Metabolite Profiling and Cardiovascular Event Risk
A Prospective Study of 3 Population-Based Cohorts

Peter Würtz, PhD; Aki S. Havulinna, DScTech; Pasi Soininen, PhD; Tuulia Tynkkynen, PhD; David Prieto-Merino, PhD; Therese Tillin, MBBS; Anahita Ghorbani, MD; Anna Artati, PhD; Qin Wang, MSc; Mika Tiainen, PhD; Antti J. Kangas, MSc; Johannes Kettunen, PhD; Jari Kaikkonen, MSc; Vera Mikkilä, PhD; Antti Jula, MD, PhD; Mika Kähönen, MD, PhD; Terho Lehtimäki, MD, PhD; Debbie A. Lawlor, MD, PhD; Tom R. Gaunt, PhD; Alun D. Hughes, MD, PhD; Naveed Sattar, MD, PhD; Thomas Illig, PhD; Jerzy Adamski, PhD; Thomas J. Wang, MD; Markus Perola, MD, PhD; Samuli Ripatti, PhD; Ramachandran S. Vasan, MD, PhD; Olli T. Raitakari, MD, PhD; Robert E. Gerszten, MD, PhD; Juan-Pablo Casas, MD, PhD; Nish Chaturvedi, MD, PhD; Mika Ala-Korpela, PhD*; Veikko Salomaa, MD, PhD*

Background—High-throughput profiling of circulating metabolites may improve cardiovascular risk prediction over established risk factors.

Methods and Results—We applied quantitative nuclear magnetic resonance metabolomics to identify the biomarkers for incident cardiovascular disease during long-term follow-up. Biomarker discovery was conducted in the National Finnish FINRISK study (n=7256; 800 events). Replication and incremental risk prediction was assessed in the Southall and Brent Revisited (SABRE) study (n=2622; 573 events) and British Women's Health and Heart Study (n=3563; 368 events). In targeted analyses of 68 lipids and metabolites, 33 measures were associated with incident cardiovascular events at $P<0.0007$ after adjusting for age, sex, blood pressure, smoking, diabetes mellitus, and medication. When further adjusting for routine lipids, 4 metabolites were associated with future cardiovascular events in meta-analyses: higher serum phenylalanine (hazard ratio per standard deviation, 1.18; 95% confidence interval, 1.12–1.24; $P=4\times10^{-10}$) and monounsaturated fatty acid levels (1.17; 1.11–1.24; $P=1\times10^{-8}$) were associated with increased cardiovascular risk, while higher omega-6 fatty acids (0.89; 0.84–0.94; $P=6\times10^{-5}$) and docosahexaenoic acid levels (0.90; 0.86–0.95; $P=5\times10^{-5}$) were associated with lower risk. A risk score incorporating these 4 biomarkers was derived in FINRISK. Risk prediction estimates were more accurate in the 2 validation cohorts (relative integrated discrimination improvement, 8.8% and 4.3%), albeit discrimination was not enhanced. Risk classification was particularly improved for persons in the 5% to

Received September 7, 2014; accepted January 2, 2015.

From Computational Medicine, Institute of Health Sciences, University of Oulu, Finland (P.W., P.S., T. Tynkkynen, Q.W., M.T., A.J.K., J. Kettunen, M.A.-K.); Department of Chronic Disease Prevention, National Institute for Health and Welfare, Finland (P.W., A.S.H., J. Kettunen, A.J., M.P., V.S.); Institute for Molecular Medicine Finland, University of Helsinki (P.W., A.S.H., M.P., S.P.); NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland, Kuopio (P.S., T. Tynkkynen, Q.W., M.T., M.A.-K.); Faculty of Epidemiology and Public Health, London School of Hygiene and Tropical Medicine, United Kingdom (D.P.-M., J.-P.C.); Institute of Cardiovascular Science, University College London, United Kingdom (T. Tillin, A.D.H., J.-P.C., N.C.); Framingham Heart Study of the National Heart, Lung, and Blood Institute and Boston University School of Medicine, Framingham, MA (A.G., R.S.V.); Genome Analysis Center, Institute of Experimental Genetics, Helmholtz Zentrum München, Neuherberg, Germany (A.A., J.A.); Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Finland (J. Kaikkonen, V.M., O.T.R.); Department of Food and Environmental Sciences, University of Helsinki, Finland (V.M.,); Department of Clinical Physiology, University of Tampere and Tampere University Hospital, Finland (M.K.); Department of Clinical Chemistry, Fimlab Laboratories, and School of Medicine, University of Tampere, Finland (T.L.); Medical Research Council Integrative Epidemiology Unit at the University of Bristol, United Kingdom (D.A.L., T.R.G., M.A.-K.); School of Social and Community Medicine, University of Bristol, United Kingdom (D.A.L., T.R.G., M.A.-K.); Institute of Cardiovascular and Medical Sciences, University of Glasgow, United Kingdom (N.S.); Hannover Medical School, Hannover Unified Biobank, Germany (T.I.); Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany (T.I.); German Center for Diabetes Research, Neuherberg, Germany (J.A.); Experimental Genetics, Technische Universität München, Freising-Weihenstephan, Germany (J.A.); Cardiovascular Research Center & Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston (T.J.W., R.E.G.); Cardiology Division, Vanderbilt University, Nashville, TN (T.J.W.); Hjelt Institute, University of Helsinki, Finland (S.P.); Wellcome Trust Sanger Institute, Hinxton, United Kingdom (S.P.); Sections of Preventive Medicine and Epidemiology and Cardiology, Department of Medicine, Boston University, MA (R.S.V.); Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Finland (O.T.R.); Broad Institute of MIT and Harvard, Cambridge, MA (R.E.G.); and Oulu University Hospital, Finland (M.A.-K.).

*Drs Ala-Korpela and Salomaa are joint senior authors.

Guest editor for this article was Gregory Y.H. Lip, MD.

The online-only Data Supplement is available with this article at http://circ.ahajournals.org/lookup/suppl/doi:10.1161/CIRCULATIONAHA.114.013116/-/DC1.

Correspondence to Peter Würtz, PhD, Computational Medicine, Institute of Health Sciences, Aapistie 5A, PO Box 5000, 90014 University of Oulu, Finland. E-mail peter.wurtz@computationalmedicine.fi

10% risk range (net reclassification, 27.1% and 15.5%). Biomarker associations were further corroborated with mass spectrometry in FINRISK (n=671) and the Framingham Offspring Study (n=2289).

Conclusions—Metabolite profiling in large prospective cohorts identified phenylalanine, monounsaturated fatty acids, and polyunsaturated fatty acids as biomarkers for cardiovascular risk. This study substantiates the value of high-throughput metabolomics for biomarker discovery and improved risk assessment. **(*Circulation*. 2015;131:774-785. DOI: 10.1161/CIRCULATIONAHA.114.013116.)**

Key Words: amino acids ■ biological markers ■ fatty acids ■ metabolomics ■ risk factors

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality worldwide. Preventive cardiovascular risk assessment relies on established risk factors, including dyslipidemia, hypertension, and diabetes mellitus; however, the first CVD events often originate from people classified as being at low or intermediate risk based on current risk algorithms.[1–4] Detailed profiling of metabolic status, termed metabolite profiling or metabolomics, can provide insights into the molecular mechanisms underlying atherosclerosis.[5–9] The quantification of large numbers of circulating metabolites across multiple pathways may also identify metabolic changes before the onset of overt disease, and thereby potentially lead to earlier and more accurate identification of individuals at high cardiovascular risk.[10–12]

Clinical Perspective on p 785

Metabolite profiling has successfully been applied to identify biomarkers for the development of type 2 diabetes mellitus[13–17]; however, few metabolite biomarkers have been consistently associated with future cardiovascular events across multiple studies.[6,10–12,18] Technological improvements in sample throughput now allow for metabolite profiling of extensive epidemiological cohorts, rather than case-control settings, to enhance biomarker discovery and replication.[8–10,16–22] Serum nuclear magnetic resonance (NMR) metabolomics enables fast, inexpensive, and reproducible quantification of circulating lipids and abundant metabolites.[8,18,19,23–25] Here, we used a high-throughput NMR platform[8] for metabolite profiling in 3 large population-based cohorts with the aim of identifying circulating biomarkers for cardiovascular risk during long-term follow-up. First, 68 lipid and metabolite measures from multiple pathways were tested for association with incident CVD. This hypothesis-generating approach was taken to discover novel biomarkers and thereby gain information on disease mechanisms. The ability to improve cardiovascular risk assessment, beyond that achieved by established risk factors, was examined by a risk score including metabolite biomarkers derived in the discovery study and tested in the 2 validation cohorts. Analytic confirmation of the identified biomarkers was provided by complementary mass spectrometry and gas chromatography. To examine the coherence across metabolomics methodologies, the NMR-based biomarker associations with CVD were further compared with those obtained from mass spectrometry in the discovery study and in the Framingham Offspring Study.

Methods

Study Populations

An overview of the study design is shown in Figure 1. This observational study examined metabolite associations with incident cardiovascular events in the population-based National Finnish FINRISK Study as discovery cohort. Metabolite biomarker candidates were replicated in 2 additional population-based cohorts. All participants provided written informed consent, and study protocols were approved by the local ethical committees. The main end point was the first incidence of a major cardiovascular event during follow-up, which includes fatal or nonfatal occurrence of myocardial infarction, ischemic stroke, cardiac revascularization (coronary artery bypass graft surgery or percutaneous transluminal coronary angioplasty), or unstable angina.[26] Individuals with prevalent CVD at enrolment were omitted from analyses.

The FINRISK 1997 study is a general population survey conducted to monitor the health of the Finnish population among persons aged 25 to 74 at recruitment.[22,26,27] In total, 8444 individuals were recruited from 5 study areas across Finland. Participants completed questionnaires on smoking status, alcohol usage, and medication. Median fasting time was 5 hours (interquartile range, 4–6 hours). Serum total cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides were measured with enzymatic methods. Metabolite profiling by high-throughput NMR was measured during 2012 for 7602 individuals with available serum samples collected at baseline in 1997.[22] Pregnant women and persons missing risk factor information (n=83) and individuals with prevalent CVD (n=263) were excluded, leaving 7256 individuals for statistical analyses. Tracking of CVD during follow-up (1997 through December 2011) was enabled by *International Classification of Diseases, 10th Revision* diagnosis codes from the Finnish National Hospital Discharge Register and Causes-of-Death Register. These registers cover all cardiovascular events that have led either to hospitalization or death in Finland. The cardiovascular diagnoses in these registers have been validated.[28]

Replication of biomarker associations with CVD and improvements in risk prediction were examined in 2 population-based studies from the United Kingdom: fasting serum samples from the Southall and Brent Revisited study (SABRE, n=2622)[29] and the British Women's Heart and Health Study (BWHHS, n=3563)[30] were profiled by the same NMR metabolomics platform as used in FINRISK. Detailed information on the study populations is provided in the expanded methods in the online-only Data Supplement.

Metabolite Quantification

A high-throughput NMR metabolomics platform[8] was used for the quantification of 68 lipid and abundant metabolite measures from baseline serum samples of the FINRISK, SABRE, and BWHHS cohorts. All metabolites were measured in a single experimental setup that allows for the simultaneous quantification of both routine lipids, total lipid concentrations of 14 lipoprotein subclasses, fatty acid composition such as monounsaturated (MUFA) and polyunsaturated fatty acids (PUFA), various glycolysis precursors, ketone bodies, and amino acids in absolute concentration units (Table I in the online-only Data Supplement).[8] The targeted metabolite profiling therefore includes both known metabolic risk factors and metabolites from multiple physiological pathways, which have not previously been examined in relation to CVD risk in large population studies. The 68 metabolite measures were assessed for an association with incident CVD events by using a hypothesis-generating biomarker discovery approach with

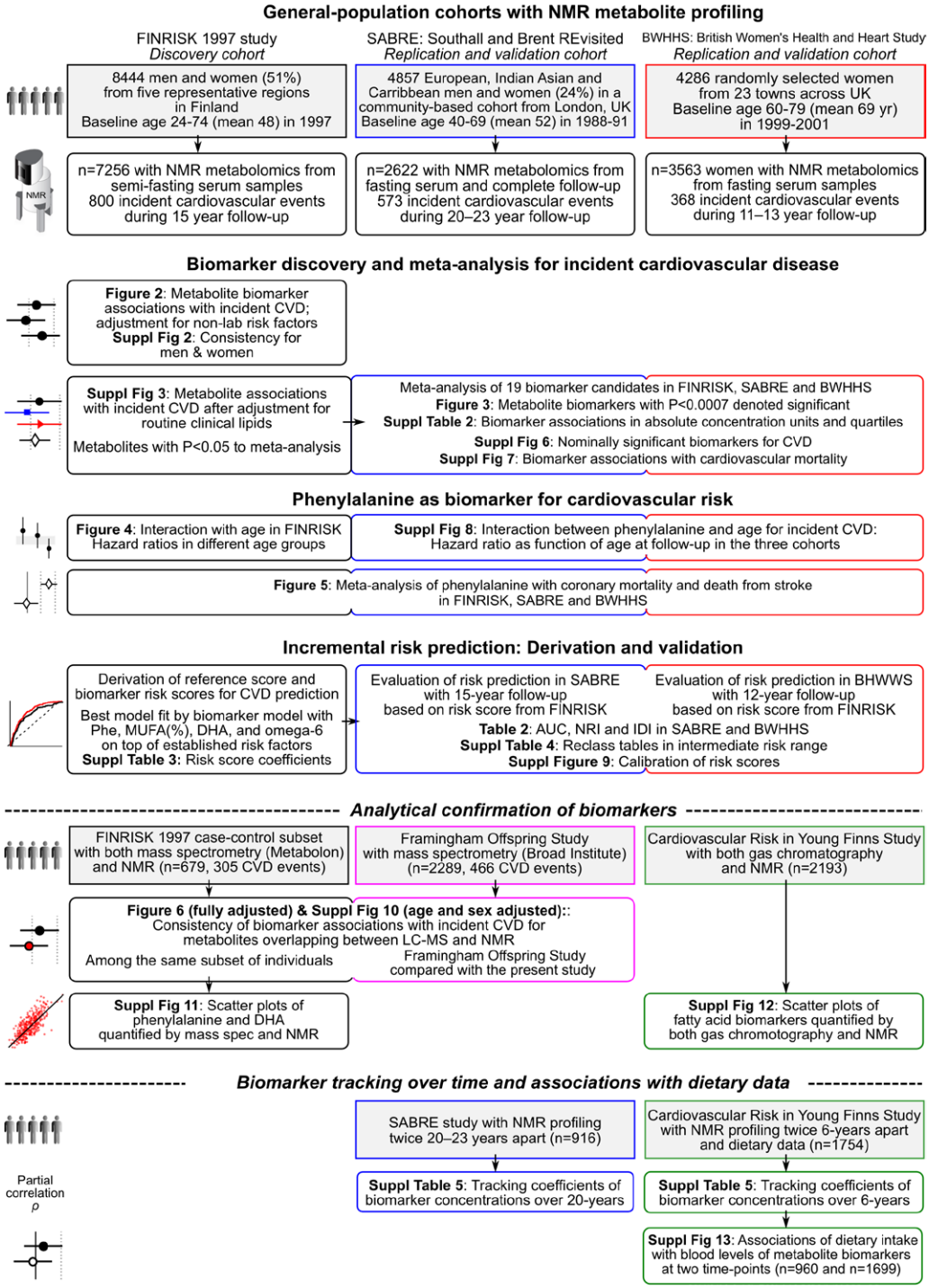


Figure 1. Overview of the study design and statistical analyses conducted. AUC indicates area under the curve; CVD, cardiovascular disease; DHA, docosahexaenoic acid; IDI, integrated discrimination improvement; LC-MS, liquid-chromatography mass spectrometry; MUFA, monounsaturated fatty acids; NMR, nuclear magnetic resonance; NRI, net reclassification improvement; and Phe, phenylalanine.

subsequent replication in 2 independent cohorts. Spearman correlations of the metabolites are shown in Figure I in the online-only Data Supplement. The NMR metabolomics platform has previously been used in various epidemiological studies,[9,10,16,17,20–22,31,32] details of the experimentation have been described,[9,24] and the method has recently been reviewed.[8,19]

A subset of 679 serum samples from the FINRISK study were additionally profiled with liquid-chromatography mass spectrometry (LC-MS) using the Metabolon platform[33] in a case-cohort design for comparison of biomarker associations with incident CVD (online-only Data Supplement Methods). The biomarker associations were further compared with those obtained by LC-MS–based profiling of the Framingham Offspring Study (fifth examination cycle, n=2289 fasting plasma samples), as described in detail previously.[13,14] Because several fatty acid biomarkers were not measured by LC-MS, the quantification was analytically confirmed by comparing NMR and gas chromatography in the Cardiovascular Risk in Young Finns Study (YFS, n=2193 fasting serum samples).[34] Metabolite profiling data collected at 2 time points in YFS[9] were used further to examine the associations of dietary intake with the

circulating biomarkers, and tracking of concentrations within the same individuals over 6 years.

Statistical Analyses

All metabolite concentrations were first log-transformed before analyses to obtain approximately normal distributions. The metabolite measures were subsequently scaled to standard deviation units separately for each cohort. Associations with incident CVD were analyzed separately for each metabolite by using Cox proportional hazards regression models. In the FINRISK discovery study, metabolite associations were first adjusted for age (as time scale), sex, systolic blood pressure, smoking, prevalent diabetes mellitus, antihypertensive treatment, lipid treatment, and geographical region, and subsequently tested with additional adjustment for routine lipid measures (total cholesterol and HDL-cholesterol). Metabolites associated with CVD at $P<0.05$ when adjusting for routine lipids were then analyzed in the 2 replication cohorts with full adjustment. Analyses in SABRE were further adjusted for ethnicity. The results from individual cohorts were combined by using inverse variance-weighted fixed-effect meta-analysis. Metabolites associated with incident CVD at $P<0.0007$ in meta-analyses were denoted significant biomarkers (Bonferroni correction of $P<0.05$ accounting for 68 independent tests). Sensitivity analyses were conducted in the FINRISK study with additional exclusion criteria and covariate adjustment. Metabolites that were nominally significant in the meta-analysis ($P<0.05$) were also tested for association with cardiovascular mortality. Since phenylalanine did not meet the proportional hazards assumptions by the scaled Schoenfeld residuals test in the FINRISK study, the association was tested for interaction with age. Phenylalanine was further examined for association with death from coronary heart disease and stroke in meta-analyses.

