Published Ahead of Print on December 2, 2020 as 10.1212/WNL.0000000000011236

OPEN

Neurology®

The most widely read and highly cited peer-reviewed neurology journal
The Official Journal of the American Academy of Neurology

AMERICAN ACADEMY OF NEUROLOGY®

Neurology Publish Ahead of Print
DOI: 10.1212/WNL.0000000000011236

NEUROLOGY/2019/013482

Association of circulating metabolites in plasma or serum and risk of stroke: Meta-analysis from seven prospective cohorts

The Article Processing Charge was funded by Erasmus Medical Center.

Dina Vojinovic, PhD[1], Marita Kalaoja, MSc[2], Stella Trompet, PhD[3,4], Krista Fischer, PhD[5], Martin J. Shipley, MSc[6], Shuo Li, MSc[7], Aki S. Havulinna, PhD[8], Markus Perola, PhD[8,9], Veikko Salomaa, PhD[8], Qiong Yang, PhD[7], Naveed Sattar, PhD[10], Pekka Jousilahti, PhD[8], Najaf Amin, PhD[1], Claudia L. Satizabal, PhD[12,21,23], Nele Taba, MSc[5], Behnam Sabayan, PhD[11], Ramachandran S. Vasan, MD[12], M. Arfan Ikram, PhD[1,13], David J. Stott, PhD[14], Mika Ala-Korpela, PhD[2,15,16,17,18,19], J. Wouter Jukema, PhD[4,20], Sudha Seshadri, MD[12,21,23], Johannes Kettunen, PhD[2,8], Mika Kivimaki, FMedSciP[6], Tonu Esko, PhD[5,22], Cornelia M. van Duijn, PhD[1,24*]

[1]Department of Epidemiology, Erasmus MC, University Medical Center, Rotterdam, the Netherlands;
[2]Computational Medicine, Faculty of Medicine, University of Oulu and Biocenter Oulu, Oulu, Finland;
[3]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands;
[4]Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands;
[5]Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia;
[6]Department of Epidemiology and Public Health, UCL, London, UK;
[7]Department of Biostatistics, School of Public Health, Boston University, Boston, MA, USA;
[8]Department of Public Health Solutions, National Institute for Health and Welfare, Helsinki, Finland;
[9]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland;
[10]BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK;
[11]Department of Neurology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA;
[12]The Framingham Heart Study, Framingham, MA, USA;
[13]Department of Radiology and Nuclear Medicine, Erasmus MC, University Medical Center, Rotterdam, the Netherlands;
[14]Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK;
[15]Systems Epidemiology, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia;
[16]NMR Metabolomics Laboratory, School of Pharmacy, University of Eastern Finland, Kuopio, Finland;

[17]Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK;
[18]Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Bristol, UK;
[19]Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Faculty of Medicine, Nursing and Health Sciences, The Alfred Hospital, Monash University, Melbourne, Victoria, Australia;
[20]Einthoven Laboratory for Experimental Vascular Medicine, LUMC, Leiden, the Netherlands;
[21]Department of Neurology, Boston University School of Medicine, Boston, MA, USA;
[22]Broad Institute of MIT and Harvard, Boston, MA, USA;
[23]Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, UT Health San Antonio, TX, USA;
[24]Nuffield Department of Population Health, University of Oxford, Oxford, UK;

*Corresponding author:
Prof. Cornelia M. van Duijn
E-mail: Cornelia.vanDuijn@ndph.ox.ac.uk

Word count abstract: 234
Word count text: 4379
Character count title: 71
Number of references: 60
Number of tables: 3
Number of figures: 1

STUDY FUNDING
No targeted funding reported.

DISCLOSURES

Dr. Dina Vojinovic, Ms. Marita Kalaoja, Dr. Stella Trompet, Dr. Krista Fischer, Mr. Martin J. Shipley, Mr. Shuo Li, Dr. Aki S. Havulinna, Dr. Markus Perola, Dr. Veikko Salomaa, Dr. Qiong Yang, Dr. Naveed Sattar, Dr. Pekka Jousilahti, Dr. Najaf Amin, Dr. Claudia L. Satizabal, Ms. Nele Taba, Dr. Behnam Sabayan, Dr. Ramachandran S. Vasan, Dr. M. Arfan Ikram, Dr. David J. Stott, Dr. Mika Ala-Korpela, Dr. J. Wouter Jukema, PhD, Dr. Sudha Seshadri, Mr. Mika Kivimaki, Dr. Tonu Esko and Dr. Cornelia M. van Duijn report no disclosures.

Dr. Johannes Kettunen possesses modest amount of stock options of Nightingale Health Ltd, a company offering NMR metabolomics profiling.

ABSTRACT

Objective: To conduct a comprehensive analysis of circulating metabolites and incident stroke in large prospective population-based settings.

Methods: We investigated the association of metabolites with risk of stroke in seven prospective cohort studies including 1,791 incident stroke events among 38,797 participants in whom circulating metabolites were measured by Nuclear Magnetic Resonance (^{1}H-NMR) technology. The relationship between metabolites and stroke was assessed using Cox proportional hazards regression models. The analyses were performed considering all incident stroke events and ischemic and hemorrhagic events separately.

Results: The analyses revealed ten significant metabolite associations. Amino acid histidine (hazard ratio (HR) per standard deviation (SD) = 0.90, 95% confidence interval (CI): 0.85, 0.94; $P = 4.45\times10^{-5}$), glycolysis-related metabolite pyruvate (HR per SD = 1.09, 95% CI: 1.04, 1.14; $P = 7.45\times10^{-4}$), acute phase reaction marker glycoprotein acetyls (HR per SD = 1.09, 95% CI: 1.03, 1.15; $P = 1.27\times10^{-3}$), cholesterol in high-density lipoprotein (HDL) 2 and several other lipoprotein particles were associated with risk of stroke. When focusing on incident ischemic stroke, a significant association was observed with phenylalanine (HR per SD = 1.12, 95% CI: 1.05, 1.19; $P = 4.13\times10^{-4}$) and total and free cholesterol in large HDL particles.

Conclusions: We found association of amino acids, glycolysis-related metabolites, acute phase reaction markers, and several lipoprotein subfractions with the risk of stroke. These findings support the potential of metabolomics to provide new insights into the metabolic changes preceding stroke.

.

INTRODUCTION

Stroke is a leading cause of death and serious long-term disability worldwide.[1] The majority of strokes are of the ischemic type, while the hemorrhagic type occurs less often but is associated with higher mortality risk.[1, 2] Stroke risk is determined by various modifiable risk factors such as hypertension, diabetes mellitus, cardiovascular disease, smoking, and obesity, whereas the association of stroke with cholesterol and its subfractions has shown inconsistent results.[1-5] Opportunities for therapeutic interventions in stroke patients depend on the type of stroke and rely on brain imaging techniques.[6] Despite advances in brain imaging techniques, costs are still high, availability is limited and not all patients show a relevant lesion on neuroimaging.[6, 7] New technology is needed to identify high-risk patients, to understand the etiology of stroke, and develop future prevention strategies. Detailed profiling of metabolic status can provide insights into metabolic changes that lead to a higher risk of stroke. As the metabolome reflects both genome and exposome including exposures to risk factors that determine the risk of stroke, this new –omics technology may open new avenues towards stroke prevention. To date, only few studies have analyzed metabolic disturbances in stroke and identified various metabolites to be associated with stroke.[8-10] These studies are based on relatively small samples or performed on participants of non-European ancestry.[11] The most comprehensive study to date was a nested case-control study conducted by Holmes *et al.* within the China Kadoorie Biobank including 1,146 patients with ischemic stroke and 1,138 patients with intracerebral hemorrhage.[11] The study reported an association between lipids and lipoprotein particles of various sizes with ischemic stroke but not with hemorrhage.[11] Furthermore, the study identified glycoprotein acetyls, ketone bodies, glucose, and docosahexaenoic acid to be associated with both ischemic and hemorrhagic stroke.[11]

As large metabolomics studies of stroke in persons of European origin are lacking and data from well-established prospective cohort studies are limited, the aim of our study is to conduct a comprehensive analysis of circulating metabolites and incident stroke in large prospective population-based settings involving 1,791 incident stroke events among 38,797 participants of European origin.

METHODS

Study population

Our study population included 38,797 participants from seven cohorts including the Rotterdam Study, Whitehall II study (Whitehall II), the national FINRISK study1997 (FINRISK97), Dietary, Lifestyle and Genetic determinants of Obesity and Metabolic Syndrome (DILGOM), PROspective Study of Pravastatin in the Elderly at Risk (PROSPER), Estonian biobank (EGCUT), and the Framingham Heart Study (FHS). Description of participating studies is available from Dryad (Additional Methods in Additional Information file): https://doi.org/10.5061/dryad.kh1893239.

