ARTICLE OPEN ACCESS

Atrial fibrillation genetic risk differentiates cardioembolic stroke from other stroke subtypes

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

Correspondence
Dr. Lubitz
slubitz@mgh.harvard.edu
or Dr. Anderson
cdanderson@mgh.harvard.edu

Neurol Genet 2018;4:e293. doi:10.1212/NXG.0000000000000293

Abstract

Objective

We sought to assess whether genetic risk factors for atrial fibrillation (AF) can explain cardioembolic stroke risk.

Methods

We evaluated genetic correlations between a previous genetic study of AF and AF in the presence of cardioembolic stroke using genome-wide genotypes from the Stroke Genetics Network (N = 3,190 AF cases, 3,000 cardioembolic stroke cases, and 28,026 referents). We tested whether a previously validated AF polygenic risk score (PRS) associated with cardioembolic and other stroke subtypes after accounting for AF clinical risk factors.

Results

We observed a strong correlation between previously reported genetic risk for AF, AF in the presence of stroke, and cardioembolic stroke (Pearson r = 0.77 and 0.76, respectively, across SNPs with $p < 4.4 \times 10^{-4}$ in the previous AF meta-analysis). An AF PRS, adjusted for clinical AF risk factors, was associated with cardioembolic stroke (odds ratio [OR] per SD = 1.40, $p = 1.45 \times 10^{-48}$), explaining ~20% of the heritable component of cardioembolic stroke risk. The AF PRS was also associated with stroke of undetermined cause (OR per SD = 1.07, $p = 0.004$), but no other primary stroke subtypes (all $p > 0.1$).

Conclusions

Genetic risk of AF is associated with cardioembolic stroke, independent of clinical risk factors. Studies are warranted to determine whether AF genetic risk can serve as a biomarker for strokes caused by AF.

*These authors contributed equally to supervision of this study.

From the Department of Genetics (S.L.P., P.I.W.d. B.), University Medical Center Utrecht, Utrecht University, The Netherlands; P.I.W.d.B. is now with Computational Genomics, Vertex Pharmaceuticals, Boston, MA; Li Ka Shing Centre for Health Information and Discovery (S.L.P.), The Big Data Institute, University of Oxford, United Kingdom; Program in Medical and Population Genetics (S.L.P., L.-C.W., S.H.C., J.R., P.T.E., S.A.L., C.D.A.), Broad Institute, Cambridge, MA; Cardiovascular Research Center (L.-C.W., P.T.E., S.A.L.), Center for Genomic Medicine (J.R., C.D.A.), J.P. Kistler Stroke Research Center (J.R., C.D.A.), and Cardiac Arrhythmia Service (P.T.E., S.A.L.), Massachusetts General Hospital, Boston; Department of Medicine (P.F.M., B.D.M.), Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore; National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study (L.T., E.J.B.); Department of Biostatistics (L.T.) and Department of Epidemiology (E.J.B.), Boston University School of Public Health, MA; Geriatrics Research and Education Clinical Center (B.D.M.), Baltimore Veterans Administration Medical Center, MD; Cardiology Preventive Medicine Sections (E.J.B.), Evans Department of Medicine, Boston University School of Medicine; Department of Neurology (S.J.K.), University of Maryland School of Medicine; and Department of Neurology (S.J.K.), Veterans Affairs Medical Center, Baltimore, MD.

The list of Atrial Fibrillation Genetics (AFGen) Consortium and Stroke Genetics Network (SiGN) Consortium members can be found at https://links.lww.com/NXG/A123.

Funding information and disclosures are provided at the end of the article. Full disclosure form information provided by the authors is available with the full text of this article at Neurology.org/NG.

The Article Processing Charge was funded by the Authors.

Glossary

AF = atrial fibrillation; **AFGen** = AF Genetics; **CCS** = Causative Classification System; **CI** = confidence interval; **GRM** = genetic relationship matrix; **GWAS** = genome-wide association studies; **MAF** = minor allele frequency; **OR** = odds ratio; **PC** = principal component; **PRS** = polygenic risk score; **SiGN** = Stroke Genetics Network.

Atrial fibrillation (AF) affects nearly 34 million individuals worldwide[1] and is associated with a fivefold increased risk of ischemic stroke,[2] a leading cause of death and disability.[3,4] AF promotes blood clot formation in the heart, which can embolize distally, and is a leading cause of cardioembolism. Secondary prevention of cardioembolic stroke is directed at identifying AF as a potential cause and initiating anticoagulation to prevent recurrences. Yet, AF can remain occult even after extensive workup owing to the paroxysmal nature and fact that it can be asymptomatic. Because both AF and stroke are heritable, and because there is a compelling clinical need to determine whether stroke survivors have AF as an underlying cause, we sought to determine whether genetic risk of cardioembolic stroke can be approximated by measuring genetic susceptibility to AF.

Recent genome-wide association studies (GWAS) have demonstrated that both AF[5] and ischemic stroke[6,7] are complex disorders with polygenic architectures. The top loci for cardioembolic stroke, on chromosome 4q25 upstream of *PITX2* and on 16q22 near *ZFHX3*, are both leading risk loci for AF.[8–10] Despite overlap in top risk loci, the genetic susceptibility to both AF and cardioembolic stroke is likely to involve the aggregate contributions of hundreds or thousands of loci, consistent with other polygenic conditions.[11]

To understand whether genetic risk of AF is an important and potentially useful determinant of overall cardioembolic stroke risk, we analyzed 13,390 ischemic stroke cases and 28,026 referents from the NINDS-Stroke Genetics Network (SiGN)[12] with genome-wide genotyping data. First, we assessed whether patients with stroke with AF have a genetic predisposition to arrhythmia, leveraging additional GWAS data from the AF Genetics (AFGen) Consortium. Second, we compared genetic risk factors for AF and stroke to ascertain the extent to which heritable risk of cardioembolic stroke is explained by genetic risk factors for AF.

