Genetic Risk Prediction of Atrial Fibrillation

Steven A. Lubitz, Massachusetts General Hospital
Xiaoyan Yin Yin, Boston University
Henry J. Lin, Harbor-UCLA Medical Center
Matthew Kolek, Vanderbilt University
J. Gustav Smith, Lund University
Stella Trompet, Leiden University
Michiel Rienstra, University of Groningen
Natalia S. Rost, Massachusetts General Hospital
Pedro L. Teixeira, Vanderbilt University
Peter Almgren, Lund University

Only first 10 authors above; see publication for full author list.

Journal Title: Circulation
Volume: Volume 135, Number 14
Publisher: American Heart Association | 2017-04-04, Pages 1311--+
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1161/CIRCULATIONAHA.116.024143
Permanent URL: https://pid.emory.edu/ark:/25593/s92w6

Final published version:
http://dx.doi.org/10.1161/CIRCULATIONAHA.116.024143

Copyright information:
©2016 American Heart Association, Inc.

Accessed December 6, 2018 6:52 PM EST
Genetic Risk Prediction of Atrial Fibrillation

Steven A. Lubitz, MD, MPH¹,²,³, Xiaoyan Yin, PhD⁴, Henry J. Lin, MD⁵,⁶, Matthew Kolek, MD⁷, J. Gustav Smith, MD, PhD⁸, Stella Trompet, PhD⁹,¹⁰, Michiel Rienstra, MD, PhD¹¹, Natalia S. Rost, MD, MPH¹², Pedro L. Teixeira, PhD⁷, Peter Almgren, MSc¹³, Christopher D. Anderson, MD, MMSc²,¹²,¹⁴, Lin Y. Chen, MD, MS¹⁵, Gunnar Engström, MD, PhD¹³, Ian Ford, MD PhD¹⁶, Karen L. Furie, MD, MPH¹⁷, Xiuqing Guo, PhD⁵, Martin G. Larson, ScD⁴,¹⁸,¹⁹, Kathryn L. Lunetta, PhD⁴,¹⁹, Peter W. Macfarlane, DSc²⁰, Bruce M. Psaty, MD, PhD²¹,²², Elsayed Z. Soliman, MD, MSc, MS²³, Nona Sotoodehnia, MD, MPH²⁴,²⁵, David J. Stott, MD²⁰, Kent D. Taylor, PhD⁵,⁶, Lu-Chen Weng, PhD²,³, Jie Yao, MS⁵,⁶, Bastiaan Geelhoed, PhD¹¹, Niek Verweij, PhD¹¹, Joylene E. Siland, MSc¹¹, Sekar Kathiresan, MD²,³,¹⁴, Carolina Roselli, MS², Dan Roden, MD²⁶, Pim van der Harst, MD, PhD¹¹, Dawood Darbar, MD⁷,²⁷, J. Wouter Jukema, MD, PhD²⁸, Olle Melander, MD, PhD²⁹,³⁰, Jonathan Rosand, MD, MSc²,¹²,¹⁴, Jerome I. Rotter, MD⁵,⁶, Susan R. Heckbert, MD, PhD²⁴,²⁵,†, Patrick T. Ellinor, MD, PhD¹,²,³,†, Alvaro Alonso, MD, PhD³¹,†, and Emelia J. Benjamin, MD, ScM¹,⁴,³²,³³,³⁴ on behalf of the AFGen Consortium

