DOI:
10.1002/ana.24840

Document Version
Peer reviewed version

Genetic Variation at 16q24.2 is associated with small vessel stroke.

Running head: 16q24.2 and small vessel stroke

Matthew Traylor[1], Rainer Malik[2], Mike A. Nalls[3], Ioana Cotlarciuc[4], Farid Radmanesh[5,6,7], Gudmar Thorleifsson[8], Ken B. Hanscombe[1], Carl Langefeld[9], Danish Saleheen[10], Natalia S. Rost[6], Idil Yet[11], Tim D. Spector[11], Jordana T. Bell[11], Eilis Hannon[12], Jonathan Mill[12,13], Ganesh Chauhan[14,15], Stephanie Debette[14,15], Joshua C. Bis[16], WT Longstreth Jr.[17], M. Arfan Ikram[18], Lenore J. Launer[19], Sudha Seshadri[20.21], METASTROKE, UK Young Lacunar DNA Study, NINDS Stroke Genetics Network, Neurology Working Group of the CHARGE Consortium, Monica Anne Hamilton-Bruce[22], Jordi Jimenez-Conde[23], John W. Cole[24], Reinhold Schmidt[25], Agnieszka Słowik[26], Robin Lemmens[27,28,29], Arne Lindgren[30,31], Olle Melander[32], Raji P. Grewal[33], Ralph L. Sacco[34], Tatjana Rundek[34], Kathryn Rexrode[35], Donna K. Arnett[36], Julie A. Johnson[37,38], Oscar R. Benavente[39], Sylvia Wassertheil-Smoller[40], Jin-Moo Lee[41], Sara L. Pulit[42], Quenna Wong[43], Stephen S. Rich[44], Paul I.W. de Bakker[45,46], Patrick F. McArdle[47], Daniel Woo[48], Christopher D. Anderson[5,6,7,49], Huichun Xu[50], Laura Heitsch[51], Myriam Fornage[52], Christina Jern[53], Kari Stefansson[8,54], Unnur Thorsteinsdottir[8,54], Solveig Gretarsdottir[9], Cathryn M. Lewis[1,13], Pankaj Sharma[4], Cathie L.M. Sudlow[55], Peter M. Rothwell[56], Giorgio B. Boncoraglio[57], Vincent Thijs[21,22,58], Chris Levi[59], James F. Meschia[60], Jonathan Rosand[5,6,7,48], Steven J. Kittner[24], Braxton D. Mitchell[61,62], Martin Dichgans[2,63], Bradford B. Worrall[64], Hugh S. Markus[65]; on behalf of the International Stroke Genetics Consortium

[1] Department of Medical and Molecular Genetics, King's College London, London, UK

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/ana.24840

[2] Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians University, Munich, Germany

[3] Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA

[4] Institute of Cardiovascular Research Royal Holloway University of London (ICR2UL), UK

[5] Division of Neurocritical Care and Emergency Neurology, Massachusetts General Hospital, Boston, MA, USA

[6] J. Philip Kistler Stroke Research Center, Department of Neurology, Massachusetts General Hospital, Boston, MA, USA

[7] Center for Human Genetic Research, Massachusetts General Hospital, Boston. MA, USA

[8] deCODE genetics/AMGEN, Reykjavik, Iceland

[9] Center for Public Health Genomics and Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA

[10] Department of Genetics, Perelman School of Medicine, University of Pennsylvania, PA, USA

[11] Department of Twin Research & Genetic Epidemiology, King's College London, London, UK

[12] University of Exeter Medical School, University of Exeter, Exeter, UK

[13] Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[14] Inserm Research Center for Epidemiology and Biostatistics (U897) – Team Neuroepidemiology, Bordeaux, France

[15] University of Bordeaux, Bordeaux, France

[16] Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA

[17] Departments of Neurology and Epidemiology, University of Washington, Seattle, WA, USA

[18] Department of Neurology, Epidemiology and Radiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands

[19] Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Bethesda, MD, USA

[20] Boston University School of Medicine, Boston, MA, USA

[21] Framingham Heart Study, Framingham, MA, USA

[22] Department of Neurology, Royal Adelaide Hospital, Adelaide, South Australia, Australia

[23] Neurovascular Research Group (NEUVAS). Neurology Department. Institut Hospital del Mar d'Investigació Mèdica. Barcelona. Spain

[24] Department of Neurology, University of Maryland School of Medicine and Baltimore VAMC, MD, USA

[25] Department of Neurology, Clinical Division of Neurogeriatrics, Medical University Graz, Austria

[26] Department of Neurology, Jagiellonian University, Krakow, Poland

[27] KU Leuven - University of Leuven, Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), Leuven, Belgium

[28] VIB, Vesalius Research Center, Laboratory of Neurobiology, Department of Neurology, Leuven, Belgium

[29] University Hospitals Leuven, Department of Neurology, Leuven, Belgium

[30] Department of Clinical Sciences Lund, Neurology, Lund University, Lund, Sweden

[31] Department of Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund, Sweden

[32] Department of Clinical Sciences, Lund University, Malmö, Sweden

[33] Neuroscience Institute, Saint Francis Medical Center, School of Health and Medical Sciences, Seton Hall University, South Orange, New Jersey, USA

[34] Department of Neurology, Miller School of Medicine, University of Miami, USA

[35] Harvard Medical School, Boston, USA Center for Faculty Development and Diversity, Brigham and Women's Hospital, Boston, MA, USA

[36] College of Public Health, University of Kentucky, Lexington, KY, USA

[37] Department of Pharmacotherapy and Translational Research and Center for Pharmacogenomics, University of Florida, College of Pharmacy, Gainesville, FL, USA

[38] Division of Cardiovascular Medicine, College of Medicine, University of Florida, Gainesville, FL, USA

[39] Department of Neurology, University of British Columbia, Vancouver, British Columbia, Canada

[40] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, NY, USA

[41] Stroke Center, Department of Neurology, Washington University School of Medicine, WA, USA

[42] Department of Medical Genetics, University Medical Center Utrecht, Utrecht, Netherlands

[43] Department of Biostatistics, University of Washington, WA, USA

[44] Center for Public Health Genomics, University of Virginia School of Medicine, VA, USA

[45] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

[46] Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, the Netherlands

[47] Department of Medicine, University of Maryland School of Medicine, MD, USA

[48] University of Cincinnati College of Medicine, Cincinnati, OH, USA

[49] Program in Medical and Population Genetics, Broad Institute, Boston, MA, USA

[50] Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, MD, USA

[51] Division of Emergency Medicine, Washington University School of Medicine, St Louis, MO, USA

[52] The University of Texas Health Science Center at Houston, Houston, TX, USA

[53] Institute of Biomedicine, the Sahlgrenska Academy at University of Gothenburg, Sweden

[54] Faculty of Medicine, University of Iceland, Reykjavik, Iceland

[55] Center for Clinical Brain Sciences & Institute of Genetics and Molecular Medicine, University of Edinburgh, UK

[56] Nuffield Department of Clinical Neurosciences, University of Oxford, UK

[57] Department of Cerebrovascular Diseases, Fondazione IRCCS Istituto Neurologico "Carlo Besta", Milano, Italy

[58] Department of Neurology, Austin Health and Florey Institute of Neuroscience and Mental Health, Heidelberg, Australia

[59] John Hunter Hospital, Hunter Medical Research Institute and University of Newcastle, Newcastle, NSW, Australia

[60] Department of Neurology, Mayo Clinic Jacksonville, Jacksonville, FL, USA

[61] Department of Medicine, University of Maryland School of Medicine, MD, USA

[62] Geriatrics Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, MD, USA

[63] Munich Cluster of Systems Neurology, SyNergy, Munich, Germany

[64] Departments of Neurology and Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, USA

[65] Stroke Research Group, Division of Clinical Neurosciences, University of Cambridge, UK

Corresponding Author:

Matthew Traylor (matthew.traylor@kcl.ac.uk)

Department of Medical and Molecular Genetics,

King's College London

8th Floor, Tower Wing,

Guys Hospital, Great Maze Pond

London SE1 9RT

Title Characters: 76; Running Head Characters: 31

Word Count: 4,090; Abstract: 247; Tables: 2; Figures: 5; References: 50

Running head: 16q24.2 and small vessel stroke

Abstract

Objective: Genome-wide association studies (GWAS) have been successful at identifying associations with stroke and stroke subtypes, but have not yet identified any associations solely with small vessel stroke (SVS). SVS comprises a quarter of all ischaemic stroke and is a major manifestation of cerebral small vessel disease, the primary cause of vascular cognitive impairment. Studies across neurological traits have shown younger onset cases have an increased genetic burden. We leveraged this increased genetic burden by performing an age-at-onset informed GWAS meta-analysis, including a large younger onset SVS population, to identify novel associations with stroke.

Methods: We used a three-stage age-at-onset informed GWAS to identify novel genetic variants associated with stroke. On identifying a novel locus associated with SVS, we assessed its influence on other small vessel disease phenotypes, as well as on mRNA expression of nearby genes, and on DNA methylation of nearby CpG sites in whole blood and in the fetal brain.

Results: We identified an association with SVS in 4,203 cases and 50,728 controls on chromosome 16q24.2 (OR(95% CI)=1.16(1.10-1.22); $p=3.2x10^{-9}$). The lead SNP (rs12445022) was also associated with cerebral white matter hyperintensities (OR(95% CI)=1.10(1.05-1.16); $p=5.3x10^{-5}$; N=3,670), but not intracerebral haemorrhage (OR(95% CI)=0.97(0.84-1.12); p=0.71; 1,545 cases, 1,481 controls). rs12445022 is associated with mRNA expression of *ZCCHC14* in arterial tissues ($p=9.4x10^{-7}$), and DNA methylation at probe cg16596957 in whole blood ($p=5.3x10^{-6}$).

Interpretation: 16q24.2 is associated with SVS. Associations of the locus with expression of *ZCCHC14* and DNA methylation suggest the locus acts through changes to regulatory elements.

Introduction

Genome-wide association studies (GWAS) enable identification of common genetic variants that influence disease risk and have proved successful in elucidating pathophysiological mechanisms underlying diseases with a genetic influence.[1] A number of GWAS associations have recently been identified with ischaemic stroke, almost all of which have been associated with specific stroke subtypes.[2-4] A number of genetic associations have been reported with cardioembolic (CE) and large artery stroke (LAS), but in contrast there have been no robust associations solely with small vessel stroke (SVS). This is despite epidemiological data that suggest genetic factors are particularity important for SVS. For example, there are a number of monogenic stroke disorders associated with SVS,[5] and family history studies have shown a strong association between SVS and a family history of stroke.[6] Similarly, related traits, including white matter hyperintensities, have been shown to have high heritability.[7]

SVS itself comprises a quarter of all ischaemic stroke and is one of the clinically overt manifestations of cerebral small vessel disease (SVD), the major cause of vascular cognitive impairment. Other radiological features of SVD include white matter hyperintensities (WMH) best seen on T2-weighted MRI, cerebral microbleeds - seen on gradient echo MRI, and intracerebral hemorrhages (ICH).[8] Despite its importance, the pathogenesis of SVD remains poorly understood and this limits the development of proven treatments for established disease.

One consistent finding across adult-onset neurological complex diseases including Parkinson's disease,[9] Alzheimer's disease,[10] and stroke,[11] is that younger onset cases have a stronger genetic burden from common disease-associated SNPs. Leveraging this increased burden, by focussing on younger onset cases in analysis of genetic data, can lead

to detection of novel trait-associated variants.[11] This may be particularly relevant for SVS, as epidemiological studies have shown stronger associations with SVS and a family history of stroke in younger stroke cases.[6]

Here, we perform an age-at-onset informed GWAS meta-analysis in stroke, including a large population of younger onset (age<70) small vessel stroke (SVS) cases. We perform analysis for all ischaemic stroke (IS) and its three subtypes: cardioembolic (CE), large artery stroke (LAS) and SVS. Using this approach we identify a novel association with SVS, seek further validation of the locus in other SVD phenotypes, and assess the influence of SNPs at the locus on mRNA expression of nearby genes and DNA methylation at nearby CpG sites.

Methods

Study design

We employed a three-stage design for the association analysis (Figure 1). In brief, in stage I we performed association analysis of stroke phenotypes in 10,210 cases and 12,285 controls of European ancestry from Europe, United States, and Australia; most of which contributed to the METASTROKE ischaemic stoke GWAS meta-analysis – and all of which have been described previously (Table 1). [2, 12, 13] In all cases, diagnosis of stroke was based on clinical evaluation with radiological confirmation. Subtyping of stroke cases was based on the TOAST criteria; in this analysis we considered the CE, LAS and SVS subtypes.[14] Of note, our SVS analysis included a large sample (1,012 cases, 970 controls) of younger onset (age<70) MRI-confirmed lacunar strokes, meaning that although we investigated all subtypes, we had most power to identify associations with SVS.

In stage II, we took 3 SNPs from the top 25 loci from each phenotype forward for a first in silico replication in the NINDS Stroke Genetics Network (SiGN), [2] which consisted of 7,743 cases and 17,790 controls. We meta-analysed stages I and II together and identified 3 loci with $p<5x10^{-7}$. Finally in stage III, we determine whether these 3 SNPs were associated with the phenotype in which they were identified (CE or SVS) by *in silico* replication in a large Icelandic population (deCODE; 520 SVS cases, 1,100 CE cases, 50,728 controls; stage III).

Genotyping and Imputation

Genotyping, quality control and imputation of all studies has been described previously.[2, 3, 13] All studies were genotyped on commercially available arrays from Illumina or Affymetrix and imputed to 1000 Genomes phase 1 reference panels using IMPUTE or MACH.[15] Imputation quality score was assessed by calculating the ratio of the observed to the expected binomial variance of the allele dosage.

Association analysis

Association analysis was performed using a covariate-informed approach,[11, 16] which we, and others, have implemented previously.[11, 17] Briefly, the approach uses case/control status and a covariable – in this case age-at-onset – to estimate each individual's stroke liability, which can be interpreted as their underlying propensity to stroke, on a normally-distributed scale. In this analysis cases with an earlier age-at-onset take more extreme positive values than late onset cases as, due to the lower prevalence of stroke at younger ages, they are assumed to have higher stroke liability. Conversely, controls who are older and stroke-free at age-at-observation take more extreme negative value than younger controls as they have been stroke-free for a longer time and are therefore assumed to have a lower stroke liability.

