Geographic Disparities in the Incidence of Stroke among Patients with Atrial Fibrillation in the United States

J'Neka S. Claxton, MPH,* Pamela L. Lutsey PhD,† Richard F. MacLehose, PhD,†
Lin Y. Chen, MD, MS,‡ Tené T. Lewis, PhD,* and Alvaro Alonso, MD, PhD*

Aim: To determine whether regional variation in stroke incidence exists among individuals with AF. *Methods:* Using healthcare utilization claims from 2 large US databases, MarketScan (2007-2014) and Optum Clinformatics (2009-2015), and the 2010 US population as the standard, we estimated age-, sex-, race- (only in Optum) standardized stroke incidence rates by the 9 US census divisions. We also used Poisson regression to examine incidence rate ratios (IRR) of stroke and the probability of anticoagulation prescription fills across divisions. *Results:* Both databases combined included 970,683 patients with AF who experienced 15,543 strokes, with a mean follow-up of 23 months. In MarketScan, the age- and sex-standardized stroke incidence rate was highest in the Middle Atlantic and East South Central divisions at 3.8/1000 person-years (PY) and lowest in the West North Central at 3.2/1000 PY. The IRR of stroke and the probability of anticoagulation fills were similar across divisions. In Optum Clinformatics, the age-, sex-, and race-standardized stroke incidence rate was highest in the East North Central division at 5.0/1000 PY and lowest in the New England division at 3.3/1000 PY. IRR of stroke and the probability of anticoagulation fills differed across divisions when compared to New England. *Conclusions:* These findings suggest regional differences in stroke incidence among AF patients follow a pattern that differs from the hypothesized trend found in the general population and that other factors may be responsible for this new pattern. Cross-database differences provide a cautionary tale for the identification of regional variation using health claims data.
Key Words: Atrial fibrillation—stroke—stroke incidence—epidemiology—health services research—geographic disparities

Introduction

In the United States, stroke is the fifth leading cause of death—accounting for 1 in 20 deaths—[1,2] and the leading cause of serious, long-term disability.[3−5] It has been characterized as one of the most expensive chronic diseases to treat, with an annual estimated economic burden of $17.9 billion from direct medical costs.[6] The prevalence of stroke increases with age, with two-thirds of stroke hospitalizations occurring in those aged 65 years and older.[4] With the aging of the US population, its prevalence is expected to dramatically increase over time.[4]

Atrial fibrillation (AF) is the most common arrhythmia in clinical practice, affecting between 2.7 and 6.1 million people in the United States.[7] Prevalence of AF also

increases with age,[8] affecting approximately 9% of people aged 65 years or older compared to 2% of people under age 65.[7] AF increases the risk of stroke[9] and is independently associated with a 5-fold increase in stroke incidence.[9,10] Stroke risk attributable to AF increases with age from 1.5% at 50-59 years of age to 23.5% at 80-89 years of age.[6] In patients with stroke, AF is associated with increased severity, functional impairment, and mortality.[11−13] Adjusted mean costs of stroke hospitalizations for AF patients compared to non-AF patients were significantly higher for ischemic stroke (∼$5000),[4,11] hemorrhagic stroke ($7824), and transient ischemic stroke ($1890), respectively.[11] Nationally, stroke constitutes a sizable health and economic burden for the AF population.

In the general US population, the burden of stroke has substantial geographical variation that has persisted for decades.[14,15] Since at least the 1940s, higher stroke mortality rates have been found within the Southeastern region of the United States when compared to the rest of the country.[14,16] During the 1960s, the term "stroke belt" was coined to describe this region of concentrated high stroke mortality rates.[14,17] The National Heart, Lung, and Blood Institute has defined the stroke belt as states with 1980 age-adjusted stroke mortality rates more than 10% higher than the US average.[16,18] These states are Alabama, Arkansas, Georgia, Indiana, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Virginia.[18] Analogous to stroke mortality, regional differences have been observed for stroke incidence[19-21] and stroke case-fatality.[22] However, it is currently unknown whether geographic variation in stroke incidence exists in individuals with AF and whether it can be explained by oral anticoagulant (OAC) patterns or patient characteristics. Identifying which states carry the heaviest burden of incident stroke among AF patients could contribute to understanding overall geographic disparities in stroke and may lead to the evaluation of existing strategies of AF management in regions at elevated risk. In this study, we used healthcare utilization data from 2 large US databases, Optum Clinformatics and MarketScan, to determine stroke incidence rates among patients diagnosed with AF by US census division.

Methods

Study Population

This study uses healthcare utilization data from 2 large US databases, MarketScan and Optum Clinformatics.

We included claims data from the Truven Health MarketScan Commercial Claims and Encounters Database and the Medicare Supplemental and Coordination of Benefits Database Truven Health Analytics Inc., Ann Arbor, MI) for the period January 1, 2007 to December 31, 2014. The MarketScan Commercial Database includes health insurance claims and enrollment data collected from large US employers and health plans that provide private healthcare coverage for employees, their spouses, and dependents. The Medicare Supplemental Database includes claims from individuals and their dependents with employer-sponsored Medicare supplemental coverage.

We also used deidentified claims data from Optum Clinformatics Data Mart, a product of OptumInsight, Inc. (Eden Prairie, MN), for the period January 1, 2009 to September 30, 2015. The database includes enrollment data and administrative health claims data for commercial and managed Medicare enrollees from a large national health insurer with enrollees throughout the United States. 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, expenditure, and outcomes across inpatient and outpatient services and outpatient pharmacy services.

In each database, the analysis was restricted to individuals with available medical and outpatient pharmaceutical data, a history of nonvalvular AF, and at least 6 months of continuous enrollment prior to their AF diagnosis. History of AF was defined as at least one inpatient or 2 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. Individuals with any inpatient ICD-9-CM codes for mitral stenosis or mitral valve disorders were excluded. The Institutional Review Board at Emory University reviewed and approved this study.

