Edinburgh Research Explorer

Ischemic stroke is associated with the ABO locus: the EuroCLOT study.

Citation for published version:
Williams, FMK, Carter, AM, Hysi, PG, Surdulescu, G, Hodgkiss, D, Soranzo, N, Traylor, M, Bevan, S, Dichgans, M, Rothwell, PM, Sudlow, C, Farrall, M, Silander, K, Kaunisto, M, Wagner, P, Saarela, O, Kuhlasmää, K, Virtamo, J, Salomaa, V, Amouyel, P, Arveiler, D, Ferrieres, J, Wiklund, P, Arfan Ikram, M, Hofman, A, Boncoraglio, GB, Parati, EA, Helgadottir, A, Gretarsdottir, S, Thorsteinsdottir, U, Thorleifsson, G, Stefansson, K, Seshadri, S, Destefano, A, Gschwendtner, A, Psaty, B, Longstreth, W, Mitchell, BD, Cheng, Y, Clarke, R, Ferrario, M, Bis, JC, Levi, C, Attia, J, Holliday, EG, Scott, RJ, Fornage, M, Sharma, P, Furie, KL, Rosand, J, Nalls, M, Meschia, J, Mosely, TH, Evans, A, Palotie, A, Markus, HS, Grant, PJ & Spector, TD 2013, 'Ischemic stroke is associated with the ABO locus: the EuroCLOT study.', Annals of Neurology, vol. 73, no. 1, pp. 16-31. https://doi.org/10.1002/ana.23838

Digital Object Identifier (DOI):
10.1002/ana.23838

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Annals of Neurology

Publisher Rights Statement:
Copyright © 2012 American Neurological Association
Re-use of this article is permitted in accordance with the Creative Commons Deed, Attribution 2.5, which does not permit commercial exploitation.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 17. Sep. 2023
Ischemic Stroke Is Associated with the ABO Locus: The EuroCLOT Study

Frances M. K. Williams, PhD, FRCP(E), 1 Angela M. Carter, PhD, 2 Pirro G. Hysi, PhD, 1 Gabriela Surdulescu, MSc, 1 Dylan Hodgkiss, MSc, 1 Nicole Soranzo, PhD, 3,4,5 Matthew Traylor, MSc, 6 Steve Bevan, PhD, 6 Martin Dichgans, MD, PhD, 7 Peter M. W. Rothwell, MD, PhD, FMedSci, 8 Cathie Sudlow, DPhil, FRCP(E), 9 Martin Farrall, PhD, FRCPath, 10 Kaisa Silander, PhD, 11 Mari Kaunisto, PhD, 11 Peter Wagner, BSc, 11 Olli Saarela, PhD, 12 Kari Kuulasmaa, PhD, 12 Jarmo Virtamo, MD, PhD, 12 Veikko Salomaa, MD, PhD, 12 Philippe Amouyel, MD, PhD, 13 Dominique Arveiler, MD, 14 Jean Ferrieres, MD, 15 Per-Gunnar Wiklund, MD, PhD, 16 M. Arfan Ikram, MD, PhD, 17 Albert Hofman, MD, PhD, 17 Giorgio B. Boncoraglio, MD, 18 Eugenio A. Parati, MD, 18 Anna Helgadottir, MD, PhD, 19 Solveig Gretarsdottir, PhD, 19 Unnur Thorsteinsdottir, PhD, 19 Gudmar Thorleifsson, PhD, 19 Kari Stefansson, PhD, Dr med, 19

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.23838

Received Apr 27, 2012, and in revised form Sep 21, 2012. Accepted for publication Oct 29, 2012.

Address correspondence to Dr Williams, Department of Twin Research and Genetic Epidemiology, King’s College London, St Thomas’ Hospital Campus, 3rd Floor South Wing Block 8, Westminster Bridge Road, London, SE1 7EH. E-mail: frances.williams@kcl.ac.uk

Additional Supporting Information may be found in the online version of this article.

Wellcome Trust Case Control Consortium 2 members are listed in the Appendix on page 29.
Objective: End-stage coagulation and the structure/function of fibrin are implicated in the pathogenesis of ischemic stroke. We explored whether genetic variants associated with end-stage coagulation in healthy volunteers account for the genetic predisposition to ischemic stroke and examined their influence on stroke subtype.

Methods: Common genetic variants identified through genome-wide association studies of coagulation factors and fibrin structure/function in healthy twins (n = 2,100, Stage 1) were examined in ischemic stroke (n = 4,200 cases) using 2 independent samples of European ancestry (Stage 2). A third clinical collection having stroke subtyping (total 8,900 cases, 55,000 controls) was used for replication (Stage 3).

Results: Stage 1 identified 524 single nucleotide polymorphisms (SNPs) from 23 linkage disequilibrium blocks having significant association (p < 5 x 10^{-8}) with 1 or more coagulation/fibrin phenotypes. The most striking associations included SNP rs5985 with factor XIII activity (p = 2.6 x 10^{-166}), rs10665 with FVII (p = 2.4 x 10^{-47}), and rs505922 in the ABO gene with both von Willebrand factor (p = 4.7 x 10^{-57}) and factor VIII (p = 1.2 x 10^{-36}). In Stage 2, the 23 independent SNPs were examined in stroke cases/noncases using MOnica Risk, Genetics, Archiving and Monograph (MORGAM) and Wellcome Trust Case Control Consortium 2 collections. SNP rs505922 was nominally associated with ischemic stroke (odds ratio = 0.94, 95% confidence interval = 0.88-0.99, p = 0.023). Independent replication in Meta-Stroke confirmed the rs505922 association with stroke, beta (standard error, SE) = 0.066 (0.02), p = 0.001, a finding specific to large-vessel and cardioembolic stroke (p = 0.001 and p = < 0.001, respectively) but not seen with small-vessel stroke (p = 0.811).

Interpretation: ABO gene variants are associated with large-vessel and cardioembolic stroke but not small-vessel disease. This work sheds light on the different pathogenic mechanisms underpinning stroke subtype.
whether such associations differed between stroke subtypes.

**Subjects and Methods**

We used a 3-stage study design to identify common variants influencing coagulation and fibrin structure/function in the normal population and then tested genome-wide significant independent single nucleotide polymorphisms (SNPs) for association with stroke in subjects of Northern European extraction (Fig 1). To study the broad range of hemostatic variables contributing to end-stage coagulation, GWA studies of fibrin structure/function ex vivo, fibrin turnover (D-dimer) in vivo, and individual hemostatic components were performed in a healthy volunteer cohort of twins (Stage 1). In Stage 2, those variants found to be independently associated with coagulation or fibrin structure/function were assessed as risk factors for ischemic stroke in cases and controls. In Stage 3, the top 4 SNPs from the meta-analysis of ischemic stroke were examined for replication in a third clinical collection of stroke having information on whether stroke resulted from occlusion of large-vessel, small-vessel, or cardiac emboli. Detailed methods are provided below. Written informed consent was obtained from participants in the study, and each individual study group obtained local ethics approval.

