Genetic Associations With White Matter Hyperintensities Confer Risk of Lacunar Stroke

Matthew Traylor, PhD; Loes C.A. Rutten-Jacobs, PhD; Vincent Thijs, MD; Elizabeth G. Holliday, PhD; Chris Levi, MD; Steve Bevan, PhD; Rainer Malik, PhD; Giorgio Boncoraglio, MD; Cathie Sudlow, FRCP(E); Peter M. Rothwell, FMedSci; Martin Dichgans, MD; Hugh S. Markus, DM

Background and Purpose—White matter hyperintensities (WMH) are increased in patients with lacunar stroke. Whether this is because of shared pathogenesis remains unknown. Using genetic data, we evaluated whether WMH-associated genetic susceptibility factors confer risk of lacunar stroke, and therefore whether they share pathogenesis.

Methods—We used a genetic risk score approach to test whether single nucleotide polymorphisms associated with WMH in community populations were associated with magnetic resonance imaging–confirmed lacunar stroke (n=1,373), as well as cardioembolic (n=1,331) and large vessel (n=1,472) Trial of Org 10172 in Acute Stroke Treatment subtypes, against 9,053 controls. Second, we separated lacunar strokes into those with WMH (n=568) and those without (n=787) and tested for association with the risk score in these 2 groups. In addition, we evaluated whether WMH-associated single nucleotide polymorphisms are associated with lacunar stroke, or in the 2 groups.

Results—The WMH genetic risk score was associated with lacunar stroke (odds ratio [OR; 95% confidence interval [CI]]=1.14 [1.06–1.22]; P=0.0003), in patients both with and without WMH (WMH: OR [95% CI]=1.15 [1.05–1.26]; P=0.003 and no WMH: OR [95% CI]=1.11 [1.02–1.21]; P=0.019). Conversely, the risk score was not associated with cardioembolic stroke (OR [95% CI]=1.03 [0.97–1.09]; P=0.63) or large vessel stroke (OR [95% CI]=0.99 [0.93,1.04]; P=0.39). However, none of the WMH-associated single nucleotide polymorphisms passed Bonferroni-corrected significance for association with lacunar stroke.

Conclusions—Genetic variants that influence WMH are associated with an increased risk of lacunar stroke but not cardioembolic or large vessel stroke. Some genetic susceptibility factors seem to be shared across different radiological manifestations of small vessel disease.

Key Words: cerebral small vessel diseases ■ genetics ■ genetic association studies ■ leukoencephalopathies ■ stroke, lacunar

Cerebral small vessel disease (SVD) affects the small perforating arteries of the brain and is characterized radiologically by several features, including white matter hyperintensities (WMH), subcortical lacunar infarcts, intracerebral hemorrhages, and cerebral microbleeds. Despite the considerable impact of SVD on health through increased risk of stroke and vascular dementia, the pathophysiological mechanisms underlying SVD remain largely unknown. Pathological findings in diseased vessels include lipohyalinosis and microatheroma, whereas in the parenchyma findings include myelin pallor, enlargement of perivascular spaces, and gliosis. Many of these findings are common...
to both WMH and lacunar stroke. However, patholog-
ical studies have been hampered by methodological and phenotypic inconsistencies. In addition, little is known about the extent to which underlying pathogenesis is shared across the radiological manifestations. WMH are increased in lacunar stroke, more so than in other pathological subtypes of stroke, which may indicate that shared pathological processes underlie the 2. In addition, both confluent WMH and lacunar infarcts are a common finding in Mendelian forms of SVD such as cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, although the underlying arterial pathology is considerably different to that of sporadic SVD. However, aside from these exceptions and shared cardiovascular risk factors such as hypertension, few molecular processes have been robustly shown to impact on both lacunar stroke and WMH.

