Evaluation of a Prediction Model for the Development of Atrial Fibrillation in a Repository of Electronic Medical Records

Matthew J. Kolek, M.D., M.S.C.I.[*], **Amy J. Graves, S.M., M.P.H.**[†], **Meng Xu, M.S.**[†], **Aihua Bian, M.P.H.**[†], **Pedro Luis Teixeira, M.S.**[‡], **M. Benjamin Shoemaker, M.D., M.S.C.I.**[*], **Babar Parvez, M.D., M.S.C.I.**[*], **Hua Xu, Ph.D.**[§], **Susan R. Heckbert, M.D., Ph.D.**[‖], **Patrick T. Ellinor, M.D., Ph.D.**[¶], **Emelia J. Benjamin, M.D., Sc.M.**[#,††], **Alvaro Alonso, M.D., Ph.D.**[‡‡], **Joshua C. Denny, M.D.**[‡], **Karel G.M. Moons, Ph.D.**[†,§§], **Ayumi K. Shintani, Ph.D., M.P.H.**[‖‖], **Frank E. Harrell Jr., Ph.D.**[†], **Dan M. Roden, M.D.**[*], and **Dawood Darbar, M.D.**[*,¶¶]

[*]Division of Cardiovascular Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee [†]Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee [‡]Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, Tennessee [¶¶]Division of Cardiology, University of Illinois at Chicago, Chicago, IL [§]School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, Texas [‖]Department of Epidemiology, University of Washington, Seattle, Washington [¶]Cardiac Arrhythmia Service, Massachusetts General Hospital, Boston, Massachusetts [#]National Heart Lung and Blood Institute's and Boston University's Framingham Heart Study, Framingham, Massachusetts [††]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts [‡‡]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia [§§]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands [‖‖]Department of Biostatistics, Osaka University, Suita, Japan

Abstract

Importance—Atrial fibrillation contributes to substantial morbidity, mortality, and healthcare expenditures. Accurate prediction of incident atrial fibrillation would enhance patient management and potentially improve outcomes.

Objective—We aimed to validate the atrial fibrillation risk prediction model originally developed by the CHARGE-AF investigators utilizing a large repository of electronic medical records.

Address for correspondence: Dawood Darbar, M.D., 840 S. Wood St., 920S (MC 715), University of Illinois at Chicago, Chicago, IL 61612, Office: 312 413 8870, darbar@uic.edu.

Design—Using a database of de-identified medical records, we conducted a retrospective electronic medical record study of subjects without atrial fibrillation followed in Internal Medicine outpatient clinics at our institution. Individuals were followed for incident atrial fibrillation from 2005 until 2010. Adjusting for differences in baseline hazard, we applied the CHARGE-AF Cox proportional hazards model regression coefficients to our cohort. A simple version of the model, with no ECG variables was also evaluated.

Setting—Outpatient clinics at a large academic medical center.

Participants—33,494 subjects of age 40 years, white or African American, and no previous history of atrial fibrillation.

Predictors—Predictors in the model included age, race, height, weight, systolic and diastolic blood pressure, treatment for hypertension, smoking status, diabetes, heart failure, history of myocardial infarction, left ventricular hypertrophy, and PR interval.

Main outcome—Incident atrial fibrillation.

Results—The median age was 57 years (25th to 75th percentile: 49 to 67), 57% of patients were women, 85.7% were white, 14.3% were African American. During the mean follow-up period of 4.8 ± 0.85 years, 2455 (7.3%) subjects developed atrial fibrillation. Both models had poor calibration in our cohort, with under-prediction of AF among low-risk subjects and over-prediction of AF among high-risk subjects. The full CHARGE-AF model had a C-index of 0.71 (95% confidence interval [CI]: 0.70 to 0.72) in our cohort. The simple model had similar discrimination (C-index: 0.71, 95% CI: 0.70 to 0.72, P = 0.71 for difference between models).

Conclusions and Relevance—Despite reasonable discrimination, the CHARGE-AF models showed poor calibration in our EMR cohort. Our study highlights the difficulties of applying a risk model derived from prospective cohort studies to an EMR cohort and suggests that these AF risk prediction models be used with caution in the EMR setting. Future risk models may need to be developed and validated within EMR cohorts.

Introduction

Atrial fibrillation (AF), the most common sustained cardiac arrhythmia, is becoming increasingly prevalent in the Western world.[1,2] It is projected that the number of patients with AF in the United States will roughly double by the year 2050, to an estimated 12–16 million.[2,3] AF is associated with significant morbidity,[4,5] mortality,[6–9] decreased quality of life,[10] and increased healthcare expenditures.[11,12] Developing strategies for the prediction and prevention of AF in high risk individuals remains an underexplored and important area of research.[13]

Recently, the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE)-AF investigators developed and validated a risk model for prediction of incident AF.[14] The model was developed using pooled data from prospective cohort studies including the Atherosclerosis Risk in Communities (ARIC) Study, Cardiovascular Health Study (CHS), and Framingham Heart Study (FHS), and was validated in the Age Gene/Environment Susceptibility-Reykjavik (AGES) and Rotterdam Study (RS) cohorts. The model is

especially well-suited for primary care settings since it does not require laboratory or echocardiographic variables.

Novel risk models should be validated (i.e., evaluated in new settings) before they are incorporated into routine care.[15] Electronic medical records (EMRs) are becoming ubiquitous in clinical practice, and one potential use for EMR repositories in etiologic research is to validate existing risk prediction models. Additionally, risk models are unlikely to be widely utilized unless they can be incorporated into EMR systems. We therefore evaluated the CHARGE-AF risk model for incident AF in a large, de-identified EMR repository.