The potential to improve cardiovascular risk prediction was evaluated by risk scores derived based on established factors with and without the significant metabolite biomarkers in the models. A multibiomarker risk score was derived in the FINRISK discovery cohort; all combinations of the 5 significant biomarkers were tested, with established risk factors always included in the model. The model giving rise to the best fit based on the Akaike Information Criterion was selected. The logarithm of the hazard ratios in the multivariable model were used as the linear predictor for the biomarker risk score. A reference risk score was also derived in the FINRISK study by using only the conventional risk factors to define the linear predictor. Because of differences in hazard depending on age and geographical regions, the baseline hazard term was derived within each validation cohort. The predictive utility of the biomarker risk score was examined in the 2 validation cohorts in terms of risk discrimination and reclassification. Prediction estimates were calculated as 15-year absolute risk in SABRE (matching the FINRISK follow-up) and as 12-year absolute risk in BWHHS. Discrimination was assessed by the correlated censored C-statistic approach with the use of jackknife estimation accounting for censoring[35] and integrated discrimination improvement (IDI). IDI is the difference of mean predicted probabilities (absolute risk estimates) between the biomarker score and the reference score for events minus the corresponding difference in predicted probabilities for nonevents.[36,37] IDI thus denotes the average increase in absolute risk estimates by the biomarker risk score for individuals who experienced a CVD event plus the average decrease in absolute risk estimates for those who did not have a CVD event. Because the absolute risk estimates are generally low in the study populations, the relative IDI offers a more intuitive understanding of the average improvement in risk prediction accuracy achieved by the biomarker model.[36] Net reclassification improvement (NRI) was examined to determine the extent to which the biomarker risk score reassigned individuals to risk categories that more correctly reflected whether or not they experienced a CVD event during follow-up. NRI was assessed in 2 risk category ranges: (1) only for individuals in the intermediate risk range of 5% to 10% based on the reference score (clinical NRI),[3,4,38] and (2) for the whole study population by using the risk categories <5%, 5% to 10%, and >10% (categorical NRI).[36] In addition, we assessed the continuous NRI, which deems any change in predicted risk in the correct direction as appropriate without dependence on risk categories.[37] All reclassification metrics were examined separately for events and nonevents, and net reclassification denotes the unweighted sum of the 2. Model calibration within risk deciles was assessed by the Hosmer-Lemeshow goodness-of-fit test.[2]

Assessment of the consistency of biomarker associations between NMR and mass spectrometry and analytic confirmation of the biomarker quantification is described in online-only Data Supplement Methods. Details of the biomarker associations with dietary data and tracking of metabolite concentrations over time can also be found in online-only Data Supplement Methods.

Results

The discovery study included 7256 individuals from the FINRISK general-population cohort, free of CVD at baseline. During a follow-up of 15 years, 800 persons experienced an incident cardiovascular event. Metabolite biomarker candidates for CVD were replicated in the 2 United Kingdom–based population-based cohorts: 2622 individuals from SABRE and 3563 women from BWHHS, with a total of 941 incident cardiovascular events during 12- to 23-year follow-up. Baseline characteristics and cardiovascular event numbers are summarized in Table 1.

Table 1. Baseline Characteristics of the Study Populations

Clinical Characteristics	FINRISK 1997 Study (n=7256)	Southall and Brent Revisited Study (SABRE) (n=2622)	British Women's Heart and Health Study (BWHHS) (n=3563)
Women, n (%)	3678 (51)	386 (21)	3563 (100)
Location	6 regions in Finland	London, UK	23 British towns
Age, y	48±13	52±7	69±5
Body mass index, kg/m^2	26.6±4.5	26.1±3.7	27.5±4.9
Systolic blood pressure, mm Hg	136±20	124±18	147±25
Total cholesterol, mmol/L	5.5±1.1	6.0±1.1	6.7±1.2
HDL cholesterol, mmol/L	1.4±0.4	1.3±0.4	1.7±0.5
Triglycerides, mmol/L	1.5±1.0	1.8±1.2	1.8±1.0
Lipid-lowering medication, n (%)	185 (2.6)	5 (0.4)	249 (7)
Antihypertensive medication, n (%)	908 (13)	294 (11)	1033 (29)
Current smoking, n (%)	1736 (24)	592 (23)	392 (11)
Diabetes prevalence, n (%)	393 (5.4)	314 (12)	356 (10)
Follow-up time, y, range	15	20–23	11–13
Incident cardiovascular events	800	573	368

Data are number (%) or mean±standard deviation when appropriate. HDL indicates high-density lipoprotein.

Metabolite Associations With Incident Cardiovascular Events

To generate hypotheses on the roles of the metabolite measures in cardiovascular pathophysiology, we first examined each metabolite for association with incident CVD in the FINRISK discovery study with adjustment only for established nonlaboratory risk factors: age as time-scale, sex, systolic blood pressure, smoking, prevalent diabetes mellitus, geographical region, lipid medication, and antihypertensive treatment.[2–4] Associations of the 68 lipid and metabolite measures with incident cardiovascular events are illustrated in Figure 2. Hazard ratios are scaled to 1–standard deviation increments in log-transformed metabolite concentrations to facilitate comparison across metabolites. Overall, 33 metabolite measures were associated with future cardiovascular events at $P<0.0007$ (Bonferroni correction for 68 tests). Prominent associations with increased CVD risk were observed for higher circulating concentrations of aromatic amino acids, glycolysis metabolites, MUFA relative to total fatty acids (MUFA%), glycoprotein acetyls, as well as the lipid concentrations within medium and small very-low-density lipoproteins, intermediate-density lipoprotein, and low-density lipoproteins. The strongest inverse associations with cardiovascular risk were found for lower concentrations of lipids within large HDL particles and for HDL-cholesterol. Noticeably inverse associations were also observed for ketone bodies, and ratios of omega-3 and omega-6 relative to total fatty acids, as well. The metabolite associations were broadly consistent between men and women (Figure II in the online-only Data Supplement).

To discover metabolite biomarkers for incident CVD independent of routine lipids, analyses were further adjusted for

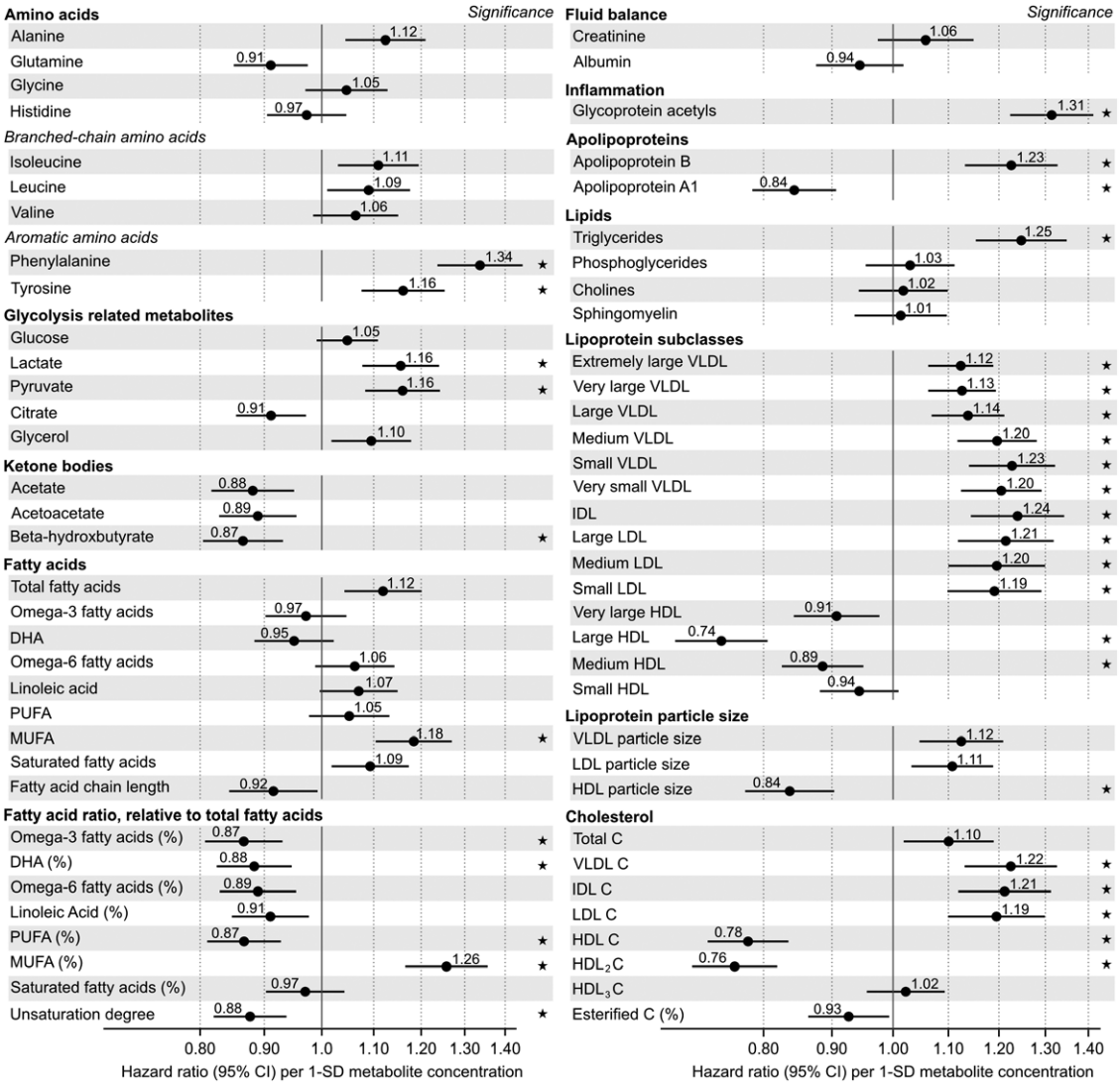


Figure 2. Metabolite associations with future cardiovascular events. Hazard ratios of 68 metabolite measures with incident cardiovascular disease during 15-year follow-up in the FINRISK study (n=7256, 800 events). Hazard ratios are per 1-SD log-transformed metabolite concentration and adjusted for age, sex, blood pressure, smoking, diabetes mellitus, geographical region, and cardiovascular medication. Error bars denote 95% confidence intervals. *$P<0.0007$ (multiple testing correction). C indicates cholesterol; CI, confidence interval; DHA, docosahexaenoic acid; HDL, high-density lipoprotein; LDL, low-density lipoprotein; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; SD, standard deviation; and VLDL, very low density lipoprotein.

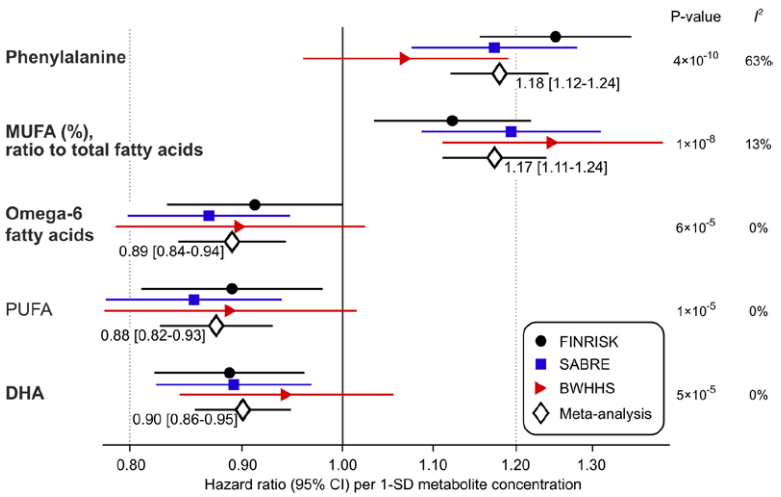


Figure 3. Meta-analysis of metabolite biomarkers for future cardiovascular events. Hazard ratios of biomarkers with incident cardiovascular events in 3 population-based studies and meta-analysis (n=13 441; 1741 events during 12–23 years follow-up). Analyses are adjusted for age, sex, blood pressure, smoking, diabetes mellitus, geographical region, cardiovascular medication, as well as total and HDL cholesterol. Hazard ratios are per 1-SD log-transformed metabolite concentration and error bars denote 95% confidence intervals. I^2 indicates heterogeneity of meta-analysis. Metabolites associated with cardiovascular events at $P<0.0007$ (multiple testing correction) are shown here; associations with $P<0.05$ are listed in Figure VI in the online-only Data Supplement. The 4 biomarkers highlighted in bold are independent of each other and were included in the risk prediction score. CI indicates confidence interval; HDL, high-density lipoprotein; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acids; and SD, standard deviation.

total- and HDL-cholesterol. In the FINRISK discovery study, 19 metabolite measures were associated with cardiovascular events at $P<0.05$ when adjusting for routine lipids in the model (Figure III in the online-only Data Supplement). These biomarker candidates were then further analyzed in the 2 replication cohorts. In meta-analyses, 5 metabolite measures were significantly associated with cardiovascular events ($P<0.0007$; Figure 3). Higher phenylalanine and MUFA% levels were associated with increased cardiovascular event risk. Higher concentrations of omega-6 fatty acids, total concentrations of PUFA, and docosahexaenoic acid (DHA; an omega-3 fatty acid) were associated with lower risk. The biomarker associations with CVD in absolute concentration units (without log-transform and standard deviation scaling), and for upper versus lower quartiles, as well, are listed in Table II in the online-only Data Supplement. The biomarker associations remained similar with additional subject exclusion criteria and adjustment factors (Figure IV in the online-only Data Supplement). In particular, associations were stronger with further adjustment for serum triglycerides. The magnitudes of the biomarker associations were comparable to those of routine lipid measures (Figure V in the online-only Data Supplement). Metabolite measures displaying nominal association with incident CVD are listed in Figure VI in the online-only Data Supplement. The metabolite associations were similar or stronger with cardiovascular mortality as outcome (Figure VII in the online-only Data Supplement).

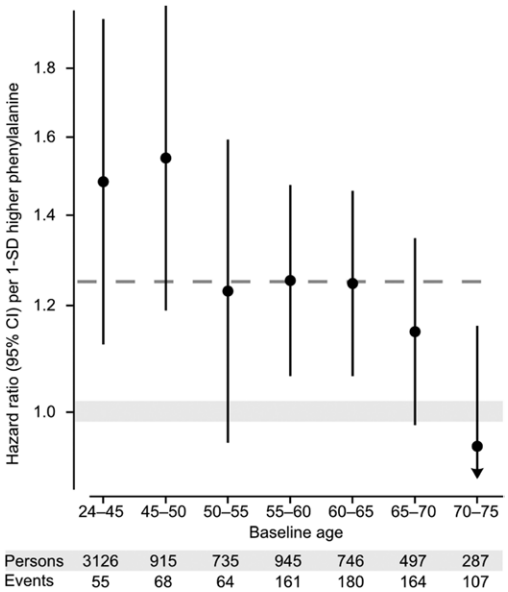


Figure 4. Phenylalanine interaction with age. Hazard ratios of phenylalanine with incident cardiovascular disease in different baseline age groups for the FINRISK study. Analyses were adjusted for age, sex, blood pressure, smoking, diabetes mellitus, geographical region, cardiovascular medication, as well as total and HDL cholesterol. Hazard ratios are per 1-SD higher log-transformed phenylalanine concentration (approximately corresponding to 14 μmol/L). Error bars denote 95% confidence intervals. The dashed line denotes the hazard ratio for the full age range. The continuous interaction of phenylalanine with age is shown for all 3 cohorts in Figure VIII in the online-only Data Supplement. CI indicates confidence interval; and HDL, high-density lipoprotein.

Although the fatty acid measures displayed coherent associations with incident CVD in each cohort, some study heterogeneity was evident in the case of phenylalanine (meta-analysis heterogeneity statistic $I^2=63\%$, $P=0.07$; Figure 3). A potential reason may be the interaction of phenylalanine with age at the end of follow-up ($P=0.001$ in FINRISK), indicating stronger associations of phenylalanine with cardiovascular events at a younger age (Figure 4). This interaction, with ≈1% weaker association per year, was consistent in SABRE but did not replicate in the older population, with narrower age range (60–79 years at baseline) from BWHHS (Figure VIII in the online-only Data Supplement). Although phenylalanine was only weakly associated with incident cardiovascular events in BWHHS, the amino acid was consistently associated with death from coronary heart disease across all 3 cohorts (Figure 5). In contrast, no association was observed for phenylalanine with death from stroke.

Cardiovascular Risk Score Validation and Reclassification

A risk prediction score including metabolite biomarkers was derived in the FINRISK discovery study. The best model

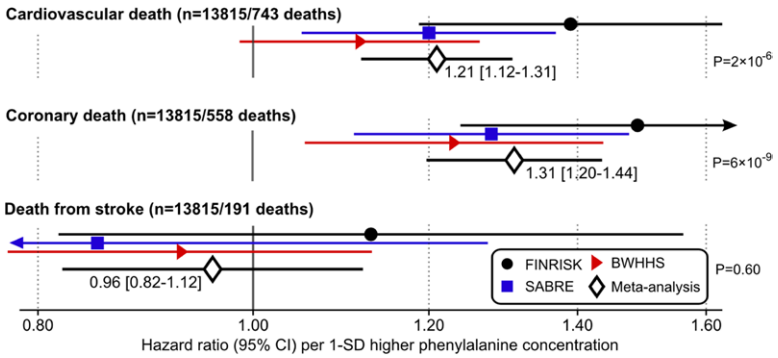


Figure 5. Phenylalanine associations with vascular mortality. Hazard ratios of phenylalanine with death from cardiovascular disease, coronary heart disease, and stroke in the 3 cohorts and meta-analysis (n=13 815). Analyses were adjusted for age, sex, blood pressure, smoking, diabetes mellitus, geographical region, cardiovascular medication, and total and HDL cholesterol. Hazard ratios are per 1-SD higher log-transformed phenylalanine concentration (approximately corresponding to 14 μmol/L). Error bars denote 95% confidence intervals. The numbers of coronary and stroke deaths sum to a slightly higher number than the cardiovascular deaths, because we have considered all causes of death written in the death certificate. CI indicates confidence interval; and HDL, high-density lipoprotein.

fit was obtained by incorporating 4 of the 5 biomarkers (Figure 3) in the score along with established risk factors: phenylalanine, MUFA%, omega-6 fatty acids, and DHA. Coefficients for calculating the risk prediction scores with and without these biomarkers are listed in Table III in the online-only Data Supplement. The potential of the biomarker score to improve risk discrimination and reclassification was tested in the SABRE and BWHHS cohorts (Table 2). Although discrimination assessed by using the C-statistic was not enhanced, the risk prediction estimates were on average more accurate in the 2 validation cohorts, both for those who had a CVD event during follow-up (mean improvement in risk prediction accuracy, ie, relative IDI 4.5% for SABRE and 3.2% for BWHHS) and for those who did not (relative IDI 4.4% for SABRE and 1.1% for BWHHS). The model calibration was modest (Figure IX in the online-only Data Supplement), in particular, for the older women from BWWHS, where both the reference and biomarker risk scores provided mediocre discrimination. The continuous risk reclassification was improved by the biomarker risk score among individuals who did not experience a cardiovascular event (26.8% for SABRE and 15.6% for BWHHS). The net categorical NRI for the whole study population was 7.6% for SABRE and 5.3% for BWHHS. Most notably, for persons classified in the intermediate risk range (5%–10%) based on the reference risk score there was a substantial upclassification by the biomarker score to >10% risk among individuals who actually developed CVD (clinical NRI among events 20.5% for SABRE and 9.8% for BWHHS). Similarly, there was a significant downclassification by the biomarker score to <5% for those who remained free of CVD during follow-up (clinical NRI among nonevents 6.6% for SABRE and 5.7% for BWHHS). Reclassification tables for the intermediate risk range are shown in Table IV in the online-only Data Supplement.

