Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: International Stroke Genetics Consortium and Wellcome Trust Case–Control Consortium 2. Failure to validate association between 12p13 variants and ischemic stroke. N Engl J Med 2010;362:1547-50.

SUPPLEMENTARY INFORMATION

The complete list of investigators is available on page 30.

METHODS

Collaboration

The International Stroke Genetics Consortium (ISGC) is a loosely formed group of investigators who first came together in 2007 to propose, fund, and complete well-powered genetic studies of cerebrovascular disease (www.strokegenetics.org). The ISGC has implemented a centralized project proposal mechanism to facilitate the completion of stroke genetics studies, as well as the submission of collaborative grant applications.

A proposal for a validation of results from the Ikram et al. report [1], was circulated to all ISGC investigators in May 2009. Collaborators were invited to contribute either *in silico* data from genome-wide platforms or direct genotype data from available samples. The Wellcome Trust Case-Control Consortium (WTCCC) joined this project in June 2009, contributing geneotype data from genome-wide datasets undergoing quality control as part of the WTCCC2 genome-wide study of ischemic stroke. Since quality control of genotype data for the WTCCC2 participants remains in process, available data have been restricted to the rs12425791 variant and to markers at the 12p13 locus required for reliable imputation of the rs11833579 variant.

Subjects

Investigators within the ISGC provided acute ischemic and/or hemorrhagic stroke targeted genotyping (Taqman/Sequenom) or genome-wide SNP data for rs12425791 and rs11833579. We performed separate analyses based on self-reported ancestry, confirmed using multi-dimensional scaling analyses for samples with genome-wide data. A summary of demographic and SNP-based summary statistics of participating studies are shown in tables S1-3.

Four hospital-based case-control studies (Edinburgh Stroke Study - ESS, University of Maryland: Stroke Prevention in the Young Adults Study – UMD:SPYAS, Middlesex County Ischemic Stroke Study – MCISS, Greater Cincinnati/Northern Kentucky Stroke Study - GCNKSS) directly genotyped both SNPs using Taqman or Sequenom assays. Direct genotyping data were also contributed by the Vitamin Intervention for Stroke Prevention (VISP) study, a multi-center secondary stroke prevention trial[2]. Single SNP data from genome-wide arrays were provided by the Wellcome Trust Case-Control Consortium 2 (WTCCC2). This included cases from two countries, which were analyzed separately: WTCCC2: UK which comprised cases from London and Oxford with controls from the UK 1958 Birth cohort study, and WTCCC2: Germany with cases from Munich (Germany) and controls from the KORA-Gen study. Two hospital-based case-control studies (ISGS[3] and MGH/MIGen[4]) and one multicenter study of affected sib-

pairs (SWISS[5]) contributed genome-wide SNP data. Genome-wide data were also available from the Women's Genome Health Study[6,7] (WGHS), a population-based cohort study.

We analyzed South Asian (Pakistani) ancestry ischemic stroke cases and controls enrolled in the Risk Assessment of Cerebrovascular Events (RACE), including 819 ischemic strokes and 2,833 controls. Han Chinese stroke cases and controls were enrolled in the Stroke Hypertension Investigation in Genetics (SHINING)[8] study, totaling 1023 ischemic stroke cases and 940 controls. A subset of 458 ischemic stroke cases and 329 controls contributed by MCISS, UMD:SPYAS and GCNKSS were self-reported African American. Given the smaller number of samples available, we only considered replication feasible for the ischemic stroke phenotype.

Participating studies were approved by relevant institutional review boards, and all participants gave informed consent for study participation, including genetic research.

Genotyping

SNP rs12425791 was directly genotyped in all samples except WGHS, which imputed both SNPs. SNP rs11833579 was imputed[9] in samples with genome-wide array data (ISGS, SWISS, MGH/MIGen, WTCCC2: UK) and directly genotyped in all European, African American and Pakistani samples without genome-wide data.

***In silico* GWAS Dataset Preparation**

For cohorts with genome-wide SNP data, we performed quality control based on the following minimum filters:

- HWE p-value < 0.0000001
- MAF < 0.01
- missingness > 5% per individual
- missingness > 5% per SNP

We identified closely-related individuals using pairwise identity-by-descent (IBD) estimates as estimated by PLINK v 1.6 [10] (http://pngu.mgh.harvard.edu/~purcell/plink/). We excluded individuals sharing pi_hat (proportion of the genome shared due to common descent, i.e. familial relatedness) > 0.15. Identity by states (IBS) based multidimensional scaling analysis using HapMap Phase 3 populations were used to identify participants significantly different from the HapMap Phase 3 CEU / TSI reference populations. Outliers were then excluded. Quality controlled data was then reclustered using only the population being analyzed to obtain "intra-European" estimates of population stratification from multidimensional scaling analyses among cases and controls. All IBD and IBS analyses were performed using an LD-pruned SNP set in order to avoid possible confounding by linkage disequilibrium (LD). SNPs

with r^2 values > 0.2 with any SNPs in a 50 SNP sliding window were removed. This yielded a final analytical SNP set of ~30,000 - 40,000 SNPs per cohort.

Results were compared to similar principal component analyses performed using the EIGENSTRAT program implemented in the EIGENTOOLS v 3.0[11,12] software package (http://genepath.med.harvard.edu/~reich/Software.htm). We further investigated population structure in our European-American case-control GWAS datasets (MGH/MIGen and ISGS-SWISS) using STRUCTURE v 2.3.1[13] (http://pritch.bsd.uchicago.edu/structure.html) to asses admixture levels within these studies.

Using MACH v 1.0.16[9], we imputed SNPs based on HapMap Phase 3 release 27 phased haplotypes for European ancestry participants (the WGHS used release 21 phased haplotypes; ISGS-SWISS and controls from the NINDS repository used release 22 haplotypes for imputation). For the association tests, we analyzed maximum likelihood dosages for imputed SNPs and integer dosages for genotyped SNPs.

Analysis of individual studies

We used logistic regression models to test associations between rs12425791 and rs11833579 and stroke in case-control studies, including as covariates age, sex, and, if available, principal components 1 and 2 to control for population stratification[10-13]. We also used an adjusted model including additional covariates measured at the time of enrollment: history of hypertension, smoking status, systolic blood pressure, hyperlipidemia, heart disease, diabetes and atrial fibrillation. We tested associations with incident ischemic stroke in WGHS and with recurrent ischemic stroke in VISP using Cox proportional hazards model (tables S4 and S5, figures S1 and S2). In order to perform case-control analysis for ischemic stroke cases enrolled in VISP we used NINDS neurologically healthy controls[14] (http://ccr.coriell.org), genotyped on the Illumina 550 platform.

Meta-analysis

We performed fixed-effects inverse variance weighted meta-analysis combining results from individual studies participating in this validation attempt[15]. Two data analysts (M.A.N., A.B.) separately performed meta-analyses using METAL (http://www.sph.umich.edu/csg/abecasis/metal) and MetABEL[16] (tables S6, S7). We systematically investigated the possibility of between-study heterogeneity in our meta-analysis (i.e. differences in effect size estimates comparing individual studies) using Cochrane's Q (and corresponding p-value) and I^2 (percent of effect size estimate attributable to study heterogeneity) statistics.[17] Relevant heterogeneity was assumed to be present for heterogeneity p-values < 0.10[17]. Galbraith (radial) plots were generated to graphically evaluate contributions of individual cohorts to meta-analysis heterogeneity[18] (table S8, figure S3).

Power

We specified a significance threshold for replication of $p < 0.05$ (two-sided). Power in this meta-analysis was >99% to detect an effect size of 1.15 for ischemic stroke (the lower bound of the 95% CI for the original finding was 1.2[1]). For atherothrombotic stroke, power was >99% to detect an effect of 1.2 (the lower bound of the 95% CI for the original finding was 1.2[1]). Power calculations are summarized in table S9.

Participating studies

Edinburgh Stroke Study (ESS)

Cases include 968 consecutive adult self-reported Caucasian patients with ischemic stroke seen as outpatient at or admitted to the Western General Hospital (Edinburgh) from April 2002 to May 2005. All ischemic strokes were subtyped according to validated TOAST criteria. Controls are 1056 ischemic stroke-free subjects drawn from The Lothian Birth Cohort 1936, currently living in the Lothian area of Scotland (mainly in Edinburgh). Both cases and controls were genotyped at the Wellcome Trust Clinical Research Facility genetics core laboratory at the Western General Hospital (Edinburgh) on the ABI 7900HT platform, using Applied Biosystems Taqman 6.0.

Greater Cincinnati/Northern Kentucky Stroke Study (GCNKSS)

The Cincinnati Ischemic Stroke Study is comprised of 224 controls and 480 cases of confirmed ischemic stroke of self reported European ancestry, and 162 cases and 94 controls self-reported as African American at enrollment. No TOAST subtype data was available from this study. Samples were genotyped using Applied Biosystems Taqman 6.0. All samples were recruited locally from clinical centers in the Cincinnati, Ohio region.

Ischemic Stroke Genetics Study (ISGS) / Siblings With Ischemic Stroke Study (SWISS)

Cases were recruited independently by investigators from the Ischemic Stroke Genetics Study (ISGS), Siblings with Ischemic Stroke Study (SWISS) and representative of geographically heterogenous European North Americans from multiple clinical centers. Total number of acute ischemic stroke cases subtyped according to TOAST criteria is 701. Cases were genotyped using Illumina 650K Quad arrays at the Department of Molecular Neuroscience and Reta Lilla Weston Laboratories, Institute of Neurology, University College London. Controls utilized in this study are participants of the Baltimore Longitudinal Study of Aging (BLSA), for a total of 702 participants. No known first degree relatives with stroke or other neurological disease are included in the controls for the present analysis. Controls were genotyped using Illumina 550Kv1 or 550Kv3 arrays at the Laboratory of Neurogenetics, National Institute on Aging, NIH (Bethesda, MD).

Massachusetts General Hospital (MGH) / Myocardial Infarction Genetics (MIGen) Consortium

Cases included 491 consecutive patients presenting to the Neurology outpatient clinics and inpatient Neurology, Medical and Vascular Surgical services of Massachusetts General Hospital with ischemic stroke from January 2003 to July 2008. Only patients of European ancestry (confirmed by principal component analysis using genome-wide SNP data) were included in the present analysis. A stroke was defined as either (1) a radiographically proven (head CT or MRI) infarct associated with the appropriate clinical stroke syndrome or (2) a fixed neurological deficit persisting more than 24 hours, consistent with a vascular pattern of involvement and without radiographic evidence of demyelinating disease, or other non-vascular structural disease. Patients with specific vascular disorders (vasculitis, subacute bacterial endocarditis, fibromuscular dysplasia, vasospasm) were excluded from the study. Diagnostic work-up included: head CT (100%), head MRI (90%), cervical and intracranial vessel imaging using CTA or MRA (75%), ultrasound (24%), echocardiography (86%), EKG (47%) and Holter monitoring (16%.Controls are 1407 individuals matched on the basis of age, sex and ancestry information obtained from principal component analysis of GWAS data obtained from the MIGen study controls. Both cases and controls were genotyped at the Broad Institute (Cambridge, MA) using the Affymetrix 6.0 platform.

Middlesex County Ischemic Stroke Study (MCISS)

The Middlesex County Ischemic Stroke Study (MCISS) was initiated in 2001 with the primary site based at John F. Kennedy Hospital (JFK)/New Jersey Neuroscience Institute, Edison, New Jersey. A secondary site at the University of North Carolina (Department of Neurology), North Carolina was included starting in 2006. At both institutions, patients are diagnosed with a clinical suspicion of stroke following WHO guidelines and admitted to a stroke unit. During this admission, as part of routine clinical care, a standardized series of investigations are performed on all patients including blood work, a cerebral MRI/MRA (if MRI examination is contraindicated, a cranial CT scan is performed), carotid duplex ultrasound, electrocardiogram and an echocardiogram. The diagnosis of cerebral infarct is confirmed by these imaging studies. The epidemiological and clinical data on these patients is collected prospectively and entered into a database. This stroke database includes age at time of ischemic stroke, sex and ethnicity (self reported by patient or family), the presence or absence of the common risk factors (diabetes mellitus, smoking, hypertension, atrial fibrillation, homocysteine levels and cholesterol levels) and the vascular distribution of the stroke. After review of all data by two independent investigators, all strokes are classified into etiological subtypes by applying the TOAST criteria and the vascular distribution by the Oxfordshire classification system. In addition to the clinical stroke database, DNA samples have been obtained from blood samples. A control cohort of stroke free patients has also been established. These patients/volunteers have been recruited from the offices of local primary care physicians in Middlesex County, New Jersey and from the neurology clinic based at the New Jersey Neuroscience Institute. After informed consent, blood samples are collected and the clinical information is tabulated in a database.