¹Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, Massachusetts, USA
²Program in Medical and Population Genetics, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA
³Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA
⁴Boston University and National Heart, Lung and Blood Institute’s Framingham Heart Study, Framingham, Massachusetts, USA
⁵Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, California, USA
⁶Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, California, USA
⁷Vanderbilt University, Nashville, Tennessee, USA
⁸Department of Cardiology, Clinical Sciences, Lund University and Skåne University Hospital, Lund, Sweden
⁹Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
¹⁰Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands
¹¹University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
¹²J. Philip Kistler Stroke Research Center, Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
¹³Department of Clinical Sciences, Lund University, Malmö, Sweden
¹⁴Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA
¹⁵Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, USA
¹⁶Robertson Center for Biostatistics, University of Glasgow, United Kingdom
¹⁷Department of Neurology, Rhode Island Hospital, Alpert Medical School of Brown University, Providence, RI, USA
¹⁸Department of Mathematics and Statistics, Boston University, Boston, Massachusetts, USA
¹⁹Department of Medicine, Massachusetts General Hospital, Boston, MA, USA
²⁰Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, California, USA
²¹University of Minnesota Medical School, Minneapolis, Minnesota, USA
²²Department of Medicine, Massachusetts General Hospital, Boston, MA, USA
²³Department of Medicine, Massachusetts General Hospital, Boston, MA, USA
²⁴Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, Minnesota, USA
²⁵Department of Cardiovascular Medicine, University of Minnesota, Minneapolis, Minnesota, USA
²⁶Department of Neurology, Rhode Island Hospital, Alpert Medical School of Brown University, Providence, RI, USA
²⁷Department of Neurology, Rhode Island Hospital, Alpert Medical School of Brown University, Providence, RI, USA
²⁸Department of Neurology, Rhode Island Hospital, Alpert Medical School of Brown University, Providence, RI, USA
²⁹Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
³⁰Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
³¹Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
³²Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
³³Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
³⁴Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands

Corresponding author, Steven A Lubitz, MD, MPH, Cardiac Arrhythmia Service and Cardiovascular Research Center, Massachusetts General Hospital, 55 Fruit Street, GRB 109, Boston, MA, 02114; (phone) 617-643-7339; (fax) 617-726-3852; slubitz@mgh.harvard.edu.

*Authors contributed equally
†Authors contributed equally
Abstract

**Background**—Atrial fibrillation (AF) is common and has a substantial genetic basis. Identification of individuals at greatest AF risk could minimize the incidence of cardioembolic stroke.

**Methods**—To determine whether genetic data can stratify risk for development of AF, we examined associations between AF genetic risk scores and incident AF in five prospective studies comprising 18,919 individuals of European ancestry. We examined associations between AF genetic risk scores and ischemic stroke in a separate study of 509 ischemic stroke cases (202 cardioembolic [40%]) and 3,028 controls. Scores were based on 11 to 719 common variants (≥5%) associated with AF at $P$-values ranging from $<1\times10^{-3}$ to $<1\times10^{-8}$ in a prior independent genetic association study.

**Results**—Incident AF occurred in 1,032 (5.5%) individuals. AF genetic risk scores were associated with new-onset AF after adjusting for clinical risk factors. The pooled hazard ratio for incident AF for the highest versus lowest quartile of genetic risk scores ranged from 1.28 (719 variants; 95%CI, 1.13–1.46; $P=1.5\times10^{-4}$) to 1.67 (25 variants; 95%CI, 1.47–1.90; $P=9.3\times10^{-15}$). Discrimination of combined clinical and genetic risk scores varied across studies and scores (maximum C statistic, 0.629–0.811; maximum ΔC statistic from clinical score alone, 0.009–0.017). AF genetic risk was associated with stroke in age- and sex-adjusted models. For example, individuals in the highest quartile of a 127-variant score had a 2.49-fold increased odds of cardioembolic stroke, versus those in the lowest quartile (95%CI, 1.39–4.58; $P=2.7\times10^{-3}$). The effect persisted after excluding individuals (n=70) with known AF (odds ratio, 2.25; 95%CI, 1.20–4.40; $P=0.01$).
Conclusions—Comprehensive AF genetic risk scores were associated with incident AF beyond clinical AF risk factors, with magnitudes of risk comparable to other clinical risk factors, though offered small improvements in discrimination. AF genetic risk was also associated with cardioembolic stroke in age- and sex-adjusted analyses. Efforts to determine whether AF genetic risk may improve identification of subclinical AF or distinguish stroke mechanisms are warranted.

Keywords
atrial fibrillation; stroke; genetic; risk; prediction

Atrial fibrillation (AF) is a heritable and common arrhythmia associated with substantial morbidity and economic costs. Approximately one in five ischemic strokes are attributable to cardioembolic events from AF. Strokes due to AF are associated with more disability and mortality than strokes from other etiologies. Since many strokes caused by AF are preventable with effective anticoagulation, and because AF may be undetected in some individuals, there is a critical need to identify those at greatest risk for the arrhythmia.

