Published in final edited form as:

Heart Rhythm. 2019 June ; 16(6): 820–826. doi:10.1016/j.hrthm.2018.12.005.

A New Model to Predict Ischemic Stroke in Patients with Atrial Fibrillation Using Warfarin or Direct Oral Anticoagulants

J'Neka S. Claxton, MPH¹, Richard F. MacLehose, PhD², Pamela L. Lutsey, PhD², Faye L. Norby, MPH, MS², Lin Y. Chen, MD, MS, FHRS³, Wesley T. O'Neal, MD, MPH⁴, Alanna M. Chamberlain, PhD⁵, Lindsay G. S. Bengtson, PhD⁶, and Alvaro Alonso, MD, PhD¹

¹Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA

²Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN

³Cardiovascular Division, Department of Medicine, University of Minnesota Medical School, Minneapolis, MN

⁴Division of Cardiology, Department of Medicine, School of Medicine, Emory University, Atlanta, GA

⁵Department of Health Sciences Research, Mayo Clinic, Rochester, MN

⁶Health Economics and Outcomes Research, Life Sciences, Optum, Eden Prairie, MN

Abstract

Background: Stroke risk stratification scores (e.g., CHA₂DS₂-VASc) are used to tailor therapeutic recommendations for patients with atrial fibrillation (AF) in different risk groups.

Objective: To develop a tool to estimate stroke risk in patients receiving oral anticoagulants (OACs) and to identify patients who remain at a high risk of stroke despite anticoagulation therapy.

Methods: Patients with non-valvular AF initiating OACs were identified in the MarketScan data from 2007-2015. Using bootstrapping methods and backward selection of 44 candidate variables, we developed a model that selected variables predicting stroke. The final model was validated in patients with non-valvular AF in the Optum database in the period 2009-2015. In both databases, the discrimination of existing stroke scores were individually evaluated and compared with our new model termed AntiCoagulaTion-specific Stroke (ACTS) score.

Results: Among 135,523 patients with AF initiating OACs in the MarketScan dataset, 2,028 experienced an ischemic stroke after anticoagulant initiation. The stepwise model identified 11

Corresponding Author: J'Neka S. Claxton, MPH, Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton RD., Atlanta, GA 30322, Phone: +1.404.727.3288, j'neka.claxton@emory.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest:

Dr. Lindsay Bengtson is an employee of Optum. All other authors have no potential conflicts of interest.

variables (including type of OAC) associated with ischemic stroke. The discrimination (c-statistic) of the model was adequate [0.68, 95% confidence interval (CI) 0.66-0.70], showing excellent calibration ($\chi^2=6.1$ p=0.73). ACTS was then applied to 84,549 AF patients in the Optum data set (1,408 stroke events), showing similar discrimination (c-statistic 0.67, 95% CI 0.65-0.69). However, previously developed predictive models had similar discriminative ability (CHA₂DS₂-VAsC 0.67, 95% CI 0.65-0.68).

Conclusion: A novel model to identify AF patients at higher risk of ischemic stroke, using extensive administrative healthcare data including type of anticoagulant, did not perform better than established simpler models.

Keywords

Atrial fibrillation; ischemic stroke; epidemiology; risk model; anticoagulation

Introduction

The risk of stroke in atrial fibrillation (AF) differs across patients and depends on the presence of various risk factors.¹ Existing risk classification schemes,²⁻⁴ developed in patients not receiving anticoagulation therapy, classify individuals as being at low, intermediate or high stroke risk. Despite their utility in identifying individuals in the AF population who are above the risk threshold and are most likely to benefit from oral anticoagulation, the existing risk scores do not estimate the actual stroke risk when receiving anticoagulation, needed to inform risk-benefit decisions by patients and providers. Also, the existing scores do not identify patients who remain at an increased stroke risk despite anticoagulation therapy. Identification of these individuals can assist clinicians in treatment decisions and overall AF management.