Risk Assessment of Cerebrovascular Events (RACE)

Risk Assessment of Cerebrovascular Events (RACE) is an existing case-control study of stroke now involving over 2000 imaging confirmed cases of stroke and 2000 controls, recruited from six centers in Pakistan. DNA, plasma, serum, whole blood and information from the baseline assessment of demographic details, habits, lifestyles and other characteristics have been collected from the participants. Cases are eligible for inclusion in the study if they: (ii) are aged at least 18 years; (ii) present with a sudden onset of neurological deficit respecting a vascular territory with sustained deficit at 24 hours verified by medical attention within 72 hours after onset (onset is defined by when the patient was last seen normal and not when found with deficit); and (iii) the diagnosis is supported by CT/MRI; and (iv) present with a Modified Rankin Score < 2 prior to the stroke. The TOAST classification method is used to classify ischemic stroke based on aetiology whereas the Oxfordshire classification is used to classify stroke neuro-anatomically. Controls are selected from the same participating hospital centres as the stroke cases so that they derive from the same catchment areas. In order to minimize any potential selection biases, controls are recruited in the following order of priority: (i) non-blood related or blood related visitors of patients of the out-patient department; (ii) non-blood related visitors of stroke patients; (iii) patients of the out-patient department presenting with minor complaints (eg. back pain, minor gastric complaints). Controls with any of the following are excluded from RACE: (i) a prior history of stroke or TIA; (ii) a prior history of CAD (MI/CABG); or (iii) who are unable to provide informed consent. A locally-piloted and validated epidemiological questionnaire has been administered to participants by medically qualified research officers that seeks >400 items of information in relation to: ethnicity (eg, personal and paternal ethnicity, spoken language, place of birth and any known consanguinity); demographic characteristics; lifestyle factors (eg, tobacco and alcohol consumption, dietary intake and physical activity); personal and family history of cardiovascular disease; and medication usage. Study level results for RACE are described in table S10.

Stroke Hypertension Investigation in Genetics (SHINING)

The Stroke Hypertension Investigation in Genetics (SHINING) study was conducted by the Beijing Hypertension League Institute between 1997 and 2000. Stroke patients and control subjects from 6 geographical regions within China were recruited (70% of the subjects came from the city of Beijing and the surrounding region). The SHINING study was comprised of subjects of Han ethnicity. Stroke was diagnosed by brain CT/MRI, and cases were considered eligible if they had experienced a stroke within the previous 5 years. Patients with ischemic stroke, cerebral hemorrhage, subarachnoid hemorrhage or transient ischemic attacks (TIA) were included in the original study. Control subjects were enrolled according to study criteria during the same period (control subjects matched to cases by sex, age within three years, geographic location and blood pressure category: <140/90 mmHg, ≥140/90 mmHg and ≤ 180/105 mmHg, >180/105 mm Hg). Data collected included age, sex, body mass index (BMI), systolic

blood pressure, diastolic blood pressure and hypertension. Study level results for SHINING are described in table S11.

University of Maryland: Stroke Prevention in the Young Adults Study (UMD-SPYAS)

Cases are 513 male individuals drawn from a population-based study of first ever ischemic stroke. Participants were aged 18 - 49 years at time of index event, and were recruited from a predefined geographic area including Baltimore/ Washington DC region. Recruitment ascertained from August 2003 to November 2007. Individuals participating in the present analysis were self-reported European-American (53%), African-American (40%), or other ethnicity/race (7%). All Ischemic Strokes were subtyped according to validated TOAST criteria. Controls are 510 males, collected over the same time-period in the same geographical area and matched to cases on the basis of age, self-reported ethnicity/race and geographic location. Both cases and controls were genotyped at the University of Maryland (Baltimore, MD) using Sequenom assays.

Vitamin Intervention for Stroke Prevention (VISP)

The Vitamin Intervention for Stroke Prevention (VISP) trial was a multi-center, double-blind, randomized, controlled clinical trial that enrolled patients aged 35 or older with homocysteine levels above the 25^{th} percentile at screening and a non-disabling cerebral infarction (defined as Modified Rankin Stroke Scale ≤3) within 120 days of randomization. TOAST subtype data were unavailable from this study. Of the 56 VISP study centers, 46 centers participated in the collection of blood samples for genetic analyses. All samples underwent whole-genome amplification using RepliG methodology through Molecular Staging Inc. (New Haven, CT) or Qiagen (Hilden, Germany). Samples were genotyped at the Center for Public Health Genomics, University of Virginia, using the ABI 7900HT platform and Applied Biosystems Taqman 6.0. The trial was designed to determine if daily intake of a multivitamin tablet with high dose folic acid, vitamin B_6 and vitamin B_{12} reduced recurrent cerebral infarction. Enrollment in VISP was completed in December 2001, with 3,680 participants enrolled. In December, 2002, the VISP Data and Safety Monitoring Board recommended to NINDS that the trial be terminated due to the unlikely event of ever demonstrating a treatment difference. All participants had been in the study long enough for at least one year of follow-up. The definition of recurrent ischemic stroke was an acute neurological ischemic event of at least 24 hours duration with focal signs and symptoms, without evidence of primary intracranial hemorrhage or other alternative explanation, together with one of the following: a one-point increase in the NIHSS in a previously normal section or, lacking this, an appropriate new or extended abnormality seen on CT or MRI. Diagnoses were reviewed by the local neurologist, two endpoint reviewers, and occasionally a full Stroke Endpoint Review Committee. In order to perform case-control analysis for ischemic stroke using the VISP data we used NINDS neurologically healthy controls, genotyped on the Illumina 550 platform. Given the lack of available GWAS data for cases in the VISP dataset, we did not include principal components obtained from multidimensional scaling analysis of

controls in our model. Given the lack of African-American controls in the NINDS dataset, we restricted the case-control analysis to European-ancestry participants in VISP.

Wellcome Trust Case-Control Consortium 2 (WTCCC2)

The WTCCC2 samples were genotyped as part of the WTCCC 2 ischemic stroke study. Stroke cases included samples recruited by investigators at St. George's University London (SGUL) and University of Oxford in the UK and the Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-University, Munich. Controls for the UK samples were drawn from shared WTCCC controls obtained from the 1958 Birth Cohort. For the German samples controls were Caucasians of German origin participating into the population KORAgen study (www.gsf.de/kora/en/english.html). This survey represents a gender- and age stratified random sample of all German residents of the Augsburg area and consists of individuals 25 to 74 years of age, with about 300 subjects for each 10-year increment. All controls were free of a history of stroke or transient ischemic attack. 817 individuals were used in the present analysis as controls for the Munich ischemic stroke cases. The SGUL collection comprised 1224 ischemic stroke samples from a hospital based setting. All cases were of self reported Caucasian ancestry. Ischemic stroke subtypes were determined according to TOAST criteria based on relevant clinical imaging and available information on cardiovascular risk factors. Cases were genotyped as part of the WTCCC2 on the Illumina 660 Quad platform. Genotypes for rs12425791 were extracted from the genome wide SNP data. The University of Oxford collection comprised 896 ischemic stroke cases, consecutively collected as part of the Oxford vascular study (OXVASC). Cases were of self reported Caucasian ancestry, and ischemic stroke subtypes were determined according to TOAST criteria based on relevant clinical imaging. The Munich samples included 1383 ischemic stroke cases. Cases were consecutive European Caucasians recruited from a single dedicated Stroke Unit at the Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-University, Munich. Ischemic stroke subtypes were determined according to TOAST criteria based on relevant clinical and imaging data. All cases were genotyped as part of the WTCCC2 ischemic stroke study on the Illumina 660 Quad platform. For the WTCCC2 UK samples a common control set was used comprising subjects from the 1958 Birth Cohort. This is a prospectively collected cohort of individuals born in 1958 (http://www.b58cgene.sgul.ac.uk/), and ascertained as part of the national child development study (http://www.cls.ioe.ac.uk/studies.asp?section=000100020003). Data from this cohort are available as a common control set for a number of genetic and epidemiological studies. 2930 controls were used in this analysis as controls for ischemic stroke cases collected in the US. Due to the non availability of clinical information (other than age and sex) in this cohort adjusted models we did not use adjusted models when exploring the association of the variants on chromosome 12p13 with ischemic stroke and TOAST subtypes. Controls were genotyped as part of the WTCCC2 on the Illumina 1M platform. For the UK case-controls samples genotypes for rs12425791 in both cases and controls were extracted from the genome wide SNP data. Genotype data for rs11833579 in the WTCCC cohort were obtained via imputation from available genotyped SNPs at the 12p13 locus, using IMPUTE v 0.0.5 and

HapMap Phase 2 (release 24) phased CEU haplotypes. Due to the lack of genome-wide SNP data for stroke samples genotyped as part of WTCCC 2 ischemic stroke study, WTCCC data were treated as if obtained from direct (targeted) genotyping, and analyzed accordingly. Specifically, since complete GWAS data were not available, multidimensional scaling analyses could not be implemented during quality control and components vectors from these analyses were not available as covariates. All participants were self-identified as European ancestry. For the German cases and controls, genotyping was performed on the Illumina 550 platform and genotypes for rs12425791 were extracted from the genome wide SNP data.

Women's Genome Health Study (WGHS)

The 22,054 individuals in the WGHS are participants in the ongoing Women's Health Study (WHS) who provided baseline blood samples and who have both genome-wide genotype data and verified self-reported European ancestry. The WHS was initiated in 1992 to evaluate low-dose aspirin and vitamin E in the primary prevention of cardiovascular disease and cancer in women over approximately 10 years, and now continues in observational mode. During follow-up, 278 incident ischemic strokes were identified in the WGHS subset. All ischemic strokes were subtyped according to validated TOAST criteria. Genotyping was performed on the Illumina HumanHap300 Duo "+" platform. Imputation of genotypes was performed using Mach v 1.0.16 using HapMap release 21.

Gene expression

We analyzed gene expression for NINJ2 and WNK1 using different approaches in two patient samples. In the first approach, we investigated gene expression levels at different time points after acute ischemic stroke onset (3h, 5h, 24h) in a sample of 15 patients, recruited as part of the CLEAR trial[19], and 8 healthy controls. In the second approach, we compared expression levels at one time point after acute ischemic stroke in a larger cohort comprised of 132 ischemic strokes and 80 controls.

Subjects

We investigated gene expression patterns for NINJ2 and WNK1 following acute ischemic stroke in 15 patients recruited through the CLEAR trial (under identifier NCT00250991 at Clinical-Trials.gov) from July 2003 to July 2004. Computed tomography brain scans were performed in all patients to exclude hemorrhage. Peripheral whole blood (15 mL) was drawn from each patient via venipuncture before treatment (< 3 h samples), approximately 2 h after the thrombolysis treatment and 5 hours after stroke (5 h samples) and 24 h after the stroke onset (24 h samples). A total of 45 blood samples were collected from the 15 stroke patients. The first cohort of control peripheral blood samples were drawn from 8 healthy volunteers. Each volunteer contributed two independent blood samples 1 day apart. Healthy controls had no history of cardiovascular or cerebral vascular disease, recent infection or hematologic disease.