Cross-Platform Biomarker Confirmation

To verify the biomarker associations across metabolomics methods, we tested metabolites that were overlapping between the NMR platform and LC-MS for association with incident CVD. In the Framingham Offspring Study (n=2289, 466 events during 12-year follow-up), the biomarker associations were consistent with those obtained from NMR, albeit phenylalanine was weaker (Figure 6). The biomarker associations were also consistent between NMR and LC-MS in comparison within the same set of individuals in a case-cohort subset of the FINRISK study (n=679, 305 incident events; Figure 6). Comparisons of biomarker associations across metabolomics platforms with adjustment for age and sex only are shown in Figure X in the online-only Data Supplement. The correspondence between phenylalanine (Pearson correlation r=0.62) and DHA (r=0.77) quantified by NMR and LC-MS is shown in Figure XI in the online-only Data Supplement. The MUFA ratio, relative to total fatty acids, and omega-6 fatty acids were not measured by mass spectrometry; quantification of these biomarkers was instead confirmed by comparison with gas chromatography in 2193 participants from the YFS cohort (Figure XII in the online-only Data Supplement). The fatty acid quantification was highly consistent between methods: r=0.92 for MUFA%, r=0.97 for omega-6 fatty acids and r=0.95 for DHA.

Biomarker Tracking and Dietary Associations

The metabolite biomarker levels were consistent over 6-year follow in YFS (tracking correlations, 0.41–0.47) and only slightly weaker over 20-year follow-up in the SABRE study (Table V in the online-only Data Supplement). We further examined the associations of dietary intake with the circulating biomarker concentrations at 2 time points in YFS. Dietary DHA was robustly associated with serum DHA levels (P<0.0001 at both time points), whereas dietary measures of phenylalanine, omega-6, and MUFA% were not strongly associated with the corresponding circulating levels (Figure XIII in the online-only Data Supplement). Circulating phenylalanine levels were not associated with aspartame in the 652 individuals from FINRISK with data available on this artificial sweetener (Spearman correlation –0.06, P=0.14).

Discussion

Using high-throughput metabolite profiling in 3 general-population studies, we identified phenylalanine and 3 measures of fatty acids as independent biomarkers for future cardiovascular events. The circulating metabolites were as strongly predictive of cardiovascular risk as the conventional lipid risk factors, and were markers of CVD onset during more than a decade follow-up. The biomarker associations

Table 2. Discrimination and Reclassification of Cardiovascular Risk in Validation Cohorts With and Without Metabolic Biomarkers in the Risk Prediction Score

Prediction Model	C-Statistic (95% CI)	Model Calibration	Reclassification	Clinical NRI* (95% CI)	Categorical NRI† (95% CI)	Continuous NRI (95% CI)	IDI (95% CI)	Relative IDI (95% CI)
Southall and Brent Revisited study (SABRE, n=2478); 379 events during 15-y follow-up								
Reference risk score	0.712 (0.695–0.745)	P=0.14	CVD events	20.5% (3.2% to 37.8%) P=0.02	–2.4% (–5.3% to 0.59%) P=0.12	–3.7% (–13.7% to 6.4%) P=0.47	0.95% (0.17% to 1.72%) P=0.02	4.5% (1.3% to 7.7%) P=0.006
Biomarker risk score	0.720 (0.687–0.738) P=0.18‡	P=0.03	Nonevents	6.6% (1.7% to 11.5%) P=0.008	10.0% (8.0% to 12.0%) $P=2\times10^{-22}$	26.8% (22.5% to 31.1%) $P=1\times10^{-34}$	0.43% (0.20% to 0.65%) P=0.0002	4.4% (3.0% to 5.7%) $P=8\times10^{-11}$
			Net	27.1% (9.1% to 45.0%) P=0.003	7.6% (4.1% to 11.2%) $P=2\times10^{-5}$	23.1% (12.2% to 34.0%) $P=3\times10^{-5}$	1.37% (0.57% to 2.2%) P=0.008	8.8% (5.4% to 12.3%) $P=6\times10^{-7}$
British Women's Health and Heart Study (BWHHS, n=3348); 354 events during 12-y follow-up								
Reference risk score	0.665 (0.636–0.695)	P=0.02	CVD events	9.8% (–1.3% to 20.9%) P=0.08	0.28% (–3.7% to 4.3%) P=0.89	–6.8% (–17.2% to 3.6%) P=0.20	0.64% (0.12% to 1.17%) P=0.02	3.2% (0.2% to 6.2%) P=0.04
Biomarker risk score	0.666 (0.637–0.694) P=0.97‡	P=0.0003	Nonevents	5.7% (2.4% to 9.0%) P=0.0007	5.0% (3.4% to 6.6%) $P=2\times10^{-9}$	15.6% (12.1% to 19.2%) $P=3\times10^{-18}$	–0.004% (–0.11% to 0.12%) P=0.94	1.1% (0.12% to 2.1%) P=0.03
			Net	15.5% (3.9% to 27.0%) P=0.009	5.3% (0.94% to 9.6%) P=0.02	8.9% (–2.1% to 19.8%) P=0.11	0.64% (0.10% to 1.17%) P=0.02	4.3% (1.2% to 7.5%) P=0.007

Risk assessment was evaluated in the SABRE (15-year absolute risk) and BWHHS (12-year absolute risk) validation cohorts based on the risk prediction scores derived in the FINRISK study (Table III in the online-only Data Supplement). The reference risk score included age, sex, smoking, diabetes mellitus, lipid and blood pressure treatment, blood pressure, total cholesterol, and HDL cholesterol. The biomarker risk score additionally included log-transformed phenylalanine, the ratio of MUFA to total fatty acids, omega-6 fatty acids, and DHA. The interpretations of NRI and IDI metrics are explained in Methods. Calibration of the models within risk deciles is shown in Figure IX in the online-only Data Supplement. CI indicates confidence interval; CVD, cardiovascular disease; DHA, docosahexaenoic acid; HDL, high-density lipoprotein; IDI, integrated discrimination improvement; MUFA, monounsaturated fatty acids; and NRI, net reclassification improvement.

*Reclassification of individuals in the 5% to 10% intermediate risk range based on the reference risk score (650 in SABRE and 1288 in BWHHS) to risk classes <5% and >10%. Reclassification tables for this intermediate risk range are shown in Table IV in the online-only Data Supplement.

†Reclassification among all study participants to risk classes <5%, 5% to 10%, and >10%.

‡P value for difference from reference model.

replicated in independent cohorts with varying baseline characteristics including age and ethnicity, and were consistent across different metabolite profiling platforms. Whereas higher circulating levels of DHA and omega-6 fatty acids have been linked with lower CVD event risk in some studies,[39–41] the blood levels of MUFA and phenylalanine have not previously been associated with higher risk for future CVD in large epidemiological studies. These results demonstrate the power of detailed metabolite profiling for biomarker discovery in large prospective cohorts, which can yield improved molecular understanding of disease mechanisms. In combination, the 4 biomarkers indicated improved cardiovascular risk assessment for people in the intermediate risk range, where clinical decision making remains ambiguous.

Phenylalanine is an essential aromatic amino acid and the precursor for tyrosine and dopamine-related neurotransmitters. The mechanisms by which blood levels of phenylalanine relate to cardiovascular risk remain unknown. Phenylalanine has been associated with insulin resistance and the risk for diabetes mellitus,[13–16] but the association with CVD remained similar after adjustment for glucose and insulin. A small case-control study using LC-MS suggested that an amino acid score including phenylalanine is associated with CVD.[42] Although phenylalanine alone was not statistically significant in that study, the hazard ratio (1.25; P=0.11) matched the magnitude observed in the present study. The ratio of phenylalanine to tyrosine has been suggested as a proxy of phenylalanine hydroxylase activity[43]; however, this ratio was not associated with CVD event risk in the present study (hazard ratio=1.03; P=0.47). Phenylalanine was more strongly associated with CVD before 60 years of age. This amino acid therefore represents a promising biomarker for the early identification of cardiovascular risk.

Dietary recommendations support MUFA intake replacing saturated fatty acid to lower cardiovascular risk.[44] In this study, high blood levels of MUFAs, relative to total fatty acids, were associated with higher cardiovascular risk. Similar results have been found for the risk of type 2 diabetes.[17] However, circulating MUFA concentrations do not directly reflect dietary intake, because MUFAs constitute the major fatty acids stored in adipose tissue and can be synthesized de novo. A potential explanation for the positive association with CVD may be that the MUFA levels originate from the desaturation of dietary saturated fatty acids.[17,45] As such, MUFA% could potentially act as a blood biomarker reflecting long-term dietary quality.

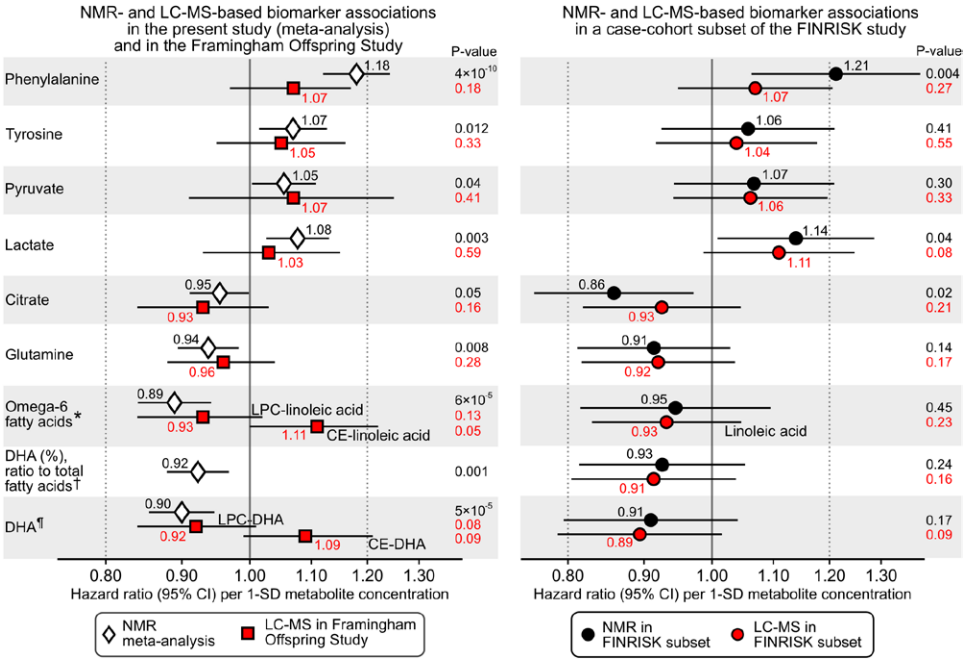


Figure 6. Consistency between NMR and LC-MS for biomarker associations with incident cardiovascular disease. Metabolites overlapping between metabolomics platforms and nominally associated with incident CVD in the present study ($P<0.05$, Figure VI in the online-only Data Supplement) are shown. **Left**, Biomarker associations with CVD risk observed in the present study based on NMR (open diamonds) in comparison with those obtained in the Framingham Offspring Study based on LC-MS (red squares; n=2289, 466 events). **Right**, Biomarker associations with CVD risk in a case-cohort subset of the FINRISK study (n=679, 305 events) profiled both by NMR (black circles) and LC-MS (red circles). Hazard ratios are per 1-SD higher log-transformed metabolite concentration. Error bars denote 95% confidence intervals. All associations were adjusted for age, sex, blood pressure, smoking, diabetes mellitus status, geographical region, cardiovascular medication, and total and HDL cholesterol. LC-MS-based associations were further adjusted for batch. The corresponding age- and sex-adjusted biomarker associations are shown in Figure X in the online-only Data Supplement. *Associations of omega-6 fatty acids were compared with lysophosphatidylcholine- and cholesterol ester-linoleic acid in the Framingham Offspring Study, and with total linoleic acid in the FINRISK subset. †Associations of DHA were compared with lysophosphatidylcholine- and cholesterol ester-DHA in the Framingham Offspring Study. ¶DHA ratio was scaled to the total fatty acid concentration quantified by NMR for both platforms. The DHA ratio was not measured in the Framingham Offspring Study. CI indicates confidence interval; CVD, cardiovascular disease; DHA, docosahexaenoic acid; HDL, high-density lipoprotein; LC-MS, liquid-chromatography mass spectrometry; NMR, nuclear magnetic resonance; and SD, standard deviation.

The role of PUFAs in CVD pathogenesis remains controversial.[39–41,46–49] Intervention trials do not suggest risk reduction by omega-3 supplementation.[46,47] Trials on omega-6 fatty acid consumption are less clear, because most studies evaluate e effects of replacing saturated fatty acids.[48,49] We observed consistent associations of higher blood levels of both DHA and omega-6 fatty acids with lower CVD risk, in agreement with some prospective studies on circulating PUFAs.[39–41] These findings contrast the results on the risk for incident type 2 diabetes, where only omega-6 fatty acids display inverse associations.[17] Circulating PUFAs might serve as more reliable markers of therapeutic target and cardiovascular risk than conventional dietary assessments.[10,39–41] Regardless of the therapeutic benefit, DHA and omega-6 fatty acid quantification could potentially augment risk assessment on top of established risk factors.

Individual biomarkers rarely improve risk prediction.[1] The biomarkers identified here were quantified by a single analytic platform, which also features the measurement of routine lipid risk factors.[8,9,20,24,31,32] The combination of 4 biomarkers yielded improvements in risk prediction accuracy when evaluated in 2 independent cohorts. The average improvement in risk prediction accuracy (relative IDI) was ≈4.5% in SABRE for both individuals who developed CVD and those who did not, and 1% to 3% in BWHHS. This modest improvement might relate to differences in age and sex distributions, population sampling, ethnicity, and geographical region between derivation and validation cohorts. The poor model calibration and lower reclassification rates in BWHHS may potentially be explained by the female-only composition, and the higher baseline age of this study cohort. Novel biomarkers for risk prediction are primarily needed for persons in the intermediate risk range, for whom treatment decisions are most challenging.[1–4] The 4 biomarkers proved particularly helpful in correctly reclassifying individuals in the 5% to 10% risk gray zone (net 27% and 15% in the 2 validation cohorts). Although there are no widely accepted definitions of intermediate risk for the composite CVD end point studied here, the continuous reclassification metrics were improved in both validation studies. Nonetheless, additional investigations should further evaluate the clinical utility of these biomarkers. In particular, metabolite profiling in clinical trials could inform on the feasibility of using the biomarkers to improve cardiovascular risk assessment. The throughput of the metabolomics

platform used now allows for profiling all samples collected in even the largest CVD prevention trials.[8] This would elucidate the role of detailed metabolite profiling in high-risk individuals and patients with prevalent disease, and address the value for risk assessment among patients already on statin medication.

Our study has several strengths, including the large number of individuals studied, biomarker replication, assessment of risk prediction in independent cohorts, and confirmation of biomarker associations across metabolomics methodologies. Some limitations should also be considered. First, the metabolite coverage by NMR is limited in comparison with that afforded by mass spectrometry, which might further improve risk prediction.[5,12] Second, blood sample collections were done before the widespread use of lipid-lowering medication, and the benefit of modern preventive treatment based on risk classification by the highlighted biomarkers remains uncertain.

Conclusion

Using NMR metabolomics profiling of 13 441 individuals with long-term follow-up, we have identified phenylalanine and MUFA as novel biomarkers of higher CVD event risk, and corroborated omega-6 fatty acids and DHA as biomarkers of lower CVD risk. The combination of the 4 independent biomarkers indicated improved prediction of cardiovascular risk for persons classified with intermediate risk based on established risk factors. Further studies are needed to elucidate the biological mechanisms underlying the associations with CVD and to clarify the clinical utility of these biomarkers to guide cardiovascular risk assessment. Overall, our investigation underscores the value of high-throughput metabolite profiling in the discovery of new and emerging biomarkers for CVD risk and their potentially cost-effective use for cardiovascular prevention.

Sources of Funding

This study was supported by the Academy of Finland (139635, 137870, 250422, 251217, 266199), the European Commission Seventh Framework Programme (BioSHaRE 261433), the Sigrid Juselius Foundation, the Yrjö Jahnsson Foundation, the Emil Aaltonen Foundation, the Paavo Nurmi Foundation, the Finnish Foundation for Cardiovascular Research, the Medical Research Fund of Tampere, the UK Medical Research Council via the University of Bristol Integrative Epidemiology Unit (IEU; MC_UU_12013/5), and Strategic Research Funding from the University of Oulu, Finland. The Wellcome Trust (WT082464AIA), British Heart Foundation (SP/07/001/23603) and Diabetes UK (13/0004774) support the SABRE Study. The British Women's Heart and Health Study was funded by the British Heart Foundation and UK Department of Health Policy Research Programme with metabolomics data funded by the UK Medical Research Council (G1000427). Metabolomics in the Framingham offspring study was supported by the US National Institutes of Health grants R01 DK 081572 and R01 HL98280. Surveillance of CVD in the Framingham Heart Study is supported by N01-HC-25195 (from the NHLBI).

Disclosures

Drs Würtz, Kangas, Soininen, and Ala-Korpela are shareholders of Brainshake Ltd, a startup company offering NMR based metabolite profiling. Drs Tynkkynen, Q. Wang, Tiainen, and Kettunen report employment (minor compensation) for Brainshake Ltd. The other authors report no conflicts.

