Supplementary information for

Shared Genetic Contributions to Atrial Fibrillation and Ischemic Stroke Risk

Sara L. Pulit, Lu-Chen Weng, Patrick F McArdle, Ludovic Trinquart, Seung Hoan Choi, Braxton D. Mitchell, Jonathan Rosand, Paul I W de Bakker, Emelia J Benjamin, Patrick T Ellinor, Steven J Kittner, Steven A Lubitz*, Christopher D Anderson*, on behalf of the Atrial Fibrillation Genetics Consortium and the International Stroke Genetics Consortium.

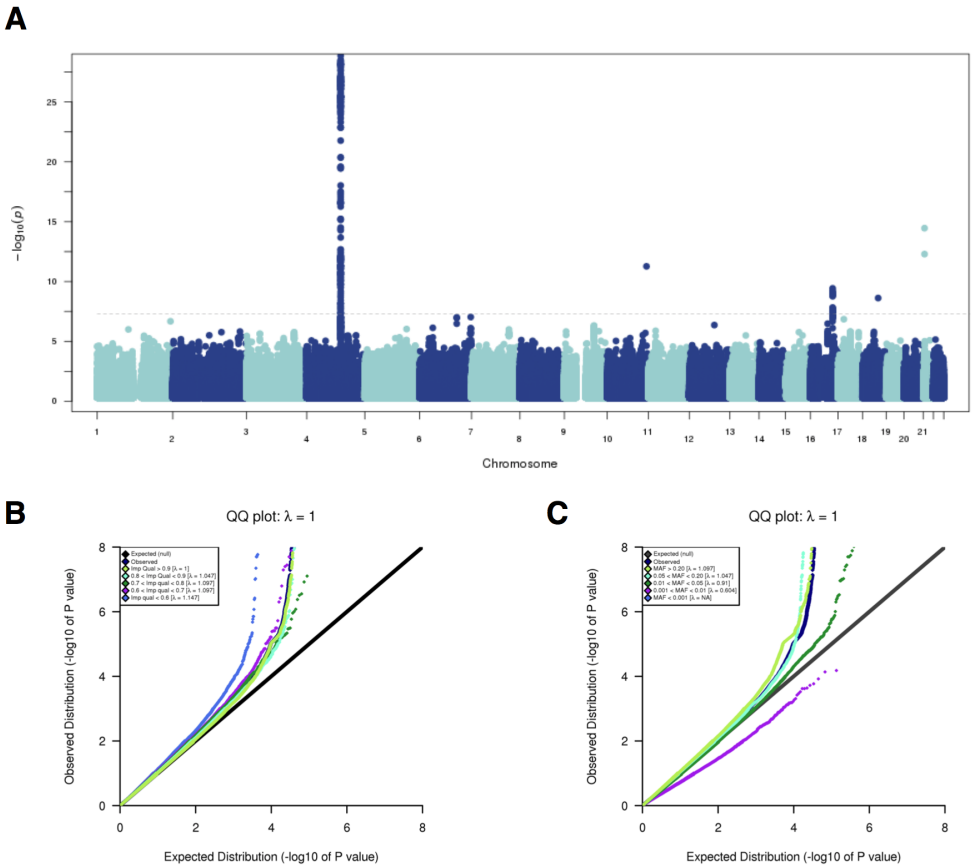
Code and data release

For code and data related to this project, including (but not limited to) sample identifiers, SNP identifiers, links to GWAS summary-level data, and SNP weights used in the construction of the polygenic risk score, please see the following GitHub repository: https://github.com/UMCUGenetics/Afib-Stroke-Overlap.

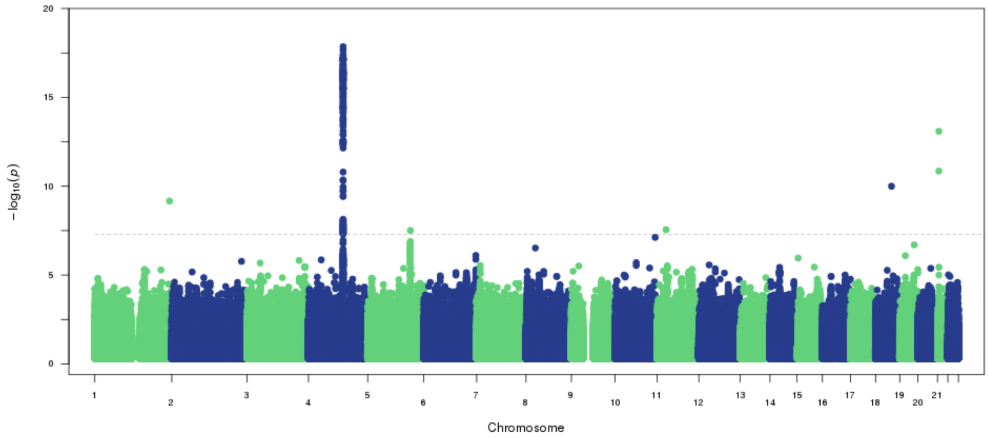
Table of Contents

Supplementary Figures

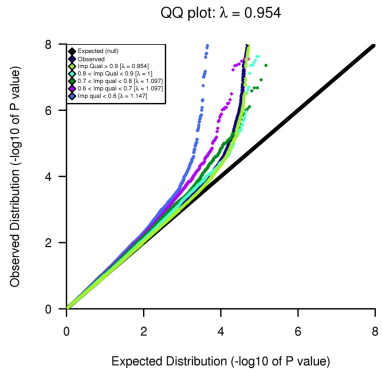
Supplementary Figure 1 | Genome-wide association study (GWAS) of atrial fibrillation in SiGN. (A) We performed a GWAS of 3,190 cases with atrial fibrillation, or paroxysmal atrial fibrillation, as well as other diagnoses suggestive of underlying atrial fibrillation, including left atrial thrombus, sick sinus syndrome, and atrial flutter. We additionally included 28,026 referents. We used a linear mixed model and adjusted the model for principal components and sex. The majority of atrial fibrillation risk loci identified through previous GWAS efforts were identified here at nominal significance or better (see **Supplementary Table 2**). The Manhattan plot only shows QC-passing SNPs with minor allele frequency > 1% and imputation quality score > 0.8. (B) Quantile-quantile (QQ) plot indicating SNPs stratified by minor allele frequency and the corresponding genomic inflation factor (lambda, λ) for each stratum. (C) QQ plot showing SNPs stratified by imputation quality and the corresponding lambda for each stratum. Figures D-F are identical to those of A-C, but for the analysis performed in atrial fibrillation cases only (N = 1,751). We performed this is an internal sensitivity analysis only, to ensure that more broadly defining the atrial fibrillation phenotype was not introducing additional phenotypic noise.



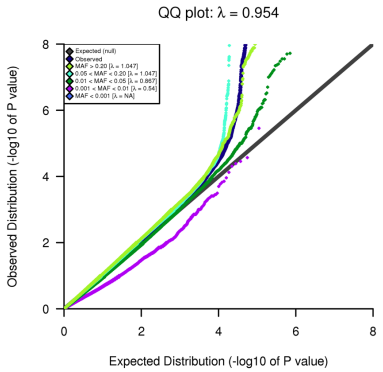
D


20
15
10
5
0
$-\log_{10}(p)$
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
Chromosome

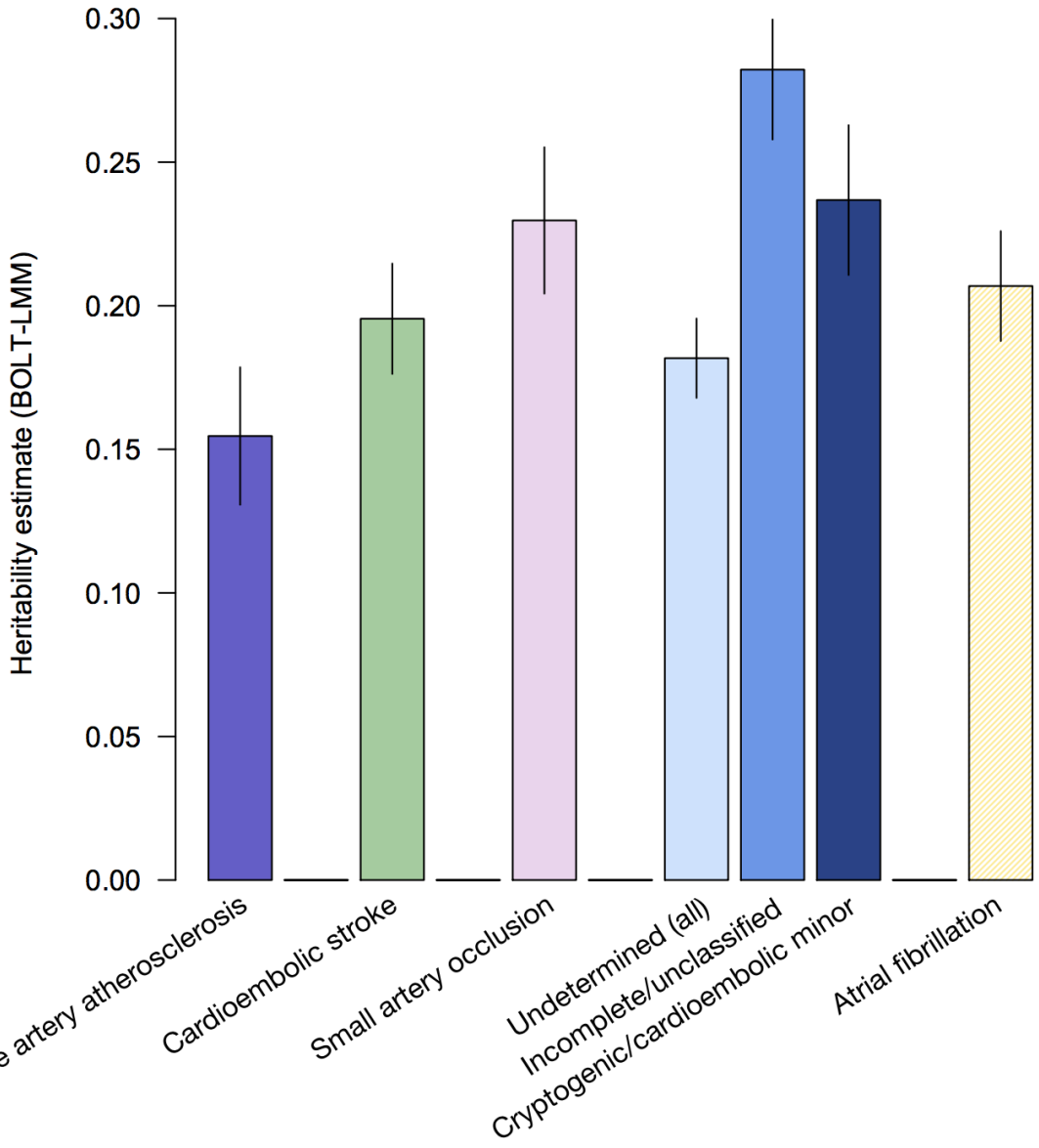
E


QQ plot: λ = 0.954
Observed Distribution (-log10 of P value)
Expected Distribution (-log10 of P value)
0
2
4
6
8

F

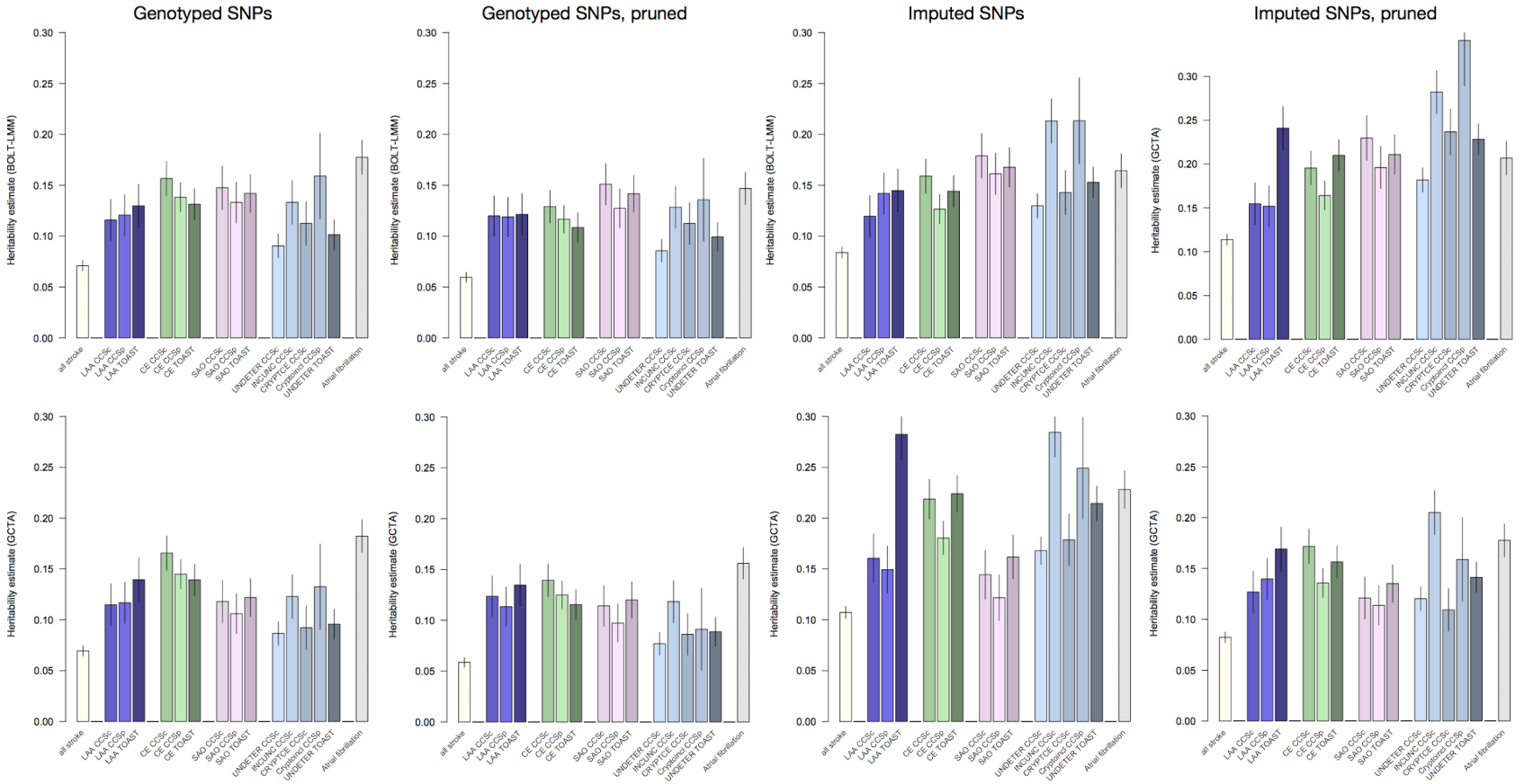

QQ plot: λ = 0.954
Observed Distribution (-log10 of P value)
Expected Distribution (-log10 of P value)
0
2
4
6
8

Supplementary Figure 2 | Estimated heritability of ischemic stroke subtypes and atrial fibrillation. Using all available stroke cases in SiGN, we estimated SNP-based heritability of the ischemic stroke subtypes (as sub-typed by the CCS Causative subtyping system) and atrial fibrillation (using the subset of 3,190 cases with atrial fibrillation) using BOLT-LMM and a genetic relationship matrix of high-quality SNPs converted to best-guess genotypes (imputation quality > 0.8, minor allele frequency > 0.01, and pruned at a linkage disequilibrium threshold of 0.2). We assumed a trait prevalence of 1% for all phenotypes. We found heritability estimates in cardioembolic stroke (green) and atrial fibrillation (yellow) to be approximately similar.