Current treatment guidelines recommend the use of vitamin K antagonists (VKA) (usually warfarin in the United States) and direct oral anticoagulants (DOACs) (i.e., dabigatran, rivaroxaban, and apixaban) for patients with a CHA₂DS₂-VAsC score of 2 or greater.⁵ Beyond the decision to initiate an oral anticoagulant (OAC), there is little guidance on the decision-making process between the available anticoagulation therapies. A model created in a population of AF patients who initiated an oral anticoagulant has the potential to improve stroke prediction in two ways: 1) Refining stroke risk prediction in those considered to be at the highest risk of stroke and 2) providing insight into an individual's risk of stroke by type of oral anticoagulant. Therefore, the objective of this analysis is to develop a risk stratification model to identify patients who are still at a high risk of stroke despite optimal OAC therapy and to provide a tool to guide a clinician's evaluation of stroke risk by oral anticoagulant, given the patient's characteristics. Using data from a large US healthcare utilization database, we developed a model for the prediction of stroke in patients who initiated OAC therapy (VKA or DOACs). We externally validated the novel model in a sample of patients in a separate large US healthcare utilization database. Finally, we assessed three existing classification schemes — CHADS₂,² CHA₂DS₂-VAsC,³ ATRIA⁴— to determine their ability to predict stroke in patients on OACs and compared their performance to our new model.

Methods

Data Source and Study Population

We used health care claims data from two large US databases: Truven Health MarketScan® Commercial Claims and Encounters Database and the Medicare Supplemental and Coordination of Benefits Database (Truven Health Analytics Inc., Ann Arbor, MI, USA) from January 1, 2007 through September 30, 2015 and the de-identified Clinformatics® Data Mart, a product of Optum (Eden Prairie, MN), from January 1, 2009 to September 30, 2015. Data from MarketScan was used to derive a predictive model of ischemic stroke among patients with AF using oral anticoagulants. The model was validated using data from Optum Clinformatics®. The MarketScan databases contain enrollment data and health insurance claims for inpatient, outpatient, and pharmacy services. These data are collected from large employers and health plans across the US that provide private healthcare coverage for employees, their spouses and dependents and for individuals and their dependents with Medicare supplemental plans. The Optum database includes Commercial Claims Data and Managed Medicare data. The Commercial Claim Databases includes enrollment data and administrative health claims collected from members of a large national health plan affiliated with Optum. The Medicare Database includes claims from individuals enrolled in a Managed Medicare health plan. In both MarketScan and Optum Clinformatics®, patient enrollment data are linked with medical and outpatient prescription drug claims and encounter data to provide individual-specific clinical use, cost, and outcomes information across inpatient and outpatient services and outpatient pharmacy services.

In each database, the analysis was restricted to individuals with at least six months of continuous enrollment prior to the first prescription of an oral anticoagulant and a history of non-valvular AF, defined by the presence of at least one inpatient or two outpatient claims at least 7 days but less than 1 year apart with the International Classification of Disease Ninth Revision Clinical Modification (ICD-9-CM) code 427.31 or 427.32 in any position, without any inpatient ICD-9-CM diagnosis codes for mitral stenosis or mitral valve disorder. Furthermore, patients were required to have a first prescription for an oral anticoagulant (warfarin or DOAC) after their initial AF diagnosis. The Institutional Review Board at Emory University reviewed and approved this study. The review board waived the need for patient consent.

Oral anticoagulant use

Each outpatient pharmaceutical claim includes information on the National Drug Code (NDC), the prescription fill date, and the number of days supplied. We identified all prescriptions for oral anticoagulants (warfarin, dabigatran, rivaroxaban, and apixaban) from January 1, 2007 to September 30, 2015 (MarketScan) and January 1, 2009 to September 30, 2015 (Optum Clinformatics®). Patients were categorized according to the first anticoagulant prescription after their AF diagnosis. The newest FDA-approved (January 2015) oral anticoagulant, edoxaban, was not included due to the small number of initiators. With the assumption that patients were prescribed the correct dose given their characteristics, all DOAC prescriptions were included independently of the dosage strength. Currently, there is

no information available on the validity of claims for DOAC prescription, however the validity of warfarin claims in administrative databases is excellent (a positive predictive value (PPV) of 99%).⁶