Genetic studies can provide novel insights into SVD and the nature of the relationship between its manifestations. In particular, genome-wide association studies have recently identified multiple genetic variants associated with WMH in community-dwelling individuals and have shown that common variants in COL4A2, a gene associated with monogenic SVD, influence sporadic SVD. In addition, genome-wide association studies provide a means of interrogating the relationship between complex traits and assessing whether such traits share pathogenesis. Polygenic risk score approaches have previously been used to assess the influence of trait-associated single nucleotide polymorphisms (SNPs) and to test whether an association with lacunar stroke was present in individuals without substantial WMH, testing the influence of WMH-associated variants on these subgroups, as well as on cardioembolic and large vessel strokes. We first used a genetic risk score approach to evaluate the overall evidence that WMH-associated variants affect stroke phenotypes, and then second evaluated whether each of the specific genetic variants is associated with lacunar stroke in both the groups with and without WMH.

Materials and Methods

Study Participants

The study data set consisted of stroke cases obtained from hospital admissions in the UK and Germany (DNA-lacunar, Genes and Ischaemic Stroke [GENESIS] and Wellcome Trust Case Control Consortium 2 [WTCCC2] study), Australia (Australian Stroke Genetics Collaborative [ASGC]), Italy (Milano—Besta Stroke Register [BSR]), and Belgium (Leuven Stroke Study [LSS]), as well as 9053 controls consisting of ancestry-matched individuals from each of the respective case populations (Table). These data sets have been described in detail in previous publications. Genotyping and