Methods

The Stroke Genetics Network

The SiGN was established with the aim of performing the largest genome-wide association study (GWAS) of ischemic stroke to date. The study design has been previously described[12] (e-Methods). Briefly, subjects in SiGN were classified into stroke subtypes using the Causative Classification System (CCS), which subtypes cases through an automated, web-based system that accounts for clinical data, test results, and imaging information.[13,14] Within the CCS, there are 2 subcategories: CCS causative, which does not allow for competing subtypes in a single sample, and CCS phenotypic, which does. In addition, ~74% of samples were subtyped using the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) subtyping system.[15] After quality control, the SiGN data set comprised 16,851 ischemic stroke cases and 32,473 stroke-free controls (e-Methods and table e-1, links.lww.com/NXG/A123). In this study, we analyze only the European- and African-ancestry samples (13,390 cases and 28,026 controls).

Standard protocol approvals, registrations, and patient consents

All cohorts included in the SiGN data set received approval from the cohort-specific ethical standards committee. Cohorts received written informed consent from all patients or guardians of patients for participation in the study, where applicable. Details on sample collection have been previously described.[12]

Identifying AF cases and controls

We defined AF in SiGN on the basis of 5 variables available in the CCS phenotyping system: (1) AF, (2) paroxysmal AF, (3) atrial flutter, (4) sick sinus syndrome, and (5) atrial thrombus. This definition yielded 3,190 AF cases for analysis. We also defined a strict case set based on "AF" only (N = 1,751 cases) for sensitivity analyses (e-Methods and figure e-1, links.lww.com/NXG/A123).

From the 28,026 controls, we established a set of 3,861 control individuals in whom AF was indicated as not present. For the remaining subjects, we assumed that individuals did not have AF because AF status for most control samples in SiGN is unknown.

Genome-wide association testing of ischemic stroke subtypes and AF in SiGN

We merged genotype dosages together and kept single nucleotide polymorphisms (SNPs) with imputation quality >0.8 and minor allele frequency (MAF) >1% (e-Methods, links.lww.com/NXG/A123). We performed association testing using a linear mixed model (LMM) implemented in BOLT-LMM.[16] We adjusted the model for the top 10 principal components (PCs) and sex, in addition to the genetic relationship matrix (GRM; e-Methods).[16] We performed GWAS in AF and each of the stroke subtypes available in SiGN. Results were unadjusted for age because adjusting for age in the AF GWAS gave results highly concordant with age-unadjusted results (e-Results).

Heritability calculations

We calculated additive SNP-based heritability estimates for ischemic stroke, stroke subtypes, and AF using restricted

maximum (REML) likelihood implemented in BOLT-REML (e-Methods, links.lww.com/NXG/A123).[16]

Genetic correlation between AF and ischemic stroke subtypes

We used summary-level data from a previous AFGen Consortium meta-analysis of AF[5] to calculate a z-score for each SNP in that GWAS. In addition, we calculated a z-score for each SNP from our SiGN GWAS of each stroke subtype and AF. As a null comparator, we downloaded SNP z-scores from a GWAS of educational attainment[17] available through LDHub (ldsc.broadinstitute.org/, accessed November 1, 2017). We calculated Pearson r between z-scores from 2 traits to evaluate correlation (e-Methods and figure e-2, links.lww.com/NXG/A123).

Constructing an AF polygenic risk score (PRS)

To construct an AF PRS, we used SNPs from a previously derived AF PRS (e-Methods, links.lww.com/NXG/A123).[18] Briefly, the PRS was derived from an AF GWAS of 17,931 cases and 115,142 controls.[5] This PRS comprised 1,168 SNPs with $p < 1 \times 10^{-4}$ and LD pruned at an r^2 threshold of 0.5.[18] Of these 1,168 SNPs, we identified 934 SNPs in the SiGN data set with imputation info >0.8 and MAF >1%. We used these 934 SNPs to construct the AF PRS in the SiGN data set. Additional details on the PRS construction can be found in the e-Methods.

Testing an AF PRS in ischemic stroke subtypes

We tested for association between the AF PRS and stroke subtypes using logistic regression (e-Methods, links.lww.com/NXG/A123). We included sex and the top 10 PCs as additional covariates. We optionally adjusted the association tests for age, diabetes mellitus, cardiovascular disease, smoking status (current smoker, former smoker, or never smoked), and hypertension.

We calculated the variance explained by the AF PRS in cardioembolic stroke by constructing a model in BOLT-REML that consisted of: (1) a variance component made up of SNPs for the GRM and (2) a variance component made up of SNPs from the PRS (e-Methods, links.lww.com/NXG/A123).

Data availability

Code, supporting data, and downloadable supplemental tables are available here: github.com/UMCUGenetics/Afib-Stroke-Overlap. The e-data (links.lww.com/NXG/A123) contain additional information regarding data access, methods, and links to summary-level data.