**Phenotyping the Cohorts**

TwinsUK. The subjects were obtained from the TwinsUK (TUK) registry (www.twinsuk.ac.uk) at King’s College London, United Kingdom, which has been ascertained by a national media campaign. For historical reasons, the majority of twin volunteers are female. TUK subjects have been shown to be representative of the wider general populations for genetic and lifestyle factors associated with a variety of traits. TUK subjects were phenotyped for fibrin structure/function, D-dimer, and hemostatic factors, according to methods described in detail elsewhere. In brief, fibrin structure/function was assessed using a turbidimetric assay, whereas D-dimer (as a measure of in vivo fibrin turnover), coagulation factors (F V, VIII, FXI, FXII A and B subunits (FXIII A, FXIII B), prothrombin, and von Willebrand Factor (vWF) were quantified by enzyme-linked immunosorbent assay, and fibrinogen, FVII, and FXIII by functional activity assays.

The MOnica Risk, Genetics, Archiving and Monograph (MORGAM) Cohort. The cohorts of the MORGAM project consist of the respondents of representative adult population samples. This study includes cohorts from a variety of centers, including Finland (FINRISK, ATBC), France (Lille, Strasbourg, Toulouse), Italy (Brienza), Northern Sweden, and Northern Ireland (Belfast) as described at http://www.ktl.fi/publications/morgam/cohorts. The participants were examined and DNA was collected at baseline, and they were followed up for stroke and acute coronary events. Genotyping was carried out in a case–cohort setting. In MORGAM cohorts, the end-point used was the subject presenting with first ischemic stroke. For some events the diagnosis was based on validation, and for some on the clinical or death certificate diagnosis (International Classification of Diseases [ICD]-9 codes 433 or 434, or ICD-10 code I63).
Health and Welfare in Finland. Several sample- and plate-specific quality control measures were implemented to minimize errors, and in addition genotyping quality was assessed from 5% blind duplicate samples in each 96-well plate. For 234 samples with low DNA yield, DNA was amplified and genotyped as previously described. Genotyping was performed using the MassARRAY System and iPLEX Gold chemistry (Sequenom, San Diego, CA) with standard protocol. Genotype clusters were manually reviewed using Typer 4.0 software (Sequenom), and genotype calls were corrected where necessary. Genotyping success rate was >95% for all but 1 SNP (rs2022309, 91.3%), with an average success rate of 95.7%. No discrepancies were identified among a total of 1,256 successful blind duplicate genotype pairs. Cox regression analysis adapted for the case-cohort data was used to assess the association between the genotypes and ischemic stroke in the MORGAM cohorts, assuming an additive genotypic effect. The analysis was stratified by cohort and sex.

WTCCC2. Stroke cases were genotyped using the Illumina 660Q platform. Shared WTCCC2 controls were genotyped using the Illumina 1M Duo platform. German controls were genotyped using the Illumina 550 platform. Analysis of the UK and German cohorts was performed independently using PLINK after quality control checking using a genotyping call rate of 98%, Hardy–Weinberg equilibrium call rate of 1e-20, and checks for individual relatedness and population stratification. The UK and German cases were then meta-analyzed using METAL. Samples were identified and removed if the genome-wide patterns of diversity differed from those of the collection at large, interpreting them as likely to be due to biases or artifacts. To do so, we used a Bayesian clustering approach to infer outlying individuals on the basis of call rate, heterozygosity, ancestry, and average probe intensity. We used a hidden Markov model to infer identity by descent along the genome and removed individuals iteratively to obtain a set with pair-wise identity by descent <5%. Samples were also removed if their inferred gender was discordant with the recorded gender or if <90% of the SNPs typed by Sequenom (Sequenom iPLEX assay for 4 gender SNPs) were concordant with the genome-wide data. For the EuroCLOT study, individual UK and German cohort and meta-analysis results were examined for the 23 available genotypes. This was performed for the phenotype of all ischemic stroke, together with the ischemic stroke subtypes of small-vessel disease, large-vessel disease, and cardioembolic stroke.

MetaStroke. Genotyping of the 13 MetaStroke contributors was performed independently by each group, using either Illumina or Affymetrix (Santa Clara, CA) platforms. Further details on cases and controls, genotyping, and imputation are available in Supplementary Table 3.

**Statistical Analysis**

Stage 1. We used multiple linear regression models to assess association between genotypes and phenotypes, using age as a covariate. The phenotypes examined in the TUK cohort were inverse-normal transformed to satisfy the assumption of normality of trait distribution of the linear models. Association
analysis was carried out using Merlin to control for family structure within the dataset. Independence of the effects conferred by SNPs in the same region was assessed by means of a backward stepwise regression analysis on the trait with which they were associated. This yielded 23 statistically independent significant SNPs (\( p < 5 \times 10^{-28} \)), associated with at least 1 quantitative outcome, which were taken forward for examination in the clinical groups at Stage 2. This stage of the analysis was performed using Stata for Windows version 10 (StataCorp, College Station, TX) with adjustment for the twins' relatedness.

Stage 2. The 23 independent SNPs remaining significant after multiple regression were carried through to investigation of association with ischemic stroke in MORGAM and WTCCC2. Results for each were meta-analyzed using a fixed effects inverse variance weighting implemented in METAL.

Stage 3. The 4 most significantly associated SNPs from Stage 2 were tested for association with overall ischemic stroke in Meta-Stroke. This international collaboration brings together GWA studies in ischemic stroke and (depending on SNP) includes 8,900 cases of ischemic stroke and 55,000 controls. In addition, subgroup analysis was possible (in Meta-Stroke and WTCCC2), as stroke events had been subphenotyped into large-vessel, small-vessel, and cardioembolic stroke by many of the contributing study groups, using the TOAST classification. Within Meta-Stroke, samples were excluded from analysis

### TABLE 1. Stage 1: Characteristics of the TwinsUK Discovery Sample

| Trait               | No. | Mean | Min | Max | SD  |
|---------------------|-----|------|-----|-----|-----|
| Age, yr             | 2,128 | 50.4 | 17.3 | 80.1 | 12.72 |
| BMI, kg/m²          | 2,124 | 25.5 | 15.2 | 52.4 | 4.7 |
| Female gender, %    | 95.6 | —    | —   | —   | —   |
| D-dimer, ng/ml      | 1,362 | 104.9 | 9   | 3,740 | 151.3 |
| FX antigen, U/ml    | 2,017 | 0.967 | 0.28 | 1.82 | 0.19 |
| FVII:C, %           | 1,403 | 107.6 | 30  | 207  | 24.4 |
| FVII antigen, IU/ml | 1,082 | 100.1 | 36  | 190  | 26.0 |
| FXIII antigen, IU/ml| 1,940 | 0.9    | 0.3 | 7.1  | 0.3 |
| FXIII activity, %   | 2,018 | 108.2 | 33  | 258  | 31.4 |
| FXIIIA, U/ml        | 2,017 | 1.0    | 0   | 3.2  | 0.3 |
| FXIIIB, U/ml        | 2,018 | 1.01   | 0.2 | 3.07 | 0.25 |
| vWF, IU/ml          | 2,006 | 1.1    | 0.21 | 3.32 | 0.4 |
| Fibrin lag time, s  | 2,484 | 81.00   | 2,555.0 | 347.0 | 8,996 |

Only those phenotypes found significant in the Stage 1 genome-wide association study are shown here (all results from Stage 1 are given in the Supplementary Material).

