ATRIAL FIBRILLATION IN OLDER ADULTS: RELATION TO PROTEOMICS, RISK PREDICTION, AND URBAN/RURAL DISPARITIES IN TREATMENT AND OUTCOMES

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

FAYE L NORBY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Aaron R. Folsom, MD, MPH

July 2020

ACKNOWLEDGEMENTS

I would like to thank my advisor, Aaron Folsom, for his continual guidance, mentorship, and support during the dissertation process. His expertise in epidemiological methods, along with clinical insight, cultivated an enjoyable learning experience. I would like to thank my dissertation committee members, Pam Lutsey, Lin Yee Chen, and Nathan Shippee for pushing me to think more deeply about my dissertation topics, answering my questions, and generously sharing their knowledge with me. I would also like to thank Alvaro Alonso for his initial mentorship and continued support in my research.

I would like to acknowledge my support during the PhD program from the UMN Regents Scholarship.

My fellow colleagues and PhD students at UMN have enriched my life and have contributed to this research. I would like to particularly thank Kelsie Full, Madison Anderson, Wendy Wang, Kim Bonner, Emily Groene, Dawn Nederhoff, Nathan Mitchell, Megan Winker, Rob Walker, and Jeff Misialek. Additionally, I would like to acknowledge those who encouraged me to enter the PhD program and supported me throughout including Lindsay Bengtson, Jill Dreyfus, Liz Bell, Kara Whitaker, Mike Bancks, Duke Appiah, Amber Fyfe-Johnson, Kristen George, and Mary Rooney.

My friends have been a constant source of support and encouragement during my training. I would like to specifically thank Erin Berglund and the members of the Setters Camp Dropouts volleyball team for never allowing me to take myself too seriously. Thank you to my siblings, Deana, Miranda, Matt and Adrienne, who molded me at a young age to form the basis for who I am today. Finally, thank you to Jeff Leuwerke, who has been a steadfast source of support and my biggest fan.

ABSTRACT

Atrial fibrillation (AF), a cardiac arrhythmia, is a major public health problem. AF is largely a disease of advancing age and contributes to other cardiovascular complications. Identification of novel protein biomarkers could advance understanding of AF mechanisms and may improve the prediction of incident AF. Additionally, it is unknown if disparities exist in AF treatment and outcomes in rural versus urban areas of the US.

For manuscripts 1 and 2, we used data from the Atherosclerosis Risk in Communities (ARIC) study, a cohort of older-aged adults in the US. For manuscripts 3 and 4, we used a sample of Medicare beneficiaries enrolled from 2011-2016 with residential zip code categorized into 4 rural/urban areas.

In the first manuscript, we examined the association of plasma proteins and identified 40 novel protein biomarkers associated with incident AF. These biomarkers provide insight into mechanistic pathways of AF development. In the second manuscript, we derived and validated a series of 5-year incident AF prediction models that are better targeted and calibrated to older populations. Incorporating biomarkers, including proteomics data, into the models improved AF risk prediction. In the third and fourth manuscripts, we examined the initiation of anticoagulation use and compared the risks of subsequent stroke, heart failure, myocardial infarction, and mortality in newly-diagnosed AF patients in rural versus urban areas. Patients in rural areas were more likely to initiate anticoagulant treatment; however, they were less likely to initiate a newer class of anticoagulants compared to those in urban areas. Those in rural areas had modestly higher risk of cardiovascular outcomes and mortality compared to those in urban areas.

Proteomics aids in understanding AF mechanisms and improves risk prediction. Future research should validate our prediction models, develop meaningful ways to incorporate protein biomarkers in clinical practice, and focus on improving AF treatment in rural areas.

1. Table of Contents

1.1. LIST OF TABLES

1.2. LIST OF FIGURES

1.3. OTHER ITEMS

A supplemental appendix consisting of an excel workbook accompanies Manuscript 1

2. INTRODUCTION

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia, with a lifetime risk in the United States (US) of 1 in 3 among whites and 1 in 5 among African Americans.[1] AF is largely a disease of advancing age[2, 3] and is associated with increased risks of adverse cardiovascular outcomes including stroke,[4] myocardial infarction (MI),[5] and mortality,[6] resulting in significant costs to the US healthcare system.[7] Once AF develops, patients have a 5-fold increased risk of stroke compared to those without AF,[4] and therefore the mainstay of stroke prevention in AF is the initiation and maintenance of anticoagulant therapies.[8]

The risk of AF increases with advancing age, European ancestry, cigarette smoking, taller height, greater weight, higher blood pressure and corresponding blood pressure medication use, diabetes, history of MI, and history of heart failure (HF).[9, 10] In addition to the traditional clinical risk factors listed above, various biomarkers have been identified as risk factors for incident AF including markers of inflammation,[11-14] oxidative stress,[15] myocardial necrosis,[11, 16, 17] myocardial stress,[11, 18-23] and mineral metabolism.[24, 25] Identification of novel biomarkers, beyond what is currently known, could advance our understanding of AF mechanisms. New technology has allowed for the systematic assessment of a large portion of the entire range of proteins measurable in plasma (the plasma proteome), commonly referred to as proteomics. The application of proteomics provides opportunities for unbiased discovery of novel markers and has the potential to advance our understanding of disease mechanisms.

The growing public health significance of AF has spurred efforts to identify individuals at higher risk of developing this arrhythmia and its complications. Several AF risk prediction scores have been developed with respectable discriminative abilities in middle-aged adults. However, these scores may have diminished discrimination in populations of older adults, in which AF is most prevalent. A well-calibrated risk score for prediction of incident AF would optimize screening in high-risk older individuals, allow for more specific clinical trial enrollment, and would lead to opportunities for targeted preventive strategies. Improving established risk prediction scores may depend on whether novel markers, such as proteomics, can add to or refine predictive models.

There are nearly 60 million people (19% of the population) living in rural areas according to the US Census Bureau. Those in rural areas have higher rates of cardiovascular risk factors such as cigarette smoking, hypertension, diabetes, and obesity.[26-28] There are rural vs. urban disparities in cardiovascular disease (CVD) in the US, such as a 40% higher heart disease prevalence in rural areas and higher risk of stroke.[29, 30] Despite the higher risk of stroke from AF, there is little known regarding anticoagulation rates in AF patients in rural vs.

urban areas. In addition, it is unknown if CVD disparities exist in AF patients living in rural versus urban areas of the US.

3. PATHOPHYSIOLOGY OF ATRIAL FIBRILLATION

3.1. NATURAL HISTORY

AF was first reported to affect humans in 1906 when 2 publications reported that "auricular fibrillation" was common in heart disease patients and that it could be identified by a new instrument, the electrocardiograph (ECG).[31] AF is an uncoordinated atrial tachyarrhythmia caused by rapid and irregular atrial depolarization which results in ineffective atrial contraction.[32] Key ECG findings are the following: a loss of P waves and replacement by fibrillatory waves; erratic activation of the ventricles resulting in an irregular, rapid heart rate (usually 90 to 170 beats per minute [bpm]); and a narrow QRS complex, unless other conduction abnormalities coexist, as seen in **Figure 3.1**.[33] Note the absence of distinct P wave, chaotic activity of atria, irregular R-R intervals with narrow QRS complex.[34]

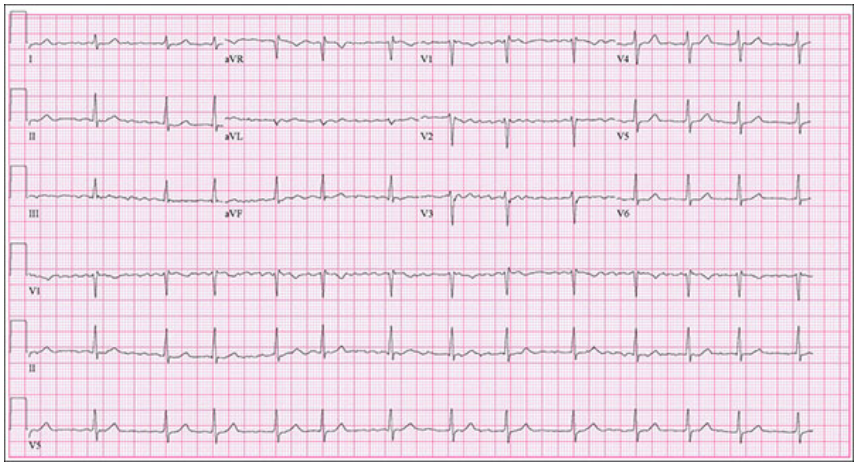


Figure 3.1. Electrocardiogram showing atrial fibrillation.

Initiation of AF requires a trigger, and in order for AF to persist, the trigger must remain or electrical remodeling that promotes AF in the absence of the trigger must occur.[33] Micro-reentry and enhanced automaticity in one or more atrial circuits are the most common triggers for AF, and can be attributed to any number or combination of mechanisms that are not mutually exclusive. Underlying atrial pathology can result in AF and atrial fibrosis and loss of atrial muscle mass are the most common pathoanatomic changes.[35] Atrial dilation can be caused by any type of cardiovascular disease (CVD) associated with AF, including hypertension, heart failure (HF) and atherosclerosis, [34] and atrial fibrosis can be triggered by

many factors including inflammation.[13] Heterogeneity of electrical conduction is caused, at least in part, by fibrotic atrial fibers juxtaposed on healthy atrial tissue.[36] Additionally, sympathetic and parasympathetic activation can provoke or worsen AF by shortening the atrial refractory period, which increases susceptibility to reentry and enhanced automaticity.[37] Once AF occurs, AF itself produces changes in atrial function and structure and provides a possible explanation for the progressive nature of this arrhythmia. [36]

3.2. TYPES OF AF

AF results from several disease processes, each with different prognoses. Nonvalvular AF occurs in the absence of rheumatic valve disease, a mechanical or bioprosthetic valve, or mitral valve abnormalities. AF can also be caused by valvular disease; however, nonvalvular AF is the most common form of AF and will be the focus on this dissertation. Paroxysmal AF is arbitrarily defined as AF that is episodic and resolves spontaneously or with intervention within 7 days. Persistent AF lasts more than 7 days. Permanent AF indicates a decision to discontinue attempts to restore or maintain sinus rhythm. Paroxysmal and permanent forms carry the same long-term risk of stroke.[38] AF can also be due to noncardiac diseases, referred to as secondary AF; treating its cause often resolves the arrhythmia.

3.3. DIAGNOSIS

Patients with AF may present with mild or no symptoms, HF, myocardial infarction (MI), stroke, or hemodynamic collapse.[34] Common symptoms, if they appear, include fatigue, palpitations, chest pain, syncope, dizziness, dyspnea, and orthopnea.[34] A classical sign of AF is an irregularly irregular pulse. A systematic review was conducted to determine the accuracy of pulse palpation to detect AF compared with an ECG diagnosis of AF. The review found checking pulse rate is 94% sensitive and 72% specific for diagnosis,[39] indicating this method is useful for ruling out AF at that time. Even suspected AF based on evaluation of pulse rate should always be confirmed with 12-lead ECG, and if the patient has AF, the ECG will show an absence of P waves, as seen in Figure 1. However, a normal test result does not completely rule out the presence of AF because an ECG may not capture a paroxysmal arrhythmia. When clinical suspicion of AF persists despite normal ECG results, a Holter monitor (typically a 24 to 48-hour recording), or ambulatory event monitor (>24 hours, up to 30 days) may be required.

4. DESCRIPTIVE EPIDEMIOLOGY-INCIDENCE AND PREVALANCE

AF is the most common sustained cardiac arrhythmia, with a lifetime risk of 1 in 3 among whites and 1 in 5 among African Americans.[1] Both the incidence and prevalence of AF have been steadily increasing, and the aging of the population and accompanying rise in prevalence have magnified its morbidity and health care costs. In 2010, AF affected between 2.7 and 6.1 million Americans; the prevalence is estimated to increase to between 5.6 and 12.1 million by 2030.[40-42] AF is largely a disease of advancing age, as risk doubles with each progressive decade of gaining and exceeds 20% by age 80 years.[2, 3] The prevalence of AF increases with older age, from 0.1% among people younger than 55 years to 9% among people 80 years or older.[40]

Several population-based cohort studies have estimated the incidence of AF in the US. In Olmsted County, Minnesota, a predominately white population, the age- and sex-adjusted incidence of AF in the county increased 12.6% in 20 years: from 3.04 (95% confidence interval [CI]: 2.78 – 3.31) per 1000 person-years in 1980 to 3.68 (95% CI: 3.42 – 3.95) per 1000 person-years in 2000.[41] AF is more prevalent in men and also more often diagnosed in whites compared to blacks. The Atherosclerosis Risk in Communities (ARIC) study (age 45-64 at baseline) found that crude incidence rates of AF were 6.7, 4.0, 3.9, and 3.0 per 1,000 persons per year in white men, white women, black men, and black women, respectively. In ARIC, compared to whites, blacks had a 41% (95% CI: 8% - 62%) lower age- and sex-adjusted risk of being diagnosed with AF.[9] Several cohort studies also report incidence rates by age. The Framingham Heart Study (FHS) found an overall incidence rate of 12.6 per 1000 person-years, and the incidence increased in age, ranging from 6.2 and 3.8 cases per 1000 person-examinations in men and women, respectively, aged 55 to 64 years, to 75.9 and 62.8 cases per 1000 person-examinations in men and women aged 85 to 94 years.[43] Similarly, the Cardiovascular Health Study (CHS), a population-based study among adults aged ≥ 65 years, calculated an overall AF incidence rate of 19.2 per 1,000 person- years.[44] Onset of AF in CHS was also strongly associated with age and male sex; the incidence rate per 1,000 person-years among men aged 65 – 74 and 75 – 84 was 17.6 and 42.7, respectively, and the corresponding rates for women were 10.1 and 21.6, respectively.[44]

In addition to utilizing cohort studies to determine AF incidence, administrative data from Medicare have been used. In the Medicare population, in those 65 years and older, the age- and sex-adjusted incidence rate of AF per 1,000 person-years was virtually unchanged from 1993 to 2007 at 27.3 and 28.3, respectively.[45] Similarly to population based studies, incidence increased substantially with age and men and whites had consistently higher rates.[45]

Administrative data from the Medicare population indicates that the prevalence of AF has increased during the last several decades, and data from 2003-2007 shows a mean 5% increase in prevalence per year.[45] The magnitude of prevalence increase was greatest among the oldest beneficiaries which were those age 90 and older.[45]

5. RISK FACTORS

A number of important risk factors for developing AF have been identified and the main clinically-based risk factors include advancing age, male sex, white race, height, weight, systolic and diastolic blood pressure, use of hypertension medications, diabetes, obesity, MI, and HF. Each of these, along with several other risk factors are discussed in this section. Risk factors in the context of risk prediction scores for incident AF are more thoroughly discussed in a later chapter.

5.1. DEMOGRAPHIC

Demographic characteristics, including age, sex and race are known to be associated with AF. Age is the most important nonmodifiable risk factor for AF, with the incidence of AF doubling with each decade of life.[43] The prevalence of AF increases from 0.5% at age 50-59 to 9% at age 80-89.[46] The median age of patients with AF is 75, and about 70% of AF patients are between 65 and 85 years old.[47, 48]

Male sex has consistently been associated with increased risk of AF.[9, 41, 43, 44] In the FHS, men had a 1.5-fold higher risk (95% CI: 1.3-1.8) of developing AF compared to women after adjustment for age and other risk factors.[43] However, the prevalence is high among both men and women at older age.[40] In CHS, AF prevalence was higher among men than women among 65-69 year olds; however, the prevalence was similar between men and women among 70-79 year olds.[44, 49]

The majority of epidemiologic data on AF in the US is based on those with white race. Despite limited data on racial/ethnic differences, there is evidence that differences do exist. In the US, incidence and prevalence of AF is lower among blacks even though blacks have a higher prevalence of risk factors for AF.[50, 51] In the ARIC study, the age and sex adjusted incidence of AF in blacks was 41% lower (95% CI: 8%-62%) compared to whites.[9] Among other racial groups, whites have an increased risk of AF when compared to blacks, Asians or Hispanics.[52, 53]

5.2. BEHAVIORAL AND CLINICAL

The association between behavioral risk factors such as alcohol and exercise and the risk of AF has been studied in observational cohort studies. Episodic heavy alcohol consumption, coined “holiday heart” is known to be associated with the onset of AF.[54] Results regarding habitual consumption of low to moderate amounts of alcohol are mixed; FHS found an increased risk of AF among those who consumed > 3 drinks/day, but no association at

lower levels.[55] CHS found no association between moderate alcohol consumption and development of AF;[56] however, a meta-analysis reported an increased risk of 1.08 (95% CI: 1.05-1.10) for each 10 gram / day increment of alcohol.[57]

The association between exercise and the risk of AF appears to vary by intensity of exercise and age of the population in the study. Numerous studies have reported an increased risk of AF, particularly AF without any underlying CVD risk factors, among elite athletes and extreme exercisers.[58-61] Among 2 studies of middle-age populations, no associations were observed between physical activity and risk of AF.[62, 63] Among adults ≥65, exercise intensity had a U-shaped association with AF; light to moderate physical activity was associated with a reduced risk of AF.[64]

Incidence rates, relative hazards and population attributable fractions for AF in the ARIC study by common risk factor profiles of blood pressure, body mass index (BMI), diabetes, and smoking are listed in **Table 5.1**, adapted from Huxley et al.[65] Hypertension is consistently one of the most important contributors to the burden of AF, and there is a linear relationship between increasing systolic and diastolic blood pressure and AF risk.[10, 43, 44] Elevated blood pressure was the most important independent contributor of AF risk, accounting for 21.6% (95% CI: 16.8-26.7) of incident AF cases in ARIC.[65]

Smoking is considered a moderate to strong risk factor for AF. In ARIC, after multivariable-adjustment, the risk of AF was 2 times higher in current smokers, and 1.3 times in former smokers compared to never smokers.[66] These results are similar to earlier findings in the Rotterdam Study[67] and the Manitoba Follow-up study.[68]

Type-2 diabetes mellitus and obesity are moderate to strong risk factors for AF. In ARIC, type-2 diabetes was associated with a 35% increased risk of incident AF (95% CI: 1.14-1.60),[69] and a meta-analysis reported that diabetes patients had a 34% greater risk of AF compared to non-diabetics (95% CI: 1.07-1.68).[70] Obesity consistently has been associated with an increased risk of incident AF. Obesity and overweight accounted for 17.9% of all incident AF cases in the ARIC study,[65] and a meta-analysis based on 5 population-based cohort studies reported a 49% increased risk (95% CI: 1.36-1.64) in obese compared to non-obese adults.[71]

Table 5.1. Incidence rate, relative hazard (95% confidence intervals) and population attributable fractions for atrial fibrillation for risk factors in the Atherosclerosis Risk in Communities study, 1987 – 2007

	No. at Risk	No. Incident AF	IR*	RH (95% CI) †	PAF %	95% CI
History of Cardiac Disease (%)						
Optimal	13398	1259	5.00	0.54 (0.46-0.62)	0.00	-
Elevated	1200	261	12.17	1 [Ref]	5.35	3.32 to 7.45
Blood pressure (%)						
Optimal	5626	381	3.93	0.55 (0.48-0.63)	0.00	-
Borderline	3317	304	4.72	0.65 (0.56-0.74)	2.89	-0.11 to 5.64
Elevated	5655	835	7.65	1 [Ref]	21.6	16.8 to 26.7
BMI (%)						
Optimal	4889	389	4.27	0.65 (0.56-0.74)	0.00	-
Borderline	5767	591	5.28	0.70 (0.62-0.79)	5.16	0.93 to 9.26
Elevated	3942	531	7.36	1 [Ref]	12.7	9.30 to 16.3
Diabetes (%)						
Optimal	7558	645	4.68	0.67 (0.58-0.78)	0.00	-
Borderline	5491	617	5.83	0.71 (0.61-0.82)	0.78	-3.52 to 4.84
Elevated	1533	253	8.77	1 [Ref]	3.08	0.91 to 5.30
Smoking (%)						
Optimal	6077	510	4.23	0.55 (0.48-0.62)	0.00	-
Borderline	4769	550	5.76	0.60 (0.52-0.68)	2.06	-2.05 to 6.05
Elevated	3752	460	7.45	1 [Ref]	9.78	6.74 to 12.9

IR = Incidence Rate of AF per 1000 person-years adjusted for age (mean age = 54.2 years); †Adjusted for age, gender, race, study site, education, income and height and each of the other risk factors.
Table published in *Huxley et al.*[65]

There are several additional clinical risk factors associated with incident AF including a higher risk of AF in those with taller height, [10, 44, 51] low estimated glomerular filtration rate (eGFR),[72] varied associations with lipid levels,[73, 74] very low or very high resting heart rate[75-77], low magnesium,[78] and high phosphorus.[79]

Individuals with major cardiovascular comorbidities such as MI and HF are at an increased risk of developing AF. In the FHS, history of MI was associated with an adjusted odds ratio (OR) of developing AF of 1.4 (95% CI: 1.0-2.0) in men and 1.2 (95% CI: 0.8-1.8) in women.[43] In the ARIC study, prevalent MI was associated in a 2.21 times higher risk (95%

CI: 1.71-2.84) of AF compared to those without MI after demographic adjustment.[29]. Other population-based cohort studies have found a similar increased risk.[10, 49, 51, 68] HF is consistently and substantially associated with an increased risk of AF. [43, 49, 51, 68, 80] In the ARIC study, prevalent HF was associated with 3 times the risk (95% CI: 2.32-3.95) of AF compared to those without HF after adjustment for age, sex and race.[51] Other cohort studies have found a similar in the range of 2 to 3 times higher risk of AF in those with HF.[10]

5.3. BLOOD BIOMARKERS

Incident AF has been associated with a diverse array of circulating biomarkers, including markers of inflammation (high-sensitivity C-reactive protein, fibrinogen),[18, 81, 82] atrial overload (atrial and B-type natriuretic peptides),[18, 81, 82] myocardial ischemia (high-sensitivity troponin T and I),[11, 16, 81] cardiac fibrosis (galectin-3),[83, 84] and others (soluble ST2, growth differentiation factor-15).[11]

5.4. ELECTROCARDIOGRAPHY

Several ECG- derived variables have been associated with an increased risk of AF including the PR interval, [10, 51, 85] ECG-based left ventricular hypertrophy,[10, 51, 85] QRS duration,[86] and a prolonged QT interval.[87] Several P wave indices have been associated with incident AF. A pooled analysis of the FHS and ARIC cohorts found associations with P wave duration, P wave area, and P wave terminal force with the incidence of AF.[88] Abnormal P wave axis was associated with AF, reporting a hazard ratio (HR) of 2.34 (95% CI: 2.12-2.58) after multivariable adjustment.[89] Finally, an analysis of 1260 participants in the CHS cohort found that a doubling of hourly premature atrial contractions count from 24-hour Holter monitoring was associated with a 17% increased risk of AF (95% CI: 1.13-1.22).[90]

5.5. IMAGING

Information on cardiac structure and function obtained from echocardiographic studies, such as left atrial diameter, left ventricular function, left ventricular mass, and left ventricular wall thickness, have been associated with incident AF.[85, 91-94] CHS found significant independent associations of measures of ventricular diastolic filling parameters including left atrial size, peak E velocity, and A wave velocity time interval with incident AF.[92]

5.6. GENETICS

Recent research has identified several common genetic variants associated with the risk of AF.[95] Genome-wide association studies (GWAS) in individuals of European descent have identified three genomic regions associated with AF on chromosomes 4q25 (*PITX2*), 16q22 (*ZFHX3*), and 1q21 (*KCNN3*). [96-98] Studies of electrocardiographic traits have also identified a number of loci associated with AF.[99, 100] A GWAS meta-analysis further identified 6 new susceptibility loci in or near plausible candidate genes involved in pacemaking activity, signal transduction and cardiopulmonary development. [95]

6. ADVERSE OUTCOMES ASSOCIATED WITH AF

6.1. STROKE

Ischemic stroke and systemic arterial occlusion in AF are primarily attributed to embolism of a thrombus as a result of blood stasis in the left atrium.[35] Individuals with AF have a 5-fold higher stroke incidence compared to those without AF, [4] and among those ≥ 75 years of age, AF is the most important single cause of ischemic stroke.[101] Among Olmsted County residents diagnosed with AF, 11% had a first ischemic stroke over a mean follow-up of 5.5 years.[102] A collaborative analysis of 5 randomized controlled trials identified age, hypertension, previous transient ischemic attack or stroke, and diabetes as independent risk factors for stroke among AF patients.[103] In the FHS, the relative risk (RR) of stroke associated with AF was fairly stable across age groups; however, the attributable risk increased significantly with age, from 1.5% among 50-59 year olds to 23.5% among 80-89 year olds. [4] Although ischemic stroke in AF patients is primarily attributed to an AF-related embolism, up to 25% of stroke in AF patients might be the result of intrinsic cerebrovascular diseases, other cardiac sources of embolism or atherosclerotic pathology in the proximal aorta.[104, 105] Treatment with anticoagulants greatly reduces the risk of stroke and is discussed in a later chapter.

6.2. MYOCARDIAL INFARCTION AND HEART FAILURE

The development of AF is also associated with subsequent increased risk of the major cardiovascular conditions of MI[5, 106] and HF.[80, 107] In ARIC, AF was associated with a 63% increased risk of MI (HR=1.63; 95% CI: 1.32-2.02); however, when type of MI was considered, AF was associated with non-ST-segment elevation MI [HR (95 CI) = 1.80 (1.39-2.31)] but was not associated with ST-segment elevation MI.[5]

Studies looking at the timing of AF and HF show a bidirectional association, often with one developing within a few years of the other.[80] The existing severity of specific cardiovascular risk factors, along with age and sex, may determine whether AF or HF occurs first.[108, 109] For example, in FHS, among 382 individuals with both AF and HF, 38% had AF first, 41% had HF first, and 21% had both diagnosed on the same day.[109] AF is one of the strongest risk factors for HF. In a recent meta-analysis, AF was associated with nearly 5-times of the risk of incident HF; RR (95%CI) = 4.99 (3.04-8.22). The absolute risk increase in incident HF associated with AF was 11.1 (5.7 to 20) events/1000 participant years.[110]

6.3. COGNITIVE IMPAIRMENT AND DEMENTIA

There is increasing evidence that AF is a risk factor for cognitive impairment and dementia. A meta-analysis restricted to studies conducted among non-stroke AF patients reported a pooled OR of dementia of 1.64 (95% CI: 1.00 – 2.71) comparing those with and without AF.[111] A prospective cohort study among the residents of Olmsted County with incident AF and without stroke found the cumulative incidence rate of dementia was 2.7% at one year, and 10.5% at 5 years after AF.[112] In ARIC, incident AF has been associated with additional cognitive decline and incident dementia, independent of clinical stroke,[113, 114] and this association might be explained by the presence or development of subclinical cerebral infarcts.[115]

6.4. BLEEDING EVENTS

Even though oral anticoagulation in AF reduces the risk of ischemic stroke and systemic thromboembolism, this benefit is accompanied by an increased bleeding risk.[116-120] Therefore gastrointestinal (GI) bleeding and other major bleeding events are important outcomes to consider when looking at AF outcomes. Identifying individuals at higher risk of bleeding complications when using oral anticoagulants may facilitate personalized treatments. To date, at least four risk scores for the assessment of bleeding risk in patients with AF treated with vitamin K antagonists have been published (summarized in **Table 6.1**).[121-124] Variables consistently associated with increased bleeding risk in this patient population include older age, renal disease, and a history of prior bleeding, with anemia, cancer, and hypertension included in several of the scores. As with the scores used for stroke risk stratification, the existing bleeding predictive models have only moderate discrimination and do not differ significantly from each other when applied to the same population.[125-127] Of note, these scores have been developed for the prediction of bleeding among persons with AF using vitamin K antagonists (warfarin). DOACs may have a different bleeding profile compared to vitamin K antagonists, and scores are currently being developed for patients using DOACs.