In this analysis, the approach was implemented in our software, CIAO (provided at https://sites.google.com/site/mtraylor263/software/covariate-informed-gwas-analysis). Specifically, the approach taken is to model phenotype data using a continuous unobserved normally distributed quantitative trait, called the disease liability ($\varphi = \sum_{j=1}^{J} c_j\left(t_j - \bar{t_j}\right) + m + \varepsilon$), where $\varepsilon = \gamma g + N(0,1)$ and g denotes the genetic effects. Then an individual is a case (*z=1*) if and only if $\varphi \geq 0$ and is a control (*z=0*) otherwise. c_j is a parameter estimating the effect of a given covariate *j* on the liability scale. m denotes the disease prevalence *p* at the covariate mean $\bar{t_j}$ under a normal cumulative distribution function ($\Phi(-m) = p$). This model is used to approximate the effect of a disease covariate – in this case age-at-onset - on the liability scale, based on estimates of risk of ischaemic stroke by age from epidemiological data, thereby estimating c_j. For this analysis, the gender-specific risk of ischaemic stroke by age from an index age of 55 was obtained from population-based estimates (1.8%, 5.4%, and 12.1% before 65, 75, and 85, respectively in women; 2.4%, 7.3%, and 12.6% before 65, 75, and 85, respectively in men).[18] We assumed that 20% of ischaemic stroke cases had each of the cardioembolic, small vessel or large vessel stroke subtypes, approximating proportions observed in population-based studies..[19] We developed two models for our analysis; one based on the risk rates for all ischaemic stroke, and secondly for the three stroke subtypes. We used these models to calculate posterior mean liabilities after conditioning on age-at-onset for the four stroke phenotypes separately ($E(\varepsilon|z,t) = \frac{\int_{-c(t-\bar{t})-m}^{\infty} \varepsilon \frac{1}{\sqrt{2\pi}} e^{\left(\frac{-\varepsilon^2}{2}\right)} d\varepsilon}{\int_{-c(t-\bar{t})-m}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\left(\frac{-\varepsilon^2}{2}\right)} d\varepsilon}, if\ z = 1$). Controls were modelled in the same way, but were assumed to take the posterior mean from the lower (unaffected) portion of the distribution in the liability threshold model ($E(\varepsilon|z,t) = \frac{\int_{\infty}^{-c(t-\bar{t})-m} \varepsilon \frac{1}{\sqrt{2\pi}} e^{\left(\frac{-\varepsilon^2}{2}\right)} d\varepsilon}{\int_{\infty}^{-c(t-\bar{t})-m} \frac{1}{\sqrt{2\pi}} e^{\left(\frac{-\varepsilon^2}{2}\right)} d\varepsilon}, if\ z = 0$). Where age data was missing, individuals were assigned the median age value (<1% of cases). Regression was then performed on posterior liabilities ($E(\varepsilon|z,t)$) by multiplying the number of samples by the

squared correlation between the expected genotype dosage and posterior mean liabilities for each of the discovery cohorts in the four ischaemic stroke phenotypes (all IS, CE, LAS, SVS). Ancestry-informative principal components were included where appropriate, using the EIGENSTRAT procedure.[20] Any residual inflation was accounted for by adjusting results by the genomic inflation factor, λ.[21] In all analyses, SNPs with imputation quality score<0.7 or minor allele frequency<0.01 were excluded and meta-analysis was performed using Stouffer's method in METAL.[22]

Further analysis of a novel locus associated with small vessel stroke

For a novel variant associated with SVS, we performed further analysis to elucidate the association for different groups based on age-at-onset. Firstly, for datasets in stage I and II, we divided the cases into quartiles based on age-at-onset and estimated the association of the SNP with each quartile using logistic regression with all controls, meta-analysing using a fixed-effects inverse variance weighted approach (data not available in BRAINS, MGH-GASROS, ISGS/SWISS). Secondly, we interrogated associations at the locus in non-European ancestry populations, comprising 657 small vessel African-American stroke cases and 3,251 matched controls from the NINDS Stroke Genetics Network and African or African-Caribbean ancestry individuals from the South London Ethnicity and Stroke Study (SLESS), [2, 23] and 314 SVS cases and 5,193 controls of Pakistani ancestry from the RACE study.[3] We used logistic regression to evaluate the association within each group, and evaluated the overall transethnic association by meta-analysing using Stouffer's method.

In addition, we explored association of the SNP with other SVD phenotypes. We evaluated association of the SNP with 1) white matter hyperintensity volumes (WMHV) measured on T2-weighted MRI in 3,670 ischaemic stroke patients of European ancestry,[24] 2) in MRI-defined small subcortical brain infarcts (SSBI) brain infarcts in 17,197 trans-ethnic individuals

(85.7% European; 8.8% African-American; 3.5% Hispanic; 1.0% Chinese; 1.0% Malay) from community studies recruited within the neuro-CHARGE consortium (mean age 68.90 ± 10.31; 1,986 with infarcts). SSBI were defined as MRI-defined brain infarcts of 3-15 mm or 3-20 mm in size, located in the basal ganglia, the white matter, or the brainstem. Association analysis was performed overall, and for the subset of cases with extensive WMH burden – defined as the top age-specific quartile of WMH volume on a quantitative scale or above the age-specific median by 5-year age-categories for studies using semi-quantitative measurements of WMH burden; N=549; 3) ICH in 1,545 European ancestry cases and 1,481 controls, described previously,[25] and stratified according to lobar or nonlobar location.

Evaluation of regulatory chromatin states, mRNA expression and DNA methylation

To investigate a novel locus, we used existing resources and performed some further analyses to characterize its regulatory potential. We interrogated chromatin states and regulatory motifs from ENCODE and Epigenomics Roadmap using Haploreg v4.1.[26] We also evaluated whether the associated SNP influences gene expression using GTEx portal.[27] Upon identifying an association between the SNP and expression of a nearby gene, we evaluated the evidence that the association signal for SVS and gene expression derives from the same causal variant using a Bayesian colocalisation test.[28] Using the R coloc package (http://cran.r-project.org/web/packages/coloc), we compared 5 models for SNPs with 50Kb of our lead SNP using the approach (H_0: No association with either trait; H_1: Association with SVS, not with expression; H_2: Association with expression, not with SVS; H_3: Association with SVS and expression, two independent SNPs; H_4: Association with SVS and expression, one shared SNP).

Next, we assessed whether the lead SNP (rs12445022), or 3 SNPs in linkage disequilibrium (LD) (rs4843625, rs12920915, rs12444224), influence DNA methylation levels in whole

blood. We evaluated genetic associations of whole blood DNA methylation levels at selected CpG-sites profiled on the Illumina Infinium HumanMethylation450 BeadChip array in a group of 660 monozygotic (MZ) female twins (mean age 59, age range 18 to 79). These individuals were research volunteers from the TwinsUK cohort in the United Kingdom.[29] All were of European ancestry. For each CpG-site of interest we calculated the normalised methylation means for the 330 MZ twin pairs as a phenotype in the genetic analysis, and took into account covariates including smoking, BMI, age, methylation plate, and blood cell count estimates. TwinsUK imputed genotypes were obtained for the 1000 genomes reference set,[30] where we excluded SNPs with Hardy–Weinberg $p<1x10^{-4}$, Minor allele frequency (MAF) <5% and those with IMPUTE info value <0.8. We tested for association with our SNP, or SNPs in close LD ($r^2>0.6$) with DNA methylation at CpG-sites. We used $p<4x10^{-5}$, equivalent to a false discovery rate (FDR) <5%,[31] to identify significant cis-mQTL associations.

Finally, we explored genetic associations at 16q24.2 (defined as within 50Kb of rs12445022) with DNA methylation profiles in 166 human fetal brain samples (92 male, 74 female) ranging from 56–166 days post-conception initially using publically available data – which holds results for mQTL associations reaching the study-wide significance threshold (http://epigenetics.essex.ac.uk/mQTL/). Methods for this study have been published in detail elsewhere.[32] Briefly, DNA methylation levels were profiled on the Illumina Infinium HumanMethylation450 BeadChip array and SNP genotypes were obtained from the Illumina HumanOmniExpress BeadChip and imputed to 1000 Genomes phase 3 using SHAPEIT and Minimac3 via the Michigan Imputation Server.[15, 33] SNP-methylation probe pairs were tested using the R package MatrixEQTL,[34] including covariates to control for age, sex and ancestry-informative principal components. Upon identifying a significant association at 16q24.2, we performed additional analyses (not publicly available: we gained access to the data) to test whether any of our 4 SNPs (rs12445022, rs4843625, rs12920915, rs12444224) were

associated with methylation at the identified probe. We again used $p<4x10^{-5}$, equivalent to a false discovery rate (FDR) <5%,[31] to identify significant cis-mQTL associations.

Results

Association analysis

In phase I association analysis we confirmed previous associations between *HDAC9* and LAS (rs2107595, $p=3.0x10^{-8}$) and between *PITX2* and CE (rs192172299, $p=2.0x10^{-9}$).[3, 4] Previous associations between *ZFHX3* and CE and between *MMP12* and LAS did not reach genome-wide significance in this analysis (rs879324, $p=5.0x10^{-7}$ and rs586701, p=0.0014; respectively).[11] A SNP in a region close to *HABP2* previously associated with young onset ischaemic stroke was also significant, albeit not genome-wide, in this analysis (rs11196288; $p=2.4x10^{-4}$).[35] Genomic inflation λ and the equivalent values scaled to 1000 cases and 1000 controls (λ_{1000}),[36] were well controlled across all analyses (IS, λ (λ_{1000}) =1.05 (1.00); CE, λ (λ_{1000}) =1.02 (1.00); LAS, λ (λ_{1000}) =1.02 (1.00); SVS. λ (λ_{1000}) =1.01 (1.00)).

We took 25 independent loci forward (3 SNPs in LD from each locus selected on p-value) from each analysis (IS, CE, LAS, SVS) for *in silico* replication in the NINDS Stroke Genetics Network study (stage II). Information on these SNPs is provided in Supplementary Tables 1-4. Following this analysis , excluding previously reported associations, three loci showed significance at $p<5x10^{-7}$ (two with SVS, one with CE) and one was genome-wide significant (rs12445022, $p=4.4x10^{-8}$, associated with SVS). We followed up all three loci in a second *in silico* replication (stage III) in a large Icelandic population (deCODE). A single SNP, rs12445022, showed evidence of replication (p=0.011). When performing a meta-analysis across all populations, rs12445022 was associated with SVS at genome-wide significance ($p=3.2x10^{-9}$; Figure 2). The SNP was either genotyped or well imputed (info>0.9) in all

cohorts and lies in an intergenic region between Junctophilin 3 (*JPH3*) and Zinc Finger, CCHC Domain Containing 14 (*ZCCHC14*). To confirm the association with rs12445022, we repeated the analysis using logistic regression; the approach taken in a conventional GWAS. The association was validated using this method, and associations were consistent across populations (OR(95% CI)=1.16(1.10-1.22); $p=1.3x10^{-8}$; heterogeneity p=0.56; Figure 3).

Further analysis of a 16q24.2 novel locus associated with small vessel stroke

We evaluated association of the lead SNP in different quantiles of age at stroke onset, using all controls in each analysis. The strongest associations were observed in younger onset cases, suggesting the influence of the SNP might be greatest in these individuals (Figure 4). However, this was not demonstrated statistically (p>0.05).

We performed further analysis to assess whether the SVS-associated SNP influenced other manifestations of cerebral SVD. The SNP (rs12445022) was also associated with increased T2-WMHV (OR(95% CI)=1.10(1.05-1.16); $p=5.3x10^{-5}$; Figures 2,5), and showed little heterogeneity across study groups (heterogeneity p=0.58). Conversely, the SNP was not associated with ICH – neither overall, nor in subgroups divided by lobar/non-lobar location. For SSBI, the direction of effect was the same as for SVS, but the effect was weaker and non-significant (OR(95% CI)=1.05(0.97-1.14); p=0.28). For the subgroup with WMH, the effect was stronger – and similar to that observed for SVS, but was again non-significant (OR(95% CI)=1.15(0.99-1.33); p=0.076).

We next evaluated the identified locus in non-European ancestry populations. The SNP had a similar frequency to Europeans in South Asians from RACE (MAF=37%), but was rarer in African ancestry populations, consisting of African-Americans from NINDS Stroke Genetics

Network and United Kingdom individuals of African or African-Caribbean ethnicity from SLESS (MAF=14%). Associations with the SNP were in the same direction as in European ancestry populations (Figure 4), but did not reach statistical significance in either ancestry, reflecting the much smaller sample sizes. However, when combining data from all populations, evidence for association at the SNP ($p=1.4x10^{-9}$) was stronger than in European ancestry populations alone, which might suggest a common association across populations. Indeed, there was not evidence of a significant difference in the strength of association between the European and non-European ancestry individuals (p=0.64).

Regulatory chromatin states, mRNA expression and DNA methylation related to 16q24.2

We used existing databases to assess the functional consequences of SNPs in the 16q24.2 region. Firstly, we used the Haploreg v4.1 database to interrogate chromatin states and regulatory motifs from ENCODE and NIH Roadmap Epigenomics Mapping Consortium. [26, 37, 38] The database showed that our lead SNP influences chromatin states in multiple tissues. The SNP is classified as a genic promoter in 9 tissues, an enhancer in 13 tissues and overlaps DNAse1 hypersensitivity sites in 21 tissues.

Secondly, we used publicly available databases to evaluate the evidence that the lead SNP influences expression of nearby genes using the GTEx portal.[27] The implicated A allele of our lead SNP (rs12445022) was associated with decreased expression of *ZCCHC14* in tibial arterial tissue ($p=9.4x10^{-7}$; Figure 2). We used a Bayesian colocalisation technique to assess whether the same variant drives the both the SVS association signal and mRNA expression of *ZCCHC14*.[28] There was overwhelming evidence in support of H_4 (Posterior probability = 99.7%), strongly indicating that a single variant - most likely to be rs12445022 - influences both SVS and expression of *ZCCHC14*.