Oral Anticoagulant Use

Each outpatient pharmaceutical claim includes information on the National Drug Code, the prescription fill date, and the number of days supplied. We considered a patient to have received an OAC (warfarin, dabigatran, rivaroxaban, and apixaban) if they had any OAC prescription in outpatient claims within the following window: 3 months prior to 6 months after AF diagnosis. We chose this time window to accommodate uncertainty about the date of AF diagnosis. Patients were categorized according to their first anticoagulant prescription during this time period. All direct oral anticoagulants prescriptions—dabigatran, rivaroxaban, and apixaban—were included regardless of the dosage strength. The validity of warfarin claims in administrative databases is excellent (positive predictive value [PPV] of 99%),[23] which is likely to be similar for direct oral anticoagulants.

Definition of Ischemic Stroke

The primary outcome in this study was hospitalization for ischemic stroke (allowing for multiple events per patient). To identify ischemic stroke, we used an established algorithm (PPV-98%) by Williams et al.[24] Ischemic

stroke was defined by 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 following their AF diagnosis.

Covariates

Demographic characteristics and comorbidities were ascertained. Race/ethnicity was additionally available in the Optum Clinformatics database and thus was considered as a covariate for those analyses. Race/ethnicity was collected directly from public records (e.g., driver's license records) for approximately 30% of individuals, while the remaining data were imputed using commercial software (E-Tech by Ethnic Technologies). The algorithm was developed with first and last names and US Census data zip codes (zip + 4). This method was validated and has 97% specificity, 48% sensitivity, and 71% PPV for predicting the race of black individuals.[25] Comorbidities were defined using ICD-9-CM codes from inpatient and outpatient claims prior to or at the time of AF diagnosis, using all information available since time of database enrollment and were based on validated algorithms.[26] They included heart failure, hypertension, diabetes mellitus, ischemic stroke, myocardial infarction, and peripheral artery disease. A patient was considered to have a comorbidity if they had an inpatient or outpatient claim with a relevant ICD-9-CM code in any position; a list of ICD-9-codes used to define each comorbidity is included in Supplementary Table S1. These comorbidities were used to compute the CHA_2DS_2-VASc score.[27] The CHA_2DS_2-VASc score is defined by the following predictors: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke/transient ischemic attack, vascular disease, age 65-75 years, and sex category.

Geographic Location

In each database, geographic location was ascertained from patient enrollment data and based on the primary beneficiary's residence. In the MarketScan database, geographic information is provided at the level of US state, division, and region, however in Optum Clinformatics geographic information is only provided at the US census level division. For consistency across datasets, we analyzed MarketScan data according to the 9 US Census Bureau-designated divisions, defined in Supplementary Table S2.

Statistical Analysis

In this analysis, we included patients with a history of nonvalvular AF. Follow-up was defined from the date of AF diagnosis and continued until December 31, 2014 (MarketScan) or September 30, 2015 (Optum Clinformatics), or patient disenrollment from a health insurance company that was included in MarketScan or Optum Clinformatics, whichever occurred earlier.

We computed descriptive statistics across the 9 US Census Bureau-designated divisions. The primary analysis estimated incidence rates per 1000 person-years (PY) for ischemic stroke across the 9 divisions. We calculated both crude and standardized rates. All rates were standardized to the relevant 2010 US census population distribution.[28] In the MarketScan database, ischemic stroke incidence rates were age- and sex-standardized within each census divisions. To examine sex differences, we also estimated sex-specific incidence rates that were age-standardized within each census divisions. In the Optum Clinformatics database, ischemic stroke incidence rates were age-, sex-, and race-standardized across divisions. Sex- and race-specific incidence rates were compared across divisions, standardizing for age. We examined the incidence rate ratios of stroke using Poisson regression. All models were initially adjusted for demographic factors (age, sex, and race) [Model 1], with the subsequent model adjusted for CHA_2DS_2-VASc score [Model 2]. In the final model [Model 3], we further adjusted for OAC use. Poisson regression models with robust variance estimates were used to compute relative risk and 95% confidence intervals (CI) of anticoagulation prescription fills across the US census divisions. Models were adjusted for age, sex, and CHA_2DS_2-VASc score; in Optum Clinformatics, we also adjusted for race. The New England census division was used as the reference.

Results

MarketScan

Cohort Characteristics

The MarketScan databases included 605,731 patients with a diagnosis of nonvalvular AF in the period of January 1, 2007 to December 31, 2014. Among those patients, a total of 595,631 had available information on geographic location. Table 1 shows demographic and clinical characteristics of the MarketScan cohort at the time of AF diagnosis by the 9 US census divisions. The average age of AF patients across divisions was lowest in the West South Central division at 66.3 years and highest in the East North Central division at 72.1 years. Generally, across divisions patients were less likely to be women, had a similar prevalence for each comorbidity, and had low rates of OAC initiation around the time of AF diagnosis (3 months prior to 6 months after diagnosis).

