

Type 2 Diabetes, Glucose, Insulin, BMI and Ischemic Stroke Subtypes Mendelian Randomization Study

Susanna C. Larsson^{1,2}, PhD; Robert A. Scott³, PhD; Matthew Traylor¹, PhD; Claudia C. Langenberg³, PhD; George Hindy⁴, MD; Olle Melander⁴, MD, PhD; Marju Orho-Melander⁴, PhD; Sudha Seshadri⁵, MD; Nicholas J. Wareham³, FRCP; Hugh S. Markus¹, DM; for the METASTROKE Collaboration and NINDS Stroke Genetics Network (SiGN)

¹Stroke Research Group, Department of Clinical Neurosciences, University of Cambridge, Cambridge, United Kingdom

²Unit of Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

³MRC Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom

⁴Department of Clinical Sciences, Lund University, Malmö, Sweden

⁵Boston University School of Medicine, Boston, MA, USA

Correspondence to: Susanna C. Larsson, Unit of Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden. Email address: Susanna.Larsson@ki.se

Email addresses of all authors:

Susanna Larsson: Susanna.Larsson@ki.se

Robert A. Scott: Robert.Scott@mrc-epid.cam.ac.uk

Matthew Traylor: mt628@medschl.cam.ac.uk

Claudia C. Langenberg: Langenberg@mrc-epid.cam.ac.uk

George Hindy: George.Hindy@med.lu.se

Olle Melander: Olle.Melander@med.lu.se

Marju Orho-Melander: Marju.Orho-Melander@med.lu.se

Sudha Seshadri: suseshad@bu.edu

Nicholas J. Wareham: Nick.Wareham@mrc-epid.cam.ac.uk

Hugh S. Markus: hsm32@medschl.cam.ac.uk

Title character count: 96

Number of references: 36

Number of tables: 0

Number of figures: 1

Word count abstract: 239

Word count paper: 2538

Supplemental data: Table e-1, Table e-2, Table e-3; Coinvestigators; STROBE checklist.

Susanna C Larsson and Robert A Scott conducted the statistical analyses.

Search Terms: body mass index; glucose; Mendelian randomization; stroke; type 2 diabetes

STUDY FUNDING

None of the funding bodies given below played any role in the design, writing, or decision to publish this article.

Hugh S Markus is supported by a National Institute for Health Research (NIHR) Senior Investigator award, and his work is supported by the Cambridge Universities NIHR Comprehensive Biomedical Research Centre.

AUTHOR CONTRIBUTIONS

Susanna C Larsson designed the study, performed the statistical analyses, wrote the first draft of the manuscript and drew the figures. Susanna C Larsson is the corresponding author and takes responsibility for the accuracy of the analysis and had authority over manuscript preparation and the decision to submit the manuscript for publication.

Robert A Scott designed the study, performed the statistical analyses, reviewed and commented on the manuscript.

Matthew Traylor reviewed and commented on the manuscript.

Claudia C. Langenberg reviewed and commented on the manuscript.

George Hindy reviewed and commented on the manuscript.

Olle Melander reviewed and commented on the manuscript.

Marju Orho-Melander reviewed and commented on the manuscript.

Sudha Seshadri reviewed and commented on the manuscript.

Nicholas J. Wareham reviewed and commented on the manuscript.

Hugh S. Markus designed the study and reviewed and commented on the manuscript.

AUTHOR DISCLOSURES

Susanna C. Larsson – Reports no disclosures

Robert A. Scott – Reports no disclosures

Matthew Traylor – Reports no disclosures

Claudia Langenberg – Reports no disclosures

George Hindy – Reports no disclosures

Olle Melander – Reports no disclosures

Marju Orho-Melander – Reports no disclosures

Sudha Seshadri – Reports no disclosures

Nicholas J. Wareham – Reports no disclosures

Hugh S. Markus – Reports no disclosures

The study is not industry-sponsored.

Abstract

Objective: To implement a Mendelian randomization (MR) approach to determine whether type 2 diabetes (T2D), fasting glucose, fasting insulin, and body mass index (BMI) are causally associated with specific ischemic stroke subtypes.