Stroke assessment

Details on stroke assessment are available from Dryad (Additional Methods in Additional Information file): https://doi.org/10.5061/dryad.kh1893239**.** The incident stroke events were assessed through follow-up of health records, while in some studies additional periodic visits to research centers were used (e.g. Rotterdam Study, FHS). Participants of the Rotterdam Study were monitored for incident stroke using an automated linkage of medical records from general practitioners with the study database.[12] Incident stroke events in the Whitehall II study were ascertained through linkage to electronic records from hospitalizations due to stroke and national statistics death registries,[13, 14] whereas in the FINRISK and DILGOM studies linkage to national health registries was used

(https://www.biorxiv.org/content/early/2018/03/12/280677). Ascertainment of incident stroke events in EGCUT was also performed through linkage to electronic records from multiple databases (https://thl.fi/publications/morgam/cohorts/ full/estonia/est-esta.htm), while information regarding domiciliary visits or hospitalizations associated with possible cardiovascular events including stroke, and information on all deaths was used for classification of study endpoints in PROSPER.[15] In the FHS, incident clinical stroke was identified as part of ongoing clinic and hospital surveillance with additional stroke surveillance by annual phone health updates and collaboration with primary care physicians and local emergency departments.[16, 17] Participants with a history of stroke at baseline were excluded from the analyses.

Baseline clinical characteristics

The baseline clinical characteristics included assessment of blood pressure, plasma glucose levels, smoking status, weight, and height. Hypertension was defined as a systolic blood pressure ≥140 mmHg, diastolic blood pressure ≥90 mmHg, or use of antihypertensive medication or based on GP diagnosis, medication reimbursement or ICD-diagnosis (I10/401/491; ICD-10/9/8) from hospital discharge register or cause of death register, or self-reported. Diabetes was defined as fasting plasma glucose levels >7 mmol/L or use of medication indicated for the treatment of diabetes. Body mass index (BMI) was calculated as weight in kilograms divided by square of heights in meters.

Metabolite quantification

Circulating metabolites were quantified using a high-throughput Nuclear Magnetic Resonance (NMR) technology. In all participating studies except the FHS, the Nightingale Health metabolomics platform (Helsinki, Finland) was used for simultaneous quantification of a wide range of metabolites, including routine lipids, 14 lipoprotein subclasses and their lipids

(esterified cholesterol, free cholesterol, total cholesterol, triglycerides, phospholipids, and total lipids), fatty acids, amino acids, ketone bodies, and various glycolysis precursors. A detailed description of the methodology has been provided previously.[18, 19] The Nightingale's quality control procedures were applied to the data from each cohort. Nightingale's automated data processing and quality control procedures included safety checks for unexpected metabolite signals. For each sample, integrated quality procedures verified the sample quality by reporting signs of degradation and contamination issues. If a metabolite concentration was below the limit of quantification, due to either biological reasons or external compounds interfering quantification, but above the limit of detection, the metabolite value was presented as zero. Nightingale's quality control procedure did not have an upper limit for quantified concentrations, except for high lactate and high pyruvate, which were established irregularities arising from suboptimal sample collection procedure. Other high values were reported as they are, meaning that high concentrations were not excluded as in some context high biomarker concentration may be biologically and physiologically relevant and provide valuable molecular insight on a disease or an outcome. Every metabolite that has been reported in the results file has passed this strict quality control procedure. No additional single cohort quality control was applied except for the EGCUT cohort in which metabolites that were detected in the large majority (at least 95%) of the cohort were included while individuals with missing values in more than 10% of the metabolites were excluded from the analyses. In the FHS, lipoprotein subclasses were measured by proton NMR spectroscopic assay (LipoScience, Raleigh, NC).[20, 21] Blood samples were collected after overnight fasting in all studied except for FINRISK97 - in which the samples were collected after 4 hours of fasting (semi-fasting state).[22, 23] The sample material was EDTA-plasma in the Rotterdam Study, FHS, and EGCUT, whereas the serum was used in FINRISK97, DILGOM, PROSPER, and Whitehall II.[22-25] The EDTA-plasma/serum samples were stored either in -80 °C or in -70

°C. The duration of sample storage was ranged from 2-9 years in EGCUT and 8 years in DILGOM to 11 years in Rotterdam Study, 15 years in FINRISK97, and 15-20 years in PROSPER. There were 147 primary non-derived metabolite measurements quantified in absolute concentration units that were further analyzed in this study (Data available from Dryad (Additional Table 1): https://doi.org/10.5061/dryad.kh1893239). The descriptive statistics of metabolites were coherent across the cohort (cohort specific descriptive statistics are available from Dryad (Additional Table 2): https://doi.org/10.5061/dryad.kh1893239).

Statistical analyses

To obtain an approximately normal distribution, all metabolites measurements were natural logarithmic transformed prior to the analyses. As some of the metabolite values in our datasets were below the limit of quantification and therefore presented as zero, one was added to all values of the metabolites before the transformation. The metabolite measurements were subsequently scaled to standard deviation (SD) units (mean = 0, SD = 1) to enable comparison of results for measures with different units and across wide ranges of concentrations. The relationship between metabolites and stroke was assessed using Cox proportional hazards regression models. The analyses were performed while adjusting for age, gender, BMI, lipid-lowering medication, and study-specific covariates if needed (Model 1). The associations were further adjusted for smoking status, diabetes, and hypertension (Model 2). Proportional hazard assumption was tested in EGCUT and FINRISK97. The violation of this assumption was observed for 7 metabolites (noted with an asterisk in Additional Table 1: https://doi.org/10.5061/dryad.kh1893239). None of these metabolites showed statistically significant association with incident stroke in our analyses. To overcome the problem of false positives due to differences in study design, sampling, storage, or metabolite assessment, we did not pool and analyze the data of different studies jointly. Rather, we have analyzed individual studies and combined the findings using the meta-analysis. The summary statistics

results of participating studies were combined using inverse variance-weighted fixed-effect meta-analysis in METAL.[26] The heterogeneity of effects was assessed by I^2 which indicates the percentage of variance in the meta-analysis attributable to study heterogeneity.[26, 27] All hazard ratios (HR) of continuous variables are expressed per one SD of the transformed variable. The analyses were performed considering all incident stroke events and ischemic and hemorrhagic events separately.

As most of 147 metabolite measures are highly correlated, we estimated the number of independent tests in the correlation matrix using the previously described method of Li and Ji.[28] Subsequently, the number of independent tests was used for calculation of Bonferroni corrected *p*-value (*p*-value=0.05/30 independent metabolites = 1.7×10^{-3}).

Standard Protocol Approvals, Registrations, and Patient Consents

Each of the participating studies received approval by local ethical committees or institutional review boards (Data available from Dryad (Additional Methods in Additional Information file): https://doi.org/10.5061/dryad.kh1893239). All participants provided written informed consent.

Data Availability

Data available upon request. Interested researchers may contact the corresponding author.

RESULTS

The baseline descriptive characteristics of study participants are shown in **Table 1**. In total there were 1,791 incident stroke events observed among 38,797 participants across the seven cohorts. The mean follow-up time ranged from 2 years in PROSPER, 6 years in the Rotterdam Study, and 7 years in EGCUT and DILGOM to 13 years in Whitehall II and 15 years in FINRISK97 and FHS.

The association analysis between circulating metabolites and all incident stroke revealed 27 significant metabolite associations ($P<1.7\times10^{-3}$) in model 1 which are shown in **Table 2**. After further adjustment for hypertension status, diabetes, and smoking, 7 metabolite associations remained significant after correction for multiple testing (**Table 2, Figure 1**). These included the amino acid histidine (HR = 0.90; 95 % confidence interval (CI): 0.85, 0.94; $P = 4.45\times10^{-5}$) and cholesterol in high-density lipoprotein (HDL) 2 (HR = 0.91; 95% CI: 0.87, 0.97; $P = 1.41\times10^{-3}$) which were associated with a lower risk of stroke, glycolysis-related metabolite pyruvate (HR = 1.09; 95% CI: 1.04, 1.14; $P = 7.45\times10^{-4}$) and acute phase reaction markers glycoprotein acetyls (HR = 1.09; 95% CI: 1.03, 1.15; $P = 1.27\times10^{-3}$) which were associated with higher risk of stroke, and several lipoprotein particles including HDL and low-density lipoprotein (LDL) subfractions (**Table 2**). Cholesterol in medium HDL was associated with lower risk (HR = 0.92; 95% CI: 0.87, 0.97; $P = 1.35\times10^{-3}$), whereas triglycerides in medium and large LDL were associated with a higher risk of stroke (HR = 1.09; 95% CI: 1.03, 1.14; $P = 1.67\times10^{-3}$ and HR = 1.09; 95% CI: 1.03, 1.14; $P = 1.19\times10^{-3}$, respectively) (**Table 2**). The direction of effect across the cohorts showed no evidence of a single cohort driving the associations (Data available from Dryad (Additional Figure 1 in Additional Information file): https://doi.org/10.5061/dryad.kh1893239). Whereas the Whitehall II study showed the opposite direction of effect for apolipoprotein A, HDL, and HDL2 cholesterol, the findings showed a general spread for most HDL subfractions.