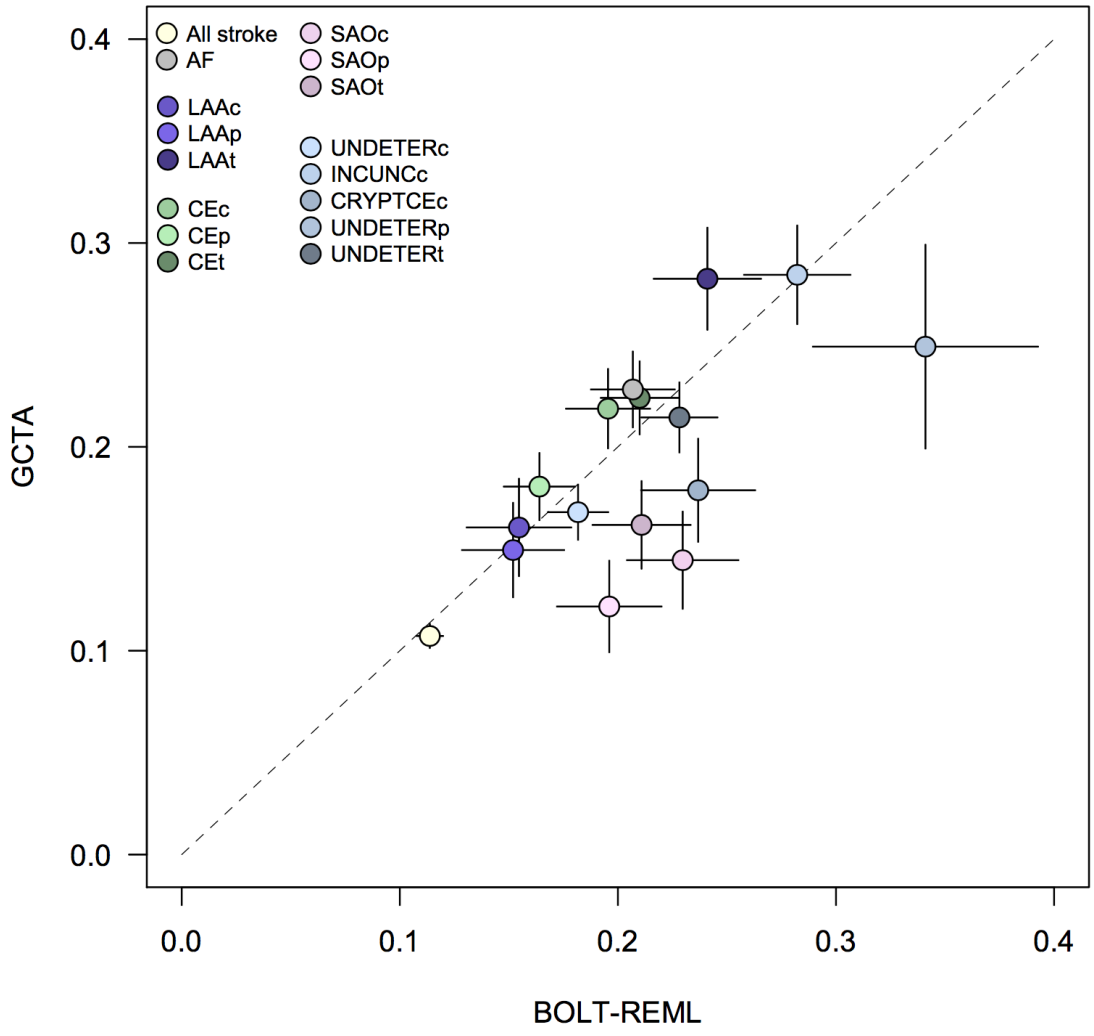
Supplementary Figure 3 | Heritability of ischemic stroke, its subtypes, and atrial fibrillation. We computed the SNP-based heritability of all stroke, all stroke subtypes, and atrial fibrillation using BOLT-LMM (top row) and GCTA (bottom row). All SNPs used for analysis had a minor allele frequency > 1% and imputation quality > 0.8 (for imputed SNPs). Imputed SNPs were converted to best-guess genotypes. We assumed a trait prevalence of 1% for all phenotypes and tested the robustness of h_g^2 estimates to SNPs included in the GRM by using four different GRMs: (a) genotyped SNPs only; (b) genotyped, pruned, and filtered (see **Supplemental Methods**); (c) imputed; and (d) imputed, pruned, and filtered. We converted the imputed SNPs to hard-call genotypes before performing heritability analyses. Estimates are shown below, including error bars. The underlying data for these figures are provided in **Supplementary Table 3**.

LAA, large artery atherosclerosis; CE, cardioembolic stroke; SAO, small artery occlusion; UNDETER, undetermined; INCUNC, incomplete/unclassified; CRYPTCE, cryptogenic and CE minor; Cryptoincl, cryptogenic; CCSc, CCS Causative subtyping system; CCSp, CCS Phenotypic subtyping system; TOAST, TOAST subtyping system.

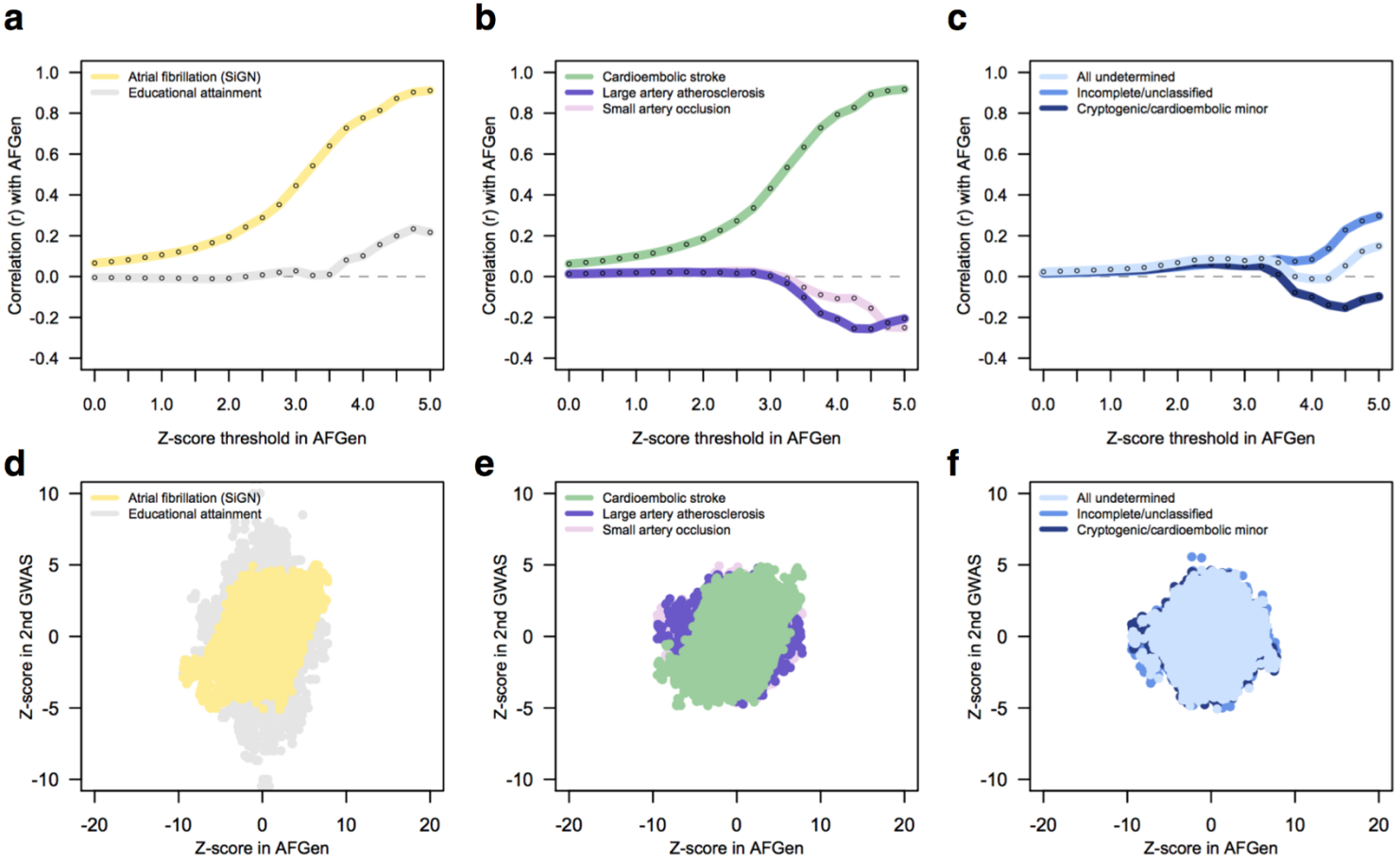


Supplementary Figure 4 | Comparison of heritability estimates from BOLT-LMM and GCTA. We computed the heritability of all stroke, all stroke subtypes, and atrial fibrillation using BOLT-LMM and GCTA, as shown in **Supplementary Figure 2**. Below, you will find a comparison of the two methods, with BOLT-REML on the x-axis and GCTA estimates on the y-axis. Error bars are shown for the respective estimates.

AF, atrial fibrillation; CE, cardioembolic stroke; LAA, large artery atherosclerosis; SAO, small artery occlusion; UNDETER, undetermined; INCUNC, incomplete/unclassified; CRYPTCE, cryptogenic/CE minor; c, CCS Causative; p, CCS Phenotypic; t, TOAST.



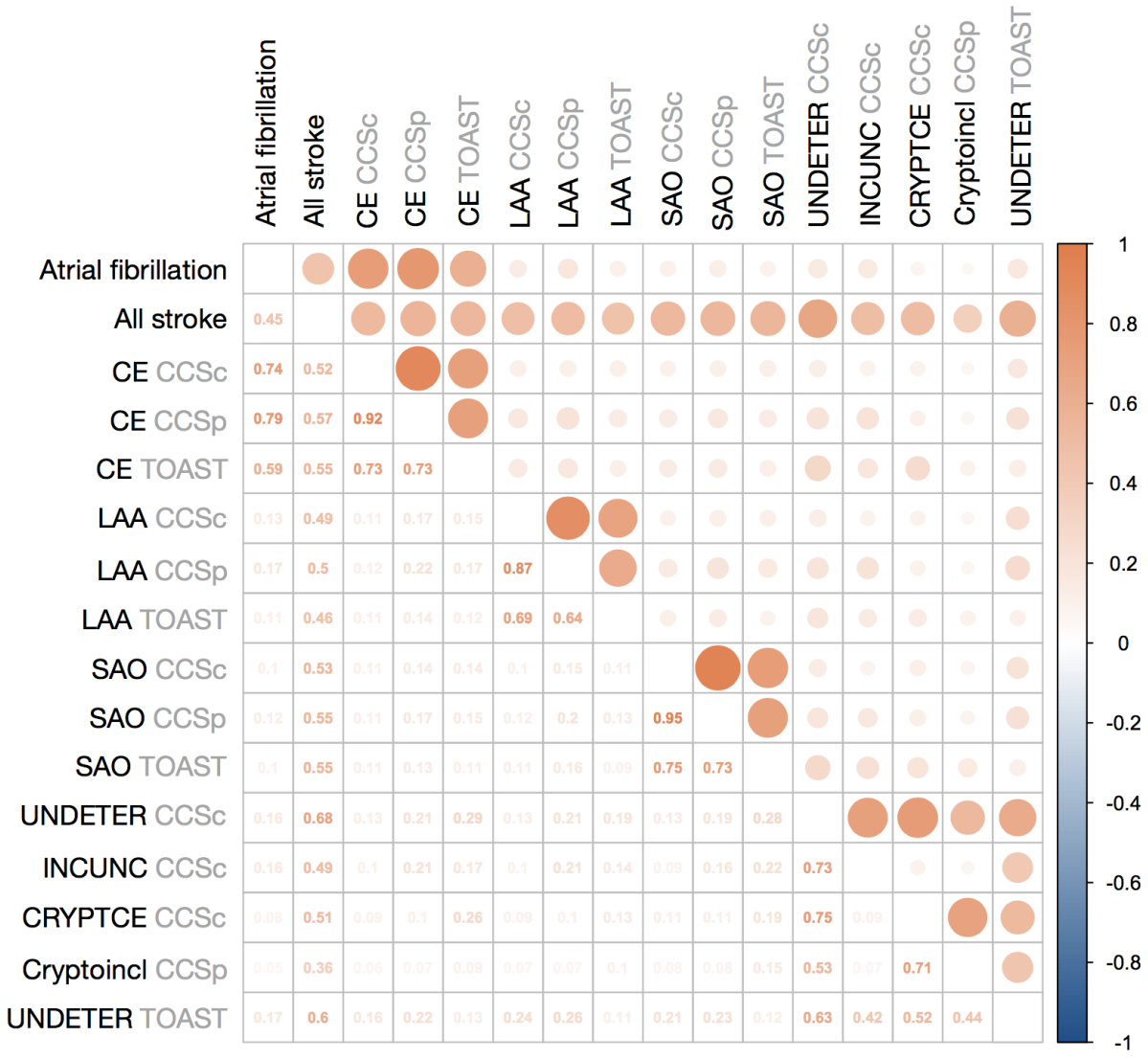
Supplementary Figure 5 | Genetic correlations between atrial fibrillation and ischemic stroke subtypes. To estimate genetic correlation between atrial fibrillation and ischemic stroke subtypes, we calculated Pearson's r between SNP z-scores in the AFGen GWAS of atrial fibrillation and in GWAS of ischemic stroke subtypes and atrial fibrillation performed here in the SiGN data. Here, we present data identical to that shown in Figure 2 of the main manuscript, but removing ±2Mb around the two most significant loci discovered in atrial fibrillation and cardioembolic stroke: the region around *PITX2* (chromosome 4) and the region around *ZFHX3* (chromosome 16). (a) Genome wide, atrial fibrillation in AFGen and in SiGN correlate with increasing strength as the z-score in AFGen increases. Educational attainment is included here as a null comparator. (b) Genetic signal in cardioembolic stroke also correlates strongly with atrial fibrillation genetic signal in AFGen, but we do not observe correlation between atrial fibrillation and the other primary stroke subtypes. (c) Removing the *PITX2* and *ZFHX3* regions leaves only somewhat modest correlation between the incomplete/unclassified undetermined subtype and atrial fibrillation. Panels (d-f) show underlying data. Correlations restricted to those SNPs used in the polygenic risk score for atrial fibrillation were: AFGen vs atrial fibrillation in SiGN, r = 0.78; AFGen vs. cardioembolic stroke in SiGN, r = 0.75.