References

1. Wang TJ. Assessing the role of circulating, genetic, and imaging biomarkers in cardiovascular risk prediction. *Circulation*. 2011;123:551–565.
2. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation*. 2010;121:1768–1777.
3. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero ST, Smith SC Jr, Watson K, Wilson PW, Eddleman KM, Jarrett NM, LaBresh K, Nevo L, Wnek J, Anderson JL, Halperin JL, Albert NM, Bozkurt B, Brindis RG, Curtis LH, DeMets D, Hochman JS, Kovacs RJ, Ohman EM, Pressler SJ, Sellke FW, Shen WK, Smith SC Jr, Tomaselli GF. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;129(25 suppl 2):S1–45.
4. Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease. NICE guidelines [CG181]. London, UK: National Institute for Health and Care Excellence; July 2014. http://www.nice.org.uk/guidance/cg181. Accessed September 4, 2014.
5. Quehenberger O, Dennis EA. The human plasma lipidome. *N Engl J Med*. 2011;365:1812–1823.
6. Shah SH, Kraus WE, Newgard CB. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. *Circulation*. 2012;126:1110–1120.
7. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung YM, Wu Y, Schauer P, Smith JD, Allayee H, Tang WH, DiDonato JA, Lusis AJ, Hazen SL. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011;472:57–63.
8. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. *Circ Cardiov Genet*. 2015. In press. doi: 10.1161/CIRCGENETICS.114.000216.
9. Inouye M, Kettunen J, Soininen P, Silander K, Ripatti S, Kumpula LS, Hämäläinen E, Jousilahti P, Kangas AJ, Männistö S, Savolainen MJ, Jula A, Leiviskä J, Palotie A, Salomaa V, Perola M, Ala-Korpela M, Peltonen L. Metabonomic, transcriptomic, and genomic variation of a population cohort. *Mol Syst Biol*. 2010;6:1–10.
10. Würtz P, Raiko JR, Magnussen CG, Soininen P, Kangas AJ, Tynkkynen T, Thomson R, Laatikainen R, Savolainen MJ, Laurikka J, Kuukasjärvi P, Tarkka M, Karhunen PJ, Jula A, Viikari JS, Kähönen M, Lehtimäki T, Juonala M, Ala-Korpela M, Raitakari OT. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J*. 2012;33:2307–2316.
11. Tang WH, 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.
12. Stegemann C, Pechlaner R, Willeit P, Langley SR, Mangino M, Mayr U, Menni C, Moayyeri A, Santer P, Rungger G, Spector TD, Willeit J, Kiechl S, Mayr M. Lipidomics profiling and risk of cardiovascular disease in the prospective population-based Bruneck study. *Circulation*. 2014;129:1821–1831.
13. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, Lewis GD, Fox CS, Jacques PF, Fernandez C, O'Donnell CJ, Carr SA, Mootha VK, Florez JC, Souza A, Melander O, Clish CB, Gerszten RE. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17:448–453.
14. Rhee EP, Cheng S, Larson MG, Walford GA, Lewis GD, McCabe E, Yang E, Farrell L, Fox CS, O'Donnell CJ, Carr SA, Vasan RS, Florez JC, Clish CB, Wang TJ, Gerszten RE. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J Clin Invest*. 2011;121:1402–1411.
15. Floegel A, Stefan N, Yu Z, Mühlenbruch K, Drogan D, Joost HG, Fritsche A, Häring HU, Hrabě de Angelis M, Peters A, Roden M, Prehn C, Wang-Sattler R, Illig T, Schulze MB, Adamski J, Boeing H, Pischon T. Identification of serum metabolites associated with risk of

type 2 diabetes using a targeted metabolomic approach. *Diabetes*. 2013; 62:639–648.
16. Stančáková A, Civelek M, Saleem NK, Soininen P, Kangas AJ, Cederberg H, Paananen J, Pihlajamäki J, Bonnycastle LL, Morken MA, Boehnke M, Pajukanta P, Lusis AJ, Collins FS, Kuusisto J, Ala-Korpela M, Laakso M. Hyperglycemia and a common variant of GCKR are associated with the levels of eight amino acids in 9,369 Finnish men. *Diabetes*. 2012;61:1895–1902.
17. Mahendran Y, Cederberg H, Vangipurapu J, Kangas AJ, Soininen P, Kuusisto J, Uusitupa M, Ala-Korpela M, Laakso M. Glycerol and fatty acids in serum predict the development of hyperglycemia and type 2 diabetes in Finnish men. *Diabetes Care*. 2013;36:3732–3738.
18. Roberts LD, Gerszten RE. Toward new biomarkers of cardiometabolic diseases. *Cell Metab*. 2013;18:43–50.
19. Rankin NJ, Preiss D, Welsh P, Burgess KEV, Nelson SE, Lawlor DA, Sattar N. The emergence of proton nuclear magnetic resonance metabolomics in the cardiovascular arena as viewed from a clinical perspective. *Atherosclerosis*. 2014;237:287–300.
20. Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikäinen LP, Kangas AJ, Soininen P, Würtz P, Silander K, Dick DM, Rose RJ, Savolainen MJ, Viikari J, Kähönen M, Lehtimäki T, Pietiläinen KH, Inouye M, McCarthy MI, Jula A, Eriksson J, Raitakari OT, Salomaa V, Kaprio J, Järvelin MR, Peltonen L, Perola M, Freimer NB, Ala-Korpela M, Palotie A, Ripatti S. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet*. 2012;44:269–76.
21. Auro K, Joensuu A, Fischer K, Kettunen J, Salo P, Mattsson H, Niironen M, Kaprio J, Eriksson JG, Lehtimäki T, Raitakari O, Jula A, Tiitinen A, Jauhiainen M, Soininen P, Kangas AJ, Kähönen M, Havulinna AS, Ala-Korpela M, Salomaa V, Metspalu A, Perola M. A metabolic view on menopause and aging. *Nat Commun*. 2014;5:4708.
22. Fischer K, Kettunen J, Würtz P, Haller T, Havulinna AS, Kangas AJ, Soininen P, Esko T, Tammesoo ML, Mägi R, Smit S, Palotie A, Ripatti S, Salomaa V, Ala-Korpela M, Perola M, Metspalu A. Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17,345 persons. *PLoS Med*. 2014;11:e1001606.
23. Ala-Korpela M, Kangas AJ, Soininen P. Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. *Genome Med*. 2012;4:36.
24. Soininen P, Kangas AJ, Würtz P, Tukiainen T, Tynkkynen T, Laatikainen R, Järvelin MR, Kähönen M, Lehtimäki T, Viikari J, Raitakari OT, Savolainen MJ, Ala-Korpela M. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst*. 2009;134:1781–1785.
25. Griffin JL, Atherton H, Shockcor J, Atzori L. Metabolomics as a tool for cardiac research. *Nat Rev Cardio*. 2011;8:630–643.
26. Blankenberg S, Zeller T, Saarela O, Havulinna AS, Kee F, Tunstall-Pedoe H, Kuulasmaa K, Yarnell J, Schnabel RB, Wild PS, Münzel TF, Lackner KJ, Tiret L, Evans A, Salomaa V. Contribution of 30 biomarkers to 10-year cardiovascular risk estimation in 2 population cohorts. *Circulation*. 2010;121:2388–2397.
27. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Männisto S, Sundvall J, Jousilahti P, Salomaa V, Valsta L, Puska P. Thirty-five-year trends in cardiovascular risk factors in Finland. *Int J Epidemiol*. 2010;39:504–518.
28. Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Räihä P, Kärjä-Koskenkari P, Mähönen M, Niemelä M, Kuulasmaa K, Palomäki P, Mustonen J, Lehtonen A, Arstila M, Vuorenmaa T, Lehto S, Miettinen H, Torppa J, Tuomilehto J, Kesäniemi YA, Pyörälä K, Salomaa V. The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil*. 2005;12:132–137.
29. Tillin T, Hughes AD, Mayet J, Whincup P, Sattar N, Forouhi NG, McKeigue PM, Chaturvedi N. The relationship between metabolic risk factors and incident cardiovascular disease in Europeans, South Asians, and African Caribbeans. *J Am Coll Cardiol*. 2013;61:1777–1786.
30. Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health*. 2003;57:134–140.
31. Kujala UM, Mäkinen VP, Heinonen I, Soininen P, Kangas AJ, Leskinen TH, Rahkila P, Würtz P, Kovanen V, Cheng S, Sipilä S, Hirvensalo M, Telama R, Tammelin T, Savolainen MJ, Pouta A, O'Reilly PF, Mäntyselkä P, Viikari J, Kähönen M, Lehtimäki T, Elliott P, Vanhala MJ, Raitakari OT, Järvelin MR, Kaprio J, Kainulainen H, Ala-Korpela M. Long-term leisure-time physical activity and serum metabolome. *Circulation*. 2012;127:340–348.
32. Würtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, Tynkkynen T, Soininen P, Havulinna AS, Kaakkinen M, Viikari JS, Savolainen M, Kähönen M, Lehtimäki T, Männistö S, Blankenberg S, Zeller T, Laitinen J, Pouta A, Mäntyselkä P, Vanhala M, Elliott P, Pietiläinen KH, Ripatti S, Salomaa V, Raitakari OT, Järvelin MR, Davey Smith D, Ala-Korpela M. Metabolic signatures of adiposity: Mendelian randomization analysis and effects of weight change. *PLoS Med*. 2014;11:e1001765.
33. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, Walter K, Menni C, Chen L, Vasquez L, Valdes AM, Hyde CL, Wang V, Ziemek D, Roberts P, Xi L, Grundberg E; Multiple Tissue Human Expression Resource (MuTHER) Consortium, Waldenberger M, Richards JB, Mohney RP, Milburn MV, John SL, Trimmer J, Theis FJ, Overington JP, Suhre K, Brosnan MJ, Gieger C, Kastenmüller G, Spector TD, Soranzo N. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014;46:543–550.
34. Kaikkonen JE, Kresanov P, Ahotupa M, Jula A, Mikkilä V, Viikari JS, Kähönen M, Lehtimäki T, Raitakari OT. High serum n6 fatty acid proportion is associated with lowered LDL oxidation and inflammation: the Cardiovascular Risk in Young Finns Study. *Free Radic Res*. 2014;48:420–426.
35. Antolini L, Nam BH, D'Agostico RB. Inference on correlated discrimination measures in survival analysis: a nonparametric approach. *Commun Statist Theory Meth*. 2004;33:2117–2135.
36. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27:157–172.
37. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30:11–21.
38. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, Guiducci C, Perola M, Jula A, Sinisalo J, Lokki ML, Nieminen MS, Melander O, Salomaa V, Peltonen L, Kathiresan S. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet*. 2010;376:1393–1400.
39. Superko HR, Superko SM, Nasir K, Agatston A, Garrett BC. Omega-3 fatty acid blood levels: clinical significance and controversy. *Circulation*. 2013;128:2154–2161.
40. Chowdhury R, Warnakula S, Kunutsor S, Crowe F, Ward HA, Johnson L, Franco OH, Butterworth AS, Forouhi NG, Thompson SG, Khaw KT, Mozaffarian D, Danesh J, Di Angelantonio E. Association of dietary, circulating, and supplement fatty acids with coronary risk: a systematic review and meta-analysis. *Ann Intern Med*. 2014;160:398–406.
41. Wu JH, Lemaitre RN, King IB, Song X, Psaty BM, Siscovick DS, Mozaffarian D. Circulating omega-6 polyunsaturated Fatty acids and total and cause-specific mortality: the cardiovascular health study. *Circulation*. 2014;130:1245–1253.
42. Magnusson M, Lewis GD, Ericson U, Orho-Melander M, Hedblad B, Engström G, Ostling G, Clish C, Wang TJ, Gerszten RE, Melander O. A diabetes-predictive amino acid score and future cardiovascular disease. *Eur Heart J*. 2013;34:1982–1989.
43. Murr C, Grammer TB, Meinitzer A, Kleber ME, März W, Fuchs D. Immune activation and inflammation in patients with cardiovascular disease are associated with higher phenylalanine to tyrosine ratios: the Ludwigshafen Risk and Cardiovascular Health Study. *J Amino Acids*. 2014;2014:783730.
44. Kris-Etherton PM. AHA Science Advisory. Monounsaturated fatty acids and risk of cardiovascular disease. American Heart Association. Nutrition Committee. *Circulation*. 1999;100:1253–1258.
45. Liu J, Cinar R, Xiong K, Godlewski G, Jourdan T, Lin Y, Ntambi JM, Kunos G. Monounsaturated fatty acids generated via stearoyl CoA desaturase-1 are endogenous inhibitors of fatty acid amide hydrolase. *Proc Natl Acad Sci U S A*. 2013;110:18832–18837.
46. Rizos EC, Ntzani EE, Bika E, Kostapanos MS, Elisaf MS. Association between omega-3 fatty acid supplementation and risk of major cardiovascular disease events: a systematic review and meta-analysis. *JAMA*. 2012;308:1024–1033.

47. Risk and Prevention Study Collaborative Group. n-3 fatty acids in patients with multiple cardiovascular risk factors. *N Engl J Med*. 2013;368:1800–1808.
48. Mozaffarian D, Micha R, Wallace S. Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: a systematic review and meta-analysis of randomized controlled trials. *PLoS Med*. 2010;7:e1000252.
49. Harris WS, Mozaffarian D, Rimm E, Kris-Etherton P, Rudel LL, Appel LJ, Engler MM, Engler MB, Sacks F. Omega-6 fatty acids and risk for cardiovascular disease. *Circulation*. 2009;119:902–907.

CLINICAL PERSPECTIVE

Metabolite profiling allows the capture of an individual's metabolic status across multiple molecular pathways. Technological advances have made metabolite profiling of blood samples feasible and affordable for large epidemiological cohorts and biobanks. The quantification of amino acids, fatty acid composition, glycolysis metabolites, and lipoprotein subclasses is hereby possible simultaneously along with the standard clinical lipids. The examination of all these metabolic measures in relation to cardiovascular disease incidence in large study collections can identify novel biomarkers and improve the understanding of the underlying mechanisms. The wide metabolic coverage also has the potential to improve the prediction of cardiovascular risk above established risk factors. We conducted high-throughput nuclear magnetic resonance profiling in 3 population-based cohorts of 13 441 individuals with 12 to 23 years of follow-up. Thirty-three of the 68 metabolic measures analyzed were associated with incident cardiovascular disease. In analyses adjusted for standard lipids, we identified circulating phenylalanine and monounsaturated fatty acids as novel biomarkers of higher cardiovascular risk. Omega-3 and omega-6 fatty acids were associated with lower cardiovascular disease risk. In combination, the 4 biomarkers improved cardiovascular risk assessment, in particular for persons classified with intermediate risk based on established risk factors. Thus, high-throughput metabolite profiling is a powerful method to uncover novel cardiovascular biomarkers and enhance risk prediction. Further studies with more detailed patient stratification and more specific end points are needed to evaluate the prospects for clinical use. Because the metabolite profiling platform quantifies amino acids and fatty acids simultaneously with the standard lipids, these novel biomarkers could be implemented to augment risk prediction without a need for additional clinical chemistry.

SUPPLEMENTAL MATERIAL

Würtz, Havulinna, Soininen, et al. Metabolite Profiling and Cardiovascular Event Risk:

A Prospective Study of Three Population-based Cohorts.

Supplemental Methods.

Study populations

FINRISK

The FINRISK 1997 study was conducted to monitor the health of the Finnish population among persons aged 25–74 at recruitment.[1] The study was approved by the ethical committee of the National Public Health Institute, Helsinki, Finland. In total, 8444 individuals were recruited to represent the working age population of five study locations across Finland. Standard laboratory measures including blood pressure and body mass index were measured, and participants completed questionnaires on smoking status and medication. Information on medication was complemented by national reimbursement records. Alcohol usage was assessed by questionnaires on beverage-specific consumption in the past week. Geographical region was defined as Eastern versus Western Finland.[1] Family history of CVD was defined as myocardial infarction or stroke in a first degree relative prior to the age of 60. Blood samples were stored at –70°C for later biomarker analyses. Samples were semi-fasting: participants were asked not to eat 4 hours prior to giving blood. High-sensitivity C-reactive protein, plasma glucose and insulin were measured by standard assays.[1] NMR metabolomics was measured during 2012 for 7610 individuals with serum samples stored. Pregnant women (n=75), persons missing risk factor information (n=8), and individuals with prevalent CVD defined as for the incident events (n=263, tracked by registries back to 1970) were excluded, leaving 7256 individuals for the analyses of incident CVD.

Tracking of incident CVD events during follow-up (spring 1997 through Dec 2011) was enabled by data obtained from the Finnish National Hospital Discharge Register and the National Causes-of-Death Register. These registers cover all cardiovascular events that have led either to hospitalization or death in Finland. The cardiovascular diagnoses in these registers have been validated.[2,3] The cardiovascular outcomes were linked to study subjects using their social security ID, which is assigned to every permanent resident of Finland. The endpoint was the first occurrence of a major cardiovascular event during follow-up, which includes fatal or nonfatal occurrence of myocardial infarction, ischemic stroke, cardiac revascularization (coronary artery bypass graft surgery or percutaneous transluminal coronary angioplasty), or unstable angina. Consistency analyses of the biomarker associations were conducted with cardiovascular mortality as outcome. In the FINRISK study, sensitivity analyses of the biomarker associations were also conducted with the exclusion of CVD events occuring during the first year of follow-up, and omission of individuals using lipid-lowering medication (n=185) at baseline (Supplemental Figure S4). Furthermore, the biomarker associations were tested with additional covariate adjustment for fasting duration, family history of CVD, habitual alcohol intake, body mass index, C-reactive protein, plasma glucose and insulin as well as serum triglycerides (Supplemental Figure S4).

In addition to NMR-based metabolite profiling, LC-MS-based metabolite profiling was performed for a case-cohort subset of the FINRISK study (n=679 with 305 incident CVD events during 15-year follow-up). The LC-MS profiling was measured at the Helmholtz Institute, Munich, Germany, with metabolite quantification in units of ion count obtained by the Metabolon platform.[4] Metabolites measured by both NMR and LC-MS, and nominally significant in the meta-analysis, were verified for consistency of biomarker association with CVD in the case-cohort subset (Figure 6). Due to the limited number of cardiovascular events in the subset, associations were also tested with adjustment for age and sex only (Supplemental Figure 10), in addition to the fully adjusted model. The correspondence between the significant

biomarkers quantified by both LC-MS and NMR was illustrated by scatter plots and Pearson's correlation coefficients estimated (Supplemental Figure 11).

SABRE: Southall and Brent Revisited

The SABRE (Southall and Brent Revisited) study examined 4857 individuals in a tri-ethnic community-based cohort from North and West London. All participants gave written informed consent. The study protocols were approved by the University College London and St. Mary's Hospital research ethics committees. Details of the cohort have been published.[4,6] Briefly, participants aged 40 to 69 at baseline (1988 through 1991) were selected randomly from 5-year age- and sex-stratified primary care physician lists. Participants included Europeans (48%), Indian Asians (47%) and a small fraction of African Caribbeans (5%). Statistical analyses in the SABRE study were adjusted for ethnicity. Participants attended a baseline clinic after an overnight fast. Fasting blood samples were drawn for Southall participants, and stored at −80°C for biomarker measurements. Study participants were invited for a follow-up clinic visit during 2008 through 2011.[4] NMR metabolomics was measured during 2013 for 3309 baseline serum samples with available stored aliquots. Fasting serum samples at follow-up were available for NMR metabolomics from 1419 individuals, of whom 916 were matched to baseline samples with NMR metabolomics data. The tracking of the metabolite biomarker concentrations within the same individuals 20 years apart was evaluated by partial correlations adjusted for age and sex (Supplemental Table 5).

Baseline metabolite data and complete data on established risk factors were available for 3236 individuals. Among these, 227 had prevalent CVD and therefore omitted in the present analysis. Prospective data on CVD incidence during 1988 through 2011 were available for 2622 individuals through linkage with general practitioner records as well as Hospital Episode Statistics. General practitioner data were reviewed independently by two senior physicians blinded to participant identity. A CVD event was identified if both physicians agreed on definite diagnosis of myocardial infarction, acute coronary syndrome, exercise-test confirmed angina, coronary interventions (coronary artery bypass graft, percutaneous transluminal coronary angioplasty, stenting), or stroke as described previously.[4] Follow-up coverage was complete for fatal events. However, hospitalization data tracking was incomplete in the period from 1988 through 1997 since the UK guidelines did not emphasize the need for hospital admission for all CVD cases at the period; classification of non-fatal event in this period relies on the adjudicated review of general practitioner records.