Finally, we performed analyses to assess whether the lead SNP, or the 3 SNPs in LD, influence DNA methylation at CpG probes in whole blood. We found evidence that the lead SNP, and 3 SNPs in close LD (r^2>0.6), influence DNA methylation at 4 nearby CpG sites (cg16596957, cg10312981, cg03020503, cg00555085; all $p<4.0x10^{-5}$, Table 2). The implicated A allele of rs12445022 was associated with decreased methylation at the cg16596957 probe (beta(SE)=-0.38(0.082); $p=5.3x10^{-6}$). The SNPs explained between 5-8% of the methylation variance at the given CpG sites. The same 16q24.2 region by CpG probe (cg16596957) association was also recently reported in another study in whole blood.[40] In addition, we looked for an association between SNPs at the 16q24.2 locus and DNA methylation levels in fetal brains, initially using publicly available data (http://epigenetics.essex.ac.uk/mQTL/). There was a strong association with SNPs in distant LD with our lead SNP (rs8047314 ~ cg08031982; $p=7.1x10^{-14}$; r^2=0.16 with rs12445022). We then performed additional analyses (not publicly available: we gained access to the data) to test if our lead SNP, or the SNPs in close LD, were associated with methylation at cg08031982. We could identify no associations that reached our significance threshold ($p<4.0x10^{-5}$). However, there was a near-significant association of rs12920915 and rs4843625 with methylation at the cg08031982 probe (both $p=7.8x10^{-5}$). Our lead SNP, rs12445022 was not associated ($p=9.9x10^{-4}$).

Discussion

Genome-wide association studies in SVS have largely been disappointing. Some studies have suggested that an association with all IS at the highly pleiotropic 12q24.12 is driven by an association with SVS, [2] but no genome-wide significant associations specifically with SVS have yet been identified. Using an age-of–onset informed analysis approach we identified a novel locus at 16q24.2 associated with SVS. The SNP was also associated using a standard logistic regression approach, but was less significant – a difference of almost an order of

magnitude ($p=3.2x10^{-9}$ compared to $p=1.3x10^{-8}$). In addition, the association was stronger with younger onset SVS, suggesting a greater influence in these individuals. We tested whether the 16q24.2 association extends to other cerebral SVD related phenotypes. We showed that the same locus also influences WMH, and may have a similar effect on MRI-defined subcortical brain infarcts from prospective studies, although the association did not reach significance in our analysis. However, the locus does not appear to influence risk of ICH. A SNP in the same 16q24.2 region (rs4081947), in partial LD with our SNP ($r^2=0.28$), was also recently reported to be associated with migraine in a large GWAS meta-analysis. [41] These data provide strong supportive evidence that this 16q24.2 locus harbours variants that influence diseases of the cerebral vasculature.

Identifying the mechanisms by which GWAS associations influence disease risk presents additional challenges. In this case, the underlying mechanism and the specific genes implicated remains uncertain. Interrogation of mRNA expression data points to the lead SNP influencing expression of the nearest gene, *ZCCHC14*. This gene is ubiquitously expressed, but is highly expressed in arterial tissues and in the brain. However, its function is not well characterized. Zinc fingers of the CCHC-type contain an 18 digit residue found in the nucleocapsid of retroviruses, and therefore may be important in viral response. Other plausible candidate genes reside nearby. The locus lies around 1Mb away from genes encoding forkhead box proteins including *FOXC2, FOXL1,* and *FOXF1.* These proteins, particularly the closely related *FOXC1* – a paralogue of *FOXC2*, have been implicated in Mendelian forms of SVS.[42] We found no evidence linking our SNP to expression of these genes. However, the function of these proteins changes dramatically between early development and in adult tissues,[43] which might explain the absence of an association. This, coupled with the fact that *FOXF2* variants have also recently been implicated in ischaemic stroke,[44] make forkhead box proteins exciting targets for follow-up experiments.

Assessing DNA methylation, the process by which methyl groups are added to DNA thereby modifying its function, offers another potential method for mechanistic insight. This epigenetic process influences gene expression and regulation in humans, and may be particularly relevant for diseases such as stroke where gene-environment interactions are likely play an important role in pathogenesis.[45] Substantial inter-individual variation exists with respect to age and tissue type.[46] However, an important emerging mechanism influencing methylation is local sequence content.[47] Notably, recent studies have shown that GWAS findings from stroke-relevant traits such as blood pressure are likely to act by influencing DNA methylation.[48] This may be particularly relevant for SVS, in which environmental and other vascular risk factors such as hypertension are important and have been shown to interact with disease risk.[49] We evaluated whether our associated SNP (rs12445022), or SNPs in close LD, influence methylation of nearby CpG sites. We found evidence from whole blood that the same genetic variation influences DNA methylation. SNPs in distant LD also influenced DNA methylation at a different probe (cg08031982) in the fetal brain. Further evidence comes from published studies in lung, breast, and kidney tissues,[31] as well as in utero,[50] all of which have shown that the genetic variation at the same 16q24.2 region influences methylation at the cg08031982 probe. Interestingly, the CpG sites influenced by the locus appear to differ by tissue, with different probes affected in whole blood compared to fetal brain. This might imply tissue-specific functional consequences of the locus and therefore highlights the importance of performing follow-up experiments in appropriate tissues. Based on the evidence presented here, we can only speculate on how genetic variation at the locus leads to increased risk of SVS. One hypothesis is that expression of *ZCCHC14*, or other proteins, is mediated through altered methylation of the probes identified. This might occur, in part, in response to environmental stimuli. Evaluating these hypotheses in a relevant tissue type will be an important future analysis to identify the causal mechanisms leading to SVS.

This study has limitations. Our results suggested that the association may be present in other ethnicities, but we had an insufficient number of cases to establish common risk conclusively. Follow-up studies are therefore required in other ethnic groups. In addition, downstream functional experiments will be required to determine the consequences of the identified association. The mRNA expression and methylation analyses presented herein were constrained by available tissue types. Validation of the findings in more disease relevant tissue types such as cerebral small vessels therefore represent important follow-up analyses, although obtaining such tissue in a state to allow mRNA studies is very challenging. We performed mRNA expression and methylation analyses using either the lead SNP (rs12445022) or 3 LD SNPs. The results should be interpreted with the limitation that we cannot be certain that any of these SNPs is the causal variant. Radiological confirmation of SVS in this study was performed using either CT or MRI. Evidence shows that MRI is considerably more reliable at identifying SVS. Replication of the association in an MRI-confirmed population may therefore provide a more accurate estimate of the effect of the locus on SVS risk. Similarly, interrogation of causative classification system (CCS) definitions of SVS may provide further insights. [51] Another method of interrogating the combined influence of age and genotype is by testing for an interaction. In this analysis, we were unable to do this as age was not available in some sets of controls (e.g WTCCC2).

In this large genome-wide meta-analysis using an age-at-onset informed approach, we have identified the first genome-wide significant locus that is associated solely with SVS. Our findings, which point to subtle changes in gene expression and DNA methylation influencing disease risk, show that strategies that account for different liability across disease related covariates such as age can identify novel associations with disease.

Acknowledgements

Matthew Traylor is funded by the NIHR Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London. Hugh Markus is supported by an NIHR Senior Investigator award and his work is supported by NIHR Comprehensive Biomedical Research Unit funding awarded to Cambridge University Hospitals Trust. Cathryn Lewis receives salary support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. Collection of the UK Young Lacunar Stroke DNA Study (DNA Lacunar) was primarily supported by the Wellcome Trust (WT072952) with additional support from the Stroke Association (TSA 2010/01). Genotyping of the DNA Lacunar samples was supported by a Stroke Association Grant (TSA 2013/01). Robin Lemmens is a senior clinical investigator of FWO Flanders. Martin Dichgans received funding from the DFG (CRC 1123, B3) and a EU Horizon 2020 grant (agreement No 666881 SVDs@target). The TwinsUK study was funded in part by the European Research Council (ERC 250157), and from the TwinsUK resource, which receives support from the Wellcome Trust and the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR. The SiGN study was funded by a cooperative agreement grant from the US National Institute of Neurological Disorders and Stroke, National Institutes of Health (U01 NS069208)

Author Contributions

M.T, R.M, C.M.L, B.B.W, H.S.M conceived and designed the study. M.T, B.B.W, H.S.M drafted the manuscript. M.T, K.B.H drew the figures. M.T, R.M, M.A.N, I.C, F.R, P.S, D.S, M.A.H-B, C.L.M.S, P.M.R, G.B, V.T, R.L, C.L, J.F.M, J.R, M.D, B.B.W, H.S.M contributed acquisition and analysis of METASTROKE datasets. H.X, L.H, M.F, C.J, J.F.M, B.D.M,

S.J.K, M.D, B.B.W J.J-C, J.W.C, R.S, A.S, R.L, A.L, O.M, R.P.G, R.L.S, T.R, K.R, D.K.A, J.A.J, O.R.B, S.W-S, J-M.L, M.T, S.S.R, P.D.B, S.L.P, Q.W, P.F.M, D.W, C.D.A, J.R contributed acquisition and analysis of the NINDS-SIGN data. J.R, D.W, C.L, C.D.A contributed acquisition and analysis of ICH data. G.C, S.D, L.J.L, S.S, J.C.B, W.T.L Jr, contributed acquisition and analysis of neuro-CHARGE data. I.Y, T.D.S, J.T.B, E.H., J.M contributed acquisition and analysis of DNA methylation data. M.T, H.S.M, N.S.R contributed acquisition and analysis of WMH data. G.T, K.S, U.T, S.G contributed acquisition and analysis of deCODE data.

Conflicts of Interest

Authors whose affiliations are listed as deCODE/Amgen are employees of deCODE/ Amgen. The remaining authors declare no conflicts of interest.

Table 1 – Ischaemic Stroke Study participants

Population	IS	CE	LAS	SVS	Controls	% cases with MRI	Age of cases (mean(s.d))
Stage I Populations							
ASGC	1,162	240	421	310	1,244	43.0%	72.9 (13.2)
WTCCC2-Germany	1,174	330	346	106	797	83.0%	66.7 (12.9)
WTCCC2-UK	2,374	474	498	460	5,175	37.2%	72.2 (12.5)
Milano	366	64	73	25	407	86.7%	57.4 (15.6)
DNA-lacunar / GENESIS	1,287	80	64	1,012	970	100.0%	59.6 (12.0)
LSS	455	157	70	55	455	89.0%	67.7 (14.5)
ISGS / SWISS	1,014	235	217	187	1,370	83.0%	66.5 (13.6)
BRAINS	361	29	120	97	444	30.8%	74.4 (14.2)
MGH-GASROS	294	106	68	23	376	60.0%	66.7 (14.5)
VISP	1,723	-	-	-	1,047	47.0%	68.0 (10.7)
Total (discovery)	10,210	1,715	1,877	2,275	12,285		
Stage II populations							
NINDS Stroke Genetics Network	7,743	2,001	1,130	1,408	17,970	62.0%	66.3 (14.8)
Stage III populations							
deCODE	-	1,100	-	520	20,473	NA	72.7 (11.6)
TOTAL	17,953	4,816	3,007	4,203	50,728		

IS, all ischaemic stroke; CE, cardioembolic stroke; LAS, large artery stroke; SVS, small vessel stroke; ASGC, Australian Stroke Genetics Collaborative; WTCCC2, Wellcome Trust Case Control Consortium 2; LSS, Leuven Stroke Study; BRAINS, Bio-repository of DNA in stroke; MGH-GASROS, The MGH Genes Affecting Stroke Risk and Outcome Study; VISP, The Vitamin Intervention for Stroke Prevention Trial; NA, information not available.

Table 2 – Significant associations between rs12445022 and LD SNPs (r^2>0.6) with cis-methylation probes in whole blood

SNP Variant	SNP BP	CpG Probe	Probe BP	RA	beta (SE)	R^2	P-value
rs12445022	87,575,332	cg16596957	87,575,151	A	-0.38 (0.082)	0.058	$5.3x10^{-6}$
rs4843625	87,576,996	cg16596957	87,575,151	C	-0.33 (0.075)	0.053	$1.3x10^{-5}$
rs4843625	87,576,996	cg10312981	87,577,304	C	0.39 (0.074)	0.077	$1.9x10^{-7}$
rs4843625	87,576,996	cg03020503	87,577,656	C	0.35 (0.075)	0.059	$5.0x10^{-6}$
rs4843625	87,576,996	cg00555085	87,616,248	C	0.34 (0.075)	0.057	$6.6x10^{-6}$
rs12920915	87,577,521	cg16596957	87,575,151	T	-0.38 (0.075)	0.069	$7.3x10^{-7}$
rs12920915	87,577,521	cg10312981	87,577,304	T	0.38 (0.075)	0.068	$1.0x10^{-6}$
rs12920915	87,577,521	cg03020503	87,577,656	T	0.34 (0.076)	0.055	$1.1x10^{-5}$
rs12920915	87,577,521	cg00555085	87,616,248	T	0.33 (0.076)	0.051	$2.2x10^{-5}$
rs12444224	87,580,855	cg16596957	87,575,151	T	-0.38 (0.075)	0.068	$8.0x10^{-7}$
rs12444224	87,580,855	cg10312981	87,577,304	T	0.38 (0.075)	0.069	$8.7x10^{-7}$
rs12444224	87,580,855	cg03020503	87,577,656	T	0.35 (0.076)	0.054	$1.1x10^{-5}$
rs12444224	87,580,855	cg00555085	87,616,248	T	0.32 (0.076)	0.050	$2.6x10^{-5}$

BP, base position; RA, reference allele; R^2, proportion of methylation variance explained by respective genotype.