Methods

Study Population

Study subjects were selected from a de-identified version of the Vanderbilt University Medical Center EMR. This resource, termed the Synthetic Derivative,[16] consists of de-identified medical records of Vanderbilt University Medical Center inpatients and outpatients, and as of December 2015 contained nearly 2.6 million individuals. The Synthetic Derivative consists of the de-identified version of the Vanderbilt Medical Center EMR that has been judged by the Vanderbilt University Institutional Review Board as falling under the designation of "non-human subjects" under Title 45, Code of Federal Regulations, Part 46; therefore, this study as well as other Synthetic Derivative research was deemed exempt by the Vanderbilt University Institutional Review Board.

In order to ensure that they had adequate follow-up, individuals met criteria for a 'medical home' model, in which they were followed in a Vanderbilt Internal Medicine Clinic with at least 3 visits documented within a 24-month period.[17] Other criteria for entry into the study included age 40 years, self-identified white or African American race, and no known previous history of AF as of December 2005. Individuals were excluded if they had billing codes for AF or mention of AF in ECG impressions or structured problem lists as determined by natural language processing.[18] Individuals were also excluded from the study if they had International Classification of Diseases, version 9 (ICD9), or Current Procedural Terminology (CPT) codes for heart transplant at the beginning of the follow-up period. Data studied included all inpatient and outpatient ICD9 and CPT codes, ECGs, and problem lists, and manual review included all inpatient and outpatient records.

Outcome: Incident AF

The follow-up period for incident AF was from December 31, 2005 until December 31, 2010. Ascertainment of incident AF was accomplished by using a validated algorithm which incorporates natural language processing and billing codes, as previously described.[18] This automated algorithm was optimized through multiple reiterations with sensitivity analyses with different cutoffs and manual review of medical records until a positive predictive value for AF >95% was achieved.[18] Cases were defined by natural language processing of cardiologist-interpreted ECG impressions, 4 occurrences of ICD9 codes for AF, or AF instances recorded in the problem list. In order to be classified as free of AF, an individual

record could not contain mention of AF in ECG impressions or structured problem lists, or ICD9 codes for AF or atrial flutter. The four ICD9 codes used in the automated AF algorithm are the most commonly used codes for AF or atrial flutter. Subjects with 1–3 ICD9 codes for AF or atrial flutter were excluded from the study cohort. In effect, these steps resulted in a more sensitive method for excluding AF at baseline and a more specific method for ascertaining incident AF during follow-up. We assessed the accuracy of the AF algorithm in the current study by manually reviewing the full EMR in a random set (enriched for incident AF by selecting roughly equal proportions of subjects with and without AF) of 200 incident AF cases and controls, block randomized to case/control ordering, as is common for validating EMR phenotypes.[19]

Predictors

CHARGE-AF model predictors were ascertained from records available from January 1st to December 31st 2005. Sex, race, age, weight, height, body mass index and systolic and diastolic blood pressure were directly extracted from structured fields in the Synthetic Derivative. History of myocardial infarction, heart failure, and diabetes mellitus were determined by using ICD9 codes, incorporating laboratory values and medication records.[20] Treatment for hypertension was assessed using a previously validated algorithm incorporating medication records in the Synthetic Derivative.[21,22] This algorithm was previously shown to have a sensitivity and positive predictive value of 88% and 93%, respectively. PR interval and left ventricular hypertrophy were obtained from outpatient ECG reports. Current smoking status was determined by using an existing algorithm with a reported positive predictive value of 93% in Vanderbilt EMR.[23]

Statistical Analysis

Baseline characteristics present at the beginning of the observation period are presented in terms of median and interquartile range for continuous variables and frequencies with percentage for categorical variables. Number of non-missing values for each variable is also given. Single imputation of missing values was performed using predictive mean matching.[24] To assess the validity of using single rather than multiple imputation, we conducted 5 separate imputations. These resulted in almost identical C-indices (Supplementary Table 1).

To evaluate the CHARGE-AF model in the Vanderbilt Synthetic Derivative cohort, we used the Cox proportional hazards model derived in the CHARGE-AF study (Table 1).[14] As in the original CHARGE-AF study, we compared discrimination using the full model and a 'simple' model, which did not incorporate ECG variables. Comparisons were made using the rcorrp.cens function in the R Hmisc package, which provides a test of whether one model gives predictions that are more concordant than the other in a way that preserves pairing of 2 predictions in the pair of patients under consideration.

We adjusted for differences in baseline hazard between the original CHARGE-AF data set and our cohort by constructing the linear predictor for each observation using the original coefficients from the CHARGE-AF model and the covariate values in the Vanderbilt Synthetic Derivative cohort data, with replacement of the mean covariate values to reflect

the Vanderbilt data.[15] We then examined the calibration and discrimination of the CHARGE-AF model in our cohort. To assess calibration, the observed risk of developing AF during the study period was plotted against predicted risk.[25] The generated curve was compared against a hypothetical ideal curve with slope = 1 and intercept = 0. For discrimination, which measures the ability of the model to distinguish subjects who will develop AF from those who will not, we calculated the continuous time C-index with censoring.[26,27] We also constructed a Kaplan-Meier curve showing cumulative incidence of AF in our cohort. Statistical analyses were conducted using R software (version 3.0.2, The R Foundation for Statistical Computing, Vienna, Austria). A two-sided P value of 0.05 was considered statistically significant. The detailed statistical code is included in the Supplementary Methods.

The funding sources had no role in the study.