Table 6.1. Risk stratification schemes for bleeding prediction in AF patients using vitamin K antagonists

Risk score	HEMORR$_2$HAGES [121]	HAS-BLED [122]	ATRIA[123]	ORBIT-AF[124]
Variables	Age Hepatic / renal disease Hypertension Prior bleeding Stroke Alcohol abuse Anemia Cancer Reduced platelet count / function Genetic factors Fall risk	Age Abnormal renal / liver function Hypertension Prior bleeding Stroke Drugs/alcohol Labile INR	Age Renal disease Hypertension Prior bleeding Anemia	Age Abnormal kidney function Prior bleeding Anemia Antiplatelet use (heart failure) (cancer) (COPD) (hip fracture / osteoporosis) (smoking)

INR: International Normalized Ratio

6.5. MORTALITY

The presence of AF has been shown to independently increase the risk of death and the mortality risk is highest during the first year after AF manifests.[6, 107, 128, 129] The age-adjusted mortality rate from AF was 6.5 per 100,000 people in 2017. As might be expected, the annual mortality rates associated with AF vary substantially depending on the population demographics. Based on medical insurance claim data, values range from 2.6% in asymptomatic untreated individuals, to 24.2% amongst an elderly population with high rates of comorbidities.[45, 130] The risk of death among those with first-detected AF is particularly high during the months immediately following diagnosis. Among residents of Olmsted County, MN , the age- and sex-matched HRs for mortality among those with new-onset AF compared to those without AF were 9.62 (95% CI: 8.93 – 10.32) in the first 4 months after diagnosis, and fell to 1.66 (95% CI: 1.59 – 1.73) in subsequent follow-up period.[6] Among Medicare beneficiaries, mortality following an AF diagnosis is 3.5 times higher than expected; 30-day and one-year mortality are 12.6% and 27.6%, respectively.[45]

Evidence indicates the risk of mortality in AF patients may differ by sex and race. In the FHS, the multivariable adjusted OR for death among those with AF compared to those without AF was of 1.5 (95% CI: 1.2 – 1.8) in men and 1.9 (95% CI: 1.5 – 2.2) in women.[128] In addition, there was a significant sex interaction such that AF appeared to diminish the survival advantage typically observed in females.[128] A meta-analysis has also found that the adjusted risk of death was significantly stronger in females than in males with AF [RR (95%

CI) =1.2 (1.07-1.17)].[131] The Women's Health Study (WHS) has additionally corroborated these findings.[132] In a Medicare unadjusted analysis, blacks and Hispanics had a higher risk of death than their white counterparts with AF; however, after adjustment for comorbidities, blacks (HR, 0.95 [95% CI, 0.93–0.96]; *P*<0.001) and Hispanics (HR, 0.82 [95% CI, 0.80-0.84]; *P*<0.001) had a lower risk of death than whites with AF.[133] In contrast, in the population-based ARIC study, the rate difference for all-cause mortality for individuals with AF versus without AF was nearly double in blacks compared to whites.[113] During AF-related hospitalizations, in-hospital mortality has been shown to be highest amongst African-Americans in comparison to other ethnic groups.[134]

Although stroke is the most feared complication of AF, the RE-LY clinical trial reported that stroke accounted for only ~7.0% of deaths in AF, with sudden cardiac death (SCD) (22.25%), progressive HF (15.1%), and noncardiovascular death (35.8%) accounting for the majority of deaths.[135] In a study that examined data from 2 population based studies, AF was associated with a doubling in the risk of SCD after accounting for baseline and time-varying confounders. In ARIC, the unadjusted incidence rate per 1000 person-years was 1.30 (95% CI: 1.14–1.47) in those without AF and 2.89 (95% CI: 2.00–4.05) in those with AF; corresponding rates in CHS were 3.82 (95% CI: 3.35–4.35) and 12.00 (95% CI, 9.45–15.25), respectively. When the 2 cohort studies combined results, the multivariable-adjusted HR associated with AF for SCD was 2.47 (95% CI: 1.95–3.13).[136] In a meta-analysis of 7 studies, individuals with AF had an RR of SCD of 1.88 (95% CI: 1.36–2.60).[110]

7. TREATMENT

7.1. OVERVIEW

The primary goals for treating patients with AF are improvement of symptoms and reduction of AF-related morbidity. The 3 basic tenets for therapy of AF are 1) control of ventricular rate responses, or rate control; 2) restoration and maintenance of sinus rhythm, or rhythm control; and 3) prevention of thromboembolism. In patients who are hemodynamically unstable, immediate evaluation and treatment are warranted, including emergency cardioversion, if necessary. In stable patients, treatment depends on the duration of AF and the presence of underlying cardiac disease or other comorbidities. **Figure 7.1** (adapted from Gutierrez et al[34], updated 2016) presents an algorithm showing the key decision-making points in the process.

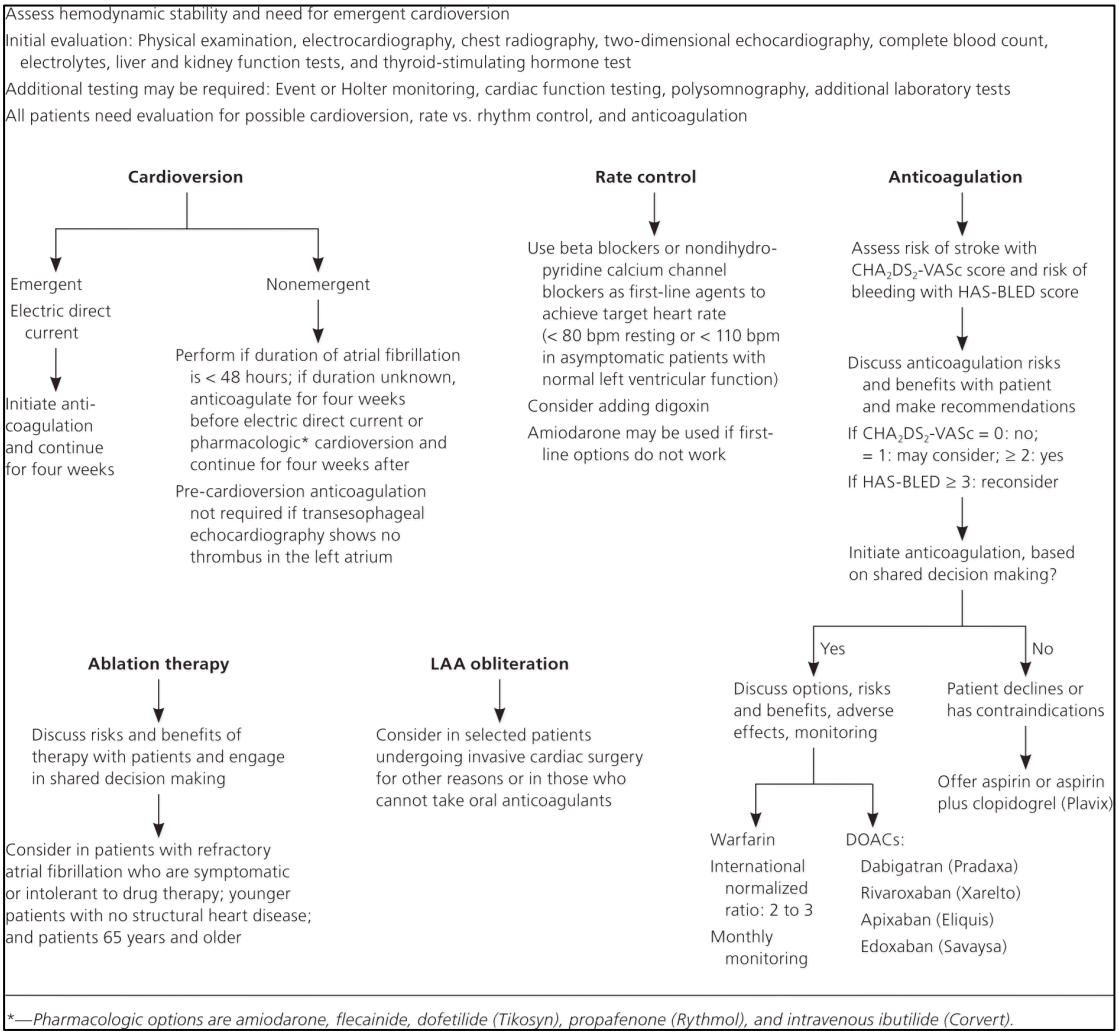


Figure 7.1. Algorithm for the evaluation and treatment of a patient with AF.

7.2. RATE CONTROL

Rate control is an essential part of AF treatment in acute and chronic settings. It promotes hemodynamic function by slowing ventricular response, improving diastolic ventricular filling, reducing myocardial oxygen demand, and improving coronary perfusion and mechanical function. Given the challenges of achieving and maintaining normal sinus rhythm and the deleterious effects of antiarrhythmic drugs, most patients with AF are treated with rate control.[34, 137, 138]

Beta blockers or nondihydropyridine calcium channel blockers are used to achieve heart rate goals. Lenient rate control to achieve a resting rate less than 110 bpm is reasonable in the majority of patients.[139] Stricter rate control (less than 80 bpm during rest) may be appropriate if needed to resolve symptoms. Beta blockers and calcium channel blockers are contraindicated in patients with preexcitation (Wolff-Parkinson-White syndrome). Non-cardioselective beta blockers are also contraindicated in patients with acute heart failure, severe chronic obstructive pulmonary disease, and asthma. Digoxin is no longer considered a first-line agent or recommended as monotherapy, but it can be added to therapy with beta blockers or calcium channel blockers. Amiodarone offers another choice for rate control when beta blockers and calcium channel blockers do not work, but its delayed action, potential toxicity, and drug interactions severely limit its use. It may also cause acute cardioversion, which could lead to a stroke if anticoagulation therapy has not been properly administered.[34]

7.3. CARDIOVERSION

The main indication for cardioversion is unstable or poorly tolerated AF that is unresponsive to drug therapy.[35, 140] Unless done emergently, or when the duration of the arrhythmia is known to be less than 48 hours, 4 weeks of pre- and post-cardioversion anticoagulation is required. Cardioversion can be attempted electrically or pharmacologically. Electrical cardioversion is usually successful in the short term, but often not in the long term. If transesophageal echocardiography shows no thrombus in the left atrium, it is safe to omit pre-cardioversion anticoagulation.[35, 140] Electrical cardioversion delivers a direct-current electric shock in synchrony with the QRS complex to avoid triggering ventricular fibrillation. One or more shocks of 200 to 300 joules may be necessary.[35, 140]

Pharmacologic cardioversion uses intravenous ibutilide, flecainide, dofetilide, propafenone, or amiodarone. Cardioversion and maintenance of normal sinus rhythm using medication are challenging because of the limited long-term effectiveness of antiarrhythmics,

the risk of triggering ventricular arrhythmias, and long-term adverse effects. In general, maintenance of normal sinus rhythm with oral medications is more successful in patients 65 years and younger with structurally normal hearts, as well as patients who have only recently developed AF.[35, 140] Contraindications to either form of cardioversion include known atrial thrombus, digitalis toxicity, multifocal atrial tachycardia, and suboptimal anticoagulation.

7.4. ABLATION THERAPY

Electrophysiologic radiofrequency ablation is a nonoperative, catheter-based procedure used to isolate and possibly destroy abnormal foci responsible for AF. Specific foci that cause AF have been found at or near the pulmonary vein ostia in the left atrium; locating these sites allows targeted ablation.[34] Some trials have shown that radiofrequency ablation is superior to antiarrhythmics in selected patients, including patients with paroxysmal AF who are symptomatic but without structural heart disease, patients who are intolerant of antiarrhythmics, and patients with inadequate pharmacologic rhythm control.[141, 142] However, data on the long-term effectiveness and safety of radiofrequency ablation are limited. AF may recur after ablation, and a repeat procedure may be required in approximately 20% of cases. Ablation of the accessory pathway is the optimal treatment for patients with Wolff-Parkinson-White syndrome and AF.[35, 140] The procedure is contraindicated in patients who cannot be anticoagulated one month before and at least several months after the procedure.

7.5. SURGERY AND PERCUTANEOUS LEFT ATRIAL APPENDAGE ISOLATION

Surgical treatments for AF are invasive, high risk, and are considered only in patients undergoing cardiac surgery for other reasons.[34] The primary surgical therapies for treating AF are the Maze procedure and left atrial appendage (LAA) obliteration. The Maze procedure aims to eliminate AF through the use of incisions in the atrial wall to interrupt arrhythmogenic wavelet pathways and reentry circuits.[143] LAA obliteration reduces stroke risk by percutaneous ligation or surgical removal of the LAA.[144-146] LAA obliteration does not correct the underlying AF; however, because approximately 90% of cardiac thrombi occur in the appendage, it decreases the subsequent risk of stroke.[34] Two percutaneously inserted devices, the Watchman and the Amplatzer Cardiac Plug, can be used to achieve occlusion of the LAA, although the latter is not yet approved for AF treatment in the US. Both are non-inferior to warfarin (Coumadin) in stroke risk reduction.[147, 148] A newer device, the Lariat, is available to ligate the LAA, but data on its long-term effectiveness and safety are limited.[147]

7.6. ANTICOAGULATION

Anticoagulation is an essential part of AF management. It significantly reduces the risk of embolic stroke, but increases the risk of bleeding. Although the benefit of anticoagulation exceeds the risk of bleeding for most patients, discussions about stroke prevention versus risk of bleeding remain challenging especially in elderly patients. Tools to aid in the assessment of the risks of stroke and bleeding are available and are useful in making decisions with patients about therapeutic options. They are discussed in other chapters.

For many years, the $CHADS_2$ (congestive heart failure; hypertension; age 75 years or older; diabetes mellitus; prior stroke, transient ischemic attack, or thromboembolism [doubled]) scoring system was used to estimate risk of stroke in patients with AF. Anticoagulation was recommended for patients with a $CHADS_2$ score of 2 or more, unless a contraindication is present.[121] More recently, the CHA_2DS_2-VASc (congestive heart failure; hypertension; age 75 years or older [doubled]; diabetes; prior stroke, transient ischemic attack, or thromboembolism [doubled]; vascular disease; age 65 to 74 years; sex category) scoring system has been recommended by the American College of Cardiology.[35] Due to a high risk for stroke in AF patients, the current ACC/AHA/HRS Guideline for the Management of Patients with AF recommends oral anticoagulation in those with a prior stroke or transient ischemic attack, or those with a moderate or greater risk of stroke (CHA_2DS_2-VASc score $\geq$1 in males or $\geq$2 in females).[35] Importantly, >80% of all AF patients are in this risk category of recommended anticoagulation.[149] CHA_2DS_2-VASc significantly increases the number of patients eligible for anticoagulation compared with $CHADS_2$.

Warfarin

Historically, warfarin (Coumadin) has been prescribed since the 1950's as an oral anticoagulant for stroke prevention in patients with AF. AF patients taking warfarin have a stroke risk reduction >60% compared to placebo.[150] However, warfarin requires frequent monitoring of international normalized ratio (INR) measures, and many patients have difficulty achieving and maintaining optimal INR measures. Even with optimal compliance, patients using warfarin are within the therapeutic range (2 to 3 for AF) only 55% to 66% of the time.[34] Additional limitations include a significant risk for bleeding complications, a narrow therapeutic range and the presence of numerous dietary and drug interactions. Therefore, many patients eligible for anticoagulation are not receiving it (up to 40% in some

studies[151]) and among those using warfarin, misuse results in inadequate protection against stroke and/or increased risk of bleeding.

Direct oral anticoagulants

There were 4 direct oral anticoagulants (DOACs) approved by the Food and Drug Administration (FDA) between 2010 and 2015 for the prevention of stroke and systemic embolism in patients with AF. Dabigatran (brand name Pradaxa), was the first to be FDA-approved (Oct. 2010), followed by rivaroxaban (Xarelto) approved Nov. 2011, apixaban (Eliquis), approved Dec. 2012), and Edoxaban (Savaysa) approved in 2015. AF patients taking DOACs experience lower rates of stroke and intracranial bleeds compared to those taking warfarin.[152-154] Additional advantages of DOACs compared to warfarin include fixed dosing, no food interactions, and no need for INR monitoring. Their major drawbacks are higher costs, and until recently, reversal agents were not available. The lack of a reversal agent is important to consider because if a provider believed an AF patient is at a higher risk of falls, injury, or bleeding, the provider may be likely to prescribe warfarin, which can be easily reversed. **Table 7.1** outlines the pharmacologic properties of DOACs and warfarin; none are recommended for patients on hemodialysis, nor are they approved for use during pregnancy or in patients with valvular AF or advanced kidney disease.

Table 7.2 compares some of the risks and benefits of DOACs vs. warfarin, using information from the Phase 3 randomized controlled trials of each drug.[152-155] Dabigatran is as effective as warfarin in preventing stroke and systemic emboli. Major bleeding events were similar to those of warfarin, with fewer intracranial bleeds, but increased GI bleeding. Rivaroxaban and edoxaban have are noninferior in preventing stroke and systemic thromboembolic events compared to warfarin, although edoxaban has a lower rate of major bleeding. Apixaban is superior to warfarin in stroke prevention and has a lower bleeding risk. Follow-up analyses of dabigatran, rivaroxaban, and apixaban in large commercially insured databases in the US indicate similar results as those in the clinical trials.[119, 120, 156-159]

Table 7.1. Pharmacologic Properties of Anticoagulants Used for the Prevention of Thromboembolism in Patients with Atrial Fibrillation

Drug	Year FDA approved	Mechanism	Dosing	Oral bio-availability	Time to effect (hours)	Half-life (hours)	Year reversal agent approved	Cost*
Apixaban (Eliquis)	2012	Factor Xa inhibitor	5 mg twice daily 2.5 mg twice daily for patients with ≥ 2 of the following: creatinine > 1.5 mg/dL, age > 80 years, weight < 132 lb	58%	3 to 4	8 to 15	2018	($445)
Dabigatran (Pradaxa)	2010	Direct thrombin inhibitor	150 mg twice daily 75 mg twice daily for CrCl 15 to 30 mL /min/1.73m^2 Not recommended if CrCl <15 mL/min/1.73m^2	3% to 7%	1 to 2	12 to 17	2015	($417)
Edoxaban (Savaysa)	2015	Factor Xa inhibitor	60 mg daily 30 mg daily if CrCl 15 to 50 15 mL/min/1.73m^2 Avoid use if CrCl >95 15 mL/min/1.73m^2 due to increased clearance Not recommended if CrCl < 1515 mL/min/1.73m^2 Avoid in Child-Pugh Class B or C liver disease	62%	1 to 2	10 to 14	none approved	($365)
Rivaroxaban (Xarelto)	2011	Factor Xa inhibitor	20 mg daily 15 mg daily for CrCl 15 to 50 15 mL/min/1.73m^2 Not recommended if CrCl < 15 15 mL/min/1.73m^2	60%	2 to 4	5 to 9	2018	($449)

Warfarin (Coumadin)	1954	Vitamin K antagonist	Variable (dose adjusted to internal normalized ratio (INR))	100%	72 to 96	40	Various reversals have been used since 1954	Varies by dose, $3 to $19

CrCl = creatinine clearance

*Estimated retail cost for one month of standard therapy based on information obtained at http://www.goodrx.com (accessed June 6, 2019). Medicare plan was arbitrarily listed as MedicareBlue Rx Standard for Minnesota in order to obtain prices. Coupons were available for use. Generic prices not available; brand price listed in parentheses. Prices vary based on Medicare plan.

Table 7.2. Risks and Benefits of Direct Oral Anticoagulants Compared with Warfarin

Selected Clinical Outcome	**Apixaban** 5mg twice daily HR (95% CI) NNT per 2 years	**Dabigatran** 150mg twice daily RR (95%CI) NNT or NNH per 2 years	**Edoxaban** 60 mg daily HR (95% CI) NNT or NNH per 3 years	**Rivaroxaban** 20mg daily HR or RR (95% CI) NNT or NNH per 3 years
Stroke or systemic emboli	HR = 0.79 (0.66 to 0.95) NNT = 168	RR = 0.66 (0.53 to 0.82) NNT = 91	HR = 0.79 (0.63 to 0.99) NNT = 141	HR = 0.79 (0.65 to 0.95) NNT = 134
Intracranial bleed	HR = 0.51 (0.35 to 0.75) NNT = 238	RR = 0.26 (0.14 to 0.49) NNT = 182	HR = 0.54 (0.38 to 0.77) NNT = 172	HR = 0.67 (0.47 to 0.93) NNT = 247
Major bleed	HR = 0.69 (0.60 to 0.80) NNT = 79	RR = 0.93 (0.81 to 1.07) nonsignificant	HR = 0.80 (0.71 to 0.91) NNT = 66	HR = 1.04 (0.90 to 1.20) nonsignificant
Gastrointestinal bleed	R = 0.89 (0.70 to 1.15) nonsignificant	RR = 1.50 (1.19 to 1.89) NNH = 100	HR = 1.23 (1.02 to 1.50) NNH = 167	RR = 1.45 NNH = 101
Any cause of death	HR = 0.89 (0.80 to 0.99) NNT = 132	RR = 0.88 (0.77 to 1.00) nonsignificant	HR = 0.92 (0.83 to 1.01) nonsignificant	HR = 0.85 (0.70 to 1.02) nonsignificant

CI = confidence interval; HR = hazard ratio; NNH = number needed to treat for a specific time to cause an adverse event; NNT = number needed to treat for a specific time to prevent an outcome; RR = relative risk.

Table adapted from Gutierrez et al,[34] using information from the main randomized controlled trials of each drug.[152-155]

7.7. AF ETIOLOGY AND TREATMENT IN OLDER VERSUS YOUNGER ADULTS

AF is a heterogeneous condition, with significant differences in its epidemiology, pathogenesis, clinical presentation and management across age groups, as shown in **Table 7.3,** adapted from *Sankaranarayanan et al.*[160] Older patients are more likely to have an abnormal substrate and present at an advanced stage with atypical symptoms and associated comorbidities. The important differences between AF in younger and older adults necessitate clearly defined diagnostic and targeted management strategies to relieve symptoms as well as to prevent complications.

Table 7.3. Common differences between AF in the young versus elderly

	AF in the young	AF in elderly patients
Causes	(i) Idiopathic	(i) Ischemic heart disease
	(ii) Genetic	(ii) Heart failure
	(iii) Alcohol, smoking	(iii) Valvular heart disease
	(iv) Personality traits	(iv) Hypertension
	(v) Body mass index	(v) Cardiomyopathies
	(vi) Endurance sports	(vi) Hyperthyroidism
	(vii) Cardiac pathologies	(vii) Secondary causes such as post operative, infection, pulmonary embolism
	(viii) Endocrine disorders	(viii) Idiopathic
Pathogenesis	Triggers/pulmonary vein Repetitive activity Substrate/atrial abnormalities	Pulmonary vein repetitive activity Atrial Abnormalities
Clinical features	Usually typical symptoms	Atypical symptoms or asymptomatic
Management	Rhythm control preferred Thromboprophylaxis usually not required unless based on CHADS2VASC	Rate control preferred Thromboprophylaxis usually required unless contraindicated

A study using Medicare data (age ≥ 65) reports race- and sex-related differences in care for newly diagnosed with AF. Females were less likely to receive oral anticoagulation compared to males, and blacks and Hispanics were less likely to receive oral anticoagulation compared to whites.[161] Possible explanations include racial differences in access, patient preferences, treatment bias, and unmeasured clinical characteristics. This study uses data from 2010-2011, and therefore most patients were prescribed warfarin instead of DOACs. It is unknown if these differences still exist in Medicare beneficiaries since the uptick in DOAC prescription.

7.8. REFERRAL TO CARDIOLOGY PROVIDERS

The treatment of nonvalvular AF should be individualized to each patient's condition, which can change over time. Referral to a cardiologist is recommended for patients with complex cardiac disease; those who cannot tolerate AF despite rate control; those who need rhythm control, require ablation therapy, or may benefit from surgical treatment; and those who need a pacemaker or defibrillator because of another rhythm abnormality.[34, 35, 140] Reports suggest cardiology providers are more likely to prescribe oral anticoagulants compared with primary care providers,[162-165] and this possibly results in a lower risk of stroke among patients who are managed by cardiology specialists.[163]

8. RURAL CARDIOVACULAR DISPARITIES IN THE UNITED STATES

There are nearly 60 million (19% of the population) people living in rural areas according to the US Census Bureau. Those in rural areas have higher rates of adverse cardiovascular risk factors such as cigarette smoking, hypertension, diabetes, and obesity.[26-28] There are known rural vs. urban disparities in cardiovascular disease (CVD) in the US, such as a 40% higher heart disease prevalence in rural areas and higher risk of stroke[29, 30] although it is thought that these CVD disparities are driven mainly by race and socioeconomic status.[166] Additionally, the life expectancy gap is widening between rural and urban areas; in 2009 it was 79.1 years in metropolitan areas, as compared with 76.7 years in rural areas.[167] In the next few sections we will discuss rural health disparities, and what evidence, if any, exists regarding rural disparities in AF treatment and outcomes.

8.1. RURALITY AND HEALTH

Rural beneficiaries make up nearly 19% of the overall US population and nearly 25% of the Medicare population.[168] Health disparities between rural and urban residents are widespread and both rural providers and patients face specific challenges with health and health care delivery. As the US becomes increasingly urbanized, there is growing concern that rural areas are at risk of falling even farther behind on health metrics.[168, 169] Rural communities face particular challenges in recruiting and sustaining an adequate health care workforce, and rural hospitals tend to have higher financial strain and slimmer margins. Low volumes, hospital market consolidation, and resulting financial pressures threaten many rural hospitals with closure, which could reduce access to care even further.[170] Understanding these patterns will be critical to ensuring that the rural-urban gaps do not widen even further.

Changes in Medicare's payment landscape have also made research on rurality increasingly important.[171] As hospitals, health plans, and providers are increasingly being held accountable for quality performance, patient outcomes, and efficiency of care under a variety of value-based payment programs and alternative payment models, the degree to which differences in context may impact performance is a critical question. A number of organizations, including the National Advisory Committee on Rural Health, the National Academy of Medicine (formerly Institute of Medicine), and the National Quality Forum, have recently issued reports suggesting that rurality may be an important element that should be addressed in the design and implementation of Medicare payment and policy changes.[169]

8.2. AF DIAGNOSIS IN RURAL POPULATIONS IN THE UNITED STATES

There is little information regarding AF diagnosis in rural populations in the US. **Figure 8.1** depicts a county-coded map of AF hospitalizations in the US Medicare population, and we see a higher concentration in the southeast section of the US, with a distribution similar to maps of the US depicting the high concentration of strokes in the "stroke-belt" region. However, this map is only accurate to a county-level and does not show rural / urban status at the individual level. In fact, very little research has been conducted looking at patient-level rural / urban AF diagnosis prevalence in the US.

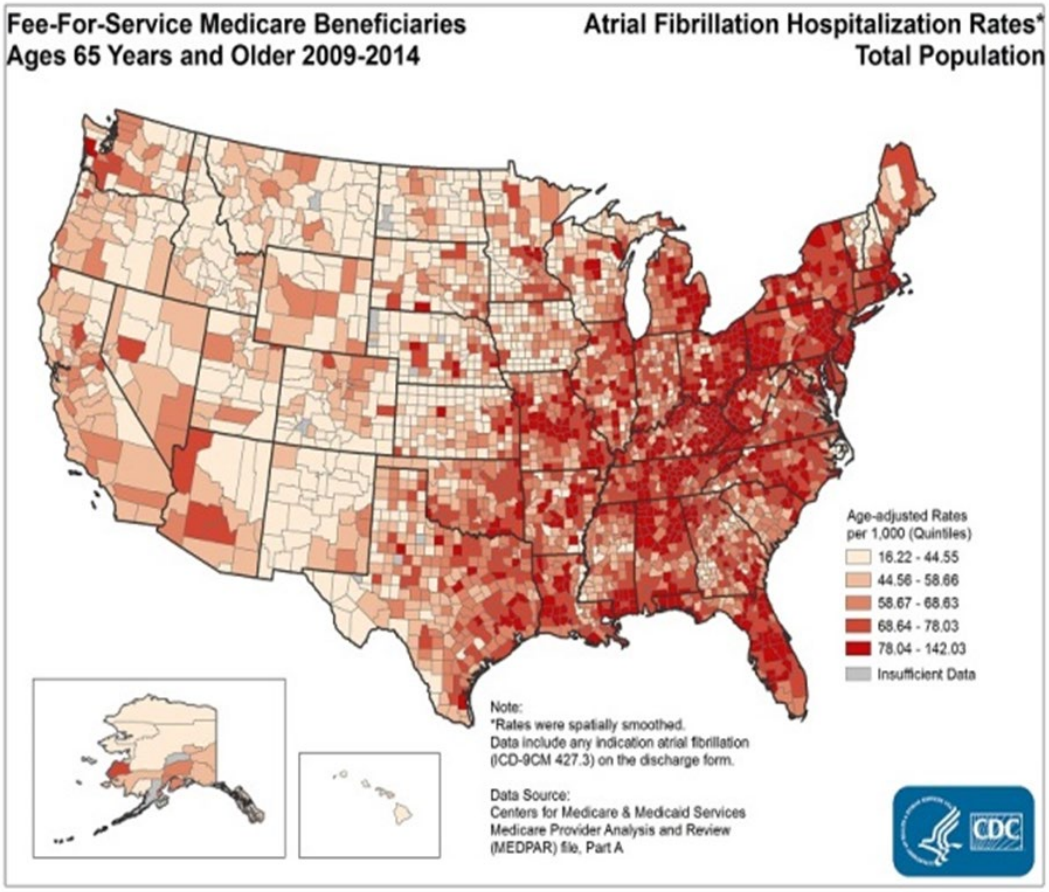


Figure 8.1. Atrial fibrillation rates for Medicare beneficiaries by county-level, 2009-2014.