Stroke Incidence Rates

Over a total follow-up time of 1,160,872 PY (mean follow-up 23.4 months), the total ischemic stroke events for the 9 divisions combined were 9026: 93.6% had 1 stroke, 5.7% had 2, and .7% had 3 or more. The resulting crude

Table 1. *Characteristics of patients with atrial fibrillation by US census bureau divisions in MarketScan, 2007-2014 and Optum Clinformatics 2009-2015 cohorts*

	New England	Middle Atlantic	East North Central	West North Central	South Atlantic	East South Central	West South Central	Mountain	Pacific
MarketScan									
N	25,892	89,086	150,251	28,863	103,349	33,069	57,360	27,824	79,937
Age at baseline, yr	68.1 (14.8)	71.1 (14.2)	72.1 (13.5)	66.7 (14.5)	68.7 (14.3)	67.6 (13.9)	66.3 (14.6)	69.4 (14.3)	71.1 (14.1)
Women	39	43	43.5	39.7	41.3	41.6	40.1	40.2	41
Anticoagulation use*	29	24.8	26.9	27.9	26.1	26.7	25.9	26.1	25.4
CHA$_2$DS$_2$-VASc	2.9 (2.0)	3.6 (2.1)	3.7 (2.1)	2.8 (2.0)	3.2 (2.0)	3.1 (2.0)	3.0 (2.0)	3.1 (2.0)	3.3 (2.0)
Chronic disease									
Chronic heart failure	22.6	27.9	30.5	22.7	23.9	25.6	25.9	24.3	26.0
Hypertension	62.8	71.7	72.5	61.7	69.6	69.6	69.4	65.7	66.5
Diabetes mellitus	23.7	29.9	31.6	25.4	28.2	28.8	26.5	24.2	24.8
Peripheral artery disease	11.5	17.5	15.7	9.9	11.9	10.5	11.3	11.1	11.6
Ischemic stroke history	16.1	23.8	23.6	16.5	20.4	19.5	19.5	18.8	19.2
Myocardial infarction	9.1	10.4	11.2	8.4	8.2	8.5	8.6	9.7	9.6
Optum Clinformatics									
N	19,469	27,905	53,663	41,948	76,731	14,234	40,988	38,682	61,432
Age at baseline, yr	74.8 (10.8)	73.5 (12.5)	72.1 (12.3)	72.9 (11.7)	70.0 (12.4)	70.8 (12.0)	70.5 (12.7)	74.0 (11.2)	77.0 (9.8)
Women	47	45.7	43.9	44.1	42.8	45	43.3	46.6	49.8
Race									
Asian	1.4	5.1	.9	.9	1.0	.9	1.4	1.2	5.4
Black	3.9	7.8	9.3	4.6	16.6	19.9	10.0	1.9	2.5
Hispanic	7.3	8.9	2.1	1.3	4.7	.8	14.6	9.4	12.9
White	87.3	78.1	87.7	93.3	77.7	78.4	74.0	87.5	79.3
Anticoagulation use*	26.3	23.0	27.0	17.2	24.6	23.8	25.5	26.7	26.7
CHA$_2$DS$_2$-VASc	4.4 (2)	4.4 (2.1)	4 (2.1)	3.9 (2)	3.8 (2.1)	4.1 (2.1)	4.1 (2.2)	4.1 (2)	4.5 (1.9)
Chronic disease									
Chronic heart failure	36.0	36.3	33.0	30.3	30.3	35.2	35.7	31.2	37.1
Hypertension	84.6	84.5	82.1	79.6	83.4	86.9	85.4	80.0	83.2
Diabetes mellitus	38.2	43.3	34.6	29.8	36.6	38.8	37.4	30.8	34.2
Peripheral artery disease	26.1	27.8	20.2	19.6	19.6	19.7	26.1	20.5	25.6
Ischemic stroke history	26.5	29.8	24.6	22.8	26.3	28.7	28.3	26.9	28.8
Myocardial infarction	16.7	13.3	13.0	13.3	12.6	13.6	15.1	13.4	16.3

Numbers correspond to mean (SD) and percentages; CHA$_2$DS$_2$-VASc: congestive heart failure, hypertension, age $\geq$ 75 years, diabetes mellitus, stroke/transient ischemic attack, vascular disease, age 65-75 years, and sex category.

*Anticoagulation 3 months prior to 6 months after AF diagnosis.

rate for the entire cohort was 7.8 ischemic strokes per 1000 PY. Patients with stroke represent less than 2% of the AF population within each of the 9 divisions. Crude rates by division are shown in Table 2. The 2010 US population age- and sex-standardized total incident stroke rate for all eligible patients was 3.6/1000 (95% CI: 3.4-3.8). By division, the total standardized rates were highest in the Middle Atlantic and East South Central (3.8/1000) and lowest in West North Central (3.2/1000) (Fig 1A). Rates for each division can be found in Supplementary Table S3. Age-standardized rates were higher for women compared to men across each division (Fig 1B,C). Compared to men, the rate of stroke for women was significantly higher in the East North Central (27%), West South Central (37%), and Mountain (77%) divisions (Supplementary Table S3).

Incidence Rate Ratios of Ischemic Stroke and Relative Risk of Anticoagulation

The rate of stroke was higher across divisions when compared to New England, with the exception of the South Atlantic division (Table 2). However, these differences were not statistically significant. The estimates decreased slightly, but remained not statistically significant with further adjustments for CHA_2DS_2-VASc score components and did not change with adjustments for OAC use. A total of 159,832 ($\sim$27%) patients filled an anticoagulant prescription across all census divisions. Compared to patients residing in the New England division, patients in each division were less likely to be anticoagulated (Table 3).

Optum Clinformatics

Cohort Characteristics

Among the 414,495 patients with a diagnosis of non-valvular AF enrolled in the Optum Clinformatics database for the period of January 1, 2009 to September 30, 2015, a total of 375,052 had available race and geographic information. The characteristics of the Optum cohort at the time of AF diagnosis by the 9 US census divisions are shown in Table 1. Across divisions, the average age of AF patients was equal to or greater than 70 years. Also, a majority of patients were white and male, with an average CHA_2DS_2-VASc score range from 3.8 to 4.5. Approximately one-fourth of patients initiated an OAC 3 months prior to 6 months after AF diagnosis.