Definition of ischemic stroke

The primary outcome variable was a hospitalization for ischemic stroke, defined based on the presence of ICD-9-CM codes 434.xx (Occlusion of cerebral arteries) and 436.xx (Acute but ill-defined cerebrovascular disease) as the primary discharge diagnosis in any inpatient claim after OAC initiation. Validation studies have shown the PPV of this definition to be >80%.⁷

Candidate predictors

Variables considered for the model were derived from literature and previously developed risk stratifications schemes (i.e., CHADS₂,² CHA₂DS₂-VAsC,³ ATRIA⁴). We identified 44 candidate predictors of ischemic stroke, which were defined using ICD-9-CM codes from inpatient, outpatient, and pharmacy claims prior to or at the time of OAC initiation and were based on available validated algorithms (Supplemental Table I).^{8,9} These variables included comorbidities, procedures, pharmacy fills, and demographic characteristics. A complete list of the covariates is provided in the online supplement (supplementary methods).

Statistical analysis

Derivation of the predictive model—We developed a predictive model of stroke using the MarketScan databases. Patients were categorized according to their first anticoagulant prescription after AF diagnosis. Follow-up started at the date of first OAC and continued until an ischemic stroke hospitalization occurred, September 30, 2015, or the patient disenrolled from their health plan, whichever occurred earlier.

Time to stroke hospitalization was the primary outcome. To identify potential predictors of ischemic stroke, we generated 1000 bootstrap samples from the MarketScan cohort and ran Cox models with stepwise selection of the 44 candidate predictors, described in the previous section, for each of the 1000 samples ($p > 0.05$ for exclusion). Variables included in at least 60% of the Cox models were selected for inclusion in the final predictive model.¹⁰ We tested for interaction between the covariates selected from the bootstrap samples and oral anticoagulant use ($p < 0.005$ after Bonferroni correction). We calculated the C-statistic to estimate the discrimination of the model and an adapted Hosmer-Lemeshow test to assess calibration using a time horizon of 12 months.^{11, 12} Due to the large sample size, calibration was also assessed qualitatively by plotting observed risk within deciles of predicted risks.¹³ Using the C-statistic, we compared the discriminative ability of our new model (termed AntiCoagulaTion-specific Stroke (ACTS) score) compared with three other stroke risk prediction scores, CHADS₂, CHA₂DS₂-VAsC, and ATRIA. Each risk score was calculated according to the definitions in their respective derivation cohorts using variables that were defined prior to or at the time of OAC initiation in the MarketScan databases (summarized in Supplemental Table II).

We performed a sensitivity analysis in which we reassessed the performance of the four models in patients followed until an ischemic stroke hospitalization occurred, September 30, 2015, a patient disenrolled from their health plan, or a switch or discontinuation of their first anticoagulant fill, whichever occurred earlier. Patients were considered to have switched OACs if they filled a prescription for an OAC other than the index OAC fill during the follow-up period. Discontinuation was defined as having no claims for the first anticoagulant fill within 30 days of the final day's supply of the last filled prescription for the index OAC.

Finally, we calculated the sensitivity, specificity, PPV, and negative predictive value (NPV) of each predictive model. These parameters were calculated using two cutoff values (1% and 2%) of predicted 1-year stroke risk, with a true positive defined as an incident stroke case with a predicted risk above the cutoff, and a true negative as a non-stroke case with a predicted risk below the cutoff value.

External model validation—The risk prediction model developed in the derivation cohort (MarketScan) was applied in the Optum Clinformatics® cohort to estimate the 1-year risk of an ischemic stroke. In the validation cohort, we replicated all model performance assessments and evaluated the discrimination of our model compared to existing models (i.e., CHADS₂, CHA₂DS₂-VASc, ATRIA). Each score was calculated in the Optum Clinformatics® database as described above. All statistical analyses were performed with SAS 9.4 (SAS Institute, Cary, NC).

Results

Derivation cohort

In the MarketScan databases, for the period of January 1, 2007 to September 30, 2015, we identified 135,523 patients with a diagnosis of non-valvular AF initiating an OAC after their AF diagnosis and with at least 180 days of enrollment data prior to OAC initiation. Table 1 presents selected demographic and clinical characteristics of these patients at the time of OAC initiation. Patients were more likely to be warfarin initiators (66%), followed by rivaroxaban (14%), dabigatran (12%), and apixaban (8%) initiators. There were small differences in demographic characteristics and disease prevalence across type of OAC.