BMI = body mass index; Fibrin lag time = lag phase of the turbidimetric clotting assay; FVII:C = factor VII clotting activity; FXIIIA = FXIII A subunit; FXIIIB = FXIII B subunit; Max = maximum value; Min = minimum value; No. = sample size; SD = standard deviation; vWF = von Willebrand Factor.

### TABLE 2. Stages 2 and 3: Characteristics of the Ischemic Stroke Collections

| Cohort | WTCCC2 | MORGAM | MetaStroke |
|--------|--------|--------|------------|
|        | Cases  | Controls | Cases | Noncases | Cases | Controls |
| No. (% females) | 3,548 (43.5) | 5,972 (49.4) | 715 (19.4) | 2,932 (14.7) | 8,884 | 55,254 |
| Age, mean yr (SD) | 70.5 (12.9) | NA | 60.4 (8.1) | 59.1 (8.3) | 67.1 (10.5) | 59.4 (9.9) |
| BMI, mean kg/m² (SD) | NA | NA | 27.9 (4.5) | 27.0 (4.1) | NA | NA |

Age and BMI refer to values at the baseline of the follow-up.

*MOnica Risk, Genetics, Archiving and Monograph (MORGAM) is a population cohort with prospective follow-up for cases. Genotyping for this study was performed in a case–cohort setting.*

BMI = body mass index; NA = not available; No. = sample size; SD = standard deviation; WTCCC2 = Wellcome Trust Case Control Consortium 2.
TABLE 3. Stage 1: Independent SNPs (n = 23) Found with $p < 5 \times 10^{-8}$ in TwinsUK

| Chr | SNP | Position | Gene | A1 | A2 | Freq A1 | Trait | Effect | SE  | $p$ |
|-----|-----|----------|------|----|----|---------|-------|--------|-----|-----|
| 6   | rs5985 | 6263794  | F13A1 | C  | A  | 0.741   | FXIII activity | $-1.077$ | 0.037 | $2.6 \times 10^{-16}$ |
| 6   | rs3024321 | 6263125 | F13A1 | A  | G  | 0.679   | FXIII activity | $-0.891$ | 0.035 | $8.4 \times 10^{-14}$ |
| 9   | rs505922 | 135139050 | ABO  | T  | C  | 0.68    | vWF      | $-0.561$ | 0.035 | $4.71 \times 10^{-5}$ |
| 9   | rs643434 | 135132176 | ABO  | G  | A  | 0.656   | vWF      | $-0.51$  | 0.034 | $1.1 \times 10^{-9}$  |
| 13  | rs10665 | 112800832 | F7/F10 | A  | G  | 0.878   | FVII:C     | $0.847$  | 0.059 | $2.37 \times 10^{-4}$  |
| 13  | rs2181540 | 112801165 | F7/F10 | T  | C  | 0.876   | FVII:C     | $0.834$  | 0.058 | $1.18 \times 10^{-4}$  |
| 13  | rs6041 | 112820708 | F7/F10 | G  | A  | 0.892   | FVII:C     | $0.877$  | 0.062 | $1.93 \times 10^{-4}$  |
| 1   | rs10922162 | 195337399 | CFHR4 | C  | T  | 0.839   | FXIII A | $0.521$  | 0.045 | $1.55 \times 10^{-30}$ |
| 1   | rs12116643 | 195239806 | CFHR4 | T  | C  | 0.84    | FXIII A | $0.515$  | 0.045 | $4.89 \times 10^{-30}$ |
| 9   | rs651007 | 195143696 | ABO  | C  | T  | 0.81    | FXII     | $-0.461$ | 0.044 | $1.94 \times 10^{-25}$ |
| 13  | rs3211770 | 112841850 | F7/F10 | G  | A  | 0.879   | FVII:C     | $0.534$  | 0.059 | $2.98 \times 10^{-19}$ |
| 1   | rs7410943 | 195421644 | CFHR4 | A  | G  | 0.531   | FXIII B | $0.301$  | 0.034 | $3.42 \times 10^{-19}$ |
| 1   | rs4915559 | 195153393 | CFHR4 | T  | C  | 0.759   | FXIII A | $0.339$  | 0.038 | $7.3 \times 10^{-19}$  |
| 6   | rs11243081 | 6282505 | F13A1 | C  | T  | 0.676   | FXIIIA    | $-0.299$ | 0.035 | $5.13 \times 10^{-18}$ |
| 6   | rs17142067 | 6286286 | F13A1 | A  | C  | 0.67    | FXIIIA    | $-0.299$ | 0.035 | $1.24 \times 10^{-17}$ |
| 9   | rs8176743 | 135121236 | ABO  | C  | T  | 0.939   | vWF      | $-0.582$ | 0.068 | $1.59 \times 10^{-17}$ |
| 6   | rs1318606 | 6276935 | F13A1 | T  | C  | 0.569   | FXIII A | $-0.273$ | 0.032 | $4.39 \times 10^{-17}$ |
| 13  | rs555212 | 112804541 | F7/F10 | G  | A  | 0.768   | FVII:C     | $-0.367$ | 0.046 | $1.07 \times 10^{-15}$ |
| 9   | rs10982156 | 116127885 | ORM1 | T  | A  | 0.939   | FX      | $0.456$  | 0.072 | $2.47 \times 10^{-10}$ |
| 1   | rs1410996 | 194963556 | CFHR4 | G  | A  | 0.585   | FXIII A | $0.203$  | 0.033 | $1.32 \times 10^{-9}$  |
| 5   | rs12518614 | 63777293 | F7/F10 | A  | G  | 0.954   | vWF      | $-0.463$ | 0.077 | $1.52 \times 10^{-9}$  |
| 13  | rs9533425 | 42668474 | ENOX1 | G  | C  | 0.801   | Fibrin lag time | $-0.308$ | 0.051 | $1.9 \times 10^{-9}$   |
| 1   | rs2022309 | 94825064 | F3   | G  | T  | 0.701   | D-dimer  | $-0.244$ | 0.045 | $4.32 \times 10^{-8}$  |

Of the 524 genome-wide significant associations identified in Stage 1, only independent SNPs are shown, and where associated with multiple traits, the most significant result is given (all Stage 1 results are listed in Supplementary Table 2). The effect size (Effect) and SE are expressed in terms of standard deviation for each phenotype. The probabilities for association ($p$) are from multivariate models using single SNP genotypes as independent variables and age as covariate. A1 = allele 1; A2 = allele 2; Chr = chromosome; Effect = beta of regression analysis; Fibrin lag time = lag phase of the turbidimetric clotting assay; Freq A1 = frequency of allele 1; FVII:C = factor VII clotting activity; FXIII A = FXIII A subunit; FXIII A subunit; FXIII B = single nucleotide polymorphism; vWF = von Willebrand Factor.

if they had call rates <80% or if reported gender was discordant with gender-specific markers. We removed pairs of samples showing concordance indicative of being duplicates. MetaStroke genotyping results were imputed to HapMap2 using MACH2. Where SNPs were imputed, $r^2$ values were >0.9. Four SNPs analyzed in these cohorts were meta-analyzed using a fixed effects model with the metan module in Stata version 10.