We also explored possible changes in expression for genes at the chromosome 12p13 locus using a second, larger cohort of acute ischemic stroke patients compared to stroke-free subjects (suffering from various other medical conditions) as controls. One hundred and thirty two subjects with acute ischemic stroke (mean age=69.6 +/- 9.1 years) were recruited between 2004 and 2008 within one week of stroke onset. These included patients with atherosclerotic (n=34), cardioembolic (n=86), unknown etiology ischemic stroke (n=7), and small vessel (n=5) subtypes. Eighty control subjects (mean age=49.5 +/-21.0 years) were recruited between 2005 and 2008. These included subjects with intracerebral hemorrhage (n=15), a*rterio-venous malformation (*AVM) with (n=22) or without hemorrhage (n=20), healthy controls (n=18), seizure (n=1) and syncope (n=1). Demographic data is presented in Supplementary Table 1. Total RNA was purified from peripheral whole blood and the same Affymetrix Human U133 Plus 2.0 GeneChip microarrays were used for expression measurements.

RNA Isolation and Affymetrix Array Hybridization

Total RNA from the whole blood was isolated according to the manufacturer's protocol (PAXgene blood RNA kit; Pre-AnalytiX). RNA samples were examined for concentration and purity using a Nanodrop ND-1000 spectrophotometer, and integrity was checked using the Agilent 2100 Bioanalyzer (260/280>2.0; 28S/18S>1.5; RIN >8). RNA samples were labeled using the One-Cycle Target Labeling protocol. Genomic microarray gene expression analysis was performed according to standard Affymetrix Protocols (Affymetrix Expression Analysis Technical Manual; Affymetrix, Santa Clara, CA, USA). Affymetrix Human U133 Plus 2.0 Arrays, which contain more than 54,000 probe sets, were used for each RNA sample. NINJ2 was represented on the array by one probe-set whereas WNK1 was represented by six probe-sets.

Microarray Data Analysis

Probe-level data were saved in Affymetrix .cel files and summarized with Robust Multi-array Average (RMA) software (http://www.bioconductor.org/)[20]. One array from one healthy control of the first cohort of subjects was found to be an outlier and was eliminated. For the analysis on the first cohort of ischemic stroke patients and control subjects, statistical analyses including a one-way analysis of variance (ANOVA) and Student's t-test were performed using Genespring software (Silicon Genetics, Redwood City, CA, USA) to determine whether the genes NINJ2 and WNK1 were significantly differentially expressed between the ischemic stroke and control groups. The Benjamini–Hochberg FDR (false discovery rate) was used to control for multiple comparisons. The methods have been described in detail in our published studies[21,22]. For the analysis on the second larger cohort of stroke patients and controls, analysis of covariance (ANCOVA) was performed in Partek Genomics Suit, version 6.4, release 6.09.0422 (Partek Inc., St. Louis, MI, USA). The factors in the ANCOVA model were: diagnosis, age, gender, race and batch. Results of these analyses may be found in tables S12-S13 and figure S4.

Table S1. Demographic characteristics

	ESS	GCNKSS	ISGS-SWISS	MGH/ MIGen	MCISS	RACE	SHINING	UMD-SPYAS	VISP[§]	WGHS	WTCCC2: Germany	WTCCC2: UK[#]
Ethnicity	EU	EU / AA [Δ]	EU	EU	EU / AA [Δ]	SA	CH-HAN	EU / AA[Δ]	EU / AA [Δ]	EU	EU	EU
Number of stroke cases												
Ischemic	968	480	700	491	550	819	1023	271	1778	278	1376	2023
- Atherothrombotic	258	n/a	241	132	174	376	n/a	65	n/a	64	469	922
- Small vessel	189	n/a	94	37	74	158	n/a	39	n/a	54	112	419
- Large artery	69	n/a	147	95	100	218	n/a	26	n/a	10	357	503
- Cardioembolic	115	n/a	193	170	145	201	n/a	53	n/a	49	351	442
Number of controls	1056	224	702	1407	536	2833	940	278	787	21777	817	2926
Female	0.46	0.51	0.44	0.49	0.0	0.30	0.40	0.0	0.42	1.0	0.40	0.47
Age at enrollment (mean, SD)	69.3 (4.7)	67.9 (13.7)	68.8 (15.2)	55.6 (4.6)	71.2 (11.8)	54.5 (11.2)	60.0 (10.4)	41.6 (6.1)	65.1 (13.4)	54.6 (7.0)	64.5 (12.1)	66.3 (11.2)
Hypertension	0.53	0.63	0.55	0.52	0.72	0.36	0.82	0.28	0.55	0.26	0.57	n/a
Type 2 diabetes	0.08	0.28	0.17	0.06	0.28	0.19	n/a	0.08	0.19	0.03	0.14	n/a
Atrial fibrillation	0.21	0.13	.12	0.12	0.19	n/a	n/a	n/a	n/a	0.09	0.18	n/a
Current smoker	0.21	0.24	0.09	0.21	0.23	0.23	n/a	0.29	0.22	0.13	0.08	n/a
Cardiovascular disease	0.27	0.15	0.12	0.11	0.30	0.11	n/a	0.08	n/a	0.10	0.10	n/a

ESS = Edinburgh Stroke Study, GCNKSS = Greater Cincinnati/Northern Kentucky Stroke Study, ISGS = Ischemic Stroke Genetics Study, SWISS = Siblings with Ischemic Stroke Study, MGH = Massachusetts General Hospital, MIGen = Myocardial Infarction Genetics Consortium, , MCISS = Middlesex County Ischemic Stroke Study, RACE = Risk Assessment of Cerebrovascular Events, SHINING = Stroke Hypertension Investigation in Genetics, UMD-SPYAS = University of Maryland: Stroke Prevention in the Young Study, VISP = Vitamin Intervention for Stroke Prevention, WGHS = Women's Genome Health Study, WTCCC2 = Wellcome Trust Case-Control Consortium 2. SD = Standard Deviation, EU = European Ancestry, AA = African-American Ancestry, SA = South Asian Ancestry, CH-Han = Han Chinese ancestry

§ European ancestry ischemic stroke cases participating in VISP were matched with neurologically healthy NINDS controls for the purpose of case-control analysis. Reported characteristics refer to the integrated case-control dataset (including only European ancestry individuals).

Additional clinical information were not available for controls from the Birth cohort 1958 matched to WTCCC 2 ischemic stroke cases. Therefore demographic characteristics for the combined case-control dataset are not reported, and we limited our analysis to the basic model for this cohort.

Δ These studies enrolled both European-ancestry and African-American ancestry stroke cases and controls. Demographic characteristics reported in this table refer to European-ancestry participants only.

Table S2. Study-specific SNP data for individuals of European ancestry

	rs12425791					rs11833579						
	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Platform	r^2
ESS	A	G	0.19	0.96	n/a	A	G	0.25	0.97	n/a	Taqman	0.73
GCNKSS	A	G	0.21	0.96	n/a	A	G	0.24	0.96	n/a	Taqman	0.72
ISGS-SWISS	A	G	0.20	1.0	n/a	A	G	0.24	1.0	0.99	Illumina 610	0.78
MCISS	A	G	0.22	1.0	n/a	A	G	0.27	1.0	n/a	Taqman	0.72
MGH/MIGen	A	G	0.21	1.0	n/a	A	G	0.23	1.0	0.96	Affymetrix 6.0	0.71
UMD-SPYAS	A	G	0.21	1.0	n/a	A	G	0.26	1.0	n/a	Sequenom	0.75
VISP*	A	G	0.20	0.96	n/a	A	G	0.26	0.95	0.99	Taqman/Illumina550	0.74
WGHS	A	G	0.18	1.0	0.60	A	G	0.23	1.0	0.86	Illumina 330	0.66
WTCCC2: Germany	A	G	0.20	0.99	n/a	-	-	-	-	-	Illumina660/Illumina550	-
WTCCC 2:UK	A	G	0.19	0.99	n/a	A	G	0.23	1.0	0.99	Illumina660/Illumina1M	0.79

ESS = Edinburgh Stroke Study, GCNKSS = Greater Cincinnati /Northern Kentucky Stroke Study, ISGS = Ischemic Stroke Genetics Study, SWISS = Siblings with Ischemic Stroke Study, MCISS = Middlesex County Ischemic Stroke Study, MGH = Massachusetts General Hospital, MIGen = Myocardial Infarction Genetics Consortium, UMD-SPYAS = University of Maryland: Stroke Prevention in the Young Study, WGHS = Women's Genome Health Study, VISP = Vitamin Intervention for Stroke Prevention, WTCCC2 = Wellcome Trust Case-Control Consortium2.
MAF = Minor Allele Frequency,
Imputation Quality = ratio of empirically observed variance of allele dosage and expected binomial variance at Hardy-Weinberg equilibrium based on HapMap data, r^2 = R squared correlation between rs12425791 and rs11833579

* Denotes cases from the VISP study based on Taqman assay combined with Illumina/imputed controls from the NINDS neurological controls dataset

Table S3. Study-specific SNP data for individuals of African-American ancestry

	rs12425791					rs11833579						
	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Platform	r^2
GCNKSS	A	G	0.11	0.98	n/a	A	G	0.22	0.99	n/a	Taqman	0.46
MCISS	A	G	0.12	1.0	n/a	A	G	0.21	1.0	n/a	Taqman	0.48
UMD-SPYAS	A	G	0.11	1.0	n/a	A	G	0.21	1.0	n/a	Sequenom	0.49

GCNKSS = Greater Cincinnati/Northern Kentucky Stroke Study, MCISS = Middlesex County Ischemic Stroke Study, UMD-SPYAS = University of Maryland: Stroke Prevention in the Young Study

Imputation Quality = ratio of empirically observed variance of allele dosage and expected bomial variance at Hardy-Weimberg equilibrium based on HapMap data, r^2 = R squared correlation between rs12425791 and rs11833579

Table S4. Cox and logistic regression for incident ischemic stroke in the Women's Genome Health Study (WGHS)

rs12425791	Cox model			Logistic model		
	Harzard Ratio	**Conf. Interval**	**p**	**Odds ratio**	**Conf. Interval**	**p**
Ischemic Stroke	1.00	0.75 – 1.33	1.00	1.01	0.75 – 1.34	0.98

rs11833579	Cox model			Logistic model		
	Harzard Ratio	**Conf. Interval**	**p**	**Odds ratio**	**Conf. Interval**	**p**
Ischemic Stroke	1.10	0.87 – 1.38	0.43	1.10	0.87 – 1.39	0.42

Table S5. Cox model and logistic model analysis for ischemic stroke recurrence in Vitamin Intervention for Stroke Prevention (VISP) depending on minor allele dosage

rs12425791	Cox model			Logistic model		
	Harzard Ratio	**Conf. Interval**	**p**	**Odds ratio**	**Conf. Interval**	**p**
European-ancestry						
- Basic model	1.09	0.83 - 1.44	0.52	1.10	0.84 - 1.45	0.52
- Adjusted model	1.09	0.83 - 1.44	0.53	1.10	0.83 - 1.45	0.50
African-American						
- Basic model	0.80	0.34 - 1.92	0.62	0.82	0.36 - 1.98	0.66
- Adjusted model	0.82	0.34 - 1.7	0.66	0.82	0.37 - 2.01	0.70

rs11833579	Cox model			Logistic model		
	Harzard Ratio	**Conf. Interval**	**p**	**Odds ratio**	**Conf. Interval**	**p**
European-ancestry						
- Basic model	1.05	0.81 - 1.36	0.71	1.06	0.81 - 1.37	0.69
- Adjusted model	1.05	0.82 - 1.37	0.68	1.07	0.82 - 1.40	0.65
African-American						
- Basic model	0.82	0.41 - 1.61	0.56	0.83	0.41 - 1.63	0.59
- Adjusted model	0.81	0.41 - 1.60	0.54	0.82	0.41 - 1.62	0.57