Stroke Incidence Rates

The total ischemic stroke events were 6517: 93.0% had 1 stroke, 6.5% had 2, and .6% had 3 or more. With a total follow-up time of 726,448 PY (mean follow-up 23.2 months), the crude rate for the entire cohort was 9.0/1000 PY. Crude rates by division are shown in Table 2. The

2010 US population age-, sex-, and race-standardized total incidence stroke rate for all eligible patients in Optum was 4.0/1000 (95% CI: 3.7-4.3). Across divisions, the total age-, sex-, and race-standardized rates were highest in the East North Central (5.0/1000) and Middle Atlantic (4.8/1000) and lowest in New England (3.3/1000) (Fig 2A). Supplementary Table S4 shows the total standardized stroke rates for each division. In contrast to results from the MarketScan databases, where women had higher stroke rates in most divisions, we did not observe a consistent pattern in the age-standardized stroke rates by sex (Fig 2B,C and Supplementary Table S4). Age- and sex-standardized stroke rates for nonwhites were higher than that of whites in most regions (Middle Atlantic [rate ratio {RR} = 1.83], East North Central [RR = 1.74], South Atlantic [RR = 1.59], West South Central [RR = 1.76], and Mountain [RR = 2.26] divisions [Supplementary Table S4]).

Incidence Rate Ratios of Ischemic Stroke and Relative Risk of Anticoagulation

In the Optum database, the rate of stroke was also lower across divisions when compared to New England, with the exception of the East South Central division (Table 2). However, with the exception of stroke rates in the Mountain and Pacific divisions, these differences were not statistically significant across any models. A total of 94,728 ($\sim$25%) patients filled an anticoagulant prescription across all census divisions. Patients residing in the West North Central division were 35% less likely to be anticoagulated, while those residing in Mountain and Pacific division were 3 and 6% more likely to be anticoagulated, respectively, when compared to New England patients (Table 3).

Discussion

In this analysis, we used 2 large administrative claims databases to determine whether there is geographic variability in stroke rates in the AF population, and whether differences in patient characteristics and OAC use would account for that variability. Age-, sex-, and race- (when applicable) standardized stroke rates did not cluster in the South as hypothesized based on the distribution of stroke rates in the general population. Across databases, rates were highest in the Middle Atlantic, East South Central, and East North Central divisions. In MarketScan, lower probabilities of anticoagulation corresponded to higher rates of stroke across census divisions when compared to New England. However, in Optum, when compared to New England, the probability of anticoagulation fills within division did not inversely correspond with stroke rates as expected. After successive adjustments for CHA_2DS_2-VASc score and use of an OAC there were no changes in the associations, suggesting that geographic variation in stroke risk could not be explained by regional differences in these risk factors.

Table 2. *Adjusted incidence rate ratios (IRR) and 95% confidence intervals (CI) of ischemic stroke hospitalization, MarketScan, 2007-2014 and Optum Clinformatics, 2009-2015*

MarketScan N	Events 9026	Crude rate* (95% CI)	IRR (95% CI)	Optum Clinformatics N	Events 6517	Crude rate* (95% CI)	IRR (95% CI)
Model 1[†]				Model 1[†]			
New England	397	7.0 (6.3-7.7)	Ref	New England	356	10.3 (9.2-11.3)	Ref
Middle Atlantic	1253	8.4 (7.9-8.9)	1.04 (.93-1.17)	Middle Atlantic	423	9.5 (8.6-10.4)	.94 (.82-1.08)
East N Central	2483	8.8 (8.4-9.1)	1.07 (.97-1.19)	East N Central	901	9.3 (8.7-9.9)	1.01 (.90-1.14)
West N Central	374	6.8 (6.1-7.5)	1.01 (.87-1.16)	West N Central	723	8.5 (7.9-9.1)	.91 (.80-1.03)
South Atlantic	1345	6.8 (6.4-7.1)	.95 (.85-1.06)	South Atlantic	1100	8.2 (7.7-8.7)	.97 (.86-1.10)
East S Central	464	7.0 (6.3-7.6)	1.08 (.95-1.24)	East S Central	273	10.9 (9.6-12.2)	1.21 (1.03-1.41)
West S Central	795	7.1 (6.6-7.6)	1.04 (.92-1.17)	West S Central	739	8.7 (8.0-9.3)	.93 (.82-1.06)
Mountain	406	7.7 (7.0-8.4)	1.03 (.89-1.18)	Mountain	638	7.7 (7.1-8.4)	.77 (.68-.88)
Pacific	1509	8.1 (7.7-8.6)	1.03 (.93-1.16)	Pacific	1364	9.9 (9.3-10.4)	.84 (.75-.94)
Model 2[‡]				Model 2[‡]			
New England	397	7.0 (6.3-7.7)	Ref	New England	356	10.3 (9.2-11.3)	Ref
Middle Atlantic	1253	8.4 (7.9-8.9)	.99 (.89,1.11)	Middle Atlantic	423	9.5 (8.6-10.4)	.92 (.8,1.05)
East N Central	2483	8.8 (8.4-9.1)	1.02 (.92,1.14)	East N Central	901	9.3 (8.7-9.9)	1.01 (.89,1.14)
West N Central	374	6.8 (6.1-7.5)	1.00 (.87,1.15)	West N Central	723	8.5 (7.9-9.1)	.93 (.82,1.06)
South Atlantic	1345	6.8 (6.4-7.1)	.92 (.82,1.03)	South Atlantic	1100	8.2 (7.7-8.7)	.95 (.84,1.07)
East S Central	464	7.0 (6.3-7.6)	1.05 (.92,1.20)	East S Central	273	10.9 (9.6-12.2)	1.15 (.98,1.35)
West S Central	795	7.1 (6.6-7.6)	1.01 (.89,1.14)	West S Central	739	8.7 (8.0-9.3)	.91 (.80,1.03)
Mountain	406	7.7 (7.0-8.4)	1.02 (.89,1.17)	Mountain	638	7.7 (7.1-8.4)	.77 (.68,.88)
Pacific	1509	8.1 (7.7-8.6)	1.03 (.92,1.15)	Pacific	1364	9.9 (9.3-10.4)	.84 (.75,.95)
Model 3[§]				Model 3[§]			
New England	397	7.0 (6.3-7.7)	Ref	New England	356	10.3 (9.2-11.3)	Ref
Middle Atlantic	1253	8.4 (7.9-8.9)	.99 (.89,1.11)	Middle Atlantic	423	9.5 (8.6-10.4)	.92 (.80,1.06)
East N Central	2483	8.8 (8.4-9.1)	1.02 (.92,1.14)	East N Central	901	9.3 (8.7-9.9)	1.01 (.89,1.14)
West N Central	374	6.8 (6.1-7.5)	1.00 (.87,1.15)	West N Central	723	8.5 (7.9-9.1)	.94 (.83,1.07)
South Atlantic	1345	6.8 (6.4-7.1)	.92 (.82,1.03)	South Atlantic	1100	8.2 (7.7-8.7)	.95 (.84,1.07)
East S Central	464	7.0 (6.3-7.6)	1.06 (.92,1.21)	East S Central	273	10.9 (9.6-12.2)	1.16 (.99,1.35)
West S Central	795	7.1 (6.6-7.6)	1.01 (.90,1.14)	West S Central	739	8.7 (8.0-9.3)	.91 (.80,1.03)
Mountain	406	7.7 (7.0-8.4)	1.02 (.89,1.18)	Mountain	638	7.7 (7.1-8.4)	.77 (.68,.88)
Pacific	1509	8.1 (7.7-8.6)	1.03 (.92,1.15)	Pacific	1364	9.9 (9.3-10.4)	.84 (.75,.95)