### Results

The characteristics of the 2,128 twin participants are shown for TUK in Table 1 and Supplementary Table 1 (Stage 1). The mean age of the twins was 50.4 years, and the sample included 87 (4.4%) males. All were of North European descent. The sample size varied between assays; for clarity, the number of subjects is included in the tables for each phenotype.

Details of the clinical collections of stroke cases and controls are shown in Table 2.

#### Stage 1

There were a number of strikingly strong genotype–phenotype associations identified in the TUK discovery group, and in total 524 associations were found having $p < 5 \times 10^{-8}$. The 524 SNPs identified as significant genome-wide were mostly associated with coagulation factor phenotypes; there was 1 association with lag time to
fibrin clot formation. After the interdependence of the SNPs had been established by backward stepwise regression analysis, 23 statistically independent SNPs were identified for examination in Stage 2 (shown in Table 3). The strongest signals were observed for SNP rs5985 in the F13A1 gene (encoding the FXIII A subunit) and FXIII activity ($p = 2.6 \times 10^{-186}$), followed by rs2731672 in the F12 gene associated with FXII concentration (encoding FXII; $p = 1.3 \times 10^{-115}$; Supplementary Table 2) and rs505922 in the ABO gene with vWF ($p = 4.7 \times 10^{-57}$; see Table 3) and factor VIII ($p = 1.2 \times 10^{-36}$; see Supplementary Table 2). Further coagulation-related phenotype–SNP associations were identified for rs10665 in F7/MCF2L and FVII clotting activity ($p = 2.4 \times 10^{-47}$), and rs2022309 in the F3 gene (encoding tissue factor) with D-dimer concentration ($p = 4.3 \times 10^{-8}$).

A clear relationship was found between plasma FXIII A subunit and SNP rs12137359 ($p = 1.0 \times 10^{-27}$) lying within the gene ZBTB41 (zinc finger and BTB domain containing 41, a highly conserved gene). However, this region on chromosome 1q is rich with candidate genes, and the SNP in question lies downstream of the CFH and CFHR1-5 genes (encoding complement factor H and CFH-related proteins 1 to 5) as well as F13B (encoding FXIII B subunit). There is also an association in this same region between rs800292 in the CFH gene and FXIII A concentration ($p = 1.5 \times 10^{-12}$).

**Stage 2**

In the MORGAM study, 6 of the 23 independent SNPs were available for lookup. None of the SNPs was significantly associated with ischemic stroke in this study group or in WTCCC2, although there was a suggestion of an effect for rs505922 in both MORGAM (T allele, beta = −0.126, $p = 0.067$) and WTCCC2 (T allele, beta = −0.054, $p = 0.097$). In the meta-analysis of WTCCC2 and MORGAM, SNP rs505922 in the ABO gene was associated with ischemic stroke (beta for T allele = −0.067, $p = 0.023$), with the major T allele being protective against stroke (Table 4).

**Stage 3**

We examined the association of the 4 ABO SNPs in the large Meta-Stroke dataset (Table 5, with genotyping details in Supplementary Table 3) and further explored their relationship with stroke subtype in Meta-Stroke and WTCCC2. Results for ischemic stroke overall are shown in Table 6 (positive results) and are illustrated by a forest plot (Fig 2). The results show an association for lead SNP rs505922 C allele with ischemic stroke (odds ratio [OR] = 1.07, 95% confidence interval [CI] = 1.03–1.11, $p = 0.006$). Two other ABO SNPs also showed significant association: rs643434 (for A allele, meta-analysis logistic regression OR = 1.06, 95% CI = 1.02–1.11, $p = 0.002$) and rs651007 (C allele, OR = 1.07, 95% CI = 1.02–1.12, $p = 0.007$; see Table 6). Analysis by stroke subtype for SNP rs505922 showed association with cardioembolic stroke (OR = 1.13, 95% CI = 1.11–1.15, $p \leq 0.001$), and large-vessel stroke (OR = 1.23, 95% CI = 1.07–1.18, $p = 0.001$), but there was no association with small-vessel disease ($p = 0.811$; Table 7).

Finally, to determine whether the genetic influence was acting through known risk factors, we performed subgroup analysis in the sample having this information, WTCCC2–Munich. Adjusting for hypertension,
hypercholesterolemia, diabetes, and smoking had a small effect on the strength of the association (unadjusted: beta = 0.159, 95% CI = 0.023–0.294, p = 0.022; adjusted: beta = 0.160, 95% CI = 0.010–0.309, p = 0.036).

**Discussion**

Ischemic stroke accounts for considerable morbidity and mortality in Western countries, and treatment is limited at present. Our 3-stage study design optimized power for discovery of common genetic variants predisposing to ischemic stroke and stroke subtype. We performed a GWA study of intermediate coagulation and fibrinolytic phenotypes in healthy volunteers to examine the genetic determinants of end-stage coagulation and went on to study their influence on stroke and stroke subtype. We identified a large number of genetic variants associated with measures of coagulation factors, both functional and antigenic, some of which have been included in GWA meta-analyses of coagulation.\(^\text{23}\) We confirmed that polymorphisms in the *ABO* gene were significantly associated with vWF and FVIII levels in healthy volunteers. Significant associations between SNPs in *ABO* and levels of vWF (rs505922, rs643434, rs8176743) and/or FVIII (rs505922, rs651007) were identified; we went on to demonstrate significant associations between *ABO* SNPs, in particular rs505922, and ischemic stroke (see forest plot in Fig 2).

The associations between FVIII levels and the *ABO* gene variant rs505922, and between *ABO* and coronary disease, suggest a possible mechanism behind the well-documented association between the ABO blood group and risk of vascular disorders. Non-O blood groups are at increased risk of stroke,\(^\text{24}\) peripheral vascular disease, and myocardial infarction (MI) but not coronary artery disease (as assessed by angina, summarized by Wu et al\(^\text{25}\)), and this suggests that end-stage coagulation is the critical determinant. The association we found with FVIII levels may account for this. Recent GWA studies of MI have identified variants within the *ABO* gene that predispose to MI,\(^\text{26,27}\) and this relationship appears to hold for common forms of thrombotic stroke; we found evidence of association in large-vessel and cardioembolic stroke, but there was no association with small-vessel disease. At present, none of the SNPs significantly associated with stroke is reported to be associated with known risk factors such as hypertension, hyperlipidemia,
### TABLE 6. Meta-Analysis of the ABO Locus in the MetaStroke Consortium: Stage 3