In recent years, risk models for AF prediction have been developed based on clinical and demographic variables. We and others have identified common genetic variants associated with AF, and some of these have been associated with incident AF and ischemic stroke after adjustment for clinical risk factors. Yet it remains unclear whether a comprehensive AF genetic risk score can facilitate identification of individuals at greatest risk for AF or cardioembolic stroke, since such individuals might benefit from stroke prevention efforts.

We therefore sought to determine whether comprehensive AF genetic risk scores are associated with incident AF beyond clinical risk factors, and might facilitate identification of individuals at greatest risk for the arrhythmia. In addition, we sought to examine whether AF genetic risk is associated with ischemic stroke, and in particular, cardioembolic stroke.

METHODS

Participants
We examined the association between AF genetic risk and incident AF in five prospective studies. Briefly, these studies were the Malmö Diet and Cancer Study (MDCS), the Multi-Ethnic Study of Atherosclerosis (MESA), the Prevention of Renal and Vascular Endstage Disease (PREVEND) study, the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER), and the Vanderbilt University de-identified DNA biobank (BioVU). We also examined the association between AF genetic risk and stroke in the Massachusetts General Hospital Genes Associated with Stroke Risk and Outcomes Study (MGH-GASROS), a hospital-based case-control study of acute ischemic stroke patients (enrolled between July 2000 and 2011) and referent individuals from the Myocardial Infarction Genetics Consortium (without a history of myocardial infarction). All stroke cases in MGH-GASROS underwent etiologic stroke subtyping in a uniform fashion, according to the Trial of Org 10172 in Acute Stroke Treatment (TOAST) criteria. Descriptions of each study are provided in the online supplement, including details on clinical risk factor and...
outcome ascertainment, genotyping, and imputation. For all analyses, samples were restricted to individuals of self-reported European ancestry. Each study was approved by its Institutional Review Board, and participants provided written informed consent.

**AF genetic risk**

To estimate genetic risk using a minimal set of single nucleotide polymorphisms (SNPs), we selected uncorrelated SNPs by pruning \(^{28}\) 2.2 million HapMap variants included in a prior independent meta-analysis of genome-wide association studies for AF from the AFGen consortium (6,707 individuals with and 53,436 without AF).\(^{15}\) We considered all SNPs that had allele frequencies ≥5% and were nominally associated with AF \((P<1\times10^{-3})\). We then selected the most significantly associated SNP within a given 250 kilobase locus that was not in linkage disequilibrium with another more significantly associated SNP at that locus \((r^2<0.1)\). In total, 719 uncorrelated SNPs were selected for construction of genetic risk scores (Supplemental Table 1).

For each individual, we calculated AF genetic risk scores by summing the dosage of each AF risk allele (ranging from 0 to 2) weighted by the natural logarithm of the relative risk for each SNP. Weights were determined in our earlier, independent meta-analysis.\(^{15}\) Thus, a genetic risk score for an individual is a single linear predictor variable. Since the optimum number of risk alleles that should be used for genetic risk scores has not been fixed, we constructed seven different scores for each individual based on the strength of association between each SNP and AF in the earlier analysis.\(^{15}\) We selected the seven different significance thresholds \textit{a priori}: \(P<1\times10^{-3}\), \(<1\times10^{-4}\), \(<1\times10^{-5}\), \(<1\times10^{-6}\), \(<1\times10^{-7}\), \(<5\times10^{-8}\), and \(<1\times10^{-8}\). Liberal inclusion of SNPs was motivated by observations that uncorrelated SNPs demonstrating less significant associations with a trait may still explain a substantial proportion of the heritability of the trait.\(^{29-32}\)