Table. Association of SNPs Associated With WMH in Lacunar Stroke Phenotypes

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene Region</th>
<th>CHR</th>
<th>BP (hg19)</th>
<th>RA</th>
<th>Lacunar Stroke</th>
<th>Lacunar Stroke w/ WMH</th>
<th>Lacunar Stroke w/out WMH</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7214628</td>
<td>TRIM65</td>
<td>17</td>
<td>73882148</td>
<td>G</td>
<td>1.02 (0.90–1.15); 0.78</td>
<td>1.08 (0.92–1.28); 0.33</td>
<td>0.94 (0.80–1.11); 0.47</td>
</tr>
<tr>
<td>rs72489890</td>
<td>NEURL</td>
<td>10</td>
<td>105319409</td>
<td>G</td>
<td>1.13 (0.99–1.29); 0.08</td>
<td>1.15 (0.97–1.38); 0.12</td>
<td>1.10 (0.92–1.30); 0.30</td>
</tr>
<tr>
<td>rs7984407</td>
<td>PDCD11</td>
<td>10</td>
<td>10571679</td>
<td>T</td>
<td>1.04 (0.94–1.15); 0.49</td>
<td>1.03 (0.90–1.18); 0.65</td>
<td>1.01 (0.88–1.15); 0.89</td>
</tr>
<tr>
<td>rs12357919</td>
<td>SH3PXD2A</td>
<td>10</td>
<td>105438112</td>
<td>T</td>
<td>1.13 (0.99–1.28); 0.07</td>
<td>1.13 (0.95–1.35); 0.15</td>
<td>1.12 (0.94–1.33); 0.19</td>
</tr>
<tr>
<td>rs7909791</td>
<td>SH3PXD2A</td>
<td>10</td>
<td>105613178</td>
<td>A</td>
<td>1.14 (1.03–1.27); 0.01*</td>
<td>1.15 (1.01–1.32); 0.04*</td>
<td>1.19 (1.04–1.36); 0.01*</td>
</tr>
<tr>
<td>rs7857879</td>
<td>EFEMP1</td>
<td>2</td>
<td>56135099</td>
<td>A</td>
<td>0.98 (0.82–1.15); 0.77</td>
<td>1.03 (0.83–1.28); 0.77</td>
<td>0.90 (0.72–1.13); 0.37</td>
</tr>
<tr>
<td>rs2984613</td>
<td>PMF1-BGLAP</td>
<td>1</td>
<td>156197380</td>
<td>C</td>
<td>1.10 (0.99–1.22); 0.07</td>
<td>1.13 (0.99–1.30); 0.07</td>
<td>1.06 (0.93–1.20); 0.41</td>
</tr>
<tr>
<td>rs11679640</td>
<td>HAOO</td>
<td></td>
<td>43141485</td>
<td>C</td>
<td>0.91 (0.81–1.03); 0.15</td>
<td>0.89 (0.76–1.04); 0.13</td>
<td>0.95 (0.81–1.12); 0.55</td>
</tr>
<tr>
<td>rs72934505</td>
<td>NBEAL1</td>
<td>2</td>
<td>203916487</td>
<td>T</td>
<td>1.24 (1.07–1.45); 0.004*</td>
<td>1.29 (1.05–1.59); 0.01*</td>
<td>1.19 (0.98–1.45); 0.08</td>
</tr>
<tr>
<td>rs17148926</td>
<td>LOC10050584</td>
<td>5</td>
<td>121510586</td>
<td>C</td>
<td>0.87 (0.76–1.00); 0.05</td>
<td>0.87 (0.73–1.04); 0.14</td>
<td>0.88 (0.74–1.05); 0.17</td>
</tr>
<tr>
<td>rs941898</td>
<td>EVL</td>
<td>14</td>
<td>100599437</td>
<td>G</td>
<td>1.12 (1.01–1.25); 0.04*</td>
<td>1.09 (0.95–1.26); 0.22</td>
<td>1.14 (0.99–1.30); 0.07</td>
</tr>
<tr>
<td>rs6942756</td>
<td>AHCYL2</td>
<td>7</td>
<td>128886821</td>
<td>G</td>
<td>1.07 (0.96–1.20); 0.22</td>
<td>1.05 (0.91–1.22); 0.50</td>
<td>1.11 (0.96–1.28); 0.15</td>
</tr>
<tr>
<td>rs2883428</td>
<td>XM_0039600</td>
<td>1</td>
<td>239571364</td>
<td>G</td>
<td>1.00 (0.89–1.11); 0.96</td>
<td>1.04 (0.90–1.20); 0.62</td>
<td>0.98 (0.85–1.14); 0.83</td>
</tr>
<tr>
<td>rs962888</td>
<td>C10L1</td>
<td>13</td>
<td>43059071</td>
<td>G</td>
<td>1.01 (0.94–1.17); 0.44</td>
<td>1.09 (0.94–1.26); 0.24</td>
<td>1.00 (0.87–1.15); 0.99</td>
</tr>
<tr>
<td>rs9515201</td>
<td>COL4A2</td>
<td>13</td>
<td>111040798</td>
<td>A</td>
<td>1.15 (1.03–1.27); 0.01*</td>
<td>1.13 (0.99–1.30); 0.07</td>
<td>1.14 (1.00–1.31); 0.06</td>
</tr>
</tbody>
</table>

Note: odds ratio is oriented to reflect direction of association with WMH (ie, the reference allele is the risk allele of WMH); BP indicates base position; CHR, chromosomes; CI, confidence interval; hg19, human genome reference 19; OR, odds ratio; RA, reference allele; SNP, single nucleotide polymorphism; and WMH, white matter hyperintensities.

*P<0.05 and in same direction of effect as WMH genome-wide association studies.
imputation of the individuals are described in the online-only Data Supplement. Briefly, all data sets were genotyped on commercially available Illumina arrays and imputed to 1000 Genomes phase 3 using SHAPEIT v2 (for phasing) and IMPUTE v2.2.2 (for imputation).