BWHHS: British Women's Heart and Health Study

The British Women's Heart and Health Study (BWHHS) recruited 4286 females between 1999 and 2001, who were randomly selected from 23 British towns and were aged between 60 and 79 years at baseline assessment.[7] Local ethics committee approvals were obtained for the BWHHS. All women provided informed written consent. Baseline data collection took place between April 1999 and March 2001. Methods used at baseline have been described in detail previously.[7] Briefly, clinical examination included blood pressure and anthropometric measurements, the collection of blood samples for biochemical analysis, the completion of medical questionnaires, and detailed reviews of medical records. Blood samples were drawn after a minimum 8-hour fast and stored at −80°C. NMR metabolomics was measured for 3822 serum samples available during 2012. Metabolite profiling and complete data on established risk factors and prospective follow-up for CVD through 2012 were available for 3719 study

participants. Of these, 156 women had prevalent CVD and were therefore omitted from the present analysis. CVD outcome is a composite of the following: myocardial infarction, ischemic or hemorrhagic stroke, unstable angina, death due to CVD or revascularization procedures including percutaneous transluminal coronary angioplasty or coronary artery bypass surgery, with information on these obtained from reviewing medical records. The prospective follow-up included all events from beginning of the study to October 2012. Only events obtained from medical records and validated by the BWHHS investigators were included.

Additional population studies analyzed for biomarker validation

Framingham Offspring Cohort

The Framingham Heart Study Offspring Cohort was formed in 1971 with the enrollment of 5124 individuals in a community-based longitudinal cohort study.[8] Study protocols were approved by the Institutional Review boards of Boston University Medical Center, Massachusetts General Hospital. All study participants provided written informed consent. Of the 2413 attendees who were free of CVD at the fifth examination cycle (1991–1995), 2289 had metabolite and lipid profiling performed from fasting plasma samples by LC-MS. The Framingham metabolomics platform differs from the commercial MS profiling platform used in the FINRISK case-cohort study; the methodology have been described previously.[9,10] All study participants also underwent a standardized physical examination and routine laboratory tests in addition to longitudinal surveillance for CVD at every Framingham offspring visit during 12 year follow-up (myocardial infarction, unstable angina pectoris, ischemic stroke, death due to CVD, or revascularization procedures including percutaneous transluminal coronary angioplasty, coronary artery bypass surgery, carotid surgery and femoral surgery).

Since the mass spectrometry platform used in the Framingham Offspring Study measures slightly different lipid species than the fatty acids quantified by NMR, we used the levels of lysophosphatidylcholine and cholesterol-ester DHA (22:6) for comparison with DHA, and lyso-phosphatidylcholine and cholesterol-ester linoleic acid (18:2) for comparison with total omega-6 fatty acids quantified by NMR. Composite measures of MUFA%, omega-3 fatty acids, PUFA and glycoprotein acetyls were not determined by mass spectrometry. All analyses of mass spectrometry data were adjusted for batch, in addition to the established cardiovascular risk factors.

Cardiovascular Risk in Young Finns Study

The Cardiovascular Risk in Young Finns Study is an ongoing population-based study on atherosclerosis risk factors from childhood to adulthood.[11] All participants gave written informed consent and the study was approved by the local Ethics Committees. In the first survey (1980), children and adolescents aged 3–18 years were chosen randomly from the national population register from all five Finnish university cities with medical schools and their rural surroundings (n=3596). Study participants were invited for adulthood follow-up studies in 2001 (n=2283) and 2007 (n=2204). Attendees were representative of the original cohort.[11,12] All participants underwent physical examination and assessment of standard risk factors. NMR metabolomics was measured by the high-throughput platform[12] from fasting serum samples from both adulthood follow-up surveys.[13] Tracking of metabolite biomarker concentrations over time within the same individuals were assessed by partial correlations adjusted for age and sex. Due to the low number of CVD events (<20), the Cardiovascular Risk in Young Finns Study was not included in the meta-analysis of biomarkers for CVD. Instead, the study was used

to confirm the quantification of fatty acid biomarkers, and examine the associations of the circulating biomarker concentrations with corresponding dietary data.

Fatty acid concentrations, relative to total fatty acids, were determined by gas chromatography (GC) using a flame ionization detector (Varian CP-3800; Varian Inc., Walnut Creek, CA, USA) for the serum samples from the 2001 field survey as described previously.[14,15] Composite measures of omega-6 fatty acids, MUFA, and PUFA were derived.[15] Absolute concentrations in mmol/L were calculated by scaling the fatty acid ratios to the amount of total fatty acids quantified by NMR. The correspondence between fatty acid concentrations quantified by GC and NMR was assessed by scatter plots and Pearson's correlation coefficients (Supplemental Figure 12).

Dietary data were assessed by 48h recall based on interviews with trained dietitians for a random half of the study population in the 2001 field survey (n=960),[16] as well as based on food-frequency questionnaires in the whole study population in the 2007 field survey (n=1699).[17] For the latter, data on habitual dietary intake over the past year were assessed with a 131-item food-frequency questionnaire. Standardized associations between dietary intake [g for absolute levels and g/total fat intake for relative levels] with the corresponding serum levels were analyzed using linear regression models adjusted for age, sex, body mass index and total energy intake (Supplemental Figure 13).

References

1. Blankenberg S, Zeller T, Saarela O, Havulinna AS, Kee F, Tunstall-Pedoe H, Kuulasmaa K, Yarnell J, Schnabel RB, Wild PS, Münzel TF, Lackner KJ, Tiret L, Evans A, Salomaa V. Contribution of 30 biomarkers to 10-year cardiovascular risk estimation in 2 population cohorts: the MONICA, risk, genetics, archiving, and monograph (MORGAM) biomarker project. *Circulation*. 2010;121:2388–2397.
2. Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Räihä P, Kärjä-Koskenkari P, Mähönen M, Niemelä M, Kuulasmaa K, Palomäki P, Mustonen J, Lehtonen A, Arstila M, Vuorenmaa T, Lehto S, Miettinen H, Torppa J, Tuomilehto J, Kesäniemi YA, Pyörälä K, Salomaa V. The validity of the Finnish Hospital Discharge Register and Causes of Death Register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil*. 2005;12:132–137.
3. Tolonen H, Salomaa V, Torppa J, Sivenius J, Immonen-Räihä P, Lehtonen A; FINSTROKE register. The validation of the Finnish Hospital Discharge Register and Causes of Death Register data on stroke diagnoses. *Eur J Cardiovasc Prev Rehabil*. 2007;14:380–385.
4. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, Walter K, Menni C, Chen L, Vasquez L, Valdes AM, Hyde CL, Wang V, Ziemek D, Roberts P, Xi L, Grundberg E; Multiple Tissue Human Expression Resource (MuTHER) Consortium, Waldenberger M, Richards JB, Mohney RP, Milburn MV, John SL, Trimmer J, Theis FJ, Overington JP, Suhre K, Brosnan MJ, Gieger C, Kastenmüller G, Spector TD, Soranzo N. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014;46:543–550.
5. Tillin T, Hughes AD, Mayet J, Whincup P, Sattar N, Forouhi NG, McKeigue PM, Chaturvedi N. The relationship between metabolic riskfactors and incident cardiovascular disease in Europeans, South Asians, and African Caribbeans. *J Am Coll Cardiol*. 2013;61:1777–1786.

6. Tillin T, Forouhi NG, McKeigue PM, Chaturvedi N. Southall And Brent REvisited: cohort profile of SABRE, a UK population-based comparison of cardiovascular disease and diabetes in people of European, Indian Asian and African Caribbean origins. *Int J Epidemiol*. 2012;41:33–42.
7. Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health*. 2003; 57:134–140.
8. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. *Lancet*. 2014;383:999–1008.
9. Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, Lewis GD, Fox CS, Jacques PF, Fernandez C, O'Donnell CJ, Carr SA, Mootha VK, Florez JC, Souza A, Melander O, Clish CB, Gerszten RE. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011; 17:448–453.
10. Rhee EP, Cheng S, Larson MG, Walford GA, Lewis GD, McCabe E, Yang E, Farrell L, Fox CS, O'Donnell CJ, Carr SA, Vasan RS, Florez JC, Clish CB, Wang TJ, Gerszten RE. Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J Clin Invest*. 2011;121:1402–1411.
11. Raitakari OT, Juonala M, Rönnemaa T, Keltikangas-Järvinen L, Räsänen L, Pietikäinen M, Hutri-Kähönen N, Taittonen L, Jokinen E, Marniemi J, Jula A, Telama R, Kähönen M, Lehtimäki T, Akerblom HK, Viikari JS. Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol*. 2008;37:1220–1226.
12. Soininen P, Kangas AJ, Würtz P, Suna T, Ala-Korpela M. Quantitative serum nuclear magnetic resonance metabolomics in cardiovascular epidemiology and genetics. Circ Cardiov Genet. 2014;In press.
13. Würtz P, Raiko JR, Magnussen CG, Soininen P, Kangas AJ, Tynkkynen T, Thomson R, Laatikainen R, Savolainen MJ, Laurikka J, Kuukasjärvi P, Tarkka M, Karhunen PJ, Jula A, Viikari JS, Kähönen M, Lehtimäki T, Juonala M, Ala-Korpela M, Raitakari OT. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J*. 2012;33:2307–2316.
14. Jula A, Marniemi J, Rönnemaa T, Virtanen A, Huupponen R. Effects of diet and simvastatin on fatty acid composition in hypercholesterolemic men: a randomized controlled trial. *Arterioscler Thromb Vasc Biol*. 2005;25:1952–1959.
15. Kaikkonen JE, Kresanov P, Ahotupa M, Jula A, Mikkilä V, Viikari JS, Kähönen M, Lehtimäki T, Raitakari OT. High serum n6 fatty acid proportion is associated with lowered LDL oxidation and inflammation: the Cardiovascular Risk in Young Finns Study. *Free Radic Res*. 2014;48:420–426.
16. Mikkilä V, Räsänen L, Laaksonen MM, Juonala M, Viikari J, Pietinen P, Raitakari OT. Long-term dietary patterns and carotid artery intima media thickness: the Cardiovascular Risk in Young Finns Study. *Br J Nutr*. 2009;102:1507–1512.
17. Skilton MR, Mikkilä V, Würtz P, Ala-Korpela M, Sim KA, Soininen P, Kangas AJ, Viikari JS, Juonala M, Laitinen T, Lehtimäki T, Taittonen L, Kähönen M, Celermajer DS, Raitakari OT. Fetal growth, omega-3 (n-3) fatty acids, and progression of subclinical atherosclerosis: preventing fetal origins of disease? The Cardiovascular Risk in Young Finns Study. *Am J Clin Nutr*. 2013;97:58–65.

Supplemental Table 1. Metabolite concentrations in the FINRISK study.

Amino acids	Mean (SD)	Hazard ratio per 1-SD metabolite concentration adjusted for age and sex
Alanine [µmol/L]	430 (67)	1.19 [1.11–1.28] $P=2\times10^{-6}$
Glutamine [µmol/L]	470 (70)	0.86 [0.81–0.92] $P=2\times10^{-5}$
Glycine [µmol/L]	320 (62)	1.06 [0.98–1.14] P=0.17
Histidine [µmol/L]	64 (9.6)	0.94 [0.87–1.01] P=0.09
Branched-chain amino acids		
Isoleucine [µmol/L]	58 (19)	1.17 [1.09–1.26] $P=4\times10^{-5}$
Leucine [µmol/L]	88 (21)	1.13 [1.04–1.21] P=0.002
Valine [µmol/L]	210 (43)	1.11 [1.02–1.19] P=0.01
Aromatic amino acids		
Phenylalanine [µmol/L]	85 (14)	1.40 [1.30–1.51] $P=1\times10^{-17}$
Tyrosine [µmol/L]	51 (13)	1.17 [1.09–1.27] $P=4\times10^{-5}$
Glycolysis related metabolites		
Glucose [mmol/L]	4.6 (1.1)	1.15 [1.09–1.21] $P=7\times10^{-8}$
Lactate [mmol/L]	1.3 (0.33)	1.24 [1.16–1.33] $P=1\times10^{-9}$
Pyruvate [µmol/L]	75 (22)	1.23 [1.16–1.32] $P=5\times10^{-10}$
Citrate [µmol/L]	110 (19)	0.91 [0.85–0.97] P=0.002
Glycerol [µmol/L]	140 (65)	1.19 [1.11–1.28] $P=2\times10^{-6}$
Ketone bodies		
Acetate [µmol/L]	54 (21)	0.87 [0.80–0.93] P=0.0002
Acetoacetate [µmol/L]	81 (53)	0.88 [0.82–0.94] P=0.0003
beta-hydroxybutyrate [µmol/L]	230 (160)	0.85 [0.80–0.92] $P=1\times10^{-5}$
Fatty acids		
Total fatty acids [mmol/L]	12 (3.7)	1.20 [1.12–1.29] $P=5\times10^{-7}$
Omega-3 fatty acids [mmol/L]	0.52 (0.21)	1.00 [0.93–1.08] P=0.93
Docosahexaenoic acid (DHA) [mmol/L]	0.21 (0.090)	0.96 [0.89–1.03] P=0.22
Omega-6 fatty acids [mmol/L]	4.1 (0.91)	1.09 [1.01–1.17] P=0.03
Linoleic acid [mmol/L]	3.5 (0.76)	1.08 [1.01–1.17] P=0.03
Polyunsaturated fatty acids [mmol/L]	4.6 (1.0)	1.08 [1.00–1.17] P=0.04
Monounsaturated fatty acids [mmol/L]	3.6 (1.5)	1.28 [1.19–1.37] $P=3\times10^{-12}$
Saturated fatty acids [mmol/L]	4.2 (1.5)	1.17 [1.09–1.26] $P=8\times10^{-6}$
Fatty acid chain length	18 (0.34)	0.89 [0.83–0.97] P=0.004
Fatty acid ratios, relative to total fatty acids		
Omega-3 fatty acids (%)	4.2 (1.2)	0.85 [0.80–0.91] $P=8\times10^{-6}$
Docosahexaenoic acid (%)	1.7 (0.65)	0.85 [0.79–0.91] $P=2\times10^{-6}$
Omega-6 fatty acids (%)	34 (4.3)	0.82 [0.77–0.88] $P=2\times10^{-8}$
Linoleic Acid (%)	28 (4.2)	0.84 [0.78–0.90] $P=6\times10^{-7}$
Polyunsaturated fatty acids (%)	38 (4.4)	0.80 [0.75–0.86] $P=4\times10^{-11}$
Monounsaturated fatty acids (%)	29 (3.7)	1.34 [1.24–1.44] $P=1\times10^{-14}$
Saturated fatty acids (%)	33 (2.8)	1.02 [0.95–1.10] P=0.54
Unsaturation degree, double bonds per fatty acid	1.3 (0.10)	0.82 [0.77–0.88] $P=6\times10^{-9}$
Fluid balance		
Creatinine [µmol/L]	58 (13)	1.05 [0.97–1.14] P=0.26
Albumin [standardized concentration units]	96 (6.2)	0.95 [0.88–1.02] P=0.14
Inflammation		
Glycoprotein acetyls [cu]	1.40 (0.24)	1.41 [1.31–1.51] $P=5\times10^{-22}$
Apolipoproteins		
Apolipoprotein B [g/L]	1.0 (0.26)	1.26 [1.16–1.36] $P=3\times10^{-8}$
Apolipoprotein A1 [g/L]	1.7 (0.23)	0.85 [0.79–0.91] $P=5\times10^{-6}$
Lipids		

Triglycerides [mmol/L]	1.3 (0.59)	1.33 [1.22–1.43] $P=2\times10^{-12}$
Phosphoglycerides [mmol/L]	1.7 (0.40)	1.07 [0.99–1.15] P=0.10
Cholines [mmol/L]	1.9 (0.42)	1.05 [0.97–1.13] P=0.22
Sphingomyelin [mmol/L]	0.29 (0.067)	1.02 [0.94–1.10] P=0.64
Total lipid concentrations in lipoprotein subclasses		
Extremely large VLDL [mmol/L]	0.033 (0.034)	1.18 [1.12–1.25] $P=2\times10^{-9}$
Very large VLDL [mmol/L]	0.064 (0.067)	1.19 [1.12–1.26] $P=6\times10^{-9}$
Large VLDL [mmol/L]	0.19 (0.19)	1.20 [1.13–1.27] $P=1\times10^{-8}$
Medium VLDL [mmol/L]	0.40 (0.25)	1.25 [1.17–1.34] $P=1\times10^{-10}$
Small VLDL [mmol/L]	0.68 (0.25)	1.28 [1.19–1.38] $P=4\times10^{-11}$
Very small VLDL [mmol/L]	0.64 (0.22)	1.25 [1.17–1.35] $P=2\times10^{-10}$
IDL [mmol/L]	1.4 (0.37)	1.27 [1.17–1.38] $P=1\times10^{-8}$
Large LDL [mmol/L]	1.6 (0.42)	1.23 [1.13–1.34] $P=1\times10^{-6}$
Medium LDL [mmol/L]	0.95 (0.26)	1.22 [1.12–1.32] $P=6\times10^{-6}$
Small LDL [mmol/L]	0.63 (0.20)	1.23 [1.13–1.34] $P=8\times10^{-7}$
Very large HDL [mmol/L]	0.79 (0.27)	0.94 [0.88–1.01] P=0.12
Large HDL [mmol/L]	0.91 (0.36)	0.73 [0.68–0.79] $P=8\times10^{-15}$
Medium HDL [mmol/L]	0.70 (0.21)	0.86 [0.80–0.92] $P=3\times10^{-5}$
Small HDL [mmol/L]	1.0 (0.14)	0.92 [0.86–0.99] P=0.02
Lipoprotein particle size		
VLDL diameter [nm]	36 (1.1)	1.18 [1.09–1.26] $P=1\times10^{-5}$
LDL diameter [nm]	23 (0.17)	1.09 [1.01–1.17] P=0.03
HDL diameter [nm]	10 (0.26)	0.86 [0.79–0.92] $P=5\times10^{-5}$
Cholesterol		
Total cholesterol [mmol/L]	5.5 (1.1)	1.11 [1.02–1.20] P=0.01
VLDL cholesterol [mmol/L]	0.86 (0.30)	1.27 [1.17–1.37] $P=8\times10^{-9}$
IDL cholesterol [mmol/L]	0.86 (0.22)	1.22 [1.13–1.33] $P=1\times10^{-6}$
LDL cholesterol [mmol/L]	2.2 (0.61)	1.20 [1.11–1.31] $P=2\times10^{-5}$
HDL cholesterol [mmol/L]	1.6 (0.36)	0.77 [0.72–0.83] $P=1\times10^{-13}$
HDL_2 cholesterol [mmol/L]	1.1 (0.36)	0.75 [0.70–0.81] $P=4\times10^{-15}$
HDL_3 cholesterol [mmol/L]	0.53 (0.047)	1.07 [1.00–1.14] P=0.06
Cholesterol esterification [%]	77 (3.0)	0.95 [0.89–1.02] P=0.19

Mean (SD) concentrations and hazard ratios [95% confidence intervals] for first incident cardiovascular event during 15-year follow-up in the FINRISK 1997 study. Hazard ratios are per 1-SD increment in log-transformed metabolite concentration. Hazard ratios were examined separately for each metabolite measure by Cox models adjusted for sex and age. All 68 metabolite measures were quantified using a high-throughput serum NMR metabolomics platform.[12] Quantification of lipoprotein subclasses was calibrated against high performance liquid chromatography on an external set of samples. The 14 lipoprotein subclass sizes were defined as follows: extremely large VLDL with particle diameters from 75 nm upwards and a possible contribution of chylomicrons, five VLDL subclasses (average particle diameters of 64.0 nm, 53.6 nm, 44.5 nm, 36.8 nm, and 31.3 nm), IDL (28.6 nm), three LDL subclasses (25.5 nm, 23.0 nm, and 18.7 nm), and four HDL subclasses (14.3 nm, 12.1 nm, 10.9 nm, and 8.7 nm). The mean size for VLDL, LDL and HDL particles was calculated by weighting the corresponding subclass diameters with their particle concentrations. For fatty acids, only the *cis* configuration was quantified since the *trans* fatty acids are below detection limit.

