



Published in final edited form as:

JACC Clin Electrophysiol. 2019 November ; 5(11): 1331–1341. doi:10.1016/j.jacep.2019.07.016.

Development and Validation of a Prediction Model for Atrial Fibrillation Using Electronic Health Records

Olivia L. Hulme, MD^{a,*}, Shaan Khurshid, MD^{b,*}, Lu-Chen Weng, PhD^a, Christopher D. Anderson, MD, MSc^{c,d}, Elizabeth Y. Wang, BA^a, Jeffrey M. Ashburner, PhD, MPH^{e,f}, Darae Ko, MD, MSc^g, David D. McManus, MD, MSc^h, Emelia J. Benjamin, MD^{g,i,j}, Patrick T. Ellinor, MD, PhD^{a,k}, Ludovic Trinquart, PhD^{i,l}, Steven A. Lubitz, MD, MPH^{a,k}

^aCardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA

^bCardiology Division, Massachusetts General Hospital, Boston, Massachusetts, USA

^cCenter for Genomic Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

^dJ.P. Kistler Stroke Research Center, Massachusetts General Hospital, Boston, Massachusetts, USA

^eDivision of General Internal Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

^fDepartment of Medicine, Harvard Medical School, Boston, Massachusetts, USA

^gDepartment of Medicine, Boston University Medical Center, Boston, Massachusetts, USA

^hDepartments of Medicine and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, Massachusetts, USA

ⁱBoston University and National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA

^jSections of Preventive Medicine and Cardiovascular Medicine, Department of Medicine, Boston University School of Medicine, Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts, USA

^kCardiac Arrhythmia Service, Massachusetts General Hospital, Boston, Massachusetts, USA

^lDepartment of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA

Abstract

Background: Early diagnosis of atrial fibrillation (AF) and treatment with anticoagulation may prevent strokes.

Address for correspondence: Steven A. Lubitz, MD, MPH, Cardiac Arrhythmia Service and Cardiovascular Research Center, Massachusetts General Hospital, 55 Fruit Street, GRB 109, Boston, MA 02114; P:617-643-7339; F:617-726-3852; slubitz@mgh.harvard.edu.

*Contributed equally to the manuscript

Tweet: #ehr estimation of #afib risk is feasible, and risk is associated with incident stroke. Risk-guided screening for #afib warrants evaluation.

Objectives: To determine whether AF risk can be estimated accurately using routinely ascertained features in the electronic health record (EHR) and whether AF risk associates with stroke.

Methods: Using a multi-institutional EHR, we identified 412,085 individuals aged 45–95 without prevalent AF between 2000–2014. We derived and validated a prediction model for five-year AF risk using split-sample validation, and compared model performance with other methods of AF risk assessment.

Results: Within five years, 14,334 individuals developed AF. In the derivation sample (n=7,216 AF events/206,042 total), the optimal risk model included: sex, age, race, smoking, height, weight, diastolic blood pressure, hypertension, hyperlipidemia, heart failure, coronary heart disease, valvular disease, prior stroke, peripheral arterial disease, chronic kidney disease, hypothyroidism, and quadratic terms for height, weight, and age. In the validation sample (n=7,118 AF events/206,043 total) the AF risk model demonstrated good discrimination (C-statistic 0.777, 95% CI 0.771–0.783) and calibration (0.99, 95% CI 0.96–1.01). Model discrimination and calibration were favorable to CHARGE-AF (C-statistic 0.753, 95% CI 0.747–0.759; calibration slope 0.72, 95% CI 0.71–0.74), C₂HEST (C-statistic 0.754, 95% CI 0.747–0.762; calibration slope 0.44, 95% CI 0.43–0.45), and CHA₂DS₂-VAsC (C-statistic 0.702, 95% CI 0.693–0.710; calibration slope 0.37, 95% CI 0.36–0.38). AF risk discriminated incident stroke (n=4,814, C-statistic 0.684; 95% CI 0.677–0.692) and stroke within 90 days of incident AF (n=327, C-statistic 0.789; 95% CI 0.764–0.814).

Conclusions: A model developed in a real-world EHR predicted AF accurately and stratified stroke risk. Incorporating AF prediction into EHRs may enable risk-guided screening for AF. Condensed Abstract: Early diagnosis of atrial fibrillation (AF) and treatment with anticoagulation may prevent strokes. We sought to determine whether AF risk can be estimated accurately using routinely ascertained features in the electronic health record (EHR). Within a large multi-institutional EHR sample including over 400,000 individuals, we derived and validated a prediction model for five-year AF risk and compared model performance to alternative methods of AF risk estimation. The model exhibited good discrimination and calibration, performing favorably to CHARGE-AF, C₂HEST, and CHA₂DS₂-VAsC. Accurate estimation of AF risk in a real-world EHR is feasible and may enable risk-guided screening for AF.

Keywords

Atrial fibrillation; stroke; risk prediction; electronic health record

INTRODUCTION

Atrial fibrillation is projected to affect over 12 million individuals in the United States by 2050 (1) and is a leading cause of stroke (2). Strokes caused by AF are especially disabling and confer an increased risk of death compared to strokes of other etiologies (3–6).

Although anticoagulation decreases the risk of stroke in patients with AF by about 60% (7), stroke may be the first manifestation of AF (8–11). Early identification of patients with AF may enable prophylactic anticoagulation, thereby preventing strokes.

Several clinical trials have demonstrated the feasibility of screening individuals for AF (12–15), but none have explicitly performed screening using a risk-guided approach, which may

be more effective. Data from community-based research cohorts demonstrate that new-onset AF can be predicted with reasonable accuracy (16–19). Yet the extent to which AF risk can be accurately estimated in clinical settings, and the relations between estimated AF risk and stroke, are unclear.

Electronic health records are well-suited for assessing the potential utility of estimating AF risk in clinical practice given their longitudinal nature, robust capture of risk factor and outcome data, and links to healthcare providers. Important risk factors for AF are routinely ascertained in healthcare settings, including age, race, height, weight, systolic and diastolic blood pressure, smoking status, diabetes, myocardial infarction and heart failure. If feasible, EHR-based risk prediction may allow for individualized management via point of care estimation, or facilitate efficient deployment of population health management initiatives for individuals at elevated risk. We utilized the EHR from a large healthcare system to develop and validate a novel prediction model for AF and assess whether predicted risk of AF was associated with stroke.