Supplementary Figure 6 | Genetic correlation and phenotypic correlation of atrial fibrillation and stroke subtypes in SiGN. (a) Using genome-wide SNP effects extracted from GWAS of atrial fibrillation, all stroke, and stroke subtypes, we calculated the Pearson's correlation (r) between each pair of available phenotypes (blue indicates strong negative correlation; orange indicates strong positive correlation). Here, we show all correlations. Correlations are indicated by circle size in the upper half of the square, and the exact correlation values are shown in the lower half of the square.

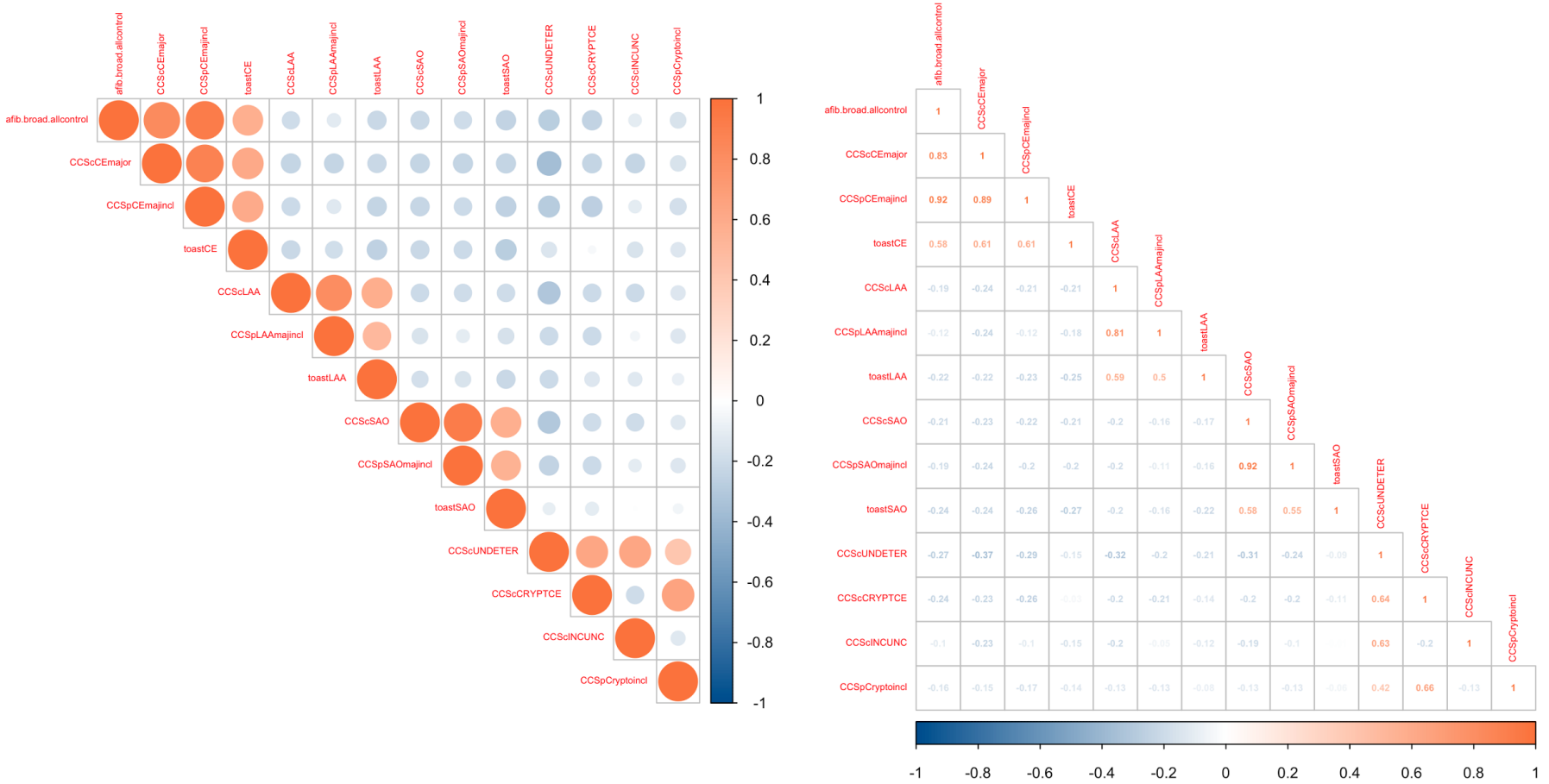
CE, cardioembolic stroke; LAA, large artery atherosclerosis; SAO, small artery occlusion; UNDETER, undetermined; INCUNC, incomplete/unclassified; CRYPTCE, cryptogenic and CE minor; Cryptoincl, cryptogenic; CCSc, CCS Causative subtyping system; CCSp, CCS Phenotypic subtyping system; TOAST, TOAST subtyping system.

a.



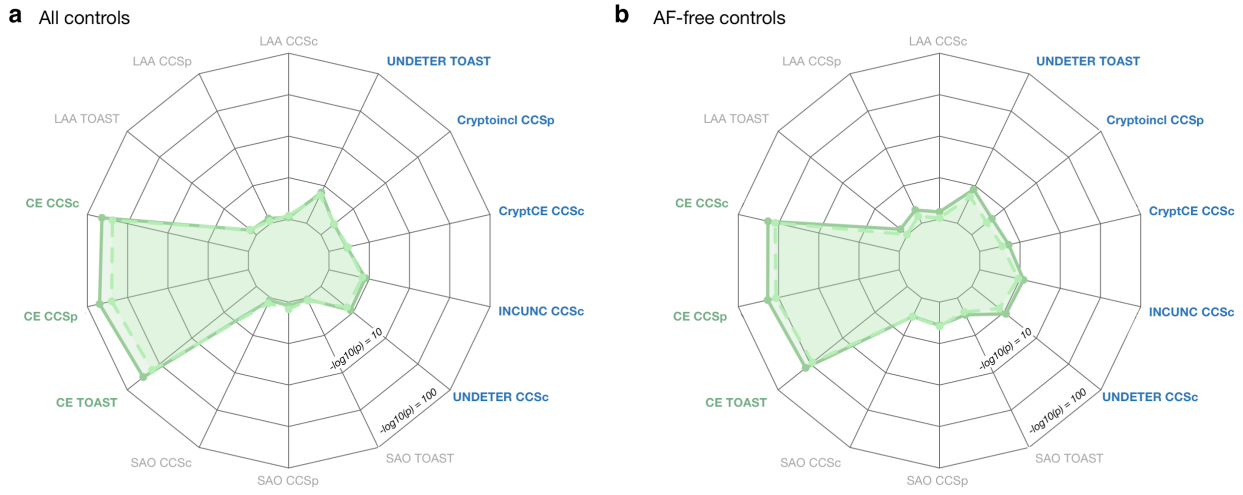
b. Same correlation calculations as in (a), but this time using the phenotypic data only (and looking in cases only, as all controls have the same phenotype). Note that the atrial fibrillation phenotypes and cardioembolic stroke phenotypes are highly correlated in the SiGN data (r = 0.83 between atrial fibrillation and cardioembolic stroke as determined by the CCS Causative subtype system).

CE, cardioembolic stroke; LAA, large artery atherosclerosis; SAO, small artery occlusion; UNDETER, undetermined; INCUNC, incomplete/unclassified; CRYPTCE, cryptogenic and CE minor; Cryptoincl, cryptogenic; CCSc, CCS Causative subtyping system; CCSp, CCS Phenotypic subtyping system; TOAST, TOAST subtyping system.



Supplementary Figure 7 | Association of atrial fibrillation polygenic risk score in ischemic stroke subtypes. We constructed a polygenic risk score (PRS) from atrial fibrillation-associated SNPs, and tested for association between the score and ischemic stroke subtypes using (a) all available controls (N = 28,026) and (b) controls without atrial fibrillation (N = 3,861). All subtypes from all available subtyping systems are shown here. The PRS strongly associated to cardioembolic stroke (subtypes highlighted in green font) in both sets of controls. In the atrial fibrillation-free set of controls (b) we observed nominal association of the PRS to incomplete/unclassified stroke. Undetermined subtypes are indicated in blue font.

CE, cardioembolic stroke; LAA, large artery atherosclerosis; SAO, small artery occlusion; UNDETER, undetermined; INCUNC, incomplete/unclassified; CRYPTCE, cryptogenic and CE minor; Cryptoincl, cryptogenic; CCSc, CCS Causative subtyping system; CCSp, CCS Phenotypic subtyping system; TOAST, TOAST subtyping system.



Supplementary Tables

Supplementary Table 1 | Atrial fibrillation cases and controls available from the Stroke Genetics Network (SiGN) Consortium.

*As classified by the CCS Causative system (note that this table is a repeat of **Table 1** from the main manuscript):*

Phenotype	Total	Cardioembolic	Large artery atherosclerosis	Small artery occlusion	Undetermined	
					Incomplete/ unclassified	Cryptogenic/ CE minor
Atrial fibrillation	1,751	1,495	63	32	151	0
Paroxysmal atrial fibrillation	1,315	1,088	52	23	138	0
Left atrial thrombus	48	37	3	3	4	0
Sick sinus syndrome	79	65	5	3	4	0
Atrial Flutter	106	90	4	2	10	0
Total	3,190	2,684	123	61	298	0

As classified by the CCS Phenotypic system (note that this system allows a case to be classified into more than one subtype):

Phenotype	Total	Cardioembolic	Large artery atherosclerosis	Small artery occlusion	Undetermined
Atrial fibrillation	1,751	1,751	161	58	0
Paroxysmal atrial fibrillation	1,315	1,315	126	61	0
Left atrial thrombus	48	48	7	4	0
Sick sinus syndrome	79	79	8	4	0
Atrial Flutter	106	106	11	3	0
Total	3,190	3,190	302	126	0

As classified by the TOAST system:

Phenotype	Total	Cardioembolic	Large artery atherosclerosis	Small artery occlusion	Undetermined
Atrial fibrillation	1,751	1,254	26	23	170
Paroxysmal atrial fibrillation	1,315	880	25	19	178
Left atrial thrombus	48	35	1	1	9
Sick sinus syndrome	79	48	0	1	13
Atrial Flutter	106	75	2	3	12
Total	3,190	2,207	54	47	371

Overlap of atrial fibrillation and cardioembolic stroke in the three subtyping systems in SiGN (CCSc, CCS Causative; CCSp, CCS Phenotypic; TOAST):

Phenotype	CCSc Cardioembolic	CCSp Cardioembolic	TOAST Cardioembolic
Atrial fibrillation	1,495	1,751	1,254
Paroxysmal atrial fibrillation	1,088	1,315	880
Left atrial thrombus	37	48	35
Sick sinus syndrome	65	79	48
Atrial Flutter	90	106	75
No atrial fibrillation phenotypes	316	418	903
Total	3,000	3,608	3,333

Supplementary Table 2 | Look-up of previously-associated atrial fibrillation SNPs in SiGN. After performing a GWAS of atrial fibrillation in the SiGN data, we looked up the 26 known genetic risk loci for atrial fibrillation, as identified in the latest GWAS.[1] Twenty-four of the 25 signals present in the SiGN data were directionally consistent with the previous GWAS. The only signal not directionally consistent was discovered through eQTL analysis. One signal, a rare variant burden signal, was absent from our data (all SNPs here have allele frequency > 1%).

Supplementary Table 2 is provided as a separate, downloadable Excel spreadsheet as well as a tab-delimited text available at the project GitHub repository:

https://github.com/saralpulit/Afib-Stroke-Overlap/blob/master/SupplementaryTable2.afib.hits.SiGN-lookup.txt

The first 14 columns are taken from *Christophersen, et al. Large-scale analysis of common and rare variants identify 12 new loci associated with atrial fibrillation. Nature Genetics, 2017.*[1] Those columns are:

SNP	single-nucleotide polymorphism; rs identifier
CHR	chromosome
BP	basepair (hg19)
Genes	Closest gene(s)
Location	Where the SNP resides relative to the listed gene
Risk	Risk allele
Ref	Reference allele
RAF	Risk allele frequency
OR	Odds ratio
CI95_1	95% confidence interval for the odds ratio (lower bound)
CI95_2	95% confidence interval for the odds ratio (upper bound)
Pval	Association p-vlaue
Mean_imp	Imputation quality
Analysis	The analysis the variant or gene was discovered in (ExWAS, expression QTL analysis; Meta, meta-analysis; RVAS, rare variant association study)

The remaining columns provided are data points extracted from the atrial fibrillation GWAS in SiGN. They are:

SiGN_RAF	Risk allele frequency in SiGN
SiGN_INFO	Imputation quality (info score) in SiGN
SiGN_BOLT_BETA	Beta of the SNP taken from BOLT-LMM; note that this is a beta that results from a linear mixed model
SiGN_LIAB_BETA	The beta, converted to the liability scale
SiGN_OR	Odds ratio in SiGN
SiGN_SE	Standard error of SIGN_BOLT_BETA
SiGN_P_	BOLT P-value from BOLT-LMM (for the infinitesimal model only)

Supplementary Table 3 | Heritability calculations in atrial fibrillation and ischemic stroke subtypes. (a) We calculated the SNP-based heritability (h_g^2) of atrial fibrillation, all ischemic stroke, and the stroke subtypes using GCTA[2]. All SNPs used had minor allele frequency > 1% and imputation quality > 0.8 (for imputed SNPs). Imputed SNPs were converted to best-guess genotypes. We assumed a trait prevalence of 1% for all phenotypes and tested the robustness of h_g^2 estimates to SNPs included in the GRM by using four different GRMs: (i) genotyped only; (ii) genotyped, pruned, and filtered (see **Supplemental Methods**); (iii) imputed; and (iv) imputed, pruned, and filtered. (b) We performed the exact same analysis but using BOLT-LMM to estimate h_g^2. BOLT-LMM estimates were converted to the liability scale (see **Supplemental Methods**).