Figure Legends

Figure 1 – Flow chart of analyses performed

Figure 2 – Associations at 16q24.2 with A) small vessel stroke, B) cerebral white matter hyperintensities, C) mRNA expression of ZCCHC14; and D) Gene Locations and associations of the locus with DNA methylation

SVS, small vessel stroke; WMH, white matter hyperintensities; ZCCHC14, Zinc Finger, CCHC Domain Containing 14; JPH3, Junctophilin 3; meQTL, methylation quantitative trait locus.

Figure 3 – Forest Plot of Associations with rs12445022 under a logistic regression model

Figure 4 – Association of rs12445022 with small vessel stroke by quartiles of age-at-stroke onset in Europeans

Figure 5 – Associations with rs12445022 for stroke and cerebral small vessel disease phenotypes

References

1. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet. 2012 Jan 13;90(1):7-24.
2. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurol. 2015 Dec 18.
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 Nov;11(11):951-62.
4. Bellenguez C, Bevan S, Gschwendtner A, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. Nat Genet. 2012 Mar;44(3):328-33.
5. Tan RY, Markus HS. Monogenic causes of stroke: now and the future. J Neurol. 2015 Dec;262(12):2601-16.
6. Jerrard-Dunne P, Cloud G, Hassan A, Markus HS. Evaluating the genetic component of ischemic stroke subtypes: a family history study. Stroke. 2003 Jun;34(6):1364-9.
7. Turner ST, Jack CR, Fornage M, Mosley TH, Boerwinkle E, de Andrade M. Heritability of leukoaraiosis in hypertensive sibships. Hypertension. 2004 Feb;43(2):483-7.
8. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 2013 Aug;12(8):822-38.
9. Escott-Price V, Nalls MA, Morris HR, et al. Polygenic risk of Parkinson disease is correlated with disease age at onset. Ann Neurol. 2015 Apr;77(4):582-91.
10. Sleegers K, Bettens K, De Roeck A, et al. A 22-single nucleotide polymorphism Alzheimer risk score correlates with family history, onset age, and cerebrospinal fluid Abeta. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2015 Jun 15.
11. Traylor M, Makela KM, Kilarski LL, et al. A novel MMP12 locus is associated with large artery atherosclerotic stroke using a genome-wide age-at-onset informed approach. PLoS Genet. 2014 Jul;10(7):e1004469.
12. Malik R, Traylor M, Pulit SL, et al. Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. Neurology. 2016 Mar 2.
13. Traylor M, Bevan S, Baron JC, Hassan A, Lewis CM, Markus HS. Genetic Architecture of Lacunar Stroke. Stroke. 2015 Sep;46(9):2407-12.
14. 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 Jan;24(1):35-41.
15. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012 Aug;44(8):955-9.
16. Zaitlen N, Lindstrom S, Pasaniuc B, et al. Informed conditioning on clinical covariates increases power in case-control association studies. PLoS Genet. 2012;8(11):e1003032.
17. Williams AL, Jacobs SB, Moreno-Macias H, et al. Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. Nature. 2014 Feb 6;506(7486):97-101.
18. Seshadri S, Wolf PA. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. Lancet Neurol. 2007 Dec;6(12):1106-14.
19. Lovett JK, Coull AJ, Rothwell PM. Early risk of recurrence by subtype of ischemic stroke in population-based incidence studies. Neurology. 2004 Feb 24;62(4):569-73.
20. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug;38(8):904-9.
21. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999 Dec;55(4):997-1004.

22. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010 Sep 1;26(17):2190-1.
23. Markus HS, Khan U, Birns J, et al. Differences in stroke subtypes between black and white patients with stroke: the South London Ethnicity and Stroke Study. Circulation. 2007 Nov 6;116(19):2157-64.
24. Traylor M, Zhang C, Adib-Samii P, et al. Genome-wide Meta-analysis of Cerebral White Matter Hyperintensities in Patients with Stroke Neurology (in press).
25. Woo D, Falcone GJ, Devan WJ, et al. Meta-Analysis of Genome-Wide Association Studies Identifies 1q22 as a Susceptibility Locus for Intracerebral Hemorrhage. Am J Hum Genet. 2014 Mar 18;94(4):511-21.
26. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012 Jan;40(Database issue):D930-4.
27. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015 May 8;348(6235):648-60.
28. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014 May;10(5):e1004383.
29. Moayyeri A, Hammond CJ, Valdes AM, Spector TD. Cohort Profile: TwinsUK and healthy ageing twin study. Int J Epidemiol. 2013 Feb;42(1):76-85.
30. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012 Nov 1;491(7422):56-65.
31. Shi J, Marconett CN, Duan J, et al. Characterizing the genetic basis of methylome diversity in histologically normal human lung tissue. Nature communications. 2014;5:3365.
32. Hannon E, Spiers H, Viana J, et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat Neurosci. 2016 Jan;19(1):48-54.
33. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nature methods. 2013 Jan;10(1):5-6.
34. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics. 2012 May 15;28(10):1353-8.
35. Cheng YC, Stanne TM, Giese AK, et al. Genome-Wide Association Analysis of Young-Onset Stroke Identifies a Locus on Chromosome 10q25 Near HABP2. Stroke. 2016 Feb;47(2):307-16.
36. de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet. 2008 Oct 15;17(R2):R122-8.
37. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012 Sep 6;489(7414):57-74.
38. Kundaje A, Meuleman W, Ernst J, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015 Feb 19;518(7539):317-30.
39. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet. 2013 Oct;45(10):1238-43.
40. Liu Y, Li X, Aryee MJ, et al. GeMes, clusters of DNA methylation under genetic control, can inform genetic and epigenetic analysis of disease. Am J Hum Genet. 2014 Apr 3;94(4):485-95.
41. Gormley P, Anttila V, Winsvold BS, et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. bioRxiv.doi: http://dx.doi.org/10.1101/030288.
42. French CR, Seshadri S, Destefano AL, et al. Mutation of FOXC1 and PITX2 induces cerebral small-vessel disease. J Clin Invest. 2014 Nov;124(11):4877-81.
43. Lehmann OJ, Sowden JC, Carlsson P, Jordan T, Bhattacharya SS. Fox's in development and disease. Trends Genet. 2003 Jun;19(6):339-44.
44. Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. Lancet Neurol. 2016 Apr 7.
45. Bell JT, Pai AA, Pickrell JK, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 2011;12(1):R10.

46. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
47. Banovich NE, Lan X, McVicker G, et al. Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. PLoS Genet. 2014 Sep;10(9):e1004663.
48. Kato N, Loh M, Takeuchi F, et al. Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. Nat Genet. 2015 Nov;47(11):1282-93.
49. Adib-Samii P, Devan W, Traylor M, et al. Genetic architecture of white matter hyperintensities differs in hypertensive and nonhypertensive ischemic stroke. Stroke. 2015 Feb;46(2):348-53.
50. Teh AL, Pan H, Chen L, et al. The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. Genome Res. 2014 Jul;24(7):1064-74.
51. Ay H, Benner T, Arsava EM, et al. A computerized algorithm for etiologic classification of ischemic stroke: the Causative Classification of Stroke System. Stroke. 2007 Nov;38(11):2979-84.

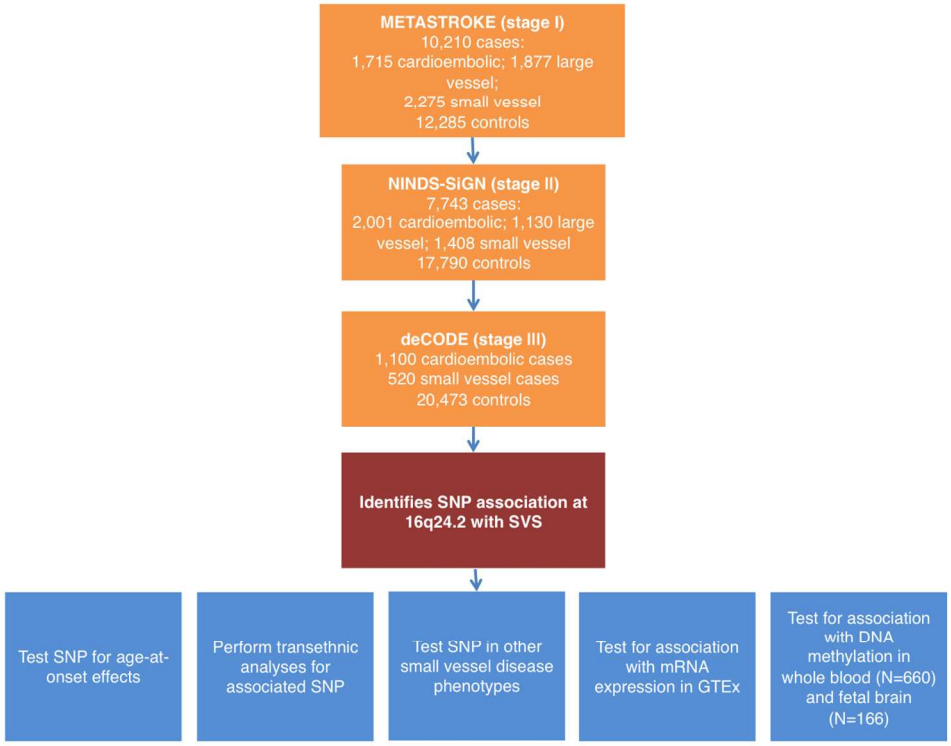

Figure 1 – Flow chart of analyses performed

508x381mm (72 x 72 DPI)

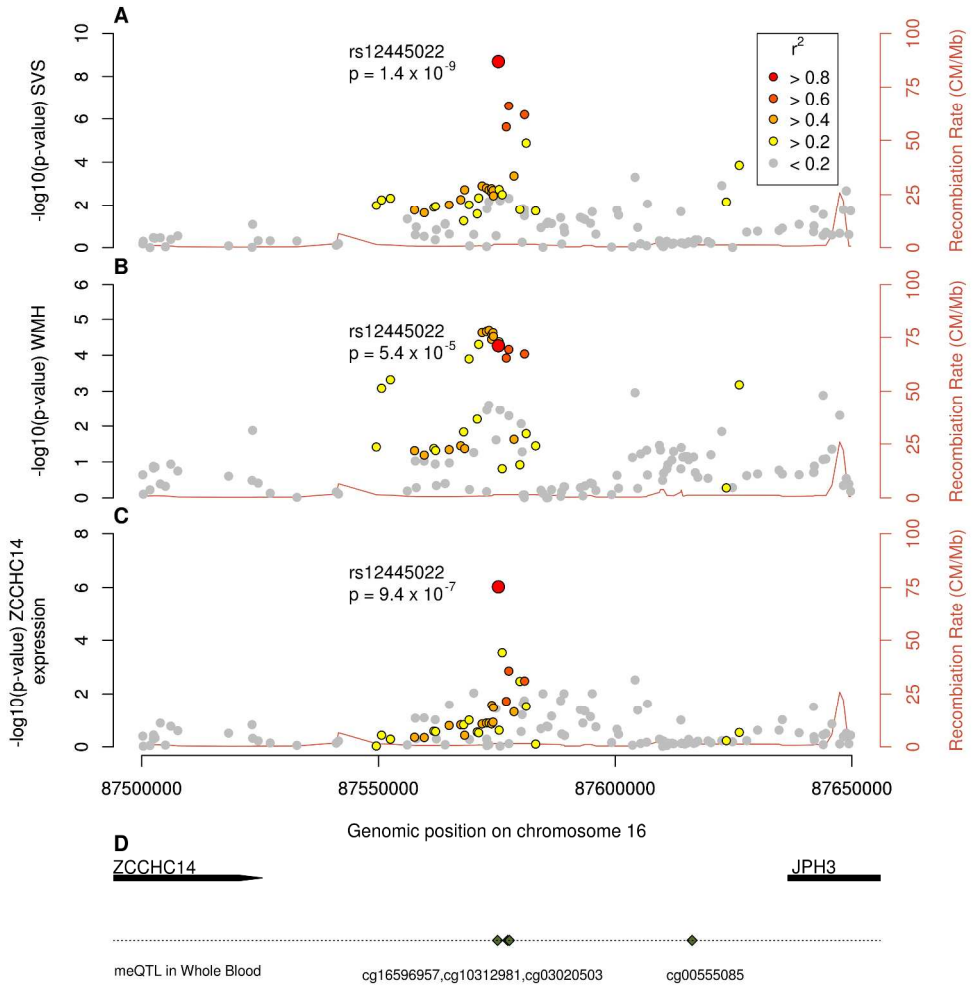

Figure 2 – Associations at 16q24.2 with A) small vessel stroke, B) cerebral white matter hyperintensities, C) mRNA expression of ZCCHC14; and D) Gene Locations and associations of the locus with DNA methylation SVS, small vessel stroke; WMH, white matter hyperintensities; ZCCHC14, Zinc Finger, CCHC Domain Containing 14; JPH3, Junctophilin 3; meQTL, methylation quantitative trait locus.