METHODS

Study sample

Study subjects were identified from the EHR using the Partners HealthCare System Research Patient Data Registry (RPDR), a data warehouse spanning nearly 7 million individuals and 7 hospitals including the Massachusetts General Hospital, Brigham and Women's Hospital, Newton-Wellesley Hospital, North Shore Medical Center, Faulkner Hospital, McLean Hospital, and Spaulding Rehabilitation Center (20). Detailed medical record information was collected for individuals with at least one outpatient visit in each of two consecutive years between January 2000 and December 2014, with longitudinal follow-up data collected through 04/29/2017. We selected a two-year baseline ascertainment window *a priori* to allow for data entry by a provider on more than one occasion, thereby increasing the probability of prevalent conditions being ascertained in the EHR. Individuals included in the study were aged 45–95 years at the first eligible baseline encounter. Individuals without follow-up, or with missing weight, height, systolic blood pressure, and diastolic blood pressure data at baseline were omitted. The characteristics of excluded individuals are found in Online Table 1. Individuals with AF at baseline were removed from the analysis (Figure 1). Study protocols complied with the Declaration of Helsinki and were approved by the Massachusetts General Hospital Institutional Review Board.

Ascertainment of clinical features

Data extracted from the EHR included 1) diagnostic codes (ICD9 and ICD10), 2) procedure codes (CPT), 3) medications, 4) cardiology test reports, 5) discharge summaries, 6) clinic notes, and 7) vital status (which is regularly updated based on the hospital registrar and Social Security Death Index). We ascertained AF utilizing a validated algorithm (20), comprising diagnostic and procedure codes, electrocardiogram reports, and medications to determine the presence of AF or atrial flutter. The positive predictive value (PPV) of the algorithm previously has been reported to be 88%, comparing favorably to other previously utilized methods of AF ascertainment in the EHR (20).

Potential risk factors were selected *a priori* based on known or proposed associations with AF and included sex, age, race, smoking status, height, weight, systolic and diastolic blood pressure, hypertension, diabetes mellitus, hyperlipidemia, heart failure, coronary heart disease, valvular disease, stroke or TIA, myocardial infarction, peripheral arterial disease, systemic and cerebral atherosclerosis, chronic kidney disease, thyrotoxicosis, and hypothyroidism. Age, demographic, anthropometric, vital sign, and vital status data were obtained directly from RPDR. We treated race as a dichotomous variable as has been previously reported (21), since we had few non-white individuals in our sample. We defined other comorbidities using an iterative, rule-based approach based on diagnostic or procedure codes, and medications (Online Table 2). Feature selection was considered satisfactory when, upon medical record review, the PPV was 85% (22,23) (the PPV ranged from 88–99% across all risk factors). We further examined pairwise agreement between two independent reviewers (O.H., E.W.) for stroke, which was strong (Cohen’s $\kappa = 0.8$).

Baseline covariates included the closest measurement prior to the start date within the two-year window for continuous variables. Dichotomous covariates were carried forward if present during or prior to the ascertainment period and assumed to persist, and assumed absent if they were not present at the beginning of follow-up.

Statistical analysis

Derivation of AF risk model—We randomly partitioned the overall sample of 412,085 individuals into derivation (n=206,042) and validation (206,043) sets. For each individual, person-time began after the initial two-year window during which baseline covariates were ascertained. For example, an individual in whom the first visit during the two-year window occurred in 2000 would have a follow-up start date of 01/01/2002. Person-time was censored at the first occurrence of death, last follow-up, or five years. We fit multivariable Cox proportional hazard models with a backwards elimination approach to retain predictors of incident AF within five years. This approach retained variables that optimized the model fit, as assessed by minimizing the Akaike Information Criterion (AIC), a penalized likelihood metric (24). Remaining variables included sex, age, race, smoking status, height, weight, diastolic blood pressure, hypertension, hyperlipidemia, heart failure, coronary heart disease, valvular disease, stroke/TIA, peripheral arterial disease, chronic kidney disease, and hypothyroidism. After variable selection, we introduced quadratic terms iteratively for each remaining continuous variable and retained the term if model fit improved further. The proportional hazards assumption for variables in the final model was assessed by inspecting Schoenfeld residuals (25).

Split-sample validation of AF risk model—We applied the best-fitting model from the derivation set to the validation set, using the parameter coefficients obtained from the derivation set to derive a weighted score for each individual. We tested the association between the score and incident AF within five years using proportional hazards regression. Person-time was censored at last follow-up, death, or five years. We evaluated the association of the score with incident AF using the Wald χ^2 statistic. We assessed discrimination of the score using the c-statistic for time-to-event data (26) and evaluated the calibration slope of the models by regressing incident AF on the linear predictor of the score

(27), and by plotting the predicted and observed risks of incident AF. The predicted five-year AF risk was estimated using the following formula: $1-s_0^{\exp(\sum\beta X - \sum\beta Y)}$ where s_0 is the average AF-free survival probability at five years, β is the regression coefficient, X is the level for each covariate, and Y is the mean value for each covariate. We displayed the cumulative risk of events using the Kaplan-Meier method, in which we stratified five-year predicted AF risk into categories of low (<2.5%), intermediate (2.5–5%), and high (>5%) risk as previously performed (16). We further assessed the risk of incident AF in the intermediate and high-risk groups, relative to the low risk group, by fitting proportional hazards models in which each group was entered as a covariate.

We compared the score from the best-fitting multivariable model to the Cohorts for Heart and Aging Research in Genomic Epidemiology AF (CHARGE-AF) (16), C₂HEST (28), and CHA₂DS₂-VASc scores (29). Although originally derived and validated to predict stroke in patients with AF, we included CHA₂DS₂-VASc as a comparator because it is easily memorized by clinicians, and some investigators have applied the score explicitly for AF prediction (30,31). Coefficients used to define the CHARGE-AF score include: factor (coefficient); age (0.508), race (0.465), height (0.248), weight (0.115), systolic blood pressure (0.197), diastolic blood pressure (–0.101), smoking (0.359), hypertension (0.349), diabetes (0.237), heart failure (0.701), myocardial infarction (0.496). We did not attempt to attribute medication use to hypertension in the EHR owing to diverse indications for many agents, and therefore substituted a diagnosis of hypertension for treatment of hypertension in the CHARGE-AF score. We defined a modified C₂HEST score based on EHR-derived features by summing one point each for coronary artery disease, pulmonary disease, hypertension, hyperthyroidism, and two points each for age of at least 75 years, and heart failure. We defined the CHA₂DS₂-VASc score based on EHR-features by summing one point each for an age between 65 and 74 years, congestive heart failure, hypertension, diabetes, vascular disease, female sex, and two points each for age of at least 75 years, or a prior stroke, TIA, or systemic embolism (29).

We regressed incident AF on each score and compared Wald χ^2 statistics, discrimination, and calibration for each as outlined above. We estimated the five-year risk of AF based on the CHARGE-AF score with the following formula: $1-0.9718412736^{\exp(\sum\beta X - 12.5815600)}$ where β is the regression coefficient and X is the level for each covariate (16). The five-year risk of AF based on the CHA₂DS₂-VASc score was estimated using the baseline hazard and mean covariate estimates from the validation sample. To assess the performance of the EHR-derived score based on age, we performed sensitivity analyses in which we regressed incident AF among individuals within the validation set aged 45–65 versus individuals aged > 65.