Results

We began by testing our ability to rediscover known AF genetic associations in the SiGN data set, assembled to study the genetics of ischemic stroke. We ran a genome-wide association study (GWAS) in SiGN using 3,190 cases, with AF or paroxysmal AF, as well as other diagnoses suggestive of underlying AF[19,20] (e-Methods, table 1 and table e-1, links.lww.com/NXG/A123), and 28,026 controls (figure e-1). We found the top associated SNPs to be highly concordant with a previous GWAS of AF performed by the AFGen Consortium (table e-2). Adjusting the GWAS for age did not substantially change our findings (r = 0.83 between SNP effects from the age-unadjusted and age-adjusted GWAS).

Table 1 AF and stroke cases in SiGN

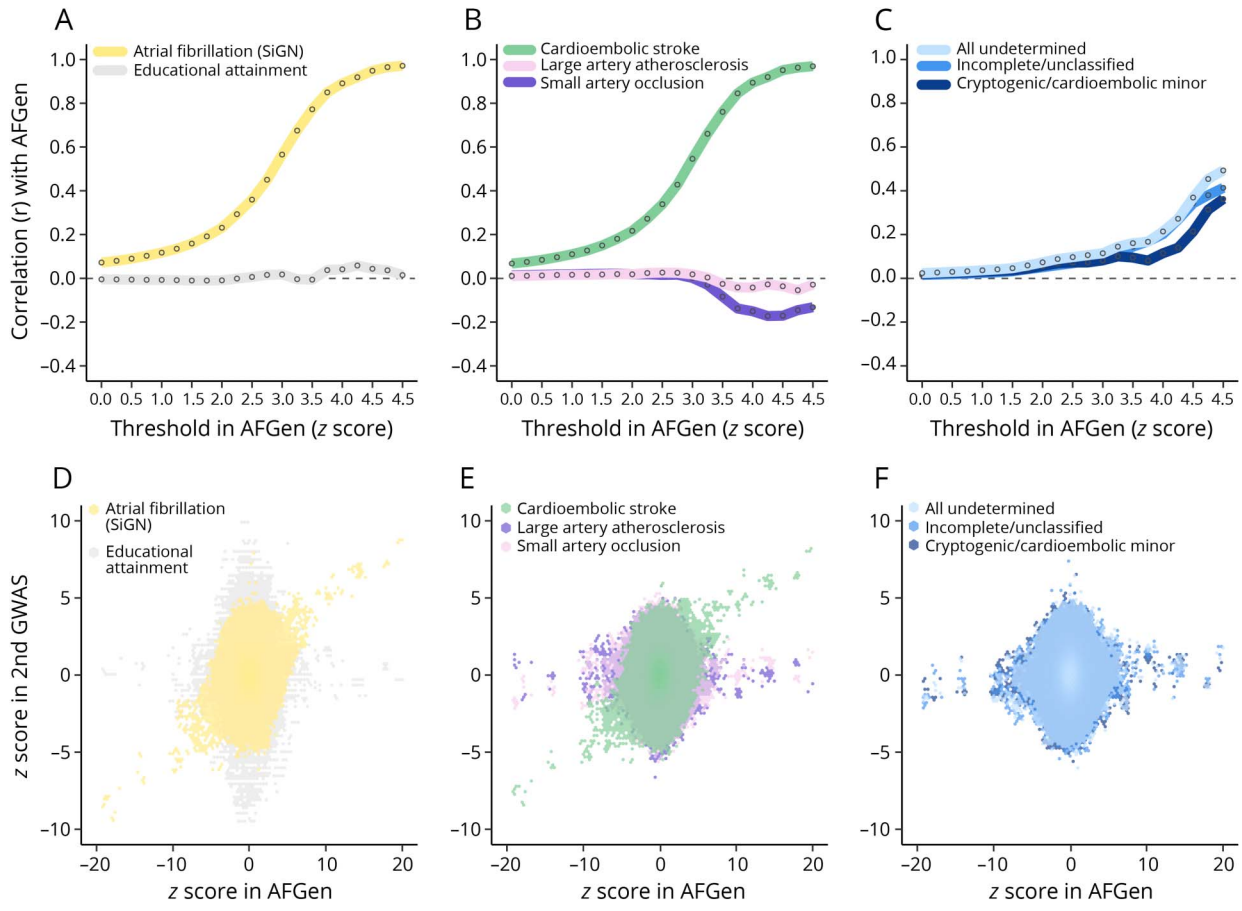
		Ischemic stroke subtype				
		Primary subtypes			Undetermined subtypes	
Phenotype	Total	Cardioembolic	Large artery atherosclerosis	Small artery occlusion	Incomplete/ unclassified	Cryptogenic/ cardioembolic minor
AF diagnosis						
AF	1,751	1,495	63	32	151	0
Paroxysmal AF	1,315	1,088	52	23	138	0
Left 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 AF cases	3,190	2,684	123	61	298	0
No AF	—	316	2,262	2,201	1,982	2,294

Abbreviations: AF = atrial fibrillation; SiGN = Stroke Genetics Network.
Of the 13,390 stroke cases available in the SiGN data set, a total of 3,190 cases had AF or other suggestive diagnoses. Although most of these cases were subtyped as having a cardioembolic stroke, a fraction was distributed among the other stroke subtypes. Samples can appear more than once per row (i.e., have more than 1 AF diagnosis), but totals represent the number of unique AF samples in each stroke subtype. There are no subjects with AF or equivalent subtyped as "cryptogenic/cardioembolic minor" because such a diagnosis would remove them from this category.