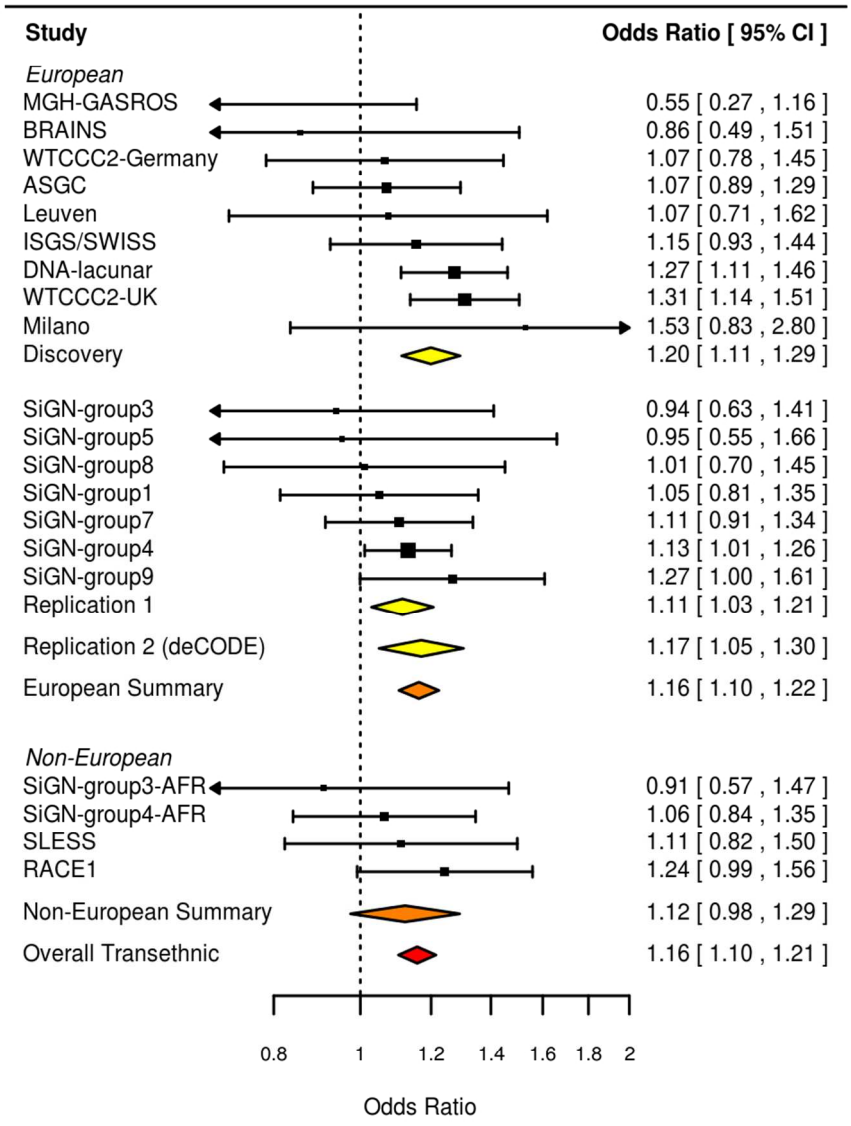

Figure 3 – Forest Plot of Associations with rs12445022 under a logistic regression model

86x113mm (300 x 300 DPI)

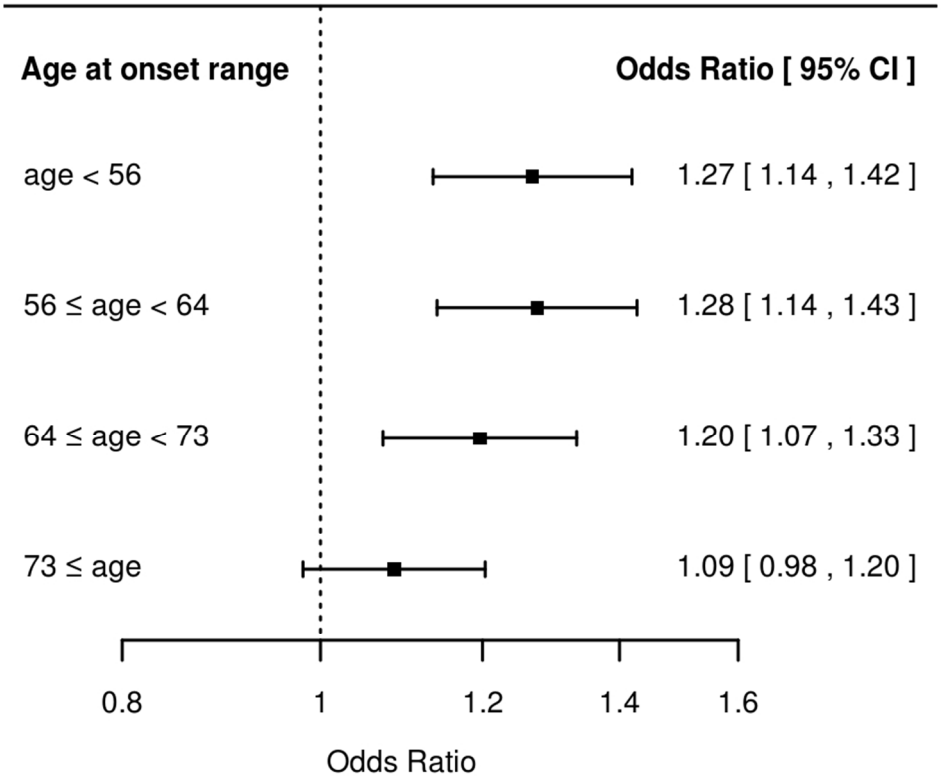

Figure 4 – Association of rs12445022 with small vessel stroke by quartiles of age-at-stroke onset in Europeans

85x75mm (300 x 300 DPI)

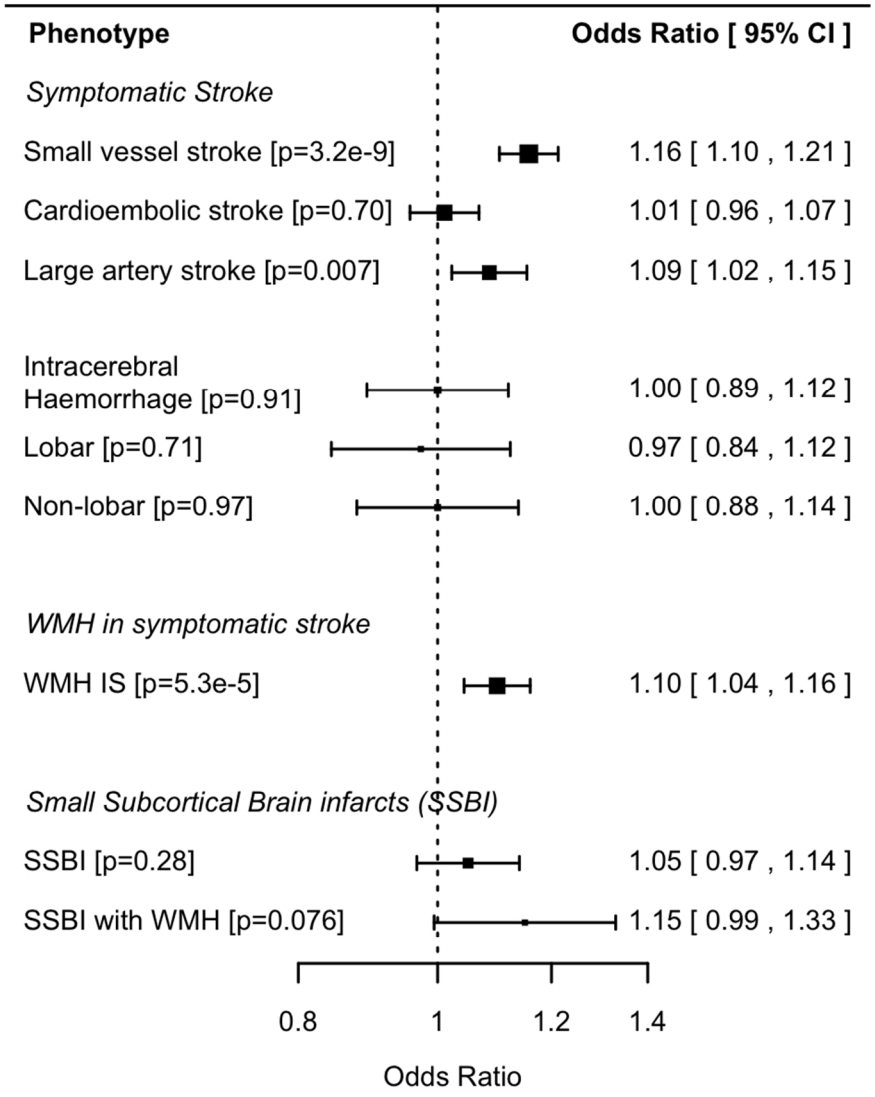

Figure 5 – Associations with rs12445022 for stroke and cerebral small vessel disease phenotypes

359x457mm (72 x 72 DPI)

Supplementary Table 1 – Small Vessel Stroke association statistics for SNPs taken forward to Stage II

					Stage I			Stage II & III			Overall		
SNP	CHR	BP	Allele1	Allele2	Freq1	Zscore	P.value	Freq1	Zscore	P.value	Freq1	Zscore	P.value
rs12445022	16	87575332	a	g	0.3378	4.813	1.49E-06	0.3233	3.981	6.87E-05	0.3273	5.923	3.17E-09
rs12920915	16	87577521	t	c	0.4041	4.255	2.10E-05	0.3999	3.353	0.0007993	0.4018	5.337	9.44E-08
rs17879928	1	92160300	a	g	0.7982	4.603	4.17E-06	0.7951	2.430	0.01509	0.7897	4.495	6.95E-07
rs11807260	1	92160761	a	g	0.2019	-4.595	4.32E-06	0.2051	-2.238	0.00466	0.2037	-5.175	2.27E-07
rs12444224	16	87580855	t	c	0.4035	4.149	3.34E-05	0.4008	3.224	0.001266	0.402	5.17	2.34E-07
rs6693438	1	92163174	a	g	0.7978	4.583	4.58E-06	0.794	2.151	0.006036	0.7957	5.105	3.31E-07
rs1961463	10	3506218	a	g	0.2904	-4.651	3.30E-06	0.2891	-2.356	0.01848	0.2897	-4.86	1.17E-06
rs2053889	10	3509300	t	g	0.2922	-4.42	9.89E-06	0.2902	-2.524	0.0116	0.2911	-4.83	1.36E-06
rs8073136	17	52654201	t	c	0.7574	-4.285	1.83E-05	0.7576	-2.594	0.0095	0.7575	-4.792	1.65E-06
rs7225084	17	52652493	t	g	0.7585	-4.266	1.99E-05	0.759	-2.513	0.01199	0.7588	-4.719	2.37E-06
rs115936795	17	2572778	t	g	0.0158	4.481	7.44E-06	0.0149	2.354	0.01856	0.0153	4.68	2.86E-06
rs12244869	10	3499052	c	g	0.7037	4.529	5.94E-06	0.7032	2.184	0.02894	0.7034	4.651	3.31E-06
rs138515134	17	2579590	t	g	0.9852	-4.501	6.77E-06	0.9856	-2.276	0.02284	0.9854	-4.633	3.60E-06
rs2439998	17	2537551	t	c	0.0189	4.47	7.81E-06	0.0181	2.188	0.02864	0.0185	4.599	4.24E-06
rs73519263	16	27123478	t	g	0.9822	-4.714	2.42E-06	0.9795	-2.085	0.03709	0.9805	-4.5	6.79E-06
rs73519272	16	27124978	t	g	0.0178	4.657	3.21E-06	0.0206	2.056	0.03983	0.0196	4.442	8.90E-06
rs896644	18	59247445	a	t	0.9107	-4.564	5.01E-06	0.9107	-1.801	0.07176	0.9107	-4.389	1.14E-05
rs1452583	18	59241941	a	t	0.0883	4.555	5.24E-06	0.0893	1.751	0.07989	0.0889	4.346	1.39E-05
rs148398060	16	27105221	a	g	0.017	4.494	7.01E-06	0.0202	1.803	0.07147	0.0189	4.24	2.24E-05
rs9827208	3	7908121	a	g	0.688	-5.102	3.36E-07	0.6925	-1.1	0.2713	0.6905	-4.227	2.37E-05
rs6567220	18	59241643	a	g	0.9098	-4.665	3.08E-06	0.9076	-1.475	0.1401	0.9086	-4.214	2.51E-05
rs2083833	17	52655010	t	c	0.192	4.353	1.35E-05	0.1892	1.742	0.08143	0.1904	4.204	2.62E-05
rs7615995	3	7923100	a	t	0.3109	5.048	4.46E-07	0.3085	1.117	0.2641	0.3096	4.203	2.63E-05
rs13386661	2	59883352	t	c	0.0196	4.61	4.03E-06	0.0189	1.528	0.1265	0.0192	4.153	3.29E-05
rs72952365	2	59859190	t	c	0.0162	4.212	2.53E-05	0.0182	1.744	0.08115	0.0174	4.059	4.93E-05
rs6737557	2	59885842	t	c	0.9834	-4.177	2.95E-05	0.981	-1.693	0.09036	0.982	-3.998	6.39E-05
rs17076342	3	44079629	a	g	0.8959	-4.171	3.03E-05	0.9035	-1.185	0.2361	0.9001	-3.668	0.0002445
rs9871447	3	44079944	t	c	0.103	4.11	3.96E-05	0.0976	1.234	0.2172	0.1	3.663	0.0002489
rs9817280	3	44083722	a	g	0.0986	4.155	3.26E-05	0.0965	1.182	0.2371	0.0974	3.655	0.0002571