| SNP       | Study         | Cases Controls/Noncases | A1 | A2 | Freq A1 | Beta | SE   | p    |
|-----------|---------------|-------------------------|----|----|---------|------|------|------|
| rs651007  | ARIC          | 385 8,803 C T           | 0.773 | -0.04 | 0.085 | 0.636 |
|           | ASGC          | 1,177 1,195 C T         | 0.796 | -0.117 | 0.073 | 0.109 |
|           | BRAINS        | 394 444 C T             | 0.803 | 0.033  | 0.123 | 0.786 |
|           | CEDIR         | 372 407 C T             | 0.776 | -0.192 | 0.123 | 0.12  |
|           | CHS           | 454 2,817 C T           | 0.788 | -0.109 | 0.106 | 0.302 |
|           | DeCODE        | 2,393 26,982 C T        | 0.852 | -0.06  | 0.044 | 0.175 |
|           | FHS           | 171 4,335 C T           | 0.79  | 0.065  | 0.149 | 0.664 |
|           | GASROS        | 516 1,202 C T           | 0.686 | -0.065 | 0.082 | 0.429 |
|           | HPS           | 578 468 C T             | 0.791 | -0.058 | 0.107 | 0.588 |
|           | HVH           | 566 1,290 C T           | 0.776 | -0.028 | 0.119 | 0.814 |
|           | ISGS          | 1,070 1,488 C T         | 0.8   | -0.158 | 0.076 | 0.037 |
|           | Rotterdam     | 367 5,396 C T           | 0.794 | 0.038  | 0.094 | 0.168 |
|           | Meta-analysis | 8,443 54,827 C T        | 0.068 | 0.025  | 0.007 |
| rs643434  | ARIC          | 385 8,803 A G           | 0.368 | 0.186  | 0.073 | 0.011 |
|           | ASGC          | 1,177 1,195 A G         | 0.34  | 0.149  | 0.062 | 0.017 |
|           | BRAINS        | 394 444 A G             | 0.309 | -0.122 | 0.109 | 0.266 |
|           | CEDIR         | 372 407 A G             | 0.378 | 0.109  | 0.104 | 0.34  |
|           | CHS           | 454 2,817 A G           | 0.366 | 0.07   | 0.068 | 0.307 |
|           | DeCODE        | 2,393 26,982 A G        | 0.283 | 0.043  | 0.033 | 0.195 |
|           | FHS           | 171 4,335 A G           | 0.348 | 0.006  | 0.115 | 0.956 |
|           | HPS           | 578 468 A G             | 0.345 | 0.054  | 0.094 | 0.567 |
|           | HVH           | 566 1,290 A G           | 0.367 | 0.054  | 0.076 | 0.48  |
|           | ISGS          | 1,070 1,488 A G         | 0.339 | 0.047  | 0.063 | 0.458 |
|           | Rotterdam     | 367 5,396 A G           | 0.351 | 0.016  | 0.076 | 0.041 |
|           | Meta-analysis | 7,927 53,631 A G        | 0.062 | 0.02   | 0.002 |
| rs505922  | ARIC          | 385 8,803 C T           | 0.345 | 0.172  | 0.074 | 0.02  |
|           | ASGC          | 1,177 1,195 C T         | 0.317 | 0.176  | 0.063 | 0.005 |
|           | BRAINS        | 394 444 C T             | 0.308 | -0.108 | 0.108 | 0.318 |
|           | CEDIR         | 372 407 C T             | 0.351 | 0.122  | 0.105 | 0.248 |
|           | CHS           | 454 2,817 C T           | 0.342 | 0.059  | 0.068 | 0.391 |
|           | DeCODE        | 2,393 26,965 C T        | 0.74  | 0.052  | 0.034 | 0.129 |
|           | FHS           | 171 4,335 C T           | 0.325 | 0.04   | 0.115 | 0.727 |
|           | GASROS        | 516 1,202 C T           | 0.3   | 0.094  | 0.081 | 0.249 |
|           | HPS           | 578 468 C T             | 0.327 | 0.047  | 0.094 | 0.619 |
|           | HVH           | 566 1,290 C T           | 0.344 | 0.054  | 0.077 | 0.484 |
|           | ISGS          | 1,070 1,488 C T         | 0.329 | 0.039  | 0.061 | 0.528 |
|           | Rotterdam     | 367 5,396 C T           | 0.666 | 0.005  | 0.077 | 0.005 |
|           | Meta-analysis | 8,443 54,810 C T        | 0.391 | 0.066  | 0.02  | 0.001 |

Meta-analysis of ischemic stroke performed in MetaStroke cohorts in the ABO locus (those 3 SNPs having positive results shown). Results of the meta-analysis of the nonsignificant SNP rs8176743 are given in Supplementary Table 2. A1 = allele 1; A2 = allele 2; ARIC = Atherosclerosis Risk in Communities; ASGC = Australian Stroke Genetics Collaborative; Beta = effect size; BRAINS = Bio-Repository of DNA in Stroke; CHS = Cardiovascular Health Study; FHS = Framingham Heart Study; Freq = frequency; GASROS = Genes Affecting Stroke Risk and Outcome Study; HPS = Heart Protection Study; HVH = Heart and Vascular Health; ISGS = Ischemic Stroke Genetics Study; SE = standard error; SNP = single nucleotide polymorphism.
## TABLE 7. Meta-Analysis of Single Nucleotide Polymorphism rs505922 in the ABO Locus by Stroke Subtype (Stage 3)