Methods: MR estimates of the association between each possible risk factor and ischemic stroke subtypes were calculated using inverse-variance weighted (conventional) and weighted median approaches, and MR-Egger regression was used to explore pleiotropy. The number of single nucleotide polymorphisms utilized as instrumental variables was 49 for T2D, 36 for fasting glucose, 18 for fasting insulin, and 77 for BMI. Genome-wide association study data of SNP–stroke associations were derived from METASTROKE and the Stroke Genetics Network (N = 18,476 ischemic stroke cases and 37,296 controls).

Results Conventional MR analysis showed associations between genetically predicted T2D and large artery stroke (odds ratio 1.28; 95% CI 1.16–1.40; $p = 3.3 \times 10^{-7}$) and small vessel stroke (odds ratio 1.21; 95% CI 1.10–1.33; $p = 8.9 \times 10^{-5}$) but not cardioembolic stroke (OR 1.06; 95% CI 0.97–1.15; $p = 0.17$). The association of T2D with large artery stroke but not small vessel stroke was consistent in a sensitivity analysis using the weighted median method and there was no evidence of pleiotropy. Genetically predicted fasting glucose and fasting insulin levels and BMI were not statistically significantly associated with any ischemic stroke subtype.

Conclusions: This study provides support for that T2D may be causally associated with large artery stroke.

Stroke is one of the leading causes of disability and death worldwide.¹ The burden of stroke is projected to increase considerably in the next decades. Hence, there is a need to develop effective prevention strategies for stroke, which necessitates a better understanding of the underlying risk factors. About 80% of all strokes are ischemic,¹ but this term describes a syndrome caused by a number of different pathologies, which may have different treatments.² The major etiologies of ischemic stroke are large artery atherosclerosis (large artery stroke), small vessel atherosclerosis (small vessel stroke) and cardioembolism (cardioembolic stroke). Recent data from genome-wide association studies (GWAS) have shown that these subtypes have distinctive risk factor profiles, and by implication diverse pathophysiological bases.³

Type 2 diabetes (T2D) is associated with an increased risk of ischemic stroke⁴ but its relative contribution to different ischemic stroke subtypes is unknown. Moreover, several risk factors that are highly correlated with T2D, including hyperglycemia, insulin resistance, and body mass index (BMI), have been associated with risk of stroke in observational studies.^{5,6} However, the causality of the association between these metabolic factors and stroke risk is uncertain because of the high correlations between these factors, and the associations may be confounded by an unhealthy diet or other behavioral or environmental risk factors. Whether elevated fasting glucose and insulin levels and BMI have different associations with different ischemic stroke subtypes are unknown.

The technique of Mendelian randomization (MR) is being increasingly used to examine causality of associations between modifiable risk factors and various diseases, particularly where confounding is of concern. It uses genetic variants (single nucleotide polymorphisms; SNPs) as instrumental variables for estimating the effect of a risk factor on a disease risk. Because genotypes are randomly allocated during meiosis, the population genotype distribution should be unrelated to any confounders.⁷ In this respect, MR can be thought of as a natural randomized controlled trial. In the present study, MR was implemented to determine

whether T2D, glucose, insulin, and BMI are causally associated with all ischemic stroke, and with its three main subtypes; large artery, small vessel, and cardioembolic stroke.

Methods

Study design and data sources

This MR analysis was conducted using summary statistics from two large ischemic stroke genetics consortia: the METASTROKE Collaboration^{3, 8} and the National Institute of Neurological Disorders and Stroke (NINDS)-Stroke Genetics Network (SiGN).⁹ GWAS from these consortia contributed a total of 18,476 ischemic stroke cases and 37,296 controls of mainly European ancestry. Overlapping participants in the two consortia were removed. The TOAST classification system was used for stroke subtyping,¹⁰ identifying 2947 large artery strokes, 2757 small vessel strokes, and 3860 cardioembolic strokes. In addition, there were 8912 undetermined strokes (cases in which the stroke mechanism had not been identified) and multiple pathologies strokes. All cases were ischemic stroke and had brain imaging to confirm this and to exclude a hemorrhage. The majority of stroke cases were prevalent cases. Genotyping was performed using a variety of Illumina or Affymetrix platforms. All cohorts underwent genotype imputation using 1000 Genomes (1 KG) phase I reference panel prior to meta-analysis.^{8, 11} Data from NINDS-SiGN were imputed to a merged reference panel that included the 1000 Genomes Project Phase I and the Genome of the Netherlands.¹²

Standard Protocol Approvals, Registrations, and Patient Consents

Each study included in the stroke consortia was approved by an Institutional Review Board, and all patients provided informed consent.