By extending our analysis beyond these top associations, we next assessed whether patients with stroke with AF have a similar overall genetic predisposition to the arrhythmia as seen in the independent AFGen GWAS. In addition, we assessed the overlap between genetic predisposition to AF and each stroke subtype, allowing for the known phenotypic concordance between cardioembolic stroke and AF (89.5% of cardioembolic stroke cases in SiGN also have AF, table e-1, links.lww.com/NXG/A123). We performed a series of GWAS in the SiGN data for AF and each of the stroke subtypes using BOLT-LMM[16] (e-Methods) and calculated the z-score (beta/standard error) of each SNP in each phenotype. We then used summary-level results available from the previous (independent) GWAS of AF[5] (from AFGen) and calculated the z-score for each SNP in that data set.

By measuring Pearson correlation (r) between AFGen z-scores and z-scores from the AF GWAS in SiGN, we found only a modest correlation (r = 0.07 across ~7.8M SNPs, figure 1, A and D). However, when we iteratively subsetted the AFGen GWAS results by the (absolute values of) z-scores of the SNPs, we found that correlation with the AF GWAS in SiGN increased as the z-score threshold became more stringent. For example, for those ~4.5M SNPs with |z| > 1 in AFGen, correlation with AF SNPs in SiGN was 0.12; for those ~1.9M SNPs with |z| > 3.5 in AFGen, correlation with the SiGN AF GWAS rose to 0.77 (figure 1, A and D, and table e-3). These correlations, calculated to include even modestly associated SNPs, indicate that AF in AFGen and AF in stroke (SiGN) share a large proportion of genetic risk factors. Removing ±2 Mb around the *PITX2* and *ZFHX3* loci only modestly affected the correlation between AFGen and AF in SiGN (r = 0.63 for SNPs with |z| > 3.5; figure e-3, links.lww.com/NXG/A123 and table e-3). Correlations between AFGen and cardioembolic stroke in SiGN were unsurprisingly highly similar to that of the results with AF in SiGN (r = 0.77 for AFGen SNPs with |z| > 3.5) likely because of the high concordance between the AF and cardioembolic stroke phenotypes (figure 1, B and E and figure e-3).

Figure 1 Genetic correlation between atrial fibrillation (AF) in the AF Genetics (AFGen) Consortium meta-analysis and AF and ischemic stroke subtypes analyzed in SiGN



Pearson r correlation between SNP z-scores in the AFGen GWAS of AF and in GWAS of selected traits performed in the SiGN data. (A) GWAS of AF in AFGen and in SiGN correlate with increasing strength as SNP z-scores in AFGen increase. Correlation with educational attainment (performed separately, shown here as a null comparator) remains approximately zero across all z-score thresholds. (B) SNP effects in AFGen also correlate strongly with cardioembolic stroke in SiGN, but not with the other primary stroke subtypes. (C) Undetermined subtypes of stroke also show modest correlation with the genetic architecture of AF in AFGen. Panels (D–F) show genome-wide z-score distributions underlying correlations. Shading of the hexagons indicates the density of the data at that point, where darker shading indicates a higher density of SNPs. GWAS = genome-wide association studies.

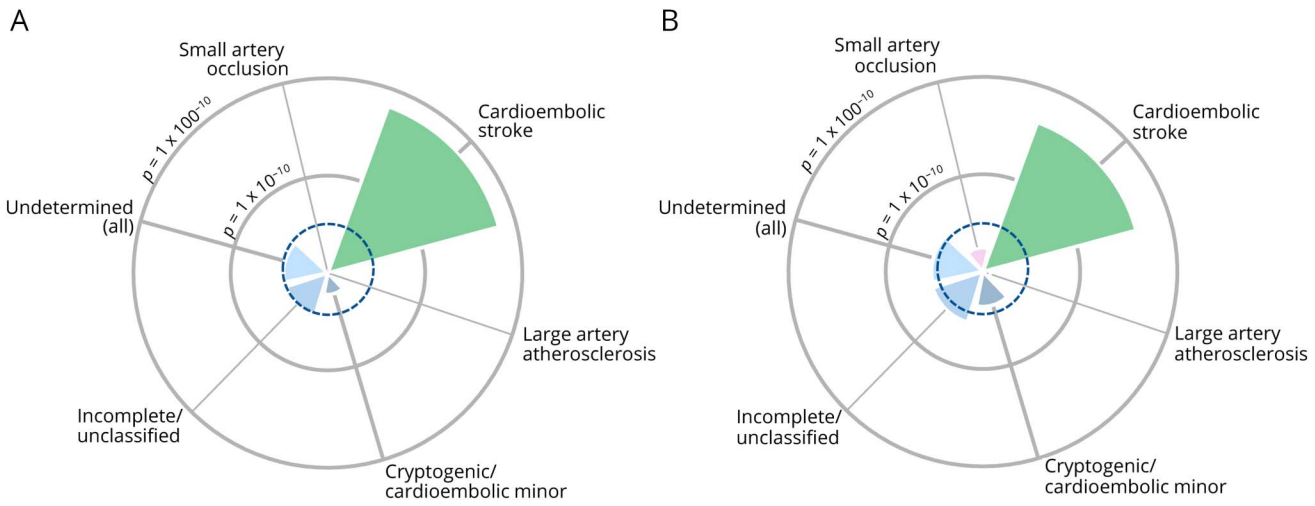
Continuing this analysis across the other stroke subtypes (large artery atherosclerosis, small artery occlusion, and undetermined stroke; figure 1, B, C, E, F), we found near-zero correlation between AFGen and either large artery atherosclerosis or small artery occlusion (figure 1, B and E), indicating no genetic overlap between the phenotypes. However, the correlation between AF and the undetermined stroke subtypes (a highly heterogeneous subset of cases[21,22] that cannot be classified with standard subtyping systems[13,15]) increased steadily as we partitioned the AFGen data by z-score (all undetermined vs AFGen r = 0.04 for AFGen SNPs with $|z| > 1$ and r = 0.16 for AFGen SNPs with $|z| > 3.5$; figure 1, C and F, and table e-3, links.lww.com/NXG/A123), indicating that genome wide, there is residual genetic correlation between AF and the undetermined stroke categories, some of which could represent causal AF stroke mechanisms in that subgroup. As an additional null comparator, we performed correlations between the AFGen results with z-scores derived from the latest GWAS of educational attainment[17] and found that correlation remained at approximately zero regardless of the z-score threshold used (figure 1, A and D, and table e-3).