| Center       | Cases | Controls/Noncases | A1 Freq | Beta | SE  | p    |
|--------------|-------|-------------------|---------|------|-----|------|
| **Cardioembolic** |       |                    |         |      |     |      |
| ARIC         | 93    | 9,095             | C T     | 0.345| 0.403| 0.148| 0.006|
| ASGC         | 240   | 1,195             | C T     | 0.305| 0.252| 0.106| 0.017|
| BRAINS       | 40    | 444               | C T     | 0.308| -0.032| 0.254| 0.900|
| CEDIR        | 65    | 407               | C T     | 0.351| 0.303| 0.196| 0.123|
| CHS          | 147   | 2,817             | C T     | 0.342| 0.182| 0.118| 0.125|
| DECODE       | 399   | 26,965            | C T     | 0.260| 0.004| 0.081| 0.960|
| FHS          | 48    | 4,335             | C T     | 0.325| 0.067| 0.241| 0.781|
| GASROS       | 169   | 1,202             | C T     | 0.300| 0.171| 0.124| 0.169|
| HVH          | 88    | 1,290             | C T     | 0.342| 0.108| 0.161| 0.502|
| ISGS         | 247   | 2,311             | C T     | 0.329| 0.077| 0.101| 0.444|
| WTCCC-G      | 330   | 797               | C T     | 0.373| 0.120| 0.095| 0.208|
| WTCCC-UK     | 460   | 5,175             | C T     | 0.325| 0.068| 0.073| 0.350|
| Meta-analysis| 2,326 | 56,033            | C T     | 0.325| 0.122| 0.010| 0.0002|
| **Large artery** |      |                    |         |      |     |      |
| ARIC         | 31    | 9,157             | C T     | 0.345| 0.490| 0.255| 0.055|
| ASGC         | 421   | 1,195             | C T     | 0.306| 0.163| 0.087| 0.060|
| BRAINS       | 118   | 444               | C T     | 0.308| -0.075| 0.161| 0.639|
| CEDIR        | 74    | 407               | C T     | 0.351| 0.235| 0.185| 0.205|
| DECODE       | 240   | 26,965            | C T     | 0.260| 0.025| 0.105| 0.815|
| GASROS       | 95    | 1,202             | C T     | 0.300| -0.076| 0.169| 0.653|
| HVH          | 61    | 1,290             | C T     | 0.342| 0.127| 0.176| 0.470|
| ISGS         | 229   | 2,329             | C T     | 0.329| 0.165| 0.103| 0.112|
| WTCCC-G      | 346   | 797               | C T     | 0.373| 0.213| 0.096| 0.026|
| WTCCC-UK     | 498   | 5,175             | C T     | 0.325| 0.075| 0.077| 0.281|
| Meta-analysis| 2,113 | 48,961            | t c     | 0.324| 0.116| 0.026| 0.001|
| **Small vessel** |      |                    |         |      |     |      |
| ARIC         | 63    | 9,125             | C T     | 0.345| -0.185| 0.193| 0.339|
| ASGC         | 310   | 1,195             | C T     | 0.305| 0.197| 0.098| 0.043|
| BRAINS       | 113   | 444               | C T     | 0.308| -0.045| 0.168| 0.787|
| CEDIR        | 25    | 407               | C T     | 0.351| 0.007| 0.312| 0.981|
| CHS          | 73    | 2,817             | C T     | 0.339| -0.110| 0.176| 0.533|
| DECODE       | 255   | 26,965            | C T     | 0.260| -0.035| 0.101| 0.728|
| GASROS       | 38    | 1,202             | C T     | 0.300| 0.320| 0.244| 0.190|
| HVH          | 173   | 1,290             | C T     | 0.342| 0.054| 0.125| 0.663|
| ISGS         | 201   | 2,357             | C T     | 0.329| -0.206| 0.115| 0.071|
| WTCCC-G      | 106   | 797               | C T     | 0.373| 0.267| 0.153| 0.082|
| WTCCC-UK     | 474   | 5,175             | C T     | 0.325| -0.096| 0.073| 0.192|
| Meta-analysis| 1,831 | 51,774            | C T     | 0.325| -0.009| 0.038| 0.811|

Subtypes of ischemic stroke were examined in MetaStroke, WTCCC-G, and WTCCC-UK. Subtyping included cardioembolic, large-artery, and small-vessel disease.

A1 = major allele; A2 = minor allele; ARIC = Atherosclerosis Risk in Communities; ASGC = Australian Stroke Genetics Collaborative; Beta = effect size; BRAINS = Bio-Repository of DNA in Stroke; CHS = Cardiovascular Health Study; FHS = Framingham Heart Study; Freq = frequency; GASROS = Genes Affecting Stroke Risk and Outcome Study; HVH = Heart and Vascular Health; ISGS = Ischemic Stroke Genetics Study; p = probability of association; SE = standard error of the effect; WTCCC-G = Wellcome Trust Case Control Consortium, German subcohort; WTCCC-UK = WTCCC, UK subcohort.
diabetes, or propensity to drink alcohol or smoke. Subgroup analysis of the study group having risk factor information (WTCCC2–Munich) attenuated the strength of the association but did not suggest that the action of the genetic variation was predominantly though 1 of these risk factors.

SNP rs505922 represents a single base pair change from T to C at position 135,139,050 and lies within the first intron of the ABO gene, although its haplotype block contains the promoter and introns 1 and 2. The minor allele frequency of this SNP is 36% in Northern Europeans. The ABO gene encodes a glycosyltransferase enzyme that catalyses the transfer of different carbohydrate groups onto the H antigen, thus forming A and B antigens of the ABO system. In support of a functional role in thrombosis (as opposed to atherosclerosis), the non-O blood group has also been shown to be a risk factor for venous thrombosis, and in a large prospective study, pulmonary embolism. A previous GWA study identified the same SNP, rs505922, to be associated with venous thromboembolism, and a recent GWA study of blood metabolites suggests that this locus may act via an effect on fibrinogen phosphorylation.

Our results demonstrate that the association between ABO SNPs and ischemic stroke is limited to large-artery and cardioembolic stroke, but absent in small-vessel stroke. Thromboembolism plays an important role in pathogenesis of both cardioembolic and large-artery stroke, with thrombus arising in the heart and on larger-artery atherosclerotic plaques, respectively, which may break off and embolize into the cerebral circulation. In both stroke subtypes, cerebral emboli can be detected in the cerebral circulation using transcranial Doppler, and antithromboembolic therapy reduces stroke risk. Recently, vWF inhibition has been shown to reduce cerebral thromboembolism in man, a clinical observation that is in keeping with our findings. In contrast, the pathogenesis of small-artery stroke is unclear, and the role of thrombosis remains uncertain. Our results suggest that thrombosis may be less important for this stroke subtype and explain why antithromboembolic medication is less effective. The subtype specificity we have identified is consistent with others’ results; of 5 GWA studies identified and replicated, 2 have been studies of cardioembolic stroke, and 1 of large-vessel stroke, and 1 of small-vessel stroke. Taken together, these data highlight that the clinical endpoint of ischemic stroke represents a varied phenotype likely resulting from multiple pathogenic mechanisms.

Other associations between SNPs and intermediate phenotypes included rs12137359 and FXIII activity and rs800292 and FXIII A subunit levels. Both variants are
found close to the gene encoding the FXIIIb subunit, which acts as a carrier protein for FXIIIa in the circulation and stabilizes FXIIIa to regulate activation; however, these SNPs were not associated with MI or ischemic stroke. We also identified associations between SNPs in the vicinity of the F7 gene and FVII:C, consistent with a number of studies that have previously identified relationships between variation in the structural genes for FVII and circulating levels.39,40 No other SNPs significantly associated with coagulation intermediate phenotypes were significantly associated with ischemic stroke.

There are a number of limitations to this work. First, TUK is predominantly female in its composition, for historical reasons. Although TUK subjects are representative of the general population variation8 and there is no evidence of an effect of gender on the ABO predisposition to cardiovascular disease, the associations identified in Stage 1 are pertinent to females from Northern Europe. Second, the clinical studies used for Stage 2 were heterogeneous in many respects. We decided that it was of overriding importance to obtain a large sample, so we combined prospective and cross-sectional studies. One of the main strengths of the study design was the use of multiple novel intermediate phenotypes, as well as having the power to investigate stroke subtypes. The Stage 3 study groups had differing methods of genotyping and imputation, but methods have been shown to be broadly comparable.41

In conclusion, using end-stage coagulation intermediate traits in healthy volunteers, we identified 23 genome-wide independent coagulation-associated SNPs, which were investigated in a number of clinical collections of stroke. Genetic variant rs505922 in the ABO locus was found to be associated with ischemic stroke, and in particular the subtypes large-vessel and cardioembolic stroke, but not small-vessel disease. This SNP was highly associated with vWF and FVIII in the discovery phase, and this observation throws light on possible mechanisms underlying end-stage coagulation in cardiovascular disease. It seems that common genetic variants exert some of their influence on end-stage stroke through coagulation, and further work is needed to tease apart these complex networks of interactions. The identification of the ABO locus through its association with vWF and FVIII points the way for mechanistic work to understand better the role of these 2 coagulation factors in end-stage arterial thrombosis.