When we stratified the analysis by stroke type, we observed differences between ischemic and hemorrhagic stroke events (**Table 3**). Amino acid histidine and cholesterol in HDL2 were associated with decreased risk of ischemic but not hemorrhagic incident stroke (**Table 3**). Differences were also observed for glycolysis-related metabolite pyruvate and acute phase inflammation marker glycoprotein acetyls which were associated with increased risk of ischemic but not hemorrhagic stroke (**Table 3)**. Associations between incident stroke events

and LDL and HDL particles of various sizes were observed only in the overall analysis, suggesting contributions from both stroke subtypes (**Table 3**).

Furthermore, a significant association was observed between phenylalanine levels and increased risk of incident ischemic stroke (HR = 1.12; 95% CI: 1.05, 1.19; $P = 4.13\times10^{-4}$). We also observed association of circulating levels of cholesterol (HR = 0.89; 95% CI: 0.84, 0.95; $P = 9.00\times10^{-4}$) and free cholesterol in large HDL cholesterol (HR = 0.89; 95% CI: 0.82, 0.95; $P = 1.33\times10^{-3}$) with decreased risk of ischemic stroke. No metabolite surpassed the significant threshold in the analysis for hemorrhagic stroke.

DISCUSSION

In this study, we identified ten metabolites associated with the risk of stroke. These include amino acid histidine and cholesterol in HDL2 which were associated with decreased risk of stroke overall and ischemic stroke subtype and glycolysis-related metabolite pyruvate and acute phase reaction markers glycoprotein acetyls which were associated with increased risk of stroke overall and ischemic stroke. Cholesterol in medium HDL and triglycerides in medium and large LDL particles were associated with stroke overall, while amino acid phenylalanine and HDL subfractions including cholesterol and free cholesterol in large HDL were associated with ischemic but not with hemorrhagic stroke. This pattern of results was independent of traditional risk factors including hypertension, diabetes, smoking, and BMI.

The strongest association was found between amino acid histidine and risk of stroke. We observed that one SD increase in concentration of histidine was associated with 10% lower risk of stroke. The effect was similar across studies, with only the Finrisk97 study showing no effect. Even though the same direction of effect was observed for both ischemic and hemorrhagic stroke subtypes, the association was mainly driven by ischemic stroke. Histidine is a semi-essential amino acid as adults generally produce it while children may not. Histidine

can be converted to histamine which shows a strong effect on vasodilatation and functions as a neurotransmitter in the brain.[29, 30] Previous studies reported that oral administration of histidine can reduce blood pressure.[31-33] Plasma concentrations of histidine have been inversely associated with inflammation and oxidative stress in patients with chronic kidney disease and obese women with metabolic syndrome.[34, 35] Recent animal studies reported that histidine treatment alleviated the infarction induced by middle cerebral artery occlusion[36] and showed long term-neuroprotection after cerebral ischemia with decreased infarct volume and improved neurological function.[37] Even though our findings support the results of previous studies, in the most comprehensive study of stroke to date within the China Kadoorie Biobank, histidine was not associated with ischemic or hemorrhagic stroke. However, in that study, a nominal association was found with myocardial infarction.[11] This might be explained either by environmental or ethnic differences of studied populations, or differences in the confounders adjusted for. In the present study, we adjusted for a more comprehensive set of potential confounders including BMI, lipid-lowering medication, diabetes, and hypertension.

We also found the glycolysis-related metabolite, pyruvate, to be associated with increased risk of stroke. The analyses of stroke subtypes suggested that this association was driven by ischemic incident stroke events. Our findings suggested that one SD increase in pyruvate concentration was associated with 13% higher risk of ischemic stroke. Pyruvate is the end-product of glycolysis and it is critical for supplying energy to the cell.[38] Pyruvate has previously been shown to protect against experimental stroke possibly by blocking inflammation.[39, 40] In this light, our finding seems to contrast previously described effects of pyruvate. However, in a combined study of myocardial infarction and stroke using the same metabolomics platform as the present study, higher levels of pyruvate were also associated with a higher risk of cardiovascular disease.[41] The mechanisms linking circulating levels of pyruvate to stroke and cardiovascular disease risk are still to be elucidated.

Acute phase marker glycoprotein acetyls mainly alpha-1 glycoprotein was associated with higher risk of stroke. Our analyses suggested that the association was strongest for the ischemic subtype, for which we found that an increase of one SD in the circulating compound was associated with 13% higher risk of ischemic stroke. Our results confirmed the association of glycoprotein acetyl with ischemic stroke that was observed in individuals within the China Kadoorie Biobank.[11] Circulating levels of glycoprotein acetyls have previously been associated with cardiovascular diseases and dementia but also inflammatory disease, cancer, and mortality.[41-43]

Analyses focused on stroke subtypes revealed the association of essential amino acid phenylalanine with increased risk of ischemic stroke. One SD increase in concentration of phenylalanine was associated with 12% higher risk of ischemic stroke. Phenylalanine is a precursor for tyrosine and catecholamines including dopamine, epinephrine, and norepinephrine. Phenylalanine has previously been associated with increased risk of cardiovascular disease and diabetes.[41, 44, 45] However, the association with phenylalanine remained after adjustment for diabetes. Phenylalanine was not associated with risk of hemorrhagic stroke.

Majority of circulating biomarkers measured by NMR metabolomics technology belong to lipid concentrations and the composition of 14 lipoprotein subparticles. This provides an excellent opportunity for a comprehensive investigation of lipoprotein particles in stroke, as analyses of cholesterol and cholesterol subfractions have shown inconsistent results.[3-5] Previous metabolomic and lipidomic studies of stroke reported associations of various molecular species with ischemic stroke including ceramides, diacylglycerol, docosatrienoic acid, hydroxyeicosatetraenoic acid, hydroxyoctadecadienoic acid, lysophosphatidylcholines (LPC), and triacylglycerols.[46] Furthermore, lipidomics approach provided insights into the metabolism of stroke induced by small vessel disease as associations with

diseaseglucosylceramide, phosphatidylethanolamine, free fatty acid, and triacylglycerol were observed.[47] In our study population we observed associations of cholesterol in medium HDL with decreased risk of stroke and of triglycerides in large and medium LDL particles with increased risk of stroke. None of these lipoprotein measurements were found to be associated with stroke in the China Kadoorie Biobank.[11] However, the association of cholesterol in medium HDL with decreased risk of stroke was previously observed in the cohort of Japanese men and women.[48] Furthermore, previous studies also showed inverse association of coronary heart disease with medium-sized HDL particles.[49] Alterations in specific HDL particles provide additional evidence about the heterogeneity of HDL lipoprotein particles. However, mechanisms through which specific HDL particles might protect against stroke are not well understood yet. Previous studies reported that smaller HDL particles have a larger capacity to remove cholesterol from membranes of peripheral cells such as macrophage foam cell.[48, 50] This potentially antiatherogenic effect is one possible mechanism underlying association of medium HDL cholesterol and stroke. This is also in line with our finding that subjects with higher cholesterol levels in medium HDL particle subclasses may be protected against stroke. Another explanation for HDL protective effect could be its antioxidant function that differs across subfractions. Previous studies showed that antioxidant properties of HDL were enriched in the smallest and densest HDL particles including medium HDL.[51, 52] Similarly,

antiapoptotic and anti-inflammatory properties of HDL may also play a role in protection

against stroke. Sphingosine-1-phosphate (S1P) and other lipids carried on HDL particles are

possible molecular mediators of this anti-inflammatory effect. Previous studies reported

enrichment of S1P in small and dense HDL particles and its inverse correlation endothelial cell apoptosis.[52, 53] Furthermore, specific LDL particles, such as triglycerides in large and medium LDL were associated with increased risk of stroke in our study. Previous studies reported high LDL triglycerides levels to be associated with incident stroke and coronary heart disease.[54] As higher LDL triglycerides levels were also reported to be associated with increased hs-CRP level and white blood cell count one possible mechanism underlying association of LDL triglycerides and cerebrovascular disease may be through inflammation.[54, 55] Interestingly, the China Kadoorie Biobank reported associations of very low-, intermediate-, and low-density lipoproteins with ischemic stroke.[11] However, we were not able to confirm these results in our study population. Lack of replication might be explained by environmental and ethnic differences of studied populations or the confounders adjusted for.