To further understand the overlap between genetic risk factors for AF and cardioembolic stroke and to evaluate the degree to which cardioembolic stroke comprised risk factors beyond those for AF, we performed a restricted maximum likelihood analysis implemented in BOLT-REML[16] to estimate SNP-based heritability of AF and cardioembolic stroke. Using phenotypes derived from the CCS subtyping algorithm[23] (e-Methods, links.lww.com/NXG/A123), we estimated heritability of AF and cardioembolic stroke at 20.0% and 19.5%, respectively. These estimates are consistent with previous estimates in larger samples (figure e-4),[24,25] and the similar heritabilities suggest that cardioembolic stroke does not have a substantial heritable component beyond the primary AF risk factor. For comparison, we calculated heritability in the other stroke subtypes[15] and found estimates to be similarly modest (range: 15.5%–23.0%; figures e4-e6 and table e-4).

Up to this point, our results indicated that AF in ischemic stroke is genetically similar to that discovered in previous genetic studies of AF alone and that the bulk of the genetic risk of cardioembolic stroke seems attributable to AF genetic risk factors. Next, we sought to explicitly test what proportion of cardioembolic stroke risk could be explained by AF loci, independent of known clinical risk factors for AF. First, we identified SNPs from an AF PRS independently derived from the AFGen GWAS[5] (e-Methods, links.lww.com/NXG/A123). Of the 1,168 SNPs used to generate this pre-established PRS, we identified 934 in the SiGN data set with imputation quality >0.8 and MAF >1%. We computed the PRS per individual (e-Methods), weighting the imputed dosage of each risk allele by the effect of the SNP (i.e., the beta coefficient) as reported in AFGen.[5]

We tested the association of the AF PRS with cardioembolic stroke, using a logistic regression and adjusting for the top 10 PCs and sex (e-Methods, links.lww.com/NXG/A123). As expected from our earlier results, we found the PRS to be strongly associated with cardioembolic stroke (odds ratio [OR] per 1 SD of the PRS = 1.93 (95% confidence interval [CI]: 1.34–1.44), $p = 1.01 \times 10^{-65}$; figure 2A and table e-5), confirming the high genetic concordance of these phenotypes across SNPs that, individually, confer only a modest average association with AF. Next, we adjusted the association model for clinical covariates associated with AF including age, diabetes mellitus, cardiovascular disease, smoking, and

Figure 2 Association of atrial fibrillation (AF) polygenic risk score in ischemic stroke subtypes



We constructed an independent polygenic risk score (PRS) from AF-associated SNPs identified in the AFGen GWAS and tested associations between this PRS and ischemic stroke subtypes using (A) all available referents (N = 28,026) and (B) referents without AF (N = 3,861). The PRS strongly associated with cardioembolic stroke in both sets of samples. In the AF-free set of controls (panel B), we observed association of the PRS ($p < 5 \times 10^{-3}$, after adjusting for 5 subtypes and 2 sets of referents; indicated by the dashed dark blue line) with incomplete/unclassified stroke as well. GWAS = genome-wide association studies; PRS = polygenic risk score.

hypertension.[26] Using a (smaller) set of cases and controls with complete clinical risk factor information, we found that inclusion of these clinical risk factors in the model only modestly reduced the PRS signal in cardioembolic stroke (OR per 1 SD = 1.40 [95% CI: 1.34–1.47], $p = 1.45 \times 10^{-48}$; tables e5-e7, links.lww.com/NXG/A123). These results indicate a strong relationship between AFGen risk factors and cardioembolic stroke risk, independent of the clinical factors that associate with AF. Expanding the set of SNPs used to construct the PRS to the original 934 SNPs ±25 kb, ±50 kb, and ±100 kb (e-Methods) revealed a persistently strong, though somewhat attenuated, association between the PRS and cardioembolic stroke (PRS including SNPs within 100 kb, $p = 4.47 \times 10^{-44}$, table e-6). None of the other stroke subtypes were significantly associated with the AF PRS (all $p > 0.013$, figure 2A and figure e-7).