Association between AF risk and stroke—We examined the association between AF risk using our EHR-derived score and incident ischemic stroke within five years of follow-up in the validation set, after omitting individuals with prevalent stroke. Associations between the EHR-derived score and stroke were assessed using proportional hazards regression as outlined above. Person-time was censored at last follow-up, death, or five years. We also fit models in which we regressed the outcome on the predicted five-year AF risk categories of low, intermediate, and high as defined above. We then repeated the above

analysis with stroke within 90 days antecedent to a new diagnosis of AF (a surrogate for stroke presenting as the initial manifestation of AF) as the outcome of interest. In these models, person-time was censored at last follow-up, death, AF, 90 days after stroke, or five years. In an exploratory analysis, we also fit models with stroke within 90 days after a new diagnosis of AF as the outcome of interest. Associations between predicted AF risk groups and each outcome were displayed using the Kaplan-Meier method. We assessed discrimination of these models as described above. For graphical purposes, we further plotted the observed five-year event risks (derived from the Kaplan-Meier curves) of AF, stroke, and stroke within 90 days prior to incident AF for all estimated values of five-year predicted risk of AF in the validation set.

Two-sided p-values <0.05 were considered statistically significant. All analyses were performed with R version 3.2.2, with packages survival (32), rms (33), and survcomp (34,35).

RESULTS

The analysis included 7,216 incident AF cases within 5 years among 206,042 individuals in the derivation set and 7,118 cases among the 206,043 persons in the validation set. Baseline characteristics of individuals in the derivation and validation sets are displayed in Table 1. The average age was 61 years, 58% were women, and 85% were white.

Derivation and validation of AF prediction model

The following variables were selected as predictors of AF: male sex, age, race, smoking history, height, weight, diastolic blood pressure, hypertension, hyperlipidemia, heart failure, coronary heart disease, valvular disease, previous stroke/TIA, peripheral arterial disease, chronic kidney disease, and hypothyroidism. Results of univariable associations between candidate risk factors and incident AF at five years in the derivation set are shown in Table 2. The final multivariable prediction model is displayed in Table 3. Introducing quadratic terms for height, weight, and age further improved the model fit.

Comparison of EHR-derived model with CHARGE-AF, C₂HES₂, and CHA₂DS₂-VASc scores for AF risk

The EHR-derived score was significantly associated with five-year risk of incident AF in the 206,043 individual validation set from the Partners HealthCare System, and performed favorably compared to the CHARGE-AF, C₂HES₂, and CHA₂DS₂-VASc scores as indicated by the Wald χ^2 , c-statistic, and calibration slope (Figure 2 and Online Table 3). The C₂HES₂ score was associated with incident AF and demonstrated comparable discrimination to the CHARGE-AF score, but was not as well-calibrated as either the EHR-derived or CHARGE-AF scores (Online Table 3). The distribution of predicted five-year risk of incident AF from the EHR-derived score, and accompanying calibration plots are depicted in Figure 2.

Relative to those with low (<2.5%) predicted AF risk, individuals with intermediate (2.5–5%) risk had a 3.0-fold increased hazard for incident AF (95% CI 2.75–3.27, p<0.01), and individuals with high (>5%) risk had a 10.1-fold increased hazard for incident AF (95% CI

9.41–10.8, $p < 0.01$). Associations between AF risk and incident AF persisted in sensitivity analyses in which the EHR-derived score was assessed in individuals within the validation set subgrouped by age (Online Table 4). The cumulative risk of AF stratified by low, intermediate, and high five-year predicted AF risk is displayed in Figure 3.

Incident stroke risk

Overall, 4,814 incident strokes occurred within five years of follow-up among the 198,300 individuals in the validation cohort. Of the 6,623 individuals with a diagnosis of AF, 327 (4.9%) were diagnosed with a stroke within 90 days prior. The EHR-derived score was associated with an increased risk of incident stroke, and with stroke within 90 days antecedent to an AF diagnosis (Online Table 5). Five-year predicted risk of AF discriminated both incident stroke (C-statistic 0.684, 95% CI 0.677–0.692) and stroke within 90 days prior to an AF diagnosis (C-statistic 0.789, 95% CI 0.764–0.814). Comparisons with CHARGE-AF, C₂HES_T and CHA₂DS₂-VASc scores are provided in Online Table 5. Results of an exploratory analysis examining stroke within 90 days after an AF diagnosis are also presented in Online Table 5.

Relative to those with low predicted AF risk, individuals with intermediate risk had a 1.85-fold increased hazard for incident stroke (95% CI 1.69–2.03, $p < 0.01$), and a 2.22-fold increased hazard for stroke within 90 days prior to an AF diagnosis (95% CI 1.46–3.36, $p < 0.01$). A high predicted AF risk conferred a 4.47-fold increased hazard for incident stroke (95% CI 4.12–4.85, $p < 0.01$), and 10.5-fold increased hazard for stroke within 90 days prior to an AF diagnosis (95% CI 7.61–14.4, $p < 0.01$). The cumulative probabilities of incident stroke and stroke within 90 days prior to incident AF are displayed in Figure 3. Individuals with a high (>5%) predicted five-year risk of AF had an estimated 6.3% (95% CI 6.0–6.5) probability of developing a stroke at five years and 0.63% (95% CI 0.55–0.71) probability of stroke within 90 days prior to an AF diagnosis at five years. Observed five-year probabilities of incident AF, incident stroke, and stroke within 90 days prior to an AF diagnosis depicted across the distribution of predicted AF risk are shown in the Central Illustration.

DISCUSSION

In a large multicenter EHR, prediction of new-onset AF using routinely ascertained clinical variables was feasible and stratified stroke risk. Strokes occurring shortly before a new diagnosis of AF – potentially representing the initial manifestation of AF – occurred in about 5% of individuals and were highly associated with estimated five-year AF risk. In aggregate, our data suggest that population-based estimation of AF risk in healthcare systems has the potential to aid the early identification of AF to prevent strokes.

Our findings have three major implications. First, a risk model comprised of clinical factors readily accessible in the EHR and routinely obtained in the outpatient setting can predict five-year risk of AF. Our results demonstrate the predictive power of a large EHR data set, and the potential for implementation of risk estimation. For individuals, risk estimates can be calculated and displayed within the EHR to serve as clinical decision support tools, enabling clinicians to integrate the output from increasingly complex models (e.g., five year risk of AF) into clinical care without the need to perform calculations themselves.

Automated risk prediction can also guide population health management by identifying at-risk populations most likely to benefit from large-scale interventions to prevent or diagnose AF. It is possible that use of different EHR-derived variables, more predictors, and alternative statistical methods (e.g., LASSO, machine learning) may predict AF with even greater accuracy.