Geno, genotyped; SE, standard error; CCSc, CCS Causative; CCSp, CCS Phenotypic

a. h_g^2 estimates in GCTA

Subtype	Subtyping system	Cases	Geno h_g^2 (SE)	Geno, filtered h_g^2 (SE)	Imputed h_g^2 (SE)	Imputed, filtered h_g^2 (SE)
Large artery athero-sclerosis	CCSc	2,385	0.115 (0.020)	0.124 (0.020)	0.127 (0.020)	0.160 (0.024)
	CCSp	2,449	0.117 (0.020)	0.113 (0.019)	0.140 (0.020)	0.149 (0.023)
	TOAST	2,318	0.139 (0.021)	0.135 (0.021)	0.169 (0.022)	0.282 (0.025)
Cardio-embolic	CCSc	3,000	0.166 (0.017)	0.139 (0.016)	0.172 (0.017)	0.219 (0.019)
	CCSp	3,608	0.145 (0.014)	0.125 (0.014)	0.136 (0.014)	0.181 (0.016)
	TOAST	3,333	0.139 (0.015)	0.115 (0.015)	0.156 (0.016)	0.224 (0.018)
Small artery occlusion	CCSc	2,262	0.118 (0.021)	0.114 (0.020)	0.121 (0.021)	0.144 (0.024)
	CCSp	2,419	0.106 (0.020)	0.097 (0.019)	0.114 (0.019)	0.122 (0.022)
	TOAST	2,631	0.122 (0.019)	0.120 (0.018)	0.135 (0.019)	0.162 (0.021)
Undeter-mined	CCSc	4,574	0.087 (0.012)	0.077 (0.011)	0.120 (0.012)	0.168 (0.014)
	CCSc (INCUNC)	2,280	0.123 (0.021)	0.118 (0.021)	0.205 (0.022)	0.284 (0.024)
	CCSc (CRYPTCE)	2,294	0.092 (0.021)	0.086 (0.020)	0.109 (0.021)	0.179 (0.025)
	CCSp	1,096	0.132 (0.042)	0.091 (0.040)	0.159 (0.041)	0.249 (0.050)
	TOAST	3,479	0.096 (0.015)	0.089 (0.014)	0.141 (0.015)	0.214 (0.017)
--	All stroke	13,390	0.069 (0.005)	0.059 (0.005)	0.082 (0.005)	0.107 (0.006)
--	Atrial fibrillation	3,190	0.182 (0.016)	0.156 (0.015)	0.178 (0.016)	0.228 (0.019)

b. h_g^2 estimates in BOLT-LMM

Subtype	Subtyping system	Cases	Geno h_g^2 (SE)	Geno, filtered h_g^2 (SE)	Imputed h_g^2 (SE)	Imputed, filtered h_g^2 (SE)
Large artery athero-sclerosis	CCSc	2,385	0.116 (0.020)	0.120 (0.020)	0.120 (0.020)	0.155 (0.024)
	CCSp	2,449	0.121 (0.020)	0.119 (0.019)	0.142 (0.020)	0.152 (0.023)
	TOAST	2,318	0.130 (0.021)	0.121 (0.020)	0.145 (0.021)	0.241 (0.025)
Cardio-embolic	CCSc	3,000	0.157 (0.017)	0.129 (0.016)	0.159 (0.017)	0.195 (0.019)
	CCSp	3,608	0.138 (0.014)	0.117 (0.014)	0.127 (0.014)	0.164 (0.016)
	TOAST	3,333	0.131 (0.015)	0.108 (0.015)	0.144 (0.015)	0.210 (0.018)
Small artery occlusion	CCSc	2,262	0.147 (0.021)	0.151 (0.020)	0.179 (0.022)	0.230 (0.026)
	CCSp	2,419	0.133 (0.020)	0.127 (0.019)	0.161 (0.020)	0.196 (0.024)
	TOAST	2,631	0.142 (0.019)	0.142 (0.018)	0.168 (0.019)	0.211 (0.022)
Undeter-mined	CCSc	4,574	0.090 (0.012)	0.086 (0.011)	0.130 (0.012)	0.182 (0.014)
	CCSc (INCUNC)	2,280	0.133 (0.021)	0.118 (0.021)	0.128 (0.021)	0.282 (0.024)
	CCSc (CRYPTCE)	2,294	0.112 (0.021)	0.112 (0.021)	0.143 (0.021)	0.237 (0.026)
	CCSp	1,096	0.159 (0.042)	0.136 (0.041)	0.213 (0.042)	0.341 (0.052)
	TOAST	3,479	0.101 (0.015)	0.099 (0.014)	0.153 (0.015)	0.228 (0.017)
--	All stroke	13,390	0.169 (0.012)	0.059 (0.005)	0.084 (0.005)	0.114 (0.006)
--	Atrial fibrillation	3,190	0.169 (0.016)	0.140 (0.015)	0.156 (0.016)	0.200 (0.018)

Supplementary Table 4 | Genetic correlations between atrial fibrillation and ischemic stroke subtypes. To estimate genetic correlation between atrial fibrillation and ischemic stroke subtypes, we calculated Pearson's r between SNP z-scores in the Atrial Fibrillation Genetics (AFGen) GWAS of atrial fibrillation and in GWAS of ischemic stroke subtypes and atrial fibrillation performed here in the SiGN data. The correlation calculations are provided in this table, which is split into two parts and is available to download in text format here:

Part A correlations calculated across all genome-wide SNPs:
https://github.com/saralpulit/Afib-Stroke-Overlap/blob/master/SuppTable4.partA.SiGN.AFGen.trait.correlations.txt

Part B correlations calculated across all genome-wide SNPs except those ±2Mb from the PITX2 and ZFHX3 index SNPs provided in Supplementary Table 2:
https://github.com/saralpulit/Afib-Stroke-Overlap/blob/master/SuppTable4.partB.SiGN.AFGen.trait.correlations.drop-pitx2-zfhx3.txt

The headers of the two files are exactly the same:

Column	Definition
Z.threshold	Z-score threshold used to subset AFGen SNPs
EduYrs.Z	Correlation with z-scores from educational attainment GWAS
afib.broad.Z	Correlation with z-scores from atrial fibrillation (broadly defined phenotype) GWAS
allstroke.Z	Correlation with z-scores from all stroke GWAS
CCScCEmajor.Z	Correlation with z-scores from CCSc CE GWAS
CCScCRYPTCE.Z	Correlation with z-scores from CCSc CRYPTCE GWAS
CCScINCUNC.Z	Correlation with z-scores from CCSc INCUNC GWAS
CCScLAA.Z	Correlation with z-scores from CCSc LAA GWAS
CCScSAO.Z	Correlation with z-scores from CCSc SAO GWAS
CCScUNDETER.Z	Correlation with z-scores from CCSc UNDETER GWAS
CCSpCEmajincl.Z	Correlation with z-scores from CCSp CE GWAS
CCSpCryptoincl.Z	Correlation with z-scores from CCSp Cryptogenic GWAS
CCSpLAAmajincl.Z	Correlation with z-scores from CCSp LAA GWAS
CCSpSAOmajincl.Z	Correlation with z-scores from CCSp SAO GWAS
toastCE.Z	Correlation with z-scores from TOAST CE GWAS
toastLAA.Z	Correlation with z-scores from TOAST LAA GWAS
toastSAO.Z	Correlation with z-scores from TOAST SAO GWAS
toastUNDETER.Z	Correlation with z-scores from TOAST UNDETER GWAS

CCSc, CCS Causative subtyping system; CCSp, CCS Phenotypic subtyping system; TOAST, TOAST subtyping system; CE, cardioembolic stroke; LAA, large artery atherosclerosis; SAO, small artery occlusion; UNDETER, undetermined; INCUNC, incomplete/unclassified; CRYPTCE, cryptogenic and CE minor.

Supplementary Table 5 | Association between the atrial fibrillation polygenic risk score and ischemic stroke subtypes. We tested the association between a polygenic risk score (PRS) constructed from atrial fibrillation-associated SNPs and all stroke subtypes. The results of those association tests are shown here. We used two groups of controls: all available controls (N = 28,026) and all controls that were free of atrial fibrillation (AF, N = 3,860). All analyses were adjusted for sex and principal components (PCs). Regression analyses were optionally adjusted for clinical covariates (age, cardiovascular disease, type 2 diabetes status, smoking status, and hypertension).

Significant results (p = 0.0062, Bonferroni-corrected for four subtype groups and two independent subtyping classifications -- CCS and TOAST -- are bolded).

SE, standard error; CCSc, CCS Causative; CCSp, CCS Phenotypic.

Large artery atherosclerosis (LAA):

Case definition	Control definition	Cases	Controls	Logistic regression, adjusted for PCs and sex			Logistic regression, adjusted for PCs, sex, and clinical covariates		
				Beta	SE	P-value	Beta	SE	P-value
CCSc LAA	Non-AF controls	2,385	3,860	0.008	0.015	0.600	0.002	0.018	0.929
CCSc LAA	All controls	2,385	28,026	-0.002	0.012	0.885	-0.004	0.013	0.786
CCSp LAA	Non-AF controls	2,449	3,860	0.016	0.016	0.315	0.010	0.018	0.570
CCSp LAA	All controls	2,449	28,026	0.004	0.011	0.694	0.002	0.013	0.850
TOAST LAA	Non-AF controls	2,318	3,860	0.010	0.016	0.528	0.000	0.018	0.980
TOAST LAA	All controls	2,318	28,026	-0.006	0.012	0.594	-0.008	0.014	0.550
Results after standardizing PRS to a z-score									
CCSc LAA	Non-AF controls	2,385	3,860	0.016	0.030	0.600	0.003	0.035	0.929
CCSc LAA	All controls	2,385	28,026	-0.003	0.022	0.885	-0.007	0.026	0.786
CCSp LAA	Non-AF controls	2,449	3,860	0.031	0.030	0.315	0.020	0.035	0.570
CCSp LAA	All controls	2,449	28,026	0.009	0.022	0.694	0.005	0.026	0.850
TOAST LAA	Non-AF controls	2,318	3,860	0.019	0.031	0.528	-0.001	0.036	0.980
TOAST LAA	All controls	2,318	28,026	-0.012	0.023	0.594	-0.016	0.027	0.550

Cardioembolic stroke (CE):

Case definition	Control definition	Cases	Controls	Logistic regression, adjusted for PCs and sex			Logistic regression, adjusted for PCs, sex, and clinical covariates		
				Beta	SE	P-value	Beta	SE	P-value
CCSc CE	Non-AF controls	3,000	3,860	0.187	0.014	**1.59E-42**	0.218	0.018	**1.40E-34**
CCSc CE	All controls	3,000	28,026	0.169	0.010	**1.01E-65**	0.173	0.012	**1.45E-48**
CCSp CE	Non-AF controls	3,608	3,860	0.178	0.013	**6.98E-43**	0.203	0.017	**8.34E-34**
CCSp CE	All controls	3,608	28,026	0.161	0.009	**2.43E-70**	0.163	0.011	**1.05E-49**
TOAST CE	Non-AF controls	3,333	3,860	0.171	0.013	**3.17E-37**	0.172	0.015	**3.22E-29**
TOAST CE	All controls	3,333	28,026	0.149	0.009	**3.00E-56**	0.146	0.011	**4.43E-41**
Results after standardizing PRS to a z-score									
CCSc CE	Non-AF controls	3,000	3,860	0.365	0.027	**1.59E-42**	0.425	0.035	**1.40E-34**
CCSc CE	All controls	3,000	28,026	0.329	0.019	**1.01E-65**	0.337	0.023	**1.45E-48**
CCSp CE	Non-AF controls	3,608	3,860	0.348	0.025	**6.98E-43**	0.397	0.033	**8.34E-34**
CCSp CE	All controls	3,608	28,026	0.315	0.018	**2.43E-70**	0.318	0.021	**1.05E-49**
TOAST CE	Non-AF controls	3,333	3,860	0.334	0.026	**3.17E-37**	0.335	0.030	**3.22E-29**
TOAST CE	All controls	3,333	28,026	0.291	0.018	**3.00E-56**	0.284	0.021	**4.43E-41**

Small artery occlusion (SAO):

Case definition	Control definition	Cases	Controls	Logistic regression, adjusted for PCs and sex			Logistic regression, adjusted for PCs, sex, and clinical covariates		
				Beta	SE	P-value	Beta	SE	P-value
CCSc SAO	Non-AF controls	2,262	3,860	0.023	0.017	0.170	0.026	0.019	0.163
CCSc SAO	All controls	2,262	28,026	0.002	0.012	0.842	0.006	0.013	0.660
CCSp SAO	Non-AF controls	2,419	3,860	0.025	0.016	0.124	0.029	0.018	0.109
CCSp SAO	All controls	2,419	28,026	0.003	0.012	0.787	0.007	0.013	0.602
TOAST SAO	Non-AF controls	2,631	3,860	0.021	0.016	0.209	0.019	0.018	0.289
TOAST SAO	All controls	2,631	28,026	0.001	0.011	0.902	0.003	0.013	0.826
Results after standardizing PRS to a z-score									
CCSc SAO	Non-AF controls	2,262	3,860	0.046	0.033	0.170	0.051	0.036	0.163
CCSc SAO	All controls	2,262	28,026	0.005	0.023	0.842	0.012	0.026	0.660
CCSp SAO	Non-AF controls	2,419	3,860	0.049	0.032	0.124	0.057	0.035	0.109
CCSp SAO	All controls	2,419	28,026	0.006	0.023	0.787	0.013	0.025	0.602
TOAST SAO	Non-AF controls	2,631	3,860	0.040	0.032	0.209	0.037	0.035	0.289
TOAST SAO	All controls	2,631	28,026	0.003	0.022	0.902	0.005	0.025	0.826

Undetermined strokes:

Case definition	Control definition	Cases	Controls	Logistic regression, adjusted for PCs and sex			Logistic regression, adjusted for PCs, sex, and clinical covariates		
				Beta	SE	P-value	Beta	SE	P-value
CCSc UNDETER	Non-AF controls	4,574	3,860	0.036	0.013	**0.004**	0.031	0.014	0.022
CCSc UNDETER	All controls	4,574	28,026	0.021	0.009	0.013	0.021	0.010	0.030
CCSc INCUNC	Non-AF controls	2,280	3,860	0.046	0.016	**0.003**	0.045	0.017	0.010
CCSc INCUNC	All controls	2,280	28,026	0.028	0.012	0.015	0.029	0.013	0.025
CCSc CRYPTCE	Non-AF controls	2,294	3,860	0.030	0.016	0.051	0.026	0.017	0.124
CCSc CRYPTCE	All controls	2,294	28,026	0.015	0.012	0.212	0.017	0.013	0.192
CCSp Crypto	Non-AF controls	1,096	3,860	0.035	0.020	0.090	0.029	0.022	0.195
CCSp Crypto	All controls	1,096	28,026	0.019	0.016	0.258	0.021	0.018	0.245
TOAST UNDETER	Non-AF controls	3,479	3,860	0.033	0.013	0.015	0.028	0.014	0.055
TOAST UNDETER	All controls	3,479	28,026	0.021	0.010	0.027	0.022	0.011	0.042
Results after standardizing PRS to a z-score									
CCSc UNDETER	Non-AF controls	4,574	3,860	0.071	0.025	**0.004**	0.061	0.027	0.022
CCSc UNDETER	All controls	4,574	28,026	0.041	0.017	0.013	0.041	0.019	0.030
CCSc INCUNC	Non-AF controls	2,280	3,860	0.090	0.030	**0.003**	0.088	0.034	0.010
CCSc INCUNC	All controls	2,280	28,026	0.055	0.023	0.015	0.056	0.025	0.025
CCSc CRYPTCE	Non-AF controls	2,294	3,860	0.059	0.030	0.051	0.051	0.033	0.124
CCSc CRYPTCE	All controls	2,294	28,026	0.028	0.023	0.212	0.033	0.025	0.192
CCSp Crypto	Non-AF controls	1,096	3,860	0.068	0.040	0.090	0.057	0.044	0.195
CCSp Crypto	All controls	1,096	28,026	0.036	0.032	0.258	0.041	0.035	0.245
TOAST UNDETER	Non-AF controls	3,479	3,860	0.064	0.026	0.015	0.054	0.028	0.055
TOAST UNDETER	All controls	3,479	28,026	0.042	0.019	0.027	0.042	0.021	0.042

UNDETER, undetermined; INCUNC, incomplete and unclassified; CRYPTCE, cryptogenic and CE minor; Crypto, cryptogenic

Supplementary Table 6 | Sensitivity analysis for the atrial fibrillation polygenic risk score. As a sensitivity analysis for the polygenic risk score (PRS), we constructed 3 additional PRSs, including SNPs +/- 25kb, +/- 50kb, and +/- 100kb from the SNPs included in the original score. All scores remain highly significant when tested for association with cardioembolic stroke (using a logistic regression model). P-values after additionally adjusting for clinical covariates are also shown. Clinical covariates: age, cardiovascular disease, type 2 diabetes status, smoking status, and hypertension.

PCs, principal components; MAF, minor allele frequency; INFO, imputation (info) score.

PRS SNPs	**Filters**	**Total SNPs**	**PRS p-value**	
			Adjusted for PCs, sex	*Adjusted for PCs, sex, clinical covariates*
Original SNPs	MAF > 1% Info > 0.8	975	1.01×10^{-65}	1.44×10^{-48}
Original SNPs +/- 25kb	MAF > 1% Info > 0.8	146,631	9.13×10^{-50}	1.32×10^{-37}
Original SNPs +/- 50kb	MAF > 1% Info > 0.8	258,870	5.76×10^{-48}	1.40×10^{-36}
Original SNPs +/- 100kb	MAF > 1% Info > 0.8	462,146	4.47×10^{-44}	1.77×10^{-32}

Supplementary Table 7 | Clinical covariates available in the SiGN data. We adjusted our analyses of a polygenic risk score for a series of clinical covariates that are associated with atrial fibrillation. Summary-statistics on these covariates are shown below for those samples classified as (a) cardioembolic stroke or (b) undetermined stroke. The number of samples with missing data are provided in parentheses where relevant.

Cardioembolic

Phenotype	CCS Causative	CCS Phenotypic	TOAST
Female	1,588	1,859	1,618
Male	1,247	1,541	1,520
Age: mean (sd)	74.7 (12.4)	74.5 (12.3)	71.0 (15.1)
Hypertensive (missing)	2,195 (18)	2,665 (21)	2,272 (16)
Diabetes mellitus (missing)	763 (26)	950 (29)	799 (8)
CAD (missing)	989 (64)	1206 (83)	911 (119)
Smoking			
Current	379	468	513
Former	694	865	776
Never	1,737	2,055	1,905
Total	3,000	3,608	3,333

Undetermined

Phenotype	CCS Causative	CCS Causative	CCS Causative	CCS Phenotypic	TOAST
Female	1,880	1,024	856	420	1,445
Male	2,151	1,014	1,137	543	1,635
Age: mean (sd)	63.9 (15.4)	67.7 (13.9)	69.0 (15.9)	58.9 (15.7)	63.7 (16.1)
Hypertensive (missing)	2,833 (23)	1,512 (14)	1,321 (9)	612 (3)	2,110 (29)
Diabetes mellitus (missing)	958 (26)	513 (14)	445 (12)	202 (4)	708 (25)
CAD (missing)	739 (169)	421 (86)	318 (83)	115 (46)	573 (100)
Smoking					
Current	1,090	582	508	239	813
Former	1,050	516	534	235	772
Never	2,202	1,081	1,121	548	1,711
Total	4,574	2,280	2,294	1,096	3,479

Supplementary Table 8: Variance explained by the atrial fibrillation polygenic risk score in cardioembolic stroke. To determine the variance explained by the atrial fibrillation polygenic risk score (PRS) in cardioembolic stroke, we constructed a model in BOLT-LMM that consisted of two variance components: (1) a variance component made up of SNPs for the genetic relationship matrix, and (2) a variance component made up of SNPs from the PRS. After computing the estimated variance explained for each component in BOLT-LMM, we converted the estimate to the liability score. Below is variance explained for each of the cardioembolic stroke phenotypes as determined by the three subtyping systems available in SiGN: CCS Causative, CCS Phenotypic, and TOAST. Standard errors of each estimate appear in parentheses. Explained variance is shown for a PRS including the *PITX2* (chromosome 4) and *ZFHX3* (chromosome 16) loci, as well as excluding ±2Mb around these loci (see https://github.com/UMCUGenetics/Afib-Stroke-Overlap for lists of SNPs that fall in these regions). Because a large number of SNPs is needed to construct a variance component to calculate variance explained, we performed the calculation using the atrial fibrillation PRS including SNPs ±100kb from the original PRS SNPs, and then pruning SNPs a linkage disequilibrium of 0.2.

CE, cardioembolic; PRS, polygenic risk score; AF, atrial fibrillation

Subtyping System	h_g^2 CE stroke	h_g^2 atrial fibrillation PRS ±100kb	Proportion of CE h_g^2 explained by AF PRS
PRS including the PITX2 and ZFHX3 loci			
CCSc	0.195 (0.019)	0.045 (0.010)	23.1%
CCSp	0.164 (0.016)	0.040 (0.008)	24.4%
TOAST	0.210 (0.018)	0.051 (0.01)	24.3%
PRS excluding the PITX2 and ZFHX3 loci			
CCSc	0.195 (0.019)	0.037 (0.010)	19.0%
CCSp	0.164 (0.016)	0.032 (0.008)	19.5%
TOAST	0.210 (0.018)	0.044 (0.009)	21.0%

Supplementary Methods

GitHub repository and data availability

1. *GitHub repository and additional supporting data*

Relevant code for the analyses performed in this paper can be found here: https://github.com/saralpulit/Afib-Stroke-Overlap.

This repository primarily consists of:

Call to BOLT-LMM to run GWAS

Call to GCTA and BOLT-LMM to calculate heritability

Call to PLINK[3,4] to calculate the polygenic risk score (PRS)

An R script for converting observed heritability in BOLT-LMM to the liability scale (see below)

A script in R to check association between the PRS and various phenotypes.

A call to PLINK[3,4] to calculate a GRM to run GCTA

Sample identifiers for those individuals analyzed in this paper

SNP identifiers and weights for those markers included in the construction of the polygenic risk score

A complete README accompanies the GitHub repository.

2. *Sample and SNP identifiers used in these analyses*

A file containing:

the dbGaP sample identifiers

the cohort the sample is drawn from

the continental group the sample is in (as determined in the first SiGN GWAS effort[5])

a list of quality control-passing SNPs used in the initial GWAS

is available on this paper's GitHub repository.

3. *Downloadable summary-level genome-wide association study data*

The summary-level data from the original SiGN GWAS has been made publicly available through the Cerebrovascular Disease Knowledge Portal, which can be accessed here: http://www.cerebrovascularportal.org/

These summary-level results are available for cardioembolic stroke (CE), large artery atherosclerosis (LAA), small artery occlusion (SAO), and undetermined (UNDETER) stroke, for three different subtyping systems (TOAST, CCS Causative, CCS Phenotypic).

The summary-level results for the atrial fibrillation genome-wide association studies (performed in broadly-defined or strictly-defined cases versus all controls) are available here:

Broadly-defined atrial fibrillation cases vs. all referents:

https://doi.org/10.5281/zenodo.1035871

Strictly-defined atrial fibrillation cases vs. all referents:

https://doi.org/10.5281/zenodo.1035873

The Stroke Genetics Network (SiGN) and genome-wide association study of ischemic stroke subtypes

The full list of cohorts that are included in the SiGN genome-wide association study can be found in the Supplementary Material of "Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study,"[5] which can be downloaded here: https://paperpile.com/shared/nvNXQf.

SiGN is comprised of several case cohorts with pre-existing genotyping data. Newly-collected cases, as well as a small number of matched referents, were genotyped on the Illumina 5M array[6]. The majority of referents included were drawn from publicly-available genotyping data.

1. Referent datasets

Referent datasets downloaded from the Database of Genotypes and Phenotypes (dbGaP) are:

	dbGAP accession #
Genetics Resource with the Health and Retirement Study	phs000428.v2.p2
Whole Genome Association Study of Visceral Adiposity in the HABC study	phs000169.v1.p1

2. Case datasets

A large number of cases and a small number of controls (from Belgium and Poland) were genotyped at the initiation of the SiGN GWAS. These data have been uploaded to dbGaP and are available here:

> The National Institute of Neurological Disorders and Stroke (NINDS) Stroke Genetics Network (SiGN) (phs000615.v1.p1)

3. Phenotyping in SiGN

There are three primary subtype definitions of ischemic stroke: cardioembolic stroke, large artery atherosclerotic stroke, and small artery occlusion. The SiGN consortium used the CCS system to attempt to assign each case to one of these three categories. Additionally, ~74% of cases were also classified using the Trial of Org 10 172 in Acute Stroke Treatment (TOAST)[7,8] system, which classifies stroke cases based on clinical decision-making and clinically-ascertained information. The CCS and TOAST subtyping systems yield moderately-to-strongly correlated phenotyping results (**Supplementary Figure 5**)[9]. Use of these traits in a GWAS setting also yields concordant association results, as previously shown[6]. These subtypes are similarly defined in CCS and TOAST, though determined differently across the two subtyping systems.

In addition to the three primary subtypes, both the CCS and TOAST classification systems generate two additional subtypes: "undetermined" and "other." The "other" classification was small in sample size (N_{cases} = 595, 719 and 374 in CCS Causative, CCS Phenotypic and TOAST, respectively), and was therefore not included in the original SiGN GWAS and was not tested here[6]. The "undetermined" classification, though named the same in CCS and TOAST, is defined differently

across the two subtyping systems[8,10]. In TOAST, patients with conflicting subtype classifications are placed in the undetermined category[6,8]. In contrast, the CCS undetermined classification includes patients with cryptogenic embolism, other cryptogenic cases, patients with an incomplete evaluation, or samples with competing subtypes[10].

4. Brief summary of data quality control in SiGN

SiGN samples represent three continental populations (European-ancestry; African-ancestry; and non-European ancestry and non-African ancestry samples, primarily of admixed ancestry from Latin American populations, labelled 'Hispanic'). In total, the study contains 13 case-referent analysis groups: 10 of European ancestry, two of African ancestry, and one Hispanic[6].

For quality control (QC) and downstream association testing, cases and referents were matched by genotyping array and PCA-determined ancestry. European-ancestry samples were imputed with IMPUTE2[11] using a reference panel built from whole-genome sequence data collected by the 1000 Genomes Project (Phase 1)[12] and the Genome of the Netherlands[13] project; African-ancestry and Hispanic samples were imputed with the 1000 Genomes Project data only.[12] Due to data-sharing restrictions regarding the referents used for the Hispanic set of samples, only the European- and African-ancestry samples were analyzed here, totaling 13,390 cases and 28,026 referents distributed across 12 case-control analysis groups.

Before performing genome-wide association testing, for those SNPs that were genotyped in a subset of the SiGN study strata but imputed in others, we compared the frequency of the SNP across the various strata. We removed any SNP with a frequency difference > 15% within ancestral group or >50% across ancestral groups comparing imputed and genotyped data, likely induced by sequencing errors in the imputation reference panel(s).

Constructing a genetic relationship matrix for genome-wide association testing in BOLT-LMM

To construct the genetic relationship matrix (GRM) implemented in BOLT-LMM, we used SNPs that were (i) common (MAF > 5%), (ii) with missingness < 5%, (iii) linkage disequilibrium (LD) pruned at an r^2 threshold of 0.2, (iv) on the autosomal chromosomes only, (v) and not in stratified areas of the genome (i.e., not in the major histocompatibility complex (MHC), the inversions on chromosomes 8 and 17, or in the lactase (*LCT*) locus on chromosome 2). After association testing, we

additionally removed SNPs with imputation quality (info score) < 0.8, due to excess inflation of the test statistic in those SNPs (**Supplementary Figure 1**).

SNP-based heritability calculations in GCTA and BOLT-LMM

A description of the primary heritability calculations is available in the Materials and Methods section. Briefly, we computed heritability estimates in BOLT-LMM[14] using BOLT-REML. We adjusted all heritability analyses for ten PCs and sex.

To check the robustness of the heritability calculations to the SNPs included in the GRM, we calculated heritability using the GRM described above, as well as three additional GRMs: (i) using the ~1.1M SNPs with imputation quality > 0.8 and MAF > 1% (and without LD pruning); (ii) using the SNPs that were genotyped across all study strata (~155,000 SNPs); and (iii) the set of genotyped SNPs with the MHC, *LCT* locus, inversions on chromosomes 8 and 17 removed, and LD pruned at r^2 = 0.2.