The strengths of our study include its large sample size, prospective study design with detailed data collection over a long period of follow-up, and the same experimental NMR setup was used for metabolite quantification across multiple studies. Our study also has several limitations. With new improved methods available many additional metabolites can be measured, which can be of importance for stroke.[56] Included studies showed heterogeneity in terms of types of samples used across the cohorts and different times and types of sample storage. Even though different sample material was used in cohorts, previously published paper illustrated for NMR-based lipoprotein subclass measures that most metabolite levels are identical in plasma and serum samples, in particular lipid measures.[57] For some metabolites,

the absolute levels are slightly different with constant offset, but the relation is linear. This would suggest that metabolite variability in different types of samples is not problematic for cohort-specific analysis followed by meta-analysis as illustrated in many published articles that have combined analyses of plasma and serum samples in meta-analysis[58, 59] and showed consistent results between the cohorts with different blood specimens. Furthermore, highly consistent epidemiological results have been observed between studies that differ in terms of the time of sample storage. It is well-known that storage slightly modifies the lipoprotein composition of serum/plasma samples, however, these effects are due to biological changes in the actual samples and not related to the analysis method as such. These are minor in comparison to inter-individual differences in metabolite concentrations if the samples are handled according to standard clinical laboratory practices and stored at least in -70 °C. Deelen et al. and van der Lee et al. reported biomarker association with mortality or cognition using same population-based studies as in our project and the metabolite associations show consistency even though there are differences in the times how long the samples have been stored.[58, 60] There was also heterogeneity in the methods to ascertain cases of incident stroke across the cohorts. As most of the cohort studies used electronic health registries, this may have limited sensitivity which subsequently influenced power to identify novel significant associations. Statistical power was reduced in analyses of stroke subtypes as some of the cohorts were unable to distinguish between these. Limited sample size for the analysis of hemorrhagic stroke reduced our ability to detect novel associations for this stroke type. Finally, reported associations may represent signals due to other metabolites or other factors. Therefore, future studies should explore whether these metabolites play a causal role.

To conclude, we found an association between ten metabolites and risk of stroke in 1,791 incident stroke events observed among 38,797 individuals from seven prospective cohort

studies. The biological mechanisms underlying these associations should be subject of further studies.

ACKNOWLEDGMENTS

Rotterdam Study: The Rotterdam Study is supported by the Erasmus MC University Medical Center and Erasmus University Rotterdam; The Netherlands Organisation for Scientific Research (NWO); The Netherlands Organisation for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); The Netherlands Genomics Initiative (NGI); the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. Metabolomics measurements were funded by Biobanking and Biomolecular Resources Research Infrastructure (BBMRI)–NL (184.021.007) and the JNPD under the project PERADES (grant number 733051021, Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease using multiple powerful cohorts, focused Epigenetics and Stem cell metabolomics). This work has been performed as part of the CoSTREAM project (www.costream.eu) and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 667375. This work has been performed as part of the CardioVasculair Onderzoek Nederland (CVON 2012-03).

Whitehall II: The Whitehall II study is supported by the Medical Research Council (MR/R024227/1) and the British Heart Foundation (32334). M. Kivimaki is supported by the Medical Research Council (S011676/1), NordForsk, the Academy of Finland (311492), and Helsinki Institute of Life Science. We thank all of the participating civil service departments and their welfare, personnel, and establishment officers; the British Occupational Health and Safety Agency; the British Council of Civil Service Unions; all participating civil servants in the Whitehall II study; and all members of the Whitehall II study team. The Whitehall II

Study team comprises research scientists, statisticians, study coordinators, nurses, data managers, administrative assistants, and data entry staff, who make the study possible.

FINRISK 1997 & DILGOM: The FINRISK studies have received financial support from the National Institute for Health and Welfare and the Academy of Finland (139635) and the Finnish Foundation for Cardiovascular Research (to V.S). Additional funding has been obtained from the Academy of Finland grant numbers (297338, 307247), the Novo Nordisk Foundation (NNF17OC0026062) the Sigrid Juselius Foundation and University of Oulu Graduate School. MAK is supported by a Senior Research Fellowship from the National Health and Medical Research Council (NHMRC) of Australia (APP1158958). He also works in a unit that is supported by the University of Bristol and UK Medical Research Council (MC_UU_12013/1). The Baker Institute is supported in part by the Victorian Government's Operational Infrastructure Support Program. JK was supported through funds from the Academy of Finland (grant numbers 297338 and 307247) and Novo Nordisk Fonden (grant number NNF17OC0026062).

PROSPER: This work (^{1}H-NMR measurement) was supported by the European Federation of Pharmaceutical Industries Associations (EFPIA), Innovative Medicines Initiative Joint Undertaking, European Medical Information Framework (EMIF) grant #115372 and the European Commission under the Health Cooperation Work Programme of © 2017 The Authors. European Journal of Heart Failure published by John Wiley & Sons Ltd on behalf of European Society of Cardiology. 672 C. Delles et al. the 7th Framework Programme (grant #305507) "Heart 'omics' in AGEing" (HOMAGE). The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032).

EGCUT: This work was supported by Estonian Research Council [IUT20-60, PUT1660 to T.E., PUT1665P to K.F.]; European Union Horizon 2020 [692145]; European Union through

the European Regional Development Fund [2014-2020.4.01.15-0012 GENTRANSMED, 2014-2020.4.2.2] and Estonian Center of Genomics/Roadmap II (project No. 2014-2020.4.01.16-0125).

Framingham Heart Study: This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195 and No. HHSN268201500001I). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. This work was also supported by grants from the National Institute of Neurological Disorders and Stroke (NS017950 and NS100605) the National Institute of Aging (AG033040, AG033193, AG054076, AG049607, AG008122, AG016495, U01-AG049505, and U01-AG052409), and the National Institute of Diabetes and Digestive and Kidney Diseases (DK-HL081572).

APPENDIX 1: Authors

Name	Location	Contribution
Dina Vojinovic, PhD	Erasmus MC University Medical Center, Rotterdam, the Netherlands	Design and conceptualized study; analyzed and interpreted the data; drafted the manuscript for intellectual content
Marita Kalaoja, MSc	University of Oulu and Biocenter Oulu, Oulu, Finland	Analyzed the data; revised the manuscript for intellectual content
Stella Trompet, PhD	Leiden University Medical Center, Leiden, the Netherlands	Analyzed and interpreted the data; revised the manuscript for intellectual content
Krista Fischer, PhD	University of Tartu, Tartu, Estonia	Analyzed the data; revised the manuscript for intellectual content
Martin J. Shipley, MSc	UCL, London, UK	Contributed to acquisition of data; analyzed and interpreted the data; revised the manuscript for intellectual

		content
Shuo Li, MSc	Boston University, Boston, MA, USA	Analyzed the data; revised the manuscript for intellectual content
Aki S. Havulinna, PhD	National Institute for Health and Welfare, Helsinki, Finland	Contributed to acquisition of data; revised the manuscript for intellectual content
Markus Perola, PhD	National Institute for Health and Welfare, Helsinki, Finland	Contributed to acquisition of data; revised the manuscript for intellectual content
Veikko Salomaa, PhD	National Institute for Health and Welfare, Helsinki, Finland	Contributed to acquisition of data; revised the manuscript for intellectual content
Qiong Yang, PhD	Boston University, Boston, MA, USA	Interpreted the data; revised the manuscript for intellectual content
Naveed Sattar, PhD	Faculty of Medicine, Glasgow, UK	Contributed to acquisition of data; revised the manuscript for intellectual content
Pekka Jousilahti, PhD	National Institute for Health and Welfare, Helsinki, Finland	Interpreted the data; revised the manuscript for intellectual content
Najaf Amin, PhD	Erasmus MC University Medical Center, Rotterdam, the Netherlands	Interpreted the data; revised the manuscript for intellectual content
Claudia L. Satizabal, PhD	Boston University School of Medicine, Boston, MA, USA	Interpreted the data; revised the manuscript for intellectual content
Nele Taba, MSc	University of Tartu, Tartu, Estonia	Interpreted the data; revised the manuscript for intellectual content
Behnam Sabayan, PhD	Northwestern University, Chicago, IL, USA	Revised the manuscript for intellectual content
Ramachandran S. Vasan, MD	Framingham, MA, USA	Interpreted the data; revised the manuscript for intellectual content
M. Arfan Ikram, PhD	Erasmus MC University Medical Center, Rotterdam, the	Interpreted the data; revised the

	Netherlands	manuscript for intellectual content
David J. Stott, PhD	University of Glasgow, UK	Contributed to acquisition of data; revised the manuscript for intellectual content
Mika Ala-Korpela, PhD	University of Oulu and Biocenter Oulu, Oulu, Finland	Interpreted the data; revised the manuscript for intellectual content
J. Wouter Jukema, PhD	Leiden University Medical Center, Leiden, the Netherlands	Contributed to acquisition of data; revised the manuscript for intellectual content
Sudha Seshadri, MD	Boston University School of Medicine, Boston, MA, USA	Interpreted the data; revised the manuscript for intellectual content
Johannes Kettunen, PhD	University of Oulu and Biocenter Oulu, Oulu, Finland	Interpreted the data; revised the manuscript for intellectual content
Mika Kivimaki, FMedSciP	UCL, London, UK	Contributed to acquisition of data; interpreted the data; revised the manuscript for intellectual content
Tonu Esko, PhD	University of Tartu, Tartu, Estonia	Interpreted the data; revised the manuscript for intellectual content
Cornelia M. van Duijn, PhD	University of Oxford, Oxford, UK	Design and conceptualized study; interpreted data; drafted the manuscript for intellectual content

REFERENCES

1. Mozaffarian. Heart Disease and Stroke Statistics-2015 Update: A Report From the American Heart Association (vol 131, pg e29, 2015). Circulation 2015;131:E535-E535.
2. Andersen KK, Olsen TS, Dehlendorff C, Kammersgaard LP. Hemorrhagic and ischemic strokes compared: stroke severity, mortality, and risk factors. Stroke 2009;40:2068-2072.
3. Goldstein LB, Bushnell CD, Adams RJ, et al. Guidelines for the primary prevention of stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2011;42:517-584.
4. Zhang Y, Tuomilehto J, Jousilahti P, Wang Y, Antikainen R, Hu G. Total and high-density lipoprotein cholesterol and stroke risk. Stroke 2012;43:1768-1774.
5. Amarenco P, Labreuche J, Touboul PJ. High-density lipoprotein-cholesterol and risk of stroke and carotid atherosclerosis: a systematic review. Atherosclerosis 2008;196:489-496.