Second, incident AF risk is associated with and has some ability to discriminate stroke risk. Whereas models trained on ischemic stroke might have better accuracy than models trained on AF, our observations suggest that estimating AF risk will also identify individuals at risk for stroke. Presumably, strokes presenting as the initial manifestation of AF are preventable if AF is identified early, given the effectiveness of oral anticoagulation (7). Recent studies have demonstrated that screening for AF is feasible (12–15), but the effectiveness of such screening approaches for stroke prevention has not been demonstrated. Furthermore, concerns regarding misdiagnosis, follow-up testing, and inappropriate utilization of oral anticoagulation and bleeding have led to conflicting guideline endorsements for AF screening (36,37). A risk-based approach to AF screening has the potential to maximize yield and minimize misdiagnosis, and planned future studies investigating the comparative effectiveness of risk-based screening strategies relative to non-risk guided approaches will be critical in determining the value added by risk guidance. AF risk stratification may also enable preventative interventions to improve modifiable AF risk factor profiles (38,39) which may lead to prevention of AF altogether.

Third, use of locally-derived prediction models may be better suited for implementation in real-world clinical settings as compared to external risk models developed in large community-based research cohorts. For example, and similar to our findings, previous work has demonstrated poor calibration of the CHARGE-AF model when applied to an EHR-based sample (40). Our findings suggest that the predictive performance of our model owes substantially to its derivation within a large, multi-institutional EHR, accounting for patient characteristics typically encountered in a common clinical setting in which risk estimation might be utilized. Although our EHR-derived model demonstrated favorable discrimination and calibration, future implementation efforts will be necessary to determine whether the improvement in prediction we observed is clinically meaningful, cost-effective, and justifies the work required to develop prediction models across institutions using locally derived variables rather than utilize existing scores (i.e., CHARGE-AF) or other surrogates of AF risk (e.g., age thresholds, presence or absence of comorbid conditions such as heart failure, etc). Evaluation of models derived in specific populations of interest (e.g., elderly patients, individuals with recent cardiovascular events, genetic ancestry subgroups, etc) is also warranted. Though our efforts were limited to AF prediction, our findings imply that population health management interventions utilizing risk prediction may be most effective if they utilize source data from institutions, healthcare settings, or enterprises in which a particular intervention is being planned, with the inherent tradeoff that they may be less generalizable outside of the source population from which they are derived.

LIMITATIONS

Our study should be interpreted in the context of the study design. The sample is largely of European ancestry and from a single New England metropolitan area, limiting the generalizability of our findings to more diverse healthcare settings. Future prediction models based on more racially diverse samples found in other healthcare settings may improve the applicability of this score to widespread populations. The derivation of our model in a general population of ambulatory patients in a large healthcare network may limit application to hospitalized patients, patients with recent or severe cardiovascular comorbidity, or clinics without access to EHR data. However, development and validation of our model in a large sample does not necessarily preclude its use once validated in smaller systems. Measurement error of clinical features is inherent with EHR or other large-scale database research. For example, although our AF algorithm has previously been shown to have a positive predictive value of 88% for AF and has compared favorably to other methods of AF ascertainment previously utilized in large registry and cohort studies (20), approximately 10% of AF patients may have been misclassified. Misclassification may have disadvantaged scores based on highly adjudicated covariates and endpoints derived in research settings, such as the CHARGE-AF score. Similarly, although our stroke algorithm predominantly comprised codes for ischemic events, misclassification may have occurred leading to inclusion of some primary hemorrhagic events. We did not attempt to manually adjudicate events as either ischemic or hemorrhagic. Refinement in electronic phenotyping methods may improve the accuracy of certain prediction estimates. Data are limited to those captured within the Partners HealthCare EHR. Events occurring in individuals receiving care in non-Partners health settings were not accessible. Moreover, individuals who were diagnosed with AF may have had more health care encounters and differential ascertainment of AF risk factors. AF is often clinically undiagnosed and unrecognized, and we may have underestimated the incidence of AF. Reliance on EHR data limits our ability to account for the effects of certain risk factors including alcohol use, diet and physical activity. Risk factors in the current study were defined at baseline, although recent evidence suggests that a change in risk factor profile over time may have important predictive value in AF (41).

CONCLUSIONS

In conclusion, the five-year risk of AF can be estimated using readily available clinical factors in the EHR. Additionally, estimated AF risk is associated with incident stroke including stroke prior to a new diagnosis of AF. Future studies are warranted to examine whether incorporating AF risk estimation into routine clinical care may lead to early diagnosis of AF and stroke prevention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

OLH and SAL conceptualized the paper. OLH and SAL did the literature search. LCW, LT, SK and OLH performed the data analysis. Figures and tables were prepared by OLH, LCW, and SK. All authors contributed to data

interpretation. OLH, SK, and SAL wrote the manuscript and all other authors provided critical comments on the manuscript.

Funding

This work was supported by Doris Duke Charitable Foundation (New York, New York) Clinical Research Mentorship Grants 2016077 (Hulme and Lubitz) and 2017039 (Wang and Lubitz), and a Doris Duke Charitable Foundation Clinical Scientist Development Award 2014105 (Lubitz). Additional support includes NIH (Bethesda, Maryland) grants R01HL139731 (Lubitz), 2R01HL092577 (Ellinor, Benjamin), 1R01HL128914 (Ellinor and Benjamin), 1P50HL120163 (Benjamin), R01NS103924 and K23NS086873 (Anderson), R01HL104156 and K24HL105780 (Ellinor); Massachusetts General Hospital Center for Genomic Medicine Catalyst Award (Anderson); and American Heart Association (Dallas, Texas) 18SFRN34250007 (Lubitz, Anderson), 18SFRN34150007 (Benjamin), Postdoctoral Fellowship Award 17POST33660226 (Weng), and Established Investigator Award 13EIA14220013 (Ellinor).

Disclosures

Dr. Lubitz has received research support from Boehringer Ingelheim, and Bristol-Myers Squibb / Pfizer, and Bayer AG, and has received consulting support from Abbott, Bristol-Myers Squibb / Pfizer, and Quest Diagnostics. Dr. Ashburner has received research support from Boehringer Ingelheim and Bristol-Myers Squibb / Pfizer. Dr. McManus has received sponsored research support from Boehringer Ingelheim and Bristol-Myers Squibb / Pfizer and has received consulting support from Bristol-Myers Squibb, Pfizer, and Boston Biomedical Associates. Dr. Anderson has received consulting support from ApoPharma, Inc. Dr. Benjamin serves on the NHLBI Observational Safety Monitoring Board for the CARDIA study and has served as the Associate Editor for *Circulation* until 6/30/2016. Dr. Ellinor is a principal investigator on a Bayer AG grant to the Broad Institute related to the development of new therapeutics for atrial fibrillation and receives consulting support from Bayer AG, Novartis, and Quest Diagnostics. The remaining authors have nothing to disclose.