Supplemental Table 2. Biomarker associations in absolute concentration units and for upper vs. lower quartiles.

Biomarker measure, unit increment	Hazard ratios in absolute concentration units			Hazard ratio for 4th vs. 1st quartile		
	HR	95% CI	P-value	HR	95% CI	P-value
Phenylalanine, per 50 µmol/L	1.72	1.46–2.03	1×10^{-10}	1.50	1.28–1.75	3×10^{-7}
MUFA ratio to total fatty acids, per 10%	1.49	1.30—1.70	6×10^{-9}	1.45	1.25–1.68	1×10^{-6}
Omega-6 fatty acids, per 5 mmol/L	0.49	0.36-0.68	3×10^{-5}	0.71	0.61–0.84	7×10^{-5}
PUFA, per 5 mmol/L	0.49	0.35-0.68	1×10^{-5}	0.73	0.61-0.86	0.0003
DHA, per 0.5 mmol/L	0.56	0.40–0.79	0.001	0.80	0.69-0.93	0.004

Hazard ratios for first cardiovascular event with metabolites in absolute concentration units (without log-transform and SD-scaling prior to meta-analysis) and as quartiles. Results are from meta-analysis across the three cohorts (n=13441, 1741 CVD events during 12-23 years follow-up). Analyses were adjusted for age, sex, blood pressure, smoking, geographical region, diabetes, cardiovascular medication as well as total and HDL cholesterol.

Supplemental Table 3. Risk score coefficients derived in the FINRISK study for prediction of cardiovascular event risk.

Variable	Risk prediction model without biomarkers				Risk prediction model with biomarkers			
	Coef	HR	95% CI	P-value	Coef	HR	95% CI	P-value
Female gender	-0.567	0.57	0.48–0.67	2×10^{-11}	-0.613	0.54	0.46–0.64	7×10^{-13}
Systolic blood pressure [SD]	0.159	1.17	1.09–1.26	1×10^{-5}	0.145	1.16	1.08–1.24	7×10^{-5}
Current smoking (yes/no)	0.475	1.61	1.36–1.90	2×10^{-8}	0.379	1.46	1.23–1.73	1×10^{-5}
Prevalent diabetes (yes/no)	0.653	1.92	1.57–2.35	1×10^{-10}	0.665	1.94	1.59–2.38	9×10^{-11}
Antihypertensive medication (yes/no)	0.109	1.12	0.94–1.32	0.21	0.103	1.11	0.93–1.32	0.24
Lipid lowering medication (yes/no)	0.308	1.36	1.02–1.81	0.03	0.295	1.34	1.01–1.79	0.04
Total cholesterol [SD]	0.145	1.16	1.08–1.24	8×10^{-5}	0.247	1.28	1.17–1.41	2×10^{-7}
HDL cholesterol [SD]	-0.366	0.69	0.64–0.76	$<2\times10^{-16}$	-0.315	0.73	0.66–0.80	5×10^{-11}
log Phenylalanine [SD]	—	—	—	—	0.225	1.25	1.15–1.36	9×10^{-8}
log MUFA%, ratio to total fatty acids [SD]	—	—	—	—	-0.145	1.11	1.02–1.20	0.02
log Omega-6 fatty acids [SD]	—	—	—	—	-0.137	0.87	0.79–0.96	0.004
log DHA [SD]	—	—	—	—	0.100	0.86	0.80–0.94	0.0004

Regression coefficients (coef) for the linear predictors of the risk score for first cardiovascular event as derived in the FINRISK study. The corresponding hazard ratios (HR, exponential function *e* of the regression coefficients), confidence intervals and P-values are also shown for each variable. Two risk prediction scores were derived: a reference risk score composed of established risk factors, and a risk score further including the four metabolite biomarkers giving rise to the best model fit in the FINRISK study: phenylalanine, the ratio of MUFA to total fatty acids, omega-6 fatty acids, and DHA. The baseline hazard term was derived within each validation cohort due to differences in hazard depending on age and geographical regions. The two risk prediction scores were then used to estimate the absolute risk for cardiovascular events in the SABRE and BWHHS validation studies, where the predictive utility from including the four metabolites in the biomarker risk score was examined.

Supplemental Table 4. Reclassification tables for people at intermediate risk in SABRE and BWHHS.

SABRE: Reclassification for participants with 5–10% risk based on the reference risk score, n=650

Individuals who had a cardiovascular event during 15-year follow-up		**Risk Score with Established Risk Factors and Biomarkers**		
		<5% risk	5–10% risk	>10% risk
Reference Risk Score	5–10% risk	3 (7%)	29 (66%)	12 (27%)
Reclassification improvement for participants who experienced an event: 20.5±8.8%; P=0.02				
Individuals who did not have a cardiovascular event during 15-year follow-up		**Risk Score with Established Risk Factors and Biomarkers**		
		<5% risk	5–10% risk	≥10% risk
Reference Risk Score	5–10% risk	133 (21%)	380 (63%)	93 (15%)
Reclassification improvement for participants who did not have an event: 6.6±2.5%; P=0.008				
Clinical net reclassification improvement: 27.1±9.1%; P=0.003				

BWHHS: Reclassification for participants with 5–10% risk based on the reference risk score, n=1286

Individuals who had a cardiovascular event during 12-year follow-up		**Risk Score with Established Risk Factors and Biomarkers**		
		<5% risk	5–10% risk	>10% risk
Reference Risk Score	5–10% risk	9 (10%)	65 (71%)	18 (19%)
Reclassification improvement for participants who experienced an event: 9.8±5.6%; P=0.08				
Individuals who did not have a cardiovascular event during 12-year follow-up		**Risk Score with Established Risk Factors and Biomarkers**		
		<5% risk	5–10% risk	≥10% risk
Reference Risk Score	5–10% risk	236 (20%)	790 (66%)	168 (14%)
Reclassification improvement for participants who did not have an event: 5.7±1.7%; P=0.0007				
Clinical net reclassification improvement: 15.5±5.9%; P=0.009				

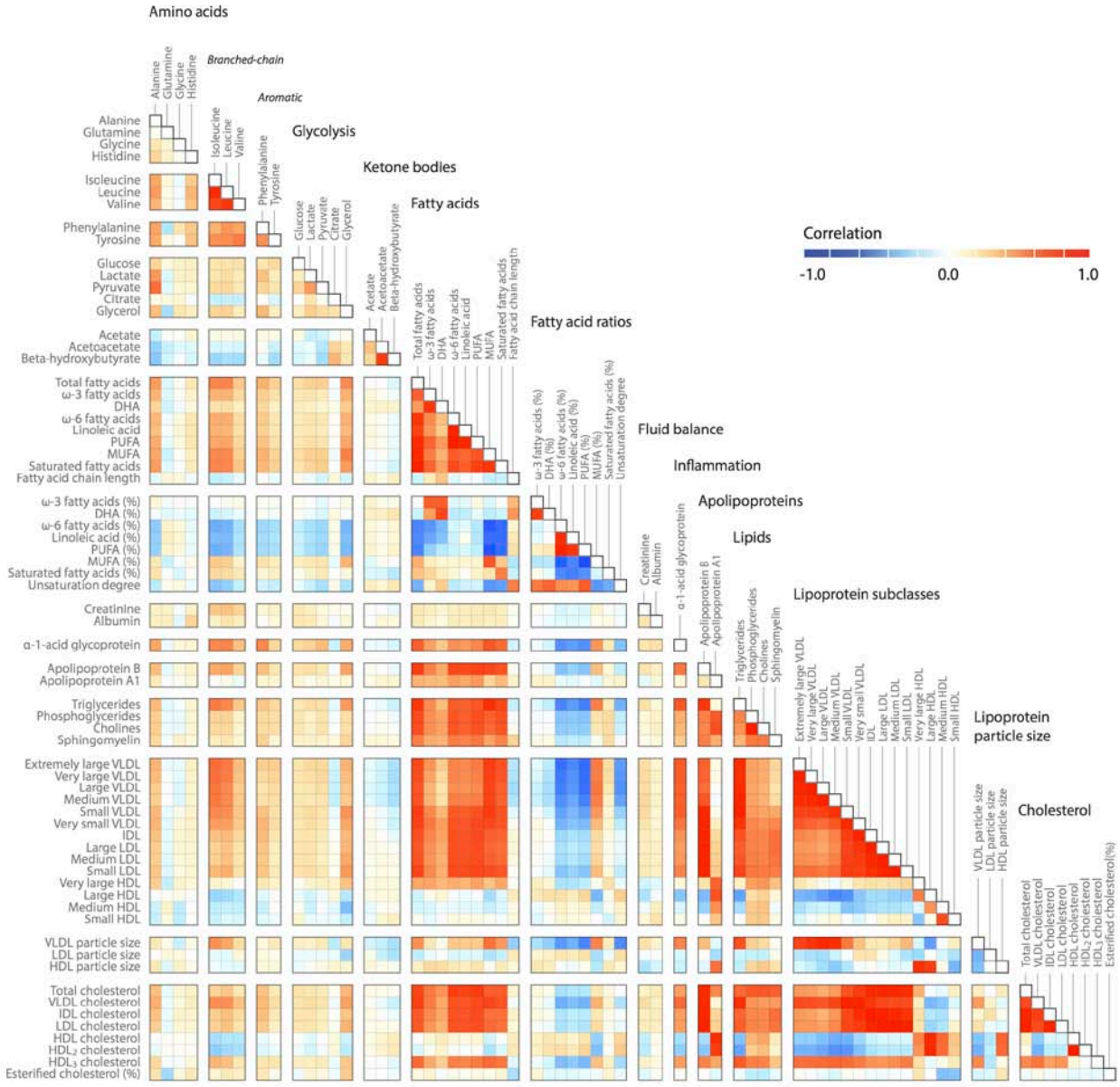
Reclassification was examined in the validation cohorts based on the risk prediction scores derived in the FINRISK study (coefficients listed in Supplemental Table 3). The established risk factors were age, sex, systolic blood pressure, smoking status, prevalent diabetes, antihypertensive medication, lipid-lowering medication, total cholesterol and HDL cholesterol. The biomarker risk score was further extended with log-transformed phenylalanine, the ratio of MUFA to total fatty acids, omega-6 fatty acids, and DHA.

Supplemental Table 5. Tracking of biomarker concentrations over time within the same individuals.

Metabolite biomarker	**6-year tracking in the YFS cohort (n=1754)** Partial correlation [95% CI]	**20-year tracking in the SABRE study (n=916)** Partial correlation [95% CI]
Phenylalanine	0.41 [0.37–0.44]	0.31 [0.25–0.36]
MUFA%, ratio to total fatty acids	0.47 [0.43–0.50]	0.45 [0.39–0.50]
Omega-6 fatty acids	0.47 [0.43–0.51]	0.20 [0.14–0.26]
PUFA	0.47 [0.44–0.52]	0.20 [0.13–0.26]
DHA	0.48 [0.45–0.52]	0.49 [0.44–0.55]

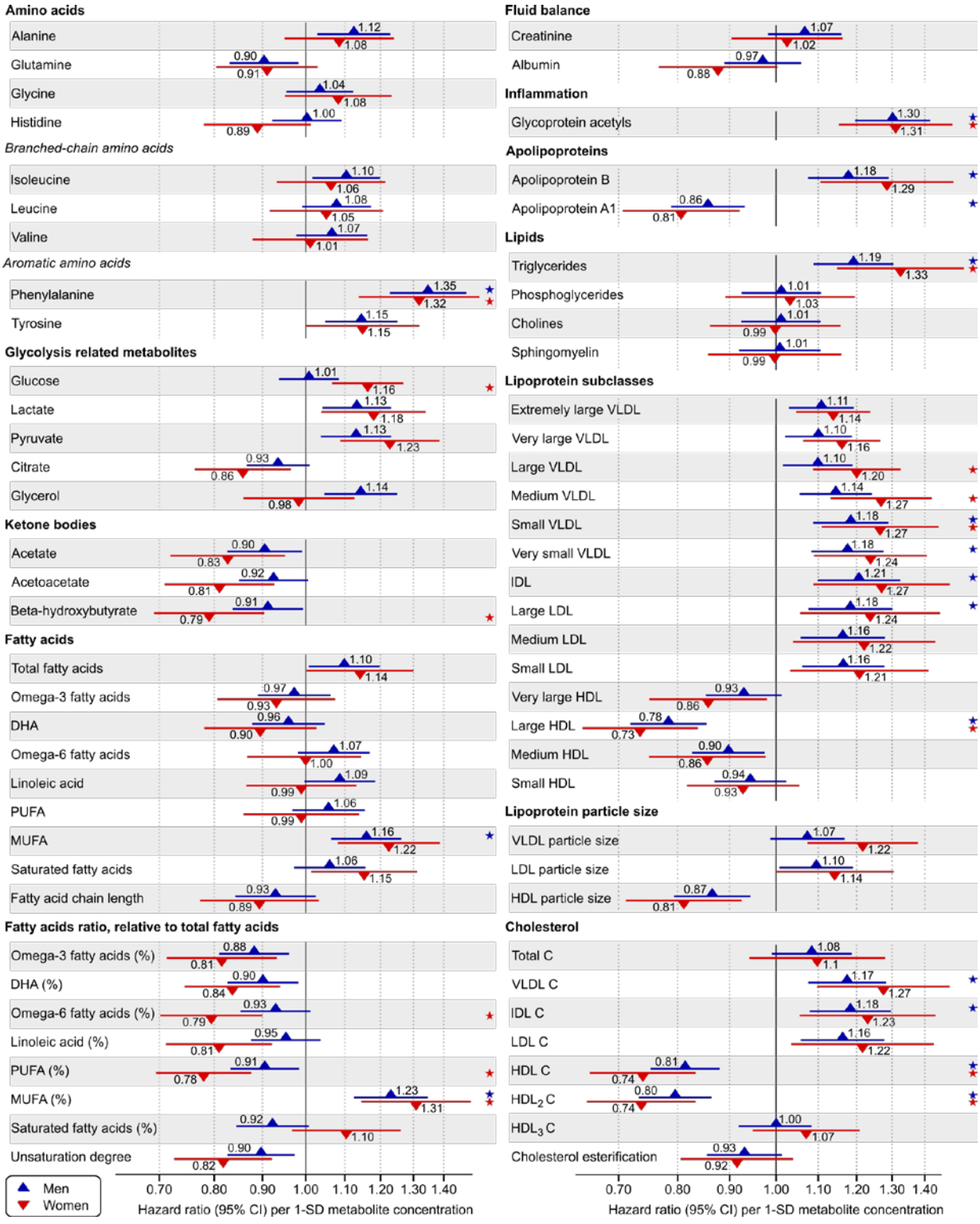
Partial Spearman's correlation coefficients [95% confidence intervals] of metabolite concentrations measured for the same individuals at two time-points, adjusted for age and sex.

Supplemental Figure 1. Correlation heatmap of the metabolite measures analyzed.



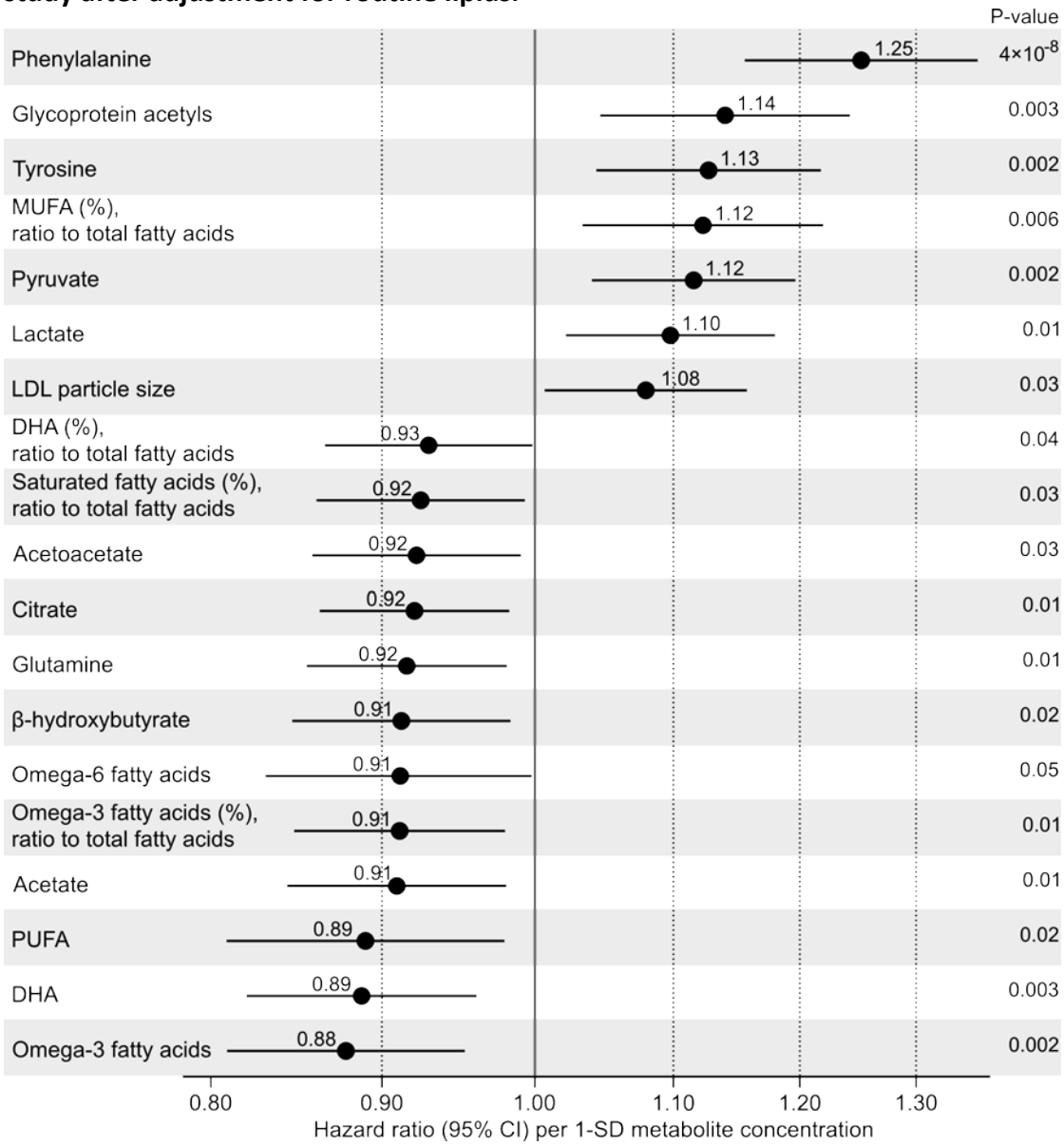
Correlation heatmap of serum metabolite measures quantified by the high-throughput NMR metabolomics platform.[12] Color coding indicates Spearman's correlation coefficients in the FINRISK study (n=7256). The correlation structure of the metabolites was similar in the other cohorts.[12,13]

Supplemental Figure 2. Metabolite associations with incident CVD for men and women in the FINRISK study.