6. Saenger AK, Christenson RH. Stroke biomarkers: progress and challenges for diagnosis, prognosis, differentiation, and treatment. Clin Chem 2010;56:21-33.
7. Makin SD, Doubal FN, Dennis MS, Wardlaw JM. Clinically Confirmed Stroke With Negative Diffusion-Weighted Imaging Magnetic Resonance Imaging: Longitudinal Study of Clinical Outcomes, Stroke Recurrence, and Systematic Review. Stroke 2015;46:3142-3148.
8. Jung JY, Lee HS, Kang DG, et al. 1H-NMR-based metabolomics study of cerebral infarction. Stroke 2011;42:1282-1288.
9. Wang D, Kong J, Wu J, Wang X, Lai M. GC-MS-based metabolomics identifies an amino acid signature of acute ischemic stroke. Neurosci Lett 2017;642:7-13.
10. Lee Y, Khan A, Hong S, Jee SH, Park YH. A metabolomic study on high-risk stroke patients determines low levels of serum lysine metabolites: a retrospective cohort study. Mol Biosyst 2017;13:1109-1120.
11. Holmes MV, Millwood IY, Kartsonaki C, et al. Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. J Am Coll Cardiol 2018;71:620-632.
12. Wieberdink RG, Ikram MA, Hofman A, Koudstaal PJ, Breteler MMB. Trends in stroke incidence rates and stroke risk factors in Rotterdam, the Netherlands from 1990 to 2008. European Journal of Epidemiology 2012;27:287-295.
13. Kivimaki M, Batty GD, Singh-Manoux A, Britton A, Brunner EJ, Shipley MJ. Validity of Cardiovascular Disease Event Ascertainment Using Linkage to UK Hospital Records. Epidemiology 2017;28:735-739.
14. Britton A, Milne B, Butler T, et al. Validating self-reported strokes in a longitudinal UK cohort study (Whitehall II): Extracting information from hospital medical records versus the Hospital Episode Statistics database. Bmc Med Res Methodol 2012;12.
15. Shepherd J, Blauw GJ, Murphy MB, et al. The design of a Prospective Study of Pravastatin in the Elderly at Risk (PROSPER). Am J Cardiol 1999;84:1192-1197.
16. Carandang R, Seshadri S, Beiser A, et al. Trends in incidence, lifetime risk, severity, and 30-day mortality of stroke over the past 50 years. JAMA 2006;296:2939-2946.
17. Seshadri S, Beiser A, Kelly-Hayes M, et al. The lifetime risk of stroke: estimates from the Framingham Study. Stroke 2006;37:345-350.
18. Soininen P, Kangas AJ, Wurtz P, Suna T, Ala-Korpela M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. Circ-Cardiovasc Gene 2015;8:192-206.
19. Soininen P, Kangas AJ, Wurtz P, et al. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 2009;134:1781-1785.
20. Freedman DS, Otvos JD, Jeyarajah EJ, et al. Sex and age differences in lipoprotein subclasses measured by nuclear magnetic resonance spectroscopy: the Framingham Study. Clin Chem 2004;50:1189-1200.
21. Kathiresan S, Manning AK, Demissie S, et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. BMC Med Genet 2007;8 Suppl 1:S17.
22. Delles C, Rankin NJ, Boachie C, et al. Nuclear magnetic resonance-based metabolomics identifies phenylalanine as a novel predictor of incident heart failure hospitalisation: results from PROSPER and FINRISK 1997. Eur J Heart Fail 2018;20:663-673.
23. Akbaraly T, Wurtz P, Singh-Manoux A, et al. Association of circulating metabolites with healthy diet and risk of cardiovascular disease: analysis of two cohort studies. Sci Rep-Uk 2018;8.
24. Kettunen J, Demirkan A, Wurtz P, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 2016;7:11122.

25. van der Lee SJ, Teunissen CE, Pool R, et al. Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. Alzheimers Dement 2018.
26. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190-2191.
27. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-560.
28. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity 2005;95:221-227.
29. Ignesti G, Pino R, Banchelli G, Ferrali C, Pirisino R, Raimondi L. Increased desensitization by picomolar phorbol ester of the endothelium-mediated effect of histamine in the perfused rat mesenteric bed. Inflamm Res 1996;45:171-175.
30. Wang L, Liu S, Yang W, et al. Plasma Amino Acid Profile in Patients with Aortic Dissection. Sci Rep 2017;7:40146.
31. Toba H, Nakamori A, Tanaka Y, et al. Oral L-histidine exerts antihypertensive effects via central histamine H3 receptors and decreases nitric oxide content in the rostral ventrolateral medulla in spontaneously hypertensive rats. Clin Exp Pharmacol Physiol 2010;37:62-68.
32. Tuttle KR, Milton JE, Packard DP, Shuler LA, Short RA. Dietary amino acids and blood pressure: a cohort study of patients with cardiovascular disease. Am J Kidney Dis 2012;59:803-809.
33. Jennings A, MacGregor A, Welch A, Chowienczyk P, Spector T, Cassidy A. Amino Acid Intakes Are Inversely Associated with Arterial Stiffness and Central Blood Pressure in Women. J Nutr 2015;145:2130-2138.
34. Watanabe M, Suliman ME, Qureshi AR, et al. Consequences of low plasma histidine in chronic kidney disease patients: associations with inflammation, oxidative stress, and mortality. Am J Clin Nutr 2008;87:1860-1866.
35. Feng RN, Niu YC, Sun XW, et al. Histidine supplementation improves insulin resistance through suppressed inflammation in obese women with the metabolic syndrome: a randomised controlled trial. Diabetologia 2013;56:985-994.
36. Adachi N, Liu K, Arai T. Prevention of brain infarction by postischemic administration of histidine in rats. Brain Res 2005;1039:220-223.
37. Liao RJ, Jiang L, Wang RR, et al. Histidine provides long-term neuroprotection after cerebral ischemia through promoting astrocyte migration. Sci Rep-Uk 2015;5.
38. Gray LR, Tompkins SC, Taylor EB. Regulation of pyruvate metabolism and human disease. Cell Mol Life Sci 2014;71:2577-2604.
39. Wang Q, van Hoecke M, Tang XN, et al. Pyruvate protects against experimental stroke via an anti-inflammatory mechanism. Neurobiol Dis 2009;36:223-231.
40. Ryou MG, Liu R, Ren M, Sun J, Mallet RT, Yang SH. Pyruvate protects the brain against ischemia-reperfusion injury by activating the erythropoietin signaling pathway. Stroke 2012;43:1101-1107.
41. Wurtz P, Havulinna AS, Soininen P, et al. Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts. Circulation 2015;131:774-785.
42. Connelly MA, Gruppen EG, Otvos JD, Dullaart RPF. Inflammatory glycoproteins in cardiometabolic disorders, autoimmune diseases and cancer. Clin Chim Acta 2016;459:177-186.
43. Lawler PR, Akinkuolie AO, Chandler PD, et al. Circulating N-Linked Glycoprotein Acetyls and Longitudinal Mortality Risk. Circ Res 2016;118:1106-1115.
44. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. Nat Med 2011;17:448-453.