8.3. RURAL DISPARITIES IN AF OUTCOMES

Few research studies have been conducted regarding rural vs. urban disparities in outcomes in AF patients in the US. A recent paper showed in-hospital mortality of AF patients is higher in rural hospitals than in urban hospitals, and these results persisted across sex, race, and region.[172] However, this study looks at the hospital location rather than the patient location, although one could say an urban patient is not as likely to be admitted to a rural hospital. Furthermore, this paper does not take into account distance to clinic, prior treatment such as anticoagulation, patient socioeconomic status, insurance, nor does it report other outcomes such as stroke. No research has been conducted looking at AF-related outcomes by individual-level geography in the US.

8.4. RURAL DISPARITIES IN AF TREATMENT

Individuals with AF have a 5-fold increased risk of stroke compared to those without AF and therefore the mainstay of stroke prevention in AF is the initiation and maintenance of anticoagulant therapies.[8] In regards to current recommended oral anticoagulants, the DOACs have fewer drug interactions, more predictable pharmacological profiles, an absence of major dietary effects, and a reduced risk of intracranial bleeding compared with warfarin.[35] Currently, DOACs account for >50% of anticoagulants prescribed for AF patients, and are associated with a higher percentage of AF patients receiving these recommended anticoagulant therapies.[173, 174] Past studies have shown that underuse of warfarin is common for Medicare beneficiaries,[175] and that elderly rural patients with AF received warfarin less frequently than elderly urban patients despite having a similar high-risk profile.[176] Due to the individualized approaches to INR monitoring needed for warfarin patients, along with numerous limitations, it has been suggested that rural patients should be considered for DOACs instead.[177] **Figure 8.2** from Hernandez et. al [178] shows the regional variation in anticoagulant use in Medicare patients in 2013-2014. The adjusted probability of receiving any anticoagulation (warfarin or DOACs) is the top panel, and the probability of receiving a DOAC is the bottom panel. Any anticoagulant use was lowest in the south, and DOAC use was lowest in the northern US. To date, no studies have looked at DOAC treatment by rural status in the US.

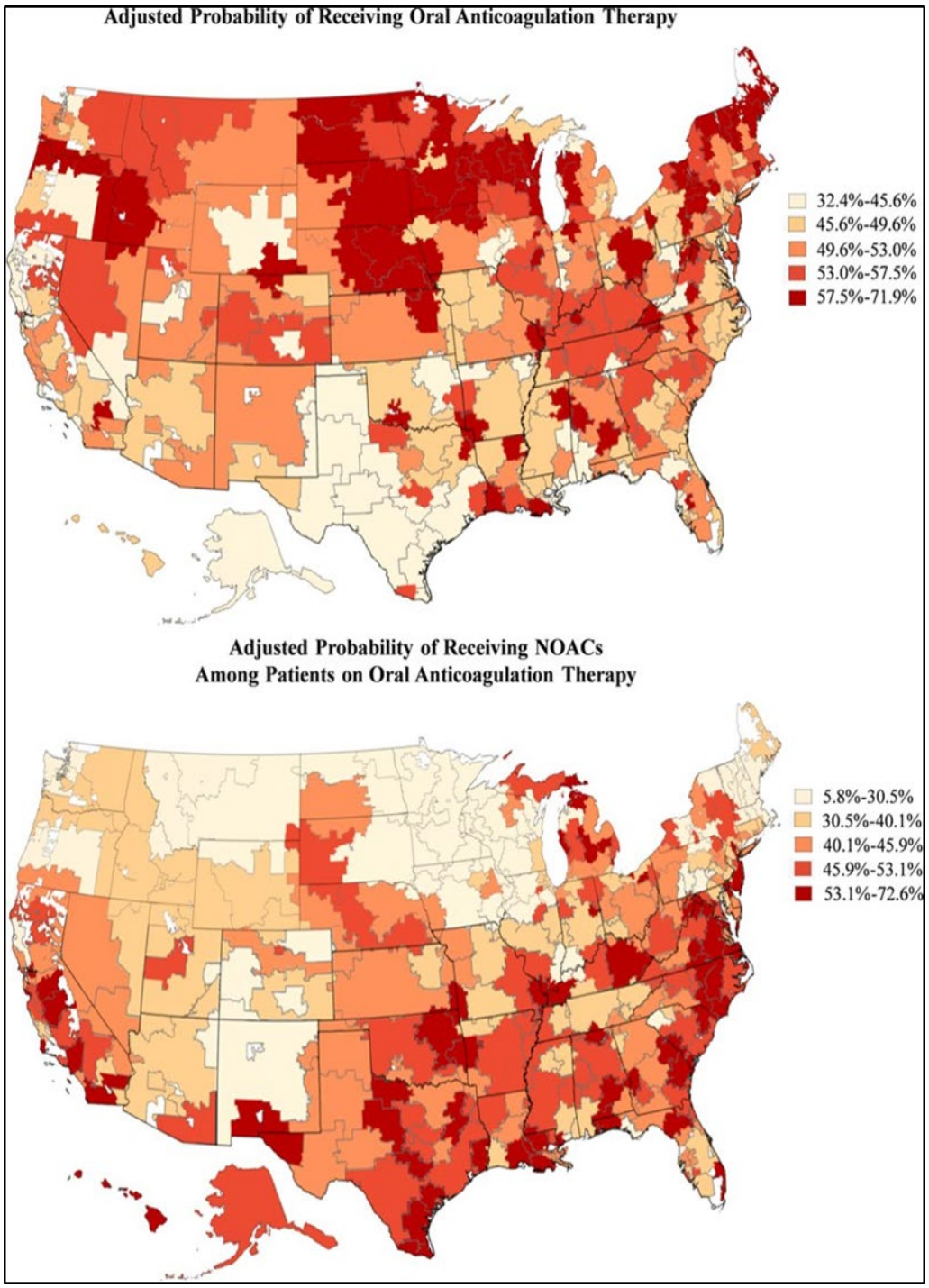


Figure 8.2. Regional variation in anticoagulant use in Medicare patients in 2013-2014.

8.5. RURAL DISPARITIES IN CARDIOLOGY INVOLVEMENT

Reports suggest cardiology providers are more likely to prescribe oral anticoagulants compared with primary care providers,[162-165] and this possibly results in a lower risk of stroke among patients who are managed by cardiology specialists.[163] Rural residents are more likely to be seen by a primary care doctor versus a cardiologist, which could be one reason for the underutilization of anticoagulants in elderly rural patients.[176] Furthermore, since early cardiology involvement after AF diagnosis increases the use of oral anticoagulants, this leads to a lower risk of stroke in patients seen by cardiology.[163] Therefore, early cardiology involvement in rural AF patients may increase access to oral anticoagulant therapy and reduce future stroke events in this high-risk population. There is little known regarding anticoagulation rates in AF patients in rural vs. urban areas, and nothing published addressing DOAC prescriptions in rural areas. A careful examination of anticoagulant use in rural vs. urban areas, taking into account provider specialty is needed. Differences in the rate of anticoagulation and/or DOAC use may identify an area of practice improvement for providers to reduce the burden of stroke in this high-risk group.

8.6. DEFINING A RURAL MEDICARE POPULATION

We captured beneficiary zip code at the time of AF diagnosis. We mapped zip codes to Rural-Urban Communing Area (RUCA) codes, which are approximation codes developed by the University of Washington Research[179] and commonly used to define rural and urban areas.[180] RUCA codes combine standard Census definitions with area commuting behaviors to capture functional and work relationships between regions.

This rural-urban taxonomy offered a more precise definition of "rural" relative to definitions based on population size alone, and has been used in other studies.[169, 181] Furthermore, it offered a more granular measure of rurality than county/-based measures such as metropolitan statistical areas, which are common measures in many claims databases. There are several ways in which rural and urban categories can be defined, and the Rural Health Research Center gives a number of suggested ways to categorize the data. We used a 4- category classification to access the rurality of beneficiaries: urban (RUCA codes 1-3, 4.1, 5.1, 7.1, 8.1, 10.1), large rural (RUCA codes 4.0, 4.2, 5.0, 5.2, 6.0, 6.1), small rural (RUCA codes 7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2) and isolated (RUCA codes 10, 10.2, 10.3, 10.4, 10.5, 10.6). **Figure 8.3** depicts an example how these 4 categories are distributed in the upper Midwest. The US population breakdown of these 4 categories across the entire Medicare population is as follows: urban = 81%; Large rural = 9.6%; small rural=5.2%; isolated = 4.2%.

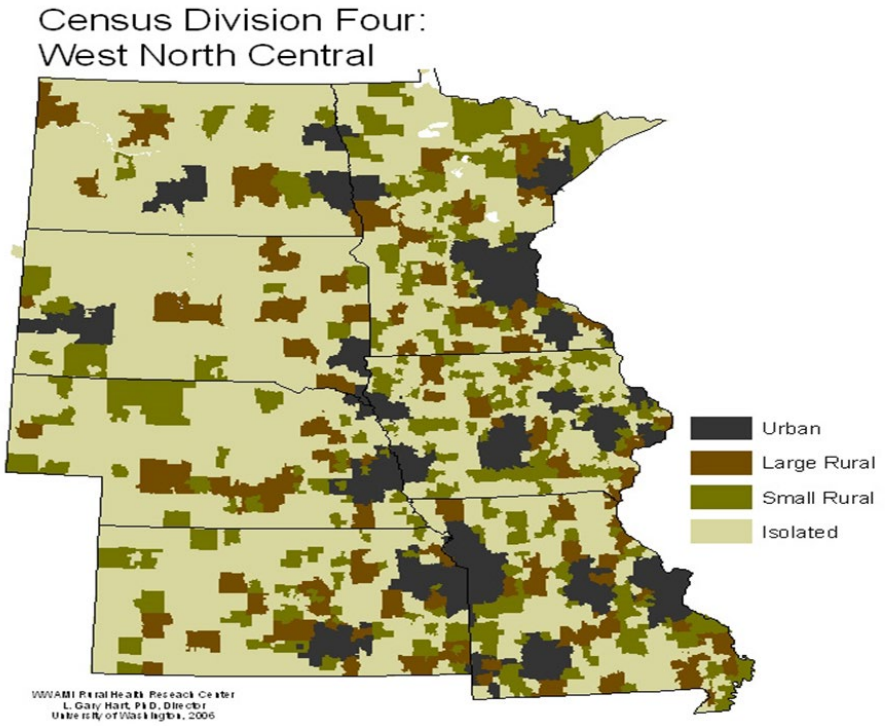


Figure 8.3. An example figure of the 4-level urban / rural classification using the Rural Health Research Center classification scheme based on ZIP codes.

9. PROTEOMICS

Cardiovascular diseases are the leading causes of death and hospitalization globally, dominated by coronary artery disease, HF, AF, and stroke.[182] Profiling of plasma proteins has been crucial for decision making in cardiovascular medicine since the introduction of many immunoassays in the 1980s, most prominently for the diagnosis of MI (creatine kinase, troponins) and HF (natriuretic peptides) and for cardiovascular risk stratification (lipoproteins). However, each of these markers suffers from limitations in diagnostic or predictive accuracy. Systematic assessment of a large portion of the entire range of proteins measurable in plasma (the plasma proteome) provides opportunities for unbiased discovery of novel markers to improve accuracy, generate pathophysiological insights, and identify therapeutic targets.

9.1. PROTEOMIC PROFILING

Proteins are vital parts of living organisms, with many functions. Although the genome remains relatively the same, the levels of proteins within various parts of the body are constantly changing in response to both internal and external factors (e.g., diet, aging, drug treatments, microorganisms, stress, etc.). Technological advances have enabled the identification of ever increasing numbers of proteins, along with their composition, structure and activity. Systematic profiling of a larger portion of the plasma proteome may provide opportunities for unbiased discovery of novel markers to improve diagnostic or predictive accuracy. In addition, proteomic profiling may inform pathophysiological understanding and point to novel therapeutic targets under the guiding hypothesis that different diseases and conditions each have novel protein profiles, including in the early stages of onset.

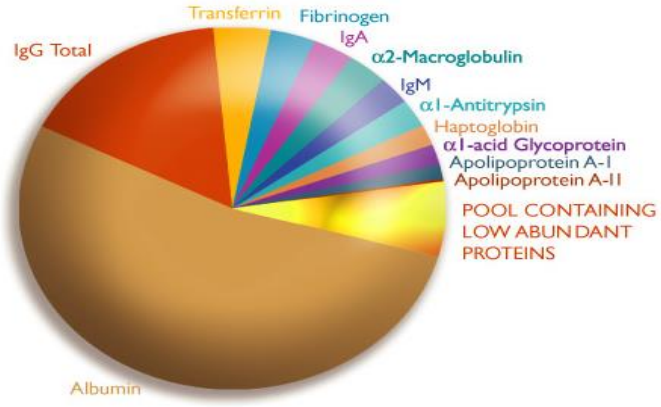


Figure 9.1. Ten proteins constitute ~ 90% of the plasma protein mass

The human plasma proteome constitutes a complex mixture of proteins derived from all tissues, which makes plasma an attractive medium for clinical analysis as a dynamic representation of the molecular states of diverse systems.[183] A wide range of proteins can thus be detected in plasma, including carrier proteins such as albumin, immune system effectors including immunoglobulins and complement factors, hemostatic factors, tissue messengers such as natriuretic peptides and interleukins, and tissue leakage products such as troponin and creatine kinase. This diversity in plasma protein function is accompanied by a diversity in protein abundance, with reference intervals for known plasma proteins in healthy subjects spanning >11 orders of magnitude.[183]

9.2. PROTEOMIC PROFILING FOR AF

Proteomic profiling enables systematic high-throughput analysis of proteins and may substantially accelerate novel biomarker discovery. Relatively unbiased proteomics approaches have the advantage of allowing simultaneous screening for large numbers of proteins involved in different biological pathways. Recently, 3 longitudinal cohort studies have reported proteomic profiling and the risk of new-onset AF. [12, 23, 184] The first study used a proximity extension assay (Olink Proseek Multiplex Cardiovascular 96 x 96 kit) to screen 92 proteins in 2 community-based cohorts of older adults in Sweden with a total of 271 incident AF cases in 1703 participants over a median follow-up of around 9 years.[23] They identified 7 proteins that were associated with incident AF after adjustment for age and sex. Two proteins, NT-proBNP and IL-6, remained significantly associated with incident AF after multivariable adjustment and Bonferroni correction.[23] The second cohort study used a community-based sample from Italy and focused on 75 inflammatory marker proteins identified from proximity extension assays (the Olink Proseek Multiplex CVD I 96 x 96 and the Proseek Multiple Inflammation I 96 x 96 kits).[12] There were 117 new AF cases among 880 participants during a 20-year follow-up. The Italian study reported the results of 75 inflammatory biomarkers including FGF-23, fatty acid binding protein 4, and IL-6, none of which were associated with AF after adjustment for age and sex.[12] The third study, from Framingham, used single-stranded DNA-based aptamers as affinity reagents (measured by the SOMAscan platform) to screen for 1373 proteins.[184] This study included 1885 participants with 349 incident AF cases during a mean follow-up of 18 years. In this study, Ko *et al.* identified 8 proteins associated with AF after adjustment for age and sex, and after further adjustment for AF risk factors, 2 proteins (ADAMTS13 and NT-proBNP) remained associated with new-onset AF.[184] The biological functions of the 8 proteins associated with AF after age and sex adjustment are listed in **Table 9.1**.

Table 9.1. Biological functions of the 8 proteins associated with the risk of incident AF in Framingham

Protein	Function
NCAM-120	Immunoglobulin-like glycoprotein. Activates fibroblast growth factor receptor and induces neurite outgrowth. Over-expression in neuroblastoma cells.
WFKN2 (WFIKKN2)	Multivalent protease-inhibitor. Inhibits growth differentiation factor and myostatin.
TrkC (Ntrk3)	One of tropomyosin receptor kinases that bind to neurotrophin-3, which in turn induces growth and differentiation of neuronal cells. Mutations in the gene associated with medulloblastoma, neuroblastoma, breast cancer, and other cancers.
EGFR (ERBB1)	Transmembrane glycoprotein kinase that acts as a receptor for epidermal growth factor. Mutations in the gene associated with different types of cancers.
ADAMTS13 (ATS13)	Multivalent protein that cleaves von Willebrand Factor. Mutations in the gene are associated with thrombotic thrombocytopenic purpura.
Angiopoietin-2	Inhibits angiopoietin-1 and endothelial TEK tyrosine kinase, thereby regulating angiogenesis and endothelial function.
NT-proBNP	Secreted by ventricular myocardium upon myocardial stretching and causes natriuresis, diuresis, vasodilation, and inhibition of the renin-angiotensin- aldosterone system.
BMPR1A	Transmembrane serine/threonine kinase receptor that binds to the members of the TGF-β superfamily. Mutations in the gene are associated with pulmonary arterial hypertension and hereditary hemorrhagic telangiectasia.

ADAMTS13: a disintegrin and metalloproteinase with thrombospondin motifs 13; BMPR1A: bone morphogenetic protein receptor type-1A; CI: confidence interval; EGFR: epidermal growth factor receptor; HR: hazard ratio; NCAM-120: Neural cell adhesion molecule 1, 120 kDa isoform; NT-proBNP: N-terminal pro-brain natriuretic peptide; TrkC: tropomyosin receptor kinase C; WFKN2: WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 2

NT-proBNP, a marker of ventricular remodeling, has been previously reported to be associated with incident AF by multiple prospective population-based studies.[18-23] ADAMTS13 is a von Willebrand factor protease, and its deficiency is found in thrombotic thrombocytopenic purpura. Previous case-control studies have shown that lower ADAMTS13 protein level was associated with chronic and paroxysmal AF.[185] Additionally, higher von Willebrand factor / ADAMTS13 ratio was significantly associated with chronic AF and left arterial remodeling,[185] and higher von Willebrand factor / ADAMTS13 ratio drawn 24 after cardioversion was associated with higher risk of AF recurrence.[186]

9.3. AN ASSAY FOR PLASMA PROTEOME

Proteomic array platforms have been developed to improve diagnostics for conditions with large unmet clinical needs, such as oncology, renal disease, and infections. Recently, a modified aptamer-based technology, SOMAscan, was developed by SomaLogic as a highly sensitive and multiplexed proteomics platform.[187-189] SOMAscan is based on Slow Off-rate Modified Aptamers (SOMAmers) that recognize specific conformational epitopes of natural 3D proteins with high sensitivity and specificity.[190] The SOMAscan assay is a hypothesis-free protein biomarker discovery tool that currently allows for the measurement of ~5,000 different proteins in an expedited fashion compared to previously existing technology, making it ideally suited for identifying clinically relevant biomarkers in a large number of samples.

A recent proteomic analysis using SomaScan among patients with stable coronary disease identified a protein-based risk score that outperformed traditional risk scores in predicting adverse cardiovascular outcomes.[191] Analysis of paired samples demonstrated that the protein-based risk score changed more than traditional CV risk markers among participants approaching new CV events.[191] In addition, the protein-based risk score generated using the follow-up sample was a stronger predictor of subsequent outcomes than the preceding baseline risk score. In addition to cardiovascular disease, validated protein biomarkers based on SOMAscan technology for chronological age, active pulmonary tuberculosis, Duchenne muscular dystrophy, and malignant pleural mesothelioma have also been reported.[192, 193] Whereas genomics are particularly useful for predicting lifelong risk, proteomics is a more dynamic approach to risk profiling that incorporates environmental and genetic influences. SOMAscan was capable of detecting dynamic changes in proteins over time as patients with latent tuberculosis approached conversion to active disease.[194]

10. THE PREDICTION OF ATRIAL FIBRILLATION

With the aging of the population, the incidence and prevalence of AF are expected to grow in future decades as well as the burden from its associated complications.[195] The growing public health significance of AF has spurred efforts to identify individuals at higher risk of developing this arrhythmia and its complications. Identifying individuals more likely to develop AF could facilitate targeting of preventive interventions and screening programs, while risk stratification schemes in AF patients can assist clinicians and patients in treatment decisions.

10.1. AVAILABLE MODELS FOR PREDICTION OF INCIDENT AF

Over the last few years, several risk scores and equations for the prediction of AF in the general population have been developed, published, and validated. **Table 10.1** enumerates in chronological order the published scores, the variables included, the characteristics of the derivation and validation samples, if any, and the performance of the model (discrimination and calibration). Discrimination refers to the ability of the model to separate subjects who develop the outcome from those who do not, while calibration refers to the agreement between observed outcomes and predictions.[196]

Table 10.1. Risk scores and equations for the prediction of atrial fibrillation in the community

Risk model	Variables	Derivation	Performance	Validation in external populations	Performance
FHS (10-year risk)[85]	Age, sex, body mass index, systolic blood pressure, treatment for hypertension, PR interval, cardiac murmur, heart failure	4764 participants, 100% white, 55% women, 45–95 years of age, mean age 61	C-statistic (95%CI): 0.78 (0.76–0.80) χ^2 = 4.2 (p = 0.09)	AGES:[197] 4238 participants, 100% white, 63% women, mean age 76	C-statistic (95%CI): 0.67 (0.64, 0.71) Recalibrated χ^2 = 16.2 (p = 0.06)
				CHS: [197] 5410 participants, 16% African-American, 84% white, 60% women, 65 and older, mean age 75	C-statistic (95%CI): 0.68 (0.66, 0.70) in whites, 0.66 (0.61, 0.71) in African Americans Recalibrated χ^2 = 46.1 (p < 0.001) in whites and 10.6 (p = 0.31) in African Americans
				ARIC:[51] 14,546 participants, 27% African-American, 73% white, 55% women, 45–64 years of age	C-statistic: 0.68 overall, 0.69 in whites, 0.65 in African Americans
				MESA:[198] 6663 participants, 38% white, 28% African-American, 22% Hispanic,12% Chinese-American, 53% women, 45–84 years of age, mean age 62	C-statistic (95%CI): 0.75 (0.72, 0.77) overall, 0.75 (0.72, 0.78) in whites, 0.74 (0.70, 0.78 in non-whites χ^2 = 57.4 (p < 0.001) overall, χ^2 = 8.1 (p = 0.53) in whites, χ^2 = 73.9 (p <0.001) in non-whites
ARIC (10-year risk) [51]	Age, race, height, smoking, systolic blood pressure, treatment for hypertension, cardiac murmur, ECG-based left ventricular hypertrophy, ECG-based left atrial enlargement, diabetes, coronary heart disease, heart failure	14,546 participants, 27% African-American, 73% white, 55% women, 45–64 years of age	C-statistic: 0.78 χ^2 = 10.0 (p = 0.35)	None	

WHS (10-year risk)[199]	Age, weight, height, systolic blood pressure, alcohol use, smoking	19,940 participants, 100% white, 100% women, median age 53	C-statistic (95%CI): 0.72 (0.68–0.75) χ^2 = 8.1 (p = 0.43)	None	
CHARGE-AF (5-year risk)[10]		18,556 participants, 81% white, 19% African-American,57% women, 46–94 years of age, mean age 65	C-statistic (95%CI): 0.77 (0.75–0.78) χ^2 = 9.3 (p = 0.41)	AGES: [10] 4469 participants, 100% white, 60% women, mean age 76	C-statistic (95%CI): 0.66 (0.63, 0.70) Recalibrated χ^2 = 12.6 (p = 0.18)
				Rotterdam Study: [10] 3203 participants, 100% white, 59% women, mean age 72	C-statistic (95%CI): 0.71 (0.66, 0.75) Recalibrated χ^2 = 16.4 (p = 0.06)
				EPIC-Norfolk:[200] 24,020 participants, >99% white, 55% women, 39–79 years of age, mean age 59	C-statistic (95%CI): 0.81 (0.75, 0.85) χ^2 = 142.2 (p < 0.001) Recalibrated χ^2 = 13.3 (p = 0.15)
				MESA[198] 6663 participants, 38% white, 28% African-American, 22% Hispanic, 12% Chinese-American, 53% women, 45–84 years of age, mean age 62	C-statistic (95%CI): 0.78 (0.74, 0.81) overall, 0.76 (0.72, 0.81) in whites, 0.78 (0.72, 0.83 in non-whites χ^2 = 25.6 (p = 0.002) overall, χ^2 = 14.6 (p = 0.10) in whites, χ^2 = 12.3 (p = 0.20) in non-whites
Ages: Age, Gene/Environment Susceptibility-Reykjavik Study; ARIC: Atherosclerosis Risk in Communities Study; CHARGE-AF: Cohorts for Aging and Research in Genomic Epidemiology—Atrial Fibrillation; CHS: Cardiovascular Health Study; CI: Confidence interval; EPIC: European Prospective Investigation into Cancer and Nutrition; FHS: Framingham Heart Study; MESA: Multi-Ethnic Study of Atherosclerosis; WHS: Women's Health Study					

The first published risk score was derived in 4764 mostly white participants in the FHS, and used basic demographic and clinical variables to predict the 10-year risk of AF.5 The discrimination of the model, assessed with the C- statistic, was good (0.78, 95% confidence interval (CI) 0.76, 0.80). This score was subsequently validated in four different cohorts: the Age, Gene/Environment Susceptibility- Reykjavik (AGES) study, the ARIC study, CHS, and the Multi-Ethnic Study of Atherosclerosis (MESA).[51, 197, 198] In these external cohorts, the discrimination of the model was acceptable, ranging from 0.67 in African American participants in CHS to 0.75 in the racially diverse MESA cohort. In most populations, however, the model required recalibration to adjust the predicted probabilities to the actual risk of AF in the different cohorts. Independently, the ARIC study also developed a 10-year risk score for AF prediction among 14,546 study participants 45–64 years of age.[51] In contrast to the FHS AF risk score, the ARIC model was based on a bi-racial cohort, including whites and African Americans. Given the well- established lower risk of AF among non-whites compared to whites,[9, 52] attention to race in AF prediction is relevant and the application of scores developed in a specific racial/ethnic group to another should be done carefully. The discrimination of the ARIC model was similar to the FHS AF risk score (C-statistic 0.78). The ARIC model, however, has not been applied in any external cohorts and, therefore, its validity outside the ARIC population is uncertain. More recently, the Women's Health Study (WHS), a cohort of mostly white, healthy women, derived and validated a 10-year model among 19,940 participants.[199] The model had good discrimination (C-statistic 0.72) and excellent calibration in the WHS cohort, but has not been validated in external populations and its applicability to men is unknown.

The FHS, ARIC, and WHS risk scores and predictive models were derived in single cohorts, and restricted in terms of race/ethnicity (FHS), age (ARIC), or sex (WHS), which may reduce generalizability to other populations. To address this limitation, the Cohorts for Aging and Research in Genomic Epidemiology (CHARGE)-AF consortium derived a new predictive model pooling data from 18,556 participants in the FHS, CHS and ARIC studies to predict the 5-year risk of AF. This model was then validated in 7,672 participants from the AGES and Rotterdam studies, showing acceptable discrimination.[10] The CHARGE-AF model, which included demographic and clinical information readily available in clinical settings, had good discrimination in the derivation cohorts (C-statistic 0.77, 95%CI 0.75, 0.78) and acceptable in AGES (0.66, 95% 0.63, 0.70) and the Rotterdam study (0.71, 95% 0.66, 0.75).[10] The CHARGE-AF risk model has been validated in two additional cohorts. In the European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk study, the model had excellent discrimination (C-statistic 0.81, 95% 0.75, 0.85) but it overestimated the

risk of AF, requiring recalibration.[200] Similarly, the CHARGE-AF model had good discrimination in the MESA cohort (C-statistic 0.78, 95%CI 0.74, 0.81), but also overestimated AF risk, particularly among those with the highest observed risk.

Finally, some studies have suggested that scores derived for prediction of stroke in patients with AF, such as the CHADS2 and CHA2DS2-VASc, could be also applied in AF prediction.[201, 202] In fact, most of the elements included in these scores (age, diabetes, hypertension, heart failure, vascular disease) are well established risk factors for AF. However, these scores perform worse than AF-specific predictive models (as assessed with c-statistics, for example) and are not adequately calibrated to predict AF, failing to provide estimates of actual predicted AF risk over a particular time period.[198]

10.2. AF PREDICTION BEYOND CLINICAL VARIABLES

Extensions of these models have evaluated whether information on blood biomarkers, echocardiographic and ECG measurements, or genetic variants would improve prediction of AF beyond the information provided by clinical variables.