CHA_2DS_2-VASc: congestive heart failure, hypertension, age ≥75 years, diabetes mellitus, stroke/transient ischemic attack, vascular disease, age 65-75 years, and sex category.
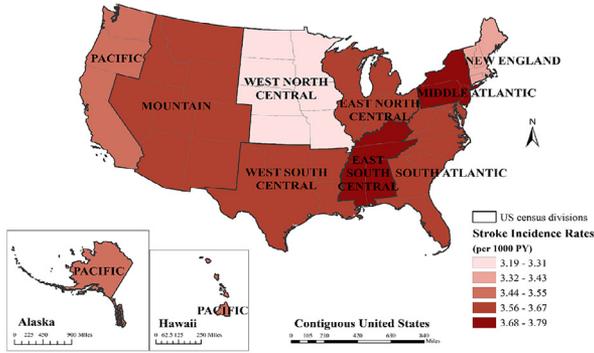*Crude stroke incidence rates per 1000 person-years.
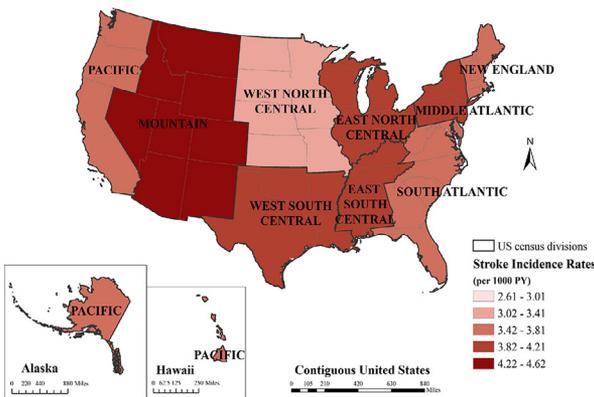[†]Model 1: Adjusted for age and sex (and race in Optum Clinformatics).
[‡]Model 2: Model 1, additionally adjusted for CHA_2DS_2-VASc variables.
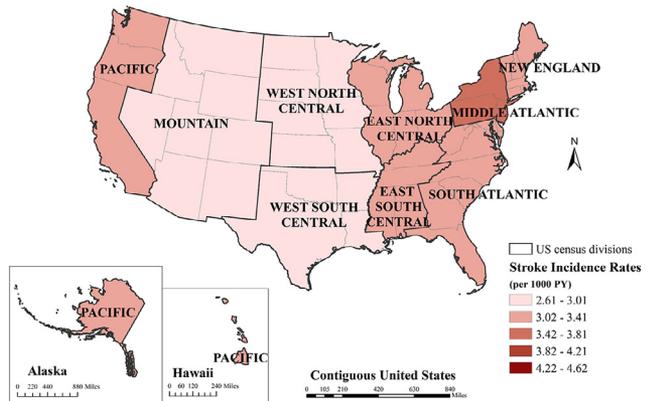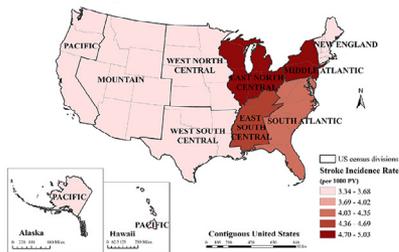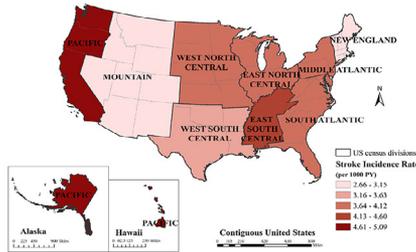[§] Model 3: Model 2, additionally adjusted for oral anticoagulant use.