**Phenotype Classification**

Subtyping of the ASGC, WTCCC2, GENESIS, BSR, and LSS groups was initially performed using the Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification,19 using clinical assessment as well as brain and vascular imaging where available. For this analysis, we considered only cardioembolic, large vessel, and small vessel subtypes. For the cases that were classified as small vessel stroke under TOAST and had accompanying MRI, as well as all cases from UK Young Lacunar Stroke Study (DNA-lacunar) which included only cases with the TOAST SVD subtype, we performed further characterization, as follows. All MRI scans were centrally reviewed by 1 physician (H.S.M.). The diagnosis of lacunar stroke was confirmed using the following criteria. Lacunar stroke was defined as a clinical lacunar syndrome,20 with an anatomically compatible lesion on MRI (subcortical infarct ≤15 mm in diameter). For MRIs performed in the acute phase, the diagnosis was made by an acute lacunar infarct on DWI. For scans not performed in the acute phase, the diagnosis was made by a lacunar syndrome in combination with a lacunar infarct visualized on T1 and fluid-attenuated inversion recovery as a cavitated lesion in an anatomically appropriate location. Exclusion criteria were as follows: stenosis >50% in the extra- or intracranial cerebral vessels; cardioembolic source of stroke, defined according to the TOAST criteria as high or moderate probability; subcortical infarct >15 mm in diameter, as these can be caused by embolic mechanisms (striatocapsular infarcts); any other specific cause of stroke (eg, lupus anticoagulant, cerebral vasculitis, dissection, monogenic forms of stroke, such as cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy). For each individual with a lacunar stroke, we characterized the degree of WMH using the semiquantitative Fazekas scale,21 which classifies individuals into 4 groups ranging from none (0) to severe (3). Based on this grading, we then divided the lacunar stroke cases into those with and without WMH: (1) no WMH—patients with only mild or absent leukoaraiosis (Fazekas grade 0 or 1) and (2) WMH—patients with moderate or severe leukoaraiosis (Fazekas grade ≥2).

**Genetic Risk Score Analyses**

For each of the 18 SNPs associated with WMH in community populations in a recent study (8 genome-wide and 10 with P<1×10−6 in Europeans or overall),3 we generated an unweighted risk score for each individual in our data set by counting the number of risk alleles and summing across all SNPs. We used an unweighted approach, rather than an approach weighted on the log of the odds ratio (OR), as effect sizes were not reported for the published associations in a recent study (8 genome-wide and 10 with P<1×10−6 in Europeans or overall).3 Three of the SNPs (rs186314186, rs150695384, and rs17126031) were rare and not well imputed in our data set so were not included. Within each data set, we then converted each individual’s risk score to a Z-score using the standard transformation. We then used logistic regression to estimate the influence of the risk score on each stroke outcome, including ancestry-informative principal components to control for population stratification and meta-analyzing the results using a fixed-effects inverse-variance weighted approach. We tested for association of the genetic risk score with lacunar, cardioembolic, and large vessel stroke. To investigate whether an association with lacunar stroke was independent of WMH, we performed the same analysis on 2 subgroups of lacunar stroke stratified on the presence of substantial WMH (Fazekas ≥2). As the purpose of this analysis was to identify whether the observed association with lacunar stroke was independent of WMH, we did not perform the same analyses in the other subtypes. All ORs reported are per 1 SD change in the normally distributed risk score. We set the criteria for statistical significance at \( P<0.01 \), Bonferroni-correcting for the 5 tests.

**Single SNP Analyses**

In addition, we tested the association of each of the 15 available SNPs with lacunar stroke and the 2 subgroups based on the presence or absence of WMH. For each SNP we performed analyses separately in the 3 batches, including the first 10 ancestry-informative principal components in each analysis. We meta-analyzed the results using a fixed-effects inverse-variance weighted approach. We set the significance threshold at \( P<0.0011 \), correcting for the 15 SNPs in each of the 3 phenotypes (45 tests in total). All analyses were performed using the R statistical software.

**Results**

**Cohort Characteristics**

The final cohort consisted of 4176 stroke cases, including 1,373 lacunar stroke cases (mean age [SD]=60.0 [11.3] years; 68.0% men), subtyped into 568 with WMH (mean age [SD]=65.1 [10.9] years; 65.0% men) and 787 without WMH (mean age [SD]=56.0 [9.9] years; 70.2% men), 1,331 cardioembolic strokes (mean age [SD]=72.8 [10.6] years; 52.7% men), and 1,472 large vessel strokes (mean age [SD]=66.9 [11.1] years; 67.3% men) and 9,053 controls (mean age [SD]=58.4 [10.7] years [age not available in 2437 of WTCCC2-UK controls]; 52.2% men). Information on WMH volumes was not available in controls. As inclusion in the MRI-informed lacunar stroke analysis depended on the availability of an MRI and confirmation of a lacunar infarct, proportions of lacunar stroke cases varied greatly between studies (Figure 1).