Blood biomarkers

The predictive value of a diverse array of circulating biomarkers, including markers of inflammation (high-sensitivity C-reactive protein, fibrinogen),[18, 81, 82] atrial overload (atrial and B-type natriuretic peptides),[18, 81, 82] myocardial ischemia (high- sensitivity troponin T and I),[11, 16, 81] cardiac fibrosis (galectin-3),[83, 84] and others (soluble ST2, growth differentiation factor-15),[11] has been assessed in the literature. Of these, only natriuretic peptides have consistently demonstrated added predictive value beyond information on clinical variables across multiple populations. For instance, in the CHARGE- AF pooled analysis, which included five separate cohorts, B-type natriuretic peptides but not C-reactive protein helped in risk reclassification of individuals, as measured by the net reclassification index (NRI).[18] Similar observations have been made in the FHS,[19] the Malmö Diet and Cancer Study,[82] the MESA cohort,[198] and in the Gutenberg Health Study.[81]

Electrocardiography

Some of the existing AF risk scores and models include ECG- derived variables, such as the PR interval in the FHS AF score, the ARIC score, and the CHARGE-AF model, or ECG-based left ventricular hypertrophy in the ARIC score and the CHARGE-AF model.[10, 51, 85] Their added predictive value beyond clinical variables, however, is only marginal. Other

ECG measurements considered as potential predictors of AF include P wave indices. A pooled analysis of the FHS and ARIC cohorts found that even though P wave indices such as P wave duration, area, and terminal force were associated with the incidence of AF, their contribution to risk prediction on top of established risk factors was minimal.[88] However, a recent analysis adding abnormal P wave axis to the CHARGE-AF score modestly improved the C-statistic from 0.719 to 0.722 in ARIC.[89] Information on atrial ectopy assessed through longer term heart rhythm monitoring could also improve AF prediction. An analysis of 1260 participants in the CHS cohort found that information on premature atrial contractions count from 24-hour Holter monitoring led to clinically significant improvements in AF prediction beyond the information provided by the FHS AF score (C-statistic of 0.65 in the FHS AF score alone vs 0.72 after adding atrial ectopy information to the statistical model).[90]

Imaging

Information on cardiac structure and function obtained from echocardiographic studies, such as left atrial diameter, left ventricular function, left ventricular mass, or left ventricular wall thickness, have not demonstrated benefit in the prediction of AF once demographic and clinical information is considered.[85, 91] Whether more novel measures of echocardiography-based left atrial function (e.g. left atrial strain by speckle tracking, tissue Doppler imaging-derived atrial conduction time)[203, 204] or other cardiac imaging modalities (e.g. periatrial epicardial adipose tissue from computerized tomography)[205] can be used for AF prediction remains to be determined.

Genetics

Recent research has identified several common genetic variants associated with the risk of AF.[95] The added value of information on these genetic variants to predict AF has been explored in at least two different populations. The WHS cohort found that a genetic risk score, calculated with information on 12 single nucleotide polymorphisms previously associated with AF, significantly improved prediction, measured with change in C-statistic and continuous NRI, beyond a clinical risk score in approximately 20,000 women: the C-statistic increased from 0.72 to 0.74, while the continuous NRI was 0.49 (95%CI 0.30–0.67).[199] In a similar analysis among 27,471 participants of the Malmö Diet and Cancer Study, however, a genetic risk score also based on 12 single nucleotide polymorphisms only minimally improved risk prediction (C-statistic changed from 0.735 to 0.738).[206] Notably, none of these analyses considered information on natriuretic peptides, which are possibly the strongest biomarkers for AF risk. Future studies should evaluate whether genetic information

improves our ability to predict AF on top of clinical variables and established AF circulating biomarkers.

10.3. APPLICATIONS OF MODELS FOR AF PREDICTION

Available risk scores, though imperfect, may play a role in identifying individuals at higher risk of developing AF, particularly the externally validated FHS and CHARGE-AF models. A follow-up question is whether this information has any clinical or public health implications. We think of two major areas in which these scores could be useful: as aids for selection of high-risk participants to screening programs and primary prevention trials, and as benchmarks for the testing of potential novel biomarkers of AF risk.

The interest in developing screening programs for identification of asymptomatic AF is growing.[207] AF is responsible for a substantial proportion of strokes, and in a number of cases, stroke is the first clinical manifestation of AF.[208] Identifying individuals with asymptomatic AF offers a unique preventive opportunity if AF diagnosis is followed by adequate antithrombotic therapy. Restricting screening programs to individuals more likely to have AF—as identified by one of the validated risk scores—would make those programs more cost-effective. A similar rationale can be applied to the selection of participants for primary prevention trials of AF. Currently, there are no established interventions for the primary prevention of AF. Trials testing such interventions will have to be conducted in subgroups at higher risk of AF, which will lead to more efficient designs.

Validated risk scores, particularly those including circulating natriuretic peptides as predictors, can also be used as benchmarks against which novel biomarkers purported to improve AF prediction can be compared. In this era of "precision medicine," rigorous comparisons with extensively validated risk scores are needed to avoid the hype that frequently surrounds the discovery of novel markers of disease. For example, as summarized above, adequate testing against a model including natriuretic peptides (BNP or NT-proBNP) showed that inflammatory markers such as CRP, despite being associated with increased risk of AF in observational studies, are not particularly useful in AF prediction.[18, 198]

11. STUDY DESIGNS AND DATA COLLECTION

Manuscript 1 and 2 utilized data from the ARIC study to study the relation of proteomics and incident AF and then to also develop a risk prediction score for incident AF in an elderly population. Manuscripts 3 and 4 used Medicare data to examine treatment and outcomes in rural versus urban AF patients

11.1. THE ARIC STUDY

The ARIC study is a prospective epidemiologic study of CVD conducted in four US communities: Forsyth County, NC; the city of Jackson, MS; eight northwestern suburbs of Minneapolis, MN; and Washington County, MD. The ARIC study has both cohort and community surveillance components and was designed to investigate the etiology and natural history of atherosclerosis, the etiology of clinical atherosclerotic disease and differences in cardiovascular risk factors, medical care and disease by race, sex, location and date.[209] The cohort component of the ARIC Study was designed to identify characteristics associated with the development and progression of clinical atherosclerosis as measured by carotid B-mode ultrasonography, to identify risk factors associated with incident atherosclerotic events and to measure variation in risk factors over time.[210]

Study design and population

The ARIC Study recruited a prospective cohort of mainly white and black adults between 45 and 64 years of age at baseline, 1987 – 1989. Approximately 4,000 participants were selected from each of the 4 communities using community specific probability sampling; households were identified by area sampling in Forsyth County, NC, driver's licenses or state identification cards were used in Jackson, MS, eligibility for jury duty (with driver's license, voter registration cards or identification cards) were applied in Minneapolis, MN, and driver's licenses or inclusion in a 1975 private county health census were utilized in Washington County, MD. Regardless of the community, all age-eligible residents of an identified household were selected as potential participants. Only blacks were recruited from the city of Jackson, MS; the other sites included both whites and blacks although < 5% of the population in Minneapolis, MN, and Washington County, MD, were black. A total of 15,792 participants enrolled at baseline (8710 women, 4314 blacks). Participants had a clinical exam at baseline and the following visits have been completed thus far: visit 2 (1990-92), visit 3 (1993-95), visit 4 (1996-1998), visit 5 (2011-2013) and visit 6 (2016-2017). Annual telephone calls are used to maintain contact with participants and identify medical events and death

throughout follow-up, and these calls have been conducted semi-annually since 2012. Follow-up is currently complete through December 31, 2018.

Data Collection

Between 1987 and 1989 baseline data were collected; the baseline exam consisted of a home interview comprised of questionnaires about cardiovascular risk factors, socioeconomic status, and family medical history as well as a clinical examination. The baseline clinical exam and each of the follow-up exams consisted of anthropometry, sitting blood pressure, venipuncture, ECG, ultrasound, physical exam and interviewer-administered questionnaires on medical history, health behaviors (alcohol and tobacco use) and social characteristics. Additional data were collected at some exams; for example, certain biomarkers were measured at select visits (NT-proBNP was measured at visits 2, 4, 5, 6). Annual telephone calls continue to maintain contact with participants and to identify any cardiovascular events, hospitalizations and death. Each center's institution review board approved the study and all participants provided written informed consent.[209] For the purposes of this dissertation, the baseline visit for manuscripts 2 and 3 is visit 5 (2011-2013). Proteomics data was measured using visit 5 samples from the entire cohort. Risk factors and covariates were also measured at visit 5.

Ascertainment of AF

Utilizing ARIC data, prevalent AF was identified by baseline ECG. Incident AF was identified by ECG during follow-up study visits, hospital discharge codes and death certificates.[9] Standard supine 12-lead resting ECGs were recorded at least one hour after consumption of caffeine or tobacco and transmitted to the ARIC ECG Reading Center for coding and interpretation. The baseline ECG had a two-minute rhythm strip and subsequent ECGs had a 10-second reading. ECGs automatically coded as AF were visually checked by a trained cardiologist to confirm the diagnosis.[211] All ECGs were recorded using MAC PC Personal Cardiographs (Marquette Electronics, Inc., Milwaukee, WI). Hospitalizations were identified by annual telephone calls to participants and through surveillance of local hospital discharges in each of the ARIC communities.[209, 210] A hospital discharge code, ICD-9-CM code of 427.3, 427.31 or 427.32, in any position, indicates AF. Starting in 2015, an ICD-10-CM code of I48 was used to ascertain AF. AF hospitalization diagnoses occurring simultaneously with heart revascularization surgery or other cardiac surgery involving heart valves or septa, without evidence of AF in subsequent hospitalizations or study examinations were excluded. AF was identified though death certificates with an ICD-10 code I48 or ICD-9

code 427.3x as the underlying cause of death. The AF incidence date was defined as the first documented occurrence of AF on ECG, hospital discharge diagnosis or death certificate.

In ARIC, two analyses were performed to determine the validity of the diagnosis of incident AF based on hospital discharge diagnosis codes.[9] First, a sample of 125 hospital discharge summaries with a first ICD-9 code for AF and ECGs performed during that hospitalization were reviewed by a study physician; the positive predictive value (PPV) for AF was 89% and for incident AF was 62%.[9] Second, a trained abstractor used information routinely collected for stroke ascertainment to complete a form with data from the complete medical record. The form includes information on the presence of AF during four weeks prior to the stroke hospitalization. Of 161 participants with AF recorded in the stroke abstraction form, 135 had an ICD code for AF (sensitivity = 84%) and of 1385 participants without AF in the abstraction form, 34 had an ICD code for AF (specificity = 98%).[9] The sensitivity of using hospital discharge codes to identify AF was similar in CHS; hospital discharge diagnoses codes (ICD-9 code of 427.3x) correctly identified 29 (70.7%) of the 41 participants with AF or AFL on at least one ECG.[44] A systematic review of algorithms used in administrative data to identify AF patients reported a median PPV of 89% (range: 70% - 96%) and a median sensitivity of 79% (range: 57% - 95%).

11.2. MEDICARE

Medicare is a health insurance program for 1) people age 65 or older, 2) people under the age of 65 with certain disabilities and 3) people of all ages with End-Stage Renal Disease. Available plans in Medicare include Part A, which is hospital insurance, Part B, which is medical insurance, and Part D, which is prescription drug coverage. The US Centers for Medicare and Medicaid Services (CMS) compiles the Medicare datasets and creates standardized datasets of a 5% random sample, a 20% random sample, and the 100% sample.

Data Collection

We obtained research identifiable claims data for a nationally representative 20% sample of Medicare beneficiaries from 2011-2016.The data include inpatient, outpatient, and carrier files from January 1, 2011 to December 31, 2016. The inpatient files contained institutional claims for inpatient services covered under Medicare Part A. The outpatient files contained institutional claims for outpatient services covered under Medicare Part B. The carrier files contained noninstitutional physician claims for services covered under Medicare Part B. All of the files contained discharge or service dates and International Classification of Diseases, Clinical Modification, 9th and 10 edition (ICD-9-CM and ICD-10-CM) diagnosis

codes. Denominator files contained the beneficiary identifier, date of birth, sex, race, date of death (if applicable), ZIP codes, concurrent enrollment in Medicaid (a proxy for low economic status), and information about program eligibility and enrollment. Additionally, the Part D Drug Event and Characteristics files were used to assess anticoagulation prescription fills, along with cardiovascular and other medication use. These files contain information on drug name, therapeutic class, prescription fill date, dose, and number of days supplied.

Medicare data contains a ZIP code for the beneficiary, which is not provided in most other claims databases. This allowed us to assess individual location and rurality. Other claims databases often provide only a hospital location or a variable for a Metropolitan Statistical Area, which are typically larger areas than ZIP codes.

Ascertainment of AF

This analysis included patients age 65+ with at least one inpatient claim for AF or 2 outpatient claims for AF 7 to 365 days apart. AF claims were identified using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis codes 427.3, 427.31, and 427.32, and ICD-10-CM codes starting October 1, 2015 of I48.x in any position, which is a standard definition used in claims analysis.[45, 212] The validity of ICD-9-CM codes for the identification of AF has been well-established with a systematic review of studies showing a positive predicted value (PPV) of approximately 90% and a sensitivity of approximately 80%.[213] We defined the diagnosis date as the earlier of 1) the earliest discharge date for an inpatient claims, or 2) the earliest service date of the outpatient or physician claim. Consistent with prior research, 2 outpatient claims were required to diagnose outpatient AF in order to minimize the impact of rule-out diagnosis and to improve specificity.[45]

Strengths and limitations of Medicare data

Within the Medicare database, we will utilize variables from hospitalization records, outpatient visits, pharmacy prescription fills, and basic information on each beneficiary, including ZIP code. Additional strengths of the Medicare dataset are the large number of beneficiaries from all areas of the US, the availability of important health variables, and key characteristics of the beneficiaries.

Limitations of using the Medicare data include that we are including only individuals age 65 and older, and we are also limiting the sample to those with exclusive stand-alone coverage so we can be sure we are capturing their medical events and prescription fills. Requiring stand-alone Part D enrollment reduces our sample size and also limits the dataset to

patients who, in general, have more comorbidities and a lower socioeconomic status compared to the entire sample.[178] Additionally, we were only able to capture prescriptions that have been filled; we have no way to track prescribed drugs that are not filled. Finally, claims datasets have inherent limitations given they are created for billing purposes, and therefore may not capture all the characteristics and intricacies of a patient's health status.

12. Manuscript 1 - Proteomics and the Risk of Incident Atrial Fibrillation in Older Adults: The Atherosclerosis Risk in Communities (ARIC) Study

12.1. OVERVIEW

Background – Plasma proteomic profiling may aid in the discovery of novel biomarkers upstream of the development of atrial fibrillation (AF) and has the potential to advance our understanding of disease mechanisms. Prior studies relating proteomic markers to incident AF have included limited numbers of proteins. We used data from the Atherosclerosis Risk in Communities (ARIC) study to examine the relationship between large-scale proteomics and incident AF in a cohort of older-aged black and white adults in the US.

Methods –We quantified 4877 plasma proteins in ARIC participants at visit 5 (2011-2013) using an aptamer-based proteomic profiling platform. We used Cox proportional hazards models to assess the association between protein levels and incident AF and explored relationships of selected protein biomarkers using annotated pathway analysis.

Results – Our study included 4668 AF-free participants (mean age 75 ± 5 years; 59% female; 20% black race) with proteomic measures. A total of 585 participants developed AF over a mean follow-up of 5.7 ± 1.7 years. After adjustment for clinical factors associated with AF, N-terminal pro-B-type natriuretic peptide (NT-proBNP) was associated with the risk of incident AF (hazard ratio, 1.82; 95% CI, 1.68-1.98; p-value=2.91 X 10^{-45}). In addition, 36 other proteins were also significantly associated with incident AF after Bonferroni correction. We further adjusted for medication use and estimated glomerular filtration rate and found 17 proteins, including Angiopoietin-2 and Transgelin, remained significantly associated with incident AF. Pathway analyses implicated the inhibition of matrix metalloproteases as the top canonical pathway in AF pathogenesis.

Conclusion – Using a large-scale proteomic platform we identified both novel and established proteins associated with incident AF, and explored mechanistic pathways of AF development

12.2. INTRODUCTION

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia, with a lifetime risk of 1 in 3 among whites and 1 in 5 among African Americans.[1] The risk of AF is higher for those with advancing age, European ancestry, cigarette smoking, taller height, greater weight, higher blood pressure and corresponding blood pressure medication use, diabetes, history of myocardial infarction, and history of heart failure.[9, 10] In addition to the traditional clinical risk factors listed above, various biomarkers have been identified as risk factors for incident AF including markers of inflammation,[11-14] oxidative stress,[15] myocardial necrosis,[11, 16, 17] myocardial stress,[11, 18-23] and mineral metabolism.[24, 25] Identification of novel biomarkers can advance our understanding of AF mechanisms, enhance opportunities for risk prediction, and may provide targeted preventive strategies for AF.

Proteomic profiling enables systematic high-throughput analysis of proteins and may aid in the discovery of novel biomarkers that are upstream of the development of AF. Proteomics approaches are relatively unbiased and have the advantage of allowing simultaneous screening for large numbers of proteins involved in different biological pathways. Recently, several longitudinal cohort studies have reported associations between plasma proteomic profiling and the risk of AF. [12, 23, 184, 214, 215] *Appendix Table 1* lists an overview of each study along with the main results. Four of the studies measured N-terminal pro-B-type natriuretic peptide (NT-proBNP) in their proteomic platform, and in all 4 studies, higher NT-proBNP was significantly associated with greater incidence of AF, even after adjustment for multiple AF risk factors. However, similarities in the results end there as each study found several different proteins associated with incident AF. These prior studies are limited by modest AF events and power, and by limited numbers of proteins included on their proteomic platforms. The only prior study that assessed a panel with >100 proteins had <1,400 participants.[184]

In this study, we used data from the Atherosclerosis Risk in Communities (ARIC) study to screen for 4877 plasma proteins and identify novel biomarkers that are associated with risk of incident AF. This community-based cohort of black and white older adults in the US has a larger number of proteins measured compared to previous studies, and nearly 600 AF events in a 6-year follow-up time, allowing us to address some limitations of previous studies.

12.3. METHODS

Study population

The Atherosclerosis Risk in Communities (ARIC) study is a prospective cohort study of cardiovascular disease and atherosclerosis risk factors.[209] Participants at baseline (1987-1989)

included 15,792 black and white men and women aged 45-64, recruited from 4 communities in the US (Washington County, Maryland; the northwest suburbs of Minneapolis, Minnesota; Jackson, Mississippi; and Forsyth County, North Carolina). Thus far, 7 study visits have been completed with visit 5 (baseline for our main analysis) occurring in 2011-2013. Additionally, ARIC participants have received annual follow-up calls (semi-annual after 2012), with response rates of ≥ 90% among survivors. The primary analysis examined the association of ARIC visit 5 protein levels with incident AF through the end of 2017 at the Jackson field center, and through the end of 2018 at the other 3 field centers. Among the 6538 participants who attended visit 5, we excluded those with prevalent AF at visit 5 (n=638), with missing (n=1170) or low quality proteomic data (n=15), with race other than white or black and non-whites in the Minneapolis and Washington County field centers (due to low numbers; n=42), having missing covariates (n=5), resulting in a study population of 4668. We also conducted a midlife replication analysis including only those proteins significantly associated with AF risk in the visit 5 primary analyses. We examined the association of proteins measured at visit 3 (1993-1995) with incident AF through the end of 2010, which was the approximate start of visit 5. After similar exclusions, 10,908 AF-free participants with protein measures at visit 3 were included in the midlife replication analysis. This study was approved by institutional review boards at each participating center, and all study participants provided written informed consent.

Ascertainment of AF

Incident AF was defined as in previous ARIC analyses.[9] A trained abstractor obtained and recorded all International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) and ICD-10-CM hospital discharge diagnoses from each participant's hospitalizations reported in the follow-up interview. AF was defined as the presence of ICD-9-CM code 427.31 or 427.32 or ICD-10-CM code I48.xx. AF hospitalization diagnoses occurring simultaneously with heart revascularization surgery or other cardiac surgery involving heart valves or septa were not included as AF events. Deceased ARIC participants were also labeled as AF cases if their underlying cause of death was AF. AF was additionally identified by study visit ECGs, performed at visits 1-5. At each ARIC study visit, a 10-second 12-lead ECG was performed using a MAC PC cardiograph (Marquette Electronics Inc, Milwaukee, WI) and transmitted to the ARIC ECG Reading Center for coding, interpretation and storage. All ECGs automatically coded as AF were visually checked by a trained cardiologist to confirm AF diagnosis.[211]

Proteomics Profiling

EDTA-plasma was obtained from blood samples that were collected during visits 3 and 5 and stored at -80 degrees C. Plasma samples were analyzed using a SOMAmer-based capture array called "SOMAscan" (Somalogic, Inc., Boulder, CO, USA). This assay was performed as described previously.[216-219] Protein levels in the plasma samples were measured by the SOMAscan platform, which uses single-stranded DNA-based aptamers to capture conformational protein epitopes. Additional information on quality control can be found in the Supplemental Methods. After all quality control measures were completed, 4877 aptamers which recognize 4697 unique human proteins or protein complexes were analyzed in this study. We examined protein distributions and applied log base 2 transformation to all SOMAmer measures to correct for skewness. We winsorized outliers that were greater or less than 5 standard deviations from the sample mean on the log 2 scale.

Covariates

Covariates for this analysis include AF risk factors from the CHARGE-AF score,[10] namely age, sex, race, cigarette smoking status, height, weight, systolic and diastolic blood pressure, anti-hypertensive medication use, diabetes, prevalent myocardial infarction, and prevalent heart failure. We additionally included several medications and estimated glomerular filtration rate (eGFR) as covariates, reasoning that, in addition to being associated with the risk of atrial fibrillation, these variables could also affect protein levels. Covariates measured at visit 5 were used in the main analysis and those measured at visit 3 were used in the midlife replication analysis. Detailed procedures for covariate measures have been published, [209] and further details can be found in the Supplemental Materials.

Statistical analysis

Baseline characteristics were described as mean (SD) for continuous covariates and counts (%) for dichotomous covariates. Our primary analysis used Cox proportional hazards regression models to relate each log base-2 protein level to incident AF (censored at the last follow-up time, death, or the end of 2017 / 2018). We used a series of models to examine the associations and to compare results with other cohorts who have made similar adjustments. A minimally adjusted model 0 accounted for age, sex, and race/center and provided comparisons with previous cohorts' results. Model 1 consisted of previously reported AF risk factors[10] and adjusted for age, sex, and race/center, current cigarette smoking, height, weight, systolic and diastolic blood pressure, the use of hypertension medications, diabetes, prevalent myocardial infarction and prevalent heart failure. Model 2 additionally adjusted for the confounders of eGFR, anticoagulant use, beta blocker use, and antiarrhythmic (Class I and III) medication use. We explored liver disease, participant fasting status, and use of statins, cardiac glycosides, calcium or channel blockers as possible confounding variables and

deemed that they were not confounders and did not include them in the final models. Bonferroni correction was used to correct for multiple tests; we considered $P<0.05 / 4877 = 1.025 \times 10^{-5}$ to be statistically significant.

We performed additional analyses on the 40 proteins that reached statistical significance in either model 1 or model 2. We explored interactions by age, sex and race using a multiplicative term in model 2. We additionally adjusted for NT-proBNP to determine the association of protein levels with incident AF, independent of the level of NT-proBNP. We assessed the proportional hazards assumption in the top 40 proteins with scaled Schoenfeld residuals using both graphical and numerical tests and found no evidence of modeling violations.

In the midlife replication analysis, we used ARIC visit 3 as baseline (1993-95) and examined the association of the 40 proteins with the risk of incident AF through the end of 2010, which was approximately the start of visit 5. We applied the same exclusion criteria as for the visit 5 analysis and used covariates measured at visit 3. For all of these analyses using the top 40 proteins, Bonferroni correction was used to account for multiple tests; we considered $P<0.05 / 40 = 1.25 \times 10^{-3}$ to be statistically significant. We performed statistical analyses using SAS v 9.4 (SAS Inc, Cary, NC).

Ingenuity Pathway Analysis

We performed network pathway analysis to 1) further explore biological mechanisms connected to the proteins associated with incident AF and 2) to identify factors upstream to AF. We analyzed data using of Ingenuity Pathway Analysis (IPA) (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis).[220] We uploaded a dataset containing the protein identifiers, hazard ratios from our primary analyses using a fully adjusted model (model 2), and corresponding p-values, to identify novel mechanisms outside of the well-known associations between our covariates and AF. We then restricted the analysis to the proteins associated with incident AF at a false discovery rate (FDR) corrected threshold of $P < 0.05$, resulting in 60 SOMAmers. Of these, 56 were successfully mapped to genes in the Ingenuity Pathways Knowledge Base; in some cases duplicated SOMAmers mapped to a single gene (e.g., SVEP1a and SVEP1b) and in other cases, more than one gene product corresponded to a single gene ID (e.g., NT-proBNP and natriuretic peptide B). In the case of duplicates, the maximum expression value of the two SOMAmers was used in the analysis.

Further details regarding IPA can be found in the Supplemental Materials. In brief, we used IPA Core Analysis to estimate the degree to which specific canonical pathways, protein networks, and upstream regulators were implicated based on the set of proteins found to be associated with AF risk. For all of the IPA analyses, only statistically significant canonical pathways, physiological systems, upstream regulators, and causal networks are reported, and only a subset are provided in our results.

12.4. RESULTS

A total of 4,688 participants with protein level measured at visit 5 were included in the main analysis (mean age = 75 ± 5 years; 59% female; 20% black race). A total of 585 (13%) participants developed incident AF during a mean (SD) follow-up time of 5.7 (2) years. Descriptive characteristics are provided in **Table 12.1** based on incident AF status. Those who developed AF were older, more likely to be male and white, and had a worse cardiovascular profile compared to those who did not develop AF.

Association of Protein Levels with Incident AF

After adjustment for age, sex, and race/center, 126 protein were significantly associated ($p<1.025 \times 10^{-5}$) with incident AF as listed in *Appendix Table 2.* After adjustment for variables included in the CHARGE-AF risk score (model 1), and further adjustment for eGFR and medication use (model 2) 37 and 17 proteins, respectively, remained significantly associated with incident AF. These proteins are listed in **Table 12.2** and ordered by the p-value (from smallest to largest) of Model 2 with p-values $<1.025 \times 10^{-5}$ considered significant. After multivariable adjustment, NT-proBNP had the most significant association; for each doubling of the protein measure, the risk of AF was 1.75 times higher (95% CI = 1.60-1.91). Transgelin had the strongest effect size in regards to the risk of incident AF; for every doubling of the protein level, the risk of AF was 2.01 times higher (95% CI = 1.56-2.59). Several proteins were inversely associated with incident AF including Protein delta homolog 1 (DLK1) and ATS 13 (ADAMTS13). Protein SET had the strongest inverse effect size; for every doubling of Protein SET the risk of AF decreased by approximately 55% (HR=0.45, 95% CI = 0.28-0.71). Two of the top proteins, SVEP1 and DLK1, are listed twice due to distinct aptamers binding to the same protein. The top 100 proteins associated with incident AF after adjustment for model 2, along with the FDR p-values are presented in *Appendix Table 3.*

We examined interactions by age, sex, and race in the 40 proteins listed in Table 2 and we did not find any statistically significant interactions. We additionally adjusted for NT-proBNP to determine the association of protein levels with incident AF independent of NT-

proBNP, and results for the main 40 proteins are listed in *Appendix Table 4*. Eight of the protein remained significantly associated with incident AF and include CMRF35-like molecule 2 (CD300E), Growth/differentiation factor 11/8 (GDF11 MSTN), DLK1 (2 aptamers), Antileukoproteinase (SLPI), Cartilage intermediate layer protein 2 (CILP2), Scavenger receptor class F member 1 (SCARF1), and Gamma-aminobutyric acid receptor-associated protein-like 1 (GABARAPL1).