Figure 1. *Standardized stroke incidence rates by US Census Bureau divisions in MarketScan, 2007-2014. (A) Age- and sex-standardized rates. (B) Age-standardized rates for women. (C) Age-standardized rates for men.*
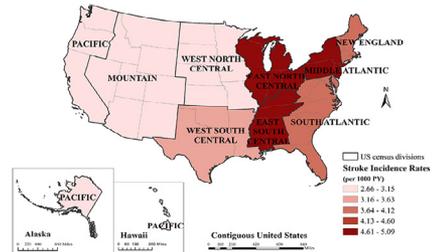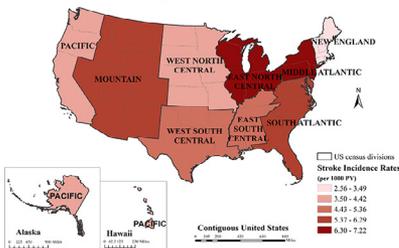

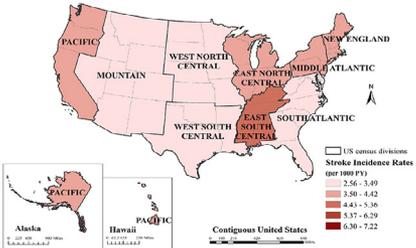**Figure 2.** *Standardized stroke incidence rates by US Census Bureau divisions in Optum Clinformatics, 2009-2015. (A) Age-, sex-, and race-standardized rates. (B and C) Age-standardized rates for women and men, respectively. (D and E) Age- and sex-standardized rates for nonwhites and whites, respectively.*

Table 3. *Anticoagulation fill patterns of patients with nonvalvular atrial fibrillation, MarketScan, 2007-2014 and Optum Clinformatics 2009-2015*

	Anticoagulation use, %	Relative Risk (95% CI)*	*P* value
MarketScan	*159,832*		
New England	29	Ref	
Middle Atlantic	24.8	.86 (.84,.88)	<.0001
East N Central	26.9	.92 (.90,.94)	<.0001
West N Central	27.9	.95 (.93,.98)	.0001
South Atlantic	26.1	.89 (.87,.91)	<.0001
East S Central	26.7	.89 (.87,.92)	<.0001
West S Central	25.9	.87 (.85,.89)	<.0001
Mountain	26.1	.89 (.87,.92)	<.0001
Pacific	25.4	.88 (.86,.90)	<.0001
Optum Clinformatics	*94,728*		
New England	26.3	Ref	
Middle Atlantic	23.0	.88 (.85,.91)	<.0001
East N Central	27.0	1.01 (.99,1.04)	.3136
West N Central	17.2	.65 (.63,.67)	<.0001
South Atlantic	24.6	.90 (.88,.92)	<.0001
East S Central	23.8	.86 (.83,.90)	<.0001
West S Central	25.5	.94 (.91,.96)	<.0001
Mountain	26.7	1.03 (1.00,1.06)	.0656
Pacific	26.7	1.06 (1.03,1.08)	<.0001

Adjusted for age, sex, CHA_2DS_2-VASc variables (and race in Optum Clinformatics) CHA_2DS_2-VASc: congestive heart failure, hypertension, age $\geq$75 years, diabetes mellitus, stroke/transient ischemic attack, vascular disease, age 65-75 years, and sex category.

*Relative risk and 95% confidence interval (CI) of anticoagulation fill.

First documented in 1965, the term stroke belt was used to describe the concentration of high stroke mortality rates in the Southeastern region compared to the rest of the United States.[14,16,17] Mortality rates can be driven by 2 components—incidence and case fatality.[22] Within recent decades, studies have found regional differences in stroke incidence[19–21,29] and stroke case fatality,[22] which correspond to the higher stroke risk in the stroke belt, and thus likely contributes to the persistent disparity in mortality rates in the Southern region of the United States. Considering AF is an important independent risk factor for stroke[9,10] and reported to cause 15%-20% of ischemic stroke events across all ages,[7] we designed this study to determine whether a similar regional pattern in stroke incidence in AF patients would emerge in the Southern region of the United States. Our findings do not support such pattern.

Many of the proposed hypotheses for regional differences in stroke incidence and mortality are based on similar regional variations in the prevalence of stroke risk factors (i.e., hypertension) and socioeconomic characteristics. Hypertension, including uncontrolled hypertension, is a major risk factor for stroke.[30] The prevalence of hypertension is higher in blacks,[31] as well as more prevalent in certain groups in the southeastern United States.[32] Howard et al found that the proportion of controlled hypertension in the United States is lower in those at older ages, and in groups with less education, lower income, or lack of access to healthcare.[33] Socioeconomic status is inversely associated with stroke incidence and stroke mortality.[34]

The complex interplay between socioeconomic status, access to healthcare and health insurance, prevalence of stroke risk factors, and the actual risk of stroke may result in differences between patterns emerging from the evaluation of individuals with health insurance those observed in the general population.

In patients with nonvalvular AF, a risk assessment scheme is used to evaluate stroke risk and help inform the decision of whether or not to initiate anticoagulation therapy. The CHA_2DS_2-VASc score is currently the recommended risk assessment scheme used to identify patients with AF who are likely to benefit from anticoagulant prophylaxis against stroke. The CHA_2DS_2-VASc score assigns 1 or 2 points to each of 7 critical risk factors, that is, congestive heart failure (C), hypertension (H), 2 points to age $\geq$75 years (A), diabetes mellitus (D), 2 points to previous or transient ischemic attack stroke (S), vascular disease (V; prior myocardial infarction, peripheral arterial disease, or aortic plaque), age 65-74 years (A) and sex category (Sc: female sex).[27] For AF patients with a CHA_2DS_2-VASc score of 2 or greater, the American College of Cardiology/American Heart Association/Heart Rhythm Society Guideline for the Management of Patients With AF recommends OACs.[35] Anticoagulation therapy reduces the risk of stroke[36] and is highly effective in stroke prevention. However, underuse of anticoagulation therapy in AF patients remains common[37] and varies across regions in the United States.[38] In this study, we found regional differences in anticoagulation fill patterns;

however, these differences did not attenuate observed stroke rates across census divisions. This suggests that regional differences in stroke rate may not be due to differences in the management of AF patients across regions but to other risk factors not measured in this study.