In the derivation cohort, over a mean (median) follow-up of 22(16) months, we identified 2,028 hospitalizations for an ischemic stroke. Of these events, there were 1,632 strokes among AF patients who had initiated warfarin, 195 among dabigatran initiators, 157 among rivaroxaban initiators and 44 among apixaban initiators. The corresponding strokes rates were 8.34, 6.81, 7.54, and 5.73 strokes per 1000 patient-years (Supplemental Table III). After creating 1000 bootstrap samples in the MarketScan cohort and performing stepwise regression, the following 14 variables were included in at least 60% of the models: age, female sex, prior history of ischemic stroke, insulin, sulfonylureas, class I and class III antiarrhythmic agents, beta blockers, calcium channel blockers, statins, other lipid lowering medications, history of intracranial bleeding, antiplatelets, and type of oral anticoagulant. In order to create a more parsimonious model, we combined medication variables used to treat the same disease or condition into one category. We combined variables in to the following categories: anti-diabetic medications (insulin and sulfonylureas), lipid lowering medications

(statins and other lipid lowering medications), and antiarrhythmic medications (type I and type III antiarrhythmic). The final model contained 11 variables. There were no significant interactions between the variables selected and type of OAC. The associations between the variables selected for the final model and the risk of ischemic stroke are in Table 2. The model had fair discrimination (C-statistic=0.68, 95%CI 0.66-0.70) and excellent calibration (Table 2 and Figure 1A). Table 3 shows the discrimination of the existing models (CHADS₂, CHA₂DS₂-VASc, ATRIA) and the ACTS model in the derivation cohort. The C-statistic is similar across models.

In our sensitivity analysis censoring at the time of switching or discontinuation of oral anticoagulation, during a mean (median) follow-up of 9(5) months, we identified 715 hospitalizations for an ischemic stroke (7.2 events per 1000, 95%CI 6.7-7.7). The models performed similarly to the original analysis, with a C-statistic of 0.69 (95%CI 0.67-0.72) and adequate calibration (Supplemental Table IV and Supplemental Figure IA).

Validation cohort

We validated the ACTS score in the Optum Clinformatics® cohort. After applying the same inclusion criteria as MarketScan, the validation sample included 84,549 patients with non-valvular AF. The majority of patients who initiated an oral anticoagulant were warfarin initiators. Overall, OAC initiators were more likely to be men, elderly (average age of 70 and above), and have moderate to high risk of stroke according to the existing stroke risk-stratification schemes (Table 1). Warfarin and apixaban initiators had higher prevalence of disease compared to dabigatran and rivaroxaban initiators. We identified 1,408 strokes (warfarin: 1,047; dabigatran: 168; rivaroxaban: 145; apixaban: 48) with a mean(median) follow-up time of 21(16) months. Ischemic stroke rates are in Supplemental Table III. The discrimination of the new model is similar to the C-statistic observed in the derivation cohort (C-statistic=0.67, 95%CI 0.65-0.69), as well as similar to the existing stroke risk scores (Table 3). We observed a similar pattern when comparing the models across sex, ablation therapy, and race (white vs. nonwhite) strata. However, the predictive ability of the model was better for patients <75 years of age compared to older individuals (Supplemental Table V). The model calibration was poor when using the adapted Hosmer-Lemeshow test (Table 2) but acceptable in a qualitative comparison of predicted vs. observed deciles of risk (Figure 1B).

When censoring individuals who switched or discontinued their first prescription OAC fill, we identified 484 hospitalizations for an ischemic stroke, during a mean(median) follow-up of 8(5) months (8.2 events per 1000, 95%CI 7.5-8.9). The models performed similarly to the original analysis (Supplemental Table IV and Supplemental Figure IB).