rs373151	3	7927037	a	g	0.5832	4.992	5.97E-07	0.593	0.421	0.6738	0.5886	3.648	0.0002644
rs2019441	18	23197880	t	c	0.9643	-4.465	8.00E-06	0.9652	-0.647	0.5178	0.9648	-3.464	0.0005324
rs870399	11	131456204	t	c	0.2958	-4.114	3.88E-05	0.2894	-0.902	0.3673	0.2922	-3.403	0.0006656
rs9356705	6	19677864	t	c	0.0358	4.859	1.18E-06	0.0337	0.334	0.7384	0.0346	3.375	0.0007389
rs73395542	18	23192332	a	g	0.9639	-4.336	1.45E-05	0.9649	-0.534	0.593	0.9645	-3.294	0.0009886
rs4800650	18	23192273	a	t	0.9639	-4.334	1.47E-05	0.9649	-0.531	0.5953	0.9645	-3.29	0.001002
rs2190179	7	79045496	a	c	0.265	4.282	1.85E-05	0.273	0.514	0.6069	0.2694	3.243	0.001181
rs10238315	7	79043013	t	g	0.2653	4.279	1.88E-05	0.2729	0.503	0.6152	0.2695	3.232	0.001229
rs2288131	12	29683089	t	c	0.3092	4.347	1.38E-05	0.3124	0.391	0.696	0.311	3.194	0.001401
rs299456	12	29695861	a	g	0.3155	4.246	2.17E-05	0.3198	0.277	0.782	0.3179	3.042	0.002348
rs299457	12	29696329	t	c	0.3156	4.245	2.18E-05	0.3198	0.271	0.7866	0.3179	3.037	0.002388
rs4730806	7	79030922	a	c	0.7351	-4.103	4.08E-05	0.7269	-0.305	0.7601	0.7306	-2.968	0.002998
rs60475239	2	219047216	a	t	0.5464	-4.811	1.50E-06	0.3428	1.062	0.2884	0.4529	-2.819	0.004813
rs9368108	6	19692570	a	t	0.9555	-4.864	1.15E-06	0.9604	0.43	0.6673	0.9584	-2.792	0.005238
rs72657608	6	19691522	t	c	0.0445	4.864	1.15E-06	0.0397	-0.433	0.6647	0.0417	2.789	0.005281
rs4937638	11	131456484	a	t	0.4282	4.219	2.46E-05	0.4153	0.221	0.825	0.4201	2.76	0.005779
rs12363659	11	131456719	a	g	0.5718	-4.211	2.54E-05	0.5847	-0.215	0.8296	0.5799	-2.751	0.005948
rs12743454	1	208339738	a	g	0.6561	4.372	1.23E-05	0.6511	-0.313	0.7544	0.6533	2.687	0.007203
rs2590685	1	208339232	a	g	0.6553	4.316	1.59E-05	0.6493	-0.287	0.7739	0.652	2.669	0.007609
rs4674267	2	219046437	t	c	0.4322	4.894	9.86E-07	0.4468	-1.073	0.2834	0.4403	2.471	0.01348
rs11894169	2	219046846	t	c	0.4332	4.802	1.57E-06	0.4474	-1.075	0.2823	0.4411	2.408	0.01605
rs13181146	5	176128648	a	t	0.8334	4.329	1.50E-05	0.8192	-0.735	0.4622	0.8252	2.247	0.02462
rs2884531	5	176129392	t	c	0.1664	-4.275	1.91E-05	0.1812	0.794	0.4273	0.175	-2.168	0.03018
rs9515350	13	111668229	a	g	0.869	-3.552	0.0003818	0.8635	0.425	0.671	0.866	-2.057	0.03971
rs139728593	8	99641925	a	c	0.0122	4.741	2.13E-06	0.0109	-1.455	0.1456	0.0115	2.023	0.04309
rs11737790	4	120751664	t	c	0.337	4.304	1.68E-05	0.3274	-1.17	0.242	0.3317	2.004	0.04503
rs4868697	5	176129823	c	g	0.8232	4.196	2.72E-05	0.8054	-0.997	0.3186	0.8129	1.962	0.04981
rs6534157	4	120753056	c	g	0.6642	-4.272	1.94E-05	0.6738	1.286	0.1983	0.6695	-1.896	0.05791
rs6858155	4	120754216	t	g	0.3358	4.27	1.96E-05	0.3266	-1.306	0.1917	0.3307	1.88	0.06006
rs10996826	10	67776227	a	g	0.5221	-4.33	1.49E-05	0.5205	1.398	0.1621	0.5212	-1.852	0.06406
rs10822693	10	67776999	c	g	0.478	4.331	1.48E-05	0.4795	-1.4	0.1614	0.4788	1.851	0.06419
rs189950918	8	99535634	c	g	0.0117	4.342	1.41E-05	0.0103	-1.36	0.1737	0.0109	1.832	0.06695

rs7098171	10	67780663	t	c	0.5133	-4.437	9.14E-06	0.518	1.527	0.1267	0.5159	-1.827	0.06772
rs6656221	1	208329618	a	g	0.6852	3.907	9.35E-05	0.687	-1.164	0.2443	0.6862	1.743	0.08129
rs67730808	13	111661710	a	g	0.1347	3.307	0.0009441	0.1362	-0.691	0.4899	0.1355	1.695	0.09009
rs146021580	3	121766140	a	c	0.0244	4.355	1.33E-05	0.0235	-1.677	0.09346	0.0239	1.661	0.09672
rs9515357	13	111683822	t	c	0.128	3.426	0.0006122	0.1217	-0.889	0.3738	0.1245	1.627	0.1038
rs35538837	3	121659202	a	g	0.0269	4.188	2.82E-05	0.0262	-1.659	0.09718	0.0265	1.563	0.1181
rs142283780	22	24334674	t	c	0.1523	4.075	4.59E-05	0.1538	-1.161	0.2458	0.1532	1.559	0.1189
rs118094794	8	99479692	a	g	0.9882	-4.039	5.38E-05	0.9898	1.47	0.1415	0.9891	-1.549	0.1213
rs140004544	3	121689502	t	c	0.0281	4.245	2.19E-05	0.0273	-1.761	0.07824	0.0277	1.525	0.1273
rs4337774	4	67107733	a	g	0.0548	4.243	2.21E-05	0.054	-1.633	0.1024	0.0543	1.508	0.1317
rs11131636	4	67108352	a	g	0.0572	4.321	1.55E-05	0.0544	-1.731	0.0834	0.0556	1.484	0.1379
rs79215719	4	67106740	c	g	0.0544	4.288	1.80E-05	0.0523	-1.734	0.08288	0.0532	1.46	0.1442
rs7291499	22	24341101	a	g	0.1545	4.17	3.04E-05	0.1542	-1.395	0.1631	0.1543	1.431	0.1524
rs182268832	22	24334133	t	g	0.1456	4.062	4.88E-05	0.1484	-1.371	0.1705	0.1474	1.384	0.1663

CHR, chromosome; BP, base position; Freq1, frequency of Allele1; Stage III results included for rs12445022 and rs17879928 only

Supplementary Table 2 – Large Artery Stroke association statistics for SNPs taken forward to Stage II

					Stage I			Stage II			Overall		
SNP	CHR	BP	Allele1	Allele2	Freq1	Zscore	P.value	Freq1	Zscore	P.value	Freq1	Zscore	P.value
rs71524263	7	19030278	t	c	0.1157	5.309	1.10E-07	0.1218	4.769	1.85E-06	0.1194	7.047	1.83E-12
rs2107595	7	19049388	a	g	0.1614	5.541	3.02E-08	0.163	4.19	2.78E-05	0.1623	6.815	9.46E-12
rs57301765	7	19052733	a	g	0.1634	5.502	3.77E-08	0.1669	4.134	3.56E-05	0.1654	6.747	1.51E-11
rs12939005	17	52986484	t	g	0.2018	4.379	1.19E-05	0.1903	2.384	0.01712	0.1954	4.693	2.69E-06
rs72749603	5	31178840	a	c	0.9636	-4.865	1.14E-06	0.9628	-1.943	0.05206	0.9632	-4.687	2.77E-06
rs72829469	17	52989920	a	g	0.2025	4.375	1.22E-05	0.1904	2.352	0.01868	0.1958	4.666	3.07E-06
rs2332312	17	52996429	a	g	0.2033	4.394	1.11E-05	0.1927	2.202	0.02769	0.1974	4.567	4.95E-06
rs62367577	5	86802658	t	g	0.8479	-4.236	2.28E-05	0.8436	-2.299	0.02152	0.8455	-4.534	5.78E-06
rs7342953	17	63245585	t	c	0.0316	4.722	2.33E-06	0.0298	1.83	0.0672	0.0306	4.508	6.54E-06
rs59192960	5	31181789	t	c	0.0385	4.569	4.91E-06	0.0377	1.903	0.05708	0.0381	4.46	8.20E-06
rs80160193	17	63243822	a	t	0.0316	4.642	3.45E-06	0.0299	1.787	0.07395	0.0307	4.422	9.76E-06
rs76490029	17	63243674	t	c	0.0316	4.64	3.48E-06	0.0299	1.787	0.074	0.0307	4.421	9.84E-06
rs11696561	20	17455185	a	g	0.2813	-4.492	7.04E-06	0.2773	-1.577	0.1149	0.2791	-4.166	3.11E-05
rs4239702	20	44749251	t	c	0.2749	4.652	3.29E-06	0.2867	1.423	0.1546	0.2815	4.157	3.22E-05
rs55862206	4	129591442	c	g	0.9211	-4.677	2.91E-06	0.9227	-1.378	0.1681	0.922	-4.14	3.47E-05
rs56299512	4	129591445	a	g	0.9211	-4.677	2.91E-06	0.9227	-1.378	0.1681	0.922	-4.14	3.47E-05
rs10811652	9	22077085	a	c	0.5057	-4.419	9.93E-06	0.4967	-1.567	0.1171	0.5007	-4.109	3.97E-05
rs72783748	5	86863605	a	g	0.7635	-4.319	1.57E-05	0.7485	-1.638	0.1014	0.7551	-4.096	4.20E-05
rs10757271	9	22076795	a	g	0.5051	-4.37	1.24E-05	0.4962	-1.586	0.1128	0.5001	-4.091	4.30E-05
rs34562050	4	129594713	c	g	0.08	4.677	2.91E-06	0.0789	1.224	0.221	0.0794	4.025	5.69E-05
rs11607832	11	102544684	t	c	0.9299	-4.008	6.13E-05	0.9332	-1.761	0.0783	0.9317	-3.981	6.88E-05
rs9632884	9	22072301	c	g	0.4918	4.389	1.14E-05	0.5007	1.241	0.2145	0.4968	3.846	0.00012
rs11905301	20	24531347	t	c	0.4669	4.274	1.92E-05	0.5327	1.404	0.1603	0.5054	3.829	0.0001289
rs7720392	5	31170385	t	c	0.9661	-4.074	4.63E-05	0.9668	-1.451	0.1467	0.9665	-3.793	0.0001486
rs4810485	20	44747947	t	g	0.2506	4.092	4.29E-05	0.2608	1.413	0.1576	0.2563	3.777	0.0001587
rs12635174	3	117008209	t	c	0.7228	4.45	8.60E-06	0.7307	1.08	0.28	0.7272	3.767	0.0001654
rs11696580	20	17455276	t	g	0.1926	-4.222	2.42E-05	0.198	-1.247	0.2122	0.1956	-3.74	0.0001838
rs1883832	20	44746982	t	c	0.2507	4.039	5.38E-05	0.2607	1.405	0.16	0.2563	3.736	0.0001872
rs74725770	7	93946272	a	g	0.989	-4.857	1.19E-06	0.989	-0.903	0.3665	0.989	-3.716	0.0002021

rs10049229	3	117001187	a	g	0.2775	-4.45	8.58E-06	0.2703	-0.962	0.3359	0.2735	-3.679	0.0002342
rs6075215	20	17454302	a	g	0.1931	-4.225	2.40E-05	0.1983	-1.142	0.2534	0.196	-3.663	0.0002492
rs13092499	3	117009819	a	c	0.7229	4.45	8.58E-06	0.73	0.863	0.388	0.7269	3.605	0.0003122
rs13039850	20	24547144	t	c	0.3574	-4.545	5.51E-06	0.348	-0.646	0.5181	0.3522	-3.506	0.0004554
rs1561734	5	86826750	a	g	0.2109	4.298	1.73E-05	0.2253	0.839	0.4015	0.2189	3.485	0.0004913
rs75684303	7	93878978	a	g	0.011	4.704	2.55E-06	0.008	0.773	0.4398	0.0091	3.442	0.0005773
rs11159150	14	76162001	a	g	0.3084	-4.34	1.43E-05	0.3038	-0.646	0.5185	0.3058	-3.369	0.0007545
rs1005224	14	76173860	a	t	0.6902	4.286	1.82E-05	0.6912	0.683	0.4944	0.6908	3.361	0.0007756
rs727409	20	24557527	t	c	0.3202	-4.252	2.12E-05	0.3164	-0.692	0.4889	0.3181	-3.345	0.0008222
rs662558	11	102718695	t	c	0.8213	-3.524	0.0004255	0.8117	-1.321	0.1865	0.8159	-3.33	0.0008671
rs2660040	10	68275581	a	g	0.294	4.797	1.61E-06	0.2961	0.142	0.8872	0.2952	3.297	0.000978
rs8073683	17	17555596	t	c	0.474	-4.525	6.05E-06	0.4749	-0.327	0.7436	0.4745	-3.254	0.001137
rs1522966	3	34843726	a	t	0.5229	4.594	4.36E-06	0.5211	0.227	0.8206	0.5219	3.225	0.001259
rs118001792	11	41905979	a	c	0.0136	4.391	1.13E-05	0.0143	0.32	0.7492	0.014	3.16	0.00158
rs586701	11	102724730	t	g	0.8236	-3.194	0.001405	0.8134	-1.372	0.1701	0.8179	-3.149	0.001639
rs61103969	11	41897789	a	g	0.0136	4.526	6.00E-06	0.0145	0.174	0.8622	0.0141	3.141	0.001685
rs525673	3	34848226	t	c	0.523	4.636	3.56E-06	0.5227	0.063	0.95	0.5228	3.131	0.001744
rs7893676	10	68270632	c	g	0.2923	4.152	3.30E-05	0.2863	0.483	0.629	0.289	3.123	0.001791
rs12258032	10	68271859	t	c	0.7076	-4.138	3.50E-05	0.7137	-0.485	0.6275	0.711	-3.115	0.001838
rs1916227	3	34847097	a	g	0.4774	-4.58	4.64E-06	0.4776	-0.069	0.9454	0.4775	-3.098	0.001947
rs11626058	14	76115445	a	c	0.3724	-4.253	2.11E-05	0.3734	-0.346	0.7296	0.373	-3.087	0.002021
rs58980959	11	41903641	c	g	0.0136	4.387	1.15E-05	0.0145	0.194	0.8461	0.0141	3.063	0.00219
rs7164538	15	60155804	a	g	0.9679	4.824	1.41E-06	0.9714	-0.337	0.7361	0.9699	2.957	0.003103
rs7164565	15	60155858	a	g	0.9679	4.823	1.41E-06	0.9714	-0.338	0.7356	0.9699	2.957	0.00311
rs191633333	7	93746175	a	c	0.0112	4.804	1.56E-06	0.011	-0.065	0.9484	0.0111	2.872	0.004074
rs12414028	10	104957629	a	t	0.0924	-4.102	4.10E-05	0.0959	-0.181	0.856	0.0944	-2.864	0.004181
rs2099580	15	60157890	t	g	0.9683	4.852	1.22E-06	0.9718	-0.449	0.6535	0.9703	2.861	0.004218
rs16870743	5	73016810	a	g	0.7238	-4.465	8.01E-06	0.7232	0.168	0.8664	0.7235	-2.845	0.004447
rs644993	11	97541637	a	g	0.1229	4.371	1.24E-05	0.1276	-0.138	0.8906	0.1255	2.805	0.005027
rs79472290	4	156589785	t	c	0.0149	4.065	4.80E-05	0.0152	0.384	0.701	0.0151	2.797	0.005153
rs680900	11	97538429	a	g	0.127	4.419	9.91E-06	0.1336	-0.231	0.8176	0.1307	2.768	0.005646
rs59216006	5	73001469	a	g	0.7174	-4.378	1.20E-05	0.7156	0.21	0.8338	0.7164	-2.756	0.005856