Results

Based on the prospectively defined criteria, 33,494 individuals were included in the analysis. Baseline characteristics for the study cohort are presented in Table 2. Median age was 57 years (25th to 75th percentile: 49 to 67), 57% of subjects were women, and 14.3% were African American. During the mean follow-up period of 4.8 (SD: 0.85) years, 2455 (7.3%) subjects developed AF. A Kaplan-Meier curve for the cumulative incidence of AF is shown in Figure 1.

The calibration curve plotting the predicted probability of AF-free survival and the observed AF-free survival in our cohort indicated that the model had poor fit (Figure 2). There was under-prediction of AF for individuals with lower (<15%) predicted probabilities. This included the vast majority of subjects in the cohort. There was also over-prediction of AF for higher probabilities. The 10th and 90th percentiles for predicted probability of incident AF were 0.005 and 0.179, respectively. Over this range, the maximum calibration error was 0.0526 (5.3%).

The full CHARGE-AF model had a C-index of 0.708 (95% confidence interval: 0.699 to 0.718) in our cohort. The simple model, which did not include ECG predictors (PR interval and left ventricular hypertrophy), had discrimination similar to the full model (C-index 0.709, 95% confidence interval [CI]: 0.699 to 0.718, P = 0.70 for difference between the models).

Because the primary study outcome, incident AF during the 5-year follow-up period, was ascertained by an automated algorithm, we conducted a manual review of individual EMRs to assess the accuracy of the algorithm. Two hundred records were selected at random after enrichment of the sample for incident AF. Manual chart review revealed that 88 individuals had incident AF during the 5-year follow-up period (Supplementary Table 2). The sensitivity of the AF algorithm was 96.5% (95% confidence interval: 90.1% to 98.8%), specificity was 94.8% (95% CI: 89.1% to 97.6%), positive predictive value was 93.2% (95% CI: 85.9% to 96.8%), and negative predictive value was 97.3% (95% CI: 92.4% to 99.1%). The 6 subjects who were not identified as having incident AF by the automated algorithm but had AF by

manual chart review were each identified as having the arrhythmia based on mentions in clinic notes, although AF billing codes were not present. It was less clear why 3 subjects were identified as having incident AF by the algorithm but not confirmed by manual review, although incomplete chart review and incorrect billing codes are possible explanations.

Since increasing age is the strongest predictor of AF in the CHARGE-AF model, we evaluated whether the model predicts incident AF better than knowledge of age alone. We generated a Cox proportional hazards model for the development of AF at 5 years in our cohort with age as the only dependent variable, resulting in a C-index of 0.684 (95% CI: 0.674 to 0.694). The C-indices for the Vanderbilt age-only model and the externally validated CHARGE-AF model could not be compared directly since the models are not nested. We also generated a scatterplot of age and cumulative predicted probability of AF at 5 years (Figure 3). This showed a broad distribution of AF risk across the spectrum of ages, indicating that age alone is an imprecise predictor of AF risk.

Discussion

We evaluated the CHARGE-AF full and simple risk models in a large cohort of individuals within the Vanderbilt EMR repository. The full model had a C-index of 0.708 (95% CI: 0.699 to 0.718). The simple CHARGE-AF model, which did not include ECG variables as predictors, had similar discrimination, with a C-index of 0.709, (95% CI: 0.699 to 0.718). However, calibration for both models was poor in our cohort, indicating a failure of validation. Our study represents a novel use of an EMR repository to evaluate an existing AF risk model and illustrates the limitations of applying a model developed in prospective cohort studies to a 'real-world' EMR context.

Our findings illustrate several potential uses for EMR repositories in biomedical research: First, EMR repositories could serve as an inexpensive and efficient complement to community cohort studies for the development of prediction models. Second, an EMR repository could be used as an independent cohort to externally evaluate an existing model, as we did here. In fact, risk models are unlikely to be widely utilized unless they can be incorporated into EMR systems. Finally, given that EMRs are integrated into clinical practice, prediction models could be incorporated into these systems to prospectively identify individuals at high risk for AF or other diseases, with the ultimate goal of developing individualized preventive strategies. Specifically, improved knowledge about individual AF risk might enable aggressive risk factor modification, more intensive screening, diligent evaluation at first sign of symptoms, and modification of stroke risk.

The CHARGE-AF model has recently been tested in additional community cohorts. When applied to the Multi-Ethnic Study of Atherosclerosis cohort, the simple CHARGE-AF model had good discrimination (C-index 0.779, 95% CI: 0.744 to 0.814) but suboptimal calibration, with over-prediction of AF in higher-risk subjects.[30] When applied to over 24,000 participants in the European Prospective Investigation of Cancer Norfolk cohort, the CHARGE-AF simple model again had good discrimination (C-index 0.81, 95% CI: 0.75 to 0.85) but also poor calibration.[31] These studies, along with our current findings, illustrate the difficulty of applying a risk model to diverse populations, particularly in an EMR setting.

