.76) | 0.08 | 1.52 (0.98–2.36) | 0.061 | 1.46 (0.88–2.41) | 0.14 |
| C50:3 triacylglycerol | 1.48 (1.12–1.96) | 0.005* | 1.33 (0.97–1.82) | 0.08 | 1.59 (0.99–2.56) | 0.057 | 1.57 (0.91–2.69) | 0.11 |
| C50:4 triacylglycerol | 1.47 (1.12–1.92) | 0.005* | 1.27 (0.94–1.71) | 0.12 | 1.76 (1.10–2.82) | 0.019* | 1.63 (0.97–2.75) | 0.07 |
| C34:1 diacylglycerol | 1.42 (1.10–1.83) | 0.006* | 1.25 (0.93–1.67) | 0.14 | 1.26 (0.82–1.94) | 0.290 | 1.18 (0.72–1.93) | 0.51 |
| C52:4 triacylglycerol | 1.48 (1.11–1.95) | 0.007* | 1.23 (0.90–1.68) | 0.20 | 1.77 (1.08–2.89) | 0.023* | 1.56 (0.91–2.66) | 0.11 |
| C36:3 diacylglycerol | 1.42 (1.10–1.84) | 0.007* | 1.22 (0.91–1.63) | 0.18 | 1.56 (1.02–2.39) | 0.042* | 1.44 (0.88–2.34) | 0.14 |
| C38:5 diacylglycerol | 1.42 (1.10–1.83) | 0.007* | 1.29 (0.96–1.72) | 0.09 | 1.46 (0.96–2.23) | 0.076 | 1.32 (0.82–2.10) | 0.25 |
| C36:4 diacylglycerol | 1.43 (1.10–1.86) | 0.008* | 1.22 (0.91–1.65) | 0.18 | 1.78 (1.14–2.78) | 0.011* | 1.60 (0.98–2.60) | 0.06 |
| C54:5 triacylglycerol | 1.43 (1.10–1.86) | 0.008* | 1.31 (0.98–1.75) | 0.07 | 1.84 (1.17–2.89) | 0.008* | 1.65 (1.02–2.68) | 0.04* |
| C48:3 triacylglycerol | 1.42 (1.09–1.87) | 0.011* | 1.28 (0.95–1.72) | 0.11 | 1.56 (0.98–2.49) | 0.061 | 1.49 (0.90–2.47) | 0.12 |
| C36:1 diacylglycerol | 1.39 (1.08–1.79) | 0.011* | 1.23 (0.91–1.65) | 0.18 | 1.24 (0.80–1.90) | 0.338 | 1.15 (0.70–1.88) | 0.59 |
| Ornithine | 1.30 (1.06–1.60) | 0.012* | 1.25 (1.00–1.56) | 0.05* | 1.30 (0.95–1.78) | 0.099 | 1.31 (0.94–1.85) | 0.12 |
| C32:1 diacylglycerol | 1.40 (1.07–1.82) | 0.013* | 1.28 (0.95–1.74) | 0.11 | 1.35 (0.87–2.11) | 0.181 | 1.33 (0.81–2.19) | 0.26 |
| C50:5 triacylglycerol | 1.39 (1.07–1.81) | 0.015* | 1.20 (0.89–1.61) | 0.22 | 1.67 (1.05–2.65) | 0.029* | 1.49 (0.91–2.46) | 0.12 |
| C52:2 triacylglycerol | 1.41 (1.07–1.85) | 0.015* | 1.20 (0.88–1.64) | 0.25 | 1.19 (0.76–1.87) | 0.451 | 1.09 (0.66–1.79) | 0.75 |
| C5 carnitine | 1.27 (1.04–1.53) | 0.016* | 1.16 (0.94–1.43) | 0.17 | 1.36 (1.01–1.81) | 0.041* | 1.34 (0.98–1.83) | 0.07 |
| C32:2 diacylglycerol | 1.40 (1.06–1.84) | 0.017* | 1.29 (0.95–1.75) | 0.10 | 1.28 (0.81–2.03) | 0.284 | 1.22 (0.74–2.01) | 0.43 |
| C36:2 diacylglycerol | 1.36 (1.06–1.76) | 0.018* | 1.16 (0.87–1.56) | 0.31 | 1.29 (0.84–2.00) | 0.250 | 1.18 (0.72–1.94) | 0.50 |
| C54:1 triacylglycerol | 1.37 (1.05–1.77) | 0.018* | 1.21 (0.90–1.63) | 0.21 | 1.28 (0.83–1.98) | 0.265 | 1.17 (0.72–1.91) | 0.53 |
| C52:3 triacylglycerol | 1.39 (1.05–1.84) | 0.020* | 1.14 (0.84–1.55) | 0.41 | 1.46 (0.90–2.37) | 0.125 | 1.29 (0.76–2.20) | 0.35 |
| C50:6 triacylglycerol | 1.36 (1.04–1.77) | 0.024* | 1.19 (0.89–1.59) | 0.23 | 1.51 (0.97–2.35) | 0.070 | 1.33 (0.83–2.14) | 0.23 |
| C48:2 triacylglycerol | 1.36 (1.04–1.78) | 0.025* | 1.26 (0.93–1.70) | 0.14 | 1.31 (0.84–2.06) | 0.234 | 1.31 (0.80–2.14) | 0.29 |
| C36:2 phosphatidylethanolamine | 1.36 (1.04–1.77) | 0.025* | 1.19 (0.88–1.60) | 0.26 | 1.49 (0.95–2.32) | 0.082 | 1.40 (0.85–2.29) | 0.18 |
| C46:2 triacylglycerol | 1.35 (1.04–1.76) | 0.027* | 1.24 (0.92–1.66) | 0.15 | 1.34 (0.86–2.09) | 0.197 | 1.28 (0.79–2.06) | 0.31 |
| C32:0 diacylglycerol | 1.32 (1.03–1.68) | 0.028* | 1.19 (0.89–1.58) | 0.23 | 1.21 (0.80–1.84) | 0.366 | 1.16 (0.72–1.86) | 0.55 |
| C34:2 phosphatidylethanolamine | 1.38 (1.03–1.84) | 0.028* | 1.26 (0.92–1.71) | 0.15 | 1.58 (0.97–2.58) | 0.066 | 1.52 (0.89–2.57) | 0.12 |
| Isoleucine | 1.25 (1.02–1.52) | 0.031* | 1.08 (0.87–1.35) | 0.47 | 1.31 (0.97–1.77) | 0.074 | 1.20 (0.86–1.67) | 0.29 |
| C50:2 triacylglycerol | 1.32 (1.02–1.72) | 0.035* | 1.18 (0.87–1.59) | 0.29 | 1.23 (0.79–1.91) | 0.351 | 1.18 (0.72–1.93) | 0.52 |
| C52:1 triacylglycerol | 1.33 (1.02–1.73) | 0.035* | 1.18 (0.87–1.60) | 0.30 | 1.18 (0.76–1.83) | 0.462 | 1.11 (0.67–1.82) | 0.69 |
| C52:0 triacylglycerol | 1.31 (1.02–1.68) | 0.037* | 1.22 (0.91–1.63) | 0.19 | 1.22 (0.80–1.87) | 0.352 | 1.17 (0.72–1.91) | 0.52 |
| C38:4 phosphatidylethanolamine | 1.31 (1.01–1.69) | 0.039* | 1.21 (0.92–1.61) | 0.18 | 1.33 (0.87–2.03) | 0.184 | 1.31 (0.83–2.08) | 0.25 |
| C54:2 triacylglycerol | 1.31 (1.01–1.70) | 0.042* | 1.09 (0.81–1.47) | 0.57 | 1.23 (0.79–1.93) | 0.358 | 1.08 (0.66–1.78) | 0.76 |
| C36:1 phosphatidylethanolamine | 1.33 (1.01–1.75) | 0.044* | 1.15 (0.85–1.56) | 0.35 | 1.32 (0.84–2.07) | 0.226 | 1.20 (0.73–1.95) | 0.47 |
| C34:0 phosphatidylinositol | 0.83 (0.65–1.07) | 0.150 | 0.82 (0.63–1.07) | 0.15 | 0.68 (0.47–0.99) | 0.046* | 0.72 (0.48–1.07) | 0.11 |
| C3 carnitine | 1.12 (0.91–1.37) | 0.277 | 1.02 (0.82–1.27) | 0.85 | 1.41 (1.04–1.91) | 0.026* | 1.34 (0.97–1.85) | 0.07 |
| C36:0 phosphatidylethanolamine | 0.91 (0.70–1.19) | 0.499 | 0.86 (0.65–1.14) | 0.28 | 0.64 (0.41–0.99) | 0.046* | 0.59 (0.36–0.94) | 0.03* |