45. Floegel A, Stefan N, Yu ZH, et al. Identification of Serum Metabolites Associated With Risk of Type 2 Diabetes Using a Targeted Metabolomic Approach. Diabetes 2013;62:639-648.
46. Au A. Metabolomics and Lipidomics of Ischemic Stroke. Adv Clin Chem 2018;85:31-69.
47. Yang L, Lv P, Ai WP, et al. Lipidomic analysis of plasma in patients with lacunar infarction using normal-phase/reversed-phase two-dimensional liquid chromatography-quadrupole time-of-flight mass spectrometry. Anal Bioanal Chem 2017;409:3211-3222.
48. Chei CL, Yamagishi K, Kitamura A, et al. High-density Lipoprotein Subclasses and Risk of Stroke and its Subtypes in Japanese Population The Circulatory Risk in Communities Study. Stroke 2013;44:327-333.
49. Ditah C, Otvos J, Nassar H, Shaham D, Sinnreich R, Kark JD. Small and medium sized HDL particles are protectively associated with coronary calcification in a cross-sectional population-based sample. Atherosclerosis 2016;251:124-131.
50. Du XM, Kim MJ, Hou LM, et al. HDL Particle Size Is a Critical Determinant of ABCA1-Mediated Macrophage Cellular Cholesterol Export. Circ Res 2015;116:1133-1142.
51. Kontush A, Chantepie S, Chapman MJ. Small, dense HDL particles exert potent protection of atherogenic LDL against oxidative stress. Arterioscl Throm Vas 2003;23:1881-1888.
52. Kim DS, Li YK, Bell GA, et al. Concentration of Smaller High-Density Lipoprotein Particle (HDL-P) Is Inversely Correlated With Carotid Intima Media Thickening After Confounder Adjustment: The Multi Ethnic Study of Atherosclerosis (MESA). J Am Heart Assoc 2016;5.
53. Kontush A, Therond P, Zerrad A, et al. Preferential sphingosine-1-phosphate enrichment and sphingomyelin depletion are key features of small dense HDL3 particles - Relevance to antiapoptotic and antioxidative activities. Arterioscl Throm Vas 2007;27:1843-1849.
54. Saeed A, Feofanova EV, Yu B, et al. Remnant-Like Particle Cholesterol, Low-Density Lipoprotein Triglycerides, and Incident Cardiovascular Disease. J Am Coll Cardiol 2018;72:156-169.
55. Marz W, Scharnagl H, Winkler K, et al. Low-density lipoprotein triglycerides associated with low-grade systemic inflammation, adhesion molecules, and angiographic coronary artery disease - The Ludwigshafen Risk and Cardiovascular Health study. Circulation 2004;110:3068-3074.
56. Wishart DS, Jewison T, Guo AC, et al. HMDB 3.0--The Human Metabolome Database in 2013. Nucleic Acids Res 2013;41:D801-807.
57. Jimenez B, Holmes E, Heude C, et al. Quantitative Lipoprotein Subclass and Low Molecular Weight Metabolite Analysis in Human Serum and Plasma by (1)H NMR Spectroscopy in a Multilaboratory Trial. Anal Chem 2018;90:11962-11971.
58. Deelen J, Kettunen J, Fischer K, et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. Nat Commun 2019;10:3346.
59. Sliz E, Kettunen J, Holmes MV, et al. Metabolomic consequences of genetic inhibition of PCSK9 compared with statin treatment. Circulation 2018;138:2499-2512.
60. van der Lee SJ, Teunissen CE, Pool R, et al. Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. Alzheimers Dement 2018;14:707-722.

Table 1. Descriptive statistics of study population

	Rotterdam Study		Whitehall II[b]		Finrisk97		DILGOM		PROSPER		EGCUT		FHS	
Variable[a]	Incident cases	Controls	Incident cases	Controls	Incident cases	Controls	Incident cases	Controls	Incident cases	Controls	Incident cases	Controls	Incident cases	Controls
N	257	2308	197	5792	474	6384	107	4424	197	4627	308	10268	251	3203
Age (years)	76.9 (6.2)	75.0 (6.1)	59.4 (5.9)	55.6 (6)	59.6 (10.4)	47.0 (12.9)	62.0 (10.4)	51.9 (13.5)	75.9 (3.7)	75.2 (3.3)	66.3 (12.5)	44.5 (17.1)	58.1 (9.0)	51.7 (10.1)
Women	54.1%	58%	25.4%	29.1%	38.6%	52.6%	42.1%	53.7%	54%	52.2%	54.9%	63.3%	47.4%	51.4%
Current Smoking	15.2%	13%	15.7%	9.4%	24.5%	23.7%	21.5%	17.4%	28.9%	27.1%	17.5%	29.9%	24.9%	24.6%
Diabetes	17.9%	14.3%	8.6%	4.4%	16.9%	4.9%	15.9%	8.9%	18.3%	10.6%	35.4%	7.7%	15.9%	5.0%
Hypertension	85.6%	81.0%	41.1%	28.0%	48.9%	21%	43%	16.5%	58.9%	62.5%	66.2%	24.4%	60.2%	34.2%
Systolic blood pressure (mmHg)	156.8 (23.9)	151.4 (20.1)	127.4 (16.3)	122.9 (16.5)	147.7 (22.3)	134.7 (19.2)	149.8 (23.7)	136.4 (20.2)	157.1 (21.9)	154.5 (21.8)	142.8 (18.8)	125.7 (16.9)	137.8 (20.8)	126.2 (18.5)
Diastolic blood pressure (mmHg)	79.7 (12.4)	79.2 (11.1)	78.3 (10.2)	77.5 (10.5)	86.1 (11.9)	81.9 (11.2)	83.1 (13.6)	79.3 (11.0)	84.6 (11.8)	83.7 (11.4)	83.4 (10.9)	77.6 (10.7)	81.6 (10.4)	78.9 (9.9)
Antihypertensive medication[c]	51%	47.1%	21.8%	12.3%	27.4%	11.5%	34.6%	22.1%	70.6%	74.4%	69.5%	24.3%	36.3%	16.4%
BMI (kg/m2)	27.2 (3.5)	27.4 (4.2)	26.2 (4.2)	26.0 (3.9)	28.4 (4.8)	26.5 (4.5)	28.0 (5.0)	27.2 (4.8)	26.5 (4.1)	26.9 (4.2)	29.1 (5.7)	26.4 (5.4)	27.6 (5.1)	26.7 (4.8)
Follow-up time (years)	5.7 (3.5)	9.8 (3.5)	12.5 (4.9)	18.2 (3.0)	15.0 (4.2)	16.9 (3.0)	7.25 (1.5)	7.75 (0.7)	1.9 (1.0)	3.3 (0.5)	6.9 (3.1)	8.9 (1.8)	14.7 (7.0)	22.4 (6.0)
Total cholesterol (mmol/l)	5.5 (1.0)	5.6 (1.0)	5.8 (1.1)	5.9 (1.1)	5.8 (1.1)	5.52 (1.1)	5.23 (1.0)	5.28 (1)	5.64 (0.9)	5.68 (0.9)	6.0 (1.2)	5.7 (1.2)	5.6 (1.1)	5.3 (1.0)
HDL cholesterol (mmol/l)	1.4 (0.4)	1.5 (0.4)	1.5 (0.4)	1.5 (0.4)	1.3 (0.3)	1.4 (0.4)	1.4 (0.4)	1.4 (0.4)	1.3 (0.3)	1.3 (0.4)	1.5 (0.4)	1.6 (0.5)	1.2 (0.4)	1.3 (0.4)
LDL cholesterol (mmol/l)	NA	NA	3.8 (1.0)	3.9 (0.9)	3.7 (0.9)	3.5 (0.9)	3.1 (0.8)	3.2 (0.9)	3.8 (0.8)	3.8 (0.8)	2.5 (0.7)	2.3 (0.6)	3.6 (1.0)	3.4 (0.9)
Triglycerides (mmol/l)	NA	NA	1.3 (0.8)	1.4 (0.9)	1.8 (1.1)	1.5 (1.0)	1.5 (0.8)	1.4 (0.9)	1.6 (0.7)	1.5 (0.7)	1.9 (1.0)	1.6 (0.9)	1.7 (1.2)	1.4 (1.2)
Lipid lowering medication	21.4%	20.6%	5.1%	3.0%	7.8%	3.1%	25.2%	14.7%	52.3%	49.5%	13.6%	4.7%	6.0%	3.7%
Coronary Heart Disease	13.6%	10.8%	11.7%	5.9%	9.1%	1.9%	5.6%	2.9%	16.8%	13.1%	35.1%	9.2%	12.4%	5.7%
Stroke														
Hemorrhagic	32 (12.5%)	-	48 (24.4%)	-	69 (14.6%)	-	23 (21.5%)	-	-	-	45 (14.6%)	-	30 (12%)	-
Ischemic	183 (71.2%)	-	126 (64.0%)	-	405 (85.4%)	-	84 (78.5%)	-	-	-	261 (84.7%)	-	219 (87.3%)	-
Not defined	42 (16.3%)	-	23 (11.6%)	-	-	-	-	-	-	-	11 (3.6%)	-	2 (0.8%)	-

[a]Values are means ± standard deviation for continuous variables and percentages for dichotomous variables;

[b]While all other cohorts included participants of European ancestry, 87.3% of Whitehall II study cases were of European ancestry, 6.6% of Asian, 5.1% of African American, and 1% of other.

[c]The percentage of participants taking antihypertensive medication is greater than the percentage of hypertension in some cohorts as study participants might use anti-hypertensive medication for other reasons than hypertension, like beta-blockers for heart rhythm problems.

Table 2. Results of association analysis showing the significant metabolite associations with overall incident stroke.