Large, prospective community cohort studies such as ARIC, CHS, FHS, AGES, and RS have been instrumental in identifying risk factors for common diseases.[32–36] In the case of coronary heart disease, the discoveries of these studies have been translated into strategies for primary and secondary prevention that have made important impacts on cardiovascular morbidity and mortality.[37–42] EMR repository studies might emerge as an important complement to prospective cohort studies. Although they might have important shortcomings including inadequate disease classifications and missing data, they also offer several attractive advantages that could be leveraged for etiologic research. Because data in EMR repositories are collected during routine clinical care, the cost of these studies is small relative to prospective cohort studies. Notably, the National Heart Lung and Blood Institute recently announced an initiative to use large EMR studies to enhance the clinical utility and reproducibility of clinical research while reducing costs.[43] We propose that automated algorithms could be deployed within EMR repositories to prospectively identify and 'flag' individuals at high risk for AF or other common diseases. These data could then be used to guide primary prevention strategies. We have pursued a similar strategy to identify and genotype individuals at risk for receiving medications that have pharmacogenetic variations in efficacy.[44] While no specific treatment for the primary prevention of AF has been established,[13] angiotensin converting enzyme inhibitors and angiotensin II receptor blockers have been associated with a decreased incidence of AF in post-hoc analyses of randomized clinical trials and in retrospective cohort studies.[28,296]

Our study does have several important limitations. One of these relates to data collection and ascertainment: although all patient variables were entered into the Vanderbilt EMR prospectively, the nature of the EMR with data entered by multiple users might lead to more inaccuracies when compared with carefully curated prospective cohorts such as those studied by the CHARGE-AF Consortium. Although there was a 12-month "run-in" period from January 1st, 2005 until December 31st, 2005 during which subjects with incident AF were excluded from the final study cohort, we did not conduct rigorous screening (e.g., ambulatory ECG monitoring) to exclude baseline AF. It is also possible that patients with AF may not seek medical attention for over 12 months and a longer run-in period may be needed. Predictor and outcome variables were extracted from the EMR repository using automated algorithms. While many of these (e.g., age, sex, race, height, weight, body mass index, blood pressure, and PR interval) are structured data fields within the repository, other predictors such as diabetes mellitus and heart failure depended on billing codes, laboratory or note data, and medication records, potentially resulting in important inaccuracies. The automated algorithm for assignment of incident AF relied primarily on natural language processing of ECG impressions, problem lists, and clinic notes, but also used billing codes. We conducted a manual review of a sample of charts that demonstrated the algorithm for ascertainment of AF status performed well with high sensitivity and specificity. It is possible that our AF algorithm with a high diagnostic accuracy that exceeds those used by the CHARGE-AF cohorts, may be one potential explanation for poor calibration of the risk prediction score in an EMR setting.

Because individuals in our study were not prospectively enrolled and followed for incident AF, our analysis is prone to indication bias, wherein individuals who developed AF may have had more clinical encounters than those who did not. Because subjects weren't

followed at pre-specified time intervals, our results might be influenced by loss of follow-up. Additionally, the classification of incident AF might have been inaccurate if individuals sought care outside of Vanderbilt, such that their AF diagnoses were not captured in the Vanderbilt EMR.

Though the CHARGE-AF models had satisfactory discrimination in our cohort, calibration was poor. One potential reason for the poor calibration is important differences in the characteristics of the CHARGE-AF discovery cohorts and our cohort (see Table 1 in Alonso et al[14]). It is difficult to directly compare baseline characteristics between our EMR cohort and the CHARGE-AF cohorts as 5 separate community cohorts were used to formulate the CHARGE-AF model. However, overall, subjects in our cohort tended to be younger, heavier, more likely to be smokers and use anti-hypertensive therapy, and generally sicker than those enrolled in the CHARGE-AF cohorts. Since age is the most important predictor of AF, the inclusion of younger subjects in our study might account, at least in part, for the poor calibration of the CHARGE-AF model in our EMR cohort. We chose the age cut-off for inclusion in our study based on previous findings that the incidence of AF begins to increase rapidly after the age of 40. Our goal was to be able to apply the CHARGE-AF model in a primary care setting and we postulate that it is particularly important to identify relatively young patients at risk for AF because they might benefit most from preventative measures, although this remains to be proven in prospective studies. Additional potential causes for failure of the models to accurately predict the development of AF include loss to follow-up (including death), differences in how predictors were defined, and more structured surveillance for incident AF in the CHARGE-AF cohorts (greater sensitivity). Due to the nature of our study, we were unable to include comprehensive data on death (i.e., using the Social Security Death Index) other that what was available in the EMR.

Conclusions

We evaluated the CHARGE-AF full and simple risk models in a large cohort of individuals without a history of AF at baseline in the Vanderbilt EMR repository. The models performed poorly in our EMR cohort, illustrating the difficulty of applying risk models developed within prospective cohort studies to a 'real-world' EMR context. Risk models for the development of AF or other complex disorders are unlikely to be widely utilized in clinical care unless they can be incorporated into EMR systems. It follows that risk models should be derived from and validated in different EMR cohorts, with the goal of prospectively and automatically identifying subjects at high risk for AF and implementing personalized strategies for primary prevention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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(8):837–847. [PubMed: 24345399]

2. 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(18):2370–2375. [PubMed: 11343485]

3. 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(11):119–125. [PubMed: 16818816]

4. Friberg J, Buch P, Scharling H, Gadsbphioll N, Jensen GB. Rising rates of hospital admissions for atrial fibrillation. Epidemiology. 2003; 14(6):666–672. [PubMed: 14569181]

5. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. Ann Intern Med. 1999; 131(7):492–501. [PubMed: 10507957]

6. Stewart S, Hart CL, Hole DJ, McMurray JJ. A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. Am J Med. 2002; 113(5):359–364. [PubMed: 12401529]

7. Krahn AD, Manfreda J, Tate RB, Mathewson FA, Cuddy TE. The natural history of atrial fibrillation: incidence, risk factors, and prognosis in the Manitoba Follow-Up Study. Am J Med. 1995; 98(5):476–484. [PubMed: 7733127]

8. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. Circulation. 1998; 98(10):946–952. [PubMed: 9737513]

9. Wolf PA, Mitchell JB, Baker CS, Kannel WB, D'Agostino RB. Impact of atrial fibrillation on mortality, stroke, and medical costs. Arch Intern Med. 1998; 158(3):229–234. [PubMed: 9472202]