The basic model adjusted for age and sex; the full model adjusted for age, sex, body mass index, alcohol consumption, smoking status, and fasting status. CI indicates confidence interval; OR, odds ratio; PE, pulmonary embolism; VTE, venous thromboembolism.

* $P < 0.05$.

the complex dynamics between BMI, metabolomic profile, and VTE, the modest effect of a single metabolite is likely to be masked by the strong main effect of BMI with our current sample size. This is also partly reflected by the lack of mediation effect (ie, the effect of BMI did not change when adjusting for individual metabolites). Therefore, larger studies are needed to confirm our preliminary findings. Larger studies would also help mitigate confounding by BMI through, for example, well-powered subgroup analysis stratified by BMI.

Our study has several strengths. The combination of phenotype, robust longitudinal data on environmental exposures, and prospectively collected plasma samples with well-measured metabolites minimizes the possibility of reverse causality. The wealth of detailed lifestyle covariates assembled by NHS and HPFS also enables us to adjust for multiple key confounders, such as age, sex, BMI, smoking, and alcohol use, which earlier studies did not have the opportunity to adjust for. Our results support the important role BMI plays both as a crucial predictor for VTE and an essential confounder in studies of lipids, and they stress the need to control metabolomic analyses for BMI. We also performed 2 sensitivity analyses, and findings from the complete case analysis are robust in both sensitivity analyses, lending support to the veracity of our results.

Nevertheless, there are several limitations to be acknowledged in addition to the power issue. It is possible that misclassification could be introduced by self-reported VTE status. However, all our subjects are health professionals and phenotype status is collected within a short period after disease onset, so we believe that the likelihood of bias is low.¹³ In addition, the need for prolonged, frequently monitored anticoagulant treatment reduces the likelihood that a health professional would misreport a VTE diagnosis. It is also possible that metabolites measured several years before a VTE event may not represent the pathophysiologic state just preceding the VTE. However, metabolites may reflect processes that increase the risk of disease over time. Moreover, in previous studies, we have assessed the ICCs to assess the stability of 233 known metabolites among 594 control subjects with 2 blood samples collected 10 years apart. The median ICC over 10 years was 0.37 (10th–90th percentile, 0.08–0.55), and 44% of metabolites had ICCs ≥ 0.4 . The branched chain amino acids (leucine, isoleucine, and valine) each had ICCs=0.5. This level of reproducibility is similar to that found for other biological variables, such as blood pressure (ICC=0.6),⁵¹ blood glucose (ICC=0.52),⁵² and serum cholesterol (ICC=0.65),⁵³ all exposures considered to be reasonably well-measured and consistent predictors of disease in epidemiologic studies. It is also possible that issues related to preprocessing of blood can attenuate metabolite estimates. We have studied this in our previous work. Our group quantified 257 metabolites from archived cohort plasma samples and found that 75% were reproducible over

delays in preprocessing (Spearman correlation or ICC ≥ 0.75 , for immediate versus 24-hour delayed processing). We also compared metabolites across fasting time, season of blood collection, and time of day of blood collection and found that multivariable-adjusted geometric mean metabolite peak areas were within 15% across these parameters for nearly all metabolites (160/166 [96%]).²⁹ Last, metabolites included in the current study are performed on the Broad Institute platform, which is weighted toward lipids. We acknowledge that lipids may not be as strongly associated with VTE as other metabolite classes, and the lack of standardization across metabolomics platforms makes the generalizability of results challenging. Future work should focus on other metabolite classes.