Abbreviations List:

AF	atrial fibrillation
EHR	electronic health record
ICD9	International Classification of Diseases, 9 th revision
ICD10	International Classification of Diseases, 10 th revision
CPT	Current procedural terminology
TIA	transient ischemic attack

References

1. Miyasaka Y, Barnes ME, Gersh BJ et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation* 2006;114:119–25. [PubMed: 16818816]
2. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;22:983–8. [PubMed: 1866765]
3. Petty GW, Brown RD Jr., Whisnant JP, Sicks JD, O’Fallon WM, Wiebers DO. Ischemic stroke subtypes : a population-based study of functional outcome, survival, and recurrence. *Stroke* 2000;31:1062–8. [PubMed: 10797166]
4. Lamassa M, Di Carlo A, Pracucci G et al. Characteristics, outcome, and care of stroke associated with atrial fibrillation in Europe: data from a multicenter multinational hospital-based registry (The European Community Stroke Project). *Stroke* 2001;32:392–8. [PubMed: 11157172]
5. Marini C, De Santis F, Sacco S et al. Contribution of atrial fibrillation to incidence and outcome of ischemic stroke: results from a population-based study. *Stroke* 2005;36:1115–9. [PubMed: 15879330]

6. Rutten-Jacobs LC, Arntz RM, Maaijwee NA et al. Long-term mortality after stroke among adults aged 18 to 50 years. *JAMA* 2013;309:1136–44. [PubMed: 23512060]
7. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Ann Intern Med* 2007;146:857–67. [PubMed: 17577005]
8. Jaakkola J, Mustonen P, Kiviniemi T et al. Stroke as the First Manifestation of Atrial Fibrillation. *PLoS One* 2016;11:e0168010. [PubMed: 27936187]
9. Borowsky LH, Regan S, Chang Y, Ayres A, Greenberg SM, Singer DE. First Diagnosis of Atrial Fibrillation at the Time of Stroke. *Cerebrovasc Dis* 2017;43:192–199. [PubMed: 28208140]
10. Lubitz SA, Yin X, McManus DD et al. Stroke as the Initial Manifestation of Atrial Fibrillation: The Framingham Heart Study. *Stroke* 2017;48:490–492. [PubMed: 28082669]
11. Lin HJ, Wolf PA, Benjamin EJ, Belanger AJ, D'Agostino RB. Newly diagnosed atrial fibrillation and acute stroke. The Framingham Study. *Stroke* 1995;26:1527–30. [PubMed: 7660392]
12. Morgan S, Mant D. Randomised trial of two approaches to screening for atrial fibrillation in UK general practice. *Br J Gen Pract* 2002;52:373–4, 377–80. [PubMed: 12014534]
13. Fitzmaurice DA, Hobbs FD, Jowett S et al. Screening versus routine practice in detection of atrial fibrillation in patients aged 65 or over: cluster randomised controlled trial. *BMJ* 2007;335:383. [PubMed: 17673732]
14. Halcox JPJ, Wareham K, Cardew A et al. Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation: The REHEARSE-AF Study. *Circulation* 2017;136:1784–1794. [PubMed: 28851729]
15. Steinhubl SR, Waalen J, Edwards AM et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. *JAMA* 2018;320:146–155. [PubMed: 29998336]
16. 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. [PubMed: 23537808]
17. Schnabel RB, Sullivan LM, Levy D et al. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. *Lancet* 2009;373:739–45. [PubMed: 19249635]
18. Schnabel RB, Aspelund T, Li G et al. Validation of an atrial fibrillation risk algorithm in whites and African Americans. *Arch Intern Med* 2010;170:1909–17. [PubMed: 21098350]
19. Sinner MF, Stepas KA, Moser CB et al. B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies. *Europace* 2014;16:1426–33. [PubMed: 25037055]
20. Khurshid S, Keaney J, Ellinor PT, Lubitz SA. A Simple and Portable Algorithm for Identifying Atrial Fibrillation in the Electronic Medical Record. *Am J Cardiol* 2016;117:221–5. [PubMed: 26684516]
21. Alonso A, Roetker NS, Soliman EZ, Chen LY, Greenland P, Heckbert SR. Prediction of Atrial Fibrillation in a Racially Diverse Cohort: The Multi-Ethnic Study of Atherosclerosis (MESA). *J Am Heart Assoc* 2016;5.
22. Newton KM, Peissig PL, Kho AN et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. *J Am Med Inform Assoc* 2013;20:e147–54. [PubMed: 23531748]
23. Ritchie MD, Denny JC, Crawford DC et al. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. *Am J Hum Genet* 2010;86:560–72. [PubMed: 20362271]
24. Akaike H A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 1974;19:716–723.
25. Grambsch PM, Therneau TM Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81:515–526.
26. Schroder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* 2011;27:3206–8. [PubMed: 21903630]

27. van Houwelingen HC. Validation, calibration, revision and combination of prognostic survival models. *Stat Med* 2000;19:3401–15. [PubMed: 11122504]
28. Li YG, Pastori D, Farcomeni A et al. A Simple Clinical Risk Score (C2HEST) for Predicting Incident Atrial Fibrillation in Asian Subjects: Derivation in 471,446 Chinese Subjects, With Internal Validation and External Application in 451,199 Korean Subjects. *Chest* 2018.
29. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest* 2010;137:263–72. [PubMed: 19762550]
30. Saliba W, Gronich N, Barnett-Griness O, Rennert G. Usefulness of CHADS2 and CHA2DS2-VASc Scores in the Prediction of New-Onset Atrial Fibrillation: A Population-Based Study. *Am J Med* 2016;129:843–9. [PubMed: 27012854]
31. Barkas F, Elisaf M, Korantzopoulos P, Tsiara S, Liberopoulos E. The CHADS2 and CHA2DS2-VASc scores predict atrial fibrillation in dyslipidemic individuals: Role of incorporating low high-density lipoprotein cholesterol levels. *Int J Cardiol* 2017;241:194–199. [PubMed: 28442233]
32. Therneau T A Package for Survival Analysis in S, R package version 2.38. <https://CRAN.R-project.org/package=survival>. 2015.
33. Harrell FE Jr.. rms: Regression Modeling Strategies. R package version 5.1–2. <http://CRAN.R-project.org/package=rms>. 2018.
34. Schroeder M, Culhane A, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* 2011;27:3206–3208. [PubMed: 21903630]
35. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G. A comparative study of survival models for breast cancer prognostication on microarray data: does a single gene beat them all? *Bioinformatics* 2008;24:2200–2208. [PubMed: 18635567]
36. Kirchhof P, Benussi S, Kotecha D et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016;37:2893–2962. [PubMed: 27567408]
37. Force USPST, Curry SJ, Krist AH et al. Screening for Atrial Fibrillation With Electrocardiography: US Preventive Services Task Force Recommendation Statement. *JAMA* 2018;320:478–484. [PubMed: 30088016]
38. Pathak RK, Elliott A, Middeldorp ME et al. Impact of CARDIOrespiratory FITness on Arrhythmia Recurrence in Obese Individuals With Atrial Fibrillation: The CARDIO-FIT Study. *J Am Coll Cardiol* 2015;66:985–96. [PubMed: 26113406]
39. Pathak RK, Middeldorp ME, Meredith M et al. Long-Term Effect of Goal-Directed Weight Management in an Atrial Fibrillation Cohort: A Long-Term Follow-Up Study (LEGACY). *J Am Coll Cardiol* 2015;65:2159–69. [PubMed: 25792361]
40. Kolek MJ, Graves AJ, Xu M et al. Evaluation of a Prediction Model for the Development of Atrial Fibrillation in a Repository of Electronic Medical Records. *JAMA Cardiol* 2016;1:1007–1013. [PubMed: 27732699]
41. Chao TF, Lip GYH, Liu CJ et al. Relationship of Aging and Incident Comorbidities to Stroke Risk in Patients With Atrial Fibrillation. *J Am Coll Cardiol* 2018;71:122–132. [PubMed: 29325634]