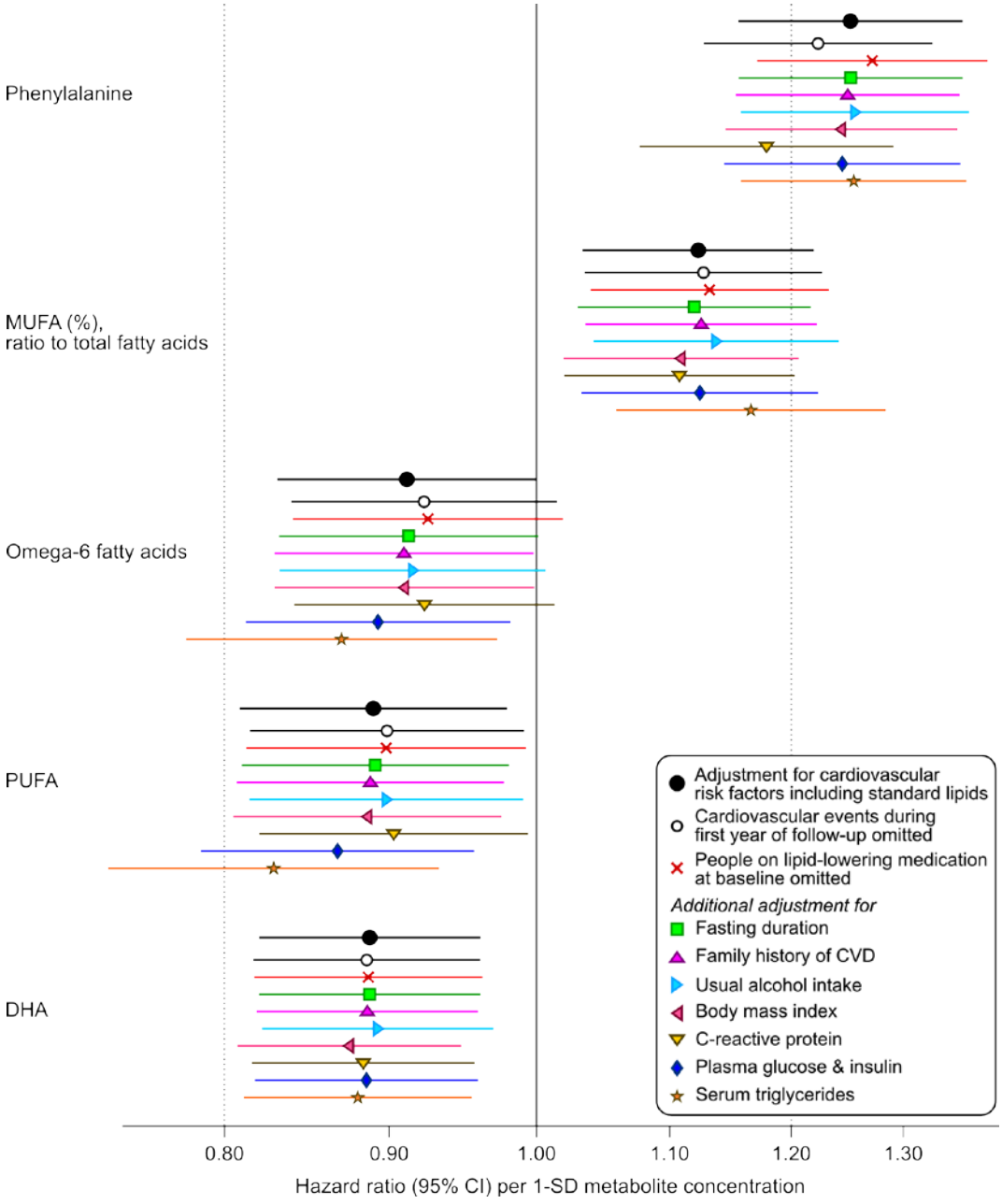
Metabolite associations with incident CVD in the FINRISK study stratified by sex (3578 men/566 events and 3678 women/234 events during 15-year follow-up). Hazard ratios are per 1-SD log-transformed metabolite concentration. Analyses were adjusted for age, blood pressure, smoking, diabetes status, geographical region, and cardiovascular medication. Errorbars denote 95% confidence intervals.
★: P<0.0007 (multiple testing correction).

Supplemental Figure 3. Metabolite biomarker candidates for incident CVD in the FINRISK study after adjustment for routine lipids.



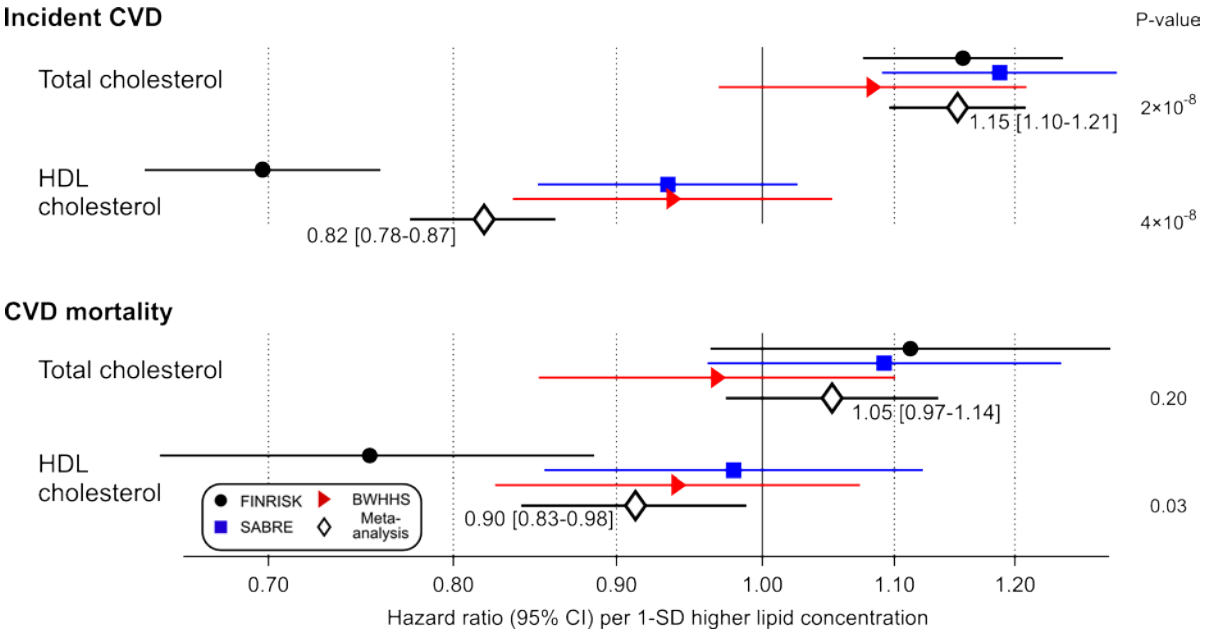
Metabolite associations with incident cardiovascular events in the FINRISK study at nominal significance (P<0.05) after adjustment for total and HDL cholesterol as well as age, sex, blood pressure, smoking, diabetes, and cardiovascular medication. Hazard ratios are per 1-SD increment in log-transformed metabolite concentration. Errorbars denote 95% confidence intervals. These metabolite biomarker candidates were meta-analyzed with the two replication studies (Figure 3 and Supplemental Figure 6).

Supplemental Figure 4. Biomarker associations with incident CVD upon additional exclusion criteria or covariate adjustment.



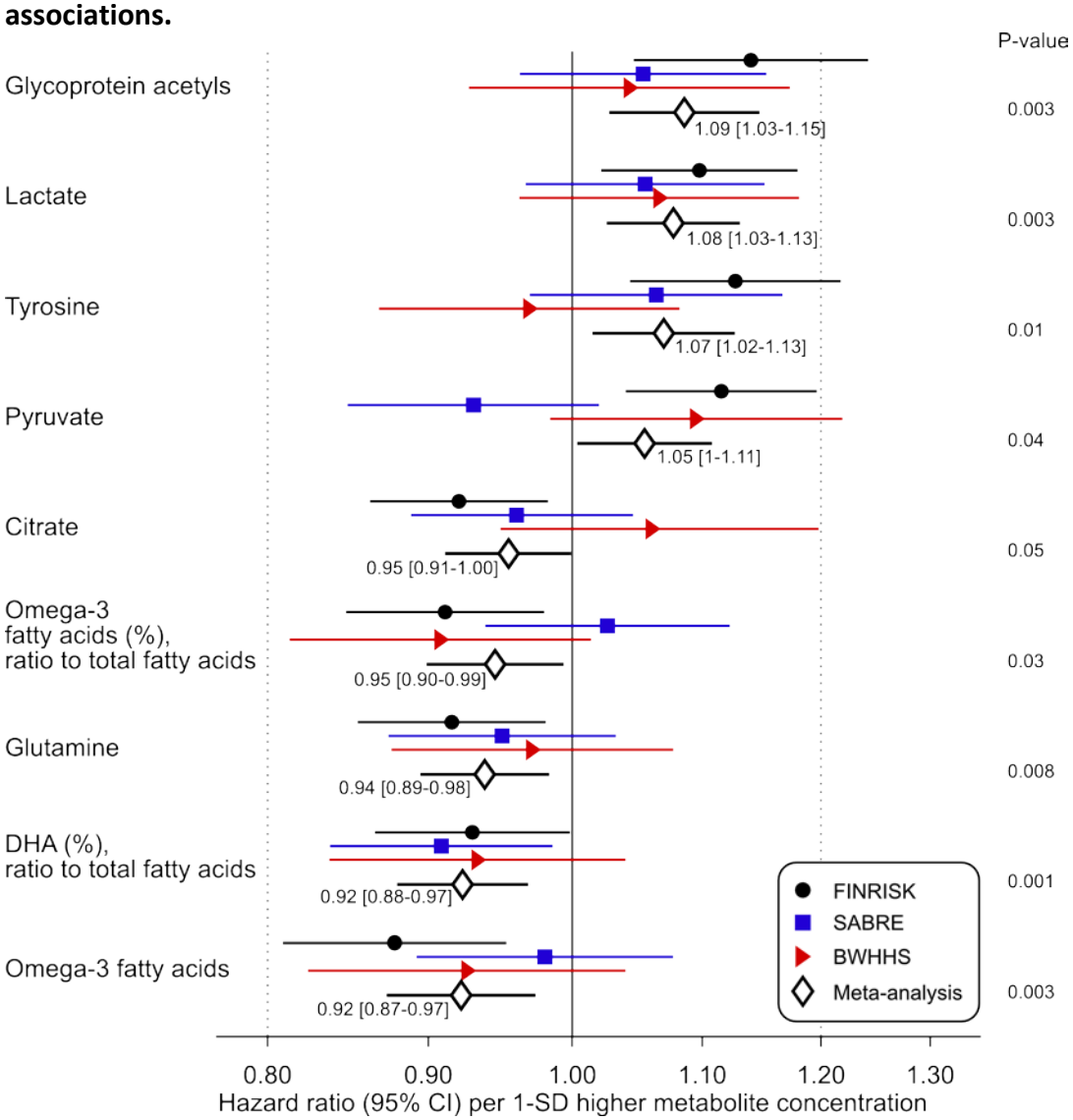
Metabolite biomarkers for incident CVD upon additional exclusion or further adjustment in the FINRISK study. Hazard ratios are per 1-SD increment in log-transformed metabolite concentration. Errorbars denote 95% confidence intervals.

Supplemental Figure 5. Associations of routine lipid measures for incident CVD and cardiovascular mortality.



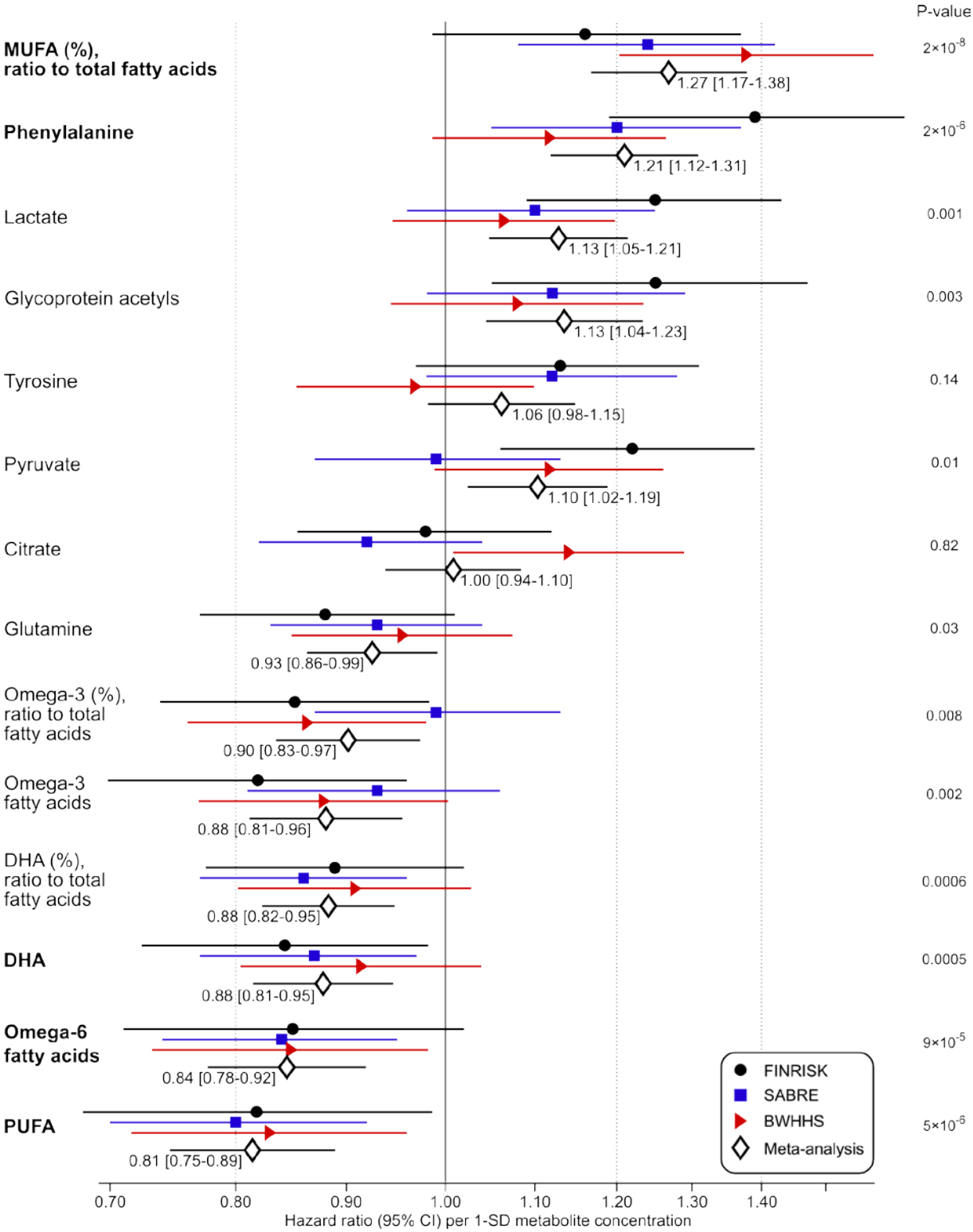
Hazard ratios of routine enzymatic lipid measures with incident CVD and cardiovascular mortality in the three cohorts. Meta-analysis results are shown by black diamonds, with hazard ratio, 95% confidence intervals, and P-value indicated. Analyses were adjusted for age, sex, blood pressure, smoking, diabetes status, geographical region, and cardiovascular medication. Hazard ratios are per 1-SD increment in lipid concentration.

Supplemental Figure 6. Meta-analysis of metabolite biomarkers for incident CVD, nominal associations.



Hazard ratios of metabolite biomarkers with incident CVD in three population-based studies (n=13441, 1741 events during 12-23 years of follow-up) and meta-analysis (diamonds). Metabolites associated with CVD at P≤0.05 in the meta-analysis, but not reaching the multiple testing corrected P≤0.0007, are shown. Analyses were adjusted for age, sex, blood pressure, smoking, geographical region, diabetes, cardiovascular medication as well as total and HDL cholesterol. Hazard ratios are per 1-SD increment in log-transformed metabolite concentration. Errorbars denote 95% confidence intervals.

Supplemental Figure 7. Meta-analysis of metabolite biomarkers with cardiovascular mortality.