**Statistical analysis**

Within each prospective study, we used proportional hazards regression to examine associations between the different AF genetic risk scores and incident AF over a 5-year time horizon. For all incident AF analyses, person-time in each cohort began at DNA collection or baseline enrollment. Individuals were treated as censored at the time of death or loss to follow-up. Models were adjusted for variables included in a previously validated composite risk score for 5-year AF risk prediction (CHARGE-AF risk score).\(^{9}\) The composite CHARGE-AF risk score included age, height, weight, systolic and diastolic blood pressures, smoking status, antihypertensive medication use, diabetes status, heart failure status, myocardial infarction status, electrocardiographic evidence of left ventricular hypertrophy, and PR interval. Electrocardiographic variables that were not available were omitted from the scores on a study-by-study basis (left ventricular hypertrophy was unavailable in MDCS, MESA, PREVEND, and PROSPER; PR interval was unavailable in MDCS, PREVEND, PROSPER, and BioVU). Race was not included in the models since we restricted our sample to individuals of European ancestry. Proportional hazards assumptions were verified with multiplicative interaction terms between covariates and the natural logarithm of follow-up time.
For each model, we calculated goodness-of-fit statistics using Akaike’s Information Criterion, a penalized likelihood metric in which lower values indicate better fit.\textsuperscript{33} We also assessed discrimination using the $C$ statistic for time-to-event data.\textsuperscript{34} Calibration of the prediction models was assessed using the Hosmer-Lemeshow statistic modified for survival analysis.\textsuperscript{35}

In exploratory analyses we combined model parameters from each study by use of an inverse variance random-effects meta-analysis approach, and calculated heterogeneity using the $I^2$ statistic.\textsuperscript{36} We utilized a random-effects approach owing to inherent differences in study design (see supplemental methods for details). We then multiplied the summary score beta coefficient by the difference between the 12.5\textsuperscript{th} and 87.5\textsuperscript{th} percentiles of AF genetic risk scores from a common reference population (Supplemental Table 2). The resulting values estimate the relative risk comparing individuals in the highest and lowest quartiles across each study and score, in a standard fashion. The common reference population used was a pooled sample of 12,801 individuals from the Framingham Heart Study (n=2,551),\textsuperscript{37} the Atherosclerosis Risk in Communities Study (n=7,278),\textsuperscript{38} and the Cardiovascular Health Study (n=2,972)\textsuperscript{39} with genome-wide genotyping data.\textsuperscript{15}

We then examined whether AF genetic risk was associated with AF, ischemic stroke, and cardioembolic stroke in MGH-GASROS using multivariable logistic regression. Since several of the identified pruned AF SNPs were not available in the MGH-GASROS sample, we utilized proxy SNPs on the basis of linkage disequilibrium when available (Supplemental Table 1). The number of SNPs in some genetic risk scores differed slightly based on inability to identify proxies. Models were adjusted for age and sex only, because extended clinical information was not available in the referent participants. Since AF was ascertained only in stroke cases, we assumed that AF was not present among referents for analyses of AF (an assumption that would be expected to bias the results toward a null association between genetic risk and AF due to the potential for misclassified individuals who have AF among the referent sample). We then examined associations between AF genetic risk and ischemic stroke, as well as the association with the TOAST cardioembolic stroke classification (a subset of ischemic stroke). We utilized the same referent sample set for analyses of ischemic and cardioembolic stroke. Because AF may occur as a subclinical condition, we examined in exploratory analyses whether AF genetic risk scores were associated with stroke in individuals without known AF, again assuming that referent subjects did not have AF.

None of the studies in our analysis of incident AF were used in any aspect of the derivation of genetic risk or the CHARGE-AF scores. The \textit{a priori} significance threshold for all analyses was $P<0.05$ using two-sided tests. Meta-analyses were conducted using the \texttt{rmeta}\textsuperscript{40} package in R.\textsuperscript{41} Other software utilized for analyses is described in the supplement.

\section*{RESULTS}

\subsection*{AF genetic risk scores and incident AF}

Among 18,919 individuals across all studies in our analyses of incident AF, the mean age ranged from 58–75 years, and the proportion of women ranged from 47–52\%. During the 5-
year follow-up window, 1,032 (5.5%) individuals developed incident AF (Table 1). AF genetic risk scores were associated with incident AF after accounting for clinical risk factors (Supplemental Figure 1 and Supplemental Table 3). Heterogeneity of effect estimates was modest between studies. Generally, the models with the best fit included scores with between 25 and 129 SNPs, as indicated by the AIC (Supplemental Table 3).