Supplemental Table VI shows the screening test performance of the ACTS model and the existing models at different cutoff values of predicted stroke risk. The sensitivity and specificity varied widely across models at each risk threshold. At a 1% predicted value threshold, the sensitivity of the models ranged from 54.9%-100%, decreasing to 15.6%-98.3% at the 2% threshold. The specificity was between 0%-68.6% at the lowest threshold and increasing to 4.6%-94% at the highest threshold. High sensitivity (100%) and low specificity (0%), as seen with the CHADS₂ model, suggests that the model perfectly

predicts stroke in those who actually develop stroke, but does not accurately identify those who will not develop disease. In other words, it classifies everyone into the high-predicted risk category, including those who will not develop stroke. Across all models, among those who had a high-predicted risk, less than 2% had a stroke event (PPV). Conversely, with the exception of the CHADS₂ model at the 1% risk threshold, those who had a low predicted risk, ~99% did not have an ischemic stroke (NPV). Finally, we evaluated the discriminative ability of the CHA₂DS₂-VASc score, the recommended score in the current treatment guidelines, followed by our new score among those considered high-risk. This approach did not improve the specificity and PPV parameters, suggesting that applying our model sequentially after the CHA₂DS₂-VASc score does not help with refining the subset of individuals who are most likely develop a stroke.

Discussion

In this analysis, we used a large administrative claims database to identify variables associated with ischemic stroke and develop a model that estimates the risk of ischemic stroke in AF patients receiving anticoagulants (i.e., warfarin or DOACs). The resulting model showed modest discrimination and adequate calibration in both the derivation and validation datasets. When compared to existing risk stratification schemes, the ACTS model did not provide improved predictive accuracy of stroke in this cohort of patients with AF using oral anticoagulants. There were also no meaningful changes in each model's performance when accounting for changes in anticoagulation therapy and medication adherence after OAC initiation. The new model performed similarly in men and women and across ablation therapy status and racial/ethnic groups. However, we observed better predictive ability among individuals younger than 75 years.

Anticoagulation therapy is effective in reducing the risk of ischemic stroke associated with AF. However, anticoagulation itself carries an increased risk of bleeding complications.¹⁴ The development of risk stratification models are important to the identification of patients with different underlying risks of stroke and bleeding. These types of models are both relevant to physicians and patients; however estimating the risk of stroke is a critical first step when assessing the risk and benefits of anticoagulation. Risk stratification models that accurately and reliably stratify stroke risk can improve AF management, decreasing underuse of OAC and thereby preventable thromboembolic events.

Several risk stratification scores exist.²⁻⁴ Current guidelines recommend the use of the CHA₂DS₂-VASc score for risk stratification in patients with AF,⁵ however the predictive ability of the model is limited (C-statistic<0.7).³ The CHA₂DS₂-VASc score and other scores (i.e., CHADS₂ and ATRIA), developed in patients not using anticoagulants, stratify patients as low, moderate, or high risk of stroke. While this is useful to identify patients that would most benefit from anticoagulation therapy, these scores do not identify patients who remain at a high level of risk despite oral anticoagulation. In addition, they do not provide information on the expected risk of stroke when receiving anticoagulation. Identifying individuals who have the highest risk of stroke and are most likely to benefit from anticoagulation therapy is crucial to AF management. Literature on the evaluation of risk stratification schemes in patients using an oral anticoagulant have demonstrated that some

patients remain at a high risk despite anticoagulation therapy and still experience a stroke.^{1, 15} Our findings demonstrate that our model, and other existing models, can be used to further stratify patients in this high risk category. How to manage these patients at high risk of ischemic stroke despite oral anticoagulation is an open question. Future research should evaluate whether other interventions, such as improved management of stroke risk factors (e.g. blood pressure control) or use of alternative preventive strategies (e.g. left atrial appendage occlusion), could be of particular benefit in this population.

Despite the potential advantages of the new model, such as derivation in a large patient population, considering numerous potential predictors, including type of anticoagulation, and validation in a separate sample, we did not observe a meaningful improvement in prediction beyond the predictive ability of existing risk stratification schemes. These observations have two significant implications. One, they offer support to the current recommendation to use the CHA₂DS₂-VASc to evaluate ischemic stroke risk and to inform decisions about oral anticoagulant initiation. Two, they highlight the need to identify better predictors of ischemic stroke in patients with AF, including circulating biomarkers and genetic factors. Given the modest predictive ability of all models evaluated (c-statistic < 0.7), there is an urgent need to develop improved predictive tools to inform treatment decisions. This is particularly important for individuals 75 and older, in which all the risk scores had worse discriminative ability than in younger individuals in our study population.