SNP selection

We selected all SNPs associated with T2D,¹³⁻¹⁶ fasting glucose,¹⁷ fasting insulin,¹⁷ and BMI¹⁸ at genome-wide significance threshold ($p < 5 \times 10^{-8}$) in individuals of European ancestry in previously published large-scale GWAS. Because the effects of *FTO* variants on T2D and fasting insulin appear to be mediated solely by BMI,¹⁷ SNPs from this locus were excluded from the analyses of T2D and fasting insulin to avoid possible confounding by BMI. Thus, the number of SNPs used as instrumental variables in the present analyses was 49 for T2D, 36 for fasting glucose, 18 for fasting insulin, and 77 for BMI (i.e., SNPs associated with BMI in the primary meta-analysis of European-descent individuals) (Table e-1). All SNPs within each trait were uncorrelated (not in linkage disequilibrium).

Statistical analysis

Summary statistics for the association of each SNP with the risk factors and ischemic stroke as a whole and the three main subtypes (large artery, small vessel, and cardioembolic strokes) were acquired, respectively, from previously published GWAS of the risk factors¹³⁻¹⁸ and two stroke consortia (METASTROKE and NINDS-SiGN).^{8,9} The “SNP–risk factor” and “SNP–stroke” associations were used to compute estimates of each “risk factor–stroke” association using an inverse-variance weighted method¹⁹ (hereafter referred to as conventional MR analysis). We conducted complementary analyses using the weighted median and MR-Egger regression methods.^{20, 21} MR-Egger regression can identify and control for bias due to directional pleiotropy (i.e., where the genetic variant has a direct effect on the outcome that is not mediated by the exposure).²⁰ In sensitivity analyses, we assessed the robustness of the results by excluding loci that overlapped between the four metabolic traits assessed in this study and by excluding SNPs associated with blood lipids^{22, 23} and smoking²⁴ (Table e-1).

Results of the MR analyses are presented per genetically predicted one-unit higher log-

odds of T2D and one standard deviation (SD) increase in fasting glucose (0.65 mmol/L) and natural log-transformed fasting insulin (0.60 pmol/L). The SDs used for scaling of the results for glucose and insulin were derived from the population-based Fenland²⁵ and Ely²⁶ studies. Because the GWAS of BMI estimated the SNP-BMI associations using inverse normal transformation, the MR estimates can be interpreted as the odds ratio (OR) per 1-SD increment in BMI. All statistical tests were 2-sided and considered statistically significant at Bonferroni-corrected significance threshold ($p = 0.05/16$ tests = 0.003). The analyses were performed in Stata version 14.1 (StataCorp, College Station, Texas, USA).

Results

The associations between each genetically predicted risk factor and all ischemic strokes and ischemic stroke subtypes, using conventional MR analysis, are presented in Figure 1. T2D was associated with a higher odds of all ischemic stroke (OR 1.12; 95% CI 1.07–1.17; $p = 3.0 \times 10^{-6}$), large artery stroke (OR 1.28; 95% CI 1.16–1.40; $p = 3.3 \times 10^{-7}$), and small vessel stroke (OR 1.21; 95% CI 1.10–1.33; $p = 8.9 \times 10^{-5}$) but not with cardioembolic stroke (OR 1.06; 95% CI 0.97–1.15; $p = 0.17$). The association of T2D with large artery stroke but not small vessel stroke was consistent in complementary analysis using the weighted median method (Table e-2). Results from the MR-Egger regression analysis, which has lower statistical power, yielded a similar OR for T2D and large artery stroke but the CI was wide and included the null (Table e-2). MR-Egger regression analysis provided no evidence of directional pleiotropy for the associations of T2D with large artery stroke (intercept = 0.003; $p = 0.82$), small vessel stroke (intercept = 0.008; $p = 0.50$), or cardioembolic stroke (intercept = 0.005; $p = 0.62$).

