

# **Supplemental Material**

## **Metabolic markers of MRI-confirmed lacunar stroke: observational and Mendelian randomization analyses**

Wen Hui Ng, Elnara Aghakishiyeva, Hugh S. Markus, and Eric L. Harshfield

### **Table of Contents**

#### **Appendices**

**eAppendix 1. Members of DNA Lacunar 1 and 2 Coordinating Centres and Recruiting Centres.**

#### **Supplemental Methods**

##### **eMethods**

- 1. Further Details of Study Population.**
- 2. Further Details of Causal Inference Analyses.**
- 3. Further Details of Interaction Analyses.**

#### **Supplemental References**

##### **eReferences**

## **Supplemental Figures**

**eFigure 1. Circular bar plot showing the results of cross-sectional univariate analyses for the association between 250 metabolic biomarkers and lacunar stroke**

**eFigure 2. Volcano plot showing the results of cross-sectional univariate analyses for the association between 250 metabolic biomarkers and lacunar stroke.**

**eFigure 3. Circular bar plot showing the results of cross-sectional univariate analyses for the association of 250 metabolic biomarkers with ILI cases versus controls.**

**eFigure 4. Circular bar plot showing the results of cross-sectional univariate analyses for the association of 250 metabolic biomarkers with MLI/LA cases versus controls.**

**eFigure 5. Heat map showing the results of cross-sectional univariate analyses for the associations of lipoprotein subclasses and relative lipoprotein lipid concentrations with lacunar stroke, MRI markers, and cognitive impairment**

**eFigure 6. Scree plot and loadings plot from principal component analysis of metabolite levels.**

**eFigure 7. Forest plot showing the results of cross-sectional univariate analyses for the association of the first twenty-two principal components of metabolite levels with lacunar stroke status, ILI vs controls, and MLI/LA vs controls.**

**eFigure 8. Heat map summarizing the causal estimates from Mendelian randomization analyses for the association of metabolic biomarkers with lacunar stroke and MRI markers of SVD using relaxed clumping threshold.**

**eFigure 9. Scatter plots of genetic associations of metabolic levels with lacunar stroke and neuroimaging markers of SVD for Mendelian randomization association analyses with significant or suggestive associations using relaxed clumping threshold.**

**eFigure 10. Forest plot summarizing the significant estimates from reverse Mendelian randomization analyses for the association of lacunar stroke and MRI markers of SVD with metabolic biomarkers.**

## **Supplemental Tables**

**eTable 1. Metabolic measures analysed using nuclear magnetic resonance spectroscopy in DNA Lacunar 1 and 2.**

**eTable 2. Comparison of characteristics of participants from DNA Lacunar 1 and 2 included in analysis and excluded from analysis.**

**eTable 3. All results for association of lacunar stroke per 1-SD higher baseline metabolite levels with adjustment for age and sex.**

**eTable 4. All results for association of lacunar stroke per 1-SD higher baseline metabolite levels with adjustment for age, sex, and additional vascular risk factors and potential confounders.**

**eTable 5. All results for association of lacunar stroke per 1-SD higher baseline metabolite levels with adjustment for age, sex, additional vascular risk factors and potential confounders, and stroke medications.**

**eTable 6. All results for association of lacunar stroke per 1-SD higher principal component of metabolite levels with adjustment for age and sex.**

**eTable 7. All results for association of neuroimaging markers of small vessel disease per 1-SD higher baseline metabolite levels with adjustment for age and sex.**

**eTable 8. All results for association of neuroimaging markers of small vessel disease per 1-SD higher baseline metabolite levels with adjustment for age, sex, and additional vascular risk factors and potential confounders.**

**eTable 9. All results for association of neuroimaging markers of small vessel disease per 1-SD higher baseline metabolite levels with adjustment for age, sex, additional vascular risk factors and potential confounders, and stroke medications.**

**eTable 10. All results for association of neuroimaging markers of small vessel disease per 1-SD higher baseline metabolite levels with adjustment for age, sex, and duration between stroke event and date blood sample was taken.**

**eTable 11. All results for association of cognitive decline per 1-SD higher baseline metabolite levels with adjustment for age and sex.**

**eTable 12. All results for association of cognitive decline per 1-SD higher baseline metabolite levels with adjustment for age, sex, and additional vascular risk factors and potential confounders.**

**eTable 13. All results for association of cognitive decline per 1-SD higher baseline metabolite levels with adjustment for age, sex, additional vascular risk factors and potential confounders, and stroke medications.**

**eTable 14. All results for association of cognitive decline per 1-SD higher baseline metabolite levels with adjustment for age, sex, and duration between stroke event and date blood sample was taken.**

**eTable 15. Sample selection process for genome-wide association analyses of metabolic biomarkers.**

**eTable 16. All significant results from genome-wide association analyses of metabolic biomarkers.**

**eTable 17. All Mendelian randomization results for association of metabolite levels with lacunar stroke and neuroimaging markers of small vessel disease.**

**eTable 18. All Mendelian randomization results for association of metabolite levels with lacunar stroke and neuroimaging markers of small vessel disease using relaxed clumping threshold.**

**eTable 19. Generalized inverse-variance weighted Mendelian randomization results for association of glycine with lacunar stroke and neuroimaging markers of small vessel disease using relaxed clumping threshold.**

**eTable 20. Reverse Mendelian randomization results for association of lacunar stroke and neuroimaging markers of small vessel disease with metabolite levels.**

**eTable 21. Results for association of lacunar stroke and cognitive decline per 1-SD higher baseline metabolite levels with adjustment for age and sex and interaction with vascular risk factors.**

# Appendices

## **eAppendix 1: Members of DNA Lacunar 1 and 2 Coordinating Centres and Recruiting Centres**

**DNA Lacunar 1 Study Managers:** Josie Monaghan; Alan Zanich, Samantha Febrey, Eithne Smith, Jenny Lennon, St George's University of London

**DNA Lacunar 1 Participating Centres (number of enrolled patients per centre; local investigators):** Aberdeen Royal Infirmary, Aberdeen (12; Mary Macleod). Addenbrooke's Hospital, Cambridge (54; Jean-Claude Baron, Elizabeth Warburton, Diana J Day, Julie White). Airedale General Hospital, Steeton (4; Samantha Mawer). Barnsley Hospital, Barnsley (3; Mohammad Albazzaz, Pravin Torane, Keith Elliott, Kay Hawley). Bart's and the London, London (2; Patrick Gompertz). Basingstoke and North Hampshire Hospital, Basingstoke (13; Elio Giallombardo, Deborah Dellafera). Blackpool Victoria Hospital, Blackpool (11; Mark O'Donnell). Bradford Royal Infirmary, Bradford (1; Chris Patterson). Bristol Royal Infirmary, Bristol (8; Sarah Caine). Charing Cross Hospital, London (12; Pankaj Sharma). Cheltenham General and Gloucester Royal Hospitals, Cheltenham and Gloucester (10; Dipankar Dutta). Chesterfield Royal Hospital, Chesterfield (4; Sunil Punnoose, Mahmud Sajid). Countess of Chester Hospital, Chester (22; Kausik Chatterjee). Derriford Hospital, Plymouth (4; Azlisham Mohd Nor). Dorset County Hospital NHS Foundation Trust, Dorchester (6; Rob Williams). East Kent Hospitals University NHS Foundation Trust, Kent (22; Hardeep Baht, Guna Gunathilagan). Eastbourne District General Hospital, Eastbourne (4; Conrad Athulathmudali). Frenchay Hospital, Bristol (1; Neil Baldwin). Frimley Park Hospital NHS Foundation Trust, Frimley (6; Brian Clarke). Guy's and St Thomas' Hospital, London (14; Tony Rudd). Institute of Neurology, London (25; Martin Brown). James Paget University Hospital, Great Yarmouth (1; Peter Harrison). King's College Hospital, London (16; Lalit Kalra). Leeds Teaching Hospitals NHS Trust, London (125; Ahamad Hassan). Leicester General Hospital and Royal Infirmary, Leicester (9; Tom Robinson, Amit Mistri). Luton and Dunstable NHSFT University Hospital, Luton (16; Lakshmanan Sekaran, Sakthivel Sethuraman, Frances Justin). Maidstone and Tunbridge Wells NHS Trust (3; Peter Maskell). Mayday University Hospital, Croydon (14; Enas Lawrence). Medway Maritime Hospital, Gillingham (5; Sam Sanmuganathan). Milton Keynes Hospital, Milton Keynes (1; Yaw Duodu). Musgrove Park Hospital, Taunton (9; Malik Hussain). Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne (12; Gary Ford).