This analysis should be interpreted in the context of several limitations. First, our assessment of predictors of ischemic stroke are dependent on risk factors available in the database. Published literature suggests that circulating biomarkers may improve stroke prediction¹⁶; however, these variables were not available in our particular data. Second, the predictive ability of our model relies on the ability to ascertain both outcome and covariates accurately in administrative data. Validated algorithms were utilized to ascertain events of interest and it is likely that any misclassification is non-differential. Third, DOAC prescriptions were included independently of dosage strength. However, as a proxy for dosing information, we included clinical indications for dosage reduction as candidate predictors in the model (e.g., kidney disease and age). Some of these variables were not selected for inclusion in the final model and, therefore, it may be reasonable to assume that they do not have a major impact on the risk of stroke. Fourth, we lack information on AF burden, which could be associated with stroke risk and influence anticoagulant treatment decisions. Finally, in the derivation and validation cohorts, 70% (93,113) and 48% (40,280), respectively, disenrolled from their health plan before the end of the study period. Unfortunately, due to the nature of the database, we are unable to distinguish or categorize the reasons for disenrollment (e.g., no information on mortality data).

Conclusion

In conclusion, we developed a novel model using administrative claims for the prediction of stroke in patients with AF using OACs. The ACTS model, despite use of extensive information, did not improve upon existing schemes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Funding

This work was supported by the National Heart, Lung, And Blood Institute and National Institute of Aging of the National Institutes of Health [grant numbers R01-HL122200, F32-HL134290, R21-AG058445]; American Heart Association [grant number 16EIA26410001]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the American Heart Association. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Poli D, Antonucci E, Grifoni E, Abbate R, Gensini GF, Prisco D. Stroke risk in atrial fibrillation patients on warfarin. Predictive ability of risk stratification schemes for primary and secondary prevention. *Thromb Haemost* 2009;101:367–372. [PubMed: 19190823]
2. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. *JAMA* 2001;285:2864–2870. [PubMed: 11401607]
3. 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;137:263–272. [PubMed: 19762550]
4. Singer DE, Chang Y, Borowsky LH, Fang MC, Pomernacki NK, Udaltsova N, Reynolds K, Go AS. A new risk scheme to predict ischemic stroke and other thromboembolism in atrial fibrillation: the ATRIA study stroke risk score. *J Am Heart Assoc* 2013;2:e000250. [PubMed: 23782923]
5. January CT, Wann LS, Alpert JS, et al. 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: Executive Summary. *J Am Coll Cardiol* 2014;64:2246–2280.
6. Garg RK, Glazer NL, Wiggins KL, Newton KM, Thacker EL, Smith NL, Siscovick DS, Psaty BM, Heckbert SR. Ascertainment of warfarin and aspirin use by medical record review compared with automated pharmacy data. *Pharmacoepidemiol Drug Saf* 2011;20:313–316. [PubMed: 21351314]
7. Andrade SE, Harrold LR, Tjia J, Cutrona SL, Saczynski JS, Dodd KS, Goldberg RJ, Gurwitz JH. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiol Drug Saf* 2012;21 Suppl 1:100–128. [PubMed: 22262598]
8. Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE, Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. *Pharmacoepidemiol Drug Saf* 2011;20:560–566. [PubMed: 21387461]
9. Quan HSV, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD., Beck CAFT, Ghali WA. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. *Med Care* 2005;43:1130–1139. [PubMed: 16224307]
10. Austin PC, Tu JV. Bootstrap Methods for Developing Predictive Models. *The American Statistician* 2004;58:131–137.
11. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;23:2109–2123. [PubMed: 15211606]
12. D'Agostino RB, Nam B-H. Evaluation of the Performance of Survival Analysis Models: Discrimination and Calibration Measures In: Balakrishnan N, Rao CR, eds. *Advances in survival analysis: Handbook of Statistics*, vol. 23 Amsterdam; San Diego and Oxford: Elsevier, North-Holland; 2004:1–25.