For each of the seven groups of genetic risk scores, we estimated hazard ratios comparing individuals in the highest quartile of each genetic risk score with those in the lowest quartile. Across the genetic risk scores, those in the highest quartile had a 1.28-fold (719 SNPs; 95% CI, 1.13–1.46; P=1.5×10^-4) to 1.67-fold (25 SNPs; 95% CI, 1.47–1.90; P=9.3×10^-15) increased hazard for AF (Figure 1). C statistics for the clinical risk factor model without AF genetic risk scores ranged from 0.615 to 0.802 across cohorts (Supplemental Table 3). Adding AF genetic risk scores to the clinical risk factor model resulted in a maximum change in the C-statistic of between 0.009 and 0.017 across all cohorts and scores. The maximum change of up to 0.065 in PROSPER may have been driven by the small sample size and was considered an outlier. To illustrate the impact of clinical and genetic risk on incident AF detection, we plotted the cumulative incidence of AF stratified by dichotomized clinical risk, as well by both clinical and genetic risk together, for one representative study (MDCS) in Supplemental Figure 2.

**AF genetic risk scores and ischemic stroke**

We examined the association between AF genetic risk scores and stroke among 509 independent individuals with stroke from MGH-GASROS and 3,028 controls (Table 2). Among the stroke cases, 202 (40%) were classified as having had a cardioembolic stroke by TOAST criteria. In total, 87 (17%) individuals with ischemic stroke had documented AF.

In MGH-GASROS, modest associations between AF genetic risk scores and AF, ischemic stroke (all subtypes), and the subset of cases with cardioembolic stroke were observed using continuous genetic risk scores (Supplemental Table 4). The most significantly associated score with AF, as judged by the score with the smallest P-value, occurred with a score constructed from 127 SNPs, corresponding to SNPs with P values <1×10^-4 for associations with AF in the prior independent AFGen analysis. Individuals in the highest quartile of the 127-SNP genetic risk score had a 3.13-fold (95% CI, 1.47–7.21; P=0.005) increased odds of AF relative to those in the lowest quartile.

In the analysis of ischemic stroke cases and referent individuals, AF genetic risk scores were also modestly associated with both ischemic stroke (all subtypes) and cardioembolic stroke (Supplemental Table 3). Those in the highest quartile of the 127-SNP genetic risk score had a 1.73-fold (95% CI, 1.15–2.61; P=9.0×10^-3) increased odds of ischemic stroke, and a 2.49-fold (95% CI, 1.39–4.58; P=2.7×10^-3) increased odds of cardioembolic stroke (after excluding other stroke subtypes, Figure 2). After omitting the 87 stroke cases with known AF (70 of whom had cardioembolic strokes), the associations between AF genetic risk and both ischemic and cardioembolic stroke remained but were slightly attenuated (Supplemental Table 5). Specifically, the relative odds of ischemic stroke comparing those in the highest with those in the lowest quartile of a 127-SNP AF genetic risk score were 1.55

* Circulation. Author manuscript; available in PMC 2018 April 04.
(95% CI, 1.03–2.36; \(P=0.04\)) for ischemic stroke, and 2.25 (95% CI, 1.20–4.40; \(P=0.01\)) for cardioembolic stroke (Figure 2).

**DISCUSSION**

In our analysis of nearly 19,000 individuals of European ancestry, scores reflecting the burden of AF risk alleles were associated with 5-year risks of new-onset AF, after adjusting for clinical risk factors. Individuals in the highest quartile of the genetic scores had up to a 67% higher risk of new-onset AF than those in the lowest quartile, although incremental discrimination beyond clinical risk factors was small regardless of the number of SNPs included in the genetic risk score. In an independent sample, individuals in the highest quartile of a score comprised of 127 AF-associated genetic markers had roughly two-fold higher odds of cardioembolic stroke, compared with those in the lowest quartile after adjustment for age and sex. Associations between AF genetic risk scores and cardioembolic stroke persisted after excluding individuals with known AF.