10. Hamer ME, Blumenthal JA, McCarthy EA, Phillips BG, Pritchett EL. Quality-of-life assessment in patients with paroxysmal atrial fibrillation or paroxysmal supraventricular tachycardia. Am J Cardiol. 1994; 74(8):826–829. [PubMed: 7942563]

11. Le Heuzey JY, Paziaud O, Piot O, et al. Cost of care distribution in atrial fibrillation patients: the COCAF study. Am Heart J. 2004; 147(1):121–126. [PubMed: 14691429]

12. Kim MH, Johnston SS, Chu BC, Dalal MR, Schulman KL. Estimation of total incremental health care costs in patients with atrial fibrillation in the United States. Circ Cardiovasc Qual Outcomes. 2011; 4(3):313–320. [PubMed: 21540439]

13. De Caterina R, Husted S, Wallentin L, et al. New oral anticoagulants in atrial fibrillation and acute coronary syndromes: ESC Working Group on Thrombosis-Task Force on Anticoagulants in Heart Disease position paper. J Am Coll Cardiol. 2012; 59(16):1413–1425. [PubMed: 22497820]

14. Calkins H. Catheter ablation to maintain sinus rhythm. Circulation. 2012; 125(11):1439–1445. [PubMed: 22431885]

15. Benjamin EJ, Chen PS, Bild DE, et al. Prevention of atrial fibrillation: report from a national heart, lung, and blood institute workshop. Circulation. 2009; 119:606–618. [PubMed: 19188521]

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(2):e000102. [PubMed: 23537808]

17. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012; 98(9):691–698. [PubMed: 22397946]

18. Roden DM, Pulley JM, Basford MA, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther. 2008; 84(3):362–369. [PubMed: 18500243]

19. Schildcrout JS, Denny JC, Bowton E, et al. Optimizing drug outcomes through pharmacogenetics: a case for preemptive genotyping. Clin Pharmacol Ther. 2012; 92(2):235–242. [PubMed: 22739144]

20. 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(4):560–572. [PubMed: 20362271]

21. 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(e1):e147–154. [PubMed: 23531748]

22. Kho AN, Hayes MG, Rasmussen-Torvik L, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. J Am Med Inform Assoc. 2012; 19(2):212–218. [PubMed: 22101970]

23. Xu H, Stenner SP, Doan S, Johnson KB, Waitman LR, Denny JC. MedEx: a medication information extraction system for clinical narratives. J Am Med Inform Assoc. 2010; 17(1):19–24. [PubMed: 20064797]

24. Doan S, Bastarache L, Klimkowski S, Denny JC, Xu H. Integrating existing natural language processing tools for medication extraction from discharge summaries. J Am Med Inform Assoc. 2010; 17(5):528–531. [PubMed: 20819857]

25. Liu M, Shah A, Jiang M, et al. A study of transportability of an existing smoking status detection module across institutions. AMIA Annu Symp Proc. 2012; 2012:577–586. [PubMed: 23304330]

26. Horton NJ, Kleinman KP. Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. Am Stat. 2007; 61(1):79–90. [PubMed: 17401454]

27. Spiegelhalter DJ. Probabilistic prediction in patient management and clinical trials. Stat Med. 1986; 5(5):421–433. [PubMed: 3786996]

28. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007; 115(7):928–935. [PubMed: 17309939]

29. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982; 247(18):2543–2546. [PubMed: 7069920]

30. 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

31. Pfister R, Brägelmann J, Michels G, Wareham NJ, Luben R, Khaw KT. Performance of the CHARGE-AF risk model for incident atrial fibrillation in the EPIC Norfolk cohort. Eur J Prev Cardiol. 2015; 22:932–939. [PubMed: 25059930]

32. Cappola AR, Fried LP, Arnold AM, et al. Thyroid status, cardiovascular risk, and mortality in older adults. JAMA. 2006; 295(9):1033–1041. [PubMed: 16507804]

33. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J 3rd. Factors of risk in the development of coronary heart disease–six year follow-up experience. The Framingham Study. Ann Intern Med. 1961; 55:33–50. [PubMed: 13751193]

34. Gress TW, Nieto FJ, Shahar E, Wofford MR, Brancati FL. Hypertension and antihypertensive therapy as risk factors for type 2 diabetes mellitus. Atherosclerosis Risk in Communities Study. N Engl J Med. 2000; 342(13):905–912. [PubMed: 10738048]

35. Kannel WB, McGee DL. Diabetes and cardiovascular disease. The Framingham study. JAMA. 1979; 241(19):2035–2038. [PubMed: 430798]

36. Kannel WB, LeBauer EJ, Dawber TR, McNamara PM. Relation of body weight to development of coronary heart disease. The Framingham study. Circulation. 1967; 35(4):734–744. [PubMed: 6024013]

37. Allen Maycock CA, Muhlestein JB, Horne BD, et al. Statin therapy is associated with reduced mortality across all age groups of individuals with significant coronary disease, including very elderly patients. J Am Coll Cardiol. 2002; 40(10):1777–1785. [PubMed: 12446061]

38. Emdin CA, Rahimi K, Patel A. Lowering blood pressure in patients with diabetes–reply. JAMA. 2015; 313(21):2183–2184.

39. Law MR, Wald NJ, Thompson SG. By how much and how quickly does reduction in serum cholesterol concentration lower risk of ischaemic heart disease? BMJ. 1994; 308(6925):367–372. [PubMed: 8043072]