	Model 1						Model 2					
Metabolite	**N**	**Ncases**	**HR**	**CI**	**I^{2**}**	**P**	**N**	**Ncases**	**HR**	**CI**	**I^2**	**P**
Phenylalanine	35091	1527	1.11	1.06;1.17	44.4	4.88E-05	35036	1524	1.08	1.03;1.14	17	3.36E-03
Histidine*	35017	1526	0.89	0.84;0.93	40.7	7.94E-06	34962	1523	0.90	0.85;0.94	30.2	4.45E-05
plasma-ApoA1	35107	1529	0.91	0.86;0.96	18.3	7.14E-04	35052	1526	0.94	0.88;0.99	30.3	1.79E-02
HDL-cholesterol	35107	1529	0.89	0.84;0.94	65.8	2.89E-05	35052	1526	0.92	0.87;0.97	68.3	3.20E-03
HDL2-cholesterol*	35107	1529	0.88	0.84;0.93	66.3	9.13E-06	35052	1526	0.91	0.87;0.97	69.1	1.41E-03
IDL-triglycerides	38561	1780	1.10	1.05;1.16	38.6	6.06E-05	38494	1775	1.07	1.02;1.12	44	9.91E-03
LDL-triglycerides	35107	1529	1.12	1.06;1.18	2.5	3.93E-05	35052	1526	1.08	1.03;1.14	20.1	2.47E-03
Glucose	34980	1524	1.15	1.10;1.20	13.7	7.81E-11	34925	1521	1.06	1.01;1.11	0	1.87E-02
Lactate	35100	1529	1.12	1.07;1.18	56.7	1.11E-05	35045	1526	1.08	1.02;1.13	29.1	5.09E-03
Pyruvate*	24423	1205	1.13	1.08;1.18	48.6	1.37E-07	24368	1202	1.09	1.04;1.14	13.6	7.45E-04
Glycoprotein acetyls*	35101	1529	1.15	1.09;1.21	43.2	1.25E-07	35046	1526	1.09	1.03;1.15	38.9	1.27E-03
HDL-diametar	35107	1529	0.89	0.84;0.94	49	3.05E-05	35052	1526	0.92	0.87;0.98	61.6	6.73E-03
S-HDL-triglycerides	35108	1529	1.11	1.06;1.17	48.4	6.80E-05	35053	1526	1.07	1.01;1.12	53.5	1.97E-02
M-HDL-cholesterol*	38560	1780	0.89	0.85;0.94	56	2.07E-05	38493	1775	0.92	0.87;0.97	50.4	1.35E-03
M-HDL-cholesterol esters	35106	1529	0.90	0.85;0.95	60.1	2.05E-04	35051	1526	0.92	0.87;0.97	57.8	3.73E-03
M-HDL-free cholesterol	35106	1529	0.91	0.86;0.96	56	7.33E-04	35051	1526	0.93	0.88;0.98	50.1	8.24E-03
L-HDL-cholesterol	38555	1780	0.89	0.84;0.94	63	2.13E-05	38488	1775	0.92	0.88;0.98	66.6	5.50E-03
L-HDL-cholesterol esters	35101	1529	0.90	0.84;0.95	69.1	2.03E-04	35046	1526	0.93	0.88;0.99	72.4	1.37E-02
L-HDL-free cholesterol	35101	1529	0.89	0.84;0.94	67.6	1.25E-04	35046	1526	0.92	0.87;0.98	70.8	9.96E-03
L-HDL-total lipids	35101	1529	0.90	0.85;0.95	69	2.12E-04	35046	1526	0.93	0.88;0.99	71.8	1.70E-02
L-HDL-phospholipids	35101	1529	0.90	0.85;0.96	71.1	6.29E-04	35046	1526	0.94	0.89;1.00	72.2	3.49E-02
L-HDL concentration	35101	1529	0.90	0.85;0.96	69.4	8.53E-04	35046	1526	0.94	0.89;1.00	71.2	4.21E-02
XL-HDL-free cholesterol	35099	1527	0.91	0.86;0.96	0	8.31E-04	35044	1524	0.94	0.89;1.00	11.7	3.55E-02
S-LDL-triglycerides	35107	1529	1.10	1.05;1.16	44.6	1.97E-04	35052	1526	1.07	1.01;1.12	51.3	1.58E-02
L-LDL-triglycerides*	35107	1529	1.12	1.06;1.17	12.2	3.00E-05	35052	1526	1.09	1.03;1.14	27.5	1.67E-03
M-LDL-triglycerides*	35106	1529	1.12	1.06;1.18	0	1.68E-05	35051	1526	1.09	1.03;1.14	1.4	1.19E-03
XL-VLDL-triglycerides	38284	1769	1.09	1.04;1.14	75.9	1.56E-04	38217	1764	1.05	1.00;1.10	73.9	4.66E-02

Abbreviations: N - Total samples size; Ncases - Number of cases; HR - Hazard Ratio; 95% CI - 95% confidence interval; I^2 - Heterogeneity parameter; P - *p*-value for association of metabolite and stroke; Model 1 - adjustment for age, gender, BMI, lipid-lowering medication and study-specific covariates if needed; Model 2- additional adjustment for smoking status, diabetes, and hypertension;

*****Associations that surpassed significance threshold in model 2;

**Random effect meta-analysis was performed in the case of statistical heterogeneity, defined as $I^2 > 50\%$. Subsequently, the correlation between effect estimates (hazard ratio (HR)) derived from fixed and random-effect meta-analysis was checked. The correlation coefficient between these two estimates was 0.99.

Table 3. Associations for incident stroke events when classified by stroke type. Metabolites that showed significant association with overall incident stroke in Table 2 were included.