In our analysis, we found differences in stroke rate by sex and race which are mostly consistent with prior literature. In MarketScan, but not Optum, we found that women had higher stroke rates when compared to men. This is consistent with the inclusion of female sex as a predictor of stroke in the CHA_2DS_2-VASc score,[27] the American College of Cardiology/American Heart Association/Heart Rhythm Society recommended stroke risk prediction score. In Optum, we observed higher stroke rates in nonwhites versus whites, in most census divisions. These findings are consistent with data from the Atherosclerosis Risk in Communities cohort which reports higher strokes in blacks versus whites.[39]

Across both databases, we observed a low proportion of patients receiving anticoagulation, <30%. Anticoagulation fill patterns have been shown to be influenced by provider specialty.[40] AF patients seen by a cardiologist are 39% more like to fill an OAC compared to patients seen by a primary care physician. The population prevalence of anticoagulant use reported in our study are substantially lower than those reported in AF registries, such as ORBIT-AF (which reports provider anticoagulant prescriptions and not actual fill rates),[41] but are consistent with other administrative databases.[42,43] As such, our data are likely more representative of what occurs in a usual practice setting.

These findings should be interpreted in the context of several limitations. The results of this analysis rely on the ability to accurately ascertain both the outcome and covariates from administrative data. Validated algorithms were used to ascertain events of interest and it is likely that any misclassification is nondifferential. Second, race/ethnicity data were imputed for more than two-thirds of patients in Optum. There is the potential for substantial misclassification of race, which could result in residual confounding in race-adjusted estimates. Third, data are presented according to census level divisions and not at the state or county-level. Care should be taken when interpreting divisional stroke rates because it is known that significantly different rates can occur within a division, state, or county.[44] Fourth, the number of stroke hospitalizations identified is a portion of all strokes available in the database. We limited our definition to ICD-9-CM codes with a high PPV in the primary diagnosis of a hospital-related claim. As a result, it is possible that we excluded true cases of strokes that occurred outside the hospital setting, identified by a different ICD-9-CM code, or occurred in a different diagnostic position (resulting in a low sensitivity definition of stroke). Finally, identifying regional differences in claims data may be dependent on the regional coverage of the insurance provider. In this analysis, we used health claims data from 2 independent databases for which different insurance companies contribute to each. Of note, we did not have information on the overall geographic distribution of either of these products across the country; however, we have no reason to assume that one product is more susceptible to variance in enrollment across the country compared to the other. Given that insurance provisions vary widely by state and different insurance companies contribute to MarketScan and Optum Clinformatics, it is difficult to tease out geography from insurance coverage and the differences between databases are to be expected.

Overall, this paper serves as a cautionary tale for the use of administrative claims data to evaluate geographical variation in an outcome, given the numerous factors that can influence this variation and the difficulty to tease them out. When geographic differences are observed in these databases, it brings to question whether the observed differences are the result of differential penetration of a given insurance product or targeted disease management programs for risk factors of the outcome, or true higher prevalence of the outcome in certain regions. This is valuable information because claims have been utilized for the identification of geographic differences across the country.

In conclusion, we found geographic variation in stroke rates in patients with AF. The highest rates were not in the Southeastern region as hypothesized. These regional differences do not appear to be attributable to differences in patient characteristics or AF treatment (i.e., OACs). This finding suggests that other factors may be responsible for this observed pattern. Further investigation into existing strategies of AF management and treatment for existing comorbidities (i.e., hypertension) related to stroke risk should be evaluated.

Disclosures

None.

Supplementary Materials

Supplementary data to this article can be found online at doi:10.1016/j.jstrokecerebrovasdis.2018.12.005.

References

1. National Center for Health Statistics (US). Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities. Hyattsville, MD: National Center for Health Statistics (US); 2016.   Available from: https://www.ncbi.nlm.nih.gov/books/NBK367640/.
2. Yang Q, Tong X, Schieb L, et al. Vital Signs: Recent Trends in Stroke Death Rates — United States, 2000-2015. MMWR Morb Mortal Wkly Rep 2017;66:933-939.
3. Raji MA, Lowery M, Lin YL, et al. National utilization patterns of warfarin use in older patients with atrial fibrillation: a population-based study of Medicare Part D beneficiaries. Ann Pharmacother 2013;47:35-42.
4. Wang G, Joo H, Tong X, et al. Hospital costs associated with atrial fibrillation for patients with ischemic stroke aged 18-64 years in the United States. Stroke 2015;46:1314-1320.
5. Wetmore JB, Ellerbeck EF, Mahnken JD, et al. Stroke and the "stroke belt" in dialysis: contribution of patient

characteristics to ischemic stroke rate and its geographic variation. J Am Soc Nephrol 2013;24:2053-2061.

6. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. Circulation 2017;135:e146-e603.

7. CDC. Atrial Fibrillation. August 22, 2017 ed: National Center for Chronic Disease Prevention and Health Promotion; 2017. p. 1-2.

8. Phillips SJ. Is atrial fibrillation an independent risk factor for stroke? Can J Neurol Sci 1990;17:163-169.

9. Brand FN, Abbott RD, Kannel WB, et al. Characteristics and prognosis of lone atrial fibrillation: 30-year follow-up in the framingham study. J Am Med Assoc 1985;254:3449-3459.

10. Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. Stroke 1991;22:983-988.

11. Sussman M, Menzin J, Lin I, et al. Impact of atrial fibrillation on stroke-related healthcare costs. J Am Heart Assoc 2013;2:e000479.