Conclusions

Circulating metabolites, including a short-chain carnitine and possibly diacylglycerols, appear to be associated with risk of incident VTE, although the associations we observed may be confounded by BMI. Larger studies involving additional individuals and with broader metabolomics coverage are needed to confirm our findings, enhance the generalizability of results, and deepen our understanding of VTE risk.

Sources of Funding

This work was funded by the National Heart, Lung, and Blood Institute–National Institutes of Health (NIH; R01 HL116854). The Nurses' Health Study, Nurses' Health Study II, and Health Professionals Follow-Up Study are supported by the National Cancer Institute (NCI)–NIH (UM1 CA186107, R01 CA49449, R01 HL034594, UM1 CA176726, R01 CA67262, UM1 CA167552, and R01 HL35464). The Health Professionals Follow-Up Study receives pivotal support for its infrastructure through an NCI-funded grant (U01 CA167552). The grant covers a range of research activities from the questionnaires, to disease follow-up, tissue collection, death follow-up, dietary databases, and sharing of the resource across consortia. Jiang is funded by an International Postdoctoral Grant from the Swedish Research Council (Vetenskapsrådet). Hagan has received an NIH grant (T32 HL098048).

Disclosures

None.

References

1. Beckman MG, Hooper WC, Critchley SE, Ortel TL. Venous thromboembolism: a public health concern. *Am J Prev Med.* 2010;38:S495–S501.
2. Heit JA, Spencer FA, White RH. The epidemiology of venous thromboembolism. *J Thromb Thrombolysis.* 2016;41:3–14.