Table S6. Meta-analysis of individuals of European ancestry

rs12425791	Minor (coded) Allele	Major (non-coded) Allele	Weighted MAF	Basic model				Adjusted model			
				Odds Ratio	95% Conf. Interval	p-value	Het. p-value	Odds Ratio	95% Conf. Interval	p-value	Het. p-value
Ischemic Stroke	A	G	0.19	0.97	0.91 - 1.04	0.41	0.72	0.96	0.89 - 1.03	0.21	0.69
- Atherothrombotic	A	G	0.19	0.98	0.88 - 1.09	0.65	0.58	0.96	0.85 - 1.08	0.49	0.37
- Large Artery	A	G	0.19	0.86	0.82 - 1.08	0.39	0.69	0.88	0.74 - 1.04	0.13	0.20
- Small Vessel	A	G	0.19	1.02	0.99 - 1.04	0.15	0.19	1.05	0.89 - 1.24	0.56	0.50
- Cardioembolic	A	G	0.19	0.97	0.86 - 1.10	0.66	0.89	1.03	0.88 - 1.21	0.72	0.97

rs11833579	Minor (coded) Allele	Major (non-coded) Allele	Weighted MAF	Basic model				Adjusted model			
				Odds Ratio	95% Conf. Interval	p-value	Het. p-value	Odds Ratio	95% Conf. Interval	p-value	Het. p-value
Ischemic Stroke	A	G	0.23	1.02	0.95 - 1.10	0.55	0.58	1.02	0.95 - 1.10	0.54	0.78
- Atherothrombotic	A	G	0.23	1.09	0.96 - 1.23	0.18	0.89	1.07	0.94 - 1.22	0.30	0.83
- Large Artery	A	G	0.23	1.12	0.95 - 1.33	0.18	0.97	1.13	0.94 - 1.36	0.19	0.96
- Small Vessel	A	G	0.23	1.05	0.90 - 1.23	0.52	0.42	1.04	0.88 - 1.22	0.66	0.45
- Cardioembolic	A	G	0.23	1.03	0.89 - 1.20	0.68	0.86	1.03	0.87 - 1.22	0.73	0.74

Weighted MAF = Meta-analysis wide Minor Allele Frequency computed as weighted average of Minor Allele Frequencies in individual studies depending on sample size. Het = Between-study heterogeneity.
Basic model included age, sex and principal components 1 and 2 from multidimensional scaling analysis. Adjusted model also included history of hypertension, smoking status, systolic blood pressure, hyperlipidemia, and atrial fibrillation.
Adjusted model does not include WTCCC-UK (no additional clinical information available for controls).

Table S7. Meta-analysis of inidividuals African-American ancestry

rs12425791	Minor (coded) Allele	Major (non-coded) Allele	Weighted MAF	Odds Ratio	95% Conf. Interval	p-value	Het. p-value
Ischemic Stroke							
- Basic Model	A	G	0.11	0.90	0.66 - 1.24	0.53	0.75
- Adjusted Model	A	G	0.11	0.97	0.69 -1.37	0.86	0.68

rs11833579	Minor (coded) Allele	Major (non-coded) Allele	Weighted MAF	Odds Ratio	95% Conf. Interval	p-value	Het. p-value
Ischemic Stroke							
- Basic Model	A	G	0.21	0.86	0.67 - 1.10	0.24	0.88
- Adjusted Model	A	G	0.21	0.92	0.70 - 1.21	0.55	0.74

Weighted MAF = Meta-analysis wide Minor Allele Frequency computed as weighted average of Minor Allele Frequencies in individual studies depending on sample size

Table S8 Fixed-effects meta-analysis for European ancestry individuals, including validation cohorts and results from Ikram et al., 2009 [1]

rs12425791	Odds Ratio	95% Confidence Interval	p-value	Heterogeneity p-value
Ischemic Stroke	1.09	1.03 - 1.15	0.003	< 0.001
- Atherothrombotic Stroke	1.17	1.09 - 1.26	1.6×10^{-5}	0.003

rs11833579	Odds Ratio	95% Confidence Interval	p-value	Heterogeneity p-value
Ischemic Stroke	1.13	1.06 - 1.19	6.3×10^{-5}	< 0.001
- Atherothrombotic Stroke	1.22	1.13 - 1.21	7.1×10^{-8}	0.0021

Table S9. Statistical Power: Meta-analysis of individuals of European ancestry

	GWAS Discovery Meta-Analysis[1]			Dutch Case-Control Replication			ISGC Replication Meta-Analysis					
rs12425791	**No. of Cases**	**Hazard Ratio**	**95% Conf. interval**	**No. of Cases**	**Hazard Ratio**	**95% Conf. interval**	**Equivalent No. of Case-Control Pairs**	**Odds Ratio**				
								1.05	**1.10**	**1.15**	**1.2**	**1.3**
Ischemic Stroke	1164	1.33	1.21 – 1.47	501	1.19	1.01 – 1.41	8756	0.48	0.98	> 0.99	> 0.99	> 0.99
- Atherothrombotic	876	1.37	1.23 – 1.54	400	1.29	1.08 – 1.54	3393	0.22	0.64	0.95	>0.99	>0.99

	GWAS Discovery Meta-Analysis[1]			Dutch Case-Control Replication			ISGC Replication Meta-Analysis					
rs11833579	**No. of Cases**	**Hazard Ratio**	**95% Conf. interval**	**No. of Cases**	**Hazard Ratio**	**95% Conf. interval**	**Equivalent No. of Case-Control Pairs**	**Odds Ratio**				
								1.05	**1.10**	**1.15**	**1.2**	**1.3**
Ischemic Stroke	1164	1.33	1.21 – 1.46	501	1.10	n/a	7744	0.48	0.97	> 0.99	> 0.99	> 0.99
- Atherothrombotic	876	1.35	1.21 – 1.50	400	1.19	1.00 – 1.40	2829	0.19	0.57	0.88	0.98	>0.99

All power calculations assume:
- Additive model
- alpha = 0.05
- prevalence = 0.03
- power calculated at locus (D' = 1)

Effective sample size was determined by computation of equivalent symmetric (i.e. case:control ratio = 1) sample size based on non-centrality parameter (NCP) for each participating study using the Genetic Power Calculator [33].
Results were confirmed using the Power for Association with Error (PAWE) [34] program.
Single-study sample sizes were weighted according to imputation quality and call rate, and summed to obtain final sample size.
ns = not significant

Table S10. Logistic regression for ischemic in the Risk Assessment for Cerebrovascular Events (RACE) study, Pakistan

rs12425791	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Odds Ratio	95% Conf. Interval	p-value
Ischemic Stroke								
- Basic model	A	G	0.09	0.99	n/a	0.81	0.68 – 1.00	0.053
- Adjusted model	A	G	0.09	0.99	n/a	0.86	0.71 – 1.05	0.20

rs11833579	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Odds Ratio	95% Conf. Interval	p-value
Ischemic Stroke								
- Basic model	A	G	0.36	0.99	n/a	1.05	0.92 - 1.19	0.45
- Adjusted model	A	G	0.36	0.99	n/a	1.03	0.90 - 1.17	0.67

Table S11. Logistic regression for ischemic in the Stroke Hypertension Investigation in Genetics (SHINING) study, China

rs12425791	Minor (coded) Allele	Major Allele	MAF	Call Rate	Imputation Quality	Odds Ratio	95% Conf. Interval	p-value
Ischemic Stroke								
- Basic model	A	G	0.27	0.98	n/a	1.04	0.90 – 1.20	0.58
- Adjusted model	A	G	0.27	0.98	n/a	1.02	0.91 – 1.15	0.69

MAF = Minor allele frequency
Basic model included age and sex. Adjusted model also included history of hypertension and systolic blood pressure.

Table S12. Gene expression activities of NINJ2 and WNK1 at different time points in peripheral blood cells of acute ischemic stroke patients compared to controls

Gene	Probe	1-way ANOVA Control, 3hr, 5hr, 24hr)		Student T test					
				3 hours		5 hours		24 hours	
		Unadjusted	FDR corrected	Unadjusted	FDR corrected	Unadjusted	FDR corrected	Unadjusted	FDR corrected
NINJ2	219594_at	0.192	0.55	0.175	0.755	0.05	0.24	0.08	0.274
WNK1	1555068_at	0.616	0.838	0.392	0.901	0.39	0.663	0.194	0.451
WNK1	202940_at	0.687	0.871	0.325	0.871	0.896	0.959	0.503	0.728
WNK1	211992_at	0.112	0.436	0.279	0.842	0.046	0.225	0.117	0.347
WNK1	211993_at	0.039	0.262	0.034	0.429	0.048	0.229	0.017	0.128
WNK1	211994_at	0.233	0.595	0.099	0.637	0.093	0.326	0.0852	0.294
WNK1	39313_at	0.719	0.885	0.452	0.923	0.585	0.80	0.273	0.537

FDR: Benjamini–Hochberg False Discovery Rate Control for Multiple Comparisons

Table S13. Gene expression activities of NINJ2 and WNK1 in peripheral blood cells of acute ischemic stroke patientscompared to controls

Gene	Affymetrix Probe Sets	4-way ANOVA (Diagnosis, Age, Gender, Race, Batch)		
		p-value		Fold Change (Ischemic Stroke/Control)
		Unadjusted	FDR-corrected	
NINJ2	219594_at	0.48	0.75	1.04
WNK1	1555068_at	0.19	0.47	1.22
WNK1	39313_at	0.22	0.51	1.06
WNK1	211994_at	0.53	0.78	1.04
WNK1	211992_at	0.72	0.89	1.02
WNK1	211993_at	0.72	0.89	-1.03
WNK1	211940_at	0.87	0.95	1.01

FDR: Benjamini–Hochberg False Discovery Rate Control for Multiple Comparisons

Figure S1. Kaplan-Meier plots for incident ischemic stroke in the Women's Genome Health Study (WGHS) stratified by minor allele dosage

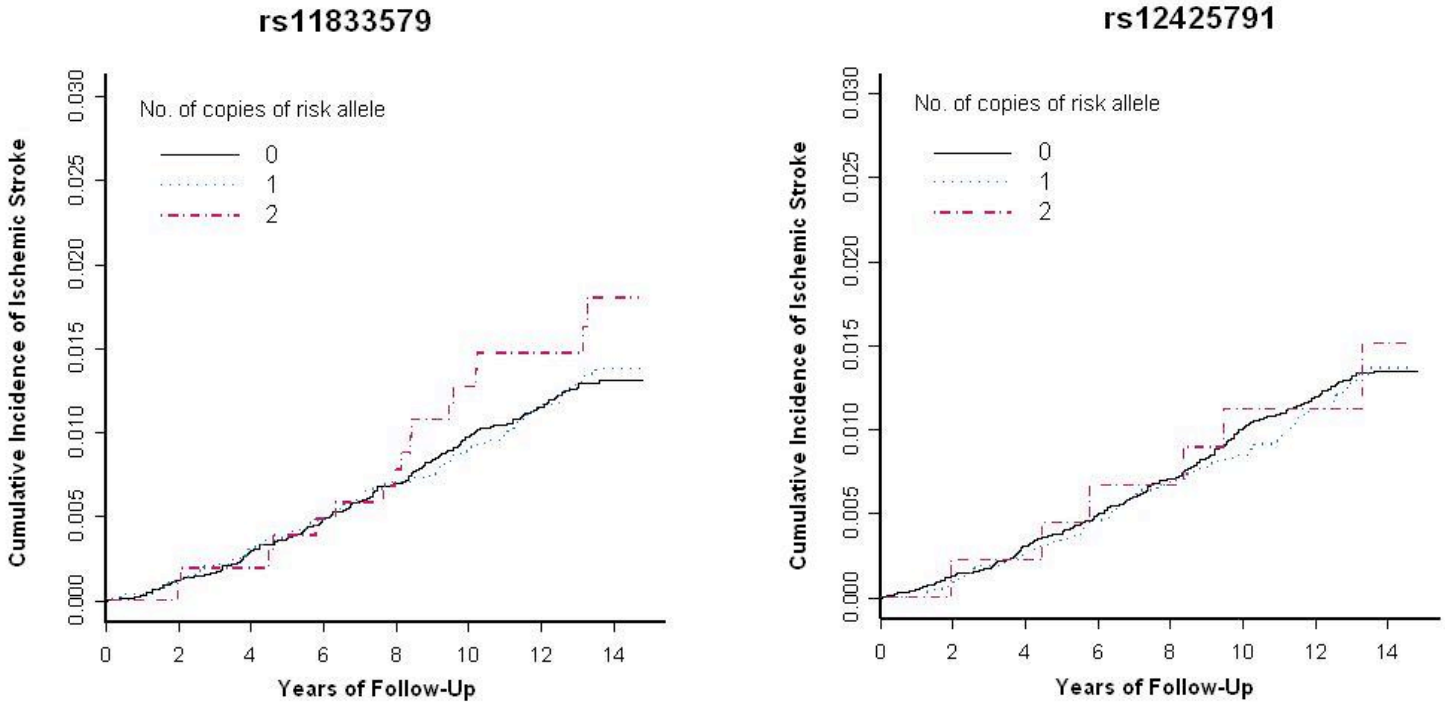


Figure S2. Kaplan-Meier plots for ischemic stroke recurrence in the Vitamin Intervention for Stroke Prevention Study (VISP) stratified by minor allele dosage in participants of European ancestry.