We ran a secondary analysis as an internal validation with 10,908 AF-free participants with protein measures at visit 3 and followed them until the end of 2010. At this visit, participants were younger with fewer comorbidities and on fewer medications (mean age = 60 ± 6 years; 55% female; 21% black race). A total of 1397 (13%) participants developed incident AF during a mean (SD) follow-up time of 13.9 (4) years. Of the top 40 proteins from the main analysis, 21 were significantly associated with incident AF (**Table 12.3)** in mid-life replication model 1, and 17 remained significant after adjustment for factors in model 2. NT-proBNP, SVEP1, Natriuretic peptides B, Transgelin, and Angiopoietin-2 were the proteins most strongly associated with incident AF in both mid-life and later-life. **Figure 12.1** depicts the beta estimates from model 2 for the top 40 proteins measured at mid-life (visit 3) plotted against those measured in later life (visit 5) for the association with incident AF. Several proteins maintained relatively consistent effect sizes at both visits, including CILP2, IGFBP-2, and Angiopoietin-2, among others.

Associations detected using IPA

Proteins associated with AF in late-life model 2 with an FDR P value <0.05 (listed in *Appendix Table 3*) were brought into the IPA environment. Of those, 56 proteins were mapped into 9 main networks and ordered from 1 to 9 by strength of association. Networks 1, 2, 3 and 5 were considered connected networks. **Figure 12.2** depicts the top 2 networks. Network 1 is centered around MMP-2, which was an upregulated protein in our analysis. Network 2 was centered around Protein Kinase B (PKB), more commonly referred to as Akt, and is a serine/threonine-specific protein kinase that plays a key role in multiple cellular processes such as glucose metabolism, apoptosis, transcription, and cell migration. A full list of molecules included in each network, along with associated diseases and functions, is listed in *Appendix Table 5*.

IPA identified canonical pathways, which are well-characterized metabolic and cell-signaling pathways, using known associations of our uploaded proteins. The 10 canonical pathways that were most significantly associated with our proteins are listed in *Appendix Table 6.* The top canonical pathway was the inhibition of matrix metalloproteases (MMPs), followed by axonal guidance signaling, and factors promoting cardiogenesis. **Figure 12.3**

depicts proteins that were up or down activated in the MMP pathway. Downstream diseases and functions identified centered around the common themes of an inhibited inflammatory response, cardiac dysfunction, kidney failure, and cell movement of cancer cells.

To identify regulators upstream of our proteins we utilized 2 IPA analyses. Upstream regulator analysis identifies molecules upstream of the proteins in the dataset that potentially explain the observed expression changes. IPA predicts which upstream regulators are activated or inhibited to explain the upregulated and down-regulated proteins observed in our dataset. Our top 20 top-identified upstream regulators are listed in *Appendix Table 7*. **Figure 12.4** depicts the mechanistic networks that are associated with the top 2 identified upstream regulators, and links the upstream regulator to our observed proteins via the intermediary molecules depicted in the figure. PTEN (phosphatase and tensin homolog) was predicted to be significantly inhibited based on the observed protein expressions in our data. PTEN inhibition was connected to 6 upregulated proteins including ANGPT2, MMP-2, NPPA, and a downregulated BMP-1 through intermediary pathways of activated ERK ½, STAT3, and IGF1, among others. P38 MAPK was predicted to be our strongest activated upstream regulator. P38 MAPK activates tumor necrosis factor (TNF) and ERK ½, leading to the upregulation of MMP-2, NPPA, and TIMP-2, and downregulation of FAS-associated death domain protein (FADD).

Next, we implemented the causal analysis algorithms which are based on a "master" network which is derived from the Ingenuity Knowledge Base. The causal algorithm identified our potential top master regulators as the GATA group (involved in thrombin signaling) and MEF2C (myocyte enhancer factor 2C; plays a role in myogenesis) and they both have the same activation z-score. MEF2C contains many GATA group members, and therefore we chose to present a figure using MEF2C as the master regulator**. Figure 12.5** depicts the relationships between MEF2C and the 19 proteins in our analysis that can be connected to MEF2C through intermediary regulators. The top 10 hypothesized master regulators based on the activation z-score are listed in *Appendix Table 8*, which depicts connected proteins through hypothesized participating regulators.

Finally, we used the regulator analysis function in IPA to determine potential pathways between upstream regulators, our measured proteins, and downstream diseases and functions. **Figure 12.6** depicts our top identified regulator network, which links 5 activated upstream regulators to 3 downstream disease functions that consist of an activated innate immune response, an inhibited accumulation of leukocytes, and inhibited death of ovarian cancer cell lines.

12.5. DISCUSSION

In this community-based prospective population study of older adults, we tested 4,877 plasma proteins and observed that 37 proteins were associated with the risk of incident AF over a nearly 6 year follow-up period at a Bonferonni corrected significance level and after adjustment for known AF risk factors. After additional adjustment for eGFR and medication use, 17 proteins remained significantly associated with an increased risk of AF. In a midlife replication sample that used proteins measured at an early ARIC visit, nearly half of the top proteins from the main analysis also demonstrated a robust association with non-overlapping incident AF events. Several proteins maintained relatively consistent effect sizes at both visits, including CILP2, IGFBP-2, and Angiopoietin-2, among others. In all analyses, NT-proBNP was the protein with the strongest association with incident AF. Using a less stringent FDR-corrected threshold, we performed network pathway analysis on the top 56 unique proteins mapped to genes and determined the top canonical pathway represented in our analysis was the inhibition of matrix metalloproteases. We identified several potential upstream regulators and mechanistic networks that provide insight into biological mechanisms involved in AF pathogenesis.

Natriuretic peptides (both NT-proBNP and mid-regional atrial natriuretic peptide) are markers of cardiac overload. Multiple prospective population-based cohort studies and previous proteomic analyses have reported higher baseline NT-proBNP concentrations predict increased incident AF.[10, 18, 20-23, 184, 214, 215] We also corroborated several other proteins that have been associated with incident AF in prior proteomic analyses including ATS13 (ADAMTS13) and Angiopoietin-2.[184] Additionally, previously reported BMP-1,[184] MMP-2, and IGFBP-7[214] associations with AF met our less-stringent FDR p value cutoff and were included in IPA. Angiopoietins are endothelial growth factors that regulate angiogenesis and vascular function and increased levels of angiopoietin-2 have been observed in several types of prevalent cardiovascular disease, including MI[221] and heart failure.[222] Similarly, the BMP signaling pathway plays an important role in the development of myocardial remodeling.[223] ATS13 is a von Willebrand factor protease that has been associated positively with incident MI, stroke, AF, and may be a marker of a prothrombotic environment.[185, 224, 225] The peptic hormone insulin-like growth factor 1 (IGF-1) and several of its binding proteins are associated positively with cardiovascular disease incidence,[226] and have additionally been linked to AF.[214, 227]

Our study reports several novel associations between circulating protein levels and incident AF which were also associated with incident AF in the mid-life replication analysis. Transgelin, a 22-kD protein of the calponin family, is exclusively and abundantly expressed

in the cytoskeleton of visceral and vascular smooth muscle cells. Transgelin influences the pulmonary arterial smooth muscle cells function which then promotes pulmonary vascular remodeling and was found to be significantly up-regulated in the lung tissue of patients with congenital heart disease and pulmonary arterial hypertension.[228] Many AF events are triggered by ectopic activation foci located in the pulmonary vein, and up-regulated transgelin may be indicative of pulmonary vascular remodeling that could result in AF. SVEP1 is a cell-adhesion molecule that acts as a ligand for integrin α9β1 and is believed to facilitate cellular adhesion in the context of pro-inflammatory signaling.[229, 230] The identification of a disease-associated missense variant in SVEP1 has been hypothesized to play a role in the development of atherosclerosis and coronary heart disease.[231] The role SVEP1 plays in contributing to AF remains to be clarified. In our causal pathway analysis, runt-related transcription factor 3 (RUNX3) was identified upstream of SVEP1. RUNX3 translocates in response to transforming growth factor (TGF)-β signaling, an important mediator of fibrosis.[232] Given that inflammation and oxidative stress are important in the pathogenesis of AF,[233] SVEP1 might increase susceptibility to AF by modulating these pathways. Additional prospective studies, using immunoassays, should verify whether Transgelin and SVEP1 are associated with AF incidence and whether the associations are causal.

Pathway analysis indicated our top canonical pathway was the inhibition of MMPs and that pathway included detected higher levels of TIMP-2, TIMP-4 and MMP-2. Atrial fibrosis is considered to be a key element of the AF substrate, with extracellular matrix (ECM) remodeling playing a major role in this process.[38] The MMPs are a family of twenty zinc-dependent enzymes that together with their specific endogenous inhibitors (tissue inhibitors of MMPs [TIMPs]), regulate the degradation of collagen and other ECM molecules. Several case-control studies have observed relationships between MMPs and AF, with the most significant associations related to MMP-9,[234, 235] and mixed results between MMP-2 and incident AF.[236] Observational studies of TIMP levels and AF have mainly shown no association, although higher TIMP-4 levels were found to be associated with prevalent AF in a few studies.[236-238] We found increased levels of both TIMP-2 and MMP-2 to both be associated with greater incident AF, and appear to be activated by several different regulators in our network analysis. Of course, associations with AF and cardiac diseases may differ according to whether levels are measured from circulating plasma or from tissue samples, as circulating levels may not reflect expression in cardiac tissue.[236] Furthermore, although expressed changes in MMPs and TIMPs occur in a number of cardiac disease states, these proteins appear to be differentially expressed in the atria and ventricles of patients with AF and end-stage heart failure.[239] Compiled, these observations provide evidence that a likely

mechanistic underpinning of interstitial atrial fibrosis with AF is changes in MMP and TIMP abundance and/or MMP and TIMP stoichiometry.[239]

IPA hypothesized relationships upstream of our target molecules along with the predicted activated / inhibited state of genes and gene products. The top upstream molecule was PTEN, which is involved in aging and tumor suppression and was predicted to be significantly inhibited based on the observed protein expressions in our data. PTEN negatively regulates intracellular levels of phosphatidylinositol-3,4,5-trisphosphate in cells and functions as a tumor suppressor by negatively regulating the AKT/PKB signaling pathway. P38 MAPK was predicted by the IPA to be activated and plays a role in apoptosis and cell differentiation. This protein kinase is also involved in a variety of binding steps, including magnesium ion binding, phosphatase binding, and transcription factor binding among other functions.

The genes implicated by our results had little overlap with previously identified AF-associated genes. This may be due to the advanced age of the participants in our study, as genetic associations with AF tend to be stronger in younger individuals. Additionally, our study had a relatively low number of AF events compared to GWAS studies and may lack power to detect some genetic associations. However, a few commonalities exist. TBX5 is a transcription factor that is critical to the formation of the cardiac electrical system and has been associated with the development of AF in several GWAS studies.[240, 241] In our analysis, TBX5 was present in our top hypothesized causal network and was regulated by MEF2C. Our results also corroborate previous findings that transcriptional regulation appears to be a key feature of AF etiology.[242] Nearly half of the identified upstream regulators from IPA are transcription regulators. Genetic variations may influence the function of transcription factors and affect the ion channels, development of cardiac conduct system or myocardium fibrosis, and play important roles in the pathogenesis of AF. Identification of the exact targets regulated by AF-related transcription factors may lead to potential new treatments for AF.

The main strengths of this study are the plethora of proteomic data in a community-based prospective sample, the quality of risk factor variables measured, and the number of AF events during follow-up. The ARIC study also includes black individuals, which have not been included in proteomic - AF analyses to date. We found no evidence of race interaction, indicating that the observed associations did not differ between blacks and whites. We were able to perform an internal mid-life replication analysis which strengthened our findings in older adults, however, replication in an external cohort would further strengthen the reproducibility and particularly establish the generalizability of these findings. Our study has several additional limitations. Incident AF was identified mainly from hospitalization

discharges, and we could be missing asymptomatic AF or AF managed exclusively in an outpatient setting. However, we and others have previously shown that the validity of AF ascertainment using hospitalizations is acceptable, and that incidence rates of AF in the ARIC study are consistent with other population-based studies.[9, 44] Additionally, we are unable to classify AF type (paroxysmal, persistent, or permanent AF) or assess the burden of AF (the percentage of time a person is in AF) accurately in the ARIC study. The possibility of protein degradation during long-term storage cannot be excluded; however, a validation study in ARIC did not support widespread protein degradation across visits.[243] Although our proteomic platform is the largest to date in cardiovascular research, we are only able to detect proteins included on this platform. Finally, SOMAscan measurements were semi-quantitative and need replication in other prospective studies.

In conclusion, we conducted proteomic profiling in a community-based population to assess the relationship between proteomics and incident AF in a cohort of older-aged black and white adults. The current results reinforced previous findings but additionally offer new observations into the biological changes that may precede AF onset and provide insight into mechanistic pathways of AF development. If replicated further, these novel proteins might be worth evaluating for AF risk scores or for possible pharmacologic targets in AF.

Table 12.1. Baseline Clinical Characteristics of Participants by Incident Atrial Fibrillation Status, ARIC, 2011-2013

	No incident atrial fibrillation through 2018	Incident atrial fibrillation through 2018
N	4083	585
Age, years	75.2 (5.1)	77.0 (5.4)
Female sex	2434 (60%)	304 (52%)
Black race	846 (21%)	73 (12%)
Height, cm	165.4 (9.3)	166.7 (9.9)
Weight, kg	78.2 (17.1)	81.0 (18.0)
Current smoker	229 (6%)	36 (6%)
Systolic blood pressure, mmHg	130.3 (17.8)	130.0 (19.0)
Diastolic blood pressure, mmHg	66.5 (10.5)	64.0 (10.9)
Antihypertensive medication use	2939 (72%)	484 (83%)
Diabetes	1257 (31%)	204 (35%)
Myocardial infarction	269 (7%)	67 (11%)
Heart failure	131 (3%)	60 (10%)
Estimated glomerular filtration rate, mL/min per m2	65.5 (17.7)	60.8 (17.9)
Anticoagulation medication use	90 (2%)	37 (6%)
Beta blocker medication use	1213 (30%)	285 (49%)
Antiarrhythmic use, class I and III	11 (0.3%)	13 (2%)

Values correspond to mean (standard deviation) or N (%)

Table 12.2. Protein Biomarkers Associated with Incident Atrial Fibrillation in Late-life, ARIC, 2011-2018

		Model 1		Model 2	
Protein Name	Gene Name	HR (95% CI)	p value	HR (95% CI)	p value
N-terminal pro-BNP	NPPB	1.82 (1.68-1.98)	**2.91E-45 ‡**	1.75 (1.60-1.91)	**4.59E-35 ‡**
Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	SVEP1	2.01 (1.71-2.36)	**2.39E-17 ‡**	1.89 (1.61-2.23)	**2.47E-14 ‡**
Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	SVEP1	1.92 (1.65-2.24)	**2.90E-17 ‡**	1.84 (1.57-2.16)	**3.31E-14 ‡**
Natriuretic peptides B	NPPB	1.52 (1.36-1.70)	**3.10E-13 ‡**	1.46 (1.30-1.65)	**4.58E-10 ‡**
Transgelin	TAGLN	1.88 (1.54-2.29)	**3.21E-10 ‡**	2.01 (1.56-2.59)	**6.41E-08 ‡**
Angiopoietin-2	ANGPT2	1.86 (1.53-2.25)	**2.88E-10 ‡**	1.74 (1.42-2.14)	**1.62E-07 ‡**
Protein delta homolog 1	DLK1	0.72 (0.63-0.84)	1.73E-05	0.68 (0.58-0.79)	**7.22E-07 ‡**
Slit homolog 2 protein	SLIT2	1.44 (1.25-1.65)	**2.90E-07 ‡**	1.41 (1.23-1.62)	**7.66E-07 ‡**
CMRF35-like molecule 2	CD300E	1.51 (1.27-1.80)	**2.28E-06 ‡**	1.52 (1.28-1.80)	**1.68E-06 ‡**
Protein delta homolog 1	DLK1	0.73 (0.63-0.85)	3.33E-05	0.68 (0.55-0.80)	**1.81E-06 ‡**
Antileukoproteinase	SLPI	1.97 (1.54-2.51)	**6.66E-08 ‡**	1.92 (1.46-2.52)	**2.43E-06 ‡**
Bone sialoprotein 2	IBSP	1.37 (1.22-1.54)	**1.05E-07 ‡**	1.33 (1.18-1.50)	**2.59E-06 ‡**
Microfibril-associated glycoprotein 4	MFAP4	1.54 (1.31-1.80)	**1.22E-07 ‡**	1.47 (1.25-1.72)	**3.13E-06 ‡**
Shadow of prion protein	SPRN	1.53 (1.26-1.84)	1.14E-05	1.57 (1.30-1.90)	**3.50E-06 ‡**
R-spondin-4	RSPO4	1.65 (1.35-2.02)	**1.45E-06 ‡**	1.63 (1.33-2.01)	**3.67E-06 ‡**
Chordin-like protein 1	CHRDL1	1.86 (1.47-2.37)	**3.17E-07 ‡**	1.79 (1.39-2.31)	**7.64E-06 ‡**
Spondin-1	SPON1	1.93 (1.49-2.49)	**6.24E-07 ‡**	1.81 (1.39-2.34)	**7.70E-06 ‡**
Endothelial cell-specific molecule 1	ESM1	1.76 (1.40-2.20)	**7.77E-07 ‡**	1.66 (1.32-2.07)	1.08E-05
R-spondin-1	RSPO1	1.65 (1.36-1.99)	**2.64E-07 ‡**	1.57 (1.28-1.92)	1.26E-05
Macrophage-capping protein	CAPG	1.53 (1.31-1.77)	**3.28E-08 ‡**	1.44 (1.22-1.70)	1.59E-05
Scavenger receptor class F member 1	SCARF1	1.81 (1.42-2.32)	**2.51E-06 ‡**	1.78 (1.37-2.31)	1.77E-05
Atrial natriuretic factor	NPPA	1.72 (1.42-2.09)	**3.03E-08 ‡**	1.54 (1.26-1.88)	2.30E-05
Insulin-like growth factor-binding protein 2	IGFBP2	1.40 (1.21-1.62)	**7.27E-06 ‡**	1.35 (1.16-1.57)	9.30E-05
Growth/differentiation factor 11/8	GDF11 MSTN	0.55 (0.42-0.72)	**9.54E-06 ‡**	0.59 (0.45-0.77)	9.78E-05
Triggering receptor expressed on myeloid cells 1	TREM1	1.56 (1.30-1.87)	**1.14E-06 ‡**	1.50 (1.22-1.84)	1.00E-04
A disintegrin and metalloproteinase with thrombospondin motifs 13	ADAMTS13	0.55 (0.43-0.71)	**2.56E-06 ‡**	0.60 (0.46-0.78)	1.15E-04
Metalloproteinase inhibitor 4	TIMP4	1.52 (1.26-1.82)	**7.03E-06 ‡**	1.43 (1.19-1.73)	1.58E-04

Ribonuclease pancreatic	RNASE1	1.29 (1.17-1.44)	**1.49E-06 ‡**	1.38 (1.17-1.64)	1.60 E-04
EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1	2.13 (1.57-2.90)	**1.24E-06 ‡**	1.94 (1.37-2.75)	1.70E-04
Regenerating islet-derived protein 3-alpha	REG3A	1.30 (1.16-1.46)	**4.81E-06 ‡**	1.26 (1.12-1.43)	2.01E-04
Lysosomal Pro-X carboxypeptidase	PRCP	0.56 (0.43-0.72)	**9.07E-06 ‡**	0.60 (0.46-0.79)	2.13E-04
Cartilage intermediate layer protein 2	CILP2	0.64 (0.53-0.78)	**6.27E-06 ‡**	0.69 (0.57-0.84)	2.15E-04
Sodium/potassium-transporting ATPase subunit beta-1	ATP1B1	0.62 (0.50-0.76)	**9.74E-06 ‡**	0.66 (0.53-0.82)	2.50E-04
Hepatitis A virus cellular receptor 2	HAVCR2	1.60 (1.31-1.95)	**3.57E-06 ‡**	1.51 (1.21-1.88)	3.05E-04
Endostatin	COL18A1	1.93 (1.47-2.55)	**3.14E-06 ‡**	1.90 (1.33-2.72)	4.07E-04
Protein SET	SET	0.36 (0.23-0.55)	**4.36E-06 ‡**	0.45 (0.28-0.71)	5.88E-04
Gamma-aminobutyric acid receptor-associated protein-like 1	GABARAPL1	1.80 (1.41-2.31)	**2.60E-06 ‡**	1.65 (1.23-2.21)	8.44E-04
Gamma-aminobutyric acid receptor-associated protein	GABARAP	1.95 (1.46-2.60)	**6.88E-06 ‡**	1.73 (1.22-2.47)	2.30E-03
Coagulation Factor X	F10	0.51 (0.40-0.64)	**3.09E-08 ‡**	0.69 (0.47-0.99)	4.50E-02
Coagulation factor Xa	F10	0.52 (0.41-0.66)	**8.99E-08 ‡**	0.71 (0.50-1.01)	5.95E-02

Model 1: adjusted for age, sex, race/center, current cigarette smoking, height, weight, systolic and diastolic blood pressure, the use of hypertension medications, diabetes, prevalent myocardial infarction and prevalent heart failure.

Model 2: adjusted for Model 1 + estimated glomerular filtration rate, antiarrhythmic medication use, beta blocker medication use, and anticoagulation use

†Hazard ratio (HR) expressed as the risk of incident AF per doubling of the protein value

‡Significance level of $P<0.05/4877 = 1.025 \times 10^{-5}$. These 40 proteins are ordered by smallest to largest p-value for Model 2.

Table 12.3. Replication Analysis of Associations of the Top 40 Late-Life Protein Biomarkers Measured in Mid-life with Incident Atrial Fibrillation, ARIC, 1993-2010

		Model 1		Model 2	
Protein Target Name	Gene Name	HR (95% CI)	p value	HR (95% CI)	p value
N-terminal pro-BNP	NPPB	1.40 (1.32-1.47)	**3.29E-35** ‡	1.37 (1.30-1.45)	**2.92E-31** ‡
Angiopoietin-2	ANGPT2	1.77 (1.55-2.02)	**3.54E-17** ‡	1.73 (1.51-1.98)	**6.42E-16** ‡
Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	SVEP1	1.57 (1.36-1.80)	**3.11E-10** ‡	1.57 (1.36-1.81)	**4.20E-10** ‡
Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	SVEP1	1.52 (1.33-1.74)	**6.99E-10** ‡	1.52 (1.33-1.74)	**9.91E-10** ‡
Triggering receptor expressed on myeloid cells 1	TREM1	1.49 (1.30-1.71)	**1.14E-08** ‡	1.46 (1.26-1.68)	**1.73E-07** ‡
Insulin-like growth factor-binding protein 2	IGFBP2	1.26 (1.16-1.37)	**7.11E-08** ‡	1.25 (1.15-1.36)	**1.91E-07** ‡
Ribonuclease pancreatic	RNASE1	1.32 (1.21-1.45)	**2.49E-09** ‡	1.31 (1.18-1.45)	**3.80E-07** ‡
EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1	1.83 (1.47-2.27)	**3.86E-08** ‡	1.75 (1.40-2.18)	**6.58E-07** ‡
Transgelin	TAGLN	1.50 (1.30-1.74)	**6.20E-08** ‡	1.46 (1.25-1.70)	**2.33E-06** ‡
Natriuretic peptides B	NPPB	1.25 (1.13-1.38)	**7.82E-06** ‡	1.23 (1.12-1.36)	**2.52E-05** ‡
Protein SET	SET	0.56 (0.42-0.73)	**2.12E-05** ‡	0.58 (0.44-0.75)	**7.90E-05** ‡
Gamma-aminobutyric acid receptor-associated protein	GABARAP	1.69 (1.36-2.11)	**2.37E-06** ‡	1.58 (1.25-2.00)	**1.17E-04** ‡
Gamma-aminobutyric acid receptor-associated protein-like 1	GABARAPL1	1.48 (1.24-1.76)	**1.14E-05** ‡	1.42 (1.18-1.70)	**2.49E-04** ‡
Hepatitis A virus cellular receptor 2	HAVCR2	1.34 (1.16-1.54)	**5.11E-05** ‡	1.29 (1.12-1.49)	**4.43E-04** ‡
Microfibril-associated glycoprotein 4	MFAP4	1.21 (1.08-1.34)	**5.41E-04** ‡	1.21 (1.09-1.35)	**4.56E-04** ‡
Cartilage intermediate layer protein 2	CILP2	0.77 (0.67-0.89)	**3.87E-04** ‡	0.77 (0.67-0.89)	**4.98E-04** ‡
Endostatin	COL18A1	1.45 (1.21-1.74)	**7.41E-05** ‡	1.39 (1.15-1.68)	**7.77E-04** ‡
Antileukoproteinase	SLPI	1.46 (1.21-1.76)	**8.97E-05** ‡	1.38 (1.13-1.68)	1.32E-03
R-spondin-4	RSPO4	1.37 (1.14-1.64)	**6.87E-04** ‡	1.34 (1.12-1.61)	1.44E-03
Scavenger receptor class F member 1	SCARF1	1.37 (1.13-1.66)	**1.10E-03** ‡	1.34 (1.11-1.63)	2.91E-03
Chordin-like protein 1	CHRDL1	1.38 (1.13-1.68)	1.52E-03	1.33 (1.09-1.63)	5.16E-03
R-spondin-1	RSPO1	1.23 (1.08-1.40)	1.64E-03	1.21 (1.06-1.38)	5.36E-03
Spondin-1	SPON1	1.29 (1.07-1.56)	6.51E-03	1.28 (1.06-1.55)	9.41E-03
Protein delta homolog 1	DLK1	0.91 (0.83-1.00)	4.96E-02	0.88 (0.80-0.97)	1.10E-02
Protein delta homolog 1	DLK1	0.91 (0.82-1.00)	5.22E-02	0.88 (0.79-0.97)	1.24E-02

A disintegrin and metalloproteinase with thrombospondin motifs 13	ADAMTS13	0.81 (0.70-0.93)	3.70E-03	0.83 (0.72-0.96)	1.39E-02
Metalloproteinase inhibitor 4	TIMP4	1.19 (1.04-1.35)	1.06E-02	1.16 (1.02-1.33)	2.29E-02
Slit homolog 2 protein	SLIT2	1.15 (1.00-1.31)	4.96E-02	1.15 (1.01-1.32)	3.96E-02
Shadow of prion protein	SPRN	1.15 (0.99-1.35)	6.63E-02	1.17 (1.00-1.36)	4.56E-02
Lysosomal Pro-X carboxypeptidase	PRCP	0.81 (0.69-0.96)	1.42E-02	0.84 (0.71-1.00)	4.84E-02
Regenerating islet-derived protein 3-alpha	REG3A	1.12 (1.02-1.22)	1.19E-02	1.09 (1.00-1.19)	5.64E-02
Growth/differentiation factor 11/8	GDF11 MSTN	0.88 (0.74-1.04)	1.30E-01	0.87 (0.73-1.03)	1.05E-01
Macrophage-capping protein	CAPG	1.12 (1.01-1.25)	2.72E-02	1.09 (0.98-1.21)	1.06E-01
Sodium/potassium-transporting ATPase subunit beta-1	ATP1B1	0.90 (0.79-1.02)	8.61E-02	0.91 (0.80-1.03)	1.26E-01
Bone sialoprotein 2	IBSP	1.05 (0.95-1.15)	2.78E-01	1.06 (0.97-1.17)	1.98E-01
CMRF35-like molecule 2	CD300E	1.08 (0.93-1.24)	3.06E-01	1.09 (0.93-1.26)	2.39E-01
Atrial natriuretic factor	NPPA	1.11 (0.92-1.34)	2.82E-01	1.11 (0.91-1.34)	2.99E-01
Coagulation Factor X	F10	0.87 (0.71-1.06)	1.71E-01	0.96 (0.77-1.20)	7.30E-01
Endothelial cell-specific molecule 1	ESM1	1.03 (0.89-1.18)	7.15E-01	1.02 (0.89-1.17)	7.52E-01
Coagulation factor Xa	F10	0.88 (0.72-1.08)	2.08E-01	0.98 (0.78-1.22)	8.24E-01

Model 1: adjusted for age, sex, race/center, current cigarette smoking, height, weight, systolic and diastolic blood pressure, the use of hypertension medications, diabetes, prevalent myocardial infarction and prevalent heart failure.

Model 2: adjusted for Model 1 + estimated glomerular filtration rate, antiarrhythmic medication use, beta blocker medication use, and anticoagulation use

†Hazard ratio (HR) expressed as the risk of incident AF per doubling of the protein value

‡Significance level of $P<0.05/40 = 1.25 \times 10^{-3}$. These 40 proteins are ordered by smallest to largest p-value for Model 2.