3. Heit JA, Armasu SM, Asmann YW, Cunningham JM, Matsumoto ME, Petterson TM, De Andrade M. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. *J Thromb Haemost.* 2012;10:1521–1531.
4. Tang W, Teichert M, Chasman DI, Heit JA, Morange P-E, Li G, Pankratz N, Leebeek FW, Paré G, de Andrade M, Tzourio C, Psaty BM, Basu S, Rutter R, Rose L, Armasu SM, Lumley T, Heckbert SR, Uitterlinden AG, Lathrop M, Rice KM, Cushman M, Hofman A, Lambert J-C, Glazer NL, Pankow JS, Wittman JC, Amouyel P, Bis JC, Bovill EG, Kong X, Tracy RP, Boerwinkle E, Rotter JJ, Trégouët D-A, Loth DW, Stricker BHC, Ridker PM, Folsom AR, Smith NL. A genome-wide association study for venous thromboembolism: the extended cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. *Genet Epidemiol.* 2013;37:512–521.
5. Germain M, Chasman DI, de Haan H, Tang W, Lindström S, Weng L-C, de Andrade M, de Visser MCH, Wiggins KL, Suchon P, Saut N, Smadja DM, Le Gal G, van Hylckama Vlieg A, Di Narzo A, Hao K, Nelson CP, Rocanin-Arjo A, Folkersen L, Monajemi R, Rose LM, Brody JA, Slagboom E, Aïssi D, Gagnon F, Delouze J-F, Deloukas P, Tzourio C, Dartigues J-F, Berr C, Taylor KD, Civelek M, Eriksson P, Cardiogenics Consortium, Psaty BM, Houwing-Duitermaat J, Goodall AH, Cambien F, Kraft P, Amouyel P, Samani NJ, Basu S, Ridker PM, Rosendaal FR, Kabrhel C, Folsom AR, Heit J, Reitsma PH, Trégouët D-A, Smith NL, Morange P-E. Meta-analysis of 65,734 individuals identifies TSPAN15 and SLC44A2 as two susceptibility loci for venous thromboembolism. *Am J Hum Genet.* 2015;96:532–542.
6. Hinds DA, Buil A, Ziemek D, Martinez-Perez A, Malik R, Folkersen L, Germain M, Mälarstig A, Brown A, Soria JM, Dichgans M, Bing N, Franco-Cereceda A, Souto JC, Dermitzakis ET, Hamsten A, Worrall BB, Tung JY; METASTROKE Consortium, INVENT Consortium, Sabater-Lleal M. Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis. *Hum Mol Genet.* 2016;25:1867–1874.
7. Klarin D, Emdin CA, Natarajan P, Conrad MF; INVENT Consortium, Kathiresan S. Genetic analysis of venous thromboembolism in UK biobank identifies the ZFPM2 locus and implicates obesity as a causal risk factor. *Circ Cardiovasc Genet.* 2017;10:e001643.
8. Kabrhel C, Varraso R, Goldhaber SZ, Rimm EB, Camargo CA. Prospective study of BMI and the risk of pulmonary embolism in women. *Obesity (Silver Spring).* 2009;17:2040–2046.
9. Lindström S, Germain M, Crous-Bou M, Smith EN, Morange P-E, van Hylckama Vlieg A, de Haan HG, Chasman D, Ridker P, Brody J, de Andrade M, Heit JA, Tang W, DeVivo I, Grodstein F, Smith NL, Tregouët D, Kabrhel C; INVENT Consortium. Assessing the causal relationship between obesity and venous thromboembolism through a Mendelian randomization study. *Hum Genet.* 2017;136:897–902.
10. Blokhin IO, Lentz SR. Mechanisms of thrombosis in obesity. *Curr Opin Hematol.* 2013;20:437–444.
11. Schäfer K, Konstantinides S. Mechanisms linking leptin to arterial and venous thrombosis: potential pharmacological targets. *Curr Pharm Des.* 2014;20:635–640.
12. Morgan ES, Wilson E, Melody T, Parmar K, Zhang Y, Gao F, Hunt BJ. An observational study of haemostatic changes, leptin and soluble endoglin during pregnancy in women with different BMIs. *Blood Coagul Fibrinolysis.* 2017;28:50–55.
13. Pun VC, Hart JE, Kabrhel C, Camargo CA, Baccarelli AA, Laden F. Prospective study of ambient particulate matter exposure and risk of pulmonary embolism in the Nurses' Health Study cohort. *Environ Health Perspect.* 2015;123:1265–1270.
14. Sumner JA, Kubzansky LD, Kabrhel C, Roberts AL, Chen Q, Winning A, Gilsanz P, Rimm EB, Glymour MM, Koenen KC. Associations of trauma exposure and posttraumatic stress symptoms with venous thromboembolism over 22 years in women. *J Am Heart Assoc.* 2016;5:e003197. DOI: 10.1161/JAHA.116.003197.
15. Ganna A, Salihovic S, Sundström J, Broeckling CD, Hedman AK, Magnusson PKE, Pedersen NL, Larsson A, Siegbahn A, Zilmer M, Prenni J, Arnlöv J, Lind L, Fall T, Ingelsson E. Large-scale metabolomic profiling identifies novel biomarkers for incident coronary heart disease. *PLoS Genet.* 2014;10:e1004801.
16. Tang WHW, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y, Hazen SL. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med.* 2013;368:1575–1584.
17. Cheng S, Shah SH, Corwin EJ, Fiehn O, Fitzgerald RL, Gerszten RE, Illig T, Rhee EP, Srinivas PR, Wang TJ, Jain M; American Heart Association Council on Functional Genomics and Translational Biology; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; and Stroke Council. Potential impact and study considerations of metabolomics in cardiovascular health and disease: a scientific statement from the American Heart Association. *Circ Cardiovasc Genet.* 2017;10:e000032.
18. Shah SH, Sun J-L, Stevens RD, Bain JR, Muehlbauer MJ, Pieper KS, Haynes C, Hauser ER, Kraus WE, Granger CB, Newgard CB, Califf RM, Newby LK. Baseline metabolomic profiles predict cardiovascular events in patients at risk for coronary artery disease. *Am Heart J.* 2012;163:844–850.e1.
19. Würtz P, Havulinna AS, Soininen P, Tynkkynen T, Prieto-Merino D, Tillin T, Ghorbani A, Artati A, Wang Q, Tiainen M, Kangas AJ, Kettunen J, Kaikkonen J, Mikkilä V, Jula A, Kähönen M, Lehtimäki T, Lawlor DA, Gaunt TR, Hughes AD, Sattar N, Illig T, Adamski J, Wang TJ, Perola M, Ripatti S, Vasari S, Raitakari OT, Gerszten RE, Casas J-P, Chaturvedi N, Ala-Korpela M, Salomaa V. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. *Circulation.* 2015;131:774–785.
20. Deguchi H, Banerjee Y, Trauger S, Siuzdak G, Kalisiak E, Fernández JA, Hoang L, Tran M, Yegneswaran S, Elias DJ, Griffin JH. Acylcarnitines are anticoagulants that inhibit factor Xa and are reduced in venous thrombosis, based on metabolomics data. *Blood.* 2015;126:1595–1600.
21. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health.* 1997;6:49–62.
22. Bao Y, Bertoia ML, Lenart EB, Stampfer MJ, Willett WC, Speizer FE, Chavarro JE. Origin, methods, and evolution of the three Nurses' Health Studies. *Am J Public Health.* 2016;106:1573–1581.
23. Varraso R, Kabrhel C, Goldhaber SZ, Rimm EB, Camargo CA. Prospective study of diet and venous thromboembolism in US women and men. *Am J Epidemiol.* 2012;175:114–126.
24. Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, Barbieri RL, Speizer FE. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. *J Natl Cancer Inst.* 1998;90:1292–1299.
25. Tworoger SS, Sluss P, Hankinson SE. Association between plasma prolactin concentrations and risk of breast cancer among predominately premenopausal women. *Cancer Res.* 2006;66:2476–2482.
26. Wei EK, Giovannucci E, Fuchs CS, Willett WC, Mantzoros CS. Low plasma adiponectin levels and risk of colorectal cancer in men: a prospective study. *J Natl Cancer Inst.* 2005;97:1688–1694.
27. Townsend MK, Bao Y, Poole EM, Bertrand KA, Kraft P, Wolpin BM, Clish CB, Tworoger SS. Impact of pre-analytic blood sample collection factors on metabolomics. *Cancer Epidemiol Biomark Prev.* 2016;25:823–829.
28. Paynter NP, Balasubramanian R, Giulianini F, Wang DD, Tinker LF, Gopal S, Deik AA, Bullock K, Pierce KA, Scott J, Martínez-González MA, Estruch R, Manson JE, Cook NR, Albert CM, Clish CB, Rexrode KM. Metabolic predictors of incident coronary heart disease in women. *Circulation.* 2018;137:841–853.
29. Townsend MK, Clish CB, Kraft P, Wu C, Souza AL, Deik AA, Tworoger SS, Wolpin BM. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. *Clin Chem.* 2013;59:1657–1667.
30. O'Sullivan JF, Morningstar JE, Yang Q, Zheng B, Gao Y, Jeanfavre S, Scott J, Fernandez C, Zheng H, O'Connor S, Cohen P, Vasari S, Long MT, Wilson JG, Melder O, Wang TJ, Fox C, Peterson RT, Clish CB, Corey KE, Gerszten RE. Dimethylguanidino valeric acid is a marker of liver fat and predicts diabetes. *J Clin Invest.* 2017;127:4394–4402.
31. Liao SG, Lin Y, Kang DD, Chandra D, Bon J, Kaminski N, Sciruba FC, Tseng GC. Missing value imputation in high-dimensional phenomic data: imputable or not, and how? *BMC Bioinformatics.* 2014;15:346.
32. Geidenstam N, Al-Majdoub M, Ekman M, Spégel P, Ridderstråle M. Metabolite profiling of obese individuals before and after a one year weight loss program. *Int J Obes (Lond).* 2017;41:1369–1378.
33. Zeleznik OA, Poole EM, Lindstrom S, Kraft P, Van Hylckama Vlieg A, Lasky-Su JA, Harrington LB, Hagan K, Kim J, Parry BA, Giordano N, Kabrhel C. Metabolomic analysis of 92 pulmonary embolism patients from a nested case-control study identifies metabolites associated with adverse clinical outcomes. *J Thromb Haemost.* 2018;16:500–507.
34. Lanfear DE, Gibbs JJ, Li J, She R, Petucci C, Culver JA, Tang WHW, Pinto YM, Williams LK, Sabbah HN, Gardell SJ. Targeted metabolomic profiling of plasma and survival in heart failure patients. *JACC Heart Fail.* 2017;5:823–832.
35. Olkowitz M, Debski J, Jablonska P, Dadlez M, Smolenski RT. Application of a new procedure for liquid chromatography/mass spectrometry profiling of plasma amino acid-related metabolites and untargeted shotgun proteomics to identify mechanisms and biomarkers of calcific aortic stenosis. *J Chromatogr A.* 2017;1517:66–78.
36. Bjerrum JT, Steenholdt C, Ainsworth M, Nielsen OH, Reed MA, Atkins K, Günther UL, Hao F, Wang Y. Metabonomics uncovers a reversible proatherogenic lipid profile during infliximab therapy of inflammatory bowel disease. *BMC Med.* 2017;15:184.
37. Ijiri Y, Naemura A, Yamashita T, Meguro S, Watanabe H, Tokimitsu I, Yamamoto J. Dietary diacylglycerol attenuates arterial thrombosis in apoE and LDLR deficient mice. *Thromb Res.* 2006;117:411–417.
38. Ijiri Y, Naemura A, Yamashita T, Ikarugi H, Meguro S, Tokimitsu I, Yamamoto J. Mechanism of the antithrombotic effect of dietary diacylglycerol in atherosclerotic mice. *Pathophysiol Haemost Thromb.* 2006;35:380–387.