Ninewells Hospital, Dundee (5; Ronald MacWalter). North Devon District Hospital, Barnstaple (8; Mervyn Dent). Nottingham University Hospitals, Nottingham (17; Philip Bath, Fiona Hammonds). Perth Royal Infirmary, Perth (2; Stuart Johnston). Peterborough City Hospital, Peterborough (1; Peter Owusu-Agyei). Queen Elizabeth Hospital, Gateshead (5; Tim Cassidy, Maria Bokhari). Radcliffe Infirmary, Oxford (5; Peter Rothwell). Rochdale Infirmary, Rochdale (4; Robert Namushi). Rotherham General Hospital, Rotherham (1; James Okwera). Royal Cornwall Hospitals NHS Trust, Truro (11; Frances Harrington, Gillian Courtauld). Royal Devon and Exeter Hospital, Exeter (22; Martin James). Royal Hallamshire Hospital, Sheffield (1; Graham Venables). Royal Liverpool University Hospital and Broadgreen Hospital, Liverpool (9; Aravind Manoj). Royal Preston Hospital, Preston (18; Shuja Punekar). Royal Surrey County Hospital, Guildford (23; Adrian Blight, Kath Pasco). Royal Sussex County Hospital, Brighton (14; Chakravarthi Rajkumar, Joanna Breeds). Royal United Hospital, Bath (6; Louise Shaw, Barbara Madigan). Salford Royal Hospital, Salford (16; Jane Molloy). Southampton General Hospital, Southampton (1; Giles Durward). Southend Hospital, Westcliff-on-Sea (26; Paul Guyler). Southern General Hospital, Glasgow (34; Keith Muir, Wilma Smith). St George's Hospital, London (108; Hugh Markus). St Helier Hospital, Carshalton (10; Val Jones). Stepping Hill Hospital, Stockport (4; Shivakumar Krishnamoorthy). Sunderland Royal Hospital, Sunderland (1; Nikhil Majumdar). The Royal Bournemouth Hospital, Bournemouth (15; Damian Jenkinson). The Walton Centre, Liverpool (15; Richard White). Torbay Hospital, Torquay (19; Debs Kelly). University Hospital Aintree, Liverpool (19; Ramesh Durairaj). University Hospital of North Staffordshire, Stoke-on-trent (16; David Wilcock). Wansbeck General Hospital and North Tyneside Hospital, Ashington and North Shields (6; Christopher Price). West Cumberland Hospital, Whitehaven (6; Olu Orugun, Rachel Glover). West Hertfordshire Hospital, Watford (20; David Collas). Western General Hospital, Edinburgh (12; Cathie Sudlow). Western Infirmary, Glasgow (33; Kennedy R. Lees, Jesse Dawson). Wycombe Hospital and Stoke Mandeville, High Wycombe (20; Dennis Briley and Matthew Burn). Yeovil District Hospital, Yeovil (46; Khalid Rashed). York Teaching Hospital, York (1; John Coyle).

**DNA Lacunar 2 Coordinating Centre:** Chief Investigator: Hugh Markus; Trial Coordinators: Rebecca Dowden, Laurence Loubiere, Elnara Aghakishiyeva; Case Reviewer: Lupei Cai.

**DNA Lacunar 2 Recruiting Centres (local investigator, number of recruited participants as of 31 January 2025):** Addenbrookes Hospital (Hugh Markus, 299), James Paget Hospital (Carlo Canepa, 3), Peterborough City Hospital (Santhosh Subramonian, 13), North Middlesex

Hospital (Robert Luder, 12), Royal Stoke Hospital (Janaka Weerathunga, 88), Royal London Hospital (Farhad Huwez, 21), West Suffolk Hospital (Abul Azim, 89), Watford General Hospital (Mohit Bhandari, 99), Royal Cornwall Hospital (Katja Adie, 20), New Cross Hospital, Wolverhampton (Nasar Ahmad, 34), Princess Royal University Hospital, Telford (Uttam Sinha, 2), King's Mill Hospital (Jahanzeb Rehan, 27), King's College Hospital (James Teo, 42), Queen Elizabeth Hospital, King's Lynn (Umesh Rai, 4), Countess of Chester Hospital (Kausik Chatterjee, 23), Leeds General Infirmary (Ahamad Hassan, 53), Royal Hallamshire Hospital (Aaizza Naqvi, 18), Wirral University Teaching Hospital (Ruth Davies, 0), University Hospital of North Tees (Ijaz Anwar, 12), Lister Hospital (Aparna Pusalkar, 4), Royal Bournemouth Hospital (Kamy Thavansan, 32), Royal Hampshire County Hospital (Nigel Smyth, 14), Derriford Hospital, Plymouth (Azlisham Mohd Nor, 25), Torbay Hospital (Biju Bhaskaran, 2), Pinderfields Hospital (Prabal Datta, 10), Queen's Hospital, Havering (Abhijit Chaudhuri, 0), Royal Devon and Exeter Hospital (Martin James, 15), University College Hospital (David Werring, 84), Nottingham City Hospital (Kailash Krishnan, 28), Salford Royal Hospital (Dwaipayan Sen, 11), James Cook University Hospital, South Tees (Samer Al-Hussayni, 31), University Hospital of North Durham (Revin Thomas, 52), Yeovil District Hospital (Khalid Rashed, 63), St Peter's Hospital (Brendan Affley, 29), Sandwell General Hospital (David Gull, 21), Guys and St Thomas' Hospital (Ajay Bhalla, 48), Royal Sussex County Hospital (Chakravarthi Rajkumar, 1), Southampton General Hospital (Emma Battersby-Wood, 31), Doncaster Royal Infirmary (Dinesh Chada, 26), St George's Hospital (Usman Khan, 123), Wycombe Hospital (Matthew Burn, 20), Musgrove Park Hospital (Malik Hussain, 37), Scunthorpe General Hospital (Amit Banerjee, 3), Ipswich Hospital (Sajid Alam, 17), Southend University Hospital (Paul Guyler, 12), Gloucestershire Royal Hospital (Dipankar Dutta, 23), Queen Elizabeth The Queen Mother Hospital, Margate (Saidu Abubakar, 9), Colchester General Hospital (Rajesh Saksena, 34), Royal Derby Hospital (Tim England, 5), Croydon University Hospital (Enas Lawrence, 66), Norfolk and Norwich University Hospital (Kneale Metcalf, 12), Leicester Royal Infirmary (Amit Mistri, 16), William Harvey Hospital (Ibrahim Balogun, 5), Medway Maritime Hospital (Samuel Sanmuganathan, 5), Kent and Canterbury Hospital (Saidu Abubakar, 16), Sunderland Royal Hospital (Naweed Sattar, 18), University Hospital Llandough (Benjamin Jelley, 11), Cheltenham General Hospital (Dipankar Dutta, 0), Royal Victoria Infirmary (Akif Gani, 16), University Hospital Coventry & Warwickshire (Munib Mirza, 16), Fairfield General Hospital (Narayanamoorthi Saravanan, 14), Milton Keynes University Hospital (Janet Costa, 9), Royal Blackburn Hospital (Anand Nair, 1), Royal Preston Hospital (Hedley Emsley, 13), Stepping Hill Hospital (Appukuttan

Suman, 6), Royal Albert Edward Infirmary (Habib Rehman, 12), Broomfield Hospital (Ramanathan Kirthivasan, 34), Basildon University Hospital (Jigisha Amin, 1), Homerton University Hospital (Thomas Harrison, 0), Queen Elizabeth Hospital, Gateshead (Louise Southern, 0), and Southmead Hospital (Sandeep Buddha, 0).

# Supplemental Methods

## eMethods

### 1. Further Details of Study Population

DNA Lacunar 1 (DL1) and DNA Lacunar 2 (DL2) are prospective studies of MRI-confirmed lacunar stroke patients recruited from stroke centers throughout the UK. DL1 recruited 1,030 patients of European ancestry aged  $\leq 70$  years with lacunar stroke from 72 specialist stroke centers between 2002 and 2012.<sup>1</sup> DL2, for which recruitment is currently ongoing, had recruited 1,319 MRI-confirmed lacunar stroke patients from 54 hospitals as of July 2023.<sup>2</sup> Both studies had identical inclusion criteria, except that DL1 only recruited patients who had experienced a stroke prior to age 70. Lacunar stroke was identified based on the presence of a clinical lacunar syndrome with an anatomically compatible lesion on MRI, defined as a diffusion-weighted imaging (DWI) positive lesion for acute infarcts, and a cavitated lacune in an anatomically corresponding lesion for cases imaged  $>3$  weeks after the acute event.<sup>1</sup> The median time from event to MRI was 5 days. Each patient underwent comprehensive stroke investigations including brain MRI, carotid artery imaging and ECG, as well as echocardiography where necessary. Exclusion criteria included stenosis  $>50\%$  in the extracranial or intracranial cerebral vessels, history of carotid endarterectomy, high or moderate probability of cardioembolic stroke sources defined by the Trial of Org 10172 Acute Stroke Therapy (TOAST) classification criteria,<sup>3</sup> cortical infarcts on MRI, subcortical infarcts  $>15\text{mm}$  in diameter, monogenic cause of stroke, or other specific causes of stroke.<sup>1</sup> All clinical histories and MRI scans were reviewed centrally by an experienced physician to confirm eligibility.