13. Kramer AAZ, Jack E. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited*. Crit Care Med 2007;35:2052–2056. [PubMed: 17568333]
14. Rowan SB, Bailey DN, Bublitz CE, Anderson RJ. Trends in Anticoagulation for Atrial Fibrillation in the U.S.: An Analysis of the National Ambulatory Medical Care Survey Database. J Am Coll Cardiol 2007;49:1561–1565. [PubMed: 17418296]
15. Baruch L, Gage BF, Horrow J, Juul-Moller S, Labovitz A, Persson M, Zabalgaitia M. Can patients at elevated risk of stroke treated with anticoagulants be further risk stratified? Stroke 2007;38:2459–2463. [PubMed: 17673721]
16. Hijazi Z, Oldgren J, Siegbahn A, Wallentin L. Application of Biomarkers for Risk Stratification in Patients with Atrial Fibrillation. Clin Chem 2017;63:152–164. [PubMed: 27811208]

Figure 1.

Calibration of final model in derivation and validation cohorts. Calibration curve relating observed and predicted ischemic stroke rates across deciles of risk in **A**. Derivation Cohort (MarketScan) **B**. Validation Cohort (Optum Clinformatics®). The diagonal dashed line indicates perfect fit.

Table 1.

Characteristics of patients with atrial fibrillation according to initial prescribed anticoagulant in the derivation (MarketScan,2007-2015) and validation (Optum Clinformatics®, 2009-2015) cohorts

	Derivation Cohort (MarketScan)				Validation Cohort (Optum Clinformatics®)			
	Warfarin	Dabigatran	Rivaroxaban	Apixaban	Warfarin	Dabigatran	Rivaroxaban	Apixaban
N	90271	15778	19048	10426	52130	9368	14520	8531
Demographics								
Age, years	70.6(12.7)	67.8(12.6)	67.9(12.7)	69.4(12.8)	73.9(10.3)	70.3(11.3)	71.5(11.2)	73.6(10.7)
Women, %	40.7	36.7	39.5	41	44.3	40.1	42.6	47.2
Existing stroke risk scores								
CHADS ₂	2.5(1.6)	2.2(1.5)	2.1(1.5)	2.3(1.5)	2.9(1.5)	2.5(1.5)	2.6(1.5)	2.8(1.5)
CHA ₂ DS ₂ -VASc	3.8(2.1)	3.3(2.1)	3.3(2.1)	3.6(2.1)	4.6(2.0)	3.9(2.0)	4.1(2.1)	4.4(2.0)
ATRIA	6.0(3.5)	5.2(3.6)	5.2(3.6)	5.6(3.6)	7.2(3.1)	6.2(3.4)	6.5(3.3)	7.1(3.2)
Existing bleeding risk scores								
HAS-BLED	2.3(1.4)	2.1(1.3)	2.1(1.3)	2.2(1.3)	2.8(1.3)	2.5(1.3)	2.7(1.3)	2.8(1.3)
Prevalent disease, %								
Heart failure	36.0	29.1	27.5	31.1	45.4	33.9	35.0	38.8
Hypertension	77.1	79.4	80.3	82.8	89.0	87.9	87.9	89.4
Diabetes	32.8	29.7	28.9	30.7	39.9	34.5	35.6	37.1
Stroke	27.9	23.8	22.5	24.1	33.2	27.7	28.8	32.6
Peripheral artery disease	16.5	14.1	14.2	15.2	25.6	18.8	23.1	24.2
Prior gastrointestinal bleed	10.4	10.0	9.3	8.9	12.4	10.8	12.2	13.5
Prior other bleed	12.6	12.1	12.0	11.5	15.9	13.9	17.0	17.3
Prior intracranial bleed	1.7	1.0	1.2	1.5	2.1	1.1	1.4	1.9

Numbers correspond to mean (SD) and percentages

CHADS₂: congestive heart failure, hypertension, age >75, diabetes, prior stroke; CHA₂DS₂-VASc: congestive heart failure, hypertension, age 75 years, diabetes mellitus, stroke/transient ischemic attack, vascular disease, age 65–75 years, and sex category; ATRIA: Age, sex category, diabetes, congestive heart failure, proteinuria, reduced kidney function or end-stage renal disease, prior stroke; HAS-BLED: hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, labile international normalized ratio, elderly (age >65 years), drugs/alcohol concomitantly; SD: standard deviation

Table 2.