rs17400036	6	69707211	t	g	0.9298	-4.285	1.83E-05	0.9294	0.133	0.8939	0.9296	-2.751	0.005944
rs56120605	6	69714649	a	g	0.9299	-4.32	1.56E-05	0.9304	0.174	0.8622	0.9302	-2.744	0.006072
rs6878052	5	72999965	a	g	0.2824	4.371	1.24E-05	0.2842	-0.229	0.8188	0.2834	2.737	0.006202
rs11191548	10	104846178	t	c	0.919	3.734	0.0001888	0.9099	0.297	0.7662	0.9139	2.706	0.006816
rs9901291	17	17541750	a	g	0.4376	4.577	4.73E-06	0.447	-0.482	0.6295	0.4428	2.684	0.007267
rs597928	11	97545967	a	c	0.1251	4.444	8.85E-06	0.1282	-0.375	0.708	0.1268	2.676	0.007442
rs3740390	10	104638480	t	c	0.0841	-3.774	0.0001605	0.092	0.004	0.9965	0.0885	-2.507	0.01216
rs12972736	19	6656777	t	c	0.2588	-4.538	5.67E-06	0.2626	0.702	0.4829	0.261	-2.389	0.01688
rs12977994	19	6655261	t	c	0.2543	-4.307	1.66E-05	0.2564	0.742	0.4579	0.2555	-2.209	0.02717
rs56111161	1	204695345	t	c	0.014	4.945	7.63E-07	0.0133	-1.322	0.1863	0.0136	2.177	0.02946
rs11240252	1	204691325	a	g	0.0143	4.98	6.35E-07	0.0138	-1.703	0.08851	0.014	1.909	0.05632
rs55878455	6	69639391	t	c	0.9259	-3.878	0.0001054	0.9211	0.824	0.4097	0.9231	-1.877	0.06059
rs3738157	1	204684426	a	g	0.0119	4.941	7.76E-07	0.012	-1.653	0.09839	0.012	1.818	0.06912
rs62066209	17	17520784	t	c	0.388	4.531	5.88E-06	0.3868	-1.609	0.1077	0.3873	1.813	0.06984
rs12608923	19	6659855	a	g	0.2464	-4.435	9.22E-06	0.2482	1.191	0.2337	0.2475	-1.724	0.08466

CHR, chromosome; BP, base position; Freq1, frequency of Allele1

Supplementary Table 3 – Cardioembolic Stroke association statistics for SNPs taken forward to Stage II

					Stage I			Stage II & III			Overall		
SNP	CHR	BP	Allele1	Allele2	Freq1	Zscore	P.value	Freq1	Zscore	P.value	Freq1	Zscore	P.value
rs192172299	4	111681501	t	g	0.2148	5.994	2.05E-09	0.2217	6.832	8.39E-12	0.2187	9.088	1.01E-19
rs2466455	4	111685615	t	c	0.7979	-5.725	1.04E-08	0.7763	-7.052	1.76E-12	0.7858	-9.076	1.13E-19
rs2634074	4	111677041	a	t	0.7975	-5.739	9.55E-09	0.7759	-7.018	2.25E-12	0.7854	-9.059	1.32E-19
rs879324	16	73068678	a	g	0.1721	5.026	5.02E-07	0.1681	3.371	0.0007501	0.1699	5.853	4.83E-09
rs12932445	16	73069888	t	c	0.8278	-4.866	1.14E-06	0.8313	-3.293	0.0009929	0.8298	-5.689	1.28E-08
rs4499262	16	73059159	a	c	0.168	4.893	9.94E-07	0.165	3.043	0.002345	0.1663	5.519	3.40E-08
rs8134198	21	45369535	t	g	0.3246	4.717	2.40E-06	0.3302	2.308	0.02102	0.3335	4.402	1.07E-05
rs139223992	5	31110574	t	c	0.9116	-4.593	4.37E-06	0.9115	-2.319	0.02037	0.9115	-4.779	1.77E-06
rs72747745	5	31111931	a	g	0.0879	4.498	6.87E-06	0.088	2.372	0.01771	0.088	4.755	1.99E-06
rs183560286	6	105692807	a	g	0.0769	4.254	2.10E-05	0.0719	2.64	0.008287	0.0739	4.721	2.35E-06
rs4941428	13	43008177	a	g	0.2564	4.269	1.96E-05	0.2494	2.52	0.01172	0.2525	4.715	2.42E-06
rs11740741	5	31114581	t	c	0.0897	4.573	4.81E-06	0.0888	2.198	0.02792	0.0892	4.675	2.94E-06
rs7100632	10	134653951	t	c	0.5884	-4.463	8.08E-06	0.6015	-2.287	0.02218	0.5958	-4.669	3.03E-06
rs17065682	6	105704149	t	g	0.9235	-4.091	4.29E-05	0.9269	-2.683	0.007303	0.9256	-4.653	3.28E-06
rs17065694	6	105704658	t	c	0.0766	4.101	4.11E-05	0.073	2.648	0.008106	0.0744	4.631	3.63E-06
rs3750580	10	134650440	t	c	0.6018	-4.296	1.74E-05	0.6147	-2.367	0.01794	0.609	-4.618	3.88E-06
rs4880434	10	134648569	a	g	0.5867	-4.424	9.69E-06	0.6007	-2.155	0.03113	0.5946	-4.544	5.52E-06
rs2147160	13	43021639	a	c	0.7445	-4.036	5.43E-05	0.7516	-2.491	0.01275	0.7485	-4.539	5.66E-06
rs1351832	13	43057549	t	c	0.255	4.064	4.82E-05	0.2483	2.417	0.01564	0.2512	4.502	6.73E-06
rs11868827	17	69773732	t	c	0.1269	-4.4	1.08E-05	0.1239	-1.94	0.05237	0.1252	-4.367	1.26E-05
rs552148	9	136153481	t	c	0.2482	-4.133	3.59E-05	0.2414	-2.126	0.03353	0.2444	-4.329	1.50E-05
rs663367	9	136153451	a	g	0.2481	-4.125	3.71E-05	0.2413	-2.117	0.03428	0.2443	-4.317	1.58E-05
rs12453161	17	69752498	a	g	0.1291	-4.431	9.36E-06	0.1252	-1.84	0.06581	0.1269	-4.312	1.62E-05
rs500428	9	136155343	a	g	0.2483	-4.115	3.88E-05	0.2415	-2.11	0.03483	0.2445	-4.305	1.67E-05
rs1953522	1	216837241	a	c	0.3773	-4.493	7.04E-06	0.3647	-1.697	0.08962	0.3702	-4.246	2.18E-05
rs7252834	19	16075211	a	g	0.7727	-4.205	2.61E-05	0.7502	-1.774	0.07608	0.7601	-4.113	3.90E-05
rs73226947	21	45363522	t	g	0.3552	3.848	0.0001189	0.3598	2.042	0.04116	0.3578	4.065	4.79E-05
rs2838444	21	45360862	a	g	0.3507	3.898	9.70E-05	0.3525	1.982	0.04745	0.3517	4.053	5.06E-05
rs7252929	19	16075678	t	c	0.2274	4.257	2.07E-05	0.2493	1.627	0.1037	0.2397	4.038	5.40E-05

rs1594692	19	16075015	t	c	0.7724	-4.182	2.90E-05	0.7495	-1.673	0.09437	0.7595	-4.022	5.78E-05
rs10958962	9	10229902	t	c	0.7405	-4.381	1.18E-05	0.7464	-1.451	0.1469	0.7438	-3.987	6.69E-05
rs2429937	17	69717560	a	g	0.1482	-4.397	1.10E-05	0.1462	-1.356	0.1752	0.1471	-3.927	8.62E-05
rs922982	9	84348195	c	g	0.4474	4.319	1.57E-05	0.4505	1.342	0.1796	0.4491	3.865	0.0001112
rs4434691	9	84348762	a	c	0.4475	4.304	1.67E-05	0.4505	1.344	0.1788	0.4492	3.857	0.0001148
rs7380303	5	168722937	t	c	0.7424	-5.203	1.96E-07	0.7314	-0.406	0.685	0.7362	-3.748	0.0001783
rs4242184	5	168726086	a	g	0.7436	-5.09	3.58E-07	0.7328	-0.466	0.6415	0.7375	-3.718	0.0002007
rs1455181	9	3774843	a	c	0.4018	4.726	2.29E-06	0.4009	0.712	0.4767	0.4013	3.661	0.0002509
rs10061289	5	168718340	c	g	0.7666	-5.142	2.71E-07	0.7587	-0.337	0.7362	0.7622	-3.656	0.0002559
rs80129168	5	56481753	a	g	0.1082	-4.049	5.14E-05	0.1057	-1.227	0.2198	0.1068	-3.583	0.0003403
rs1997571	7	116198621	a	g	0.5916	4.2	2.67E-05	0.5846	1.033	0.3016	0.5877	3.554	0.0003789
rs1997572	7	116198828	a	g	0.4084	-4.2	2.67E-05	0.4154	-1.033	0.3017	0.4123	-3.554	0.0003794
rs10867804	9	84341050	a	c	0.4322	4.326	1.52E-05	0.4379	0.903	0.3666	0.4354	3.54	0.0003999
rs13361596	5	56540855	t	c	0.1088	-4.007	6.15E-05	0.1078	-1.207	0.2272	0.1082	-3.54	4.00E-04
rs3807989	7	116186241	a	g	0.4068	-4.09	4.32E-05	0.4149	-1.064	0.2875	0.4114	-3.504	0.0004579
rs2065070	9	10222790	c	g	0.7354	-4.117	3.83E-05	0.7412	-0.986	0.3239	0.7387	-3.465	0.0005308
rs12698965	7	70441601	t	c	0.6205	4.433	9.31E-06	0.6186	1.009	0.3128	0.6193	3.461	0.0005382
rs3786778	19	48285809	t	c	0.9559	-4.468	7.90E-06	0.9508	-0.661	0.5088	0.953	-3.453	0.0005551
rs12983568	19	48272855	t	c	0.9551	-4.495	6.96E-06	0.9528	-0.63	0.5284	0.9538	-3.448	0.0005651
rs10269258	7	70440091	t	c	0.3795	-4.438	9.06E-06	0.3817	-0.982	0.326	0.3809	-3.443	0.0005758
rs6460609	7	70437060	t	c	0.6203	4.394	1.11E-05	0.6182	0.965	0.3344	0.619	3.403	0.0006674
rs10471998	5	56447742	a	g	0.8859	4.458	8.26E-06	0.8908	0.911	0.3621	0.8891	3.366	0.0007637
rs75781828	3	6994624	t	c	0.1182	-4.347	1.38E-05	0.1123	-0.548	0.5837	0.1149	-3.288	0.001008
rs12973532	19	48264879	a	g	0.9346	-4.432	9.36E-06	0.9295	-0.454	0.6501	0.9317	-3.273	0.001063
rs11071630	15	62098001	t	c	0.3935	4.466	7.98E-06	0.3968	0.421	0.6736	0.3954	3.272	0.001069
rs10809038	9	10229301	t	c	0.8813	-3.966	7.32E-05	0.8788	-0.83	0.4067	0.8799	-3.247	0.001167
rs1472433	10	4661676	a	g	0.2765	4.461	8.18E-06	0.2749	0.388	0.6982	0.2756	3.243	0.001182
rs10973456	9	3774359	t	c	0.5761	-4.452	8.52E-06	0.5761	-0.367	0.7139	0.5761	-3.222	0.001275
rs7910959	10	4646222	a	t	0.2755	4.466	7.97E-06	0.2743	0.354	0.7232	0.2748	3.222	0.001275
rs149397829	3	6991641	t	c	0.1161	-4.267	1.98E-05	0.1121	-0.521	0.6021	0.1139	-3.215	0.001303
rs1021632	10	4652819	t	g	0.2764	4.451	8.54E-06	0.2753	0.357	0.721	0.2758	3.214	0.001309
rs116972146	13	69632241	t	c	0.9641	-4.219	2.46E-05	0.9659	-0.562	0.5743	0.9651	-3.214	0.001311

rs13379934	15	62096931	a	t	0.7639	-3.906	9.37E-05	0.7501	-0.771	0.441	0.7561	-3.163	0.00156
rs181957491	13	69635572	a	g	0.9672	-4.11	3.95E-05	0.9682	-0.543	0.5869	0.9678	-3.128	0.00176
rs2163735	3	6995426	a	g	0.116	-4.171	3.04E-05	0.1117	-0.434	0.6643	0.1136	-3.086	0.00203
rs1268153	6	109035696	a	g	0.0324	4.286	1.82E-05	0.0319	0.512	0.6086	0.0321	3.081	0.002062
rs34716044	15	62097134	t	g	0.762	-3.767	0.0001655	0.7487	-0.729	0.4658	0.7545	-3.04	0.002367
rs6901124	6	109035704	t	c	0.0314	4.092	4.27E-05	0.0299	0.584	0.5594	0.0305	3.016	0.002562
rs78031559	13	69621457	t	g	0.9607	-4.412	1.02E-05	0.9612	0.056	0.9553	0.961	-2.878	0.003996
rs11699273	20	58257735	a	g	0.0466	4.238	2.25E-05	0.0463	0.056	0.9553	0.0464	2.847	0.00441
rs10814589	9	3772674	a	g	0.3173	-2.646	0.008143	0.3204	-1.436	0.1511	0.319	-2.828	0.00469
rs77749998	20	58259307	t	c	0.0418	4.193	2.76E-05	0.0438	-0.272	0.7859	0.0429	2.572	0.01012
rs11696365	20	58255101	t	c	0.0415	4.024	5.72E-05	0.0438	-0.295	0.7677	0.0428	2.442	0.0146
rs9384677	6	108863432	a	c	0.0398	3.86	0.0001133	0.042	-0.278	0.7812	0.041	2.347	0.01893
rs11847901	14	26862055	a	g	0.016	4.279	1.87E-05	0.0179	-0.519	0.6035	0.0172	2.272	0.02306
rs12566058	1	216182811	a	g	0.5483	-4.216	2.48E-05	0.5329	0.7	0.4838	0.5396	-2.266	0.02345
rs11120712	1	216183467	a	g	0.465	4.296	1.74E-05	0.473	-0.78	0.4353	0.4695	2.259	0.02391
rs79205855	14	26855727	c	g	0.983	-4.021	5.80E-05	0.9806	0.407	0.6841	0.9815	-2.199	0.02791
rs11845727	14	26850849	t	c	0.9834	-3.945	7.99E-05	0.9814	0.418	0.676	0.9822	-2.142	0.03218
rs142258091	1	37772341	a	t	0.9807	-4.762	1.92E-06	0.9837	1.405	0.16	0.9824	-2.064	0.03905
rs142632566	1	37747395	t	c	0.9807	-4.814	1.48E-06	0.9838	1.486	0.1373	0.9825	-2.037	0.04169
rs141334771	1	37737497	a	g	0.0191	4.781	1.74E-06	0.0161	-1.525	0.1273	0.0174	1.986	0.04704