DL1 also recruited 1,973 unrelated hospital controls of European ancestry without cerebrovascular disease from general practices in the same geographical area by random sampling.<sup>1</sup> Both cases and controls completed standardized clinical assessments and a questionnaire. Brain MRI was not conducted for the controls. Blood was drawn from cases and controls, including serum and plasma, which was centrifuged immediately and stored at  $-80^\circ\text{C}$ . The consent date was used as an approximation for the date of blood draw for the metabolomics assay since samples were usually taken on the date of consent. The day of consent was a median of 25 days (interquartile range 4-105 days) after stroke onset, so most blood samples were drawn from cases during the acute or subacute phases.

## 2. Further Details of Causal Inference Analyses

We conducted a genome-wide association analysis for each of the 250 metabolites using genetic data from DL1 and DL2. Genotyping was conducted by University College London Genomics at the Zayed Centre for Research into Rare Disease in Children using the Infinium Global Screening Array-24 kit. Samples with a low genotype call rate ( $< 0.90$ ) were removed, and the genotyped datasets were harmonized against the reference panel 'PASS.Variantsbravo-dbsnp-all\_hg37.tab.gz' and pre-imputation quality control (QC) filters were applied. Variants were excluded if they were missing from a substantial number of samples ( $\geq 3\%$ ), if they had low minor allele frequencies ( $< 1\%$ ) or Hardy-Weinberg equilibrium  $P$ -values ( $< 1 \times 10^{-7}$ ), or if they were non-autosomal. Samples were also excluded if they had a substantial number of missing SNPs ( $\geq 3\%$ ). Imputation was performed on the TOPMed Imputation Server against the TOPMed-r2 reference panel utilizing Minimac4.<sup>4</sup> Imputed SNPs with sufficiently high imputation scores ( $R_{sq} > 0.3$ ) were kept. Imputed genotypes were sorted and stored according to their chromosomal positions on build GRCh38.

Post-imputation harmonization was carried out using Genotype Harmonizer<sup>5</sup> and post-imputation QC filters were applied. Samples were excluded if they were genotyped for less than 95% of SNPs, and variants were excluded if they were genotyped in less than 97% of samples or if their allele frequencies were less than 0.5%. Kinship analysis was performed with the KING algorithm in PLINK v2 and pi-hat test in PLINK v1.9,<sup>6</sup> and one sample from each pair of samples with a kinship coefficient  $> 0.345$  or pi-hat test  $> 0.2$  was removed from further analysis. LD pruning was carried out to remove highly correlated SNPs and a principal component analysis (PCA) was performed on the pruned dataset to identify the top 20 principal components for genetic ancestry.

Association testing was conducted using Regenie v3.4.1<sup>7</sup> with adjustment for age, sex, and the top three principal components for ancestry, and the GWAS results were visualized using Manhattan and QQ plots. A false discovery rate (FDR) threshold of 5% was applied to correct for multiple testing and select significantly associated variants. For the primary Mendelian randomization (MR) analyses, strict linkage disequilibrium (LD) clumping ( $r^2 < 0.01$ ) was applied to ensure the independence of the genetic instruments.

We excluded palindromic genetic variants with ambiguous allele frequencies and genetic variants with potential strand issues that could not be resolved. For each metabolite, all

independent variants were harmonized with the outcome data to ensure that the effect estimates of each variant with the exposure and outcome were reported with respect to the same effect allele. Proxy variants from the 1000 Genomes European reference population were used for several exposure–outcome pairs where the lead SNPs after clumping were not found in the outcome dataset.

Genetic association summary statistics for MRI-confirmed lacunar stroke (n=3,199), ILI (n=1,133), MLI/LA (n=1,658), white matter hyperintensities (WMH) (n=55,291), cerebral microbleeds (CMB) (n=3,556 cases; 22,306 controls), lower fractional anisotropy (FA) (n=36,533), mean diffusivity (MD) (n=36,460), and peak width of skeletonized mean diffusivity (PSMD) (n=36,012) were obtained from published studies, which used data from DNA Lacunar 1 and 2, UK Biobank, and the Cohorts of Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium.<sup>8,9</sup> All summary statistics were obtained from European-ancestry individuals, except CMB, which included 3% of participants from other ancestries.

Because 12 metabolites yielded no available genetic instruments under the strict clumping threshold, we repeated the MR analysis using a relaxed clumping threshold ( $r^2 < 0.5$ ). One metabolite (glycine) yielded a statistically significant association using this relaxed threshold. To ensure this finding was not an artifact of instrument correlation, we validated this result using the generalized IVW method, which incorporates the LD correlation matrix (which we estimated from the 1000 Genomes European reference panel) into the variance calculation. This validation was conducted in R using the MendelianRandomization package v0.10.0.

### **3. Further Details of Interaction Analyses**

As an additional sensitivity analysis, we examined whether there was significant evidence of interaction with vascular risk factors, namely hypertension, smoking, alcohol consumption, BMI, hyperlipidemia, type 2 diabetes status, and age category. The interaction analyses were conducted for metabolites that were significantly associated with lacunar stroke or its subtypes in both the observational and MR analyses, and for metabolites that were significantly associated with the BMET EF/PS subdomain in the observational analyses.