**Genetic Risk Score Analyses**

A genetic risk score comprises 15 SNPs associated with WMH in community populations was significantly associated with lacunar stroke (OR [95% confidence interval (CI)]=1.14 [1.06–1.22]; \( P=0.0003 \); Figure 2). The association was slightly stronger, although not significantly so, in the group with WMH (OR [95% CI]=1.15 [1.05–1.26]; \( P=0.003 \)) and slightly weaker and not reaching Bonferroni-corrected significance in the group without substantial WMH (OR [95% CI]=1.11 [1.02–1.21]; \( P=0.019 \)). Conversely, the risk score was not associated with cardioembolic (OR [95% CI]=1.03 [0.97–1.09]; \( P=0.39 \)) or large vessel stroke (OR [95% CI]=0.99 [0.93–1.04]; \( P=0.63 \)).

**Single SNP Analyses**

No SNP reached the a priori significance threshold after Bonferroni correction (Table). Four SNPs (rs7909791[SH3PXD2A], OR [95% CI]=1.14 [1.03–1.27]; \( P=0.01 \); rs72934505[NBEAL1], OR [95% CI]=1.24 [1.07–1.45]; \( P=0.004 \); rs941898[EVL], OR [95% CI]=1.12 [1.01–1.25]; \( P=0.04 \); and rs9512020[COL4A2], OR [95% CI]=1.15 [1.03–1.27]; \( P=0.01 \)) reached a nominal significance threshold in the all lacunar stroke analysis.

**Discussion**

We used a genetic risk score approach to determine whether genetic variants associated with WMH confer risk of lacunar stroke, and therefore whether WMH and lacunar stroke share pathogenesis. We found strong evidence that genetic variants associated with WMH in community populations also influence risk of lacunar stroke. This provides further evidence to support the long-held view that neuroimaging features
When dividing our lacunar stroke population into those with moderate to severe WMH and those without, we found some evidence for association with both groups, although the association was marginally (and not significantly) stronger in the group with WMH and the association in the group without WMH did not reach Bonferroni-corrected significance. This suggests that variants influencing WMH confer risk of lacunar stroke even for lacunar strokes without substantial WMH. In addition, 2 of the SNPs, rs9515201 [COL4A2] and rs2984613 [PMF1-BGLAP], are also associated with intracerebral hemorrhage.10,22 This serves to emphasize that shared pathophysiological processes seem to underlie many of the clinical manifestations of cerebral SVD and suggests that a coordinated attempt to identify cerebral SVD associations will likely be fruitful. Four SNPs reached nominal significance for association with lacunar stroke (rs7909791 [SH3PXD2A], rs72934505 [NBEAL1], rs941898 [EVL], and rs9515201 [COL4A2]). With the exception of COL4A2, which has been linked to SVD, none of these loci have formerly been linked to ischemic stroke. As discussed above, similar arterial changes have been described in patients with lacunar stroke or WMH,5 including diffuse arteriosclerosis and a more focal microatheroma. Other studies have shown that mechanisms including blood–brain barrier dysfunction,22–24 and endothelial dysfunction,25 are important in both.23 As our results show a shared molecular basis to the 2 traits, they might suggest that these findings are because of the fact that WMH and lacunar stroke are outward manifestations of a shared underlying pathological process, namely, cerebral SVD.

In contrast, a genetic risk score comprising the same 15 SNPs was not associated with large vessel or cardioembolic stroke. Some studies have shown a relationship between subclinical atherosclerosis and WMH,23,24 whereas others have found an increased risk of all stroke in individuals with WMH.25 Our results, in a well-characterized population, suggest that the relationship between WMH and ischemic stroke is limited to lacunar stroke. This finding might suggest that previously reported associations between nonlacunar strokes and WMH may be because of shared risk factors such as hypertension rather than shared pathogenesis.