**Authorship**

A.M.C. and P.G.H. contributed jointly to the work.

---

**Potential Conflicts of Interest**

F.M.K.W.: grants/grants pending, Wellcome Trust Fellowship 2007-12. M.D.: board membership, Boehringer Ingelheim, Ever Pharma; consultancy, Bayer Vital, Boehringer Ingelheim, Biol Heilmittel Heel, Bristol-Myers Squibb; grants/grants pending, EU FP7, DFG/DLR, Fondation Leducq, Vasc. Dem. Res. Foundation, Jackstaaedt Foundation, Corona Foundation; paid manuscript preparation, Georg Thieme Verlag, UpToDate, W. Kohlhammer Verlag. P.A.: board membership, Fondation Plan Alzheimer; consultancy, Servier, Total, Alzprotec; grants/grants pending, Ipsen, Sanofi-Aventis, AstraZeneca; stock/stock options, Genoscreen. P.-G.W.: board membership, Boehringer-Ingelheim. K.St.: board membership, stock/stock options, DeCODE Genetics. B.P.: Yale University, Service on Steering Committee for YODA funded by Medtronic; University of San Francisco, Service on DSMB for a clinical trial funded by Zoll LifeCor. Y.-C.C.: grants/grants pending, US Department of Veterans Affairs.

---

**Acknowledgment**

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

TwinsUK acknowledges the Wellcome Trust funding of TwinsUK as well as Intermediate Fellow F.M.K.W., EU Framework 6 support for funding EuroCLOT (QLRT-2001-01254), and genotyping by Wellcome Trust Sanger Centre and Sequenom/Gemini genomics, Cambridge. TwinsUK is also supported by an National Institute of Health Research (NIHR) Biomedical Resource Centre grant to Guy’s and St Thomas’ National Health Service (NHS) Foundation Trust and King’s College London. M. Boothby and J. Surr from The Division of Cardiovascular and Diabetes Research in Leeds, United Kingdom are thanked for their contributions to this program of work. The principal funding for the WTCCC2 stroke study was provided by the Wellcome Trust, as part of the WTCCC2 project (085475/B/08/Z, 085475/Z/08/Z, and WT084724/MA).

For MORGM, we thank M. Suvela and A. Vikman for MassARRAY SNP genotyping, the DNA core of the National Institute for Health and Welfare for DNA sample handling, and the Wellcome Trust Sanger Centre for Metabochip genotyping.

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

The MORGAM component of this research was in part funded through the European Community’s Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413. K.Si. has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics. MORGAM thanks K. Kaliva and T. Tuovinen for help with statistical analyses and K. Kristiansson, who processed and assessed the Metabochip data.

Grant support was provided by the Wellcome Trust (grant number 098051 to A.P.); Academy of Finland (grant number 251704 to A.P.); Academy of Finland, Center of Excellence in Complex Disease Genetics (grant numbers 213506 and 129680 to A.P.); European Community’s Seventh Framework Program (FP7/2007-2013), ENGAGE Consortium (grant agreement HEALTH-F4-2007-201413); EU/SYNSYS-Synaptic Systems (grant number 242167 to A.P.); and Sigrid Juselius Foundation, Finland (to A.P.).

The Heart and Vascular Health study research was supported by NIH National Heart, Lung, and Blood Institute (NHLBI) grants R01 HL085251 and R01 HL073410.

The Atherosclerosis Risk in Communities Study was carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C) and grants (R01HL087641, R01HL59367, and R01HL086694); National Human Genome Research Institute contract U01HG004402; and NIH contract HHSN268200625226C. Infrastructure was partly supported by grant number UL1RR025005, a component of the NIH and NIH Roadmap for Medical Research. Atherosclerosis Risk in Communities analyses performed as part of the MetaStroke project were supported by grant HL093029 to M.Fa. The authors thank the staff and participants of the Atherosclerosis Risk in Communities study for their important contributions.

The Australian Stroke Genetics Collaborative (ASGC) GWAS was funded by grants from the Australian National and Medical Health Research Council (NHMRC Project Grant ID: 569257), the Australian National Heart Foundation (NHF Project Grant ID: G04S 1623), the University of Newcastle, the Gladys M Brawn Fellowship scheme and the Vincent Fairfax Family Foundation in Australia. Australian population control data was derived from the Hunter Community Study. We thank the University of Newcastle for funding and the men and women of the Hunter region who participated in this study. EGH is supported by the Australian NHMRC Fellowship scheme.

Bio-Repository of DNA in Stroke (BRAINS) is an international study recruiting highly phenotyped patients with stroke. For the purposes of the current work, all patients were Caucasians. Diagnosis of stroke was confirmed using positive imaging (MR imaging or CT). BRAINS is partly funded by a Senior Fellowship from the UK Department of Health to Dr Pankaj Sharma, the Henry Smith Charity, and the UK–India Education Research Institute from the British Council.

The Besta Cerebrovascular Diseases Registry (CEDIR) was supported by the Italian Ministry of Health, years 2007-2010 (Annual Research Funding; Grant Numbers: RC 2007/LR6, RC 2008/LR6; RC 2009/LR8; RC 2010/LR8). Controls for CEDIR were provided by PROCARDIS, which acknowledges support from EC FP6 (LSHM-CT-2007-037273) and AstraZeneca AB. M.Fa. is supported by the British Heart Foundation Centre for Research Excellence. A.He. was supported by an Intermediate Fellowship award from the Oxford BHF Centre for Research Excellence. M.Fe. acknowledges support from the Wellcome Trust core award to the Wellcome Trust Centre for Human Genetics (075491/Z/04 and 090532/Z09/Z).

Cardiovascular Health Study research was supported by NHLBI contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, and N01-HC-45133 and NHLBI grants HL080295, HL075366, HL087652, and HL105756, with an additional contribution from the NIH National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIH National Institute on Aging (NIA). See also http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping were supported in part.
by National Center for Research Resources Clinical and Translational Sciences Institute grant UL 1RR033176 and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Framingham Heart Study research was supported by NHLBI contracts N01-HC-25195 and N02-HL-6-4278; and NINDS (NS17950), NHLBI (HL093029) and NIA (AG08122, AG16495, AG033193, AG031287).

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

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

The Massachusetts General Hospital Genes Affecting Stroke Risk and Outcome study was supported by grant number U01NS069208 from the NINDS. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NINDS or the NIH. Additional funding was provided by the American Heart Association/Bugher Foundation Centers for Stroke Prevention Research.

Rotterdam Study: The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810, Erasmus Medical Center and Erasmus University Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), the Municipality of Rotterdam, and the Netherlands Heart Foundation (Nederlands Hartstichting) 2009B102.

Nicole Soranzo’s team is supported by the Wellcome Trust (Grant Codes WT098051 and WT091310), the EU FP7 (EPIGENESYS Grant Code 257082 and BLUEPRINT Grant Code HEALTH-F5-2011-282510), C.L. and E.G.H. received funding from the Australian National Health and Medical Research Council (project grant 569257).