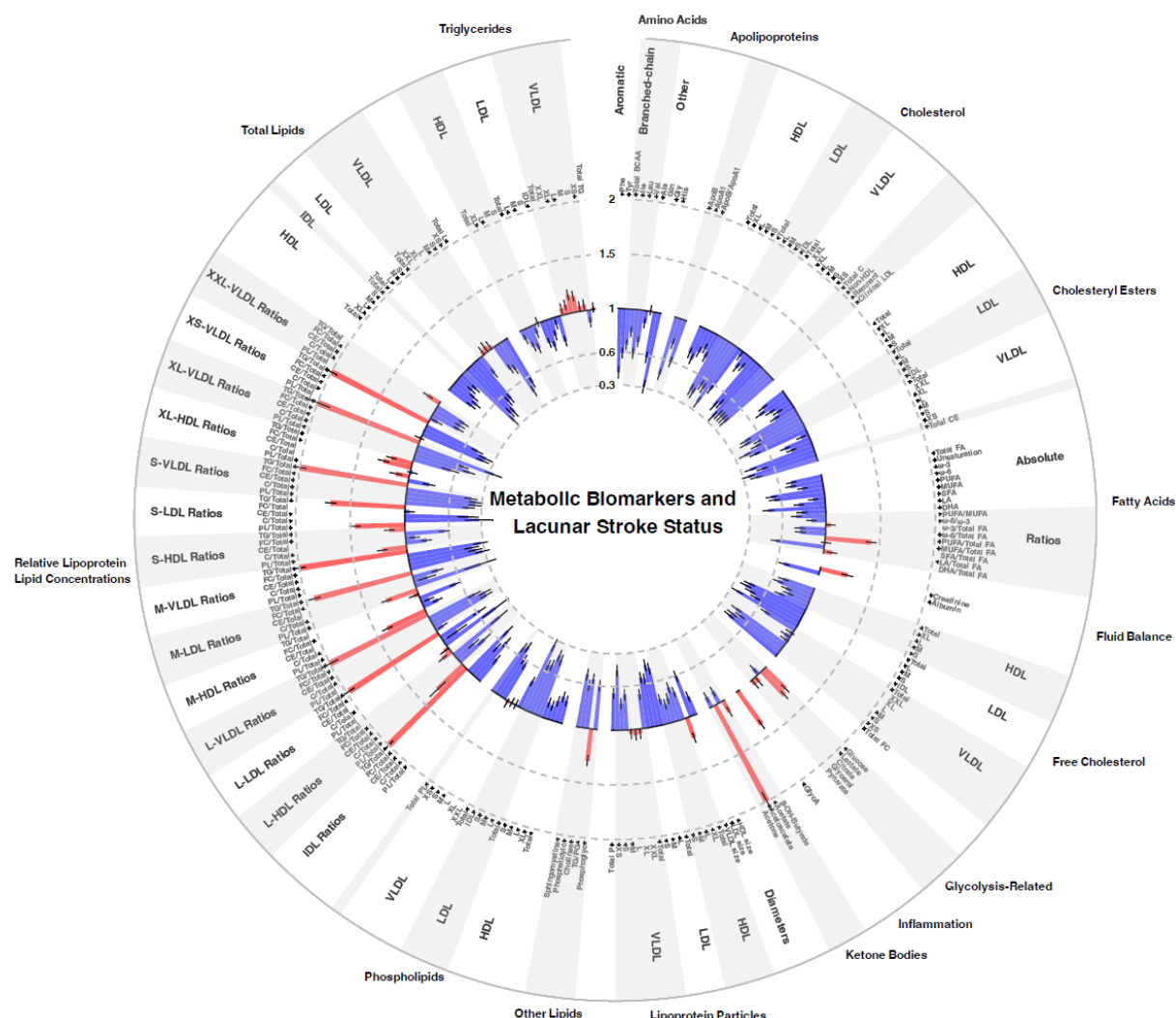
# Supplemental References

## eReferences

1. Kilarski LL, Rutten-Jacobs LCA, Bevan S, et al. Prevalence of CADASIL and Fabry disease in a cohort of MRI defined younger onset lacunar stroke. *PLoS One*. 2015;10(8):e0136352. doi:10.1371/journal.pone.0136352
2. Traylor M, Persyn E, Tomppa L, et al. Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies. *Lancet Neurol*. 2021;20(5):351-361. doi:10.1016/S1474-4422(21)00031-4
3. Adams HP, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial. TOAST: Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24(1):35-41. doi:10.1161/01.STR.24.1.35
4. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287. doi:10.1038/ng.3656
5. Deelen P, Bonder MJ, van der Velde KJ, et al. Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. *BMC Res Notes*. 2014;7(1):901. doi:10.1186/1756-0500-7-901
6. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi:10.1186/s13742-015-0047-8
7. Mbatchou J, Barnard L, Backman J, et al. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet*. 2021;53(7):1097-1103. doi:10.1038/s41588-021-00870-7
8. Koohi F, Harshfield EL, Shatunov A, Markus HS. Does thrombosis play a causal role in lacunar stroke and cerebral small vessel disease? *Stroke*. 2024;55(4):934-942. doi:10.1161/STROKEAHA.123.044937
9. Knol MJ, Lu D, Traylor M, et al. Association of common genetic variants with brain microbleeds: a genome-wide association study. *Neurology*. 2020;95(24):e3331-e3343. doi:10.1212/WNL.0000000000010852

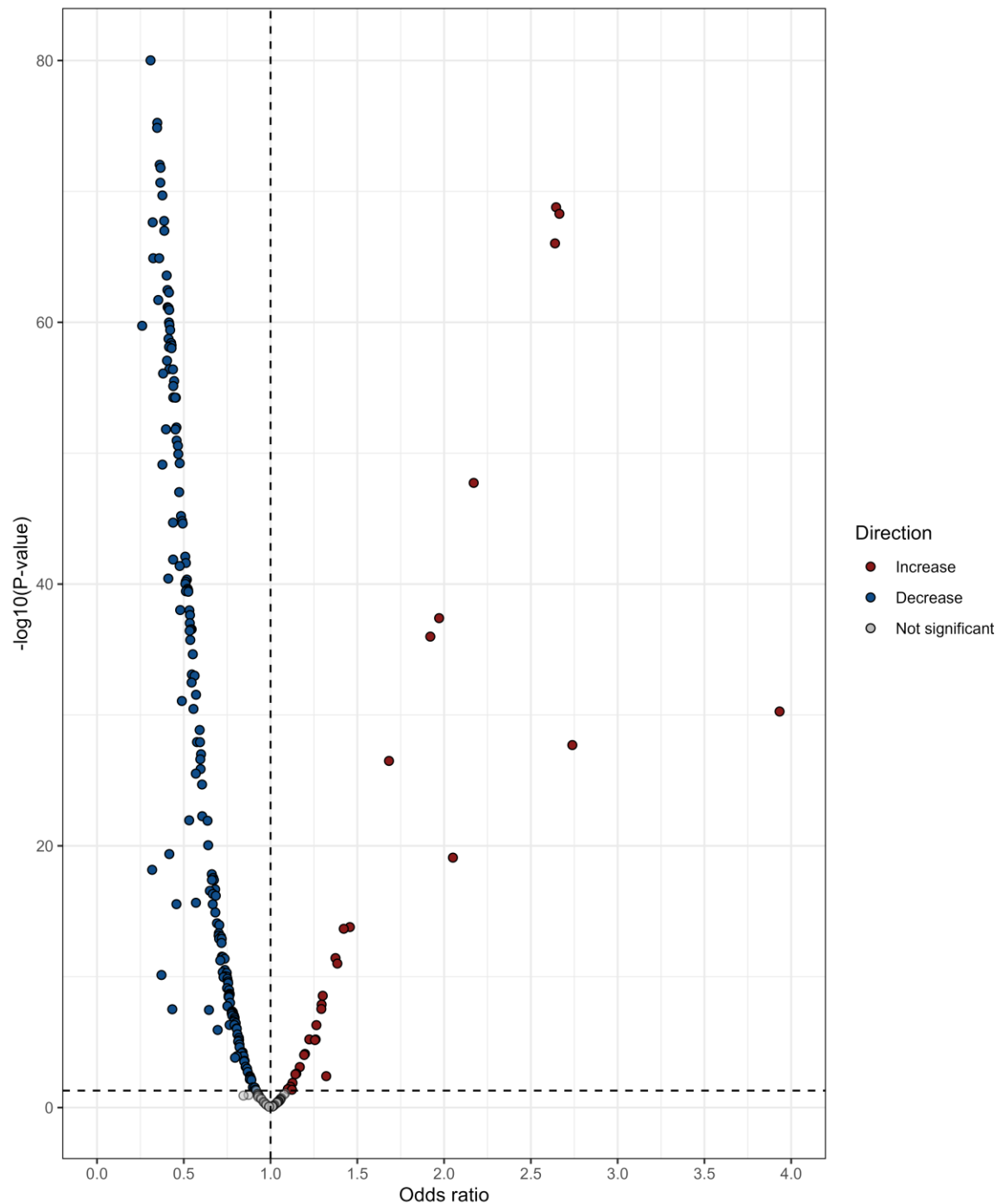
# Supplemental Figures

**eFigure 1. Circular bar plot showing the results of cross-sectional univariate analyses for the association between 250 metabolic biomarkers and lacunar stroke**



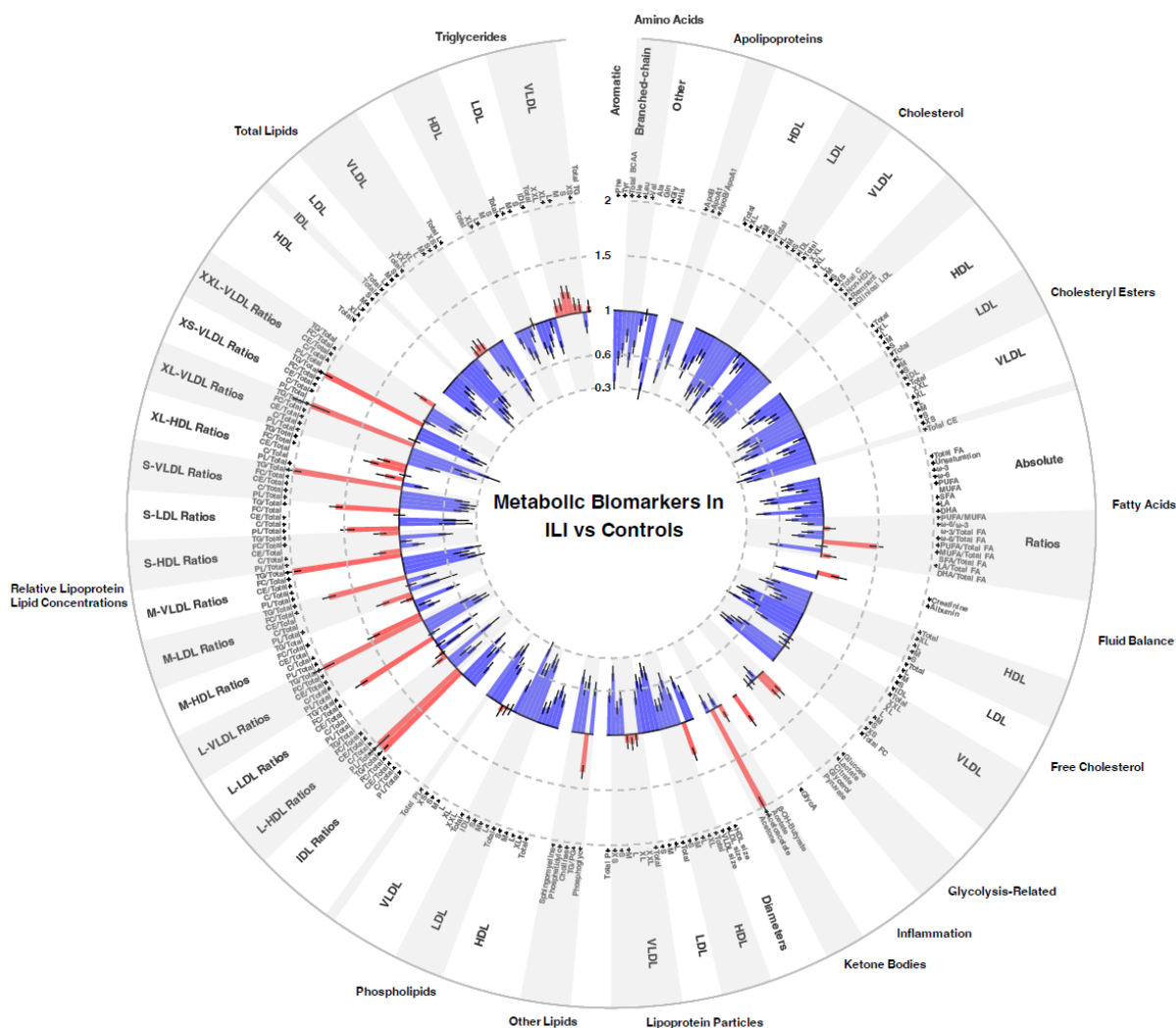
Abbreviations: FDR: false discovery rate. Odds ratios with 95% confidence intervals for the association with lacunar stroke versus controls are presented per 1-SD increase in metabolite concentration with adjustment for age, sex, hypertension, and recruitment center. Associations shown in red indicate increased odds of lacunar stroke, and in blue indicate reduced odds of lacunar stroke. Associations significant at an FDR-corrected  $P < 0.05$  are indicated with an asterisk (\*) alongside the names of the respective metabolites.