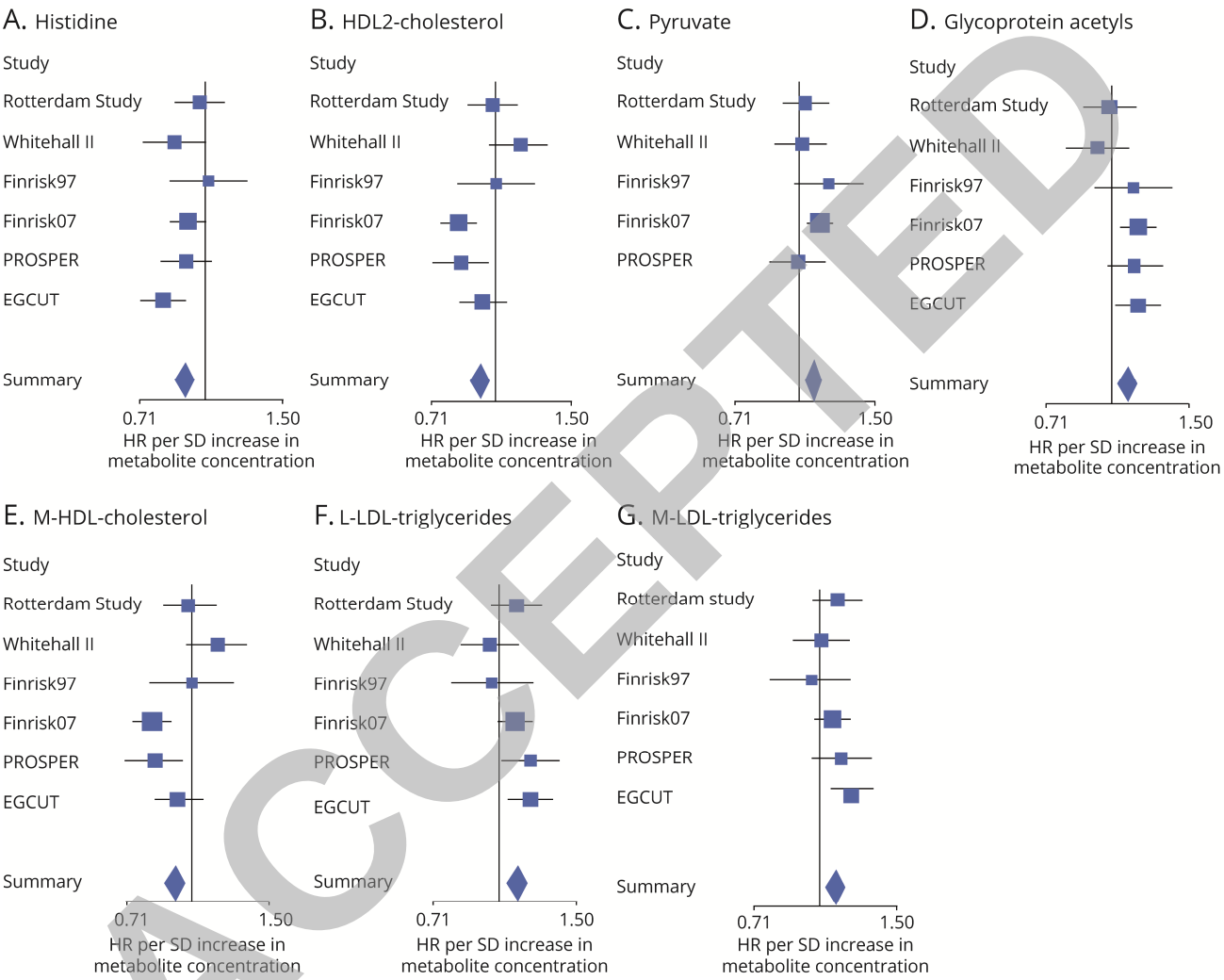
		Model 1					Model 2				
Metabolite	**Type**	N	Ncases	HR	CI	P	N	Ncases	HR	CI	P
Phenylalanine	hemorrhagic	30144	214	0.94	0.81;1.09	3.90E-01	30092	214	0.91	0.78;1.05	2.00E-01
	ischemic*	30290	1051	1.16	1.09;1.23	**3.15E-06**	30236	1049	1.12	1.05;1.19	**4.13E-04**
Histidine	hemorrhagic	30070	214	0.95	0.82;1.1	4.78E-01	30018	214	0.96	0.83;1.11	6.16E-01
	ischemic*	30216	1050	0.88	0.82;0.94	**1.18E-04**	30162	1048	0.89	0.84;0.95	**4.94E-04**
Apolipoprotein A1	hemorrhagic	30155	216	1.03	0.89;1.19	7.32E-01	30103	216	1.04	0.9;1.2	6.17E-01
	ischemic	30301	1051	0.89	0.83;0.95	**5.87E-04**	30247	1049	0.92	0.86;0.98	1.43E-02
HDL-cholesterol	hemorrhagic	30155	216	1.04	0.9;1.21	5.63E-01	30103	216	1.07	0.92;1.24	3.97E-01
	ischemic	30301	1051	0.86	0.81;0.92	**1.82E-05**	30247	1049	0.9	0.84;0.96	1.89E-03
HDL2-cholesterol	hemorrhagic	30155	216	1.04	0.9;1.21	5.75E-01	30103	216	1.07	0.92;1.24	3.90E-01
	ischemic*	30301	1051	0.85	0.8;0.91	**2.85E-06**	30247	1049	0.89	0.83;0.95	**5.29E-04**
IDL-triglycerides	hemorrhagic	33609	246	0.92	0.81;1.06	2.63E-01	33545	246	0.89	0.78;1.02	1.02E-01
	ischemic	33755	1270	1.13	1.07;1.2	**6.91E-06**	33689	1266	1.09	1.03;1.15	2.01E-03
LDL-triglycerides	hemorrhagic	30155	216	1.02	0.88;1.18	8.27E-01	30103	216	0.99	0.85;1.14	8.62E-01
	ischemic	30301	1051	1.14	1.07;1.21	**2.89E-05**	30247	1049	1.1	1.04;1.17	1.82E-03
Glucose	hemorrhagic	30033	214	1.13	0.99;1.28	7.07E-02	29981	214	1.09	0.96;1.24	2.20E-01
	ischemic	30179	1048	1.17	1.12;1.23	**5.37E-11**	30125	1046	1.07	1.01;1.13	2.13E-02
Lactate	hemorrhagic	30153	216	1.06	0.92;1.22	4.04E-01	30101	216	1.04	0.9;1.19	6.13E-01
	ischemic	30299	1051	1.16	1.09;1.24	**1.11E-06**	30245	1049	1.1	1.04;1.17	1.98E-03
Pyruvate	hemorrhagic	19481	167	0.99	0.84;1.16	8.70E-01	19429	167	0.96	0.81;1.12	5.94E-01
	ischemic*	19627	778	1.17	1.11;1.23	**2.86E-10**	19573	776	1.13	1.07;1.19	**1.93E-05**
Glycoprotein acetyls	hemorrhagic	30154	216	1.02	0.88;1.18	8.06E-01	30102	216	0.96	0.83;1.11	6.28E-01
	ischemic*	30300	1051	1.2	1.13;1.28	**8.55E-09**	30246	1049	1.13	1.06;1.2	**2.17E-04**
Mean diameter of HDL	hemorrhagic	30155	216	1.03	0.89;1.2	6.98E-01	30103	216	1.07	0.92;1.24	4.08E-01
	ischemic	30301	1051	0.86	0.8;0.92	**1.28E-05**	30247	1049	0.9	0.84;0.96	2.93E-03
S-HDL-triglycerides	hemorrhagic	30156	216	1	0.86;1.15	9.71E-01	30104	216	0.96	0.83;1.11	5.84E-01
	ischemic	30302	1051	1.14	1.08;1.22	**1.99E-05**	30248	1049	1.09	1.02;1.16	8.32E-03
M-HDL-cholesterol	hemorrhagic	33608	246	1	0.87;1.15	9.88E-01	33544	246	1.02	0.88;1.17	8.27E-01
	ischemic	33754	1270	0.88	0.83;0.93	**3.11E-05**	33688	1266	0.91	0.85;0.97	1.95E-03
M-HDL-cholesterol esters	hemorrhagic	30154	216	0.99	0.86;1.14	8.84E-01	30102	216	1	0.87;1.16	9.72E-01

	ischemic	30300	1051	0.89	0.83;0.95	**4.60E-04**	30246	1049	0.91	0.86;0.98	6.99E-03
M-HDL-free cholesterol	hemorrhagic	30154	216	1.02	0.88;1.18	8.25E-01	30102	216	1.02	0.88;1.18	7.79E-01
	ischemic	30300	1051	0.9	0.84;0.96	1.75E-03	30246	1049	0.92	0.86;0.98	1.56E-02
L-HDL-cholesterol	hemorrhagic	33603	246	1.07	0.93;1.24	3.19E-01	33539	246	1.11	0.96;1.28	1.56E-01
	ischemic*	33749	1270	0.85	0.8;0.91	**2.03E-06**	33683	1266	0.89	0.84;0.95	**9.00E-04**
L-HDL-cholesterol esters	hemorrhagic	30149	216	1.07	0.92;1.24	3.88E-01	30097	216	1.1	0.95;1.28	2.13E-01
	ischemic	30295	1051	0.86	0.8;0.92	**4.28E-05**	30241	1049	0.9	0.84;0.97	3.59E-03
L-HDL-free cholesterol	hemorrhagic	30149	216	1.09	0.93;1.26	2.81E-01	30097	216	1.12	0.96;1.3	1.44E-01
	ischemic*	30295	1051	0.85	0.79;0.91	**1.04E-05**	30241	1049	0.89	0.82;0.95	**1.33E-03**
L-HDL-total lipids	hemorrhagic	30149	216	1.06	0.91;1.23	4.71E-01	30097	216	1.09	0.93;1.27	2.77E-01
	ischemic	30295	1051	0.87	0.81;0.93	**5.91E-05**	30241	1049	0.91	0.84;0.97	6.00E-03
L-HDL-concentration	hemorrhagic	30149	216	1.09	0.94;1.27	2.73E-01	30097	216	1.12	0.96;1.3	1.44E-01
	ischemic	30295	1051	0.87	0.81;0.93	**1.30E-04**	30241	1049	0.91	0.85;0.98	1.04E-02
L-HDL-phospholipids	hemorrhagic	30149	216	1.08	0.93;1.26	3.21E-01	30097	216	1.11	0.95;1.29	1.85E-01
	ischemic	30295	1051	0.87	0.81;0.93	**9.62E-05**	30241	1049	0.91	0.85;0.98	9.32E-03
XL-HDL-free cholesterol	hemorrhagic	30147	216	1.07	0.93;1.24	3.52E-01	30095	216	1.09	0.94;1.26	2.38E-01
	ischemic	30293	1049	0.88	0.82;0.94	**3.46E-04**	30239	1047	0.92	0.86;0.99	1.75E-02
S-LDL-triglycerides	hemorrhagic	30155	216	1.00	0.87;1.16	9.84E-01	30103	216	0.97	0.84;1.12	6.99E-01
	ischemic	30301	1051	1.12	1.06;1.19	**1.18E-04**	30247	1049	1.08	1.02;1.15	1.10E-02
L-LDL-triglycerides	hemorrhagic	30155	216	1.01	0.87;1.17	9.12E-01	30103	216	0.98	0.85;1.13	7.87E-01
	ischemic	30301	1051	1.13	1.07;1.2	**4.39E-05**	30247	1049	1.1	1.03;1.17	2.20E-03
M-LDL-triglycerides	hemorrhagic	30154	216	1.04	0.9;1.2	5.70E-01	30102	216	1.02	0.88;1.17	8.35E-01
	ischemic	30300	1051	1.14	1.07;1.21	**2.84E-05**	30246	1049	1.1	1.04;1.17	1.80E-03
XL-VLDL-triglycerides	hemorrhagic	33352	242	0.98	0.85;1.12	7.66E-01	33288	242	0.96	0.83;1.1	5.23E-01
	ischemic	33499	1263	1.12	1.06;1.18	**2.00E-05**	33433	1259	1.07	1.01;1.13	1.41E-02

Abbreviations: N – total sample size; Ncases – Number of cases; HR - Hazard Ratio; 95% CI - 95% confidence interval; P - p-value for association of metabolite and stroke; Model 1 - adjustment for age, gender, BMI, lipid-lowering medication and ethnicity if needed; Model 2 - additional adjustment for smoking status, diabetes, and hypertension; Associations that passed threshold for multiple testing are shown in bold; *Associations that surpassed significance threshold in model 2.

Figure legend

Figure 1. Forest plots for metabolites associated with overall incident stroke. The associations were significant after adjustment for age, gender, BMI, lipid-lowering medication, smoking status, diabetes, and hypertension.



Association of circulating metabolites in plasma or serum and risk of stroke: Meta-analysis from seven prospective cohorts

Dina Vojinovic, Marita Kalaoja, Stella Trompet, et al.

Neurology published online December 2, 2020
DOI 10.1212/WNL.0000000000011236