12. Lin HJ, Wolf PA, Kelly-Hayes M, et al. Stroke severity in atrial fibrillation. The Framingham Study. Stroke. 1996;27:1760-1764.

13. Penado S, Cano M, Acha O, et al. Stroke severity in patients with atrial fibrillation. Am J Med 2002;112:572-574.

14. El-Saed A, Kuller LH, Newman AB, et al. Geographic variations in stroke incidence and mortality among older populations in four US communities. Stroke 2006;37:1975-1979.

15. Liao Y, Greenlund KJ, Croft JB, et al. Factors explaining excess stroke prevalence in the US Stroke Belt. Stroke 2009;40:3336-3341.

16. Lanska DJ, Kuller LH. The geography of stroke mortality in the United States and the concept of a stroke belt. Stroke 1995;26:1145-1149.

17. Casper ML B.E., Williams G.I. Jr., Halverson J.A., et al. Atlas of Stroke Mortality: Racial, Ethnic, and Geographic Disparities in the United States.

18. National Heart L, and Blood Institute Stroke Belt Initiative. 1996:1-23.

19. Howard VJ, Kleindorfer DO, Judd SE, et al. Disparities in stroke incidence contributing to disparities in stroke mortality. Ann Neurol 2011;69:619-627.

20. Nguyen-Huynh MN, Johnston SC. Regional variation in hospitalization for stroke among Asians/Pacific Islanders in the United States: a nationwide retrospective cohort study. BMC Neurol 2005;5:21.

21. Lackland DT, Bachman DL, Carter TD, et al. The geographic variation in stroke incidence in two areas of the southeastern stroke belt: the Anderson and Pee Dee Stroke Study. Stroke 1998;29:2061-2068.

22. Labarthe DR, Howard G, Safford MM, et al. Incidence and case fatality at the county level as contributors to geographic disparities in stroke mortality. Neuroepidemiology 2016;47:96-102.

23. Garg RK, Glazer NL, Wiggins KL, et al. Ascertainment of warfarin and aspirin use by medical record review compared with automated pharmacy data. Pharmacoepidemiol Drug Saf 2011;20:313-316.

24. Williams LS, Rotich J, Qi R, et al. Effects of admission hyperglycemia on mortality and costs in acute ischemic stroke. Neurology 2002;59:67-71.

25. DeFrank JT, Michael BJ, Rimer BK, et al. Triangulating differential nonresponse by race in a telephone survey. Prev Chronic Dis. 2007;4:A60.

26. Quan H SV, Halfon P, Fong A, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130-1139.

27. Lip GY, Nieuwlaat R, Pisters R, et al. 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.

28. Howden LMaM, Julie A. Age and sex composition. 2010. In: Bureau USC, editor. 2010 Census Briefs. Onlin 2011. p. 1-16.

29. Gillum RF, Ingram DD. Relation between residence in the southeast region of the United States and stroke incidence. The NHANES I epidemiologic follow-up study. Am J Epidemiol 1996;144:665-673.

30. Benjamin EJ, Virani SS, Callaway CW, et al. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. Circulation 2018;137:e67-e492.

31. Hertz RP, Unger AN, Cornell JA, et al. Racial disparities in hypertension prevalence, awareness, and management. Arch Intern Med 2005;165:2098-2104.

32. Obisesan TO, Vargas CM, Gillum RF. Geographic variation in stroke risk in the United States. Region, urbanization, and hypertension in the Third National Health and Nutrition Examination Survey. Stroke 2000;31:19-25.

33. Howard G, Prineas R, Moy C, et al. Racial and geographic differences in awareness, treatment, and control of hypertension: the REasons for Geographic And Racial Differences in Stroke study. Stroke 2006;37:1171-1178.

34. Bravata DM, Wells CK, Gulanski B, et al. Racial disparities in stroke risk factors: the impact of socioeconomic status. Stroke 2005;36:1507-1511.

35. 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.

36. Lip GY, Lim HS. Atrial fibrillation and stroke prevention. Lancet Neurol 2007;6:981-993.

37. Ogilvie IM, Newton N, Welner SA, et al. Underuse of oral anticoagulants in atrial fibrillation: a systematic review. Am J Med 2010;123:638-645. e4.

38. Hernandez I, Saba S, Zhang Y. Geographic variation in the use of oral anticoagulation therapy in stroke prevention in atrial fibrillation. Stroke 2017;48:2289-2291.

39. Magnani JW, Norby FL, Agarwal SK, et al. Racial Differences in atrial fibrillation-related cardiovascular disease and mortality: the Atherosclerosis Risk in Communities (ARIC) study. JAMA Cardiol 2016;1:433-441.

40. O'Neal WT, Sandesara PB, Claxton JS, et al. Provider specialty, anticoagulation prescription patterns, and stroke risk in atrial fibrillation. J Am Heart Assoc 2018;7:e007943.

41. Fosbol EL, Holmes DN, Piccini JP, et al. Provider specialty and atrial fibrillation treatment strategies in United States community practice: findings from the ORBIT-AF registry. J Am Heart Assoc 2013;2:e000110.

42. Lang K, Bozkaya D, Patel AA, et al. Anticoagulant use for the prevention of stroke in patients with atrial fibrillation: findings from a multi-payer analysis. BMC Health Serv Res 2014;14:329.

43. Willey V, Franchino-Elder J, Fu AC, et al. Treatment and persistence with oral anticoagulants among newly diagnosed patients with non-valvular atrial fibrillation: a retrospective observational study in a US commercially insured and Medicare Advantage population. BMJ Open 2018;8:e020676.

44. Roth GA, Dwyer-Lindgren L, Bertozzi-Villa A, et al. Trends and patterns of geographic variation in cardiovascular mortality among US counties, 1980-2014. JAMA 2017;317:1976-1992.