40. National Cholesterol Education Program Expert Panel on Detection E, Treatment of High Blood Cholesterol in A. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation. 2002; 106(25):3143–3421. [PubMed: 12485966]

41. Pedersen TR, Kjekshus J, Berg K, et al. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). 1994 Atheroscler Suppl. 2004; 5(3):81–87. [PubMed: 15531279]

42. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. N Engl J Med. 1998; 339(19):1349–1357. [PubMed: 9841303]

43. Gibbons, GH., Department of Health and Human Services. National Institutes of Health. Fiscal Year 2014 Budget Request. 2013. Available at: http://olpa.od.nih.gov/PDFs/CongretionalHearings/Final.pdf

44. Peterson JF, Bowton E, Field JR, et al. Electronic health record design and implementation for pharmacogenomics: a local perspective. Genet Med. 2013; 15(10):833–841. [PubMed: 24009000]

45. Zhang Y, Zhang P, Mu Y, et al. The role of renin-angiotensin system blockade therapy in the prevention of atrial fibrillation: a meta-analysis of randomized controlled trials. Clin Pharmacol Ther. 2010; 88(4):521–531. [PubMed: 20811347]

46. Schneider MP, Hua TA, Bohm M, Wachtell K, Kjeldsen SE, Schmieder RE. Prevention of atrial fibrillation by Renin-Angiotensin system inhibition a meta-analysis. J Am Coll Cardiol. 2010; 55(21):2299–2307. [PubMed: 20488299]
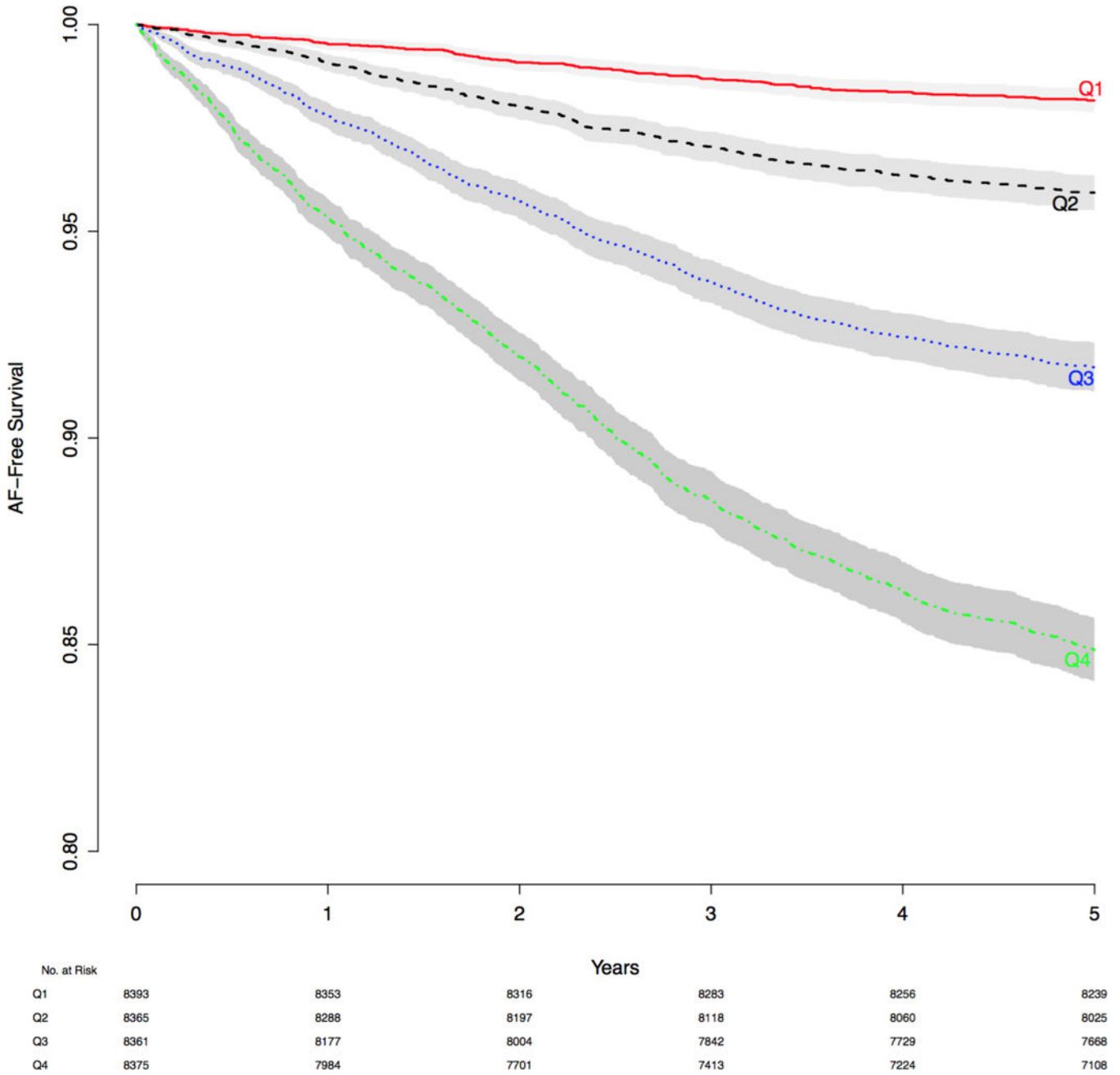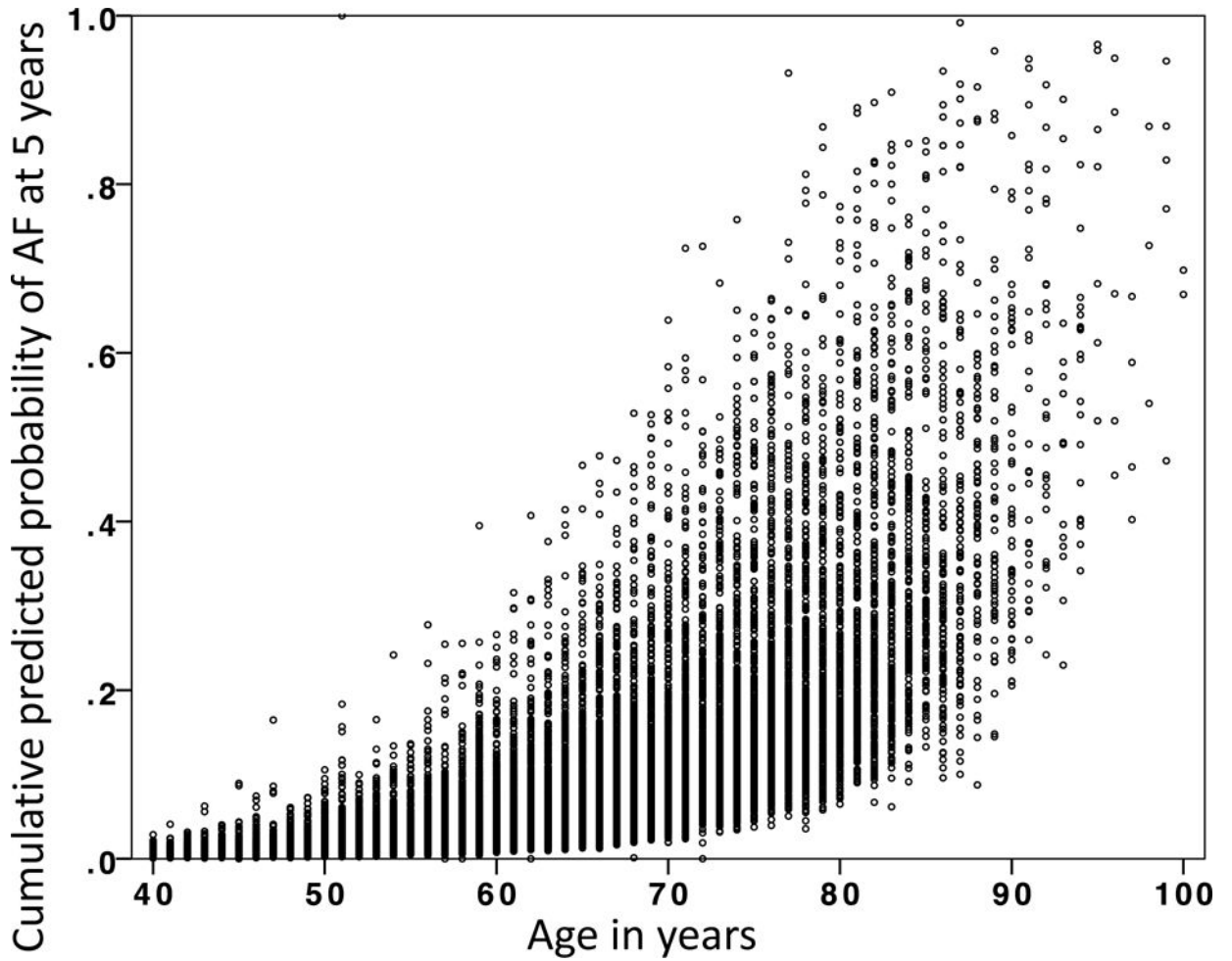
Figure 1.
Kaplan-Meier curve showing the cumulative incidence of atrial fibrillation in the Vanderbilt electronic medical record cohort. The shaded area represents 95% confidence interval. AF: atrial fibrillation.
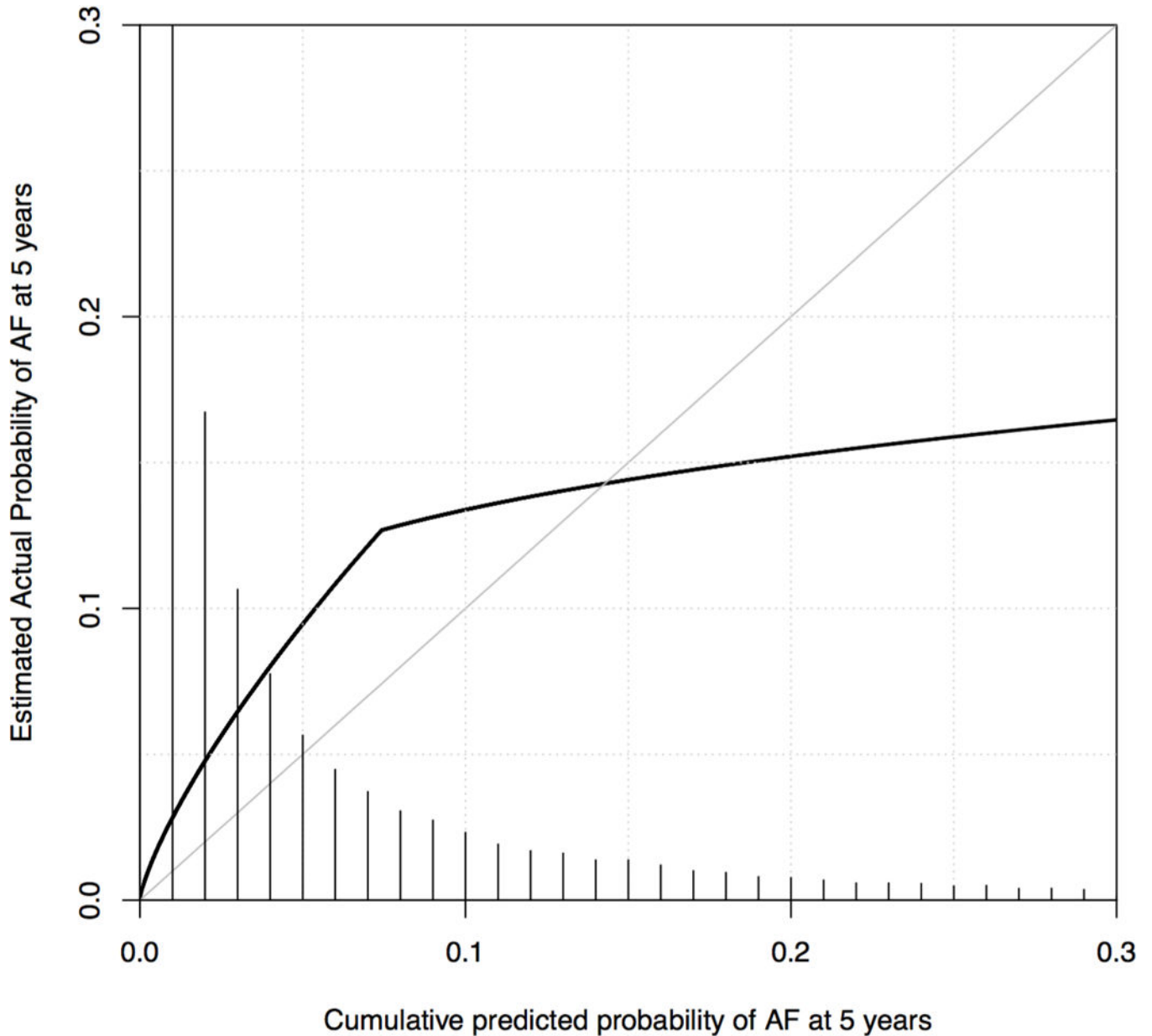
Figure 2.
Calibration curve depicting predicted and observed cumulative incidence of atrial fibrillation using the CHARGE-AF model in the Vanderbilt electronic medical record cohort. The figure is restricted to 95% of subjects with predicted cumulative incidence of atrial fibrillation from 0 to 0.3. The diagonal line represents a hypothetical ideal curve where predicted and actual atrial fibrillation probability match perfectly for all levels of predicted risk, representing perfect calibration. The curve shows under-prediction for subjects with a predicted probability of atrial fibrillation 0 to 0.15 and over-prediction for subjects with a predicted probability 0.15. A histogram showing the distribution of predicted risk is shown. AF: atrial fibrillation.