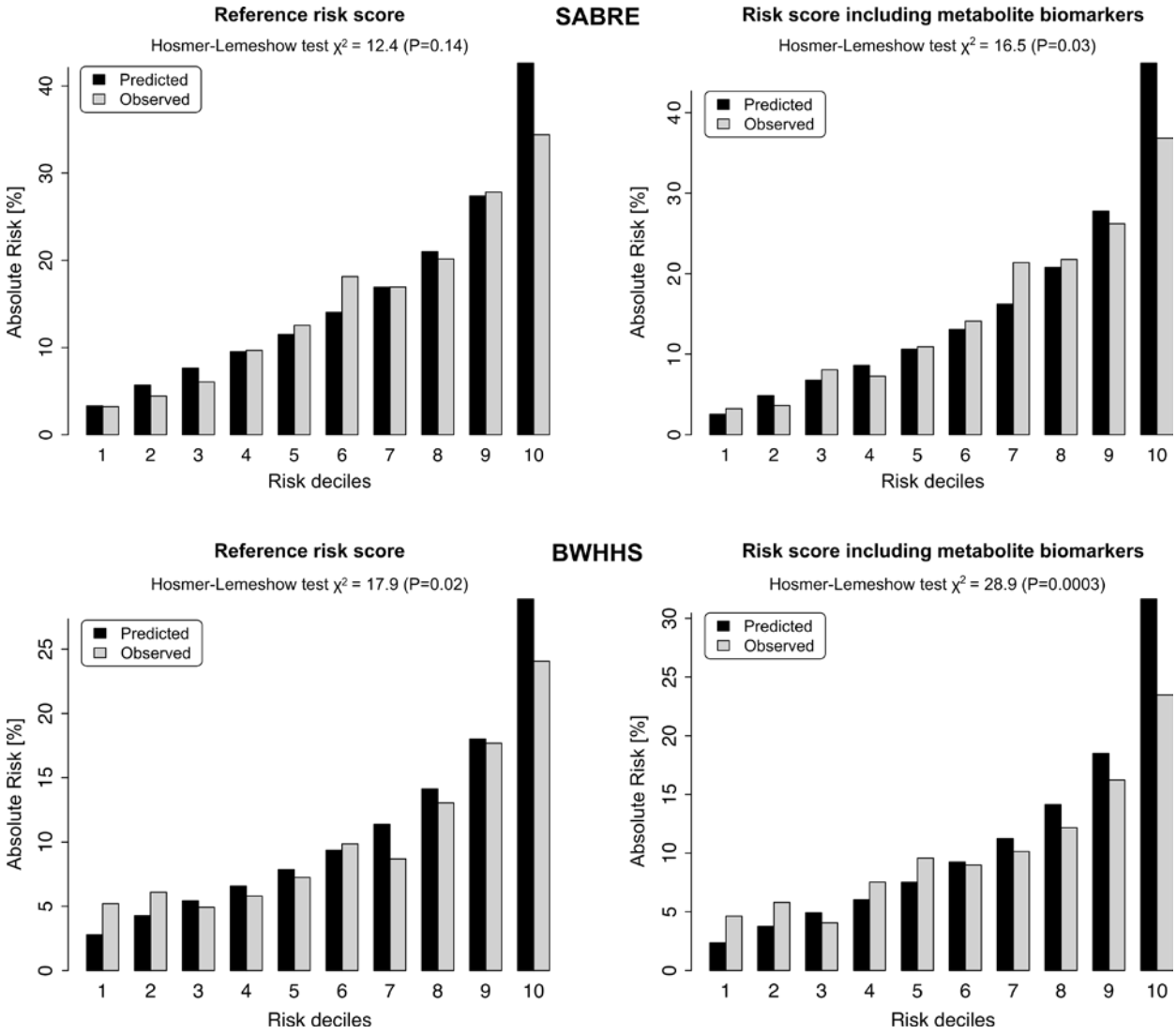


Associations of metabolite biomarkers with cardiovascular mortality (n=13821, 743 deaths). Meta-analysis results are shown by diamonds, with hazard ratio, 95% confidence intervals and P-value indicated. Analyses were adjusted for age, sex, blood pressure, smoking, geographical region, diabetes, cardiovascular medication as well as total and HDL cholesterol. Hazard ratios are per 1-SD increment in log-transformed metabolite concentration. Metabolites highlighted in bold reach the multiple testing correction of P<0.0007.

Supplemental Figure 8. Interaction between phenylalanine and age of incident CVD.

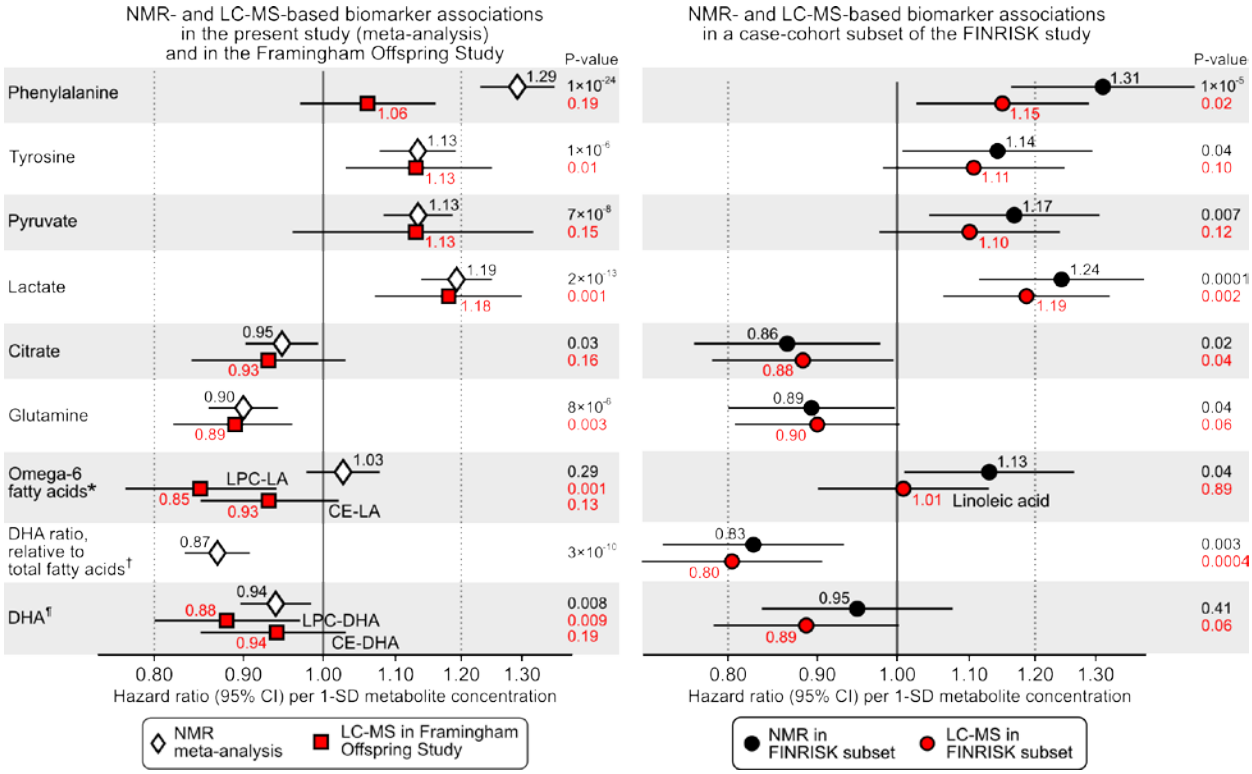
Hazard ratio per 1-SD log-transformed phenylalanine concentration as function of age at incident CVD or end of follow-up for non-events. Dashed lines indicate 95% confidence intervals of the hazard ratio estimates. The association magnitudes and P-values for phenylalanine*age-at-follow-up interaction term are indicated. Models were adjusted for age, sex, blood pressure, smoking, geographical region, diabetes, cardiovascular medication as well as total and HDL cholesterol.

Supplemental Figure 9. Calibration of risk scores for predicting first cardiovascular event.



Calibration within risk deciles of the reference risk score composed of established risk factors, and the biomarker risk score additionally including the four biomarkers (log-transformed phenylalanine, MUFA ratio to total fatty acids, omega-6 fatty acids, and DHA). The absolute risk of incident cardiovascular event was estimated in the SABRE and BWHHS studies based on the risk prediction scores derived in the FINRISK study. Calibration was examined using the Hosmer-Lemeshow goodness-of-fit test, which compares the observed number of death with those predicted from the risk score.

Supplemental Figure 10. Consistency between NMR and LC-MS for biomarker associations with incident cardiovascular disease: age- and sex-adjusted associations.



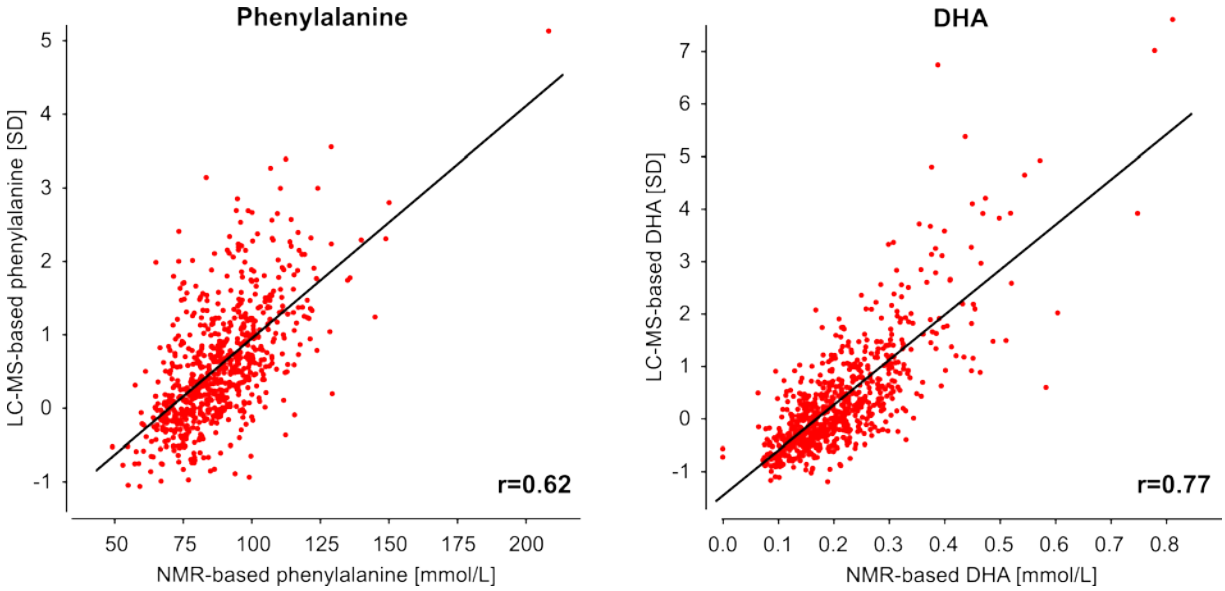
Biomarker associations of metabolites overlapping between metabolomics platforms and nominally associated with incident CVD in the present study (P<0.05, Supplemental Figure 6). Hazard ratios are per 1-SD log-tranformed metabolite concentration. Errorbars denote 95% confidence intervals. Associations were adjusted for age and sex. LC-MS-based associations were further adjusted for batch.
Left panel: Biomarker associations observed in the present study compared with those obtained in the Framingham Offspring Study based on LC-MS (n=2289, 466 events).
Right panel: Biomarker associations in a case-cohort subset of the FINRISK study (n=679, 305 events) profiled both by NMR and LC-MS.

* Associations of omega-6 fatty acids were compared with lysophosphatidylcholine- and cholesterolester-linoleic acid in the Framingham Offspring Study and with linoleic acid in FINRISK.

† DHA ratio was scaled to the total fatty acid concentration quantified by NMR for both NMR and LC-MS platforms.

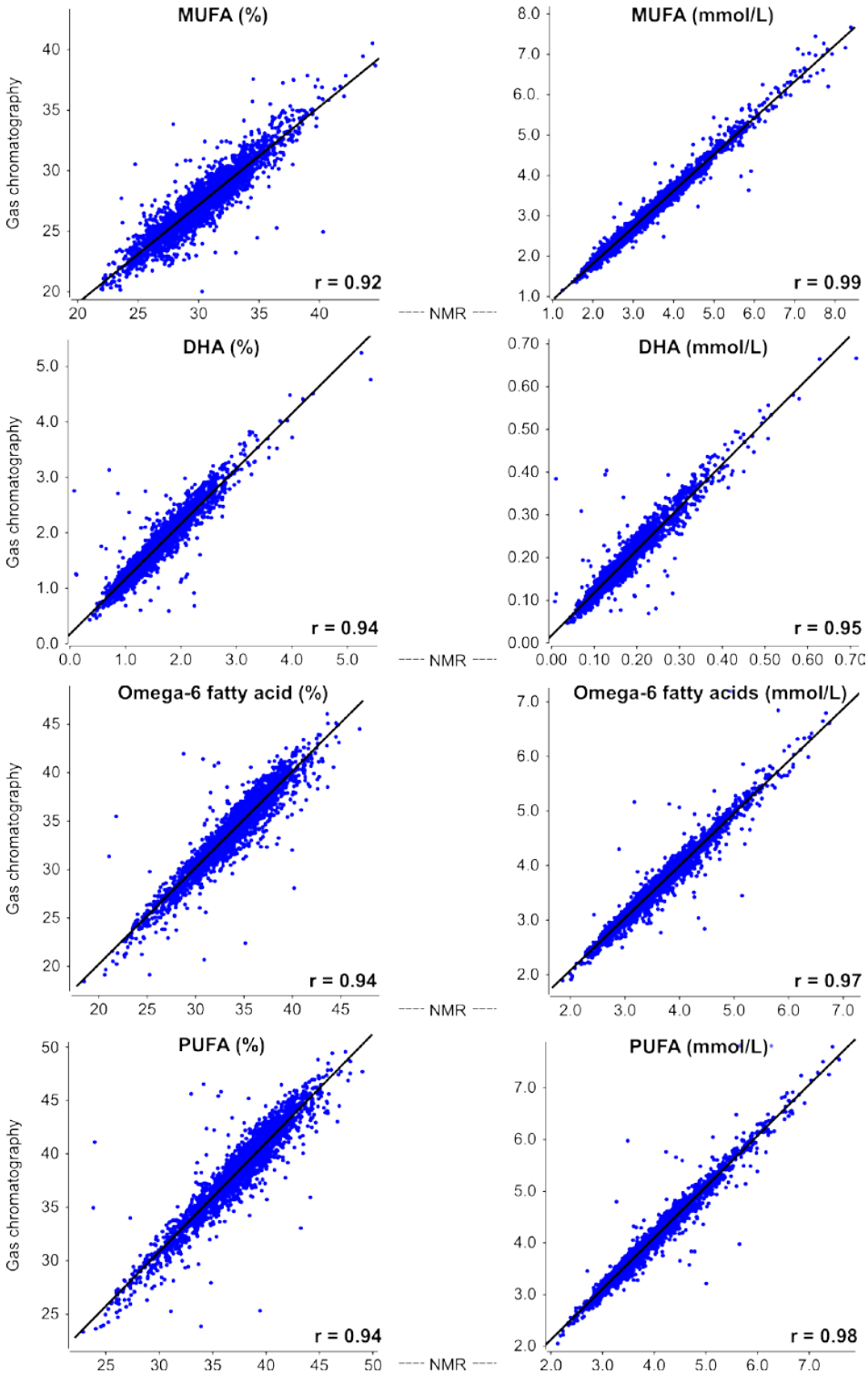
¶ Associations of DHA were compared with lysophosphatidylcholine- and cholesterolester-DHA in the Framingham Offspring Study.

Supplemental Figure 11. Scatter plots of metabolite biomarkers assayed by both NMR spectroscopy and liquid-chromatography mass spectrometry.



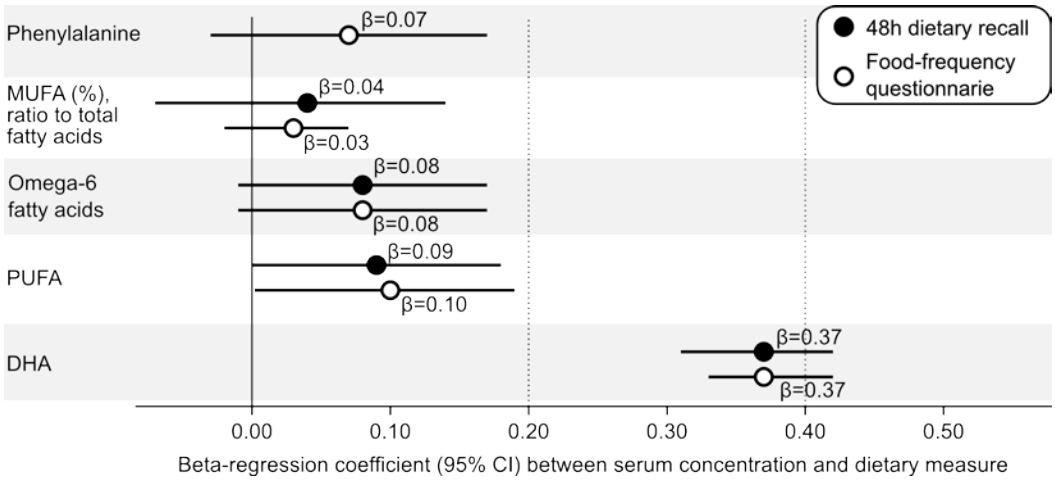
Concentrations of phenylalanine and DHA were measured by both NMR and LC-MS in a case-cohort subset of the FINRISK study (n=679). The linear fit of the analytical correspondence is indicated. *r* denotes Pearson's correlation coefficients. Metabolite concentrations from LC-MS were scaled to SD-units since the data were not quantified in absolute concentration units. Based on the NMR quantification, 1 SD corresponds to 14.4 µmol/L for phenylalanine and 90 µmol/L for DHA (*cf*. Supplementary Table 1). Quantification of the significant biomarkers for incident CVD in the present study (Figure 3) that were not measured by LC-MS was confirmed by gas chromatography (Supplemental Figure 12).

Supplemental Figure 12. Scatter plots of biomarkers assayed by both NMR and gas chromatography.



Fatty acids associated with incident CVD in the present study (Figure 3) were measured by both NMR and gas chromatography in the 2001-survey of the Cardiovascular Risk in Young Finns Study (n=2193). Fatty acid concentrations are measured relative to total fatty acids by gas chromatography; concentrations in mmol/L were obtained by scaling to the total fatty acid concentration determined by NMR. The linear fit of the analytical correspondence is indicated, and *r* denotes Pearson's correlation coefficients.

Supplemental Figure 13. Associations of dietary intake with blood levels of metabolite biomarkers.



Associations of dietary intake with circulating biomarkers for CVD at two time points in the Cardiovascular Risk in Young Finns Study. Associations were adjusted for age, sex, body mass index and total energy intake. Serum metabolite concentrations and dietary measures were determined in the Cardiovascular Risk in Young Finns Study in the 2001 field survey by 48h dietary recall (n=960)[16], as well as by food-frequency questionnaire in the 2007 field survey (n=1699).[17] Dietary phenylalanine was only determined by the food-frequency questionnaire.

Metabolite Profiling and Cardiovascular Event Risk: A Prospective Study of 3 Population-Based Cohorts

Peter Würtz, Aki S. Havulinna, Pasi Soininen, Tuulia Tynkkynen, David Prieto-Merino, Therese Tillin, Anahita Ghorbani, Anna Artati, Qin Wang, Mika Tiainen, Antti J. Kangas, Johannes Kettunen, Jari Kaikkonen, Vera Mikkilä, Antti Jula, Mika Kähönen, Terho Lehtimäki, Debbie A. Lawlor, Tom R. Gaunt, Alun D. Hughes, Naveed Sattar, Thomas Illig, Jerzy Adamski, Thomas J. Wang, Markus Perola, Samuli Ripatti, Ramachandran S. Vasan, Olli T. Raitakari, Robert E. Gerszten, Juan-Pablo Casas, Nish Chaturvedi, Mika Ala-Korpela and Veikko Salomaa

Circulation. 2015;131:774-785; originally published online January 8, 2015;
doi: 10.1161/CIRCULATIONAHA.114.013116
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7322. Online ISSN: 1524-4539