39. Doggen CJM, Smith NL, Lemaitre RN, Heckbert SR, Rosendaal FR, Psaty BM. Serum lipid levels and the risk of venous thrombosis. *Arterioscler Thromb Vasc Biol.* 2004;24:1970–1975.
40. McColl MD, Sattar N, Ellison J, Tait RC, Walker ID, Packard CJ, Greer IA. Lipoprotein (a), cholesterol and triglycerides in women with venous thromboembolism. *Blood Coagul Fibrinolysis.* 2000;11:225–229.
41. Vayá A, Mira Y, Ferrando F, Contreras M, Estelles A, España F, Corella D, Aznar J. Hyperlipidaemia and venous thromboembolism in patients lacking thrombophilic risk factors. *Br J Haematol.* 2002;118:255–259.
42. Kawasaki T, Kambayashi J, Ariyoshi H, Sakon M, Suehisa E, Monden M. Hypercholesterolemia as a risk factor for deep-vein thrombosis. *Thromb Res.* 1997;88:67–73.
43. Lowe GD, Rumley A, Woodward M, Reid E, Rumley J. Activated protein C resistance and the FV:R506Q mutation in a random population sample: associations with cardiovascular risk factors and coagulation variables. *Thromb Haemost.* 1999;81:918–924.
44. de Visser MC, Rosendaal FR, Bertina RM. A reduced sensitivity for activated protein C in the absence of factor V Leiden increases the risk of venous thrombosis. *Blood.* 1999;93:1271–1276.
45. Chadarevian R, Bruckert E, Dejager S, Presberg P, Turpin G. Relationship between triglycerides and factor VIIIc and plasminogen activator inhibitor type-1: lack of threshold value. *Thromb Res.* 1999;96:175–182.
46. Tsai AW, Cushman M, Rosamond WD, Heckbert SR, Tracy RP, Aleksic N, Folsom AR. Coagulation factors, inflammation markers, and venous thromboembolism: the longitudinal investigation of thromboembolism etiology (LITE). *Am J Med.* 2002;113:636–642.
47. Sanmarco M, Alessi MC, Harle JR, Sapin C, Aillaud MF, Gentile S, Juhan-Vague I, Weiller PJ. Antibodies to phosphatidylethanolamine as the only antiphospholipid antibodies found in patients with unexplained thromboses. *Thromb Haemost.* 2001;85:800–805.
48. Staub HL, Bertolaccini ML, Khamashta MA. Anti-phosphatidylethanolamine antibody, thromboembolic events and the antiphospholipid syndrome. *Autoimmun Rev.* 2012;12:230–234.
49. Sanmarco M, Gayet S, Alessi M-C, Audrain M, de Maistre E, Gris J-C, de Groot PG, Hachulla E, Harlé J-R, Sié P, Boffa M-C. Antiphosphatidylethanolamine antibodies are associated with an increased odds ratio for thrombosis: a multicenter study with the participation of the European Forum on antiphospholipid antibodies. *Thromb Haemost.* 2007;97:949–954.
50. Chen Y, Wang Y, Xie Z, Ming X, Li Z, Kong Y. A tryptophan derivative TD-26 attenuates thrombus formation by inhibiting both PI3K/Akt signaling and binding of fibrinogen to integrin α IIb β 3. *Biochem Biophys Res Commun.* 2015;465:516–522.
51. Rosner B, Hennekens CH, Kass EH, Miall WE. Age-specific correlation analysis of longitudinal blood pressure data. *Am J Epidemiol.* 1977;106:306–313.
52. Gordon T, Shurtleff D. The Framingham Study: an epidemiologic investigation of cardiovascular disease. Section 29: Means at each examination and inter-examination variation of specified characteristics: Framingham Study Exam 1 to Exam 10. 1973. DHEW publication (NIH) 74-478.
53. Shekelle RB, Shryock AM, Paul O, Lepper M, Stamler J, Liu S, Raynor WJ. Diet, serum cholesterol, and death from coronary heart disease: the Western Electric study. *N Engl J Med.* 1981;304:65–70.