Clinical Perspectives:**Competency in Patient Care:**

Risk of AF can be estimated accurately using routine clinical factors. Patients at high risk for AF may benefit from targeted interventions to diagnose AF early and initiate treatment with oral anticoagulation to prevent stroke.

Competency in Systems-Based Practice:

Implementation of automated AF risk estimation into EHRs as clinical decision support may allow clinicians to leverage the precision and objectivity afforded by complex models in order to provide individualized risk estimation to patients at the point of care. Additionally, AF risk prediction at the population level may inform population health management through identification of at-risk populations most likely to benefit from large-scale interventions to prevent or diagnose AF.

Translational Outlook:

Further investigation is required to assess whether routine utilization of AF risk estimation to inform AF screening interventions improves diagnostic yield, minimizes false positives, and improves outcomes in AF.

- Initially seen in **outpatient setting** within Partners HealthCare system between **2000-2014**
- **≥1 outpatient visit** in each of two consecutive calendar years (i.e., resulting in ≥ 2 outpatient visits) between January 2000 and December 2014
- **Ages 45-95** years at first documented encounter

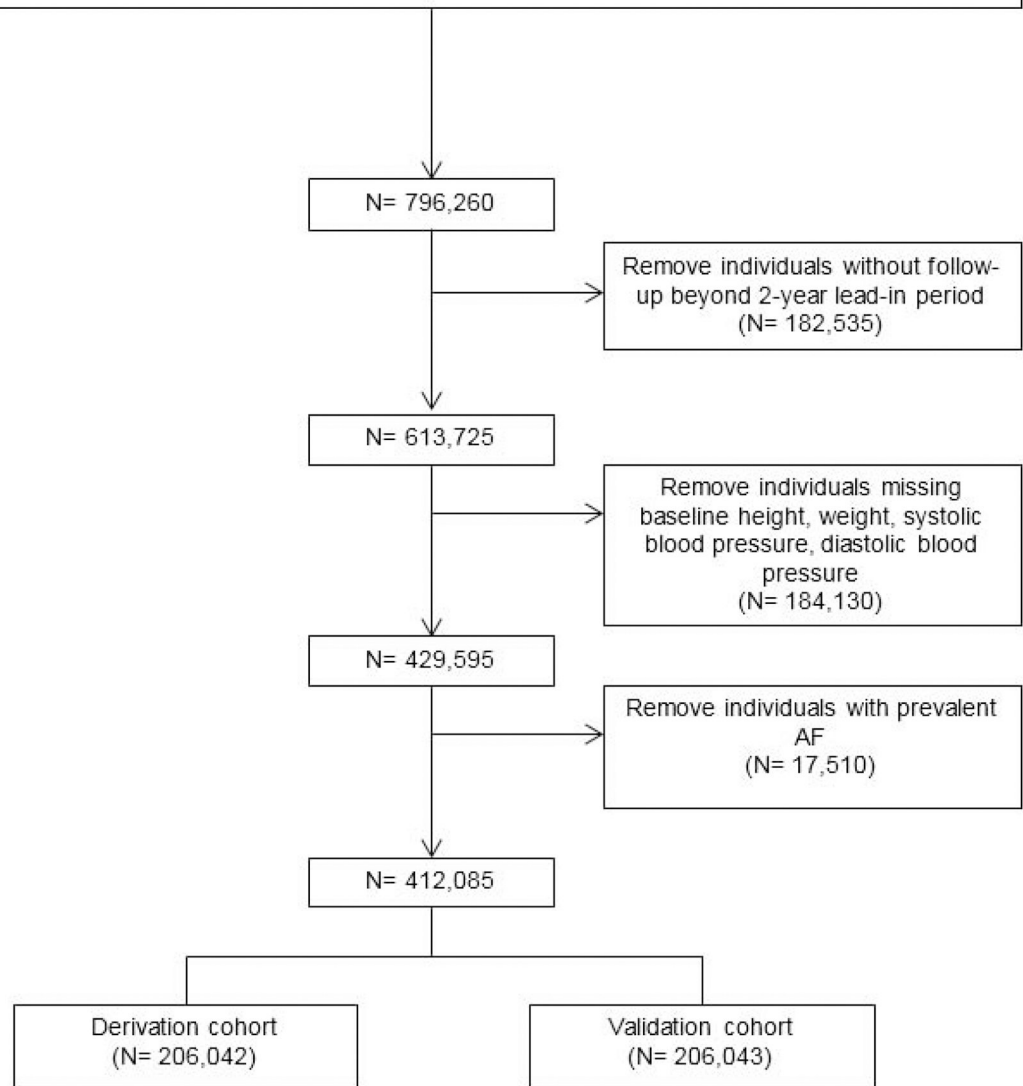


Figure 1. Flow diagram depicting patient inclusion.
AF = atrial fibrillation.