Genetically predicted higher fasting glucose was associated with a higher odds of large artery stroke in conventional MR analysis (Figure 1) but the result did not remain statistically

significant after Bonferroni correction and no statistically significant association was observed in the weighted median and MR-Egger regression analyses (Table e-2). A similar, non-significant association was found between fasting insulin and large artery stroke (Figure 1 and Table e-2). Fasting glucose and fasting insulin were not associated with small vessel or cardioembolic strokes, and there were no associations between BMI and all ischemic strokes or any subtype (Figure 1 and Table e2).

We conducted a series of sensitivity analyses to assess the robustness of the results (Table e-3). First we excluded the *TCF7L2* locus, which was associated with large artery stroke as well as strongly with T2D and weakly with the other three traits (Table e-1). Exclusion of this locus did not appreciably change the results for T2D and large artery stroke (OR 1.25; 95% CI 1.13–1.39; $p = 3.3 \times 10^{-5}$). Removing loci that overlapped between the four traits also did not materially alter the results for T2D and large artery stroke (OR 1.32; 95% CI 1.14–1.53; $p = 2.5 \times 10^{-4}$) or any of the other findings (Table e-3). In a further sensitivity analysis, we excluded SNPs associated with lipids (low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, total cholesterol, or triglycerides) at genome-wide significance level. In this analysis the association between T2D and large artery stroke remained (OR 1.29; 95% CI 1.17–1.42; $p = 6.0 \times 10^{-7}$) and the associations of the other three traits with ischemic stroke did not change appreciably (Table e-3). Finally, the lack of association between BMI and all ischemic stroke and its subtypes remained after removing the *BDNF* locus, which has been found to be associated with smoking (Table e-3).

Discussion

This MR study showed a robust association between genetic predisposition to T2D and large artery stroke but not small vessel or cardioembolic strokes. It did not provide clear evidence for a causal role of fasting glucose, fasting insulin, and BMI in any ischemic stroke subtype.

Studies of the association between T2D and etiologic subtypes of ischemic stroke are rare, but several previous observational studies have investigated the association between T2D and risk of all ischemic strokes or total stroke (ischemic and hemorrhagic strokes combined). For example, in a prospective cohort of 1.9 million people, T2D was associated with a 72% increased risk of all ischemic strokes.⁴ Furthermore, in a meta-analysis of 64 observational prospective studies, representing 775,385 individuals and 12,539 stroke cases, the overall relative risks of total stroke for individuals with diabetes versus non-diabetics were 2.28 (95% CI 1.93–2.69) in women and 1.83 (95% CI 1.60–2.08) in men.²⁷

T2D is associated with insulin resistance and elevated insulin levels during the early stages of the disease. Eventually the pancreas is no longer able to compensate the insulin resistance by secreting sufficient insulin to keep blood glucose level within the normal range, thereby resulting in hyperglycemia. A meta-analysis of 14 prospective studies reported a statistically significant 44% increased risk of total stroke for high versus low fasting glucose levels.⁶ Insulin resistance was associated with a 76% increased risk of stroke (based on four studies) and high versus low fasting insulin was associated with a non-significant 90% higher risk (based on two studies).⁶

The value of intensive glucose-lowering therapy for cardiovascular disease prevention in patients with T2D was questioned since two randomized controlled trials failed to show a reduction in major vascular events, including stroke,^{28,29} and even increased mortality²⁹ with the use of intensive glucose control compared with standard therapy. The present MR study found no clear support for a causal role of fasting glucose in any ischemic stroke subtype. Genetically predicted fasting glucose was associated with large artery stroke at nominal statistical significance in the conventional MR analysis but this association did not remain in the MR-Egger regression analysis.

A randomized controlled trial that investigated the effect of pioglitazone, which improves insulin sensitivity, on cardiovascular events in patients with insulin resistance but without T2D showed a 24% reduced risk of the composite primary outcome of stroke and myocardial infarction in the pioglitazone arm compared with placebo over a median follow-up of 4.8 years.³⁰ In this MR analysis, genetically predicted higher fasting insulin was associated with a non-significant higher odds of large artery stroke.