Supplemental Material

Table S1. The 211 named metabolites and their category.

| Category | Metabolites |
|---------------------------------------|---|
| Carnitine | C2 carnitine, C16 carnitine, C10 carnitine, C5 carnitine, C6 carnitine, C8 carnitine, C3 carnitine, C18 carnitine, C4 carnitine, C14:1 carnitine, C12 carnitine, C5:1 carnitine, C18:1 carnitine, C14 carnitine, C26 carnitine, C18:2 carnitine, C4-OH carnitine, C9 carnitine, C10:2 carnitine, C12:1 carnitine, C14:2 carnitine, carnitine |
| Cholesteryl ester (CE) | C18:2 CE, C16:1 CE, C16:0 CE, C18:1 CE, C14:0 CE, C20:4 CE, C20:5 CE, C22:6 CE, C20:3 CE, C18:0 CE, C18:3 CE, C22:5 CE |
| Diacylglycerol (DAG) | C30:0 DAG, C32:0 DAG, C32:1 DAG, C34:1 DAG, C34:2 DAG, C32:2 DAG, C34:3 DAG, C38:5 DAG, C36:1 DAG, C36:2 DAG, C36:3 DAG, C36:4 DAG |
| Lysophosphatidylcholine (LPC) | C18:1 LPC, C14:0 LPC, C16:0 LPC, C16:1 LPC, C18:0 LPC, C18:2 LPC, C20:4 LPC, C20:5 LPC, C22:6 LPC |
| Lysophosphatidylethanolamine (LPE) | C18:0 LPE, C16:0 LPE, C18:1 LPE, C18:2 LPE, C20:4 LPE, C22:6 LPE |
| Phosphatidylcholine (PC) | C30:0 PC, C30:1 PC, C32:0 PC, C32:1 PC, C32:2 PC, C34:4 PC, C34:0 PC, C34:1 PC, C34:2 PC, C36:4 PC-A, C38:6 PC, C34:3 PC, C36:0 PC, C36:1 PC, C36:2 PC, C38:3 PC, C38:4 PC, C40:6 PC, C36:3 PC, C36:4 PC-B, C38:2 PC, C40:10 PC, C40:9 PC, C34:1 PC plasmalogen-A, C34:2 PC plasmalogen, C34:3 PC plasmalogen, C34:5 PC plasmalogen, C36:5 PC plasmalogen-B, C36:5 PC plasmalogen-A, C38:7 PC plasmalogen, C36:1 PC plasmalogen, C36:2 PC plasmalogen, C36:3 PC plasmalogen, C38:4 PC plasmalogen, C40:7 PC plasmalogen, C36:4 PC plasmalogen, C34:0 PE, C34:2 PE, C36:4 PE, C38:2 PE, C34:2 PE plasmalogen, C36:0 PE, C36:1 PE, C36:2 PE, C38:4 PE, C40:6 PE, C36:3 PE, C38:5 PE, C36:2 PE plasmalogen, C38:6 PE, C34:3 PE plasmalogen, C38:3 PE plasmalogen, C38:5 PE plasmalogen, C38:6 PE plasmalogen, C40:7 PE plasmalogen, C36:5 PE plasmalogen, C38:7 PE plasmalogen, C36:3 PE plasmalogen, C36:4 PE plasmalogen |
| Lysophosphatidylethanolamine (PE) | C18:0 SM, C16:0 SM, C24:0 SM, C14:0 SM, C20:0 SM, C22:0 SM, C22:1 SM |
| Sphingomyeline (SM) | C48:0 TAG, C50:0 TAG, C48:1 TAG, C50:1 TAG, C52:4 TAG, C52:0 TAG, C52:1 TAG, C52:2 TAG, C54:4 TAG, C48:2 TAG, C50:2 TAG, C52:3 TAG, C54:5 TAG, C56:8 TAG, C54:1 TAG, C54:2 TAG, C56:2 TAG, C54:3 TAG, C56:5 TAG, C56:3 TAG, C58:8 TAG, C48:3 TAG, C50:3 TAG, C50:4 TAG, C52:6 TAG, C56:9 TAG, C56:6 TAG, C56:7 TAG, C58:9 TAG, C58:7 TAG, C60:12 TAG, C46:0 TAG, C46:1 TAG, C46:2 TAG, C50:5 TAG, C50:6 TAG, C54:9 TAG, C56:10 TAG, C52:7 TAG, C54:8 TAG, C58:11 TAG, C44:0 TAG |
| Triacylglycerol (TAG) | N-carbamoyl-beta-alanine |
| β -alanine derivative | creatine |
| Nitrogenous organic acid | dimethylglycine |
| Amino acid glycine derivative | glycine, glutamate, tyrosine, phenylalanine, alanine, threonine, asparagine, isoleucine, histidine, lysine, serine, ornithine, arginine, glutamine, leucine, methionine, kynurenic acid, pipecolic acid, hydroxyproline, valine, citrulline, tryptophan |
| Amino acid | taurine |
| organic compound, constituent of bile | allantoin |
| Diureide of glyoxylic acid | 1-methylnicotinamide |
| Others | Metabolite of nicotinamide |
| Carboxylic acid of piperidine | pipecolic acid |
| TMAO | trimethylamine-N-oxide |
| Alkaloid found in tobacco | cotinine |
| Vitamin | niacinamide |
| Modified nucleosides | 1-methyladenosine |
| Ceramide | C16:0 Ceramide (d18:1), C22:0 Ceramide (d18:1), C24:0 Ceramide (d18:1) |
| Phosphatidylserine | C40:6 PS, C34:0 PS, C34:0 PI |
| N-monomethyl-L-arginine | NMMA |