Figure 3.
Scatterplot of cumulative predicted probability of atrial fibrillation at 5 years and age. The C-index for a Cox proportional hazards model using only age as a dependent variable was 0.684 (95% confidence interval: 0.674 to 0.694). While age is the most powerful predictor of incident atrial fibrillation, this figure demonstrates that for any given age there is a wide variability in the predicted probability of developing atrial fibrillation. AF: atrial fibrillation.

Table 1

CHARGE-AF simple and full risk model coefficients β and standard errors (SE).

Predictor	Simple model coefficient β (SE)	Full model coefficient β (SE)
Age, 5 years	0.51 (0.02)	0.50 (0.02)
Race, white	0.47 (0.09)	0.49 (0.09)
Height, 10cm	0.25 (0.04)	0.24 (0.04)
Weight, 15kg	0.12 (0.03)	0.12 (0.03)
Systolic BP, 20mmHg	0.20 (0.03)	0.19 (0.03)
Diastolic BP, 10mmHg	–0.10 (0.03)	–0.10 (0.03)
Smoking, current	0.36 (0.09)	0.37 (0.09)
Treatment for HTN	0.35 (0.06)	0.34 (0.06)
Diabetes	0.24 (0.07)	0.24 (0.07)
Heart failure	0.70 (0.11)	0.68 (0.11)
Myocardial infarction	0.50 (0.09)	0.47 (0.09)
LVH by ECG criteria	——	0.40 (0.13)
PR interval <120msec	——	0.65 (0.20)
PR interval >200msec	——	0.12 (0.08)

BP: blood pressure; HTN: hypertension; LVH: left ventricular hypertrophy. The Cox proportional hazards models derived in CHARGE-AF were specified as $1 - 0.9718412736^{\exp(\Sigma bX\ 12.5815600)}$ for the simple model and $1 - 0.9719033184^{\exp(\Sigma \beta X - 12.4411305)}$ for the full model.

Table 2

Baseline patient characteristics.

Variable	Number of non-missing values	Frequency or median [25th, 75th percentile]
Age, years	33,494	57 [49,67]
Sex, women	33,494	57.0%
Race	33,494	
African American		14.3%
White		85.7%
Height, cm	29,372	170 [163,178]
Weight, kg	29,418	83 [69,98]
Body mass index, kg/m^2	29,399	28 [25,33]
Systolic BP, mmHg	31,636	130 [118,142]
Diastolic BP, mmHg	31,636	76 [69,82]
Treatment for Hypertension	33,494	85.6%
Current smoking	33,494	33.0%
Diabetes	33,494	17.3%
Heart failure	33,494	11.7%
Myocardial infarction	33,494	7.0%
LVH by ECG	33,494	5.8%
PR interval, msec	10,620	162 [148,180]

AF: atrial fibrillation; BP: blood pressure; LVH: left ventricular hypertrophy.