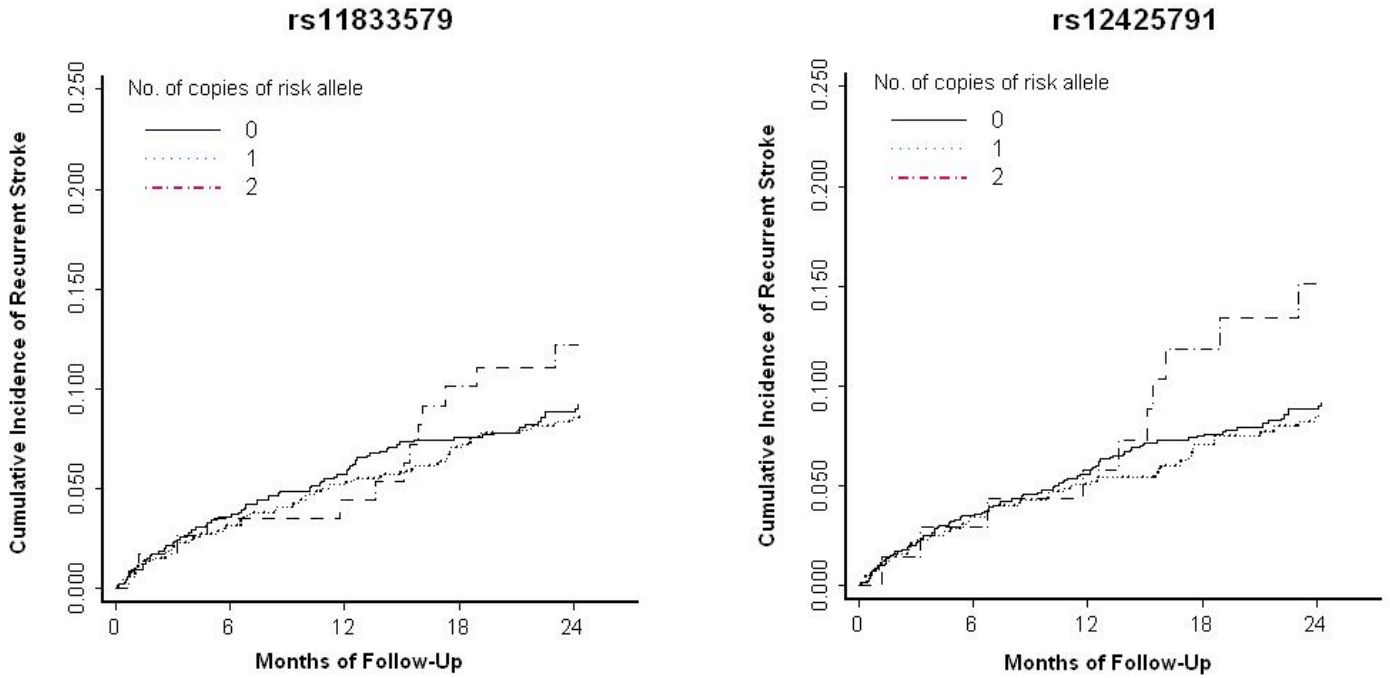
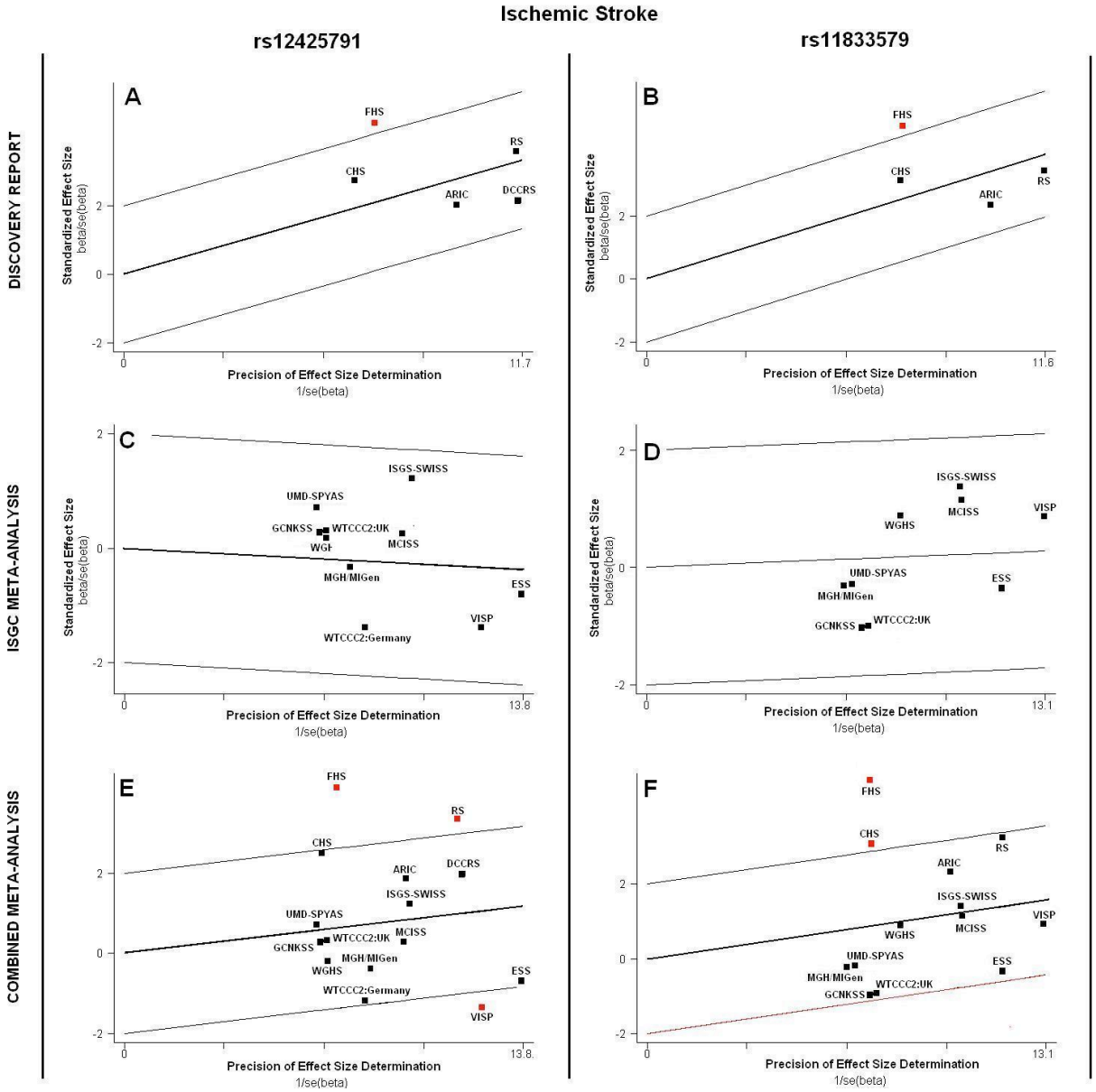


Figure S3. Galbraith plots of meta-analyses of ischemic stroke from Ikram et al. demonstrate significant heterogeneity

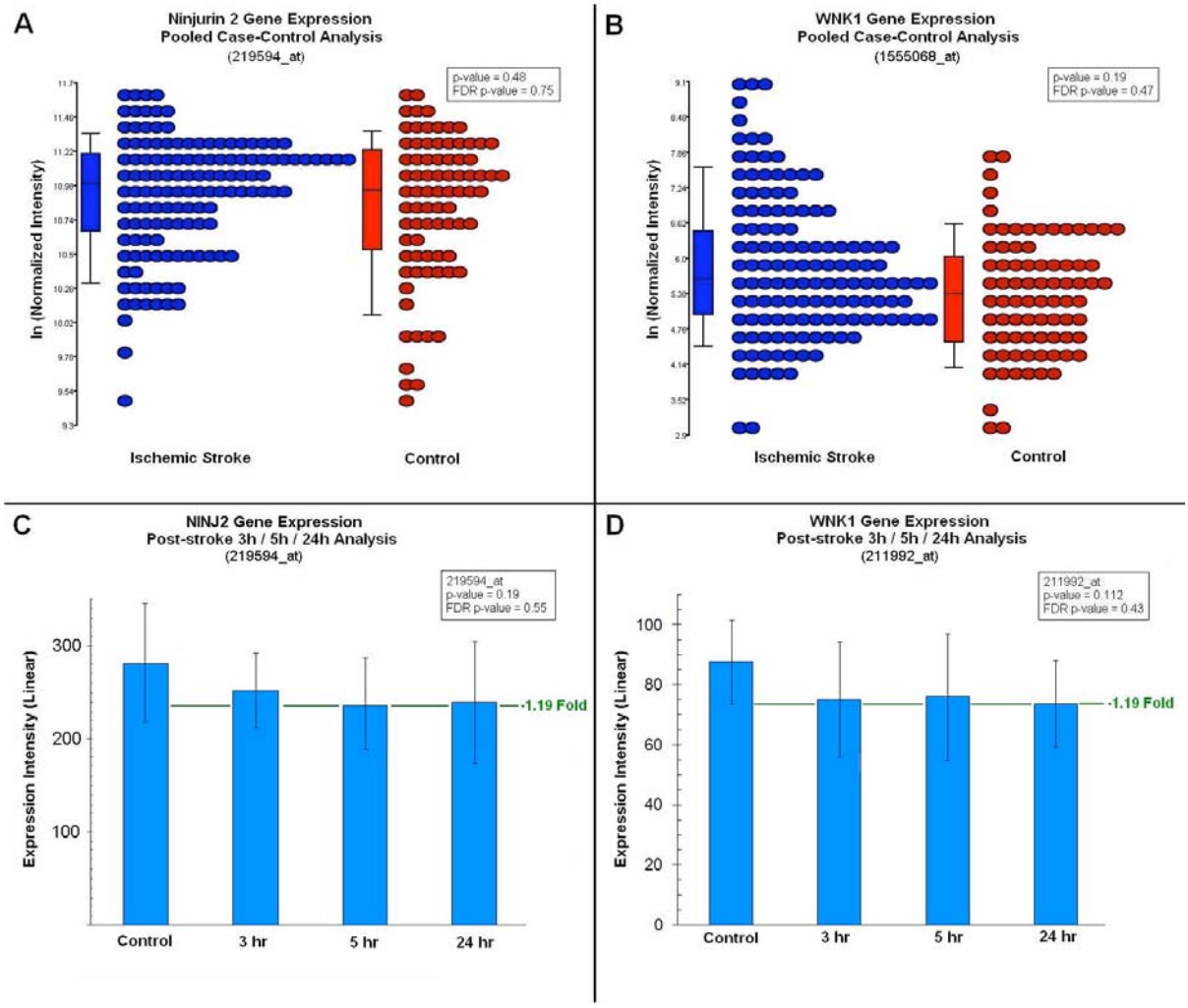


Galbraith plots examine meta-analysis heterogeneity by plotting the standardized effect size (as determined by the ratio of the effect size and the standard error) and the precision of the effect size determination (inverse of the standard error) for each study. The horizontal position of the graph corresponds to increasing weight assigned to each study due to increased precision (smaller standard error). The vertical position corresponds to the degree of heterogeneity compared to the thick line (fixed-effects meta-analysis point estimate of effect size) and the thin lines (95% Confidence Interval boundaries for the fixed-effects estimate). In the absence of significant heterogeneity all studies should be positioned within the 95% confidence interval estimate provided by the fixed-effects meta-analysis.
Meta-analysis heterogeneity is graphically inspected in the original discovery report (A: rs12425791, B: rs11833579), in the ISGC replication meta-analyses (C: rs12425791, D: rs11833579) and combining all available data on each variant (E: rs12425791, F: rs11833579).