The lack of clear association between genetically predicted fasting glucose and insulin and large artery stroke suggests that other pathways might explain the association with T2D. Observational studies have reported that T2D is associated with increased arterial stiffness,^{31, 32} which is a risk factor for stroke.³³ In addition, a causal association between T2D and arterial stiffness was found in a recent MR study.³⁴ Stiffening of the arteries may result in the development of atherosclerotic carotid plaques³⁵ and increased risk of large artery stroke.³⁶

The present study did not detect a statistically significant association between genetically predicted BMI and ischemic stroke or its subtypes. Results from a recent meta-analysis of 6 cohort studies conducted in Europe, the United States, and Australia, with a total of 14,443 ischemic stroke cases, showed a 22% increase in ischemic stroke risk per 5-kg/m² increment of BMI.⁵ Because our MR study had limited statistical power to detect such a small increase in risk, we cannot rule out the possibility of a causal association between BMI and ischemic stroke. A previous MR study, including 23,782 participants and 3813 stroke cases, observed no association between a genetic instrument comprising 14 BMI-associated SNPs and risk of all strokes (OR 1.03; 95% CI 0.95–1.12, per 1 kg/m² increase in BMI).³⁷ However, another MR study, including 20 055 participants and 1500 ischemic stroke cases, found a statistically significant association of genetically predicted higher BMI, based on a genetic instrument comprising 32 BMI-associated SNPs, with increased risk of ischemic stroke (hazard ratio 1.83; 95% CI 1.05–3.20, per 4.5 kg/m² increment of BMI).³⁸

Chief strengths of this study include the large number of total ischemic stroke cases and data on etiologic subtypes of ischemic stroke. Moreover, because of the MR design, reverse causation bias was prevented and potential confounding was reduced because genetic variants are not associated with self-selected dietary and lifestyle factors that may affect results from observational studies.

The MR approach also has limitations. First, pleiotropy (i.e., the genetic association with the outcome is through a different causal pathway and not via the risk factor of interest) could have influenced our results. However, our main finding for T2D and large artery stroke was consistent in sensitivity analyses using the weighted median and MR-Egger methods, and remained after excluding pleiotropic SNPs associated with lipids. MR-Egger regression provides a valid test of directional pleiotropy, and a valid test of the causal null hypothesis, under the InSIDE assumption that the direct pleiotropic effects of the genetic variants on the outcome are distributed independently of the genetic association with the exposure.²⁰ Second, we could not assess whether canalization (i.e. compensatory mechanisms)⁷ may have affected our results. Canalization tends to attenuate the estimates and thus cannot explain the association between T2D and large artery stroke but might have contributed to the lack of association of the other risk factors with ischemic stroke. Third, we cannot entirely rule out population stratification as a source of bias in this study. Nevertheless, restriction of the data to largely European-descent individuals reduced potential bias due to population stratification. Fourth, the MR assumptions are somewhat stricter for binary exposures like T2D than for continuous exposures. For MR analysis of binary exposures, the monotonicity assumption must hold.³⁹ Within the context of MR, the monotonicity assumption implies that any effect of the genetic instrument on the exposure is in the same direction for all persons. We cannot rule out the possibility that our results for T2D may have been biased due to monotonicity violations.

A further limitation is that the genetic variants associated with the risk factors under study only explain a small fraction of the variation in the risk factors. For example, the BMI-associated SNPs account for less than 3% of the variation in BMI.¹⁸ This low variability together with the relatively small sample sizes for the analyses of ischemic stroke subtypes, resulted in low precision. Thus, we cannot exclude that the lack of significant associations of fasting glucose, fasting insulin, and BMI with ischemic stroke subtypes are due to insufficient statistical power. Another limitation is that the genetic effects on fasting glucose and insulin were estimated in individuals without diabetes, whereas our analyses included both diabetics and non-diabetics. Finally, we could not examine potential non-linear relationships for the continuous risk factors.

This MR study indicates that T2D may be a causal risk factor for large artery stroke, but did not confirm the results of observational studies suggesting that higher fasting glucose, fasting insulin, and BMI are associated with an increased risk of ischemic stroke. Further MR studies with larger sample sizes are required to determine whether glucose, insulin, and BMI are causally associated with any ischemic stroke subtype.