Figure 12.1. Beta estimates for Associations of the Top 40 Protein Biomarkers Measured in Mid-life (visit 3) and Late-Life (visit 5) with Incident Atrial Fibrillation, ARIC, 1993-2018.

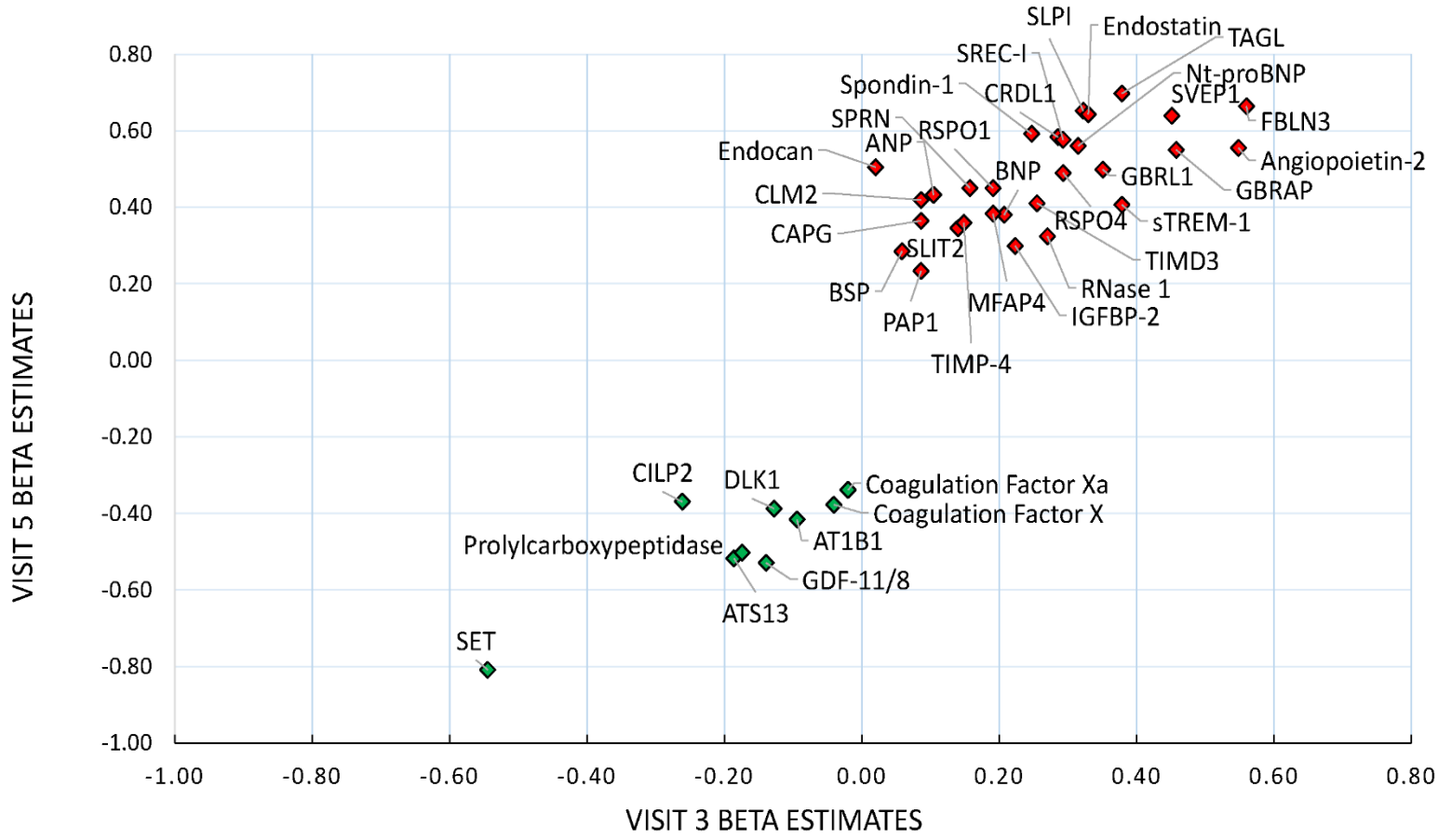
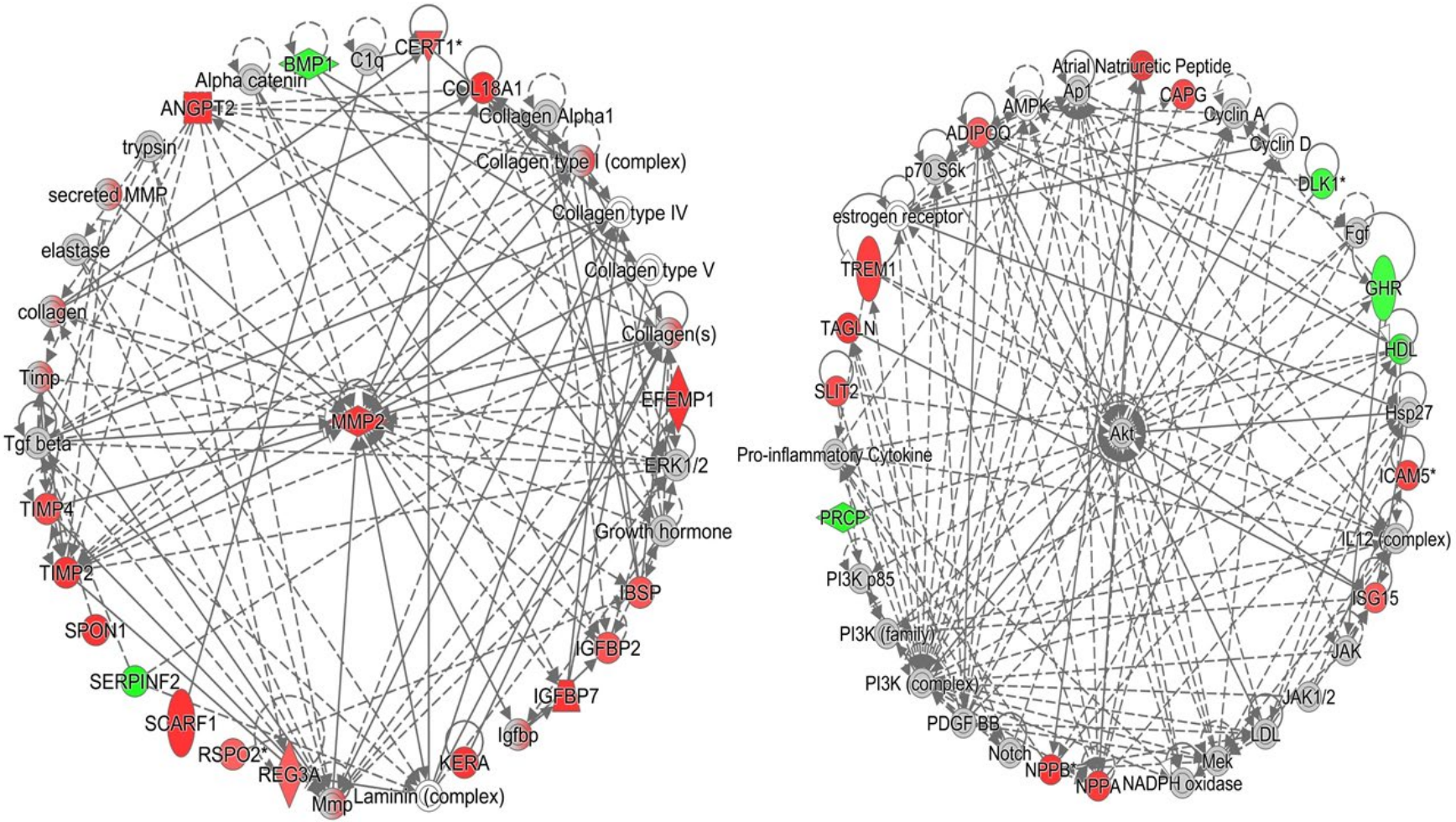


Figure 12.2. The top 2 protein networks identified using Ingenuity Pathway Analysis (IPA) for the association with incident AF.



Each network is depicted radially, with the protein most central to the figure in the center.

Figure 12.3. A depiction of the top identified canonical pathway, the inhibition of matrix metalloproteases.

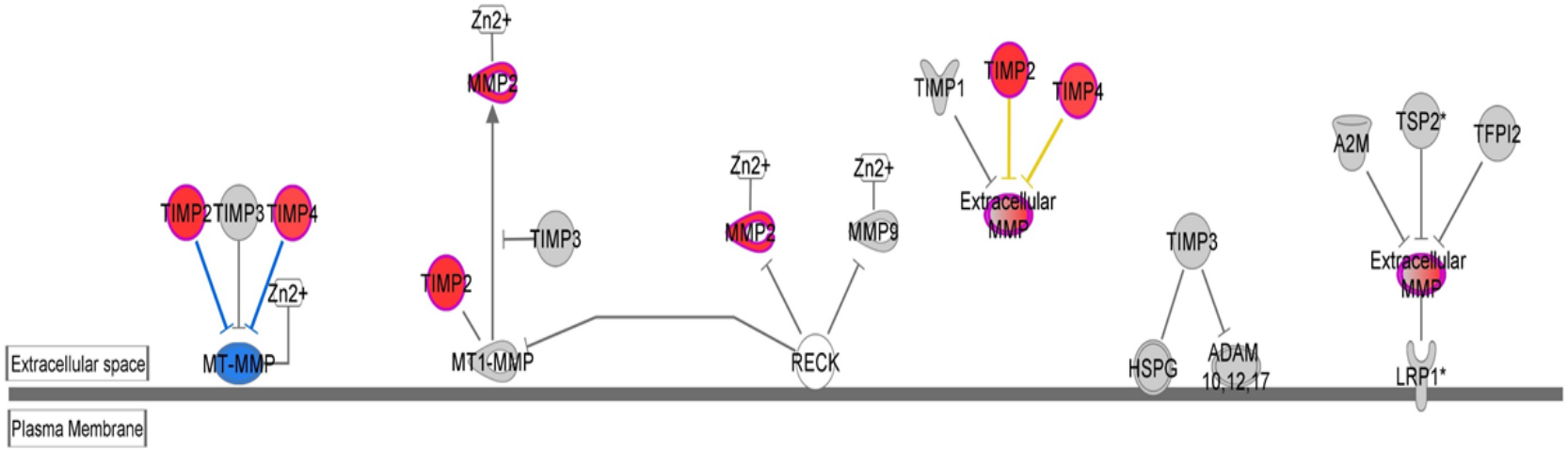
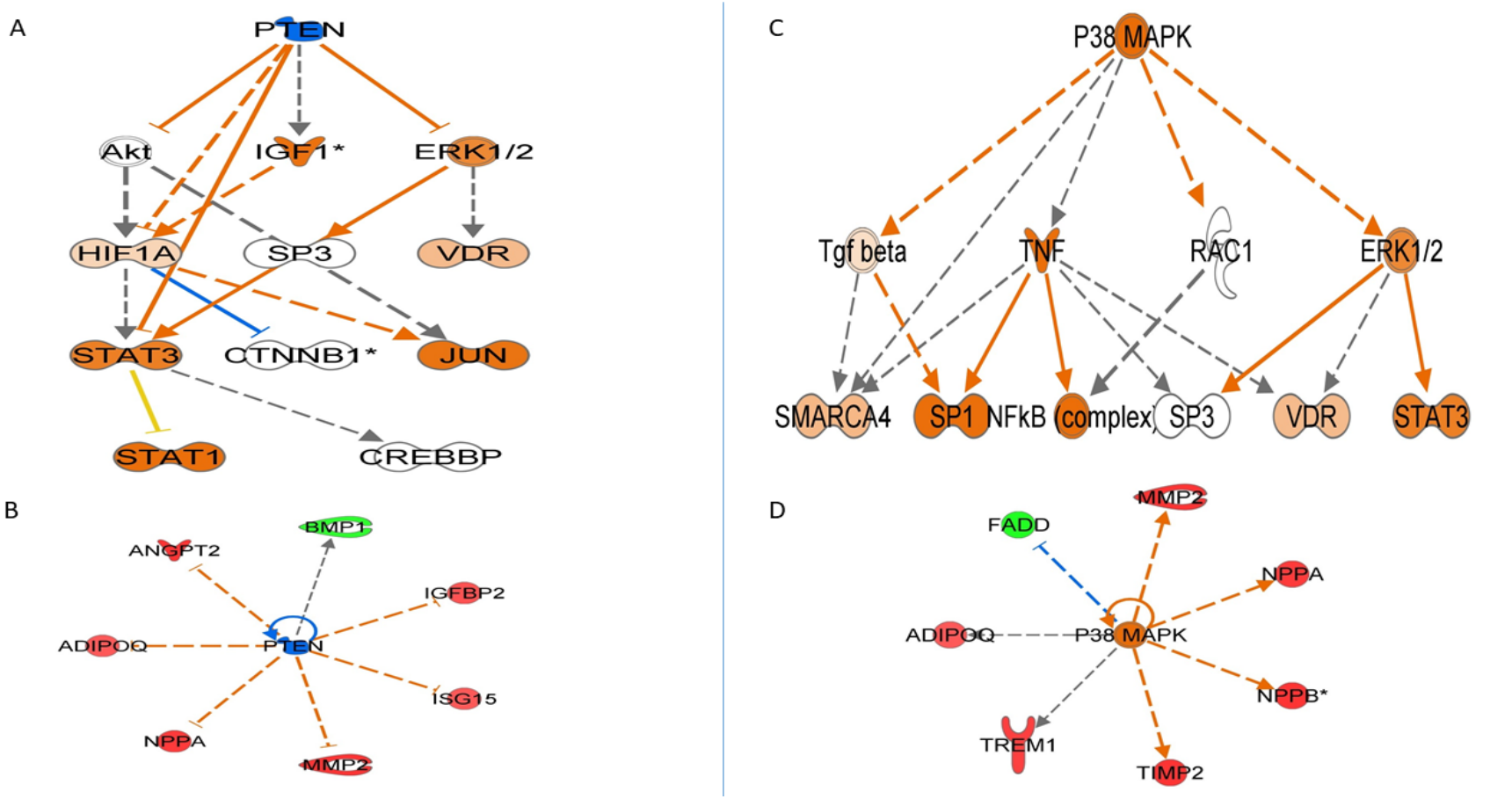


Figure 12.4. The top upstream regulators identified using IPA, based on experimentally observed relationships between regulators and genes or gene products.



Panel A depicts the hierarchical associations between PTEN and its expected downstream regulators to produce the associations we observe in the proteins in Panel B. Panel C depicts the hierarchical associations we would expect between P38 MAPK and intermediate regulators to produce the effects on the proteins observed in Panel D.

Figure 12.5. The top causal network identified using IPA, centered around MEF2C as the master regulator and connecting 19 observed protein states through intermediate regulators.

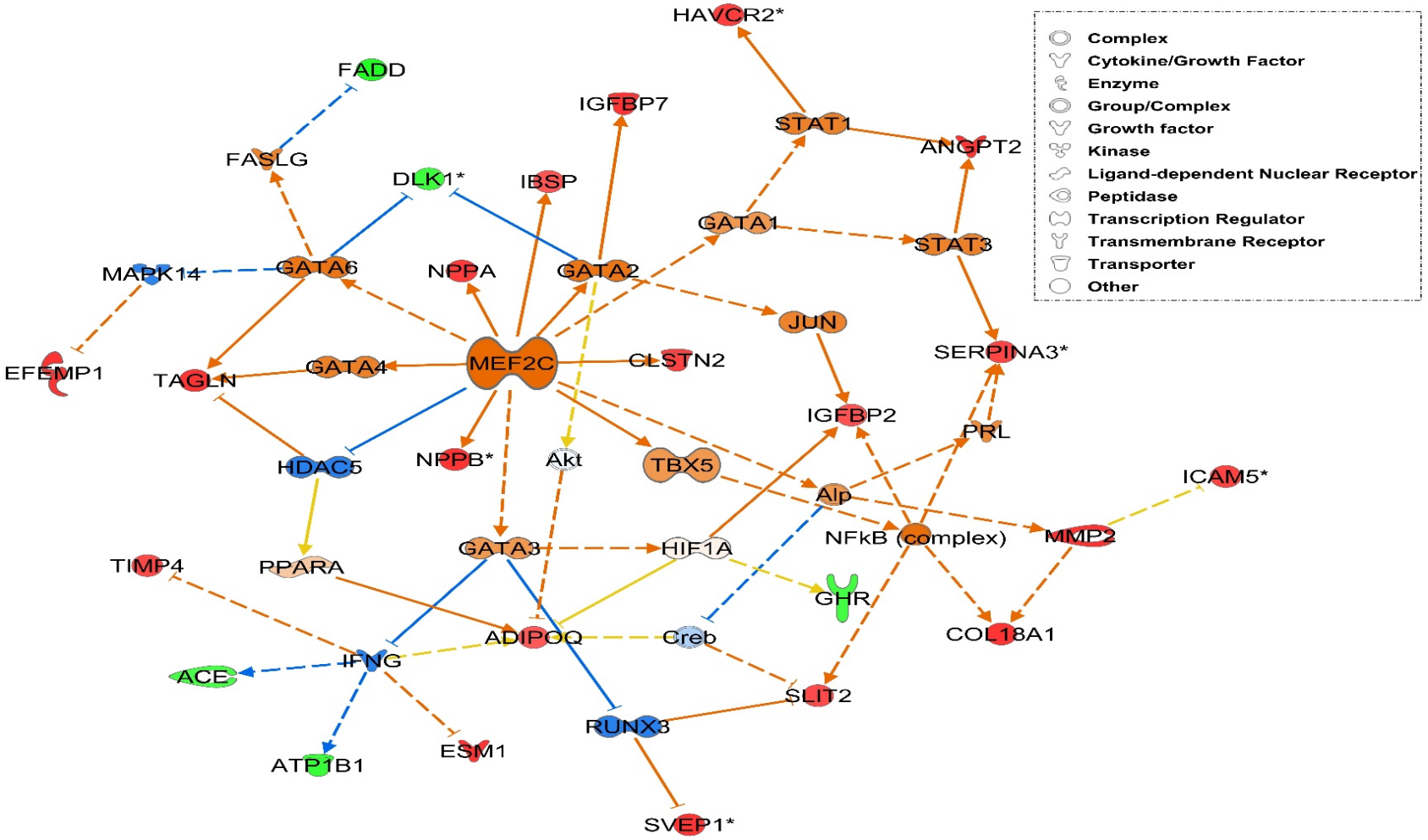
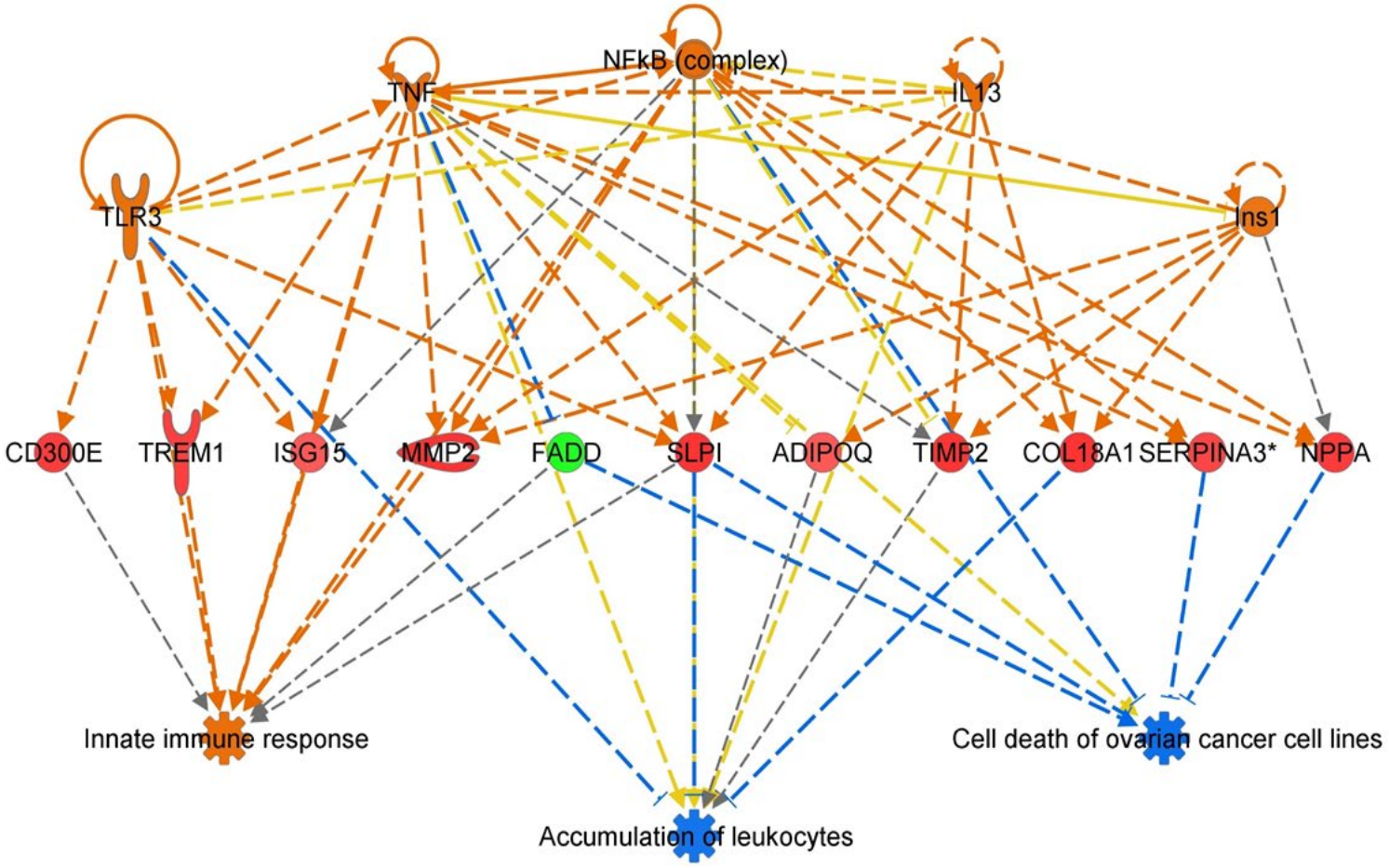


Figure 12.6. The top regulator effect network identified using IPA which depicts how predicted activated upstream regulators might cause increases or decreases in phenotypic or functional outcomes downstream.



Supplemental Methods

SomaLogic Quality Control

All measures are reported as relative fluorescent units (RFU). Each measure has been validated for its specificity, upper and lower limits of detection, and intra- and inter-assay variability.[244] Previous work indicates median intra- and inter-run coefficients of variation of approximately 5% and an intra-class correlation coefficient of ~0.9.[243, 245] A list of all the 5,284 modified aptamers in the v.4 SOMAscan menu can be found in the supplement to a publication by Williams *et al*.[246]

Protein analyte measurements underwent the regular SOMAscan data standardization and normalization process.[243, 247] Briefly, hybridization control normalization was first applied to each sample based on a set of hybridization control sequences to correct for systematic biases during hybridization. Second, median signal normalization was applied to measures within a plate to remove sample or assay biases that may be because of pipetting variation, variation in reagent concentrations, assay timing, and other sources of systematic variability within a single plate run. Finally, each plate contained calibrator samples for each SOMAmer reagent, which was used to correct for plate-to-plate variation based on established global reference standards. Protein analytes with a calibration factor greater or less than the median calibration factor (0.4) were excluded from all analyses.

ARIC quality control

We inserted blind split-sample duplicate plasma aliquots for 197 of the 5327 (3%) participants with available SOMAmer data at visit 5 and 422 of 11,565 (4%) participants with SOMAmer data available at visit 3. The median inter-assay coefficient of variation for SOMAmers measured from visit 5 plasma (calculated using the Bland-Altman method because proteins levels are measured on a relative scale [CV_{BA}]) was 4.7%. The median inter-assay CV_{BA} for SOMAmers measured from visit 3 plasma was 6.3%. Thus, the older samples collected at ARIC visit 3 show good performance overall for many proteins, but were inferior to those collected at visit 5. The median split sample reliability coefficient was 0.85 at visit 3 and 0.94 at visit 5, after excluding quality control outliers, as described below.

Of the 5284 available SOMAmers, we excluded 94 that had a CV_{BA} >50% or a variance of < 0.01 on the log scale at either visit 5 or visit 3. Additionally, we excluded 313

SOMAmers because of binding to non-proteins, including hybridization control elution, non-human proteins, non-biotin, non-cleavable, and spuriomer products.

In a previous study on a subset of ARIC participants, we also were able to validate the measurement of three AF-associated aptamers, compared with immunoassays in the ARIC central laboratory. SomaScan and traditional immunoassay measurements were highly correlated: NT-proBNP (n=5168, r=0.90), B2M (n=5313, r=0.92), and GDF15 (n=142, *r*=0.94).[243]

ARIC covariate measures

Participants reported information on smoking, history of cardiovascular disease, use of medications, and underwent a physical exam at each visit that included height and weight. Seated blood pressure was measured using a random-zero sphygmomanometer after 5 minutes rest, and was defined as the average of the 2nd and 3rd measurements taken. Diabetes mellitus was defined as fasting glucose ≥ 126 mg/dL (7.0 mmol/L), non-fasting glucose ≥ 200 mg/dL (11.1 mmol/L), treatment for diabetes mellitus, or self-reported physician diagnosis of diabetes. MI was ascertained by study visit ECGs or the ARIC Morbidity and Mortality Classification Committee, by using data from follow-up calls, hospitalization records and death certificates.[210] Prevalent HF was defined as the reported use of HF medication in the previous two weeks, presence of HF according the Gothenburg criteria (only at the baseline ARIC visit), or having had a HF hospitalization during follow-up.[248,249] Plasma creatinine and cystatin C were measured, and eGFR was calculated as mL/min/1.73 m^2 using the CKD Epidemiology Collaboration (CKD-EPI) combined creatinine-cystatin C equation.[250]

Ingenuity Pathway Analysis

IPA is a knowledge database relying on published literature related to protein function, localization, relevant interactions, and biological mechanisms. We ran the core analysis using the Ingenuity Knowledge Base as the reference set and included both direct and indirect experimentally confirmed relationships from all species. Networks were then algorithmically generated based on their connectivity. Methods used to determine the overlap of p-value, the activation z-score and causal analysis in IPA have been previously published.[220]

The IPA Core Analysis calculates *P*-values using a right-tailed Fisher's exact test to quantify the probability of overlap between a set of AF-associated proteins identified in current analysis and a set of proteins known to exist within a specific pathway or process due to random chance. We used a *P*-value of <0.05 as the threshold for statistical significance after applying Benjamini-Hochberg FDR adjustment for multiple comparisons. A z-score was also calculated, which quantifies the likelihood and directionality of the expression of canonical pathways and upstream regulators, considering the direction of the protein-specific association in our dataset and the known directional effect of one molecule on another molecule or on a process. A z-score <-2 or >2 has been recommended as the threshold for statistical significance when interpreting directionality. We used IPA network analysis to identify interactions between groups of highly connected proteins associated with AF risk, and the program generated algorithmically based networks on known genetic or molecular connectivity with other gene or gene products. Highly connected proteins or genes are first identified as focus molecules or "seeds." Focus molecules identified as having the most interactions with other focus molecules are connected to form a network. We used IPA Upstream Regulator Analysis to identify the cascade of upstream transcriptional regulators that may be responsible for gene expression changes observed in our analysis. IPA predicts which upstream regulators are activated or inhibited to explain the upregulated and down-regulated proteins that we observed in our dataset. Mechanistic networks are provided for many of the upstream regulators and identify signaling cascades that connect upstream regulators to visualize how they may work together to elicit expression changes observed in our dataset.

We implemented 2 additional advanced analytics components of IPA.[220] We performed a causal network analysis which uses algorithms based on a "master" network derived from the Ingenuity Knowledge Base. The power of causal network analysis is its ability to detect novel master upstream regulators that operate through other regulators, especially in cases where few or no relationships exist directly between it and the dataset genes. Finally, we used the Regulator Effects analysis function in IPA to determine potential pathways between upstream regulators, our measured proteins, and downstream diseases and functions. Regulator Effects explains how predicted activated or inhibited upstream regulators might cause increases or decreases in phenotypic or functional outcomes downstream. These causal hypotheses take the form of directionally coherent networks formed from the merger of Upstream Regulator networks with Downstream Effects networks.