ARIC = Atherosclerosis Risk In Communities, CHS = Cardiovascular Health Study, ESS = Edinburgh Stroke Study, FHS = Framingham Heart Study, GCNKSS = Greater Cincinnati/ Northern Kentucky Stroke Study, ISGS = Ischemic Stroke Genetics Study, SWISS = Siblings with Ischemic Stroke Study, MCISS = Middlesex County Ischemic Stroke Study, MGH = Massachusetts General Hospital, MIGen = Myocardial Infarction Genetics Consortium, RS = Rotterdam Scan Study, UMD-SPYAS = University of Maryland: Stroke Prevention in the Young Study, VISP = Vitamin Intervention for Stroke Prevention, WGHS = Women's Genome Health Study, WTCCC2 = Wellcome Trust Case-Control Consortium2.

MAF = Minor Allele Frequency, se = standard error

Figure S4. Gene expression analyses for NINJ2 and WNK1



A. Pooled case-control analysis for Ninjurin2 in 132 Ischemic Stroke cases and 80 Stroke-free Controls.

B. Pooled case-control analysis for WNK1 (With No lysine K 1) in 132 Ischemic Stroke cases and 80 Stroke-free Controls.

C. Gene expression activities of NINJ2 in blood cells of acute ischemic stroke patients within 24hr of onset, represented by Affymetrix probes set 219594_at. X-axis shows the conditions: control (Healthy Control); 3hr, 5hr and 24hr (3hr, 5hr and 24hr after onset of ischemic stroke). Y-axis shows the linear expression intensity. The maximum fold of gene expression differences in ischemic stroke patients compared to control is marked to the right of each panel. Significance computed by one-way ANOVA (p = 0.19) and adjusted for multiple comparisons via False Discovery Rate (FDR) method (FDR p-value = 0.55). Unadjusted p-values at individual time points were as follows: 3 hours: p = 0.17 (FDR p = 0.76); 5 hours: p = 0.052 (FDR p = 0.24); 24 hours: p = 0.075 (FDR p = 0.27).

D. Gene expression activity of WNK1 in blood cells of acute ischemic stroke patients within 24hr of onset, represented by Affymetrix probes set 211993_at. X-axis shows the conditions: control (Healthy Control); 3hr, 5hr and 24hr (3hr, 5hr and 24hr after onset of ischemic stroke). Y-axis showed the linear expression intensity. The maximum fold of gene expression differences in ischemic stroke patients compared to control is marked to the right of each panel. Significance computed by one-way ANOVA (p = 0.04) and

adjusted for multiple comparisons via False Discovery Rate (FDR) method (FDR p-value = 0.26). Unadjusted p-values at individual time points were as follows: 3 hours: p = 0.034 (FDR p = 0.64); 5 hours: p = 0.048 (FDR p = 0.23); 24 hours: 211992_at p = 0.117 (FDR p = 0.35) and 211993_at p = 0.017 (FDR p = 0.13).

Supplementary References:

1. Ikram MA, Seshadri S, Bis JC, Fornage M, et al. Genomewide association studies of stroke. New England Journal of Medicine. 2009;360:1718-1728
2. Toole JF, Malinow MR, Chambless LE, et al. Lowering homocysteine in patients with ischemic stroke to prevent recurrent stroke, myocardial infarction, and death: The vitamin intervention for stroke prevention (VISP) randomized controlled trial. JAMA. 2004;291:565-575
3. Matarín M, Brown WM, Scholz S, et al. A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. Lancet Neurol. 2007;6(5):414-20.
4. Kathiresan S, Voight BF, Purcell S et al. or the Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009;41(3):334-41.
5. Meschia JF, Kissela BM, Brott TG, et al. The Siblings With Ischemic Stroke Study (SWISS): a progress report. Clin Med Res. 2006;4(1):12-21
6. Ridker PM, Chasman DI, Zee RY, et al. for the Women's Genome Health Study Working Group Clin Chem. 2008;54(2):249-55. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women.
7. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. Ann Intern Med 2009;150:65-72.
8. Wu L, Shen Y, Liu, Wang X. et al. The 1425G/A SNP in PRKCH Is Associated With Ischemic Stroke and Cerebral Hemorrhage in a Chinese Population. Stroke. 2009 Jun 11.
9. Li Y , Abecasis GR. Mach 1.0: Rapid Haplotype Reconstruction and Missing Genotype Inference. Am J Hum Genet 2006;S79:2290
10. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-7
11. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006 Dec;2(12):e190.
12. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38(8):904-9.
13. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data.Genetics. 2000 Jun;155(2):945-59.

14. Fung HC, Scholz S, Matarin M, et al. Genome-wide genotyping in Parkinson's disease and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol. 2006;5(11):911-6.

15. Zeggini E, Ioannidis JP. Meta-analysis in genome-wide association studies. Pharmacogenomics. 2009 Feb;10(2):191-201.

16. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007;23(10):1294-6

17. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;15;21(11):1539-58.

18. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. Stat Med. 1988;7(8):889–894

19. CLEAR Trial Investigators. The combined approach to lysis utilizing eptifibatide and rt-PA in acute ischemic stroke: the CLEAR stroke trial. Stroke. 2008 ;39(12):3268-7

20. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 2003;19(2):185-93.

21. Tang Y, Xu H, Du X, et al. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study. J Cereb Blood Flow Metab 2006;26(8):1089-102.

22. Xu H, Tang Y, Liu DZ, et al. Gene expression in peripheral blood differs after cardioembolic compared with large-vessel atherosclerotic stroke: biomarkers for the etiology of ischemic stroke. J Cereb Blood Flow Metab 2008;28(7):1320-8.

Investigators

Michael A. Nalls, Ph.D.[#], Alessandro Biffi, M.D.[#], Mar Matarin, Ph.D., Christopher D. Anderson, M.D., Daniel I. Chasman, Ph.D., Steve Bevan, Ph.D., Chris C.A. Spencer, Ph.D., Andreas Gschwendtner, M.D., Michele M. Sale, Ph.D., Lijun Wu, Ph.D., Danish Saleheen, M.D., Ian J. Deary, Ph.D., Carl Langefeld, Ph.D., Leema Reddy Peddareddygari M.D, John W. Cole, M.D., M.Sc., Sampath K. Arapelli, M.Sc., Natalie Aucutt-Walter, M.D., Thomas G. Brott, M.D., Joseph P. Broderick, M.D., Devin L. Brown, M.D. M.Sc., Mark Brown, Ph.D., Robert D. Brown, Jr., M.D., Julie E. Buring, Sc.D., Lynelle Cortellini, M.Sc., John Danesh FRCP, D.Phil., Ph.D., Ranjan Deka, Ph.D., Gustavo A. Elias, Luigi Ferrucci, M.D., Ph.D., Philippe Frossard, Ph.D., Karen Furie, M.D., M.P.H., J Raphael. Gibbs, B.Sc., Steven M. Greenberg, M.D., Ph.D., Maryam Gul, M.D., John Hardy, Ph.D., Sarah E. Harris, Ph.D., Dena G. Hernandez, M.Sc., Fang-Chi Hsu, Ph.D., Subba Indugula, Ph.D., Caroline Jackson, Ph.D., Brett Kissela, M.D., Lisheng Liu, M.D., E Jeffrey. Metter, M.D., Braxton D. Mitchell, Ph.D., Lee Murphy, M.Sc., Josyf C. Mychaleckyj, Ph.D., Jeffrey R. O'Connell, Ph.D., Guillaume Paré, M.D., M.Sc., Alex N. Parker, Ph.D., Rohan K. Patel B.Sc., Deborah Poole, Asif Rasheed, M.D., Alexa Richie, M.P.H., Peter M. Rothwell, M.D., Ph.D., Lynda Rose, M.Sc., Owen A. Ross, Ph.D., Natalia S. Rost, M.D., Kristin Schwab B.A., Souvik Sen, M.D., M.Sc., Scott L. Silliman, M.D., Alexandra I. Soto-Ortolaza, B.Sc., John M. Starr, M.B.B.S.., O. Colin Stine, Ph.D., Boryana Stamova, Ph.D., Huichun Xu, M.D., Ph.D., H. Erich Wichmann, M.D, Ph.D., Lauren E. Young, B.Sc., William Young, M.D., Moazzam Zaidi, M.D., Hongye Zhang, M.D., Steven J. Kittner, M.D., M.P.H., Raji P. Grewal, M.D., Daniel Woo, M.D., M.Sc., Cathie Sudlow, Ph.D., Ayeesha K. Kamal, M.D., Xingyu Wang, Ph.D., Frank R. Sharp M.D., Bradford B. Worrall, M.D., M.Sc., Martin Dichgans, M.D., Hugh S. Markus, D.M., Paul M. Ridker, M.D., M.P.H., Paul I.W. de Bakker, Ph.D., Wellcome Trust Case-Control Consortium 2, Andrew B. Singleton, Ph.D.*, James F. Meschia, M.D.*, Jonathan Rosand, M.D., M.Sc.*

on behalf of the International Stroke Genetics Consortium

[#] These investigators should be considered equal first

* Address for correspondence

Affiliations are as follows, ordered alphabetically by study: **ESS** - Division of Clinical Neurosciences, University of Edinburgh, UK (C.S., C.J.); Institute of Genetics and Molecular Medicine, University of Edinburgh, UK (C.S., L.M.); Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, UK (I.D.,S.H.,J.S.,C.S.); Department of Geriatric Medicine, University of Edinburgh (J.S.), Department of Psychology, University of Edinburgh, UK (I.D., S.H.); Genetics Core Laboratory, Wellcome Trust Clinical Research Facility, Western General Hospital, University of Edinburgh, UK (L.M.); **GCNKSS** - Section on Statistical Genetics and Bioinformatics, Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest University Health Sciences, Winston-Salem, North Carolina, USA (M.B., C.L.); Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio, USA (R.D., S.I.); and Department of Neurology, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA (D.W.); **GENE EXPRESSION IN STROKE** - Department of Neurology, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA (J.P.B.); Department of Neurology, MIND Institute, University of California at Davis, Sacramento, CA, USA (F.R.S.,B.S.,H.X.); Department of Anesthesiology, University of California at San Francisco, San Francisco, CA, USA (W.Y.); **ISGS-SWISS** - Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, Bethesda, MD, USA (S.K.A., J.R.G., D.G.H.,M.M.,M.A.N.,A.B.S.); Department of Neurology, Mayo Clinic Jacksonville, Jacksonville, FL, USA (T.G.B,A.R.,O.A.R., J.F.M, A.I.S.); Department of Neurology, Mayo Clinic Rochester, Rochester, MN, USA (R.D.B.jr.); Longitudinal Studies Section, Clinical Research Branch, Intramural Research Program, National Institute on Aging, Baltimore, MD, USA (L.F.,E.J.M.); Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, University College London, London, UK (J.R.G.,J.H.,D.G.H.), Department of Neurology, University of Cincinnati College of Medicine, Cincinnati, Ohio, USA (B.K.), Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA (R.K.P.); .); **MCISS** - New Jersey Neuroscience Institute, JFK Medical Center, Edison, New Jersey, USA (R.P.G.,L.R.P.); UNC Stroke Program, Departments of Neurology and Cardiology, University of North Carolina, Chapel Hill, North Carolina, USA (N.A-W., S.S.); **MGH/MIGen** - Stroke Service, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA (C.D.A., A.B., L.C., K.L.F., S.M.G., J.R., N.S.R., K.S.); Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA (A.B., L.C., J.R., N.S.R.); Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA (A.B., L.C., J.R., N.S.R.); Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA (P.I.W.de B.), Stroke Program, Department of Neurology, University of Michigan Health System, Ann Harbor, MI, USA (D.L.B), Department of Neurology, University of Florida College of Medicine, Jacksonville, FL, USA (S.L.S.), Dept Neurology, University of Virginia, Charlottesville, VA, USA (B.B.W.); **RACE** - Center for Non-Communicable Diseases, Karachi, Pakistan (D.S., P.M.F., M.G., A.R., M.Z.); Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK (J.D., D.S.); Section of Neurology, Department of Medicine, Aga Khan University, Karachi, Pakistan (A.K.K.); **SHINING** - Beijing Hypertension League Institute, Bejing, China (L.L., H.Z.); Laboratory of Human Genetics, Beijing Hypertension League Institute, Bejing, China (X.W., L.W.); **UMD-SPYAS** - Department of Veterans Affairs Medical Center, Baltimore, MD, USA (J.W.C., S.J.K.); University of Maryland at Baltimore, Baltimore, MD, USA (J.W.C., S.J.K., B.D.M., J.R.O., O.C.S.); **VISP** - Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA (G.A.E., J.C.M., M.M.S., B.B.W.); Dept Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA (F.S.); Dept Public Health Sciences, University of Virginia, Charlottesville, VA, USA (J.C.M., B.B.W.); Dept Neurology, University of Virginia, Charlottesville, VA, USA (B.B.W.); Dept Medicine, University of Virginia, Charlottesville, VA, USA (M.M.S.); Dept Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA, USA (M.M.S.); **WGHS** - Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA (J.E.B., D.I.C., P.G., P.M.R., L.R.); Harvard Medical School, Cambridge, MA, USA (J.E.B. D.I.C., P.G., P.M.R.); Harvard School of Public Health, Cambridge, MA, USA (J.E.B., P.M.R.); Genetic and Molecular Epidemiology Laboratory, McMaster University, Hamilton, ON Canada (P.G.); Amgen, Cambridge, MA, USA (A.N.P., L.E.Y.). **WTCCC2** - Stroke phenotype group: Centre for Clinical Neuroscience, St. George's, University of London, London, United Kingdom (S.B., H.S.M.), Department of Neurology, Klinikum Großhadern, Ludwig-Maximilians-Universitat, Munich, Germany, (M.D., A.G.); Stroke Prevention Research Unit, Department of Clinical Neurology, University of Oxford, Oxford, UK (P.M.R., D.P.), Department of Statistics, University of Oxford, Oxford OX1 3TG, UK (C.A.A.S).