Because AF status was missing for most controls in the SiGN data set, we performed sensitivity analyses using only the 3,861 controls confirmed as having no AF. Although reducing the set of controls to this refined group did not substantially change the results for the primary stroke subtypes, we found that the AF PRS was modestly associated ($p < 5 \times 10^{-3}$, after adjusting for 5 subtypes and 2 control groups) with the overall undetermined subtype (OR per 1 SD = 1.07 [95% CI: 1.02–1.13], $p = 4.15 \times 10^{-3}$) (figure 2B and table e-5, links.lww.com/NXG/A123). Further examination of the 2 mutually exclusive subgroups of the undetermined group revealed that the PRS associated significantly with the incomplete/unclassified categorization (OR per 1 SD = 1.09 [95% CI: 1.03–1.16], $p = 3.17 \times 10^{-3}$) (figure 2B) but not with cryptogenic/cardioembolic minor (OR per 1 SD = 1.06 [95% CI: 1.00–1.13], $p = 5.10 \times 10^{-2}$). Correcting for clinical covariates only modestly changed the signal in the incomplete/unclassified phenotype ($p = 9.7 \times 10^{-3}$, figure 2), supporting the robustness of the observed association, independent of clinical risk factors.

Last, we created a model in BOLT-LMM, fitting 2 genetic variance components: 1 component including SNPs for the GRM and the second component including the original PRS SNPs from the AF PRS (including ±100 kb around these SNPs to include a sufficient number of markers to estimate the variance explained) (table e-8). We found that the SNPs from the AF PRS explained 4.1% of the total (20.0%) heritability in AF. In evaluating the variance explained in cardioembolic stroke, we found a nearly identical result: the component representing the AF risk score explained 4.5% (SE = 1.00%) of the total 19.5% genetic heritability in cardioembolic stroke. Thus, AF genetic risk accounts for 23.1%, or approximately one-fifth, of the total heritability of cardioembolic stroke.

Discussion

Our results suggest that individuals with cardioembolic strokes have an enrichment for AF genetic risk, despite the fact that cardioembolic stroke often affects older adults with multiple clinical comorbidities[27] that could increase risk of AF because of nongenetic factors. The fact that cardioembolic stroke and AF share a highly similar genetic architecture extends our understanding of the morbid consequences of heritable forms of arrhythmia. Furthermore, the observation that AF genetic risk was only associated with cardioembolic stroke, and (consistently) lacked association in large artery atherosclerosis or small artery occlusion,[28] increases the possibility that AF genetic risk may be informative in the management of ischemic stroke survivors in whom the mechanism may be unclear.

The use of PRSs for complex traits has proved an efficient means of understanding how genetic predisposition to diseases can overlap. Given the onslaught of genotyping data available for common diseases, PRSs can now be used to stratify patients by risk (e.g., in breast cancer[29,30]) or predict outcome (e.g., in neuropsychiatric disease[29]). More recently, PRSs have been used to identify individuals in the general population with a four-fold risk of coronary disease,[31] proposed for inclusion in clinical workups of individuals with early-onset coronary artery disease,[32] and used to identify patients for whom lifestyle changes or statin intervention would be beneficial.[33,34] Although previous work has also shown an association between an AF PRS and cardioembolic stroke,[28] we have extended this work to formally quantify the extent to which an AF PRS captures genetic risk of cardioembolic stroke. These findings lay the groundwork for future work that can potentially leverage this overlap to develop AF PRSs that could be used to predict individuals at highest risk of cardioembolic stroke (to improve diagnostic resource allocation) or help distinguish between clinical subtypes of stroke.

Although our analysis was aimed at understanding the genetic overlap between cardioembolic stroke and AF, we additionally observed genetic correlation between AF and undetermined stroke, a finding not observed in a previous investigation of AF PRS in ischemic stroke subtypes, albeit in a smaller sample.[28] Perhaps contrary to expectation, we specifically found the AF PRS to be more strongly associated with the subset of etiology-undetermined strokes with an incomplete clinical evaluation, as opposed to those with cryptogenic stroke of a presumed, but not demonstrated, embolic source. These associations could be due to physician biases in diagnostic workups, rather than supporting a low prevalence of occult AF in presumed embolic strokes of undetermined source. Identifying patients with stroke with AF is an important clinical challenge because occult AF is well known to cause strokes[35,36] and because such patients are at high risk of recurrent stroke, which is preventable with anticoagulation.[37,38] Together, our findings indicate that AF genetic risk may augment clinical algorithms to determine stroke etiology, but will require further study.

The work presented here benefits from a number of improvements, including increased sample size; analysis of samples from a multicenter consortium, potentially enhancing

the generalizability of the findings; and use of the CCS subtyping system, which provides more nuanced phenotyping, particularly in the cryptogenic subtype. Nevertheless, some limitations remain. Stroke is a heterogeneous condition that occurs later in life and has a high lifetime prevalence (>15%[39]), features that can reduce statistical power. Furthermore, sample sizes have lagged behind other GWAS efforts, a challenge further compounded by subtyping (nearly one-third of all cases are categorized as undetermined[23]). Reduced sample sizes affect power for discovery and make other analytic approaches—such as standard approaches for measuring trait correlation[16]—unfeasible. Also, our sample primarily comprised European-ancestry samples, and work in non-Europeans, particularly in African-ancestry samples where risk of stroke is double that of European samples, is crucial. Finally, the current analysis does not analyze rare variations, which also likely contributes to disease susceptibility.[5]

We have shown that the cumulative genetic risk of AF in individuals with a stroke is similar to that reported in a larger population-based cohort.[25] Genome-wide variation related to AF is substantially associated with cardioembolic stroke risk. Moreover, AF genetic risk was specific for cardioembolic stroke and was not associated with the other primary stroke subtypes. The observation that AF genetic risk associated with strokes of undetermined cause supports the notion that undetected AF underlies a proportion of stroke risk in these individuals. Further work will need to incorporate emerging discoveries of rare genetic variants in AF and explore the potential of genetic risk tools, including PRSs performed via clinical-grade genotyping, to assist in the diagnostic workup of individuals with ischemic stroke.