Additionally, we computed heritability in GCTA[2] using the same GRMs and assuming a trait prevalence of 1%. We compared the results to the BOLT-based h_g^2 estimates (**Supplementary Table 3** and **Supplementary Figures 2-3**). As genome-wide heritability estimates need a large number of SNPs to be accurate, we report in the paper all estimates using a GRM containing imputed, pruned SNPs. Estimates resulting from all GRMs are presented here, in the **Supplementary Information**.

To test the effect of chaniging the GRM (referred to by the --bfile and 'modelSNPs' option in BOLT-LMM), we selected SNPs for the GRM in four ways:

(1) Genotyped SNPs only (minor allele frequency > 1%)
115,553 SNPs total

(2) Genotyped SNPs, pruned at a linkage disequilibrium threshold (r2 threshold) of 0.2, and removing the MHC, *LCT* locus, and two chromosomal inversions.
60,432 SNPs total

(3) Imputed SNPs (minor allele frequency > 1% and imputation info > 0.8) converted to best-guess genotypes.
1,128,985 SNPs total

(4) Imputed SNPs (minor allele frequency > 1% and imputation info > 0.8); pruned at a linkage disequilibrium threshold (r2 threshold) of 0.2; removing

the MHC, *LCT* locus, and two chromosomal inversions; and converted to best-guess genotypes.

250,209 SNPs total

The GRM in (4) is the GRM used for all heritability results presented in the main manuscript.

As calculating GRMs in GCTA can be extremely computationally intensive, we calculated the GRMs using PLINK 1.9 and then used those GRMs to estimate heritability. A script that shows how to do this is included in the GitHub repository noted above.

The genomic locations (hg19) for excluded markers are as follows:

The lactase (*LCT*) locus
Chromosome 2, positions 129,883,530 - 140,283,530

The major histocompatibility complex (MHC)
Chromosome 6, positions 24,092,021 - 38,892,022

Inversion 1
Chromosome 8, positions 6,612,592 - 13,455,629

Inversion 2
Chromosome 17, positions 40,546,474 - 44,644,684

All non-autosomal SNPs

BOLT-LMM produces heritability estimates on the observed scale. To convert to the liability scale (i.e., the scale on which GCTA produces heritability estimates) we performed a conversion in R. Running the conversion requires knowing the trait prevalence, total cases analyzed, total controls analyzed, and the heritability on the observed scale. This code snippet is available in the accompanying GitHub repository for this paper.

Quality control in genome-wide data for correlation calculations

We used summary-level data from the latest Atrial Fibrillation Genetics (AFGen) Consortium meta-analysis of atrial fibrillation[1] to calculate a z-score for each SNP in that GWAS. Additionally, we calculated a z-score for each SNP in a GWAS of each stroke subtype in SiGN as well as in the GWAS of atrial fibrillation we performed in the SiGN data. Finally, as a null comparator, we downloaded SNP z-scores from a

GWAS of educational attainment[15] available through LDHub (http://ldsc.broadinstitute.org/, accessed 11-1-2017). We aligned z-score signs based on the risk allele reported in each study. SNPs with an allele frequency difference >5% between AFGen and SiGN (all stroke analysis) were removed from the AFGen data (25,784 SNPs); similarly, SNPs with an allele frequency difference >5% between the educational attainment GWAS and SiGN (all stroke) were also removed (27,866 SNPs). Finally, we calculated Pearson's r between z-scores from two traits to evaluate correlation.

Appendix I

Members of the Atrial Fibrillation Genetics (AFGen) Consortium

Please note that the AFGen Consortium participants evolve over time. Further information on the AFGen Consortium can be found at www.afgen.org.

Ingrid E. Christophersen, MD, PhD[1–3]
Michiel Rienstra, MD, PhD[4]
Carolina Roselli, MSc[1,5,6]
Xiaoyan Yin, PhD[7,8]
Bastiaan Geelhoed, PhD[4]
John Barnard, PhD[9]
Honghuang Lin, PhD[7,8]
Dan E. Arking, PhD[10]
Albert V. Smith, PhD[11,12]
Christine M. Albert, MD, MPH[13]
Mark Chaffin, MSc[1]
Nathan R. Tucker, PhD[1,2]
Molong Li, MD[2]
Derek Klarin, MD[1]
Nathan A Bihlmeyer, BS,[14]
Siew-Kee Low, PhD[15]
Peter E. Weeke, MD, PhD[16,17]
Martina Müller-Nurasyid, PhD[5,18,19]
J. Gustav Smith, MD, PhD[1,20]
Jennifer A. Brody, BA[21]
Maartje N. Niemeijer MD[22]
Marcus Dörr, MD[23,24]
Stella Trompet, PhD[25]
Jennifer Huffman, PhD[26]
Stefan Gustafsson, PhD[27]
Claudia Schurmann, PhD[28,29]
Marcus E. Kleber, PhD[30]
Leo-Pekka Lyytikäinen, MD[31]
Ilkka Seppälä, MD[31]
Rainer Malik, PhD[32]
Andrea R. V. R. Horimoto, PhD[33]
Marco Perez, MD[34]
Juha Sinisalo, MD, PhD[35]
Stefanie Aeschbacher, MSc[36,37]
Sébastien Thériault, MD, MSc[38,39]
Jie Yao, MS[40]
Farid Radmanesh, MD, MPH[1,41]
Stefan Weiss, PhD[24,42]
Alexander Teumer, PhD[24,43]
Seung Hoan Choi, PhD[1]
Lu-Chen Weng, PhD[1,2]
Sebastian Clauss, MD[2,18]
Rajat Deo, MD, MTR[44]
Daniel J. Rader, MD[44]
Svati Shah, MD, MHS,[45]
Albert Sun, MD[45]
Jemma C. Hopewell, PhD[46]
Stephanie Debette, MD, PhD[47–50]
Ganesh Chauhan, PhD[47,48]
Qiong Yang, PhD[51]
Bradford B. Worrall, MD, MSc[52]
Guillaume Paré, MD, MSc[38,39]
Yoichiro Kamatani, MD, PhD[15]
Yanick P. Hagemeijer, MSc[4]
Niek Verweij, PhD[4]
Joylene E. Siland, BSc,[4]
Michiaki Kubo, MD, PhD[53]
Jonathan D. Smith, PhD[9]
David R. Van Wagoner, PhD[9]
Joshua C. Bis, PhD[21]
Siegfried Perz, MSc[54]
Bruce M. Psaty, MD, PhD[21,55–57]
Paul M. Ridker, MD, MPH[13]
Jared W. Magnani, MD, MSc[7,58]
Tamara B. Harris, MD, MS[59]
Lenore J. Launer, PhD[59]
M. Benjamin Shoemaker, MD, MSCI[16]
Sandosh Padmanabhan, MD[60]
Jeffrey Haessler, MS[61]
Traci M. Bartz, MS[62]
Melanie Waldenberger, PhD[19,54,63]
Peter Lichtner, PhD[64]
Marina Arendt, MSc[65]
Jose E. Krieger, MD, PhD[33]
Mika Kähönen, MD, PhD[66]
Lorenz Risch, MD, MPH[67]
Alfredo J. Mansur, MD, PhD[68]
Annette Peters, PhD[19,54,69]
Blair H. Smith, MD[70]
Lars Lind, MD, PhD[71]
Stuart A. Scott, PhD[72]
Yingchang Lu, MD, PhD[28,29]
Erwin B. Bottinger, MD[28,73]
Jussi Hernesniemi, MD, PhD[31,74]
Cecilia M. Lindgren, PhD[75]
Jorge A Wong, MD[76]
Jie Huang, MD, MPH[77]
Markku Eskola, MD, PhD[74]
Andrew P. Morris, PhD[75,78]
Ian Ford, PhD[79]
Alex P. Reiner, MD, MSc[61,80]
Graciela Delgado, MSc[30]
Lin Y. Chen, MD, MS[81]
Yii-Der Ida Chen, PhD[40]
Roopinder K. Sandhu, MD, MPH[82]
Man Li, PhD[83,84]
Eric Boerwinkle, PhD[85]
Lewin Eisele, MD[65]
Lars Lannfelt, MD, PhD[86]
Natalia Rost, MD, MPH, FAAN,[1,87]
Christopher D. Anderson, MD, MMSc[1,41]
Kent D. Taylor, PhD[40]
Archie Campbell, MA,[88]
Patrik K. Magnusson, PhD[89]
David Porteous, PhD[88]
Lynne J. Hocking, PhD[90]
Efthymia Vlachopoulou, PhD[91]
Nancy L. Pedersen, MA, PhD[89]
Kjell Nikus, MD, PhD[74]
Marju Orho-Melander, PhD[92]
Anders Hamsten, MD, PhD[93]
Jan Heeringa, MD, PhD[22]
Joshua C. Denny, MD[16]
Jennifer Kriebel, PhD[54,63,69]
Dawood Darbar, MD[94]
Christopher Newton-Cheh, MD, MPH[1,2]
Christian Shaffer, BS,[16]
Peter W. Macfarlane, PhD, DSc[95]
Stefanie Heilmann, PhD[96,97]
Peter Almgren, MSc[92]
Paul L. Huang, MD, PhD[2]
Nona Sotoodehnia, MD, MPH[98]
Elsayed Z. Soliman, MD, MSc, MS[99]
Andre G. Uitterlinden, PhD[100]
Albert Hofman, MD, PhD[22]
Oscar H. Franco, MD, PhD[22]
Uwe Völker, PhD[24,42]
Karl-Heinz Jöckel, PhD[65]
Moritz F. Sinner, MD, MPH[18,19]
Henry J. Lin, MD[40]
Xiuqing Guo, PhD[40]
Martin Dichgans, MD[32,101,102]
Erik Ingelsson, MD, PhD[27,103]
Charles Kooperberg, PhD[61]
Olle Melander, MD, PhD[104]
Ruth J. F. Loos, PhD[28,29,105]
Jari Laurikka, MD, PhD[106]
David Conen, MD, MPH[36–38]
Jonathan Rosand, MD, MSc[1,41]
Pim van der Harst, MD, PhD[4]
Marja-Liisa Lokki, PhD[91]
Sekar Kathiresan, MD[1]
Alexandre Pereira, MD, PhD[107]
J. Wouter Jukema, MD, PhD[25,108,109]
Caroline Hayward, PhD[26]
Jerome I. Rotter, MD[110]
Winfried März, MD[111]
Terho Lehtimäki, MD, PhD[31]
Bruno H. Stricker, MD, PhD[112]
Mina K. Chung, MD[9]
Stephan B. Felix, MD[23,24]
Vilmundur Gudnason, MD, PhD[11,12]
Alvaro Alonso, MD, PhD[113]
Dan M. Roden, MD[16]
Stefan Kääb, MD, PhD[18,19]
Daniel I. Chasman, PhD[1,114]
Susan R. Heckbert, MD, PhD[55,56]
Emelia J. Benjamin, MD, ScM[7,58,115]
Toshihiro Tanaka, MD, PhD[116,117]
Kathryn L. Lunetta, PhD[7,8]
Steven A. Lubitz, MD, MPH[1,2,118]
Patrick T. Ellinor, MD, PhD[1,2,118]

AFGen Consortium Member Affiliations

1. Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA, USA.
2. Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA.
3. Department of Medical Research, Bærum Hospital, Vestre Viken Hospital Trust, Norway.
4. Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.
5. Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
6. Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany.
7. NHLBI and Boston University's Framingham Heart Study, Framingham, MA, USA.
8. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.
9. Departments of Cardiovascular Medicine, Cellular and Molecular Medicine, Molecular Cardiology, and Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH, USA.
10. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
11. Icelandic Heart Association, Kopavogur, Iceland.
12. Faculty of Medicine, University of Iceland, Reykavik, Iceland.
13. Divisions of Preventive and Cardiovascular Medicine, Brigham and Women's Hospital & Harvard Medical School, Boston, MA, USA.
14. Predoctoral Training Program in Human Genetics, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
15. Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
16. Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA.
17. The Heart Centre, Department of Cardiology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark.
18. Department of Medicine I, University Hospital Munich, Ludwig-Maximilians-University, Munich, Germany.
19. DZHK (German Centre for Cardiovascular Research), partner site: Munich Heart Alliance, Munich, Germany.
20. Molecular Epidemiology and Cardiology, Clinical Sciences, Lund University, Lund, Sweden.
21. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA.
22. Department of Epidemiology, Erasmus University Medical Center Rotterdam, Rotterdam, the Netherlands.
23. Department of Internal Medicine B, University Medicine Greifswald, Greifswald, Germany.
24. DZHK (German Centre for Cardiovascular Research), partner site: Greifswald, Germany.
25. Department of Cardiology, Leiden University Medical Center, The Netherlands.
26. MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, UK.
27. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden.
28. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
29. The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
30. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Germany.
31. Department of Clinical Chemistry, Fimlab Laboratories and University of Tampere School of Medicine, Tampere, Finland.
32. Institute for Stroke and Dementia Research, Klinikum der Universität München, Ludwig-Maximilians University, München, Germany.
33. Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil.
34. Stanford University, Stanford, CA, USA.
35. Heart and Lung Center HUS, Helsinki University Central Hospital, Helsinki, Finland.
36. University Hospital Basel, Switzerland.
37. Cardiovascular Research Institute Basel, Switzerland.
38. Population Health Research Institute, Hamilton, Canada.
39. Department of Pathology and Molecular Medicine, McMaster University, Hamilton, Canada.
40. Institute for Translational Genomics and Population Sciences, Department of Pediatrics, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
41. Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA.
42. Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany.
43. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany.
44. Division of Cardiovascular Medicine, Department of Medicine, Perelman School of Medicine at the