**eFigure 2. Volcano plot showing the results of cross-sectional univariate analyses for the association between 250 metabolic biomarkers and lacunar stroke.**



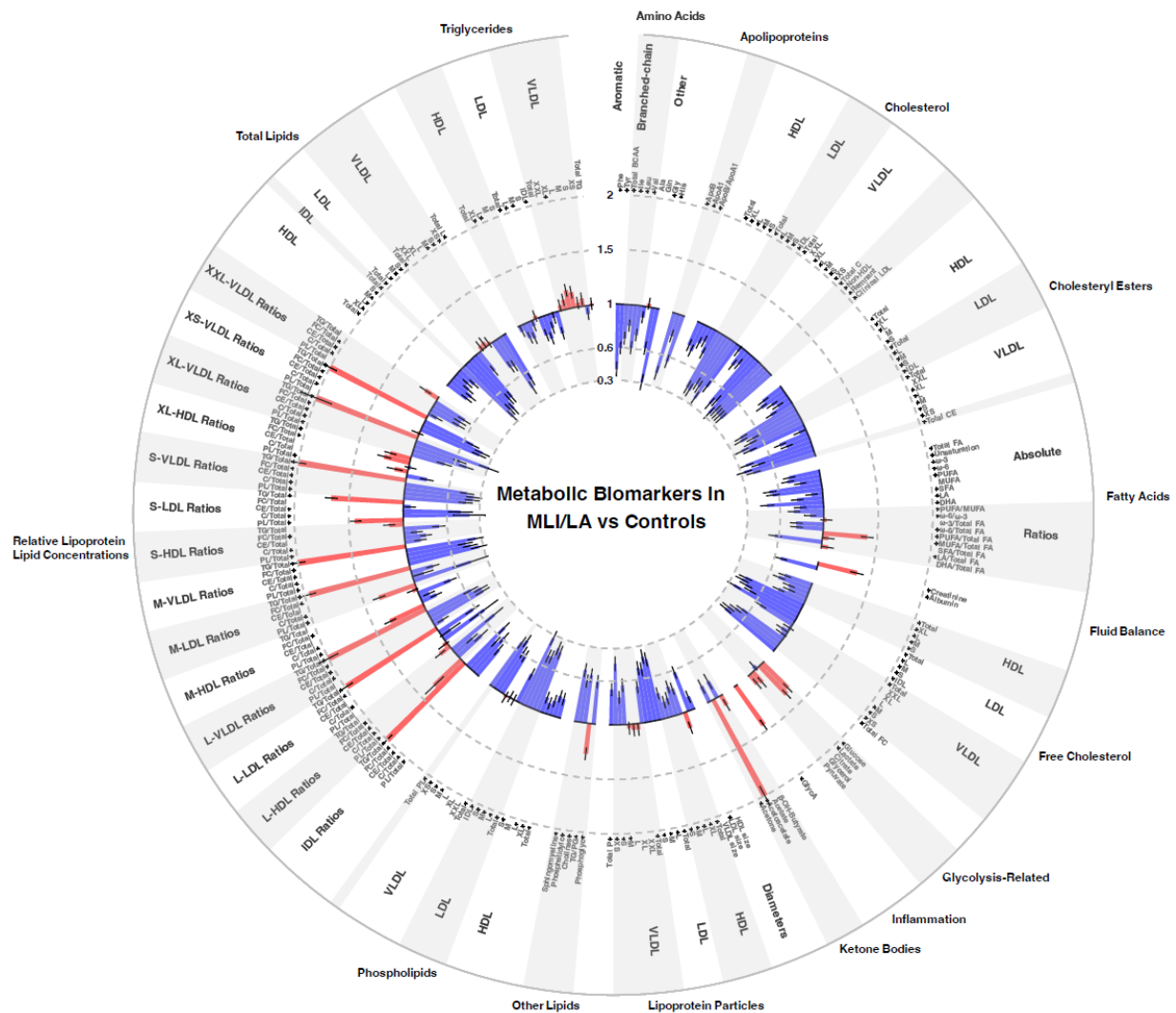
Odds ratios are shown on the x-axis and  $-\log_{10} P$ -values are shown on the y-axis. The FDR-corrected threshold of  $P < 0.05$  is shown as a dashed horizontal line. Metabolites significantly associated with increased risk of lacunar stroke are shown in red and metabolites significantly associated with decreased risk of lacunar stroke are shown in blue.

**eFigure 3. Circular bar plot showing the results of cross-sectional univariate analyses for the association of 250 metabolic biomarkers with ILI cases versus controls.**