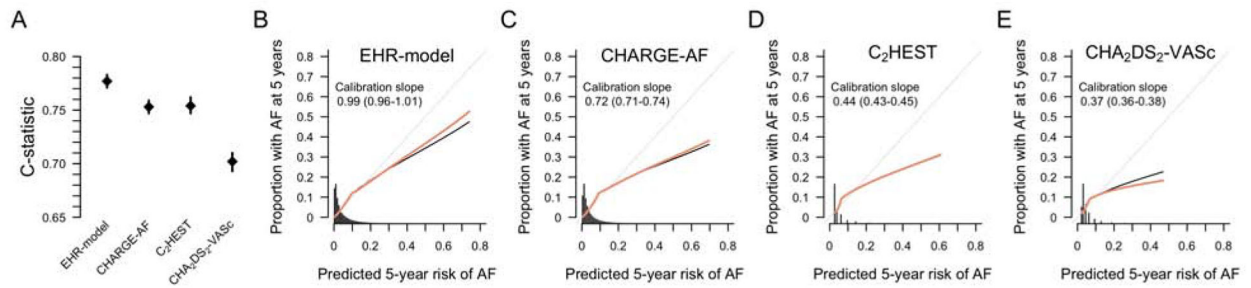


Figure 2. Discrimination and calibration of scoring systems for predicting incident atrial fibrillation.

Panel A displays the C-statistic point estimate and 95% confidence interval for each of the three prediction models compared for predicting incident atrial fibrillation. Panels B-E display the predicted (x-axis) versus observed (y-axis) probability of incident atrial fibrillation for the EHR-derived score, CHARGE-AF score, C₂HEST, and CHA₂DS₂-VASc scores, respectively. For Panels B-E, the histograms represent the distribution of five-year predicted risk of atrial fibrillation. Perfect calibration is represented by the gray diagonal line through the origin. The black line corresponds to the observed calibration, and the orange line to the optimism-corrected calibration. Calibration slopes (and 95% confidence intervals) are depicted on the graphs. Calibration plots generated using the rms package (33).

CHARGE-AF = Cohorts for Heart and Aging Research in Genomic Epidemiology Atrial Fibrillation; EHR = electronic health record.

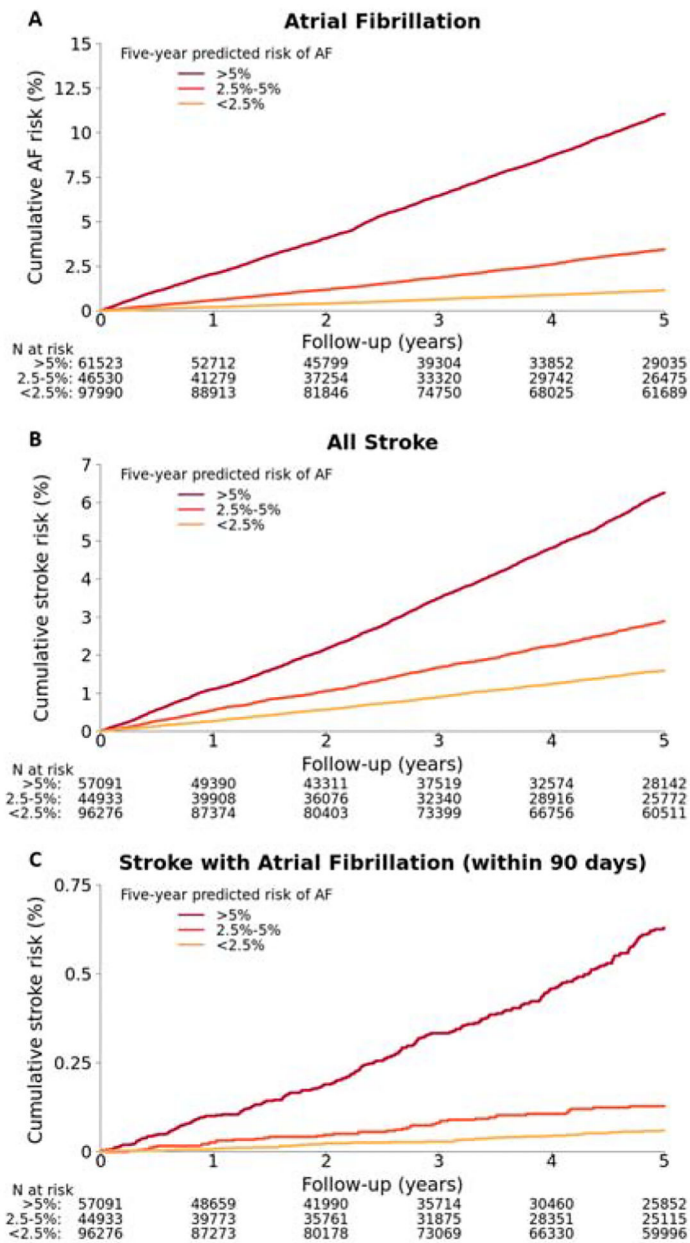
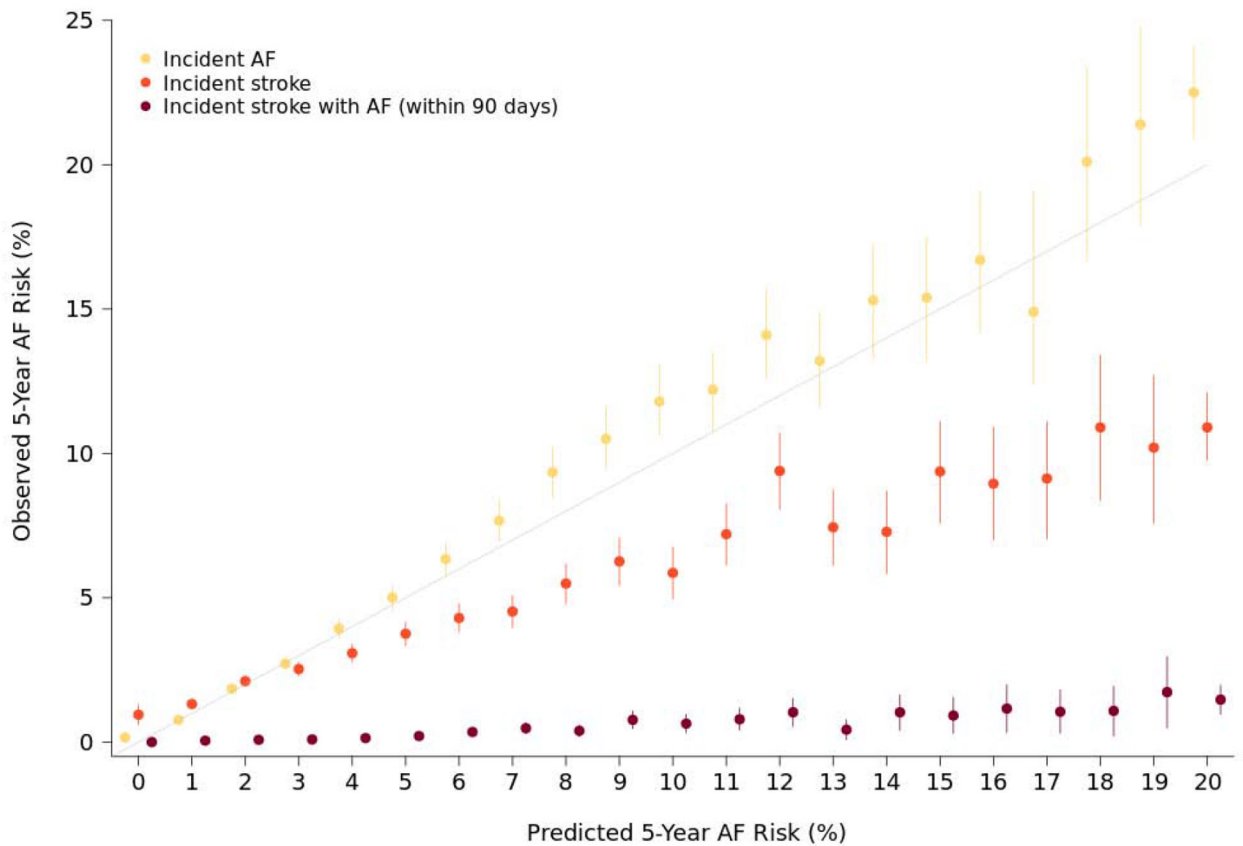