University of Pennsylvania, Philadelphia, PA, USA.
45. Division of Cardiology, Department of Medicine, Duke University School of Medicine, Durham, NC, USA.
46. CTSU - Nuffield Department of Population Health, University of Oxford, Oxford, UK.
47. Inserm Center U1219 (Bordeaux Population Health Centre), Bordeaux, France.
48. University of Bordeaux, Bordeaux, France.
49. Department of Neurology, Bordeaux University Hospital, Bordeaux, France.
50. Department of Neurology, Boston University School of Medicine, Boston, MA, USA.
51. Biostatistics Department, School of Public Health, Boston University, Boston, MA, USA.
52. University of Virginia Health System, Departments of Neurology and Public Health Science, Charlottesville, VA, USA.
53. RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
54. Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
55. Department of Epidemiology and Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA.
56. Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA.
57. Department of Health Services, University of Washington, Seattle, WA, USA.
58. Department of Medicine, Boston University School of Medicine, Boston, MA, USA.
59. Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Bethesda, MD, USA.
60. Institute of Cardiovascular and Medical Sciences, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, UK.
61. Fred Hutchinson Cancer Research Center, University of Washington, Seattle, WA, USA.
62. Cardiovascular Health Research Unit, Departments of Medicine and Biostatistics, University of Washington, Seattle, WA, USA.
63. Research unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
64. Institute of Human Genetics, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
65. Institute for Medical Informatics, Biometry, and Epidemiology, University Hospital, University Duisburg-Essen, Germany.
66. Department of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
67. University Institute of Clinical Chemistry, University of Bern, Switzerland and labormedizinisches zentrum Dr. Risch, Schaan, Liechtenstein.
68. Heart Institute, University of Sao Paulo, Sao Paulo, Brazil.
69. German Center for Diabetes Research, Neuherberg, Germany.
70. Division of Population Health Sciences, University of Dundee, Scotland, UK.
71. Department of Medical Sciences, Cardiovascular Epidemiology, Uppsala University, Uppsala, Sweden.
72. Department of Genetics and Genomic Sciences , Icahn School of Medicine at Mount Sinai, New York, NY, USA.
73. Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
74. Department of Cardiology, Heart Hospital, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
75. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
76. Division of Cardiology, Hamilton Health Sciences, McMaster University, Hamilton, Ontario, Canada.
77. Boston VA Research Institute, Inc., Boston, MA, USA.
78. Department of Biostatistics, University of Liverpool, Liverpool, UK.
79. Robertson Center for Biostatistics, University of Glasgow, Glasgow, UK.
80. Department of Epidemiology, University of Washington, Seattle, WA, USA.
81. Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, MN, USA.
82. Division of Cardiology, University of Alberta, Edmonton, Canada.
83. Department of Epidemiology, Johns Hopkins University, Baltimore, MD, USA.
84. Division of Nephrology & Hypertension, Internal Medicine, School of Medicine, University of Utah, UT, USA.
85. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA.
86. Department of Public Health and Caring Sciences, Geriatrics, Uppsala University, Uppsala, Sweden.
87. Acute Stroke Services, Massachusetts General Hospital, Boston, MA, USA.
88. Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, UK.
89. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
90. Musculoskeletal Research Programme, Division of Applied Medicine, University of Aberdeen, Aberdeen, UK.
91. Transplantation Laboratory, Medicum, University of Helsinki, Helsinki, Finland.
92. Department of Clinical Sciences, Lund University, Malmö, Sweden.
93. Cardiovascular Genetics and Genomics Group, Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden.

94. University of Illinois, Chicago, IL, USA.
95. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, UK.
96. Institute of Human Genetics, University of Bonn, Germany.
97. Department of Genomics, Life & Brain Research Center, University of Bonn, Germany.
98. Cardiovascular Health Research Unit, University of Washington Medical Center, Seattle, WA, USA.
99. Epidemiological Cardiology Research Center (EPICARE), Wake Forest School of Medicine, Winston Salem, NC, USA.
100. Department of Epidemiology and Internal Medicine, Erasmus University Medical Center Rotterdam, the Netherlands.
101. Munich Cluster for Systems Neurology (SyNergy), München, Germany.
102. German Center for Neurodegenerative Diseases (DZNE), Munich, Germany.
103. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA.
104. Department of Internal Medicine, Clinical Sciences, Lund University, Malmö, Sweden.
105. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
106. Department of Cardio-Thoracic Surgery, Heart Hospital, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland.
107. Laboratory of Genetics and Molecular Biology, Heart Institute, University of Sao Paulo, Sao Paulo, Brazil and Department of Genetics, Harvard Medical School, Boston, MA, USA.
108. Durrer Center for Cardiogenetic Research, Amsterdam, The Netherlands.
109. Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands.
110. Institute for Translational Genomics and Population Sciences, Departments of Pediatrics and Medicine, LABioMed at Harbor-UCLA Medical Center, Torrance, CA, USA.
111. Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria and Synlab Academy, Synlab Services GmbH, Mannheim, Germany.
112. Department of Epidemiology and Internal Medicine, Erasmus University Medical Center Rotterdam, the Netherlands and Inspectorate of Health Care, Utrecht, the Netherlands.
113. Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA.
114. Divisions of Preventive Medicine and Genetics, Brigham and Women's Hospital & Harvard Medical School, Boston, MA, USA.
115. Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.
116. Laboratory for Cardiovascular Diseases, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan.
117. Department of Human Genetics and Disease Diversity, Tokyo Medical and Dental University Graduate School of Medical and Dental Sciences, Tokyo, Japan.
118. Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, MA, USA.

Appendix II

Members of the International Stroke Genetics Consortium (ISGC)

Please note that ISGC participants evolve over time. Further information on the ISGC can be found at http://www.strokegenetics.org/.

Sylvia Smoller, PhD[1]
John Sorkin, MD[2]
Xingwu Wang, MD[3]
Magdy Selim, MD, PhD[4]
Aleksandra Pikula, MD, PhD[5]
Philip Wolf, MD, PhD[5]
Stephanie Debette, MD[5]
Sudha Seshadri, MD[5]
Paul de Bakker, PhD[6]
Sara L. Pulit, PhD[6]
Daniel Chasman, MD[7]
Kathryn Rexrode, MD[7]
Ida Chen, MD[8]
Jerome Rotter, MD[8]
May Luke, MD[9]
Michelle Sale, MD[10]
Tsong-Hai Lee, MD[11]
Ku-Chou Chang, MD[11]
Mitchell Elkind, MD, MS[12]
Larry Goldstein, MD, PhD[13]
Michael Luke James, MD[13]
Monique Breteler, MD[14]
Chris O'Donnell, MD[15]
Didier Leys, MD[16]
Cara Carty, MD[17]
Chelsea Kidwell, MD[18]
Jes Olesen, MD[19]
Pankaj Sharma, MD, PhD[20]
Stephen Rich, MD, PhD[21]
Turgot Tatlisumak, MD[22]
Olli Happola, MD[22]
Philippe Bijlenga, MD[23]
Carolina Soriano, MD[24]
Eva Giralt, MD[24]
Jaume Roquer, MD[24]
Jordi Jimenez-Conde, MD[24]
Ioana Cotlarcius, MD[25]
John Hardy, MD[26]
Michal Korostynski, MD[27]
Giorgio Boncoraglio, MD[28]
Elena Ballabio, MD[28]
Eugenio Parati, MD[28]
Adamski Mateusz, MD[29]
Andrzej Urbanik, MD[29]
Tomasz Dziedzic, MD[29]
Jeremiasz Jagiella, MD[29]
Jerzy Gasowski, MD[29]
Marcin Wnuk, MD[29]
Rafael Olszanecki, MD[29]
Joanna Pera, MD[29]
Agnieszka Slowik, MD[29]
Karol Jozef Juchniewicz, MD[29]
Christopher Levi, MD[30]
Paul Nyquist, MD, PhD[31]
Iscia Cendes, MD[32]
Norberto Cabral, MD[32]
Paulo Franca, MD[32]
Anderson Goncalves, MD[32]
Lina Keller, MD[33]
Milita Crisby, MD[33]
Konstantinos Kostulas, MD[33]
Robin Lemmens, MD[34]
Kourosh Ahmadi, MD[34]
Christian Opherk, MD[35]
Marco Duering, MD[35]
Martin Dichgans, MD[35]
Rainer Malik, PhD[35]
Mariya Gonik, MD[35]
Julie Staals, MD[36]
Olle Melander, MD, PhD[37]
Philippe Burri, MD[37]
Ariane Sadr-Nabavi, MD[38]
Javier Romero, MD, PhD[39]
Alessandro Biffi, MD[39]
Chris Anderson, MD[39]
Guido Falcone, MD[39]
Bart Brouwers, MD[39]
Jonathan Rosand, MD, MSc[39]
Natalia Rost, MD, MSc[39]
Rose Du, MD[39]
Christina Kourkoulis, BA[39]
Thomas Battey, BA[39]
Steven Lubitz, MD, PhD[39]
Bertram Mueller-Myhsok, MD[40]
James Meschia, MD[40]
Thomas Brott, MD, PhD[41]
Guillaume Pare, MD[42]
Alexander Pichler, MD[43]
Christian Enzinger, MD[43]
Helena Schmidt, MD[43]
Reinhold Schmidt, MD[43]
Stephan Seiler, MD[43]
Susan Blanton, MD[44]
Yoshiji Yamada, MD[45]
Anna Bersano, MD[46]
Tatjana Rundek, MD[47]
Ralph Sacco, MD[47]
Yu-Feng Yvonne Chan, MD[48]
Andreas Gschwendtner, MD, PhD[35]
Zhen Deng, MD[49]
Taura Barr, MD[50]
Katrina Gwinn, MD[50]
Roderick Corriveau, MD[50]
Andrew Singleton, MD, PhD[50]
Salina Waddy, MD[50]
Lenore Launer, MD[50]
Christopher Chen, MD[51]
Kim En Le, MD[51]
Wei Ling Lee, MD[51]
Eng King Tan, MD[51]
Akintomi Olugbodi, MD[52]
Peter Rothwell, MD, PhD[53]
Sabrina Schilling, MD[54]
Vincent Mok, MD[55]
Elena Lebedeva, MD[56]
Christina Jern, MD[57]
Katarina Jood, MD[57]
Sandra Olsson, MD[57]
Helen Kim, MD[58]
Chaeyoung Lee, MD[59]
Laura Kilarski, MD[60]
Hugh Markus, MD[60]
Jennifer Peycke, MD[60]
Steve Bevan, PhD[60]
Wayne Sheu, MD[61]
Hung Yi Chiou, MD[62]
Joseph Chern, MD[62]
Elias Giraldo, MD[63]
Muhammad Taqi, MD[63]
Vivek Jain, MD[64]
Olivia Lam, MD[65]
George Howard, MD[66]
Daniel Woo, MD[67]
Steven Kittner, MD[68]
Braxton Mitchell, PhD, MPH[68]
John Cole, MD[68]
Jeff O'Connell, MD[68]
Dianna Milewicz, MD[69]
Kachikwu Illoh, MD[70]
Bradford Worrall, MD[21]
Colin Stine, MD[70]
Bartosz Karaszewski, MD[71]
David Werring, MD[71]
Reecha Sofat, MD[71]

June Smalley, MD[71]
Arne Lindgren, MD[72]
Bjorn Hansen, BA[72]
Bo Norrving, MD[72]
Gustav Smith, MD[72]
Juan Jose Martin, MD[73]
Vincent Thijs, MD[74]
Karin Klijn, MD[75]
Femke van't Hof, MD, PhD[75]
Ale Algra, MD[75]
Mary Macleod, MD[76]
Rodney Perry, MD[77]
Donna Arnett, MD[77]
Alessandro Pezzini, MD[78]
Alessandro Padovani, MD[78]
Steve Cramer, MD, PhD[79]
Mark Fisher, MD[79]
Danish Saleheen, MD[80]
Joseph Broderick, MD[81]
Brett Kissela, MD[81]
Alex Doney, MD[82]
Cathie Sudlow, MD[83]
Kristiina Rannikmae, MD[83]
Scott Silliman, MD[84]
Caitrin McDonough, MD[84]
Matthew Walters, MD[85]
Annie Pedersen, MD[86]
Kazuma Nakagawa, MD[87]
Christy Chang, MD[88]
Mark Dobbins, MD[88]
Patrick McArdle, PhD[88]
Yu-Ching Chang, MD[88]
Robert Brown, MD[89]
Devin Brown, MD[89]
Elizabeth Holliday, MD[90]
Raj Kalaria, MD[91]
Jane Maguire, MD[91]
John Attia, MD[91]
Martin Farrall, MD[92]
Anne-Katrin Giese, MD[93]
Myriam Fornage, MD[94]
Jennifer Majersik, MD[95]
Mary Cushman, MD[96]
Keith Keene, MD[97]
Siiri Bennett, MD[98]
David Tirschwell, MD, MSc[98]
Bruce Psaty, MD[98]
Alex Reiner, MD[98]
Will Longstreth, MD[99]
David Spence, MD[100]
Joan Montaner, MD[101]
Israel Fernandez-Cadenas, MD[102]
Carl Langefeld, MD[102]
Cheryl Bushnell, MD[102]
Laura Heitsch, MD[103]
Jin-Moo Lee, MD, PhD[103]
Kevin Sheth, MD[104]

ISGC Consortium Member Affiliations

1. Albert Einstein College of Medicine, Bronx, NY, USA
2. Baltimore VA Medical Center, Baltimore, MD, USA
3. Beijing Hypertension League Institute, Beijing, China
4. Beth Israel Deaconess Medical Center, Boston, MA, USA
5. Boston University Medical Center, Boston, MA, USA
6. University Medical Center Utrecht, Utrecht, The Netherlands
7. Brigham and Women's Hospital, Boston, MA, USA
8. Cedars Sinai Medical Center, Los Angeles, CA, USA
9. Celera, Alameda, CA, USA
10. University of Virginia, Charlottesville, VA, USA
11. Chang Gung Memorial Hospital, Linkou Medical Center, Guishan District, Taoyuan City, Taiwan
12. Columbia University, New York, NY, USA
13. Duke University, Durham, NC, USA
14. Erasmus University, Rotterdam, Zuid Holland, The Netherlands
15. Framingham Heart Study, Framingham, MA, USA
16. Université du Droit et de la Santé Lille, Lille, France
17. Fred Hutchinson Cancer Research Center, Seattle, WA, USA
18. Georgetown University, Georgetown, MD, USA
19. Glostrup Hospital, Glostrup, Denmark
20. Hammersmith Hospitals & Imperial College London, London, UK
21. University of Virginia Health System, Charlottesville, VA, USA
22. Helsinki University Central Hospital, Helsinki, Finland
23. Hìpitaux Universityersitaires de Genäve, Geneva, Switzerland
24. IMIM-Hospital del Mar, Barcelona, Spain
25. Imperial College London, London, UK
26. Institute of Neurology, University College London, London, UK
27. Institute of Pharmacology, Krakow, Poland
28. IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
29. Jagiellonian University, Krakow, Poland
30. John Hunter Hospital, University of Newcastle, Newcastle, New-South-Wales, Australia
31. Johns Hopkins School of Medicine, Baltimore, MD, USA
32. Joinville Biobank, Joinville, Brazil
33. Karolinska Institutet, Karolinska, Sweden
34. Leuven University, Leuven, Belgium