Wellcome Trust Case-Control Consortium 2 :

Management Committee: Elvira Bramon[1], Matthew Brown[2], Juan Casas[3], Aiden Corvin[4], Nicholas Craddock[5], Panos Deloukas[6], Janus Jankowski[7], Hugh Markus[8], Christopher G Mathew[9], Mark McCarthy[10], Colin Palmer11, Robert Plomin[12], Stephen Sawcer[13], Richard C Trembath9, Ananth Viswanathan[14], Nick Wood[15] , Leena Peltonen[6] (Deputy Chair) & Peter Donnelly[16] (Chair).

Data and Analysis Group: Chris C Spencer[16], Céline Bellinguez[17], Daniel Davison[16], Colin Freeman[17], Amy Strange[17], Peter Donnelly[16]

DNA, Genotyping, Data QC and Informatics Group: Cordelia Langford[6], Sarah E Hunt[6], Sarah Edkins[6], Rhian Gwilliam[6], Hannah Blackburn[6], Suzannah J. Bumpstead[6], Serge Dronov[6], Matthew Gillman[6], Emma Gray[6], Naomi Hammond[6], Alagurevathi Jayakumar[6], Owen T McCann[6], Jennifer Liddle[6], Marc L Perez[6], Simon Potter[6], Radhi Ravindrarajah[6], Michelle Ricketts[6], Matthew Waller[6], Paul Weston[6], Sara Widaa[6], Pamela Whittaker[6], Panos Deloukas[6], Leena Peltonen[6].

Publications Committee: Jenefer Blackwell[18], Matthew Brown[2], Aiden Corvin[4], Mark I McCarthy[10], Chris C Spencer[16] & Christopher G Mathew[9] (Chair).

UK Blood Services Controls: Antony P Attwood[19], Jonathan Stephens[20], Jennifer Sambrook[20] and Willem H Ouwehand[19].

1958 Birth Cohort Controls: Wendy L McArdle[21], Susan M Ring[22] and David P Strachan[23].

People of the British Isles Controls: Bruce Winney[24], Walter F Bodmer[24].

KORA-gen (Cooperative Health Research in the Region of Augsburg, Southern Germany): : Norman Klopp [25], H.-Erich Wichmann [25].

1. Dept Psychological Medicine, King's College London Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK
2. Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Nuffield Orthopaedic Centre, Oxford OX3 7LD, UK;
3. Dept Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK;
4. Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire;
5 Dept Psychological Medicine, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, UK;
6 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK;
7 Centre for Gastroenterology, Bart's and the London School of Medicine and Dentistry, London E1 2AT, UK;
8 Division of Cardiac and Vascular Sciences, Dept Clinical Neurosciences, St George's Hospital, London SW17 0RE, UK;
9 Dept Medical and Molecular Genetics, King's College London School of Medicine, Guy's Hospital, London SE1 9RT, UK;
10 Oxford Centre for Diabetes, Endocrinology and Metabolism (ICDEM), Churchill Hospital, Oxford OX3 7LJ, UK;
11 Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK;
12 Social, Genetic and Developmental Psychiatry Centre, King's College London Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK
13 Dept Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK;
14 Glaucoma Research Unit, Moorfields Eye Hospital NHS Foundation Trust, London EC1V 2PD,UK;
15 Dept Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK;
16 Dept Statistics, University of Oxford, Oxford OX1 3TG, UK ;
17 Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7LJ, UK;
18 Genetics and Infection Laboratory, Cambridge Institute for Medical Research, Addenbrooke's Hospital, Cambridge CB2 0XY, UK ;
19 Dept Haematology, University of Cambridge and National Health Service Blood and Transplant, Long Road, Cambridge CB2 2PT, UK, and Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK;
20 Dept Haematology, University of Cambridge and National Health Service Blood and Transplant, Long Road, Cambridge CB2 2PT, UK;
21 ALSPAC DNA Bank, Dept Social Medicine, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK;
22 ALSPAC Laboratory, Dept Social Medicine, Clifton, Bristol BS8 2BN, UK;
23 Division of Community Health Sciences, St George's Hospital, London SW17 0RE, UK;
24 Cancer and Immunogenetics Laboratory, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DS, UK.
25 KORA-Gen, Helmholtz Zentrum München,German Research Center for Environmental Health, Ingolstädter Landstrasse 1, D-85764 München/Neuherberg

FUNDING

ESS: The Edinburgh Stroke Study was supported in part by the Wellcome Trust (clinician scientist award to Dr. Sudlow), the Binks Trust (research fellowship to Dr. Jackson)and REMiND (Edinburgh Medical Neurology Research Fund). The Lothian Birth Cohort Study was supported in part by the Research into Aging, Help the Aged (Sidney De Haan Award and The Disconnected Mind Major Gift Campaign), Medical Research Council, United Kingdom Biotechnology and Biological Sciences Research Council. **GCNKSS and Cincinnati Intracebral Hemorrhage GWAS** : These studies were supported by NIH grants NS36695 (Genetic and Environmental Risk Factors for Hemorrhagic Stroke) and NS30678 (Hemorrhagic and Ischemic Stroke among Blacks and Whites) and by the Greater Cincinnati Foundation Grant (Cincinnati Control Cohort). **GENE EXPRESSION IN STROKE:** This work was supported by NIH grant NS056302 (F Sharp), the American Heart Association/Bugher Foundation Centers for Stroke Prevention Research 0775014N (F Sharp) and by PO21040N635110 (Joseph Broderick). **ISGS/NIA/UCL**: This work was supported in part by the American Heart Association (AHA) and the Myron and Jane Hanley Award in Stroke research and by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services; project number Z01 AG000954-06. This work used samples and clinical data from the NINDS Human Genetics Resource Center DNA and Cell Line Repository (http://ccr.coriell.org/ninds), human subjects protocol numbers 2003-081 and 2004-147. The inclusion of BLSA samples was supported in part by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services; project number Z01 AG000015-50, human subjects protocol number 2003-078. Dr. Hardy is supported by an MRC Returning scientist award. **MCISS:** This study was funded by intramural funds from the New Jersey Neuroscience Institute and the Neurogenetics Foundation. **MGH-MIGen and ICH GWAS:** These studies were funded by the American Heart Association/Bugher Foundation Centers for Stroke Prevention Research 0775010N, NIH-NINDS grant 1R01 NS059727 and by the Deane Institute for Integrative Research in Atrial Fibrillation and Stroke. The MIGen study was funded by the U.S. National Institutes of Health and National Heart, Lung, and Blood Institute'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. Drs. Biffi, Rost and Anderson were supported in part by the American Heart Association/Bugher Foundation Centers for Stroke Prevention Research. **RACE:** Genotyping in this project was supported through funds available to the investigators at the University of Cambridge. The RACE is supported by grants from Higher Education Commission, Pakistan and from the National Institute of Neurological Disorders And Stroke (Award Number R21NS064908). **SHINING:** SHINING was funded through the Beijing Hypertension League Institute, in part through the National Infrastructure Program of Chinese Genetic Resource (2005DKA21300) and an unrestricted educational grant from F. Hoffmann-La Roche. **UMD-SPYAS:** Drs. Kittner and Cole were supported in part by the Office of Research and Development, Medical Research Service, the Baltimore Research Enhancement Award Program in Stroke and the Baltimore Geriatrics Research, Education, and Clinical Center of the Department of Veterans Affairs; a Cooperative Agreement with the Division of Adult and Community Health, Centers for Disease Control and Prevention; the National Institute of Neurological Disorders and Stroke (NINDS) and the NIH Office of Research on Women's Health (ORWH) (Grant R01 NS45012); the National Human Genome Research Institute (NHGRI) (Grant U01 HG004436); the National Institute on Aging (NIA) Pepper Center (Grant P60 12583); the American Heart Association (Beginning Grant-in-Aid 0665352U); and the University of Maryland General Clinical Research Center (Grant M01 RR 165001), National Center for Research Resources, NIH. This study used samples from the NINDS Human Genetics Resource Center DNA and Cell Line Repository (http://ccr.coriell.org/ninds), as well as clinical data.
VISP: This work was funded by R25 HL088724 (GAS - SRIP program grant), K08 NS045802 and R01 NS 42733 (B.B.W - ISGS). The VISP study was funded by NS34447 and the WFU GCRC M01 RR07122. **WGHS:** The WGHS is supported by HL 043851 and HL69757 from the National Heart, Lung, and Blood Institute and CA 047988 from the National Cancer Institute, the Donald W. Reynolds Foundation and the Fondation Leducq, with collaborative scientific support and funding for genotyping provided by Amgen. **WTCCC2.** The Wellcome Trust Case Control Consortium 2 was funded by the Wellcome Trust. We acknowledge use of the British 1958 Birth Cohort DNA collection funded by the Medical Research council grant G0000934 and the Wellcome Trust grant 068545/Z/02. WTCCC2:Germany was supported by

the Vascular Dementia Research Foundation (to M.D.) and the German Ministry of Education and Research (PGE-S04T13 and PGES04T10 to H.E.W.).

ACKNOWLEDGMENTS

We thank the patients and families who contributed to these studies.