ACKNOWLEDGMENT

METASTROKE 1000K Discovery Constituent Studies

Australian Stroke Genetics Collaboration (ASGC) Australian population control data were derived from the Hunter Community Study. We also thank the University of Newcastle for funding and the men and women of the Hunter region who participated in this study. This research was funded by grants from the Australian National and Medical Health Research Council (NHMRC Project Grant ID: 569257), the Australian National Heart Foundation (NHF Project Grant ID: G 04S 1623), the University of Newcastle, the Gladys M Brawn Fellowship scheme, and the Vincent Fairfax Family Foundation in Australia. Elizabeth G Holliday was supported by a Fellowship from the National Heart Foundation and National Stroke Foundation of Australia (ID: 100071).

Bio-Repository of DNA in Stroke (BRAINS) is partly funded by a Senior Fellowship from the Department of Health (UK) to P Sharma, the Henry Smith Charity and the UK-India Education Research Institute (UKIERI) from the British Council.

Genes Affecting Stroke Risk and Outcome Study (GASROS) was supported by NINDS (U01 NS069208), the American Heart Association/Bugher Foundation Centers for Stroke Prevention Research 0775010N, the NIH and NHLBI's STAMPEED genomics research program (R01 HL087676), and a grant from the National Center for Research Resources. The Broad Institute Center for Genotyping and Analysis is supported by grant U54 RR020278 from the National Center for Research resources.

Genetics of Early Onset Stroke (GEOS) Study, Baltimore, USA was supported by the NIH Genes, Environment and Health Initiative (GEI) Grant U01 HG004436, as part of the GENEVA consortium under GEI, with additional support provided by the Mid-Atlantic Nutrition and Obesity Research Center (P30 DK072488), and the Office of Research and Development, Medical Research Service, and the Baltimore Geriatrics Research, Education, and Clinical Center of the Department of Veterans Affairs. Genotyping services were provided by the Johns Hopkins University Center for Inherited Disease Research (CIDR), which is fully funded through a federal contract from the NIH to the Johns Hopkins University (contract number HHSN268200782096C). Assistance with data cleaning was provided by the GENEVA Coordinating Center (U01 HG 004446; PI Bruce S Weir). Study recruitment and assembly of datasets were supported by a Cooperative Agreement with the Division of Adult and Community Health, Centers for Disease Control and Prevention and by grants from NINDS and the NIH Office of Research on Women's Health (R01 NS45012, U01 NS069208-01).

Heart Protection Study (HPS) (ISRCTN48489393) was supported by the UK Medical Research Council (MRC), British Heart Foundation, Merck and Co (manufacturers of simvastatin), and Roche Vitamins Ltd (manufacturers of vitamins). Genotyping was supported by a grant to Oxford University and CNG from Merck and Co. Jemma C Hopewell acknowledges support from the British Heart Foundation (FS/14/55/30806).

Ischemic Stroke Genetics Study (ISGS)/Siblings With Ischemic Stroke Study (SWISS) was supported in part by the Intramural Research Program of the NIA, NIH project Z01 AG-000954-06. ISGS/SWISS used samples and clinical data from the NIH-NINDS Human Genetics Resource Center DNA and Cell Line Repository (<http://ccr.coriell.org/ninds>), human subjects protocol numbers 2003-081 and 2004-147. ISGS/SWISS used stroke-free participants from the Baltimore Longitudinal Study of Aging (BLSA) as controls. The inclusion of BLSA samples was supported in part by the Intramural Research Program of the NIA, NIH project Z01 AG-000015-50, human subjects protocol number 2003-078. The ISGS study was funded by NIH-NINDS Grant R01 NS-42733 (J F Meschia). The SWISS study was funded by NIH-NINDS Grant R01 NS-39987 (J F Meschia). This study used the high-performance computational capabilities of the Biowulf Linux cluster at the NIH (<http://biowulf.nih.gov>).

Milano - Besta Stroke Register Collection and genotyping of the Milan cases within CEDIR were supported by Annual Research Funding of the Italian Ministry of Health (Grant Numbers: RC 2007/LR6, RC 2008/LR6; RC 2009/LR8; RC 2010/LR8). FP6 LSHM-CT-2007-037273 for the PROCARDIS control samples.