35. Ludwig-Maximilians-Universitat, Munchen, Germany
36. Maastricht University Medical Centre, Maastricht, the Netherlands
37. Malmo University Hospital, Malmo, Sweden
38. Mashhad University of Medical Sciences, Masshad, Iran
39. Massachusetts General Hospital, Boston, MA, USA
40. Max Planck Institute of Psychiatry, Munich, Germany
41. Mayo Clinic, Rochester, MN, USA
42. McMaster University, Hamilton, Canada
43. Medical University Graz, Graz, Austria
44. Miami Institute of Human Genomics, University of Miami Miller School of Medicine, Miami, FL
45. Mie University, Tsu, Japan
46. Milan University, Milan, Italy
47. University of Miami, Miami, FL, USA
48. Mount Sinai Medical Center, Miami Beach, FL, USA
49. Nanfang Hospital, Southern Medical University, Guangdong, China
50. National Institutes of Health, Bethesda, MD, USA
51. National Neuroscience Institute, Singapore General Hospital, Singapore
52. Obafemi Awolowo University, Ile-Ife, Nigeria
53. Radcliffe Infirmary, Oxford University, Oxford, UK
54. University of Bordeaux, Bordeaux, France
55. Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong
56. Ulm University, Ulm, Germany
57. Sahlgrenska University Hospital, Gothenburg, Sweden
58. Center for Cerebrovascular Research, San Francisco General Hospital, San Francisco, CA, USA
59. Soongsil University, Seoul, South Korea
60. St. George's University of London, London, UK
61. Taichung Veterans General Hospital, Taichung City, Taiwan
62. Taipei Medical University, Taipei City, Taiwan
63. The University of Tennessee Health Science Center at Memphis, Memphis, TN, USA
64. University of California Irvine Medical Center, Irvine, CA, USA
65. University of California San Francisco, San Francisco, CA, USA
66. University of Alabama School of Public Health
67. University of Cincinnati, Cincinnati, OH, USA
68. University of Maryland School of Medicine, Baltimore, MD, USA
69. University of Texas Medical School at Houston, Houston, TX, USA
70. University of Texas-Houston, Houston, TX, MA
71. University College London, London, UK
72. University Hospital Lund, Lund, Sweden
73. University Hospital Sanatorio Allende, Cordoba, Argentina
74. University Hospital Leuven, Leuven, Belgium
75. University Medical Center Utrecht, Utrecht, The Netherlands
76. University of Aberdeen, Aberdeen, Scotland
77. University of Alabama at Birmingham School of Public Health, Birmingham, AL, USA
78. University of Brescia, Brescia, Italy
79. University of California Irvine, Irvine, CA, USA
80. University of Pennsylvania, Philadelphia, PA, USA
81. University of Cincinnati, Cincinnati, OH, USA
82. University of Dundee, Dundee, Scotland
83. University of Edinburgh, Western General Hospital, Edinburgh, Scotland
84. University of Florida, Gainesville, FL, USA
85. University of Glasgow, Glasgow, Scotland, UK
86. University of Gothenburg, Gothenburg, Sweden
87. University of Hawaii, Honolulu, HI, USA
88. University of Maryland, Baltimore, MD, USA
89. University of Newcastle, New-South-Wales, Australia
90. Wellcome Trust Center for Human Genetics, University of Oxford, Oxford, UK
91. University of Rostock, Rostock, Germany
92. University of Texas-Houston, Health Sciences Center, Houston, TX, USA

93. University of Utah, Salt Lake City, UT, USA
94. University of Vermont and Fletcher Allen Health Care, Burlington, VT, USA
95. University of Virginia, Charlottesville, VA, USA
96. University of Washington, Seattle, WA, USA
97. University of Washington, Harborview Medical Center, Seattle, WA, USA
98. University of Western Ontario, Robarts Research Institute, Ontario, Canada
99. Vall d'Hebron Hospital, Barcelona, Spain
100. Wake Forest University, Winston-Salem, NC, USA
101. Washington University of St. Louis, St. Louis, MO, USA
102. Yale New Haven Hospital, Yale School of Medicine, Yale, CT, USA

References

1. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J, Lin H, Arking DE, Smith AV, Albert CM, Chaffin M, Tucker NR, Li M, Klarin D, Bihlmeyer NA, Low S-K, Weeke PE, Müller-Nurasyid M, Smith JG, Brody JA, Niemeijer MN, Dörr M, Trompet S, Huffman J, Gustafsson S, Schurmann C, Kleber ME, Lyytikäinen L-P, Seppälä I, Malik R, Horimoto ARVR, Perez M, Sinisalo J, Aeschbacher S, Thériault S, Yao J, Radmanesh F, Weiss S, Teumer A, Choi SH, Weng L-C, Clauss S, Deo R, Rader DJ, Shah SH, Sun A, Hopewell JC, Debette S, Chauhan G, Yang Q, Worrall BB, Paré G, Kamatani Y, Hagemeijer YP, Verweij N, Siland JE, Kubo M, Smith JD, Van Wagoner DR, Bis JC, Perz S, Psaty BM, Ridker PM, Magnani JW, Harris TB, Launer LJ, Shoemaker MB, Padmanabhan S, Haessler J, Bartz TM, Waldenberger M, Lichtner P, Arendt M, Krieger JE, Kähönen M, Risch L, Mansur AJ, Peters A, Smith BH, Lind L, Scott SA, Lu Y, Bottinger EB, Hernesniemi J, Lindgren CM, Wong JA, Huang J, Eskola M, Morris AP, Ford I, Reiner AP, Delgado G, Chen LY, Chen Y-DI, Sandhu RK, Li M, Boerwinkle E, Eisele L, et al. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nat Genet*. 2017;57:289.

2. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.

3. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M a. R, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.

4. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:1–16.

5. NINDS Stroke Genetics Network (SiGN), International Stroke Genetics Consortium. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol*. 2015;15:4–7.

6. Pulit SL, McArdle PF, Wong Q, Malik R, Gwinn K, Achterberg S, Algra A, Amouyel P, Anderson CD, Arnett DK, Arsava EM, Attia J, Ay H, Bartz TM, Battey T, Benavente OR, Bevan S, Biffi A, Bis JC, Blanton SH, Boncoraglio GB, Brown Jr RD, Burgess AI, Carrera, Smith SNC, Chasman DI, Chauhan G, Chen WM, Cheng YC, Chong M, Cloonan LK, Cole JW, Cotlarciuc I, Cruchaga C, Cuadrado-Godia E, Dave T, Dawson J, Debette S, Delavaran H, Dell CA, Dichgans M, Doheny KF, Dong C, Duggan DJ, Engström G, Evans MK, Pallejà XE, Faul JD,Fernández-Cadenas I, Fornage M, Frossard PM, Furie K, Gamble DM, Gieger C, Giese AK, Giralt-Steinhauer E, González HM, Goris A, Gretarsdottir S, Grewal RP, Grittner R, Gustafsson S, Han B, Hankey GJ, Heitsch L, Higgins P, Hochberg MC, Holliday E, Hopewell EC, Horenstein RB, Howard G, Ikram MA, Ilinca A, Ingelsson E, Irvin MR, Jackson RD, Jern C, Conde JJ, Johnson JA, Jood K, Kahn MS, Kaplan R, Kappelle LJ, Kardia SLR, Keene KL, Kissela BM, O Kleindorfer D, Koblar S, Labovitz D, Launer LJ, Laurie CC, Laurie CA, Lee CH, Lee JM, Lehm M, Lemmens R, Levi C, Leys D, Lindgren A, Longstreth Jr WT, Maguire J, Manichaikul Ani, Markus HS, McClure LA, McDonough CW, Meisinger C, Melander O, Meschia JF, Mola-Caminal M, Montaner J, Mosley TH, Müller-Nurasyid M, Nalls MA, O'Connell JR, O'Donnell M, Ois A, Papanicolaou GJ, Paré G, Peddareddygari LR, Pedersén A, Pera J, Peters A, Poole D, Psaty BM, Rabionet R, Raffeld MR, Rannikmäe K, Rasheed A, Redfors P, Reiner AP, Rexrode K, Ribasés M, Rich SS, Robberecht W, Rodriguez-Campello A, Rolfs A, Roquer J, Rose LM,

Rosenbaum D, Rost NS, Rothwell PM, Rundek T, Ryan KA, Sacco RL, Sale MM, Saleheen D, Salomaa V, Sánchez-Mora C, Schmidt CO, Schmidt H, Schmidt R, Schürks M, Scott R, Segal HC, Seiler S, Seshadri S, Sharma P, Shuldiner AR, Silver B, Slowik A, Smith JA, Söderholm M, Soriano C, Sparks MJ, Stanne T, Stefansson K, Stine OC, Strauch K, Sturm J, Sudlow CLM, Tajuddin SM, Talbert RL, Tatlisumak T, Thijs V, Thorleifsson G, Thorsteindottir U, Tiedt S, Traylor M, Trompet S, Valant V, Waldenberger M, Walters M, Wang L, Wassertheil-Smoller S, Weir DR, Wiggins KL, Williams SR, Wloch-Kopec D, Woo D, Woodfield D, Wu O, Xu H, Zonderman AB, Australian Stroke Genetics Consortium, Cervical Artery Dissection and Ischemic Stroke Patients (CADISP) study, Cohorts of Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium, Consortium of Minority Population genome-wide Association Studies of Stroke (COMPASS), METASTROKE consortium, Wellcome Trust Case-Control Consortium, Worrall BB, de Bakker PIW, Kittner SJ, Mitchell BD, Rosand J. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurology*. 2015;15:174–184.

7. Goldstein LB, Jones MR, Matchar DB, Edwards LJ, Hoff J, Chilukuri V, Armstrong SB, Horner RD. Improving the reliability of stroke subgroup classification using the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) criteria. *Stroke*. 2001;32:1091–1098.

8. Adams HP Jr, Birgitte H. Bendixen P, Jaap Kappelle L, Biller J, Love BB, Gordon DL, Eugene MarshIII E, and the TOAST Investigators. Classification of Subtype of Acute Ischemic Stroke: Definitions for Use in a Multicenter Clinical Trial. *Stroke*. 1993;24:35–41.

9. McArdle PF, Kittner SJ, Ay H, Brown RD, Meschia JF, Rundek T, Wassertheil-Smoller S, Woo D, Andsberg G, Biffi A, Brenner DA, Cole JW, Corriveau R, de Bakker PIW, Delavaran H, Dichgans M, Grewal RP, Gwinn K, Huq M, Jern C, Jimenez-Conde J, Jood K, Kaplan RC, Katschnig P, Katsnelson M, Labovitz DL, Lemmens R, Li L, Lindgren A, Markus HS, Peddareddygari LR, Pedersén A, Pera J, Redfors P, Roquer J, Rosand J, Rost NS, Rothwell PM, Sacco RL, Sharma P, Slowik A, Sudlow C, Thijs V, Tiedt S, Valenti R, Worrall BB. Agreement between TOAST and CCS ischemic stroke classification: the NINDS SiGN study. *Neurology*. 2014;83:1653–1660.

10. Ay H, Benner T, Arsava EM, Furie KL, Singhal AB, Jensen MB, Ayata C, Towfighi A, Smith EE, Chong JY, Koroshetz WJ, Sorensen AG. A computerized algorithm for etiologic classification of ischemic stroke: The causative classification of stroke system. *Stroke*. 2007;38:2979–2984.

11. 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;44:955–959.

12. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB, Gibbs RA, Green ED, Hurles ME, Knoppers BM, Korbel JO, Lander ES, Lee C, Lehrach H, Mardis ER, Marth GT, McVean GA, Nickerson DA, Schmidt JP, Sherry ST, Wang J, Wilson RK, Boerwinkle E, Doddapaneni H, Han Y, Korchina V, Kovar C, Lee S, Muzny D, Reid JG, Zhu Y, Chang Y, Feng Q, Fang X, Guo X, Jian M, Jiang H, Jin X, Lan T, Li G, Li J, Li Y, Liu S, Liu X, Lu Y, Ma X, Tang M, Wang B, Wang G, Wu H, Wu R, Xu X, Yin Y, Zhang D, Zhang W, Zhao J, Zhao M, Zheng X, Gupta N, Gharani N, Toji LH, Gerry NP, Resch AM, Barker J, Clarke L, Gil L, Hunt SE, Kelman G, Kulesha E, Leinonen R, McLaren WM, Radhakrishnan R, Roa A, Smirnov D, Smith RE, Streeter I, Thormann A, Toneva I, Vaughan B, Zheng-Bradley X, Grocock R,

Humphray S, James T, Kingsbury Z, Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Lienhard M, Mertes F, Sultan M, Timmermann B, Yaspo M-L, et al. A global reference for human genetic variation. *Nature*. 2015;526:68–74.

13. Francioli LC, Menelaou A, Pulit SL, van Dijk F, Palamara PF, de Bakker PIW, Swertz MA, Wijmenga C. Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet*. 2014;46:818–825.

14. Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, Chasman DI, Ridker PM, Neale BM, Berger B, Patterson N, Price AL. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet*. 2015;47:284–290.

15. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen G-B, Emilsson V, Meddens SFW, Oskarsson S, Pickrell JK, Thom K, Timshel P, de Vlaming R, Abdellaoui A, Ahluwalia TS, Bacelis J, Baumbach C, Bjornsdottir G, Brandsma JH, Pina Concas M, Derringer J, Furlotte NA, Galesloot TE, Girotto G, Gupta R, Hall LM, Harris SE, Hofer E, Horikoshi M, Huffman JE, Kaasik K, Kalafati IP, Karlsson R, Kong A, Lahti J, Lee SJ van der, deLeeuw C, Lind PA, Lindgren K-O, Liu T, Mangino M, Marten J, Mihailov E, Miller MB, van der Most PJ, Oldmeadow C, Payton A, Pervjakova N, Peyrot WJ, Qian Y, Raitakari O, Rueedi R, Salvi E, Schmidt B, Schraut KE, Shi J, Smith AV, Poot RA, St Pourcain B, Teumer A, Thorleifsson G, Verweij N, Vuckovic D, Wellmann J, Westra H-J, Yang J, Zhao W, Zhu Z, Alizadeh BZ, Amin N, Bakshi A, Baumeister SE, Biino G, Bønnelykke K, Boyle PA, Campbell H, Cappuccio FP, Davies G, De Neve J-E, Deloukas P, Demuth I, Ding J, Eibich P, Eisele L, Eklund N, Evans DM, Faul JD, Feitosa MF, Forstner AJ, Gandin I, Gunnarsson B, Halldórsson BV, Harris TB, Heath AC, Hocking LJ, Holliday EG, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*. 2016;533:539.