Author contributions

S.L. Pulit: conception of research design, data analysis, drafting of the manuscript, and critical revision of the manuscript. L.-C. Weng: data analysis and critical revision of the manuscript. P.F. McArdle: data acquisition, data analysis, and critical revision of the manuscript. L. Trinquart and S.H. Choi: data acquisition and critical revision of the manuscript. B.D. Mitchell and J Rosand: data acquisition, study supervision, and critical revision of the manuscript. P.I.W. de Bakker: study supervision and critical revision of the manuscript. E.J. Benjamin, P.T. Ellinor, and S.J. Kittner: data acquisition, study supervision, and critical revision of the manuscript. S.A. Lubitz and C.D. Anderson: conception of research design, study supervision, drafting of the manuscript, and critical revision of the manuscript.

Study funding

S.L. Pulit, B.D. Mitchell, P.F. McArdle, and S.J. Kittner are supported by the NIH grant R01NS100178. The NINDS-SiGN Consortium is supported by the NIH grants R01NS100178 and U01NS069208. S.L. Pulit is supported by Veni Fellowship 016.186.071 (ZonMW) from the Dutch Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO). C.D. Anderson is supported in part by K23NS086873, R01NS103924, an American Heart Association Strategically Focused Research Network in Atrial Fibrillation Award, and a Massachusetts General Hospital Center for Genomic Medicine Catalyst Award. S.A. Lubitz is supported by the NIH grant K23HL114724, an American Heart Association Strategically Focused Research Network in Atrial Fibrillation Award, and a Doris Duke Charitable Foundation Clinical Scientist Development Award 2014105. L.-C. Weng is supported by an American Heart Association Postdoctoral Fellowship Award 17POST33660226. Drs. Ellinor and Benjamin are supported by the NIH grants R01HL092577 and R01HL128914 and an American Heart Association Strategically Focused Research Network in Atrial Fibrillation Award. E.J. Benjamin is additionally supported by the NIH grants 1RC1HL101056 and 1R01HL102214. P.T. Ellinor is additionally supported by the NIH grants R01HL104156 and K24HL105780; the National Heart, Lung, and Blood Institute (NHLBI); American Heart Association Established Investigator Award 13EIA14220013; and the Fondation Leducq 14CVD01.

Disclosure

S.L. Pulit has received speaker honoraria from the commercial entity Illumina. L.-C. Weng has received research support from the American Heart Association. P.F. McArdle has received research support from the commercial entity Regeneron and has received governmental research support the National Institute of Neurological Disorders and Stroke. L. Trinquart and S.H. Choi report no disclosures. B.D. Mitchell has served on the External Scientific Review Board for the Integrative Cardiac Health Project (ICHP) at Walter Reed National Military Medical Center; has served on the editorial boards of *Current Genetic Medicine Reports* and *Frontiers in Neurology*; and has received governmental research support from the National Institutes of Health. J. Rosand has served on the scientific advisory boards of Pfizer and the Data and Safety Monitoring Board; has served on the editorial boards of *Lancet Neurology* and *Stroke*; has received governmental research support from the National Institutes of Health; and has been involved in legal proceedings with Boehringer Ingelheim. P.I.W. de Bakker is employed by Vertex Pharmaceuticals. E.J. Benjamin has served on the scientific advisory boards of non-profit organizations, the National Institutes of Health, and the National Heart, Lung, and Blood Institute; served on the editorial board of the American Heart Association; has been employed by Boston University School of Medicine; has received governmental research support from the National Institutes of Health, American Heart Association, and the Robert Wood Johnson Foundation; and has received foundation/society research support from the Boston University School of Medicine and from the American Council on Education. P.T. Ellinor has served on the scientific advisory board for Bayer AG Novartis Quest Diagnostics; has received research support (commercial) from Bayer AG; has received governmental research support from the National Institutes of Health; and has received foundation/society

research support from the American Heart Association and the Fondation Leducq. S.J. Kittner serves on the editorial board of *Neuroepidemiology* and has received governmental research support from the National Institute of Neurological Disorders and Stroke. S.A. Lubitz has acted as a consultant for Abbott, Quest Diagnostics, and Bristol-Myers Squibb; has received commercial research support from Bristol-Myers Squibb, Bayer HealthCare, Biotronik, and Boehringer Ingelheim; has received governmental research support from the National Institutes of Health; and has received foundation/society research support from the National Institutes of Health, the American Heart Association, and the Doris Duke Charitable Foundation. C.D. Anderson has served as a consultant for ApoPharma, Inc. and has received governmental research support from the National Institutes of Health and the American Heart Association. Full disclosure form information provided by the authors is available with the full text of this article at Neurology.org/NG.

Publication history

Received by *Neurology: Genetics* June 23, 2018. Accepted in final form September 9, 2018.