This study has several strengths. The sample size was large and all lacunar strokes were confirmed by MRI, reducing the possibility of misclassification which might occur when using CT. In addition, the design of the study, which made use of genetic data, means that the results are less susceptible to the residual confounding and reverse causation that observational studies can suffer from, although other sources of confounding, such as technical artifacts, may arise. Similarly, this study has weaknesses. We were unable to evaluate 3 rare SNPs which were associated with WMH in the previous publication,9 which may have affected our results. Some lacunar infarcts were diagnosed as acute lesions on DWI, but others were diagnosed from MRI scans performed after the acute

---

**Figure 1.** Cohort characteristics. ASGC indicates Australian Stroke Genetics Collaborative; CE, cardioembolic stroke; GENESIS, Genes and Ischaemic Stroke; LSS, Leuven Stroke Study; LVD, large vessel disease; MRI, magnetic resonance imaging; SVD, small vessel disease; WMH, white matter hyperintensities; and WTCCC2, Wellcome Trust Case Control Consortium 2.
stroke phase as cavities on T1 or fluid-attenuated inversion recovery images. The inclusion of patients defined using these different radiological criteria may introduce a subtle bias, and it is possible that some of these cavities could have resulted from hemorrhage rather than ischemia. Although MRI was performed for all lacunar stroke subjects, it was not performed in all nonlacunar subtypes, which could lead to some degree of misclassification. In addition, controls were historical and MRI was not performed to rule out cerebrovascular disease.

Conclusions

Genetic factors that affect WMH are also associated with risk of lacunar stroke, but not other stroke subtypes. This sheds light on processes that are implicated in lacunar stroke, but not other stroke subtypes. This sheds new light on processes that are implicated in lacunar stroke and provides further evidence that shared pathophysiological processes underlie different manifestations of SVD.

Sources of Funding

H. Markus was supported by an National Institute for Health Research (NIHR) Senior Investigator award. H. Markus and Dr Bevan were supported by the NIHR Cambridge University Hospitals Comprehensive Biomedical Research Centre. Collection of the UK Young Lacunar Stroke Resource was primarily supported by the Wellcome Trust (WT072952) with additional support from the Stroke Association (TSA 2010/01). Genotyping and Dr Traylor were supported by a project grant from the Stroke Association (TSA 2013/01). The research was also supported by the NIHR Biomedical Research Centre based at Guy’s and St Thomas’ NHS Foundation Trust and King’s College London. Dr This was supported by a Flemish Fund of Scientific Research (FWO) Clinical Investigator Grant.

Disclosures

None.

References

White Matter Hyperintensities and Lacunar Stroke


Genetic Associations With White Matter Hyperintensities Confer Risk of Lacunar Stroke

Stroke. 2016;47:1174-1179; originally published online April 12, 2016;
doi: 10.1161/STROKEAHA.115.011625

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://stroke.ahajournals.org/content/47/5/1174
Free via Open Access

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2016/04/22/STROKEAHA.115.011625.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/
SUPPLEMENTAL MATERIAL

Genotyping and imputation

Cohorts genotyped on the Illumina 610K (Australian cases and controls, Italian cases), 660W (UK cases and controls, German cases and Italian controls) and 1M arrays (German controls) were treated as one batch. Quality control was performed on each dataset separately, as previously described. A consensus set of 381,428 SNPs was then identified that was consistent across the four populations and the four populations were merged for these SNPs. We then performed principal components analysis using EIGENSTRAT on an LD-pruned set of SNPs from the combined dataset removing any population outliers, defined as greater than 6 standard deviations from the mean on the first 5 principal components. 37 individuals were removed in total (29 Australian, 8 Italian; 31 cases, 6 controls). The remaining individuals were then imputed to 1000 Genomes phase 3: SHAPEIT v2 was used to phase the haplotypes and IMPUTE v2.2.2 was used to perform the imputation.