13. Manuscript 2 -- Developing a Prediction Model for Atrial Fibrillation in the Elderly: The Atherosclerosis Risk in Communities (ARIC) study

13.1. OVERVIEW

Background –Existing atrial fibrillation (AF) risk prediction scores are not well-calibrated to older populations and improving these scores may depend on whether novel markers, such as proteomics, can add to or refine predictive models. We used data from the Atherosclerosis Risk in Communities (ARIC) study, a cohort of older-aged black and white adults in the US, to derive and internally validate 5-year prediction scores for incident AF.

Methods – Our analysis included 4308 AF-free participants (mean age 75 ± 5 years; 58% female; 19% black race) with clinical, proteomic, ECG, and echocardiograph measures, who attended visit 5 in 2011-13. Using Cox regression models in 1000 bootstrapped samples, we developed a series of models from simple to involved that selected variables predicting incident AF within a 5 year period. The models were internally validated using 1000 bootstrapped samples and adjusted for optimism.

Results –A total of 394 participants developed AF over a 5-year follow-up. The final simple predictive stepwise model included the variables of age, race, weight, myocardial infarction, heart failure, and use of beta-blockers, anti-arrhythmic agents, and anticoagulants, and had moderate discrimination (c-statistic 0.697; 95% CI 0.671-0.723). The addition of blood biomarkers plus 16 proteins from proteomic analysis greatly improved the discrimination (c-statistic 0.795; 95% CI 0.773-0.816) while still showing excellent calibration (χ^2 = 7.6; P = 0.58). Addition of abnormal P wave axis, left atrial diameter, and septal E/e prime moderately increased the c-statistic to 0.806 (95% CI: 0.785-0.827) in the full-developed model that contained 30 variables total. Using internal validation adjusted for optimism, discrimination of the fully-developed prediction model was acceptable (c-statistic 0.795).

Conclusion –We developed a series of AF prediction models that are better targeted and calibrated to older populations. Results from our study should be externally validated. The addition of biomarkers, including proteomics data, improved prediction, suggesting it may be worthwhile to explore developing cost-effective and time-efficient ways to quantify the predictive protein biomarkers.

13.2. INTRODUCTION

Atrial fibrillation (AF), a cardiac arrhythmia, has emerged as a major public health problem. AF is largely a disease of advancing age[2, 3] and is associated with increased risks of adverse cardiovascular outcomes including stroke,[4] myocardial infarction,[5] and mortality,[6] resulting in significant costs to the US healthcare system.[7] The aforementioned complications and financial burden associated with AF underscore the importance of accurate AF risk assessment. A well-calibrated risk score for prediction of incident AF would optimize screening in high-risk older individuals, allow for more specific clinical trial enrollment, and would lead to opportunities for targeted preventive strategies.

Several AF risk prediction scores have been developed to predict AF with respectable discriminative abilities in middle-aged adults, including the Cohorts for Aging and Research in Genomic Epidemiology (CHARGE)-AF risk score.[10] However, these scores may have diminished discrimination in populations of older adults, in which AF is most prevalent. A risk score that remains accurate in elderly populations is important; the burden that AF places on the health care system is increasing with the growth in the number of individuals in older age categories.

Improving established risk prediction scores may depend on whether novel markers can add predictive value to existing models. New technology has allowed for the systematic assessment of a large portion of the entire range of proteins measurable in plasma (the plasma proteome), commonly referred to as proteomics. The application of proteomics provides opportunities for unbiased discovery of novel markers to improve accuracy in the prediction of AF. Furthermore, since the publication of the CHARGE-AF risk score in 2013 (derivation c-statistic =0.765), additional variables have been linked with an increased risk of incident AF and should be considered when re-evaluating an AF risk prediction equation. Current scores lack variables incorporating left atrial function and the addition of these variables along with biomarkers and proteomic profiles could improve individual AF risk prediction. We evaluated the performance of the existing CHARGE-AF risk score[10] in a cohort of older black and white men and women. We then developed and internally validated 4 predictive scores for incident AF that ranged from a simple model to more complex models that included variables from lab measures, proteomics, ECGs and echocardiograms.

13.3. METHODS

Study population

The Atherosclerosis Risk in Communities (ARIC) study is a prospective cohort study of cardiovascular disease and atherosclerosis risk factors.[209] Participants at baseline (1987-

1989) included 15,792 black and white men and women aged 45-64, recruited from 4 communities in the US (Washington County, Maryland; the northwest suburbs of Minneapolis, Minnesota; Jackson, Mississippi; and Forsyth County, North Carolina). Thus far, 7 study visits have been completed with visit 5 occurring in 2011-2013. Additionally, ARIC participants have received annual follow-up calls (semi-annual after 2012), with response rates of ≥ 90% among survivors. We chose ARIC visit 5 as baseline due to the older-age population and the availability of lab, proteomic, ECG and echocardiogram variables obtained at this visit. Among the 6538 participants that attended visit 5, we excluded those with prevalent AF at visit 5 (n=631), missing proteomics measures (n=1159), missing or indeterminate ECG or echocardiograph measures (n=342), race other than white or black and non-whites in the Minneapolis and Washington County field centers (due to small numbers; n= 42), and those missing covariates (n=56). This study was approved by institutional review boards at each participating center, and all study participants provided written informed consent.

Ascertainment of AF

We defined incident AF as in previous ARIC analyses.[9] A trained abstractor obtained and recorded all International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) and ICD-10-CM hospital discharge diagnoses from each participant's hospitalizations reported in the follow-up interview. AF was defined as the presence of ICD-9-CM code 427.31 or 427.32 or ICD-10-CM code I48.xx. AF hospitalization diagnoses occurring simultaneously with heart revascularization surgery or other cardiac surgery involving heart valves or septa were not included as AF events. Deceased ARIC participants were also labeled as AF cases if their underlying cause of death was AF. We identified 95% of the incident AF events from hospitalization records. Validity of ICD codes for AF is adequate as approximately 90% of the cases were confirmed in a physician review of discharge summaries from 125 possible AF cases.[9] We ascertained incident AF events through a 5-year time period after visit 5, and the date of the first AF event was considered the outcome date.

Candidate Prediction variables

We identified candidate predictors of incident AF from the literature and other prediction models (i.e. CHARGE-AF[10] and its augmented models[18]). Candidates included clinical variables, blood measures and biomarkers, ECG variables, and echocardiographic variables measured at visit 5. We included several P wave indices from the including abnormal P wave axis,[88, 251] and echocardiograph variables[91, 252] associated with AF. We considered proteomics data consisting of 4877 proteins recently measured in the ARIC

cohort. Detailed procedures for ARIC measures have been published,[209] and further details can be found in the Supplement Materials, along with a full list of all variables considered for inclusion. In brief, participants reported information on smoking, history of cardiovascular disease, use of medications, and underwent physical assessments and blood draw at the study visit. We measured protein levels in plasma samples using the SomaScan platform, which uses single-stranded DNA-based aptamers to capture conformational protein epitopes. Participants underwent ECGs and echocardiograms. Where appropriate, we evaluated candidate predictors as continuous variables, and if clinical cutpoints existed we also evaluated the variable by established cutpoints. We log base 2 transformed each proteomic variable and winsorized outliers that were greater or less than 5 standard deviations from the sample mean on the log 2 scale.

Statistical analysis

We calculated person-years of follow-up from exam 5 (2011-2013) until first AF diagnosis, death, loss to follow-up, or a follow-up of 5 years, whichever came first. All predictor variables were from exam 5 and not updated during follow-up. Time to incident AF was the outcome for all models. All statistical analyses were performed with SAS 9.4 (SAS Institute, Cary, NC).

Performance of the CHARGE-AF risk score

First, we assessed the performance of the original CHARGE-AF risk score to determine the risk of AF in an elderly cohort. We evaluated model performance using the c-statistic,[253] and Nam and D'Agostino's modified Hosmer-Lemeshow chi-square statistic for survival analysis (calibration).[254] Calibration was also qualitatively assessed by plotting the observed risk within deciles of predicted risks.

Overview of the derivation of predictive models

We performed an analysis to identify predictors and we created new AF risk scores in the elderly by deriving a 5-year predictive models. We followed guidelines from the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement.[255] The TRIPOD Statement is a set of recommendations for the reporting of studies developing, validating, or updating a prediction model, whether for diagnostic or prognostic purposes. A TRIPOD checklist is attached in the

Supplemental Materials. Currently, other potential validation cohorts lack the proteomic measures we included in our prediction model, and therefore, in concordance with the TRIPOD guidance, we developed and validated the model using the entire data set, but then used bootstrap resampling techniques to evaluate the performance and calibration of the developed model, thus internally validating our model.[256]

To facilitate the use of our score in clinical settings with limited access to blood and diagnostic tests, we first developed a predictive model that did not require information from a blood draw, ECG, or electrocardiogram, which we labeled the "simple model". We then developed a sequence of more complex models, first by considering measures obtained from a blood draw. We included routine measures from a lab blood draw, including common biomarkers, which we labeled the "lab biomarker model". We then created another model that contained proteomics variables measured from a blood plasma sample, which we labeled "proteomics model". Finally, we developed a more complex model by adding in candidate ECG and echocardiographic variables, which we labeled "ECG and echo model".

Derivation of the predictive models

We provide full details for each the derivation of each predictive model in the Supplemental Methods. Briefly, we initially ran minimally-adjusted Cox proportional hazard models to assess the individual predictors of AF. Variables significantly associated with AF were then considered candidate predictors and selected for inclusion into the next step. We generated 1000 bootstrap samples and ran Cox models with backward selection of the candidate predictors in each of the 1000 models. Based on recommendations from the literature, we selected variables included in at least 60% of the Cox models for inclusion in the final predictive model.[257] As we progressed from the simple model to the more complex, we forced predictors from the previous model to stay in the model during backward selection in order to assess the added predictive value of the new variables.

Assessment of the derived predictive models

Once the final variables were selected for each model, we calculated the model-based individual 5-year risk of AF using the beta estimates and ARIC mean reference values from each variable. We evaluated model performance using the c-statistic,[253] and calibration chi-square using Nam and D'Agostino's modified Hosmer-Lemeshow statistic.[254] We calculated the added predictive value of the complex models versus the simple model by assessing the discrimination slopes,[258] and the categorical net reclassification improvement (NRI) using the

following risk categories: <5%, 5 to 15%, >15%. [258] We arbitrarily chose these risk categories based on our simple score where approximately 25% of our sample had a risk <5% and 20% had a risk > 15% and we wanted to see meaningful movement across categories. In addition, we estimated relative integrated discrimination improvement (IDI), which is the ratio of absolute difference in discrimination slopes of the 2 models over the discrimination slope of the original model.[259]

Validation analysis

When prediction models are developed in relatively small samples, they may be overfitted and may show optimistic performance. To adjust for overfitting and optimistic performance of the model, we used bootstrap resampling for internal validation.[255] Bootstrapping methods provide more stable estimates with a lower bias compared to other methods of internal validation.[256] We generated 1,000 bootstrap samples, sampling with replacement. A prognostic model was developed in each sample, and the performance was evaluated in the bootstrap samples and applied to the original sample. We used the same cohort for developing the score and for validation and therefore adjusted for optimism in our c-statistic obtained from the internal validation.[260]

13.4. RESULTS

Our analysis included 4308 AF-free participants (mean age 75 ± 5 years; 5% female; 19% black race) with clinical, proteomic, ECG and echocardiographic measures, who attended visit 5 in 2011-13. A total of 394 participants developed AF during a 5-year follow-up. Descriptive characteristics are provided in **Table 13.1** based on incident AF status. Those who developed AF were older, more likely to be male, white, use listed medications, and had a worse cardiovascular profile, including higher levels of biomarkers, compared to those who did not develop AF.

We applied the original, simple CHARGE-AF risk score in our cohort and obtained a modest c-statistic (95% CI) of 0.660 (0.634-0.686) with a poor calibration value of 27.4 (p=0.001). We plotted the predicted vs. observed risk by decile and found the CHARGE-AF risk score over-predicted the risk in the lower deciles of risk.

In the derivation of our simple clinical model, the following 8 variables were selected in at least 60% of the predictive bootstrap samples: age, race, weight, prevalent heart failure, prevalent MI, use of beta-blockers, anticoagulants, and anti-arrhythmic medications. Hazard ratios (HR) and 95% CI for each predictor are listed in **Table 13.2**. Heart failure was one of

the strongest predictors of AF in the simple model [HR (95% CI) = 2.27 (1.59-3.25)], along with the use of anti-arrhythmic agents and anticoagulants.

For derivation of the lab biomarker model, plasma troponin-T and NT-proBNP were selected in at least 60% of the predictive bootstrap samples and were added to the simple prediction model, along with eGFR. The variable of eGFR was inversely associated with the risk of incident AF in age and sex-adjusted analysis, however, when included in the model with NT-proBNP, the association reversed direction. eGFR was considered borderline significant in this analysis (selected in 54% of the predictive bootstrap samples) and knowing that eGFR was an important variable to consider when measuring blood protein levels, we made the modeling decision to allow eGFR to stay and move forward in the sequence of deriving prediction models.

In developing the proteomics model, 16 proteins were selected in at least 60% of the prediction bootstrap samples and are listed in **Table 13.2**. Although the lab biomarker of NT-proBNP (per 1 log-transformed SD) was already included in the model, the proteomic measure of NT-proBNP (log-base 2 transformed) was also significantly associated with incident AF and added predictive value in the model. Antileukoproteinase (SLPI) was strongly associated with AF; the risk of AF was 2.24 times higher for every doubling of the protein level [HR (95% CI) = 2.24 (1.55-3.25)]. Finally, when evaluating ECG and echocardiographic variables, the measures of abnormal P wave axis, left atrial diameter, and septal E/e-prime were included in the final predictive model.

Table 13.3 lists the c-statistic, calibration, NRI and IDI for each model. In the simple model, the c-statistic of the simple prediction score was 0.697 (95% CI, 0.671-0.723) with appropriate calibration chi-square of 9.4 (p=0.40). Driven mainly by NT-proBNP, the addition of the lab biomarkers raised the c-statistic to 0.742 (95% CI, 0.717-0.767) and calibration remained adequate (16.0; p-value =0.07). The NRI significantly increased 0.233 (95% CI, 0.181, 0.285) as did the IDI. The NRI tables for each model compared to the simple model are shown in **Figure 13.1.** The lab biomarker model improved upon the simple model by correctly moving those with AF events up a risk category and also by correctly moving those without an AF event down a risk category. The addition of proteomics variables increased the c-statistic to 0.795 (95% CI, 0.773-0.816) while still showing excellent calibration (χ^2 = 7.6; P = 0.58). NRI and IDI showed significant improvements over both the simple model and the lab biomarker model. The NRI tables in **Figure 13.1** indicates the proteomics variables correctly moved up those with AF events and moved down those without AF. Finally, the addition of abnormal P wave axis, LA diameter, and septal E/e prime showed a modest improvement in the c-statistic to 0.806 (95% CI, 0.785-0.827), and

indicated a significant improvement in NRI over the simple model and also a slight improvement over the proteomics model. **Figure 13.2** compares the NRI of each of the complex models to one another.

The beta coefficients, baseline survival, and ARIC mean variable values used to derive the prediction models are listed in **Table 13.4.** The formula to calculate the 5-year risk of incident AF based on these variables is included as a table footnote.

Finally, validation c-statistics for each model using internal bootstrapping with adjustment for optimism are listed in **Table 13.3**. The adjusted c-statistic of 0.692 for the simple model and 0.737 for lab biomarkers model, which were only slightly lower than our derived c-statistics, indicated that these scores would perform well in individuals from populations similar to the ARIC cohort. The validated c-statistic for the proteomics model was 0.784 and for the ECG and echo model was 0.795, and indicated that while there was most likely some overfitting of our derivation models, the validated predictive value in both scores remained good to excellent.

13.5. DISCUSSION

In this community-based prospective population study of older adults, we created a simple risk model calibrated to this age group (66-90 years old) that included variables routinely collected in a primary care setting and predicted future risk of AF. We incrementally added variables to this model including lab biomarkers, proteomics, ECG and echocardiographic measures and developed more complex models that increased the discrimination ability of each model. The derived models performed well using internal validation.

The need for the accurate prediction of AF has given rise to the development of several population-based prediction equations.[117] Risk scores for AF have been developed in the Framingham Heart Study (FHS),[85] ARIC,[51] the Women's Health Study,[199] and the CHARGE-AF consortium.[10] The CHARGE-AF risk score is a 5-year predictive model that used pooled data from 18,556 participants from ARIC, FHS, and the Cardiovascular Health Study (CHS), and was validated in the Age, Gene/Environment Susceptibility Reykjavik study (AGES) and the Rotterdam Study (RS), and in a separate study was again validated in MESA.[198] Investigators developed a simple model, which incorporated common clinically-measured variables, and an augmented model, which incorporated additional ECG measures and blood tests. The advanced-age cohort CHS (mean age 73) was used in deriving the prediction models, however the model did not perform as well in the elderly AGES cohort

(mean age 76) as it did in the pooled derivation sample; c-statistic of the simple model = 0.664 in AGES vs. 0.765 in the pooled derivation cohort. Similarly, when we applied the simple CHARGE-AF score to our older ARIC cohort, the c-statistic was 0.660 with a poor calibration of 27.4 (p=0.001) indicating that a prediction model that was more accurate in the elderly was warranted.

Since the publication of the CHARGE-AF risk score in 2013, additional variables have been associated with an increased risk of incident AF and most notably, the addition of basal levels of the biomarker NT-proBNP significantly increased the predictive value.[18] In that study conducted by Sinner et al, participants of the AGES and RS cohorts were markedly older than participants of the other cohorts and the authors observed the predictive performance of NT-proBNP was better in these two older cohorts compared to the younger participants, presumably reflecting a higher prevalence of subclinical disease in AGES and RS. In our study of AF-free older adults, elevated levels of NT-proBNP, even once adjusted for overt heart disease, predicted AF. Another biomarker, troponin T, is highly predictive for myocardial damage and is the pathological hallmark of acute MI or myocardial injury. Basal troponin T concentration has been associated with incident AF in several cohorts but it has not added predictive value in existing AF prediction scores.[16, 18] However, in our analysis in this elderly cohort, troponin T remained as a significant predictor of incident AF, even after the addition of proteomics to the prediction model. Based on the results from our study, there may be significant clinical utility in measuring NT-proBNP and troponin T to determine risk of AF in older individuals.

This is the first study to determine the predictive value of adding novel protein measures to an AF risk prediction score. Importantly, all protein measures in our final score were predictive of incident AF independent of NT-proBNP, troponin T and independent of each other. The protein with the strongest independent predictive value of AF was Antileukoproteinase, also known as secretory leukocyte protease inhibitor (SLPI), which modulates the inflammatory and immune responses. SLPI functions as a non-redundant alarm anti-protease and is considered important in the defense against proteolytic attack from liberated granulocyte proteases.[261] Apart from its anti-protease activity, SLPI has antibacterial, antiviral, and anti-inflammatory properties and promotes wound healing.[262] Immunoassays exist for several, but not all of the proteomic measures significant in our study. Moving forward, strategies to streamline the proteomic variables into clinically useful measures would include developing cost-effective and time-efficient assays. Another approach could be to develop disease-specific proteomic chips that measure pre-determined proteins in a time-efficient manner. For example, a chip that included the 16 protein measures included in our model would aid in prediction of AF in older populations. Alternatively, with

the uptick of proteomics measures applied to risk prediction scores, a single plasma sample could be processed to provide individual health information and simultaneously predict the future risk of diseases such as diabetes[246] and cardiovascular diseases.[245]

Abnormal P wave axis is associated with increased risk of AF and has added predictive value to the CHARGE-AF risk score in a previous study.[251] Adverse atrial remodeling is associated with increased risk of AF and can be detected by this shift in the P wave axis. Left atrial diameter has been independently associated with incident AF in several cohorts.[44, 91, 92] Increased left atrial size has been thought to increase AF risk as a result of stretch of the atrial appendage which leads to remodeling of the anatomy and physiology of the left atrium and increases dispersion of atrial refractoriness. Septal E/e prime, a measure of diastolic dysfunction, has been associated positively with AF,[263] along with higher NT-proBNP, incident HF and death,[264] but its addition to a risk prediction score has not been evaluated until now. These 3 measures added modest value to an AF prediction score for older ages that already included clinical, biomarker, and proteomic variables.

The main strength of this study is the plethora and quality of candidate predictor variables, including proteomics data, available for AF prediction in this older cohort. Additionally, this study included black and white men and women from a community sample followed 5 years after baseline with nearly 400 AF events. Nevertheless, our findings need to be evaluated in the context of limitations in the study design. Foremost, our prediction model was not able to be externally validated in a comparable cohort at this time. As proteomics become more widespread, studies should attempt to replicate and validate our results. Our cohort contained individuals from 66 to 90 years of age and results may not be generalizable to those outside this age range, and similarly, might not be generalizable to those individuals with a race other than white or black. Next, incident AF was identified mainly from hospitalization discharges, and we could be missing asymptomatic AF or AF managed exclusively in an outpatient setting. However, we and others have previously shown that the validity of AF ascertainment using hospitalizations is acceptable, and that incidence rates of AF in the ARIC study are consistent with other population-based studies.[9, 44] Additionally, we will be unable to classify AF type (paroxysmal, persistent, or permanent AF) or assess the burden of AF (the percentage of time a person is in AF) in the ARIC study. It is currently unknown if the predictive variables differ by AF type or by AF burden.

In conclusion, we have developed a series of risk prediction models for the prediction of AF in older adults. The simple model uses information readily available in a primary care setting and includes variables more tailored and calibrated than previous prediction models to older individuals. The more complex models, which include blood biomarkers, proteomics,

ECG, and echocardiographic variables, greatly improve AF prediction. These well-calibrated risk scores can optimize screening in high-risk older individuals, allow for more specific clinical trial enrollment, and can lead to opportunities for targeted preventive strategies. Future research should replicate our study results and should also develop simple and cost-effective ways to quantify and evaluate novel protein measures so they might be readily included in AF prediction.

Table 13.1. Selected baseline characteristics of ARIC participants measured in 2011-2013, stratified by incident AF status within a 5-year follow-up period

	No incident AF (n=3914)	Incident AF (n=394)
Age	75.0 (4.9)	77.0 (5.4)
Female sex	2303 (59%)	204 (52%)
Black race	761 (19%)	49 (12%)
Height, cm	165.5 (9.4)	166.6 (10.0)
Weight, kg	78.4 (17.1)	80.6 (17.6)
Current smoker	226 (6%)	24 (6%)
Current drinker	1982 (51%)	195 (49%)
Systolic BP, mmHg	130.0 (17.6)	129.8 (19.4)
Diastolic BP, mmHg	66.3 (10.4)	64.1 (11.3)
Beta-blocker use	1164 (30%)	194 (49%)
Diuretic use	1217 (31%)	121 (31%)
Other antihypertensive medication use	2390 (61%)	262 (67%)
Diabetes	1195 (31%)	131 (33%)
Heart failure	112 (3%)	37 (9%)
Myocardial infarction	411 (11%)	81 (21%)
Stroke	108 (3%)	15 (4%)
Statin use	2022 (52%)	223 (57%)
Antiarrhythmic use	10 (0.3%)	11 (2.8%)
Anticoagulant use	74 (1.9%)	23 (5.8%)
eGFR mL/min per m2	67.0 (17.5)	60.8 (18.2)
C-reactive protein, mg/L, Median (25-75%)	1.9 (0.9-4.1)	1.9 (1.0-4.2)
Ln (c-reactive protein), mean (SD)	0.7 (1.1)	0.8 (1.1)
Troponin-T, ng/L, Median (25-75%)	10 (7-15)	13 (8-21)
Ln (troponin-T), mean (SD)	2.3 (0.6)	2.6 (0.7)
NT-proBNP, pg/mL, Median (25-75%)	111 (59-211)	237 (130-451)
Ln (NT-proBNP), mean (SD)	4.7 (1.1)	5.5 (1.2)
Abnormal P-wave axis (<0 or >75)	430 (11%)	74 (19%)
Left atrial diameter, cm	3.5 (0.5)	3.8 (0.5)
Septal E/e-prime, cm/sec	9.5 (1.9)	8.4 (2.1)

*Values correspond to mean ± standard deviation or N (%) unless indicated
eGFR = estimated glomerular filtration rate; NT-proBNP = N-terminal pro B-type natriuretic peptide

Table 13.2. Hazard Ratios (95% confidence intervals) for the final variables included in the derived multivariable models for the prediction of the 5-year risk of incident AF, derived in ARIC.

Variables	Simple model	+ Lab biomarkers	+ Proteomics	+ ECG and echo
Age, per 5 years	1.43 (1.29-1.57)	1.21 (1.09-1.34)	1.14 (1.02-1.27)	1.14 (1.02-1.27)
White race	1.68 (1.24-2.27)	1.75 (1.28-2.38)	1.36 (0.97-1.92)	1.29 (0.91-1.83)
Weight, per 15 kg	1.14 (1.05-1.24)	1.14 (1.05-1.24)	1.17 (1.08-1.28)	1.06 (0.96-1.16)
Prevalent heart failure	2.27 (1.59-3.25)	1.56 (1.08-2.25)	1.44 (1.00-2.07)	1.33 (0.92-1.92)
Prevalent myocardial infarction	1.50 (1.16-1.95)	1.19 (0.91-1.55)	1.25 (0.96-1.64)	1.15 (0.88-1.51)
Anti-arrhythmic agent use	5.79 (3.05-11.01)	5.21 (2.76-9.81)	5.35 (2.76-10.4)	4.76 (2.39-9.50)
Beta-blocker use	1.72 (1.39-2.12)	1.41 (1.14-1.74)	1.29 (1.04-1.61)	1.14 (0.92-1.43)
Anticoagulant use	2.38 (1.52-3.74)	2.11 (1.36-3.28)	2.23 (1.43-3.47)	2.16 (1.37-3.40)
eGFR, per 10 mL/min per m2		1.06 (1.00-1.13)	1.16 (1.08-1.25)	1.13 (1.05-1.22)
Ln (Troponin-T), per 0.65		1.28 (1.14-1.44)	1.29 (1.15-1.46)	1.26 (1.12-1.43)
Ln (NT-proBNP), per 1.04		1.70 (1.52-1.91)	1.19 (1.03-1.37)	1.16 (1.00-1.33)
Proteomics measures (per doubling)				
NT-proBNP			1.50 (1.29-1.74)	1.29 (1.10-1.50)
CLM2			1.50 (1.24-1.83)	1.44 (1.18-1.77)
ID-1			1.46 (1.27-1.69)	1.49 (1.29-1.72)
CV015			0.60 (0.45-0.89)	0.60 (0.45-0.80)
SIA10			0.74 (0.62-0.89)	0.77 (0.65-0.92)
NBR1			1.63 (1.25-2.13)	1.58 (1.20-2.07)
MIA			0.47 (0.32-0.69)	0.47 (0.32-0.69)
EMIL3			0.59 (0.44-0.80)	0.62 (0.47-0.82)
PGP			0.55 (0.40-0.77)	0.57 (0.41-0.80)
SAP18			1.40 (1.15-1.70)	1.31 (1.08-1.60)
sICAM-5			1.45 (1.18-1.79)	1.44 (1.16-1.78)
SLPI			2.24 (1.55-3.25)	2.12 (1.47-3.05)
PAP1			1.29 (1.10-1.50)	1.28 (1.10-1.50)
GLP1R			0.53 (0.36-0.79)	0.55 (0.36-0.82)

CD244	0.65 (0.47-0.91)	0.65 (0.47-0.91)
RAP2A	0.87 (0.78-0.97)	0.86 (0.77-0.96)
Abnormal P wave axis (<0 or >75)		1.42 (1.09-1.85)
Left atrial diameter, per 0.5 cm		1.29 (1.14-1.45)
Septal E/e prime, per 1.89 cm/s		0.77 (0.69-0.86)

CLM2=CMRF35-like molecule 2, ID-1=DNA-binding protein inhibitor ID-1, CV015=Uncharacterized protein C22orf15, SIA10=Type 2 lactosamine alpha-2,3-sialyltransferase, NBR1=Next to BRCA1 gene 1 protein, MIA=Melanoma-derived growth regulatory protein, EMIL3=EMILIN-3, PGP=Glycerol-3-phosphate phosphatase, SAP18=Histone deacetylase complex subunit SAP18, sICAM-5=Intercellular adhesion molecule 5, SLPI=Antileukoproteinase, PAP1=Regenerating islet-derived protein 3-alpha, GLP1R=Glucagon-like peptide 1 receptor, CD244=Natural killer cell receptor 2B4, RAP2A=Ras-related protein Rap-2a. The model derivation included 4308 participants and 394 AF events

Table 13.3. Model Discrimination and Calibration for the derived and validated models for the prediction of the 5-year risk of incident AF

	Derivation Models			
	Simple model	+ Lab biomarkers	+ Proteomics	+ ECG and echo
C-statistic (95% CI)	0.697	0.742	0.795	0.806
	(0.671-0.723)	(0.717-0.767)	(0.773-0.816)	(0.785-0.827)
Calibration chi-square (p-value)	9.4 (0.40)	16.0 (0.07)	7.59 (0.58)	9.44 (0.40)
Net reclassification index	0 (Ref)	0.233 (0.181, 0.285)	0.446 (0.386, 0.501)	0.506 (0.445, 0.567)
	--	0 (Ref)	0.246 (0.191, 0.295)	0.289 (0.232, 0.332)
	--	--	0 (Ref)	0.053 (0.012, 0.095)
Discrimination slope (IDI)	0 (Ref)	0.469 (0.313, 0.641)	1.17 (0.880, 1.51)	1.29 (0.960, 1.67)
	--	0 (Ref)	0.479 (0.340, 0.639)	0.559 (0.393, 0.741)
	--	--	0 (Ref)	0.054 (-0.005, 0.108)
	Validation Models (internal validation and adjusted for optimism)			
	Simple model	+ Lab biomarkers	+ Proteomics	+ ECG and echo
C-statistic (95% CI)	0.692	0.737	0.784	0.795
	(0.666-0.718)	(0.712-0.762)	(0.762-0.805)	(0.774-0.816)

Table 13.4. Beta estimates, baseline survival and mean variable values used for the prediction scores for the prediction of the 5-year risk of incident AF, derived in ARIC.