Table S2. HMDB IDs and metabolite names.

| HMDB | METABOLITE |
|-----------|--------------------------|
| HMDB00026 | N-carbamoyl-beta-alanine |
| HMDB00062 | carnitine |
| HMDB00064 | creatine |
| HMDB00092 | dimethylglycine |
| HMDB00123 | glycine |
| HMDB00148 | glutamate |
| HMDB00158 | tyrosine |
| HMDB00159 | phenylalanine |
| HMDB00161 | alanine |
| HMDB00167 | threonine |
| HMDB00168 | asparagine |
| HMDB00172 | isoleucine |
| HMDB00177 | histidine |
| HMDB00182 | lysine |
| HMDB00187 | serine |
| HMDB00201 | C2 carnitine |
| HMDB00214 | ornithine |
| HMDB00222 | C16 carnitine |
| HMDB00251 | taurine |
| HMDB00252 | sphingosine |
| HMDB00462 | allantoin |
| HMDB00517 | arginine |
| HMDB00610 | C18:2 CE |
| HMDB00641 | glutamine |
| HMDB00651 | C10 carnitine |
| HMDB00658 | C16:1 CE |
| HMDB00687 | leucine |
| HMDB00688 | C5 carnitine |
| HMDB00696 | methionine |
| HMDB00699 | 1-methylnicotinamide |
| HMDB00705 | C6 carnitine |
| HMDB00715 | kynurenic acid |
| HMDB00716 | pipecolic acid |
| HMDB00725 | hydroxyproline |
| HMDB00791 | C8 carnitine |
| HMDB00824 | C3 carnitine |
| HMDB00848 | C18 carnitine |
| HMDB00883 | valine |
| HMDB00885 | C16:0 CE |
| HMDB00904 | citrulline |
| HMDB00918 | C18:1 CE |
| HMDB00925 | trimethylamine-N-oxide |
| HMDB00929 | tryptophan |
| HMDB01046 | cotinine |
| HMDB01348 | C18:0 SM |
| HMDB01406 | niacinamide |
| HMDB02013 | C4 carnitine |
| HMDB02014 | C14:1 carnitine |
| HMDB02250 | C12 carnitine |
| HMDB02366 | C5:1 carnitine |
| HMDB02815 | C18:1 LPC |
| HMDB03331 | 1-methyladenosine |
| HMDB04949 | C16:0 Ceramide (d18:1) |
| HMDB04952 | C22:0 Ceramide (d18:1) |
| HMDB04956 | C24:0 Ceramide (d18:1) |
| HMDB05065 | C18:1 carnitine |

| | |
|-----------|-----------------|
| HMDB05066 | C14 carnitine |
| HMDB05356 | C48:0 TAG |
| HMDB05357 | C50:0 TAG |
| HMDB05359 | C48:1 TAG |
| HMDB05360 | C50:1 TAG |
| HMDB05363 | C52:4 TAG |
| HMDB05365 | C52:0 TAG |
| HMDB05367 | C52:1 TAG |
| HMDB05369 | C52:2 TAG |
| HMDB05370 | C54:4 TAG |
| HMDB05376 | C48:2 TAG |
| HMDB05377 | C50:2 TAG |
| HMDB05384 | C52:3 TAG |
| HMDB05385 | C54:5 TAG |
| HMDB05392 | C56:8 TAG |
| HMDB05395 | C54:1 TAG |
| HMDB05403 | C54:2 TAG |
| HMDB05404 | C56:2 TAG |
| HMDB05405 | C54:3 TAG |
| HMDB05406 | C56:5 TAG |
| HMDB05410 | C56:3 TAG |
| HMDB05413 | C58:8 TAG |
| HMDB05432 | C48:3 TAG |
| HMDB05433 | C50:3 TAG |
| HMDB05435 | C50:4 TAG |
| HMDB05436 | C52:6 TAG |
| HMDB05448 | C56:9 TAG |
| HMDB05456 | C56:6 TAG |
| HMDB05462 | C56:7 TAG |
| HMDB05463 | C58:9 TAG |
| HMDB05471 | C58:7 TAG |
| HMDB05478 | C60:12 TAG |
| HMDB06347 | C26 carnitine |
| HMDB06469 | C18:2 carnitine |
| HMDB06725 | C14:0 CE |
| HMDB06726 | C20:4 CE |
| HMDB06731 | C20:5 CE |
| HMDB06733 | C22:6 CE |
| HMDB06736 | C20:3 CE |
| HMDB07011 | C30:0 DAG |
| HMDB07098 | C32:0 DAG |
| HMDB07099 | C32:1 DAG |
| HMDB07102 | C34:1 DAG |
| HMDB07103 | C34:2 DAG |
| HMDB07128 | C32:2 DAG |
| HMDB07132 | C34:3 DAG |
| HMDB07199 | C38:5 DAG |
| HMDB07216 | C36:1 DAG |
| HMDB07218 | C36:2 DAG |
| HMDB07219 | C36:3 DAG |
| HMDB07248 | C36:4 DAG |
| HMDB07869 | C30:0 PC |
| HMDB07870 | C30:1 PC |
| HMDB07871 | C32:0 PC |
| HMDB07873 | C32:1 PC |
| HMDB07874 | C32:2 PC |
| HMDB07883 | C34:4 PC |
| HMDB07970 | C34:0 PC |