The DNA-lacunar dataset was genotyped on the Illumina HumanExomeCore array. This array contains both exome content (~250,000 SNPs) and common tag SNPs (~250,000 SNPs) found on conventional GWAS arrays. Post-imputation, this array provides comparable coverage of common SNPs to larger arrays (MAF>5%, r2≥0.8 compared to 86.6% for Illumina OmniExpress). SNPs were excluded with MAF<0.01, genotype missingness>3%, HWE p<1e-6 in controls, strand ambiguity (A/T or C/G) or evidence of differential missingness by case-control status (p<0.05). Individuals were excluded if they had missingness>3%, excess or reduced heterozygosity, showed evidence of relatedness with another individual (pi-hat>0.1875), or failed a “sex-check” in PLINK. EIGENSTRAT was used to remove non-caucasian individuals, and was then repeated to calculate ancestry-informative principal components. The remaining 269,691 autosomal SNPs and 2,603 individuals were then imputed to 1000 Genomes phase 3: SHAPEIT v2 was used to phase the haplotypes and IMPUTE v2.2.2 was used to perform the imputation.

The Leuven dataset (Leuven Stroke Study) was genotyped on the Illumina Omni 5M array. SNPs were excluded with MAF<0.01, genotype missingness>3%, HWE p<1e-6 in controls, strand ambiguity (A/T or C/G) or evidence of differential missingness by case-control status (p<0.05). Individuals were excluded if they had missingness>3%, excess or reduced heterozygosity, showed evidence of relatedness with another individual (pi-hat>0.1875), or failed a “sex-check” in PLINK. EIGENSTRAT was used to remove non-caucasian individuals. The remaining individuals were then imputed to 1000 Genomes phase 3: SHAPEIT v2 was used to phase the haplotypes and IMPUTE v2.2.2 was used to perform the imputation.
References


Appendix

UK Young Lacunar Stroke DNA Study collaborators

Study managers: Josie Monaghan; Alan Zanich, Samantha Febrey, Eithne Smith, Jenny Lennon, St George’s University of London

Database cleaning: Loes Rutten-Jacobs, University of Cambridge

Participating centres (number of enrolled patients per centre; local investigators):