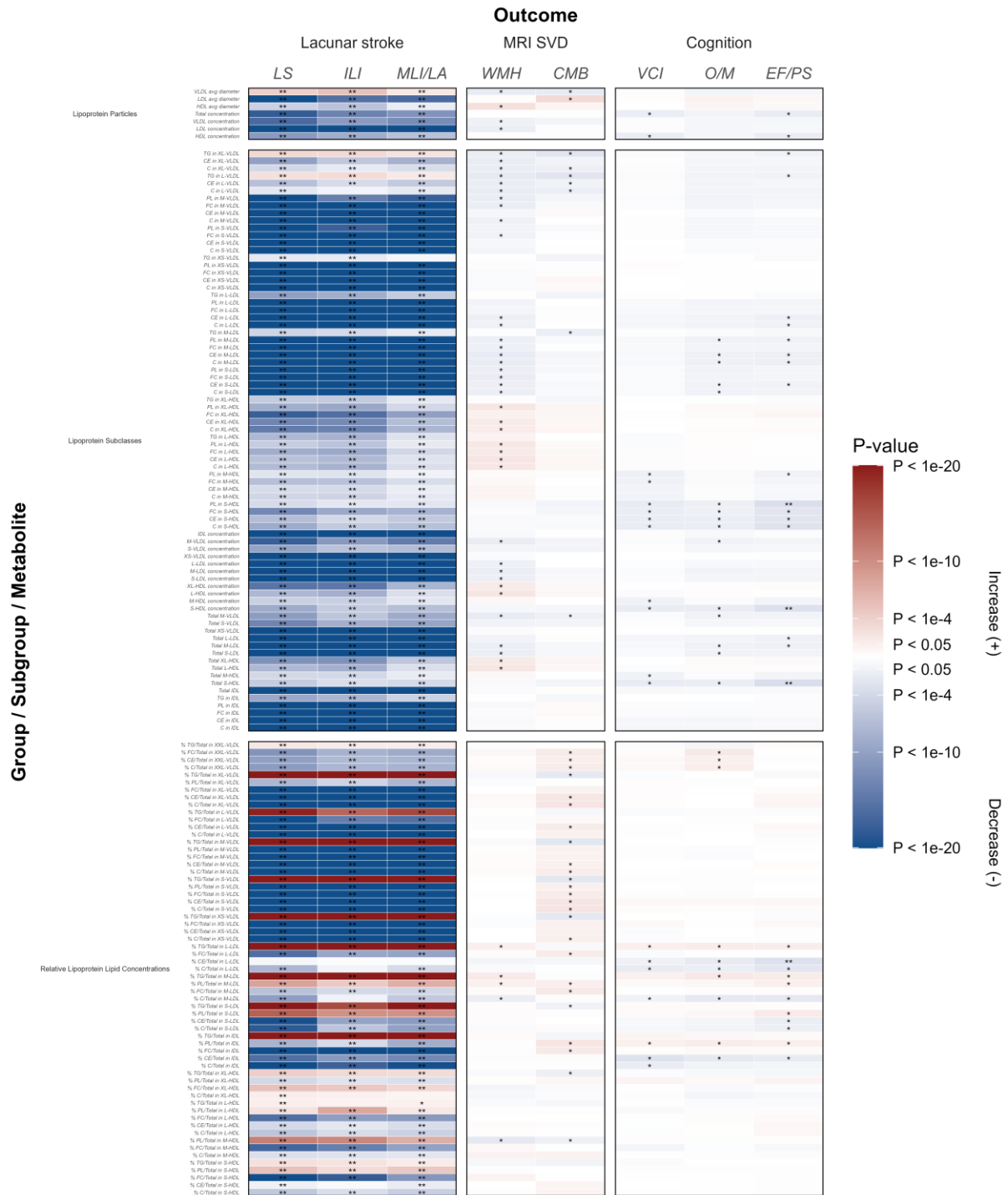
Abbreviations: FDR: false discovery rate; ILI: isolated lacunar infarcts. Odds ratios with 95% confidence intervals are presented for the association with ILI versus controls per 1-SD increase in metabolite concentration with adjustment for age, sex, hypertension, and recruitment centre. Odds ratios  $>1$  showing positive associations are represented in red, and odds ratios  $<1$  showing inverse associations are in blue. Associations significant at an FDR-corrected  $P < 0.05$  are indicated with an asterisk (\*) alongside the names of the respective metabolites.

**eFigure 4. Circular bar plot showing the results of cross-sectional univariate analyses for the association of 250 metabolic biomarkers with MLI/LA cases versus controls.**



Abbreviations: FDR: false discovery rate; MLI: multiple lacunar infarcts; LA: leukoaraiosis. Odds ratios with 95% confidence intervals are presented for the association with MLI or LA versus controls per 1-SD increase in metabolite concentration with adjustment for age, sex, hypertension, and recruitment centre. Odds ratios >1 showing positive associations are represented in red, and odds ratios <1 showing inverse associations are in blue. Associations significant at an FDR-corrected  $P < 0.05$  are indicated with an asterisk (\*) alongside the names of the respective metabolites.

**eFigure 5. Heat map showing the results of cross-sectional univariate analyses for the associations of lipoprotein subclasses and relative lipoprotein lipid concentrations with lacunar stroke, MRI markers, and cognitive impairment.**

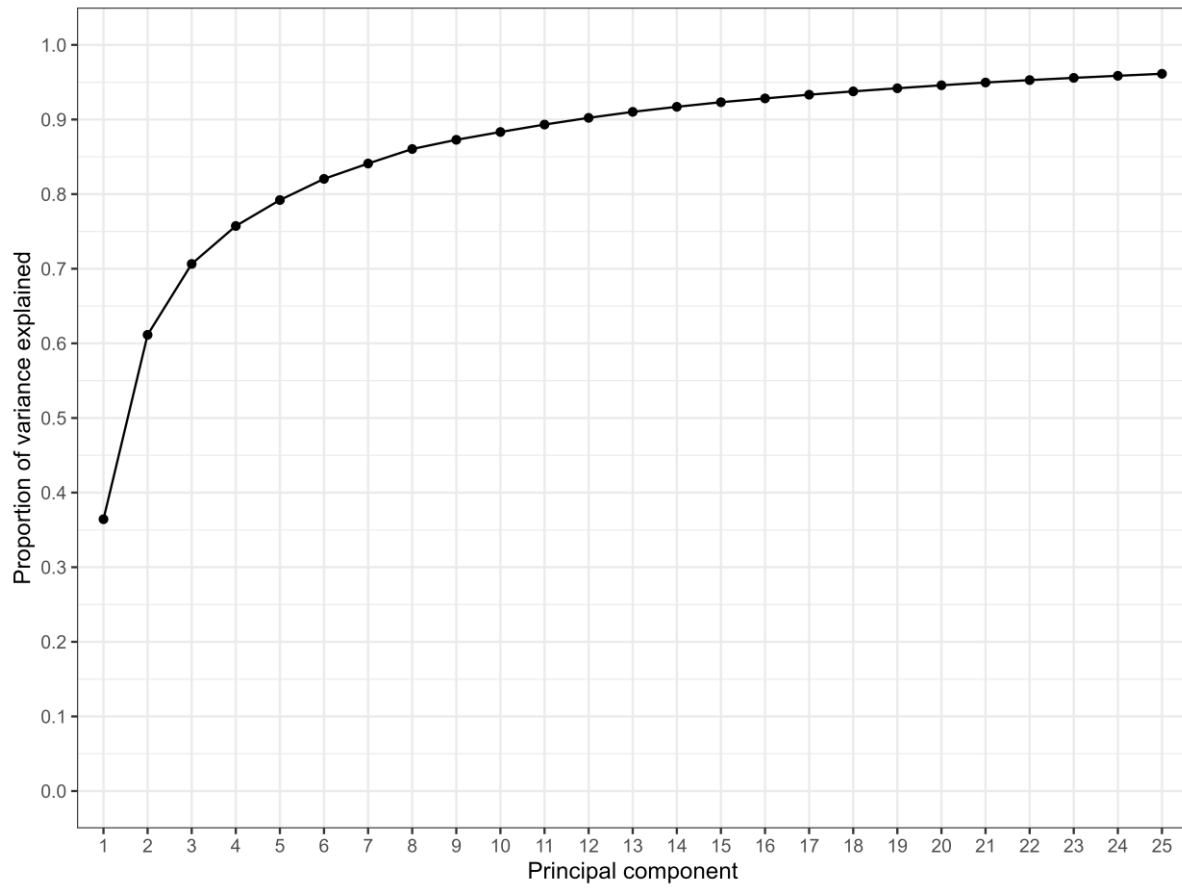


Abbreviations: BMET: Brief Memory and Executive Test; EF/PS: executive functioning/processing speed; FDR: false discovery rate; ILI: isolated lacunar infarcts; LA: leukoaraiosis; MLI: multiple lacunar infarcts; O/M: orientation/memory; SVD: small vessel disease; WMH: white matter hyperintensities. Beta estimates and *P*-values were obtained from