Wellcome Trust Case-Control Consortium 2 (WTCCC2) was principally funded by the Wellcome Trust, as part of the Wellcome Trust Case Control Consortium 2 project (085475/B/08/Z and 085475/Z/08/Z and WT084724MA). The Stroke Association provided

additional support for collection of some of the St George's, London cases. The Oxford cases were collected as part of the *Oxford Vascular Study* which is funded by the MRC, Stroke Association, Dunhill Medical Trust, National Institute of Health Research (NIHR) and the NIHR Biomedical Research Centre, Oxford. The *Edinburgh Stroke Study* was supported by the Wellcome Trust (clinician scientist award to C Sudlow), and the Binks Trust. Sample processing occurred in the Genetics Core Laboratory of the Wellcome Trust Clinical Research Facility, Western General Hospital, Edinburgh. Much of the neuroimaging occurred in the Scottish Funding Council Brain Imaging Research Centre (www.sbirc.ed.ac.uk), Division of Clinical Neurosciences, University of Edinburgh, a core area of the Wellcome Trust Clinical Research Facility and part of the SINAPSE (Scottish Imaging Network—A Platform for Scientific Excellence) collaboration (www.sinapse.ac.uk), funded by the Scottish Funding Council and the Chief Scientist Office. Collection of the Munich cases and data analysis was supported by the Vascular Dementia Research Foundation.

Vitamin Intervention Stroke Prevention (VISP) The GWAS component of the VISP study was supported by the United States National Human Genome Research Institute (NHGRI), Grant U01 HG005160 (PI Michèle Sale & Bradford Worrall), as part of the Genomics and Randomized Trials Network (GARNET). Genotyping services were provided by the Johns Hopkins University Center for Inherited Disease Research (CIDR), which is fully funded through a federal contract from the NIH to the Johns Hopkins University. Assistance with data cleaning was provided by the GARNET Coordinating Center (U01 HG005157; PI Bruce S Weir). Study recruitment and collection of datasets for the VISP clinical trial were supported by an investigator-initiated research grant (R01 NS34447; PI James Toole) from the United States Public Health Service, NINDS, Bethesda, Maryland. Control data for comparison with VISP stroke cases were from the dbGaP study High Density SNP Association Analysis of Melanoma: Case-Control and Outcomes Investigation (phs000187.v1.p1; R01CA100264, 3P50CA093459, 5P50CA097007, 5R01ES011740, 5R01CA133996, HHSN268200782096C; PIs Christopher Amos, Qingyi Wei, Jeffrey E. Lee).

Women's Health Initiative (WHI) Funding support for WHI-GARNET was provided through the NHGRI GARNET (Grant Number U01 HG005152). Assistance with phenotype harmonisation and genotype cleaning, as well as with general study coordination, was provided by the GARNET Coordinating Center (U01 HG005157). Funding support for genotyping, which was performed at the Broad Institute of MIT and Harvard, was provided by the NIH Genes, Environment, and Health Initiative (GEI; U01 HG004424).

REFERENCES

1. Endres M, Heuschmann PU, Laufs U, Hakim AM. Primary prevention of stroke: blood pressure, lipids, and heart failure. *Eur Heart J* 2011;32:545-552.
2. Markus HS, Bevan S. Mechanisms and treatment of ischaemic stroke--insights from genetic associations. *Nat Rev Neurol* 2014;10:723-730.
3. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol* 2012;11:951-962.
4. Shah AD, Langenberg C, Rapsomaniki E, et al. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet Diabetes Endocrinol* 2015;3:105-113.
5. Kroll ME, Green J, Beral V, et al. Adiposity and ischemic and hemorrhagic stroke: Prospective study in women and meta-analysis. *Neurology* 2016;87:1473-1481.
6. Gast KB, Tjeerdema N, Stijnen T, Smit JW, Dekkers OM. Insulin resistance and risk of incident cardiovascular events in adults without diabetes: meta-analysis. *PLoS One* 2012;7:e52036.
7. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1-22.
8. Malik R, Traylor M, Pulit SL, et al. Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. *Neurology* 2016;86:1217-1226.
9. Meschia JF, Arnett DK, Ay H, et al. Stroke Genetics Network (SiGN) study: design and rationale for a genome-wide association study of ischemic stroke subtypes. *Stroke* 2013;44:2694-2702.