CHR, chromosome; BP, base position; Freq1, frequency of Allele1; Stage III results included for rs8134198 only

Supplementary Table 4 – All Ischaemic Stroke association statistics for SNPs taken forward to Stage II

					Stage I			Stage II			Overall		
SNP	CHR	BP	Allele1	Allele2	Freq1	Zscore	P.value	Freq1	Zscore	P.value	Freq1	Zscore	P.value
rs11513729	12	112273499	t	c	0.4412	4.712	2.46E-06	0.4394	4.693	2.69E-06	0.4402	6.63	3.35E-11
rs635634	9	136155000	t	c	0.1936	4.48	7.47E-06	0.2077	3.818	0.0001348	0.2009	5.86	4.64E-09
rs532436	9	136149830	a	g	0.1928	4.435	9.19E-06	0.2078	3.86	0.0001132	0.2005	5.859	4.65E-09
rs507666	9	136149399	a	g	0.1928	4.433	9.29E-06	0.2078	3.862	0.0001124	0.2005	5.859	4.66E-09
rs58799304	1	156080444	a	t	0.2187	-5.05	4.42E-07	0.2348	-1.684	0.09225	0.2275	-4.644	3.42E-06
rs12035615	1	156055344	t	c	0.2231	-4.767	1.87E-06	0.2227	-1.753	0.07954	0.2229	-4.579	4.68E-06
rs1440410	4	144158309	a	t	0.5364	-4.481	7.43E-06	0.5306	-2.024	0.04296	0.5334	-4.573	4.80E-06
rs3814314	1	156045183	a	t	0.2263	-4.774	1.81E-06	0.2262	-1.739	0.08197	0.2262	-4.573	4.80E-06
rs181541997	6	47107964	a	c	0.0251	4.577	4.71E-06	0.0224	1.816	0.06939	0.0237	4.491	7.09E-06
rs4322086	9	85787718	a	g	0.5302	4.382	1.18E-05	0.5399	2.003	0.04522	0.5352	4.489	7.16E-06
rs761521	6	47112002	t	c	0.0252	4.618	3.88E-06	0.0224	1.746	0.08086	0.0238	4.469	7.86E-06
rs4618266	4	144036127	t	c	0.5608	-4.642	3.45E-06	0.5547	-1.71	0.08726	0.5577	-4.46	8.18E-06
rs6808694	3	20186297	t	c	0.7309	4.42	9.89E-06	0.728	1.923	0.05451	0.7294	4.458	8.27E-06
rs6805673	3	20185804	a	g	0.7271	4.374	1.22E-05	0.7238	1.944	0.05188	0.7254	4.441	8.94E-06
rs9999826	4	144036457	a	g	0.4391	4.635	3.56E-06	0.445	1.678	0.09335	0.4421	4.433	9.30E-06
rs11759394	6	47115212	t	c	0.9747	-4.637	3.53E-06	0.9774	-1.66	0.09688	0.9761	-4.421	9.81E-06
rs77215829	12	112618346	a	c	0.8808	4.519	6.21E-06	0.8751	1.689	0.09122	0.8775	4.225	2.39E-05
rs138239252	3	41764014	t	c	0.1585	-4.2	2.68E-05	0.172	-1.71	0.08727	0.1658	-4.105	4.05E-05
rs56729913	3	20184853	a	t	0.2693	-4.29	1.79E-05	0.2669	-1.509	0.1312	0.2681	-4.071	4.68E-05
rs6807015	3	41754514	t	c	0.8337	4.074	4.62E-05	0.8291	1.769	0.07698	0.8312	4.063	4.85E-05
rs10010247	4	138466213	t	c	0.5888	-5.39	7.04E-08	0.5844	-0.297	0.7664	0.5865	-3.968	7.26E-05
rs2711814	11	29376425	t	c	0.3867	-4.608	4.07E-06	0.3874	-1.054	0.2918	0.3871	-3.966	7.31E-05
rs2711824	11	29384629	t	c	0.6127	4.585	4.54E-06	0.6123	1.055	0.2914	0.6125	3.951	7.79E-05
rs28695597	4	138458960	a	c	0.4098	5.233	1.66E-07	0.4151	0.401	0.6888	0.4125	3.933	8.40E-05
rs12002585	9	81958890	t	c	0.928	4.146	3.39E-05	0.9264	1.353	0.176	0.9272	3.858	0.0001141
rs56291600	17	38638260	t	c	0.9326	-4.613	3.97E-06	0.9278	-0.883	0.3772	0.9301	-3.847	0.0001197
rs2151155	11	29452159	t	c	0.3883	-4.645	3.41E-06	0.3339	-0.85	0.3956	0.3603	-3.845	0.0001207
rs10867369	9	81980758	t	c	0.0731	-4.159	3.20E-05	0.0746	-1.273	0.2031	0.0739	-3.81	0.0001389
rs17472276	17	38634605	t	c	0.9331	-4.658	3.20E-06	0.9275	-0.726	0.4679	0.9302	-3.765	0.0001664

rs72819648	17	38635876	t	c	0.067	4.662	3.13E-06	0.0725	0.721	0.4708	0.0698	3.765	0.0001667
rs12551670	9	81957986	a	c	0.0725	-4.035	5.46E-05	0.0736	-1.287	0.1981	0.0731	-3.734	0.0001885
rs1575506	9	85788822	a	g	0.3229	4.512	6.44E-06	0.325	0.815	0.4148	0.324	3.728	0.0001933
rs9817510	3	41813108	t	c	0.1676	-3.941	8.10E-05	0.1556	-1.417	0.1565	0.1611	-3.714	0.0002039
rs61493282	11	102754868	t	c	0.0788	4.41	1.04E-05	0.0818	0.859	0.3903	0.0803	3.688	0.0002261
rs6967828	7	123254157	t	c	0.9192	-4.778	1.77E-06	0.9218	-0.414	0.6788	0.9205	-3.625	0.000289
rs76386877	4	138422617	t	c	0.613	-5.539	3.03E-08	0.5862	0.336	0.7367	0.5992	-3.617	0.0002978
rs73718437	7	123253277	t	c	0.0809	4.808	1.52E-06	0.0779	0.352	0.725	0.0794	3.602	0.0003163
rs6970517	7	123252866	a	g	0.0808	4.802	1.57E-06	0.0778	0.349	0.7271	0.0793	3.595	0.0003244
rs13276017	8	134741381	a	c	0.2573	5.16	2.48E-07	0.2528	-0.022	0.9828	0.255	3.578	0.0003458
rs7865702	9	85790620	t	c	0.6731	-4.387	1.15E-05	0.6702	-0.715	0.4745	0.6716	-3.569	0.0003585
rs72983521	11	102730033	t	c	0.9216	-4.265	2.00E-05	0.9205	-0.744	0.4566	0.921	-3.505	0.0004565
rs113242701	8	73598626	a	g	0.9813	-4.817	1.46E-06	0.9838	-0.082	0.9346	0.9826	-3.397	0.0006815
rs60549974	12	112591821	t	c	0.9029	4.097	4.19E-05	0.9062	0.75	0.4531	0.9046	3.392	0.0006941
rs186861709	8	73630793	a	g	0.0171	4.668	3.05E-06	0.0165	0.201	0.8403	0.0168	3.379	0.0007265
rs2958828	8	134776896	t	c	0.2783	4.493	7.03E-06	0.275	0.335	0.7374	0.2766	3.37	0.0007517
rs10107182	8	59392737	t	c	0.6623	-4.13	3.63E-05	0.6526	-0.685	0.4933	0.6573	-3.368	0.0007563
rs72985562	11	102800278	t	g	0.9195	-4.152	3.30E-05	0.9172	-0.61	0.5418	0.9183	-3.33	0.0008697
rs4738684	8	59393273	a	g	0.3401	4.045	5.24E-05	0.348	0.713	0.476	0.3442	3.329	0.0008726
rs113145661	8	73510416	c	g	0.9794	-4.634	3.59E-06	0.9829	-0.231	0.8172	0.9813	-3.327	0.0008782
rs9297994	8	59392324	a	g	0.6614	-4.09	4.32E-05	0.6522	-0.633	0.5268	0.6567	-3.303	0.0009578
rs41276920	15	90347920	a	g	0.0798	-3.96	7.51E-05	0.0764	-0.755	0.45	0.078	-3.3	0.000967
rs2833496	21	33140026	t	c	0.2049	4.673	2.97E-06	0.2137	0.049	0.9611	0.2094	3.29	0.001003
rs72754570	15	90348979	a	g	0.0834	-4.176	2.96E-05	0.0801	-0.5	0.6169	0.0817	-3.284	0.001025
rs204746	21	33139745	c	g	0.7951	-4.676	2.93E-06	0.7862	-0.032	0.9743	0.7905	-3.28	0.001038
rs9978172	21	33140566	t	c	0.2049	4.678	2.90E-06	0.2138	0.031	0.9756	0.2095	3.28	0.001038
rs2922495	8	134735858	t	c	0.3205	4.607	4.08E-06	0.3218	0.022	0.9822	0.3212	3.225	0.00126
rs7922120	10	115051007	t	c	0.0366	4.075	4.61E-05	0.0358	0.55	0.5822	0.0362	3.213	0.001315
rs118187259	10	115121967	t	c	0.0121	4.009	6.10E-05	0.0103	0.675	0.4994	0.0111	3.185	0.001445
rs7168849	15	90346227	a	g	0.9041	3.977	6.98E-05	0.9141	0.581	0.5609	0.9096	3.102	0.001922
rs12218673	10	115053330	t	c	0.0378	4.026	5.67E-05	0.0351	0.425	0.6711	0.0364	3.088	0.002014
rs8113355	19	49062163	a	g	0.7063	-4.427	9.54E-06	0.7067	0.292	0.77	0.7065	-2.835	0.00458

rs8100256	19	49062157	t	g	0.7063	-4.426	9.62E-06	0.7067	0.292	0.7703	0.7065	-2.834	0.004593
rs4981631	14	27312452	a	c	0.7508	4.65	3.33E-06	0.7514	-0.581	0.5611	0.7511	2.822	0.004779
rs1956818	14	27311163	c	g	0.262	-4.584	4.57E-06	0.259	0.55	0.5825	0.2605	-2.798	0.00514
rs111955440	5	121004011	t	c	0.073	-4.808	1.53E-06	0.0736	0.451	0.652	0.0733	-2.789	0.005284
rs3848542	19	49061724	a	g	0.2956	4.365	1.27E-05	0.2954	-0.327	0.7435	0.2955	2.767	0.005655
rs11696019	2	220664236	t	c	0.2147	4.393	1.12E-05	0.2113	-0.46	0.6457	0.2129	2.73	0.006336
rs982458	2	220663477	t	c	0.2138	4.35	1.36E-05	0.2118	-0.468	0.6401	0.2128	2.694	0.007057
rs13340342	5	120989002	a	t	0.9285	4.576	4.75E-06	0.9256	-0.548	0.5836	0.9268	2.564	0.01034
rs13340387	5	120986986	t	c	0.9286	4.42	9.85E-06	0.9257	-0.484	0.6281	0.9269	2.511	0.01202
rs6720347	2	220645599	t	c	0.7667	-4.375	1.21E-05	0.7671	0.779	0.436	0.7669	-2.489	0.01282
rs111810710	15	24066424	a	g	0.0431	4.393	1.12E-05	0.0449	-0.976	0.3291	0.0441	2.262	0.02372
rs78186832	15	24069618	t	c	0.0436	4.403	1.07E-05	0.045	-1.033	0.3017	0.0444	2.227	0.02597
rs80245324	15	24069688	a	t	0.9564	-4.404	1.06E-05	0.955	1.037	0.2997	0.9556	-2.224	0.02614
rs58673065	7	1885600	a	g	0.7641	4.022	5.76E-05	0.7464	-0.823	0.4108	0.7544	2.098	0.03586
rs11762803	7	1886805	a	g	0.7644	3.975	7.05E-05	0.7464	-0.811	0.4171	0.7546	2.075	0.03803
rs11766945	7	1888054	a	g	0.2137	-4.385	1.16E-05	0.2264	1.445	0.1484	0.2206	-1.882	0.05989
rs2064738	14	27465839	a	g	0.7711	4.503	6.69E-06	0.7731	-1.751	0.07992	0.7721	1.88	0.06009

CHR, chromosome; BP, base position; Freq1, frequency of Allele1