Aberdeen Royal Infirmary, Aberdeen (12; Mary Macleod). Addenbrooke’s Hospital, Cambridge (54; Jean-Claude Baron, Elizabeth Warburton, Diana J Day, Julie White). Airedale General Hospital, Steeton (4; Samantha Mawer). Barnsley Hospital, Barnsley (3; Mohammad Albazzaz, Pravin Torane, Keith Elliott, Kay Hawley). Bart’s and the London, London (2; Patrick Gompertz). Basingstoke and North Hampshire Hospital, Basingstoke (13; Elio Giallombardo, Deborah Dellafera). Blackpool Victoria Hospital, Blackpool (11; Mark O'Donnell). Bradford Royal Infirmary, Bradford (1; Chris Patterson). Bristol Royal Infirmary, Bristol (8; Sarah Caine). Charing Cross Hospital, London (12; Pankaj Sharma). Cheltenham General and Gloucester Royal Hospitals, Cheltenham and Gloucester (10; Dipankar Dutta). Chesterfield Royal Hospital, Chesterfield (4; Sunil Punnoose, Mahmud Sajid). Countess of Chester Hospital, Chester (22; Kausik Chatterjee). Derriford Hospital, Plymouth (4; Azlisham Mohd Nor). Dorset County Hospital NHS Foundation Trust, Dorchester (6; Rob Williams). East Kent Hospitals University NHS Foundation Trust, Kent (22; Hardeep Baht, Guna Gunathamilan). Eastbourne District General Hospital, Eastbourne (4; Conad Athulathmudali). Frenchay Hospital, Bristol (1; Neil Baldwin). Frimley Park Hospital NHS Foundation Trust, Frimley (6; Brian Clarke). Guy’s and St Thomas’ Hospital, London (14; Tony Rudd). Institute of Neurology, London (25; Martin Brown). James Paget University Hospital, Great Yarmouth (1; Peter Harrison). King's College Hospital, London (16; Lalit Kalra). Leeds Teaching Hospitals NHS Trust, London (125; Ahamad Hassan). Leicester General Hospital and Royal Infirmary, Leicester (9; Tom Robinson, Amit Mistri). Luton and Dunstable NHS University Hospital, Luton (16; Lakshmanan Sekaran, Sakthivel Sethuraman, Frances Justin). Maidstone and Tunbridge Wells NHS Trust (3; Peter Maskell). Mayday University Hospital, Croydon (14; Enas Lawrence). Medway Maritime Hospital, Gillingham (5; Sam Sanmuganathan). Milton Keynes Hospital, Milton Keynes (1; Yaw Duodu). Musgrove Park Hospital, Taunton (9; Malik Hussain). Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne (12; Gary Ford). Ninewells Hospital, Dundee (5; Ronald MacWalter). North Devon District Hospital, Barnstaple (8; Mervyn Dent). Nottingham University Hospitals, Nottingham (17; Philip Bath, Fiona Hammonds). Perth Royal Infirmary, Perth (2; Stuart Johnston). Peterborough City Hospital, Peterborough (1; Peter Owusu-Agyei). Queen Elizabeth Hospital, Gateshead (5; Tim Cassidy, Maria Bokhari). Radcliffe Infirmary, Oxford (5; Peter Rothwell). Rochdale Infirmary, Rochdale (4; Robert Namushi). Rotherham General Hospital, Rotherham (1; James Okwera). Royal Cornwall Hospitals NHS Trust, Truro (11; Frances Harrington, Gillian Courtauld). Royal Devon and Exeter Hospital, Exeter (22; Martin James). Royal Hallamshire Hospital, Sheffield (1; Graham Venables). Royal Liverpool University Hospital and Broadgreen Hospital, Liverpool (9; Aravind Manoj). Royal Preston Hospital, Preston (18; Shuja Punekar). Royal Surrey County Hospital, Guildford (23; Adrian Blight, Kath Pasco). Royal Sussex County Hospital, Brighton (14; Chakravarthi Rajkumar, Joanna Breeds). Royal United Hospital, Bath (6; Louise Shaw, Barbara Madigan). Salford Royal Hospital, Salford (16; Jane Molloy). Southampton General Hospital, Southampton (1; Giles Durward). Southend Hospital, Westcliff-on-Sea (26; Paul Guyler). Southern General Hospital, Glasgow (34; Keith Muir, Wilma Smith). St George’s Hospital, London (108; Hugh Markus). St Helier Hospital, Carshalton (10; Val Jones). Stepping Hill Hospital, Stockport (4; Shivakumar Krishnamoorthy). Sunderland Royal Hospital, Sunderland (1; Nikhil Majumdar). The Royal Bournemouth Hospital, Bournemouth (15; Damian Jenkinson). The Walton Centre, Liverpool (15; Richard White). Torbay Hospital, Torquay (19; Debs Kelly). University Hospital Aintree, Liverpool (19; Ramesh Durairaj). University Hospital of North Staffordshire, Stoke-on-trent (16; David Wilcock). Wansbeck General Hospital and North Tyneside Hospital, Ashington and North Shields (6; Christopher Price). West Cumberland
Hospital, Whitehaven (6; Olu Orugun, Rachel Glover). West Hertfordshire Hospital, Watford (20; David Collas). Western General Hospital, Edinburgh (12; Cathie Sudlow). Western Infirmary, Glasgow (33; Kennedy R. Lees, Jesse Dawson). Wycombe Hospital and Stoke Mandeville, High Wycombe (20; Dennis Briley and Matthew Burn). Yeovil District Hospital, Yeovil (46; Khalid Rashed). York Teaching Hospital, York (1; John Coyle).