10. Adams HP, Jr., 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;24:35-41.
11. NINDS Stroke Genetics Network (SiGN), International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol* 2016;15:174-184.
12. Genome of the Netherlands Consortium. Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet* 2014;46:818-825.
13. Kong A, Steinthorsdottir V, Masson G, et al. Parental origin of sequence variants associated with complex diseases. *Nature* 2009;462:868-874.
14. Voight BF, Scott LJ, Steinthorsdottir V, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 2010;42:579-589.
15. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010;42:105-116.
16. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981-990.
17. Scott RA, Lagou V, Welch RP, et al. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet* 2012;44:991-1005.
18. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197-206.

19. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 2013;37:658-665.
20. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512-525.
21. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;40:304-314.
22. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;466:707-713.
23. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013;45:1274-1283.
24. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010;42:441-447.
25. Rolfe Ede L, Loos RJ, Druet C, et al. Association between birth weight and visceral fat in adults. *Am J Clin Nutr* 2010;92:347-352.
26. Forouhi NG, Luan J, Hennings S, Wareham NJ. Incidence of Type 2 diabetes in England and its association with baseline impaired fasting glucose: the Ely study 1990-2000. *Diabet Med* 2007;24:200-207.
27. Peters SA, Huxley RR, Woodward M. Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes. *Lancet* 2014;383:1973-1980.
28. ADVANCE-Collaborative-Group, Patel A, MacMahon S, et al. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med* 2008;358:2560-2572.

29. Action to Control Cardiovascular Risk in Diabetes Study G, Gerstein HC, Miller ME, et al. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008;358:2545-2559.
30. Kernan WN, Viscoli CM, Furie KL, et al. Pioglitazone after Ischemic Stroke or Transient Ischemic Attack. *N Engl J Med* 2016;374:1321-1331.
31. Henry RM, Kostense PJ, Spijkerman AM, et al. Arterial stiffness increases with deteriorating glucose tolerance status: the Hoorn Study. *Circulation* 2003;107:2089-2095.
32. Salomaa V, Riley W, Kark JD, Nardo C, Folsom AR. Non-insulin-dependent diabetes mellitus and fasting glucose and insulin concentrations are associated with arterial stiffness indexes. The ARIC Study. Atherosclerosis Risk in Communities Study. *Circulation* 1995;91:1432-1443.
33. van Sloten TT, Sedaghat S, Laurent S, et al. Carotid stiffness is associated with incident stroke: a systematic review and individual participant data meta-analysis. *J Am Coll Cardiol* 2015;66:2116-2125.
34. Xu M, Huang Y, Xie L, et al. Diabetes and Risk of Arterial Stiffness: a Mendelian Randomization Analysis. *Diabetes* 2016;65:1731-1740.
35. Selwaness M, van den Bouwhuijsen Q, Mattace-Raso FU, et al. Arterial stiffness is associated with carotid intraplaque hemorrhage in the general population: the Rotterdam study. *Arterioscler Thromb Vasc Biol* 2014;34:927-932.
36. Derdeyn CP. Mechanisms of ischemic stroke secondary to large artery atherosclerotic disease. *Neuroimaging Clin N Am* 2007;17:303-311, vii-viii.
37. Holmes MV, Lange LA, Palmer T, et al. Causal effects of body mass index on cardiometabolic traits and events: a Mendelian randomization analysis. *Am J Hum Genet* 2014;94:198-208.

38. Hagg S, Fall T, Ploner A, et al. Adiposity as a cause of cardiovascular disease: a Mendelian randomization study. *Int J Epidemiol* 2015;44:578-586.
39. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* 2015.

FIGURE LEGEND

Figure 1. Mendelian randomization estimates of the association between each genetically predicted risk factor and ischemic stroke and its subtypes^a

BMI, body mass index; CI, confidence interval; OR, odds ratio; T2D, type 2 diabetes.

^aAnalyses were performed using conventional Mendelian randomization analysis (inverse-variance weighted method). The scaling of the odds ratios was one unit higher log-odds for type 2 diabetes and one standard deviation increase for fasting glucose, fasting insulin, and BMI.

^bNumber of cases/controls in parenthesis.