References

1. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study. Circulation 2014;129:837–847.
2. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. Stroke 1991;22:983–988.
3. World Health Organization. WHO|the Top 10 Causes of Death [online]. World Health Organization; 2014. Available at who.int/mediacentre/factsheets/fs310/en/. Accessed November 1, 2017.
4. Feigin VL, Lawes CM, Bennett DA, Barker-Collo SL, Parag V. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. Lancet Neurol 2009;8:355–369.
5. Christophersen IE, Rienstra M, Roselli C, et al. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. Nat Genet 2017;57:289.
6. Bevan S, Traylor M, Adib-Samii P, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genome-wide associations. Stroke 2012;43:3161–3167.
7. Lubitz SA, Ozcan C, Magnani JW, Kääb S, Benjamin EJ, Ellinor PT. Genetics of atrial fibrillation: implications for future research directions and personalized medicine. Circ Arrhythm Electrophysiol 2010;3:291–299.
8. Gudbjartsson DF, Arnar DO, Helgadottir A, et al. Variants conferring risk of atrial fibrillation on chromosome 4q25. Nature 2007;448:353–357.
9. Gretarsdottir S, Thorleifsson G, Manolescu A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. Ann Neurol 2008;64:402–409.
10. Gudbjartsson DF, Holm H, Gretarsdottir S, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. Nat Genet 2009;41:876–878.
11. Visscher PM, Wray NR, Zhang Q, et al. 10 years of GWAS discovery: biology, function, and translation. Am J Hum Genet 2017;101:5–22.
12. Pulit SL, McArdle PF, Wong Q, Malik R, et al. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurol 2015;15:174–184.
13. Ay H, Benner T, Arsava EM, et al. A computerized algorithm for etiologic classification of ischemic stroke: the causative classification of stroke system. Stroke 2007;38:2979–2984.
14. Arsava EM, Ballabio E, Benner T, et al. The causative classification of stroke system: an international reliability and optimization study. Neurology 2010;75:1277–1284.
15. Adams HP Jr, Birgitte H, Bendixen P, Jaap Kappelle L, et al. Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial. Stroke 1993;24:35–41.
16. Loh PR, Tucker G, Bulik-Sullivan BK, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet 2015;47:284–290.
17. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 2016;533:539–542.
18. Weng L-C, Preis SR, Hulme OL, et al. Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. Circulation 2018;137:1027–1038.
19. Franz MR, Karasik PL, Li C, Moubarak J, Chavez M. Electrical remodeling of the human atrium: similar effects in patients with chronic atrial fibrillation and atrial flutter. J Am Coll Cardiol 1997;30:1785–1792.
20. Lamas GA, Lee KL, Sweeney MO, et al. Ventricular pacing or dual-chamber pacing for sinus-node dysfunction. N Engl J Med 2002;346:1854–1862.
21. Petty GW, Brown RD, Whisnant JP, Sicks JD, O'Fallon WM, Wiebers DO. Ischemic stroke subtypes: a population-based study of incidence and risk factors. Stroke 1999;30:2513–2516.
22. Kolominsky-Rabas PL, Weber M, Gefeller O, Neundoerfer B, Heuschmann PU. Epidemiology of ischemic stroke subtypes according to TOAST criteria: incidence, recurrence, and long-term survival in ischemic stroke subtypes: a population-based study. Stroke 2001;32:2735–2740.
23. 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.
24. Go AS, Hylek EM, Phillips KA, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. JAMA 2001;285:2370–2375.
25. Weng LC, Choi SH, Klarin D, et al. Heritability of atrial fibrillation. Circ Cardiovasc Genet 2017;10:e001838.
26. Alonso A, Krijthe BP, Aspelund T, et al. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. J Am Heart Assoc 2013;2:e000102.
27. Henninger N, Goddeau RP Jr, Karmarkar A, Helenius J, McManus DD. Atrial fibrillation is associated with a worse 90-day outcome than other cardioembolic stroke subtypes. Stroke 2016;47:1486–1492.
28. Lubitz SA, Parsons OE, Anderson CD, et al. Atrial fibrillation genetic risk and ischemic stroke mechanisms. Stroke 2017;48:1451–1456.
29. Lewis CM, Vassos E. Prospects for using risk scores in polygenic medicine. Genome Med 2017;9:96.
30. Mavaddat N, Pharoah PDP, Michailidou K, et al. prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 2015;107. doi: 10.1093/jnci/djv036.
31. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases to identify individuals with risk equivalent to monogenic mutations. Nat Gen 2018;50:1219–1224.
32. Thériault S, Lali R, Chong M, Velianou JL, Natarajan MK, Paré G. Polygenic contribution in individuals with early-onset coronary artery DiseaseClinical perspective. Circ Genom Precis Med 2018;11:e001849.
33. Khera AV, Emdin CA, Drake I, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. N Engl J Med 2016;375:2349–2358.
34. Mega JL, Stitziel NO, Smith JG, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. Lancet 2015;385:2264–2271.
35. Gladstone DJ, Spring M, Dorian P, et al. Atrial fibrillation in patients with cryptogenic stroke. N Engl J Med 2014;370:2467–2477.
36. Sanna T, Diener HC, Passman RS, et al. Cryptogenic stroke and underlying atrial fibrillation. N Engl J Med 2014;370:2478–2486.
37. ACTIVE Investigators, Connolly SJ, Pogue J, Hart RG, et al. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. N Engl J Med 2009;360:2066–2078.
38. Diener H-C, Eikelboom J, Connolly SJ, et al. Apixaban versus aspirin in patients with atrial fibrillation and previous stroke or transient ischaemic attack: a predefined subgroup analysis from AVERROES, a randomised trial. Lancet Neurol 2012;11:225–231.
39. Seshadri S, Beiser A, Kelly-Hayes M, et al. The lifetime risk of stroke: estimates from the Framingham study. Stroke 2006;37:345–350.

Atrial fibrillation genetic risk differentiates cardioembolic stroke from other stroke subtypes

Sara L. Pulit, Lu-Chen Weng, Patrick F. McArdle, et al.

Neurol Genet 2018;4;
DOI 10.1212/NXG.0000000000000293