Figure 3. Cumulative incidence of atrial fibrillation and stroke stratified by five-year predicted risk of atrial fibrillation using EHR-derived score. Cumulative incidence of A) atrial fibrillation, B) stroke, and C) stroke within 90 days of an atrial fibrillation diagnosis, stratified by five-year predicted risk of atrial fibrillation. Panel A includes all 206,043 individuals from the validation set, while panels B and C include 198,300 individuals without prevalent stroke.



Central Illustration. Observed event rates stratified by predicted five-year risk.

Observed five-year incidence of atrial fibrillation, stroke, and stroke within 90 days of an atrial fibrillation diagnosis, stratified by percentage point increase in five-year predicted risk of atrial fibrillation. Five-year predicted risk of atrial fibrillation was rounded to the nearest whole number. Perfect correlation between predicted and observed atrial fibrillation incidence is represented by the gray line. Point estimates indicate observed five-year event risk, and bars represent 95% confidence intervals.

Table 1.

Characteristics of individuals in the derivation and validation sets.

	Derivation (n=206,042)	Validation (n=206,043)
	% or Mean (SD)	% or Mean (SD)
Demographics		
Female Sex, No. (%)	58	58
Age	61 (11)	61 (11)
Race		
White	85	85
Black	3.9	4.0
Hispanic/Latino	3.1	3.1
Asian	2.3	2.2
Other	1.2	1.2
Mixed	0.038	0.035
Unknown	4.0	4.1
Potential AF risk factors		
Smoking	9.7	9.8
Height, cm	167 (10)	167 (10)
Weight, kg	79 (19)	79 (19)
Systolic blood pressure, mmHg	129 (17)	129 (17)
Diastolic blood pressure, mmHg	76 (10)	76 (10)
Hypertension	28	29
Diabetes	9.4	9.4
Hyperlipidemia	30.2	30.2
Heart failure	3.1	3.2
Coronary heart disease	9.3	9.4
Valvular disease	1.4	1.5
Previous stroke/TIA	3.8	3.8
Vascular disease*		
Myocardial infarction	4.1	4.0
Peripheral artery disease	3.9	3.8
Systemic atherosclerosis	1.4	1.4
Cerebral atherosclerosis	3.5	3.5
Chronic kidney disease	3.4	3.4
Thyrotoxicosis	1.5	1.5
Hypothyroidism	8.7	8.6

* Note that presence of at least one of these constituted a criterion for CHA₂DS₂-VASc

Table 2.

Univariable associations between candidate risk factors and five-year risk of incident atrial fibrillation in the derivation sample.

Baseline characteristic (N=206,042)	Hazard ratio (95% CI) for incident atrial fibrillation
Demographics	
Female Sex	0.54 (0.52–0.57)
Age (HR per 10-year increase)	2.19 (2.14–2.23)
Race: White vs. Nonwhite	1.47 (1.36–1.58)
Potential AF risk factors	
Smoking	1.34 (1.25–1.44)
Height (per 10-cm increase)	1.19 (1.16–1.22)
Weight (per 15-kg increase)	1.17 (1.15–1.19)
Systolic blood pressure 130 mmHg	1.36 (1.30–1.43)
Diastolic blood pressure 80 mmHg	0.83 (0.81–0.85)
Hypertension	1.83 (1.75–1.92)
Diabetes	1.82 (1.71–1.94)
Hyperlipidemia	1.47 (1.40–1.54)
Heart failure	3.85 (3.56–4.15)
Coronary heart disease	2.72 (2.57–2.87)
Valvular disease	3.38 (3.03–3.78)
Previous stroke/TIA	2.29 (2.10–2.50)
Vascular disease*	
Myocardial infarction	2.73 (2.53–2.94)
Peripheral artery disease	2.75 (2.55–2.98)
Systemic atherosclerosis	3.08 (2.73–3.48)
Cerebral atherosclerosis	2.62 (2.41–2.85)
Chronic kidney disease	2.56 (2.34–2.80)
Thyrotoxicosis	0.91 (0.75–1.10)
Hypothyroidism	0.95 (0.87–1.03)

* Note that presence of at least one of these constituted a criterion for CHA₂DS₂-VASc

Table 3.

Final multivariable model for five-year risk of incident atrial fibrillation in the derivation sample.

Baseline characteristic (N=206,042)	Estimated β (SE)
Demographics	
Female Sex	-0.137 (0.035)
Age (per 10-year increase)	1.494 (0.125)
Age ² (per 10-year increase)	-0.048 (0.009)
Race: White vs. Nonwhite	-0.208 (0.039)
AF risk factors	
Smoking	0.152 (0.039)
Height (per 10-cm increase)	-0.231 (0.279)
Height ² (per 10-cm increase)	0.012 (0.008)
Weight (per 15-kg increase)	-0.050 (0.059)
Weight ² (per 15-kg increase)	0.021 (0.005)
Diastolic blood pressure \geq 80 mmHg	-0.104 (0.025)
Hypertension	0.106 (0.030)
Hyperlipidemia	-0.156 (0.030)
Heart failure	0.563 (0.045)
Coronary heart disease	0.210 (0.036)
Valvular disease	0.487 (0.060)
Previous stroke/TIA	0.132 (0.047)
Peripheral artery disease	0.126 (0.044)
Chronic kidney disease	0.279 (0.049)
Hypothyroidism	-0.138 (0.044)

For an individual with baseline characteristics X_i , the predicted 5-year risk of AF can be calculated as $1 - S_0 \exp(\sum \beta_i X_i - 6.728)$, where $S_0 = 0.9712209$ (average AF-free survival probability at 5 years). In this equation, the β reported in the table for age, height, and weight must be divided by the reported increment (e.g., $-0.050/15$ for weight in kg).