Our findings support and extend prior observations that AF genetic risk is associated with both AF and stroke. We previously observed an association between familial AF and incident AF in the Framingham Heart Study, beyond associations for clinical risk factors.\(^1\) Subsequently, we observed an approximately 4 to 5-fold gradient in risk between those in the highest versus lowest tails of a 12-SNP AF genetic risk score (based on nine loci) in case-referent and cohort studies.\(^16\) The Women’s Genome Health Study reported an association between an AF genetic risk score based on 12 SNPs and occurrence of incident AF,\(^18\) although the AF-associated SNPs used in the analysis were identified in a previous discovery study using the same study sample. Earlier work also described associations between the top AF-associated variants on chromosomes 4q25 and 16q22 with ischemic (and in particular, cardioembolic) stroke.\(^13,26,42–44\) Recently, we and others reported a 2-fold increased hazard of AF and a 1.23-fold increased hazard of ischemic stroke for individuals in the highest versus lowest quintiles of scores based on a 12-SNP genetic risk model during an average follow-up of 14 years in the MDCS, subjects of which were included in the present analysis of incident AF.\(^19\) Thus, by using well-characterized independent study samples, our current findings extend prior reports that AF genetic risk is associated with incident AF, as well as ischemic stroke.

Our observations have three major implications. First, our finding that AF genetic risk is associated with incident AF beyond the effects observed for accepted clinical risk factors highlights the ability of common genetic variation to capture complementary information. Indeed, the 28%-67% increased risk of AF among individuals in the highest versus the lowest quartile of genetic risk is comparable to the magnitude of risk conferred by traditional clinical risk factors for AF.\(^9\) Nevertheless, even by including a large number of genetic variants and assessing associations with incident AF in large cohorts, the magnitudes of risk associated with genetic risk improved discrimination minimally beyond clinical factors. Such findings underscore the challenges of improving clinical prediction models even when including highly associated predictors.\(^45\)
Second, our observations, coupled with prior findings that AF genetic risk may be preferentially associated with cardioembolic stroke,\textsuperscript{13,42,43} raise the possibility that AF genetic risk may serve as a signature for strokes caused by thromboembolism due to AF. Our observation that AF genetic risk was associated with an increased risk of cardioembolic stroke even after excluding individuals with known AF is consistent with the hypothesis that AF genetic risk may be a clinically relevant marker for subclinical, or previously undiagnosed, AF. Although AF genetic risk has a limited impact beyond knowledge of clinical risk factors on AF prediction over a 5-year time horizon, it is possible that such genetic profiling may provide insights into stroke mechanisms and therefore screening and treatment options for secondary prevention. Future analyses are warranted to determine if AF genetic risk discriminates effectively between different stroke subtypes, to assess whether AF genetic risk can identify cryptogenic stroke patients at elevated risk for recurrent stroke due to AF, and whether estimating AF risk can enhance secondary stroke prevention efforts.

Third, our observation that genetic risk scores constructed from liberally selected SNPs were nevertheless associated with AF and AF-related stroke emphasizes the polygenic nature of AF. Therefore, true AF susceptibility variants are likely to exist even though they may not meet the stringent genome-wide significance criteria currently utilized. Future genetic discovery efforts in larger samples with better power are warranted to identify additional AF susceptibility signals. Indeed, since publication of the most recent AFGen meta-analysis,\textsuperscript{15} additional \textit{bona fide} subthreshold AF signals have been identified, and some appear to be associated with stroke.\textsuperscript{17} It remains to be determined whether future assessment of AF genetic risk based on associations derived from larger samples will enhance specificity of prediction models.