Final model predictors of ischemic stroke hospitalization among patients with atrial fibrillation initiating oral anticoagulation in the derivation (MarketScan, 2007-2015) and validation (Optum Clinformatics®, 2009-2015) cohorts

VARIABLE	Derivation Cohort (MarketScan)		Validation Cohort (Optum Clinformatics®)	
	HR	(95%CI)	HR	(95%CI)
Age, per 1 year	1.04	(1.03,1.04)	1.04	(1.03,1.05)
Sex (female)	1.24	(1.13,1.35)	1.36	(1.22,1.51)
Prior history of ischemic stroke	1.90	(1.73,2.09)	1.84	(1.65,2.06)
Anti-diabetic medications	1.64	(1.47,1.83)	1.43	(1.25,1.64)
Antiarrhythmic medications	0.75	(0.67,0.84)	0.76	(0.66,0.88)
Beta blockers	1.23	(1.11,1.37)	1.07	(0.95,1.21)
Calcium channel blockers	1.11	(1.02,1.21)	1.04	(0.93,1.17)
Lipid lowering medications	0.83	(0.75,0.91)	0.90	(0.80,1.01)
History of intracranial bleeds	1.48	(1.15,1.91)	0.96	(0.68,1.37)
Antiplatelet therapy	1.17	(1.05,1.30)	1.17	(1.02,1.35)
DOAC (vs warfarin)				
Dabigatran	0.94	(0.81,1.09)	0.99	(0.84,1.16)
Rivaroxaban	1.04	(0.88,1.23)	0.85	(0.71,1.01)
Apixaban	0.72	(0.53,0.97)	0.60	(0.45,0.81)
Calibration (χ^2 , (p value))	6.1	(p=0.73)	21.9	(p=0.01)

The 1-year risk of ischemic stroke can be calculated as $1 - (0.9937)^{\text{exp}[0.03814 \cdot (\text{Age}-69.8346) + 0.64192 \cdot (\text{Ischemic stroke} - 0.26356) + 0.21136 \cdot (\text{Female sex} - 0.40083) + 0.49492 \cdot (\text{Antidiabetic medications} - 0.14609) - 0.1923 \cdot (\text{Lipid lowering medications} - 0.56891) + 0.21049 \cdot (\text{Beta blockers} - 0.70587) + 0.39245 \cdot (\text{Intracranial bleed history} - 0.01529) - 0.2841 \cdot (\text{Antiarrhythmic use} - 0.24112) + 0.153 \cdot (\text{Antiplatelet use} - 0.1578) + 0.10354 \cdot (\text{Calcium channel blockers} - 0.39906) - 0.0633 \cdot (\text{Dabigatran}) + 0.03916 \cdot (\text{Rivaroxaban}) - 0.3289 \cdot (\text{Apixaban})]}$

Table 3.

Model discrimination [c-statistic (95% confidence interval)] by derivation and validation cohorts.

Score	Derivation cohort(MarketScan)	Validation cohort (Optum Clinformatics®)
ACTS	0.68(0.66,0.70)	0.67(0.65,0.69)
CHADS ₂	0.66(0.65,0.68)	0.66(0.63,0.68)
CHA ₂ DS ₂ -VASC	0.67(0.65,0.68)	0.66(0.64,0.68)
ATRIA	0.68(0.66,0.70)	0.67(0.65,0.69)

ACTS: AntiCoagulation-specific Stroke score; CHADS₂: congestive heart failure, hypertension, age >75, diabetes, prior stroke; CHA₂DS₂-VASC: congestive heart failure, hypertension, age 75 years, diabetes mellitus, stroke/transient ischemic attack, vascular disease, age 65–75 years, and sex category; ATRIA: Age, sex category, diabetes, congestive heart failure, proteinuria, reduced kidney function or end-stage renal disease, prior stroke.