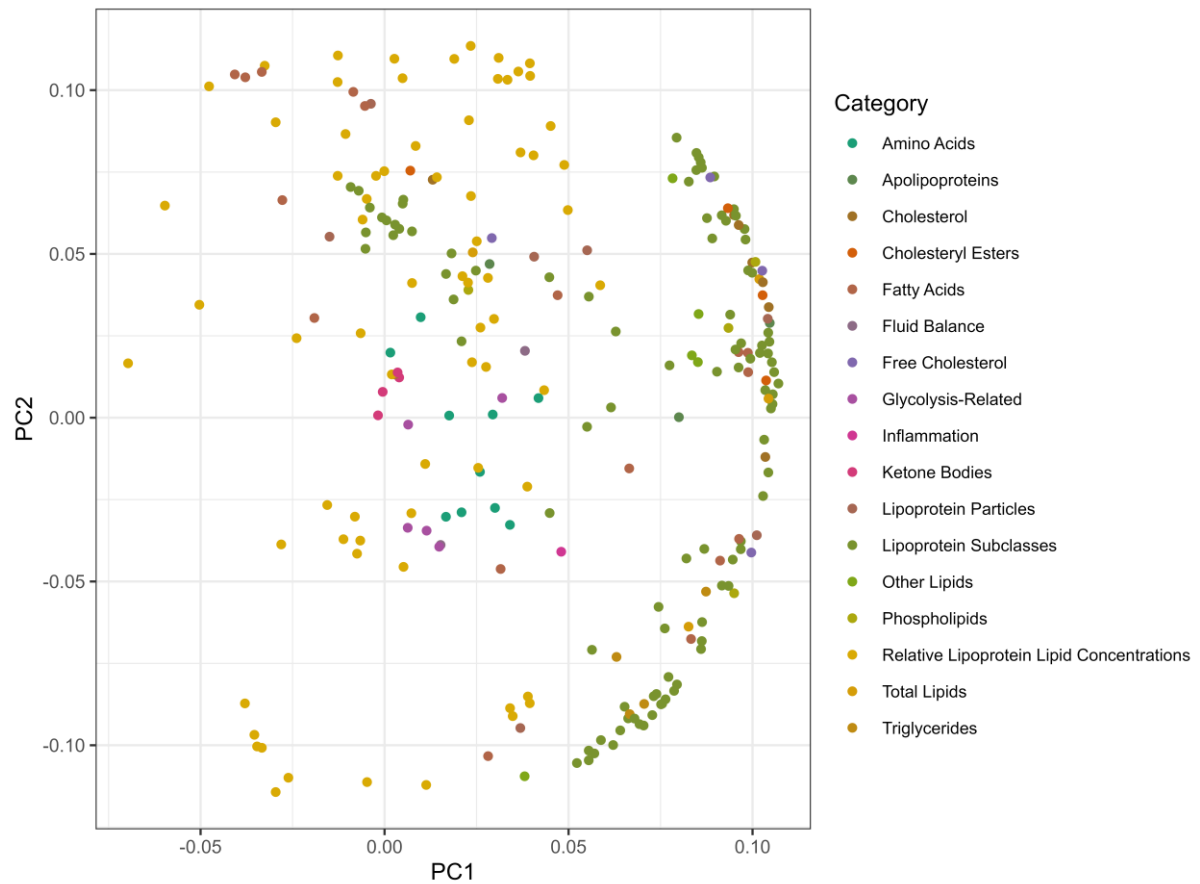
logistic or linear regression models adjusted for age and sex. Colors show magnitude and direction of  $P$ -value for association of metabolite with each outcome. Associations with beta coefficients  $>1$  are represented in red, and associations with beta coefficients  $<1$  are shown in blue. Associations significant at an unadjusted  $P < 0.05$  are indicated with a single asterisk (\*) and associations significant at an FDR-corrected  $P < 0.05$  are indicated with two asterisks (\*\*).

**eFigure 6. Scree plot and loadings plot from principal component analysis of metabolite levels. (A) Scree plot. (B) Loadings plot.**

**(A) Scree plot.**

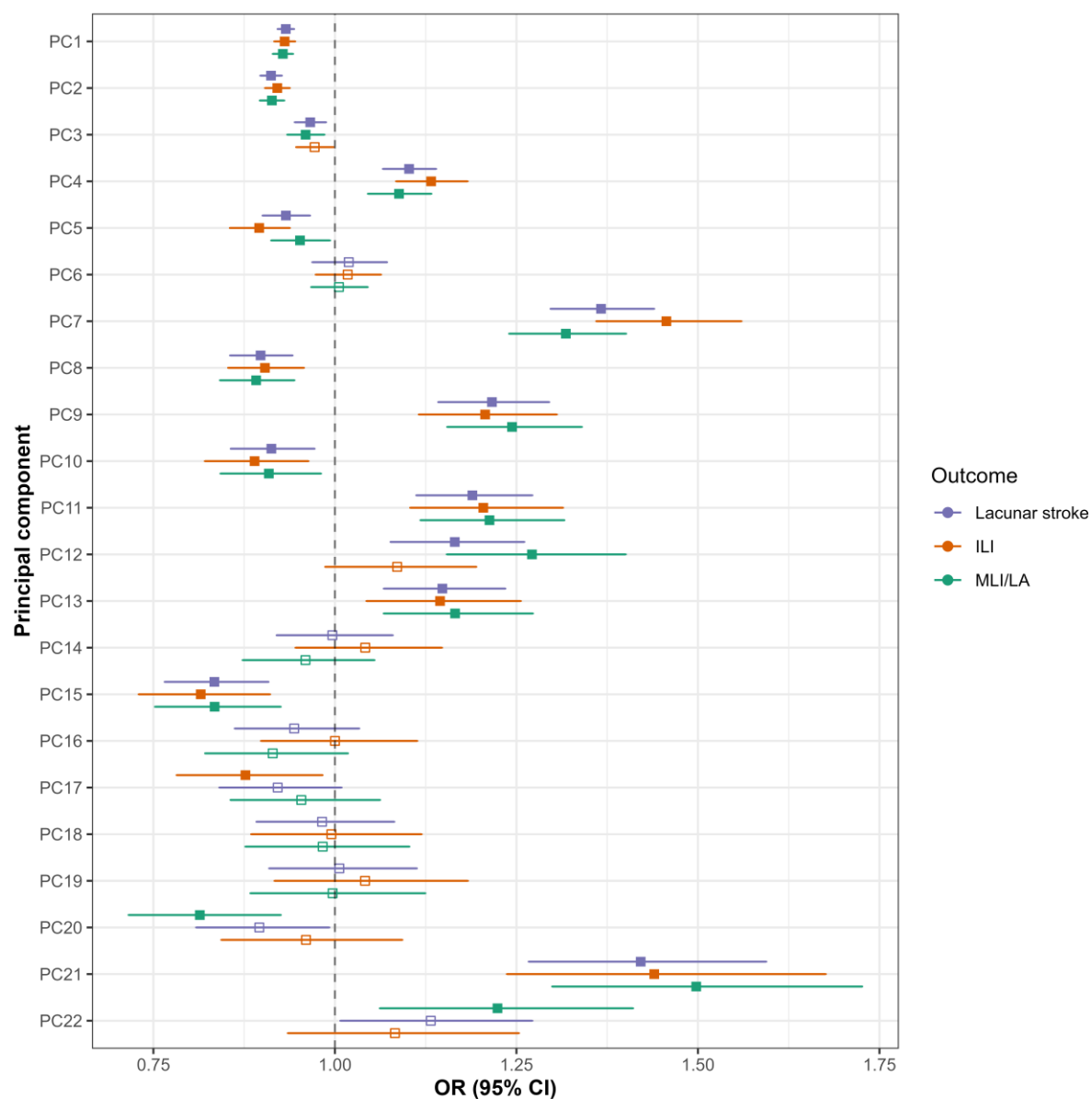


**(B) Loadings plot.**



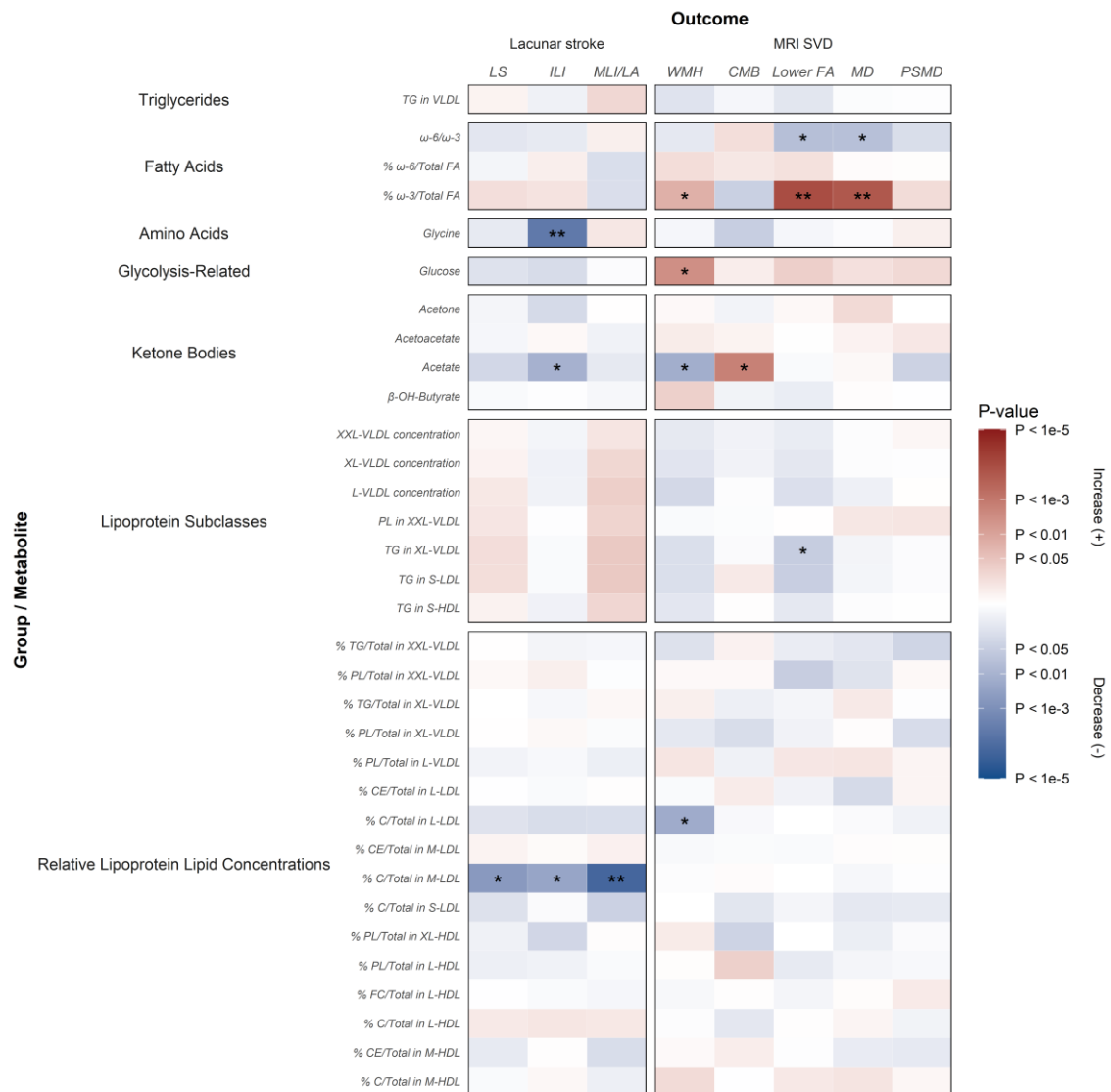
The scree plot **(A)** shows the cumulative proportion of variance explained by each successive principal component of the metabolite concentrations. The loadings plot **(B)** shows a scatter plot of the matrix loadings for the first and second principal components of the metabolite concentrations.