Our study should be interpreted in the context of the study design. First, all participants were of European descent, and therefore our findings may not be generalizable to individuals of other ancestral groups. Second, the genetic risk models were linear in nature with a single predictor variable, and did not account for potential non-additive genetic effects, interactions between genetic variants, or interactions between genetic variants and environmental factors. Additional modeling methods, including penalized regression or other techniques, may yield more precise genetic risk models. Third, other important determinants of AF risk were not available in our study, including plasma biomarkers such as brain natriuretic peptide.\textsuperscript{46} Similarly, in analyses of ischemic stroke, clinical covariates beyond age and sex were unavailable, so we could not evaluate whether the genetic risk score adds appreciably to prediction afforded by the CHA\textsubscript{2}-DS\textsubscript{2}-VASc score\textsuperscript{47} or individual stroke risk factors. Future studies are warranted to determine whether genetic risk adds additional information to other clinical and biomarker factors related to AF and stroke. Fourth, our genetic risk models were comprised of common SNPs genotyped in the HapMap reference populations,\textsuperscript{48} many of which are likely tag-SNPs and serve as proxies for true causal variation. Through the use of larger sample sizes and newer techniques to comprehensively assess genomic variation, such as whole genome sequencing, we anticipate better power to identify causal variants underlying AF in the future. Inclusion of causal variants in genetic risk scores may improve the specificity of the models. Fifth, the genetic predictors of prevalent stroke may not be identical to those of incident stroke due to potential survival biases. Therefore, the clinical
utility of AF genetic risk factors for identifying individuals at risk for incident stroke merits future study.

Conclusions

We observed that comprehensive AF genetic risk scores were associated with incident AF, exceeding effects of clinical risk factors, in individuals of European ancestry. We further observed that AF genetic risk is associated with both ischemic and cardioembolic stroke after adjustment for age and sex, even among individuals with cardioembolic stroke without established AF. Our findings underscore the polygenic nature of AF and the independent value of genetic information beyond clinical risk factors for the identification of individuals at risk for AF. However, although genetic risk scores are highly associated with AF, genetic information currently affords small improvements in discrimination of AF risk, and therefore does not yet need to be incorporated into routine clinical decision-making. Future clinical trials are necessary to rigorously assess whether AF genetic risk is an effective clinical marker of cardioembolic stroke etiology, and can identify individuals with subclinical AF.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Sources of funding: Please refer to supplemental material for detailed funding support. The sponsors did not have any input into the study design or conduct; data collection, management, analysis, or interpretation; nor did they influence the preparation, review, or approval of the manuscript.

Disclosures: Dr. Ellinor is a principal investigator on a grant from Bayer HealthCare to the Broad Institute.

References


28. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, 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. [PubMed: 17701901]


Thijs V. The association of the 4q25 susceptibility variant for atrial fibrillation with stroke is limited to stroke of cardioembolic etiology. Stroke; a journal of cerebral Circulation. 2010; 41:1850–1857.


## CLINICAL PERSPECTIVE

### What is new?
- Studies have identified several genetic loci associated with AF, yet it is unclear whether genetic profiling can identify individuals at greatest risk for AF or cardioembolic stroke.
- Using genome-wide data from an independent large-scale analysis, we tested comprehensive AF genetic risk scores for association with new-onset AF in five prospective studies, and with stroke in a separate stroke case-control sample.
- Genetic risk scores were associated with AF beyond established clinical risk factors, but improved prediction minimally.
- AF genetic risk was strongly associated with cardioembolic stroke, suggesting that elevated AF genetic risk might serve as a surrogate for thromboembolism from AF.

### What are the clinical implications?
- Our findings underscore the complementary information provided by both clinical and genetic factors.
- However, since genetic information currently affords small improvements in discrimination of AF risk, widespread use of genetic risk profiling does not need to be incorporated into routine clinical decision-making at this time.
- Our findings raise the possibility that AF genetic risk may serve as a signature for strokes caused by thromboembolism from AF.
- Future studies are warranted to determine whether AF genetic risk can distinguish stroke etiologic mechanisms, or identify individuals with strokes that have subclinical AF.
Figure 1.
Pooled 5-year relative hazard of incident atrial fibrillation among individuals in the highest quartile of AF genetic risk relative to those in the lowest quartile.
SNPs included in scores were derived using different thresholds of association between each SNP and atrial fibrillation in an earlier, independent study.15
Figure 2.
Risk of cardioembolic stroke in MGH-GASROS according to atrial fibrillation genetic risk. Odds ratios for cardioembolic stroke in relation to atrial fibrillation genetic risk scores among cardioembolic stroke cases and 3,028 controls. Blue histograms show distributions of genetic risk scores among cases and controls. Black dots indicate odds ratios for stroke for each quartile of genetic risk scores (bars indicate 95% confidence intervals). For panels A–C, genetic risk scores were based on 45 (A), 127 (B), and 701 (C) SNPs among 202 cardioembolic stroke cases (including 70 with known AF) and controls. For panels D–F,
genetic risk scores were based 45 (D), 127 (E), and 701 (F) SNPs among 152 cardioembolic stroke cases (none with known AF) and controls. SNP totals may not equal those used in the incident atrial fibrillation analysis since some SNPs were unavailable in MGH-GASROS, in which case proxies were used when available (Supplemental Table 1).
Table 1