ESS
For the Edinburgh Stroke Study – we thank Profs Martin Dennis, Peter Sandercock, Charles Warlow, Drs Simon Hart and Sarah Keir (patient recruitment and characterization), Prof Joanna Wardlaw, Drs Andrew Farrall and Gillian Potter (neuroimaging interpretation), Aidan Hutchison and Mike McDowall (programming and database support). For the Lothian Birth Cohort – we thank Dr Alan Gow (coordination, subject assessment, database support), Caroline Brett, Janie Corley, Ross Henderson, Catherine Murray and Alison Pattie (subject assessment and database support), Lorna Houlihan and Michelle Luciano (genetics). Paula Davies (administrative support)

MGH/MIGen and ICH GWAS
We thank the Myocardial Infarction Genetics Consortium (MIGen) study for the use of their genotype data as control data in our study. The MIGen study was funded by the U.S. National Institutes of Health and National Heart, Lung, and Blood Institute's STAMPEED genomics research program (R01 HL087676) and a grant from the National Center for Research Resources. We thank all research staff from the Stroke Service at Massachusetts General Hospital and other participating Institutions for patient enrollment and clinical/neuroimaging data management. We also thank Mark J. Daly, Benjamin F. Voight, Benjamin M. Neale and members of the Medical and Population Genetics group at the Broad Institute for valuable discussion, and members of the Broad Biological Samples and Genetic Analysis Platforms for sample management and genotyping.

MCISS
We thank Drs .Martin S. Gizzi, Spozhmy Panezai and Nicholas Papamitsakis for patient recruitment and characterization. We would also like to thank Drs Harry Collins, Machia Uthappa, and Mark Mayer for recruiting control patients, the Hematology department at JFK Medical Center for assistance in processing of blood samples. We are grateful to Ana Virginia Dutra PhD for database support when the study was initiated.

SWISS
The SWISS investigators include **Executive Committee:** James F. Meschia, chair; Thomas G. Brott, Robert D. Brown, Jr, Brett Kissela, John Hardy, Stephen S. Rich, Andy Singleton, Bradford Worrall; **Statistical Center:** University of Virginia, Charlottesville, VA. Stephen S. Rich, Director of Center for Public Health Genomics; **DNA Repository:** Coriell Institute for Medical Research, Camden, New Jersey, Roderick Corriveau, PhD, Director; and **Data Management:** Mayo Clinic, Rochester, Minnesota, Barry Bisbee, Data Management Specialist. SWISS Centers and Investigators Listed by Proband Enrollment as of July 2009: 1. *Neurological Associates, Inc, Richmond, Virginia* (probands enrolled, 72): Principal Investigator (PI): Alan Schulman, MD; coordinator: Gwendolyn H. Darby; *Mayo Clinic, Rochester, Minnesota* (probands enrolled, 67): PI: Robert D. Brown, Jr., MD; coordinator: Colleen S. Albers; *Mayo Clinic, Jacksonville, Florida* (probands enrolled, 64): PI: Thomas G. Brott, MD; coordinators: Alexa Richie; Dale Gamble; Sothear Luke; *University of Virginia Health System, Charlottesville, Virginia* (probands enrolled, 59): PI: Bradford Worrall, MD, MSc; coordinator: Colleen Harmon; *University of Cincinnati, Cincinnati, Ohio* (probands enrolled, 51): PI: Brett Kissela, MD; coordinator: Kathleen Alwell; *Shands HealthCare, affiliated with the University of Florida, Jacksonville, Florida* (probands enrolled, 50): PI: Scott Silliman, MD; coordinator: Rhonda Calhoun; *Mercy Ruan Neurology Clinic, Des Moines, Iowa* (probands enrolled, 45): PI: Michael Jacoby, MD; coordinator: Cheryl Inman; *Washington University School of Medicine, St. Louis, Missouri* (probands enrolled, 34): PI: Jin-Moo Lee, MD, PhD; coordinators: Julie Rickmann; Jill Newgent; *University of Maryland, Baltimore, Maryland* (probands enrolled, 27): PI: John Cole, MD; coordinator: Mark Dobbins; Mary J. Sparks; *University of Pennsylvania MedicalCenter, Philadelphia, Pennsylvania* (probands enrolled, 22): PI: Scott Kasner, MD; coordinator: Mary Liz DeSanto; *Main Line Health Stroke Program, Bryn Mawr, Pennsylvania* (probands enrolled,22): PI: Gary H. Friday, MD; coordinator: Joan Brown; *Kaleida Stroke Care Center, Millard Fillmore Gates Circle Hospital, Buffalo, New York* (probands enrolled, 22): PI: R. Ferguson, MD; coordinator: Kathey

Parkes; *Centre Hospitalier Affilie Universitaire de Quebec,Quebec City, Province of Quebec, Canada* (probands enrolled, 20): PI: Ariane Mackey MD; coordinator: Annette Hache; *Mayo Clinic, Scottsdale, Arizona* (probands enrolled, 20): PI: Bart Demaerschalk, MD; coordinator: Erica L. Boyd; *Metro Health Medical Center, Cleveland, Ohio* (probands enrolled, 20): PI: Joseph Hanna, MD; coordinator: Dana Cook; *Ohio State University, Columbus, Ohio* (probands enrolled, 19): PI: Andrew Slivka, MD; coordinator: Peggy Notestine; *Luther Midelfort Hospital, Eau Claire,*
Wisconsin (probands enrolled, 18): PI: Felix Chukwudelunzu, MD; coordinator: Karen Snobl; *OSF Stroke Center, Peoria, Illinois* (probands enrolled, 17): PI: Arun Talkad, MD; coordinator: Mary Buttice;
Rochester General Hospital, Rochester, New York (probands enrolled, 17): PI: W. Scott Burgin, MD; coordinator: Cheryl Weber; *University of Iowa Hospital, Iowa City, Iowa* (probands enrolled, 15): PI: Patricia Davis, MD; coordinator: Jeri Sieren; *University of South Alabama, Mobile, Alabama* (probands enrolled, 15): PI: J. Ivan Lopez, MD; coordinator: Mel Parnell; *Hôpital Charles LeMoyne, Greenfield Park, Province of Quebec, Canada* (probands enrolled, 14): PI: Leo Berger, MD; coordinators: Martine Mainville; Louise-Marie Béliveau; *Marshfield Clinic, Marshfield, Wisconsin* (probands enrolled, 14): PI: Percy Karanjia, MD; coordinator: Kathy Mancl, Rob Bohl; *University of California, Davis School of Medicine, Sacramento, California* (probands enrolled, 13): PI: PieroVerro, MD; coordinator: John Bautista; *Mercy Hospital, Sioux City, Iowa* (probands enrolled, 13): PI: Jennifer Pary, MD; coordinator: Deb Motz; Darla Bauer; *University of Texas Southwestern Medical Center at Dallas, Dallas, Texas* (probands enrolled, 12): PI: Mark Johnson, MD, coordinator: April Blair; *Cleveland Clinic Florida, Weston, Florida* (probands enrolled, 11): PI: Virgilio Salanga, MD; coordinator: Kilaun Thompson; *Wake Forest University School of Medicine,*
Winston-Salem, North Carolina (probands enrolled, 11): PI: David Lefkowitz, MD; coordinator: Emily Smith; *University of Utah Stroke Center, Salt Lake City, Utah* (probands enrolled, 10). PI: Elaine J. Skalabrin, MD; coordinator: Jill Austin; *Helen Hayes Hospital, West Haverstraw, New York* (probands enrolled, 10): PI: Jason Greenberg, MD; coordinator: Laura Tenteromano; *Stroke Prevention and Atherosclerosis Research Centre ,London, Ontario, Canada* (probands enrolled, 10): PI: David Spence, MD; coordinator: Rose Freitas; *Indiana University School of Medicine, Indianapolis, Indiana* (probands enrolled, 9): PI: Linda Williams, MD; coordinator: Flossy Lincoln; *Scripps Clinic, La Jolla, California* (probands enrolled, 9): PI: Mary Kalafut, MD; coordinators: Crystal Sanchez; Gail Moore; *University of Kentucky, Lexington, Kentucky* (probands enrolled, 7): PI: L. Creed Pettigrew, MD; coordinator: Deborah Taylor; *University of Wisconsin, Madison, Wisconsin* (probands enrolled, 7): PI: Robert J. Dempsey, MD; coordinator: Pam Winne; *East Bay Region Associates in Neurology, Berkeley, California* (probands enrolled, 7): PI: Brian Richardson, MD; coordinator: Lauren McCormick; *Thomas Jefferson University Hospital, Philadelphia, Pennsylvania* (probands enrolled, 7): PI: Rodney Bell, MD; coordinator: Prema Krishna; *Massachusetts General Hospital, Boston, Massachusetts* (probands enrolled 6)*:* PI: Karen Furie, MD; coordinator: Ashley Birch; *University of California Los Angeles Stroke Center, Los Angeles, California* (probands enrolled, 6): PI: Jeffery Saver, MD; coordinator: Lauren Bell; *Florida Neurovascular Institute, Tampa, Florida* (probands enrolled, 5): PI: Erfan Albakri, MD; coordinator: Valerie Walters; *Seton Medical, Austin Texas* (probands enrolled, 5): PI: Darryl Camp, MD; coordinator: Lisa Houy; *Morton Plant Hospital - Neuroscience Institute, Clearwater, Florida* (probands enrolled, 5): PI: Ajay Arora, MD; coordinators: Jo Simpson; Teresa Jones; Victoria Lindeman; *The Stroke Center at Hartford Hospital, Hartford, Connecticut* (probands enrolled, 4)*:* PI: Isaac Silverman, MD; coordinators: Martha Ahlquist; Jennifer Blum; *Southtowns Neurology* (probands enrolled, 4)*:* PI: Peterkin Lee Kwen, MD; coordinators: Sarfaraz Semy; *Vanderbilt Stroke Center, Nashville, Tennessee* (probands enrolled, 3): PI: Anne E. O'Duffy, MD; coordinator: Diane Brown; *University of North Carolina Chapel Hill, Chapel Hill, NC* (probands enrolled, 3): PI: Souvik Sen, MD; coordinator: Roxanne Poole; *Elkhart Clinic, Elkhart, IN* (probands enrolled, 3): PI: Thomas Vidic, MD; coordinator: Randall Gibson; *Lancaster General Hospital, Lancaster, Pennsylvania* (probands enrolled, 2): PI: Venkatachalam Mangeshkumar, MD; coordinator: Holly Snyder; *Tri-State Neurology, Huntington, West Virginia* (probands enrolled, 2): PI: Carl McComas, MD; coordinator: Cheryl Kane; *Albert Einstein Medical Center, Philadelphia, Pennsylvania* (probands enrolled, 2): PI: Johnathon Dissin, MD; coordinator: Rohini Bhole; *Froedtert Memorial Lutheran, Milwaukee, Wisconsin* (probands enrolled, 1)*:* PI: Michael Torbey, MD; coordinator: Erin Brandenburg; *Sparks Neurology Center, Fort Smith, Arkansas* (probands enrolled, 1): PI: Margaret Tremwel, MD; coordinators: Cheryl Hyde, LPN, CCRC; Mary Rambin; and *Legacy Health System, Tualatin, Oregon:* PI: Paul Ash, MD; coordinator: Taryn Lust.

RACE
We would like to acknowledge the following RACE investigators: Khalid Mahmood, Karimullah Maki, Muhammad Masroor Ahmed, Niaz Sheikh, Abdul Rauf Memon Naveeduddin, Furrukh Shahab Khan, Asfaq Sharif, Shoukat Ali, Jamal Ara, Tasnim Ahsan, SM Munir; Kaukab Bashir, Riaz-ul-Nabi Tayyab, Shoaib Randhawa, Waseem Iqbal, Asad Alamgir, Muhammad Nasrullah, Zafarullah Choudhary, Mushtaq Haroon, Tahira Murtaza Cheema, Abdul Hameed Anjum, Ijaz Hussain Malik, Ghulam Mohiuddin, Muhammad Arshad, Navaid Iqbal, Syed Vaqas Alam, Mohammad Nabeel, Lajpat Bansari, Fahad Hashmat, Shahan Waheed, Uzma Saleem, Mustafa Hussain, Muhammad Hussain, Sheeraz Anjum, Ather Rauf, Muhammad Khan, Mujtaba Hashir, Waleed Kayani, Salman Daood,Nabi Shah,Maria Samuel, Sadaf Raza,Naeem Khan,Saeed Ahmed,Mohammad Waqar Khan,Mohammad Tahir Khan,Mohammad Imran,Maqsood Ahmed,Rana Zulfiqar,Sarfaraz Sher Ali, Tauqeer Ahmed, Syed Waqas Ahmed and Azfar Hussain