**eFigure 7. Forest plot showing the results of cross-sectional univariate analyses for the association of the first twenty-two principal components of metabolite levels with lacunar stroke status, ILI vs controls, and MLI/LA vs controls.**



Abbreviations: CI: confidence interval; FDR: false discovery rate; ILI: isolated lacunar infarcts; MLI: multiple lacunar infarcts; LA: leukoaraiosis; OR: odds ratio. Odds ratios with 95% confidence intervals for the association with lacunar stroke status, ILI vs controls, and MLI/LA vs controls are presented per 1-SD increase in metabolite concentration with adjustment for age and sex. Associations significant at an FDR-corrected  $P < 0.05$  are indicated with filled points while associations that are not statistically significant are shown with hollow points.

**eFigure 8. Heat map summarizing the causal estimates from Mendelian randomization analyses for the association of metabolic biomarkers with lacunar stroke and MRI markers of SVD using relaxed clumping threshold.**

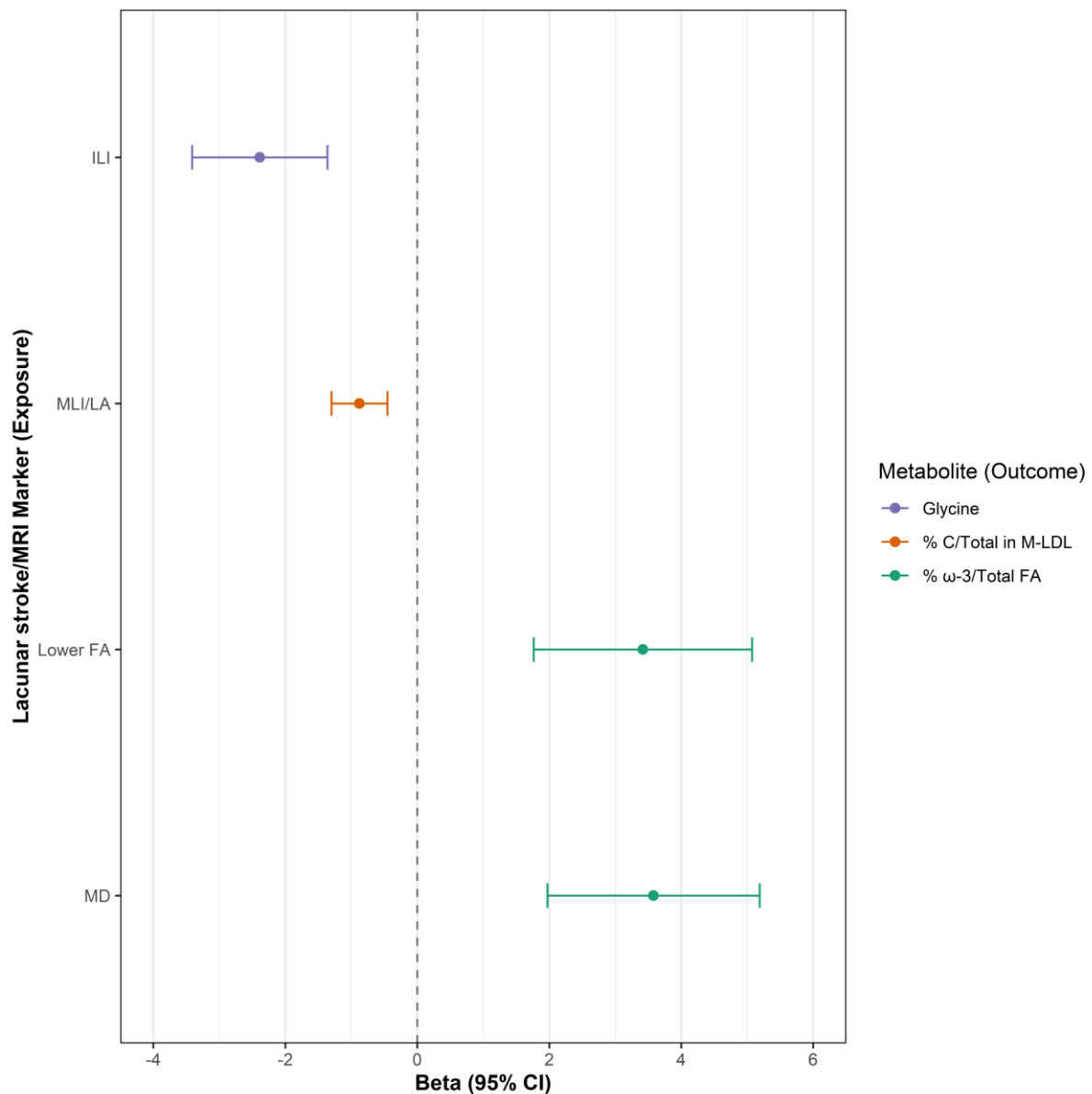


Abbreviations: CMB: cerebral microbleeds; FA: fractional anisotropy; FDR: false discovery rate; ILI: isolated lacunar infarcts; IVW: inverse-variance weighted; LA: leukoaraiosis; MD: mean diffusivity; MLI: multiple lacunar infarcts; PSMD: peak width of skeletonized mean diffusivity; SVD: small vessel disease; WMH: white matter hyperintensities. An LD clumping threshold of  $r^2 < 0.5$  was applied. Metabolites associated with at least one outcome at  $P < 0.5$  using the IVW method are shown. Beta estimates and  $P$ -values were obtained from two-sample Mendelian randomization analyses using the IVW method, or using the Wald ratio method if only a single SNP was available to be used as an instrumental variable. Colors show magnitude and direction of  $P$ -value for association of metabolite with each outcome. Associations with beta coefficients  $>1$  are represented in red, and associations with beta coefficients  $<1$  are shown in blue. Associations significant at an unadjusted  $P < 0.05$  are indicated with a single asterisk (\*) and associations significant at an FDR-corrected  $P < 0.05$  are indicated with two asterisks (\*\*).



Abbreviations: MR: Mendelian randomization; SNP: single nucleotide polymorphism; SVD: small vessel disease. Scatter plots are shown for the Mendelian randomization results with evidence of a significant or suggestive association (at an uncorrected  $P < 0.05$ ) between metabolites and lacunar stroke and MRI markers of SVD using the inverse-variance weighted or Wald ratio method. An LD clumping threshold of  $r^2 < 0.5$  was applied. The x-axis of each plot shows the association of each genetic variant with the metabolite, and the y-axis shows the association of each genetic variant with the outcome. The slope of each line corresponds to an estimate of the causal effect using the various MR methods.

**eFigure 10. Forest plot summarizing the causal estimates from reverse Mendelian randomization analyses for the association of lacunar stroke and MRI markers of SVD with metabolic biomarkers**



Abbreviations: FA: fractional anisotropy; FDR: false discovery rate; ILI: isolated lacunar infarcts; LA: leukoaraiosis; MD: mean diffusivity; MLI: multiple lacunar infarcts; SVD: small vessel disease. Beta estimates and 95% confidence intervals were obtained from two-sample Mendelian randomization analyses using the inverse-variance weighted method.