| | |
|-----------|------------------------|
| HMDB07972 | C34:1 PC |
| HMDB07973 | C34:2 PC |
| HMDB07983 | C36:4 PC-A |
| HMDB07991 | C38:6 PC |
| HMDB08006 | C34:3 PC |
| HMDB08036 | C36:0 PC |
| HMDB08038 | C36:1 PC |
| HMDB08039 | C36:2 PC |
| HMDB08047 | C38:3 PC |
| HMDB08048 | C38:4 PC |
| HMDB08057 | C40:6 PC |
| HMDB08105 | C36:3 PC |
| HMDB08138 | C36:4 PC-B |
| HMDB08270 | C38:2 PC |
| HMDB08511 | C40:10 PC |
| HMDB08731 | C40:9 PC |
| HMDB08925 | C34:0 PE |
| HMDB08928 | C34:2 PE |
| HMDB08937 | C36:4 PE |
| HMDB08942 | C38:2 PE |
| HMDB08952 | C34:2 PE plasmalogen |
| HMDB08991 | C36:0 PE |
| HMDB08993 | C36:1 PE |
| HMDB08994 | C36:2 PE |
| HMDB09003 | C38:4 PE |
| HMDB09012 | C40:6 PE |
| HMDB09060 | C36:3 PE |
| HMDB09069 | C38:5 PE |
| HMDB09082 | C36:2 PE plasmalogen |
| HMDB09102 | C38:6 PE |
| HMDB09805 | C34:0 PI |
| HMDB10167 | C40:6 PS |
| HMDB10169 | C16:0 SM |
| HMDB10368 | C18:0 CE |
| HMDB10370 | C18:3 CE |
| HMDB10375 | C22:5 CE |
| HMDB10379 | C14:0 LPC |
| HMDB10382 | C16:0 LPC |
| HMDB10383 | C16:1 LPC |
| HMDB10384 | C18:0 LPC |
| HMDB10386 | C18:2 LPC |
| HMDB10395 | C20:4 LPC |
| HMDB10397 | C20:5 LPC |
| HMDB10404 | C22:6 LPC |
| HMDB10411 | C46:0 TAG |
| HMDB10412 | C46:1 TAG |
| HMDB10419 | C46:2 TAG |
| HMDB10471 | C50:5 TAG |
| HMDB10497 | C50:6 TAG |
| HMDB10498 | C54:9 TAG |
| HMDB10513 | C56:10 TAG |
| HMDB10517 | C52:7 TAG |
| HMDB10518 | C54:8 TAG |
| HMDB10531 | C58:11 TAG |
| HMDB11130 | C18:0 LPE |
| HMDB11208 | C34:1 PC plasmalogen-A |
| HMDB11210 | C34:2 PC plasmalogen |
| HMDB11211 | C34:3 PC plasmalogen |

| | |
|-----------|------------------------|
| HMDB11214 | C34:5 PC plasmalogen |
| HMDB11220 | C36:5 PC plasmalogen-B |
| HMDB11221 | C36:5 PC plasmalogen-A |
| HMDB11229 | C38:7 PC plasmalogen |
| HMDB11241 | C36:1 PC plasmalogen |
| HMDB11243 | C36:2 PC plasmalogen |
| HMDB11244 | C36:3 PC plasmalogen |
| HMDB11252 | C38:4 PC plasmalogen |
| HMDB11294 | C40:7 PC plasmalogen |
| HMDB11310 | C36:4 PC plasmalogen |
| HMDB11343 | C34:3 PE plasmalogen |
| HMDB11384 | C38:3 PE plasmalogen |
| HMDB11386 | C38:5 PE plasmalogen |
| HMDB11387 | C38:6 PE plasmalogen |
| HMDB11394 | C40:7 PE plasmalogen |
| HMDB11410 | C36:5 PE plasmalogen |
| HMDB11420 | C38:7 PE plasmalogen |
| HMDB11441 | C36:3 PE plasmalogen |
| HMDB11442 | C36:4 PE plasmalogen |
| HMDB11503 | C16:0 LPE |
| HMDB11506 | C18:1 LPE |
| HMDB11507 | C18:2 LPE |
| HMDB11517 | C20:4 LPE |
| HMDB11526 | C22:6 LPE |
| HMDB11697 | C24:0 SM |
| HMDB12097 | C14:0 SM |
| HMDB12102 | C20:0 SM |
| HMDB12103 | C22:0 SM |
| HMDB12104 | C22:1 SM |
| HMDB12356 | C34:0 PS |
| HMDB13127 | C4-OH carnitine |
| HMDB13130 | C5-DC carnitine |
| HMDB13238 | C7 carnitine |
| HMDB13288 | C9 carnitine |
| HMDB13325 | C10:2 carnitine |
| HMDB13326 | C12:1 carnitine |
| HMDB13331 | C14:2 carnitine |
| HMDB29416 | NMMA |
| HMDB42063 | C44:0 TAG |

Figure S1. Flow chart of current study.