The EuroCLOT consortium thanks all the twin pairs who gave up their time to contribute to this study, and the staff of the NIHR BRC TwinsUK BioResource, King’s College London. The Edinburgh Stroke Study component of the WTCCC2 ischaemic stroke GWAS wishes to thank Drs C. Jackson (study coordination); M. Dennis, P. Sandercok, C. Warlow, S. Hart, and S. Keir (patient recruitment and characterization); Dr J. Wardlaw, A. Farrall, and G. Potter (neuroimaging interpretation); A. Hutchison and M. McDowall (programming and database support); L. Foley (genetics laboratory management); and the patients and caregivers who took part. We thank all the participants of the MORGAM project. The Heart Protection Study thanks J. C. Hopewell, J. Armitage, S. Parish, and R. Collins for assistance. We thank the staff and participants of the Atherosclerosis Risk in Communities study for their important contributions.

**APPENDIX**

**Membership of Wellcome Trust Case Control Consortium 2**

**MANAGEMENT COMMITTEE**, Peter Donnelly (Chair), Wellcome Trust Centre for Human Genetics and Department of Statistics, University of Oxford,
References

1. Lopez AD, Mathers CD, Ezzati M, et al. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet 2006;367:1747–1757.

2. Souto JC, Almasy L, Borrell M, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. Am J Hum Genet 2000;67:1452–1459.

3. Carter AM, Cymbalista CM, Spector TD, Grant PJ. Heritability of clot formation, morphology, and lysis. The EuroCLOT study. Arterioscler Thromb Vasc Biol 2009;29:600–605.

4. Williams FM, Carter AM, Kato B, et al. Identification of quantitative trait loci for fibrin clot phenotypes: the EuroCLOT study. Arterioscler Thromb Vasc Biol 2009;29:600–605.

5. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443–453.

6. Manolio TA. Collaborative genome-wide association studies of diverse diseases: programs of the NHGRI’s office of population genomics. Pharmacogenomics 2009;10:235–241.
7. Spector TD, Williams FM. The UK Adult Twin Registry (TwinsUK). Twin Res Hum Genet 2006;9:899–906.
8. Andrew T, Hart DJ, Snieder H, et al. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. Twin Res 2001;4:464–477.
9. Carter AM, Catto AJ, Mansfield MW, et al. Predictive variables for mortality after acute ischemic stroke. Stroke 2007;38:1873–1880.
10. de Lange M, Snieder H, Ariens RA, et al. The genetics of haemostasis: a twin study. Lancet 2001;357:101–105.
11. Freeman MS, Mansfield MW, Barrett JH, Grant PJ. Genetic contribution to circulating levels of hemostatic factors in healthy families with effects of known genetic polymorphisms on heritability. Arterioscler Thromb Vasc Biol 2002;22:506–510.
12. Freeman MS, Mansfield MW, Barrett JH, Grant PJ. Heritability of features of the insulin resistance syndrome in a community-based study of healthy families. Diabet Med 2002;19:994–999.
13. Evans A, Salomaa V, Kulaasmaa S, et al. MORGAM (an international pooling of cardiovascular cohorts). Int J Epidemiol 2005;34:21–27.
14. Kulthinal N, Karvanen J, Saarela O, Kulaasmaa K. Case-cohort design in practice—experiences from the MORGAM Project. Epidemiol Perspect Innov 2007;4:15.
15. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke 1993;24:35–41.
16. 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;44:328–333.
17. Hysi PG, Young TL, Mackey DA, et al. A genome-wide association study for myopia and refractive error identifies a susceptibility locus at 15q25. Nat Genet 2010;42:902–905.
18. Marchini J, Howie B, Myers S, et al. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39:906–913.
19. Sandler K, Alanne M, Kristiansson K, et al. Gender differences in genetic risk profiles for cardiovascular disease. PLoS One 2008;3:e3615.
20. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575.
21. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–2191.
22. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97–101.
23. Smith NL, Chen MH, Dehghan A, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von Willebrand factor: the CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) Consortium. Circulation 2010;121:1382–1392.
24. Ionescu DA, Marcu I, Bicescu E. Cerebral thrombosis, cerebral haemorrhage, and ABO blood-groups. Lancet 1976;1:278–280.
25. Wu O, Bayoumi N, Vickers MA, Clark P. ABO(H) blood groups and vascular disease: a systematic review and meta-analysis. J Thromb Haemost 2008;6:62–69.
26. Reilly MP, Li M, He J, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet 2011;377:383–392.
27. Schunkert H, Konig IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333–338.
28. Wiggins KL, Smith NL, Glazer NL, et al. ABO genotype and risk of thrombotic events and hemorrhagic stroke. J Thromb Haemost 2009;7:263–269.
29. Wolpin BM, Kabrhel C, Varraso R, et al. Prospective study of ABO blood type and the risk of pulmonary embolism in two large cohort studies. Thromb Haemost 2010;104:962–971.
30. Tregouet DA, Heath S, Saut N, et al. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. Blood 2009;113:5298–5303.
31. Suhre K, Shin SY, Petersen AK, et al. Human metabolic individuality in biomedical and pharmaceutical research. Nature 2011;477:54–60.
32. King A, Markus HS. Doppler embolic signals in cerebrovascular disease and prediction of stroke risk: a systematic review and meta-analysis. Stroke 2009;40:3711–3717.
33. Markus HS, McCollum C, Imray C, et al. The von Willebrand inhibitor ARC1779 reduces cerebral embolization after carotid endarterectomy: a randomized trial. Stroke 2011;42:2149–2153.
34. Panton L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. Lancet Neurol 2010;9:689–701.
35. Gretarsdottir S, Thorleifsson G, Manolescu A, et al. Risk variants for atrial fibrillation on chromosome 4q22 associate with ischemic stroke. Ann Neurol 2008;64:402–409.
36. Gudbjartsson DF, Holm H, Gretarsdottir S, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischaemic stroke. Nat Genet 2009;41:876–878.
37. Gschwendtner A, Bevan S, Cole JW, et al. Sequence variants on chromosome 9p21.3 confer risk for atherosclerotic stroke. Ann Neurol 2009;65:531–539.
38. Kubo M, Hata J, Ninomiya T, et al. A nonsynonymous SNP in PRKCH (protein kinase C eta) increases the risk of cerebral infarction. Nat Genet 2007;39:212–217.
39. Bernardi F, Ancieli P, Bertina RM, et al. Contribution of factor VII genotype to activated FVII levels. Differences in genotype frequencies between northern and southern European populations. Arterioscler Thromb Vasc Biol 1997;17:2548–2553.
40. Calafell F, Almasy L, Sabater-Lleal M, et al. Sequence variation and genetic evolution at the human F12 locus: mapping quantitative trait nucleotides that influence FXII plasma levels. Hum Mol Genet 2010;19:517–525.
41. Nothnagel M, Ellinghaus D, Schreiber S, et al. A comprehensive evaluation of SNP genotype imputation. Hum Genet 2009;125:163–171.