Characteristics of participants included in analyses of incident atrial fibrillation.

<table>
<thead>
<tr>
<th></th>
<th>MDCS</th>
<th>MESA</th>
<th>PREVEND</th>
<th>PROSPER*</th>
<th>BioVU</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. total</td>
<td>8,226</td>
<td>2,451</td>
<td>1,624</td>
<td>5,212</td>
<td>1,388</td>
</tr>
<tr>
<td>No. incident AF</td>
<td>190</td>
<td>76</td>
<td>34</td>
<td>503</td>
<td>229</td>
</tr>
<tr>
<td>Age, years</td>
<td>59±7</td>
<td>63±10</td>
<td>58±8</td>
<td>75±3</td>
<td>60±11</td>
</tr>
<tr>
<td>Women</td>
<td>4,275 (52)</td>
<td>1,321 (52)</td>
<td>770 (47)</td>
<td>2,716 (52)</td>
<td>678 (49)</td>
</tr>
<tr>
<td>Height, cm</td>
<td>169±9</td>
<td>169±10</td>
<td>172±9</td>
<td>165±9</td>
<td>171±11</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>75±14</td>
<td>79±16</td>
<td>80±14</td>
<td>73±13</td>
<td>86±22</td>
</tr>
<tr>
<td>Systolic blood pressure, mmHg</td>
<td>145±20</td>
<td>124±20</td>
<td>135±21</td>
<td>155±22</td>
<td>131±20</td>
</tr>
<tr>
<td>Diastolic blood pressure, mmHg</td>
<td>87±10</td>
<td>75±10</td>
<td>77±10</td>
<td>84±11</td>
<td>75±30</td>
</tr>
<tr>
<td>History of smoking</td>
<td>2,513 (31)</td>
<td>1,401 (55)</td>
<td>671 (41)</td>
<td>1,388 (27)</td>
<td>619 (45)</td>
</tr>
<tr>
<td>Antihypertensive medication</td>
<td>1,799 (22)</td>
<td>840 (33)</td>
<td>362 (22)</td>
<td>3,854 (74)</td>
<td>1,339 (96)</td>
</tr>
<tr>
<td>History of diabetes</td>
<td>542 (7)</td>
<td>151 (6)</td>
<td>98 (6)</td>
<td>540 (10)</td>
<td>359 (26)</td>
</tr>
<tr>
<td>History of heart failure</td>
<td>39 (0.5)</td>
<td>52 (2)</td>
<td>4 (0.2)</td>
<td>NA</td>
<td>161 (12)</td>
</tr>
<tr>
<td>History of myocardial infarction</td>
<td>487 (9)</td>
<td>63 (3)</td>
<td>71 (4)</td>
<td>697 (13)</td>
<td>284 (20)</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation, or No. (%).

*Maximum follow-up in PROSPER was 4 years.
Table 2

Characteristics of participants of European ancestry included in analyses of ischemic stroke from MGH-GASROS and referents.

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Referents</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>509</td>
<td>3,028</td>
</tr>
<tr>
<td>Age, years</td>
<td>66.9 ± 14.4</td>
<td>42.3 ± 7.8</td>
</tr>
<tr>
<td>Women</td>
<td>214 (24.2)</td>
<td>732 (42.0)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>87 (17)</td>
<td>–</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation, or No. (%)

Stroke etiologic subtype: cardioembolic (n=202, 39%), large artery (n=114, 22%), small vessel / lacunar (n=62, 12%), other (n=124, 24%), undetermined (n=7, 1%).

P for comparison of age and sex between cases and controls <0.001.