	Simple model	+ Lab biomarkers	+ Proteomics	+ ECG and echo	ARIC mean value
Survival	0.91896	0.92857	0.94196	0.94387	
Variables (beta estimates)					
Age, per 5 years	0.35407	0.19056	0.12795	0.12870	15.0417
White race	0.51718	0.55699	0.30730	0.25549	0.81198
Weight, per 15 kg	0.13197	0.13049	0.15899	0.05758	5.23923
Prevalent heart failure	0.82179	0.44558	0.36284	0.28161	0.03459
Prevalent myocardial infarction	0.40775	0.17172	0.22567	0.14399	0.11421
Anti-arrhythmic use	1.75647	0.34142	1.67728	1.56093	0.00487
Beta-blocker use	0.54187	1.64969	0.25505	0.13361	0.31523
Anticoagulant use	0.86877	0.74690	0.79968	0.76882	0.02252
eGFR, per 10 mL/min per m2		0.05866	0.14888	0.12295	6.64459
Ln (Troponin-T), per 0.65		0.24466	0.25614	0.23239	3.59721
Ln (NT-proBNP), per 1.04		0.53066	0.16943	0.14571	4.56756
Proteomics measures (log base 2 transformed)					
NT-proBNP			0.40584	0.25314	12.0392
CLM2			0.40813	0.36595	9.02665
ID-1			0.37848	0.39687	10.1221
CV015			-0.51101	-0.51591	7.90483
SIA10			-0.29686	-0.26044	12.7101
NBR1			0.48899	0.45704	10.1393
MIA			-0.75384	-0.76046	11.3157
EMIL3			-0.51945	-0.47931	9.62080
PGP			-0.59011	-0.55364	10.5765

SAP18	0.33548	0.27254	9.82916
sICAM-5	0.37498	0.36270	10.9893
SLPI	0.80857	0.74999	11.6821
PAP1	0.25085	0.24938	13.4666
GLP1R	-0.63018	-0.60502	9.11573
CD244	-0.42839	-0.42880	6.80953
RAP2A	-0.14089	-0.14882	9.88184
Abnormal P wave axis (<0 or >75)		0.34778	0.11699
LA diameter, per 0.5 cm		0.25340	7.02331
Septal E/e prime, per 1.89 cm/s		-0.26314	4.99017

The 5-year risk of incident AF can be calculated as 1 – survival $^{\exp(\text{beta1}*(\text{X1-mean value1})+(\text{beta2}*(\text{X2-mean value2})+\ldots.)}$ where survival is model-specific, beta is the regression coefficient, X is the level for each risk factor, and mean value is ARIC specific.

When calculating the risk, variable values must be divided by the number of units listed in the table, or for protein measures, log base 2 transformed and interpreted as per doubling of the measure. The model derivation included 4308 participants and 394 AF events

Figure 13.1. The categorical net reclassification improvement tables comparing each model to the simple model, stratified by AF status. Risk categories are <5%, 5-15%, and >15%. Shaded boxes indicate correct reclassification and patterned boxes indicates incorrect classification.

A. Comparing the lab biomarker model to the simple model

Simple Model	Lab biomarkers model				
	<5%	5-15%	>15%	Total
<5%	158	43	2	203
5-15%	22	100	29	151
>15%	0	5	35	40
Total	180	148	66	394
Participants with AF in 5 years				

Simple Model	Lab biomarkers model			
	<5%	5-15%	>15%	Total
<5%	846	207	1	1054
5-15%	709	1543	222	2474
>15%	9	140	237	386
Total	1564	1890	460	3914
Participants without AF in 5 years				

B. Comparing the proteomic model to the simple model

Simple Model	Proteomics model			
	<5%	5-15%	>15%	Total
<5%	136	46	21	203
5-15%	20	82	49	151
>15%	0	6	34	40
Total	156	134	104	394
Participants with AF in 5 years				

Simple Model	Proteomics model			
	<5%	5-15%	>15%	Total
<5%	818	221	15	1054
5-15%	1148	1031	295	2474
>15%	42	157	187	386
Total	2008	1409	497	3914
Participants without AF in 5 years				

C. Comparing the ECG and echo model to the simple model

Simple Model	ECG and echo model			
	<5%	5-15%	>15%	Total
<5%	130	56	17	203
5-15%	24	73	54	151
>15%	0	3	37	40
Total	154	132	108	394
Participants with AF in 5 years				

Simple Model	ECG and echo model			
	<5%	5-15%	>15%	Total
<5%	824	209	21	1054
5-15%	1248	946	280	2474
>15%	53	165	167	386
Total	2125	1320	469	3914
Participants without AF in 5 years				

Figure 13.2. The categorical net reclassification improvement tables comparing each complex model, stratified by AF status. Risk categories are <5%, 5-15%, and >15%. Shaded boxes indicate correct reclassification and patterned boxes indicates incorrect classification.

A. Comparing the proteomics model to the lab biomarker model

Lab biomarker	Proteomics model			
	<5%	5-15%	>15%	Total
<5%	138	36	6	180
5-15%	18	87	43	148
>15%	0	11	55	66
Total	156	134	104	394
Participants with AF in 5 years				

Lab biomarker	Proteomics model			
	<5%	5-15%	>15%	Total
<5%	1319	235	10	1564
5-15%	669	1021	200	1890
>15%	20	153	287	460
Total	2008	1409	497	3914
Participants without AF in 5 years				

B. Comparing the ECG and echo model to the lab biomarker model

Lab biomarker	ECG and echo model			
	<5%	5-15%	>15%	Total
<5%	129	42	9	180
5-15%	24	83	41	148
>15%	1	7	58	66
Total	154	132	108	394
Participants with AF in 5 years				

Lab biomarker	ECG and echo model			
	<5%	5-15%	>15%	Total
<5%	1302	248	14	1564
5-15%	784	900	206	1890
>15%	39	172	249	460
Total	2125	1320	469	3914
Participants without AF in 5 years				

C. Comparing the ECG and echo model to the proteomics model

Proteomics	ECG and echo model			
	<5%	5-15%	>15%	Total
<5%	139	17	0	156
5-15%	15	99	20	134
>15%	0	16	88	104
Total	154	132	108	394
Participants with AF in 5 years				

Proteomics	ECG and echo model			
	<5%	5-15%	>15%	Total
<5%	1854	154	0	2008
5-15%	271	1050	88	1409
>15%	0	116	381	497
Total	2125	1320	469	3914
Participants without AF in 5 years				

Supplemental Methods:

Candidate variables

The following variables from ARIC visit 5 were considered for inclusion in our prediction model, and can be found listed in Supplemental Table 1. We considered variables in the CHARGE-AF model that includes age, race, height, weight, current cigarette smoking, systolic and diastolic blood pressure, use of antihypertensive medication (split into categories of beta-blockers, diuretics and others), diabetes, heart failure, and history of myocardial infarction (MI).[10] We considered variables included in an augmented CHARGE-AF model including ECG markers (PR interval (<120; 120-199; > 200), left ventricular hypertrophy (LVH): gender-specific Cornell voltage criteria (SV3 + RaVL > 2.8mV for men, and >2.2mV for women)), and more recent augmented models include NT-proBNP and C-reactive protein.[10, 18, 198] Some clinical variables that were evaluated and not included in the final CHARGE-AF score were re-evaluated in this study as their predictive value may differ in the elderly. These included sex, fasting blood glucose, estimated glomerular filtration rate (eGFR), total cholesterol, HDL cholesterol, physical activity, triglycerides, alcohol consumption, use of lipid-lowering medications, heart rate, history of coronary artery bypass graft (CABG), history of stroke, blood troponin, and the following medications use: statins, anticoagulants, anti-arrhythmic agents, cardiac glycosides, and aspirin use. After literature review for associations with AF, we also considered the ECG and echocardiographic variables listed in Supplemental Table 1. If clinical cut-points existed for the measure, we looked at the cut-points and also at the continuous measure.

Measures of candidate variables

Procedures for measures in ARIC have been published.[209] In brief, participants reported information on smoking, history of cardiovascular disease, use of medications, and underwent a physical exam at the visit that included height and weight. Seated blood pressure was measured using a random-zero sphygmomanometer after 5 minutes rest, and was defined as the average of the 2nd and 3rd measurements taken. We defined diabetes mellitus was fasting glucose ≥ 126 mg/dL (7.0 mmol/L), non-fasting glucose ≥ 200 mg/dL (11.1 mmol/L), treatment for diabetes mellitus, or self-reported physician diagnosis of diabetes. MI was ascertained by study visit ECGs or the ARIC Morbidity and Mortality Classification Committee, by using data from follow-up calls, hospitalization records and death certificates.[210] Prevalent HF was defined as the reported use of HF medication in the previous two weeks, presence of HF according the Gothenburg criteria (only at the baseline ARIC visit), or having had a HF hospitalization during follow-up.[248,249] Plasma creatinine and

cystatin C were measured, and eGFR was calculated as mL/min/1.73 m^2 using the CKD Epidemiology Collaboration (CKD-EPI) combined creatinine-cystatin C equation.[250] Participants underwent an ECGs during each clinical exam and details for each measure have been described.[211] We defined abnormal P wave axis as any value outside of 0 to 75 degrees. Echocardiograms were obtained in all centers during visit 5 by certified study sonographers using uniform imaging equipment and following image acquisition protocol. Methods and details for each measure have been described.[265]

Proteomic profiling

EDTA-plasma was obtained from blood samples that were collected at visit 5 and stored at -80 degrees C. Plasma samples were analyzed using a SOMAmer-based capture array called "SomaScan" (SomaLogic, Inc., Boulder, CO, USA). This assay was performed as described previously.[216-219] Protein levels in the plasma samples were measured by the SomaScan platform, which uses single-stranded DNA-based aptamers to capture conformational protein epitopes. Of the 5284 available measures, we excluded 94 that had a CV_{BA} >50% or a variance of < 0.01 on the log scale at visit 5. Additionally, we excluded 313 because of binding to non-proteins, including hybridization control elution, non-human proteins, non-biotin, non-cleavable, and spuriomer products. For each measure, we winsorized outliers that were greater or less than 5 standard deviations from the sample mean on the log 2 scale. After all quality control measures were completed, 4877 SOMAmers which recognize 4697 unique human proteins or protein complexes were analyzed in this study.

In a previous study on a subset of ARIC participants, we validated the measurement of several aptamers compared with immunoassays in the ARIC central laboratory. SomaScan and traditional immunoassay measurements were highly correlated: NT-proBNP (n=5168, r=0.90).The immunoassay measure was log-transformed and interpreted as the increase in 1 SD of the log transformed measure (1.04).

Derivation of the predictive models

Derivation of the simple model

We first identified candidate predictors for the simple model by adjusting for age and sex. Variables significantly associated with incident AF ($p<0.05$) were then considered candidate predictors and selected for inclusion in the next step Next, we generated 1000

bootstrap samples and ran Cox proportional hazards models with backward selection of the candidate predictors in each of the 1000 samples. Based on recommendations from the literature, variables included in at least 60% of the Cox models were selected for inclusion in the final predictive model.[257] We tested age and race interactions for inclusion in the final model and none significantly improved the prediction of the model.

Derivation of the lab biomarker and proteomics models

Next, we wanted to develop a prediction model that would be useful in a clinical setting that included in a blood draw. Using similar steps as above, we ran a Cox model adjusting for age and sex with each lab variable and variables that were associated with incident AF at the $p<0.05$ were then considered candidate predictors and allowed into the next step. Then we generated 1000 bootstrap samples and ran Cox proportional hazards models that included adjustment for the simple model variables and had backward selection of the candidate predictor variables in each of the 1000 samples. We forced the variables from the simple model to stay in the model during backwards selection in order to determine what value, if any, the biomarkers added. Again, variables included in at least 60% of the Cox models were selected for inclusion in our prediction model.

We developed a proteomics model to determine to what extent adding proteomics variables would enhance the prediction. We started with 4877 possible proteomics variables, so we followed similar steps as above, but with slightly different criteria. Instead of just adjusting for age and sex, our first Cox model adjusted for all of the variables that were already included in the lab biomarker prediction model. We used a threshold of a p-value<0.005 to determine candidate predictors to move onto the next step. We found 63 proteins met this threshold and then became candidate predictors. Then we continued with the same steps as above, generating 1000 bootstrap samples, and ran Cox proportional hazards models that included adjustment for the simple model variables and the biomarkers above, with backward selection of the candidate predictor variables. We forced the variables from the lab biomarker model to stay in the model during backward selection. Variables included in at least 60% of the Cox models were selected for inclusion in the prediction model.

Derivation of the ECG and Echo model

Finally, using similar steps as above, we determined if any candidate variables from ECGs or echocardiograms added value to the prediction model. We ran a Cox model adjusting for the variables included in the simple model and included those with a p-value <0.05 as candidate predictors. Then we generated 1000 bootstrap samples and ran Cox models that included adjustment for variables included in prediction models thus far and

included backward selection of the candidate predictor variables in each of the 1000 samples. We forced the variables from the previous model to stay in the model during backwards selection in order to determine what value, if any, these ECG and echo variables added. Again, variables included in at least 60% of the Cox models were selected for inclusion.

Table 13.5. Supplemental Table. Candidate variables considered for inclusion in AF prediction models, measured at ARIC visit 5, 2011-2013

Clinical variables	**Lab biomarker variables**
Age, per 5 years	Blood glucose,* per 10 mg/dL
Sex, male vs female	eGFR <60 mL/min per m^2, vs ≥60
White race vs black	Total cholesterol,* per 40 mg/dL
Height, per 10 cm	HDL cholesterol,* per 15 mg/dL
Weight, per 15 kg	Triglycerides,*per 40 mg/dL
Current smoker vs non smoker	Log (hsCRP), per 1.43 (1 unit ln-transformed)
Current drinker vs. not current	Log (NT-proBNP), per 1.04 (1 unit ln-transformed)
Amount of alcohol	Log (troponin-T), per 0.65 (1 unit ln-transformed)
Systolic BP, per 20 mm Hg	
Diastolic BP, per 10 mm Hg	**Proteomic variables (log base 2 transformed, interpreted as per doubling of the measure**
Diabetes	N-terminal pro-BNP
CABG history	CMRF35-like molecule 2
Heart failure history	Origin recognition complex subunit 6
Myocardial infarction history	DNA-binding protein inhibitor ID-1
Stroke history	Netrin receptor UNC5D
Statin use	RGM domain family member B
Beta-blocker use	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1a
Diuretic use	Protein delta homolog 1
Other Antihypertensive medication use	Uncharacterized protein C22orf15
Statin use	Relaxin-3
Anticoagulant use	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1b
Antiarrhythmic agent use	DnaJ homolog subfamily A member 2
Aspirin use	Protein delta homolog 1
Cardiac glycosides	Gamma-aminobutyric acid receptor-associated protein-like 1
	Synaptotagmin-4
ECG variables	Receptor-type tyrosine-protein phosphatase delta
Heart rate, per 10 bpm	Delta and Notch-like epidermal growth factor-related receptor
Left ventricular hypertrophy, Cornell criteria	Proteasome subunit alpha type-5
QRS duration, per 20	Interleukin-18
QRS duration <90 (ref) vs. 90-120	Growth/differentiation factor 11/8
<90 (ref) vs. >120	Neutral and basic amino acid transport protein rBAT
QT interval, per 31	Smoothelin
QT interval ≥ 440 m, ≥ 460F	Type 2 lactosamine alpha-2,3-sialyltransferase
PR interval, per 30 ms	Leukotriene B4 receptor 1
PR interval (<120 vs. 120-199)	Next to BRCA1 gene 1 protein
PR interval (>199 vs. 120-199)	BH3-interacting domain death agonist
Abnormal P-wave axis (<0 or >75)	Actin-related protein 2/3 complex subunit 3
P-wave duration > 120	Ubiquitin-like protein ISG15
P-wave terminal force > 4000 uV.ms	5'-nucleotidase domain-containing protein 3
Advanced interarterial block	Melanoma-derived growth regulatory protein
	Attractin
Echocardiogram variables	MOB kinase activator 1A
LA diameter, per 0.5 cm	Shadow of prion protein

LA diameter : >4.0cm M and > 3.7cm W	EMILIN-3
LA volume index, per 8 mL/m2	GDNF family receptor alpha-1
LA volume index: ≥ 34mL/m2	Vesicle transport through interaction with t-SNAREs homolog 1A
E / E prime lateral ratio, per 3.9 cm/sec	Neuronal growth regulator 1
E / E prime lateral ratio: >11.5 M and > 13.3 W	Ephrin type-A receptor 3
E/A ratio, per 0.27	NmrA-like family domain-containing protein 1
Ejection fraction (<50%)	Tyrosine-protein kinase Fyn
LV mass index, per 20 g/m2	Glycerol-3-phosphate phosphatase
LV mass index: >115 M and > 95 W	Amyloid beta A4 precursor protein-binding family B member 3
LV diastolic diameter, per 0.5 cm	Macoilin
LV diastolic diameter (>5.8 M and >5.2 W)	Neurocan core protein
LV relative wall thickness, per 0.07	Histone deacetylase complex subunit SAP18
Mean LV wall thickness, per 0.13 cm	DnaJ homolog subfamily C member 4
RV fractional area change, per 0.08	Glutamate receptor ionotropic, delta-2
Septal E prime, per 1.45 cm/s	Intercellular adhesion molecule 5
Septal E / E prime, per 1.89 cm/s	Antileukoproteinase
	Keratocan
	Regenerating islet-derived protein 3-alpha
	Lysosomal Pro-X carboxypeptidase
	Vesicle-fusing ATPase
	Glucagon-like peptide 1 receptor
	SLIT and NTRK-like protein 1
	Natural killer cell receptor 2B4
	Cholinesterase
	Casein kinase I isoform gamma-2
	Hepatocyte nuclear factor 4-alpha
	Torsin-1A-interacting protein 1
	Ecto-ADP-ribosyltransferase 3
	Eukaryotic initiation factor 4A-II
	Ras-related protein Rap-2a

Figure 13.3. Supplemental Figure: TRIPOD statement

NUMBERS Correspond to Page number of manuscript

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	3
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	3-4
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	4
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	4
	5b	D;V	Describe eligibility criteria for participants.	4
	5c	D;V	Give details of treatments received, if relevant.	N/A
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	4
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	4-5
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	8	D;V	Explain how the study size was arrived at.	4
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	4
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	5-6
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	5-6
	10c	V	For validation, describe how the predictions were calculated.	6-7
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	6
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	N/A
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	N/A
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	7
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	7
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	N/A
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	7
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	7
	15b	D	Explain how to the use the prediction model.	8
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	7-8
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	N/A
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	10-11
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	10
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	9-11
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	9-11
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Supp
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	12

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

14. Manuscript 3 – Direct oral anticoagulants and warfarin for atrial fibrillation treatment: Rural and Urban trends in Medicare beneficiaries from 2011-2016

14.1. OVERVIEW

Background – Despite a higher risk of stroke in rural areas of the US compared to urban areas, there is little known regarding oral anticoagulation rates in atrial fibrillation (AF) patients in rural vs. urban areas. Furthermore, no data has been published addressing initiation of the recently-approved direct oral anticoagulants (DOACs) by patients' rurality. We used Medicare data to examine the initiation of anticoagulation use in newly-diagnosed AF patients in rural versus urban areas.

Methods –We identified incident AF in a 20% sample of fee-for-service Medicare beneficiaries from 2011-2016, and collected beneficiary residential zip code and covariates at the time of AF diagnosis. We identified the first anticoagulant prescription filled, if any, following AF diagnosis. We categorized beneficiaries into 4 rural/urban areas by linking zip code to rural-urban commuting area codes and used Poisson regression models to compare anticoagulant prescription fills for patients in rural vs. urban areas.

Results – Our study included 447,252 patients with diagnosed AF (mean age 79 ± 8 years) in which 82% were categorized as urban, 9% large rural, 5% small rural and 4% isolated. The percentage of those who initiated an anticoagulant rose from 34% in 2011 to 53% in 2016, driven by the uptake of DOACs. There were clear gradients of anticoagulant use by rurality. In a multivariable-adjusted analysis of beneficiaries matched by rural / urban category, those in rural areas were more likely to initiate an anticoagulant; those in isolated areas were 7% more likely (95% CI = 4-10%) compared to those in urban areas. However, those in rural areas were less likely to receive a DOAC; those in isolated areas were 18% less likely (95% CI = 15 to 22%) to initiate a DOAC compared to those in urban areas.

Conclusion – In this Medicare population with AF, anticoagulation use was low but has increased over time due to the introduction of DOACs. Those in rural areas were less likely to receive a DOAC compared to those in urban areas, with the lowest DOAC use occurring in the most isolated areas. Increasing the percentages of DOAC use in AF patients living in rural areas may reduce the burdens of stroke and healthcare utilization of older adults in rural areas.

14.2. INTRODUCTION

Individuals with atrial fibrillation (AF), a common cardiac arrhythmia, have a 5-fold increased risk of stroke compared to those without AF and therefore the mainstay of stroke prevention in AF is the initiation and maintenance of anticoagulant therapies.[8] The oral anticoagulants currently recommended include warfarin and a class of direct oral anticoagulants (DOACs). Since 2010, the Food and Drug Administration (FDA) has approved 4 DOACs for stroke prevention in AF, including the direct thrombin inhibitor dabigatran, and the direct factor Xa inhibitors, rivaroxaban, apixaban, and edoxaban. The DOACs have fewer drug interactions, more predictable pharmacological profiles, an absence of major dietary effects, and a reduced risk of intracranial bleeding and ischemic stroke compared with warfarin.[35] Currently, DOACs account for >50% of anticoagulants prescribed for AF patients and have directly contributed to the rising percentage of AF patients treated with anticoagulants.[173, 174]

There are nearly 60 million people (19% of the population) living in rural areas according to the US Census Bureau. Those in rural areas have higher rates of adverse cardiovascular risk factors such as cigarette smoking, hypertension, diabetes, obesity, coronary heart disease, and stroke.[26-28, 30] Despite the higher risk of stroke from AF, there is little known regarding anticoagulation rates in AF patients in rural vs. urban areas. Furthermore, no data have been published addressing the adoption of DOAC prescriptions in rural areas of the US. Differences in the initiation of anticoagulation and DOAC use by rural / urban status may identify an area of practice improvement for providers to reduce the burdens of stroke and healthcare utilization in a population of older adults.

Using a sample of Medicare beneficiaries, which included patient geographic location, we describe trends in oral anticoagulant prescription fills, including the initiation of the DOACs, in AF patients from 2011-2016. We also compared type of anticoagulation treatment in AF patients living in rural vs. urban areas.

14.3. METHODS

Study population

We conducted a retrospective study using health care utilization claims data from a 20% sample of Medicare beneficiaries from 2011-2016. We limited the cohort to beneficiaries receiving fee-for-service Medicare who were 65 years or older living in the US, and enrolled in a stand-alone Part D prescription drug plan. We included those continuously enrolled in traditional fee-for-service Medicare Parts A/B/D without supplemental coverage

for at least 90 days during 2011-2016. We required at least the first 90 days of a beneficiaries' follow-up time to be free of AF diagnosis codes and anticoagulation codes in order to 1) capture incident AF events, 2) capture the first anticoagulation prescription following an AF event, and 3) to serve as a run-in period to capture patient health information and comorbidities prior to an AF event. If a beneficiary enrolled in supplemental coverage we censored them at the time of supplemental enrollment. For this analysis, we required at least a 30-day follow-up period after AF diagnosis in order to allow an appropriate time window for the beneficiary to fill an anticoagulant prescription.

The initial sample included 910,649 AF patients aged 65 to 112 years. The exclusion flow chart is depicted in **Figure 14.1.** We excluded those with an AF diagnosis or prescription fill for an anticoagulant during the first 90 days of enrollment (n=412,076), those initiating edoxaban (due to small numbers; n=296), those with less than 30 days of follow-up (50,026), and those with a missing zip code or those with a zip code in a US territory (n=819). Our final analytic sample for the descriptive analysis overall was 447,252, and for the analyses comparing rural vs. urban, 210,953 of those were successfully matched.

Ascertainment of AF Patients

This analysis included patients age 65+ with at least one inpatient claim for AF or 2 outpatient claims for AF 7 to 365 days apart. AF claims were identified using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis codes 427.3, 427.31, and 427.32, and ICD-10-CM codes starting October 1, 2015 of I48.x in any position, which is a standard definition used in claims analysis.[45, 212] The validity of ICD-9-CM codes for the identification of AF has been well-established with a systematic review of studies showing a positive predicted value (PPV) of approximately 90% and a sensitivity of approximately 80%.[213] We defined the diagnosis date as the earlier of 1) the earliest discharge date for an inpatient claims, or 2) the earliest service date of the outpatient or physician claim. Consistent with prior research, 2 outpatient claims were required to diagnose outpatient AF in order to minimize the impact of rule-out diagnosis and to improve specificity.[45]

Defining rural and urban beneficiaries

We captured beneficiary zip code at the time of AF diagnosis. We mapped zip codes to Rural-Urban Communing Area (RUCA) codes, which are approximation codes developed by the University of Washington Research[179] and commonly used to define rural and urban

areas.[180] RUCA codes combine standard Census definitions with area commuting behaviors to capture functional and work relationships between regions.

We used a 4- category classification to access the rurality of beneficiaries: urban (RUCA codes 1-3, 4.1, 5.1, 7.1, 8.1, 10.1), large rural (RUCA codes 4.0, 4.2, 5.0, 5.2, 6.0, 6.1), small rural (RUCA codes 7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2) and isolated (RUCA codes 10, 10.2, 10.3, 10.4, 10.5, 10.6). In a secondary analysis, we reported rural-urban trends in oral anticoagulation use by splitting the US into 4 US Census Bureau Regions: Northeast (CT, ME, MA, NH, RI, VT, NJ, NY, PA), Midwest (IN, IL, MI, OH, WI, IA, KA, MN, MO, NE, ND, SD), South (DE, D.C., FL, GA, MD, NC, SC, VA, WV, AL, KY, MS, TN, AR, LA, OK, TX), West (AZ, CO, ID, NM, MT, UT, NV, WY, AK, CA, HI, OR, WA).

Anticoagulation treatment definitions

We identified filled prescriptions for oral anticoagulation using Part D pharmaceutical claims data which included the prescription fill date, the strength and number of days supplied. Beneficiaries were assigned to the first anticoagulant filled in either the 30 days prior to, and anytime following their first AF claim. We included prescriptions initiated for warfarin, dabigatran, rivaroxaban, and apixaban in this analysis. We excluded edoxaban users due to small numbers. Validity of warfarin claims in administrative databases is excellent with a sensitivity of 94% and a PPV of 99%.[266] Validation studies of DOAC claims have not yet been conducted.

Covariates

Using the Medicare datasets, we identified covariates prevalent at the time of AF diagnosis. Race was self-reported and we categorized it into the race categories of white, black, and other/unknown (due to small numbers). We defined pre-determined covariates based on inpatient, outpatient, carrier, and pharmacy claims using validated published algorithms.[120, 267, 268] These included demographic characteristics, comorbidities, and pharmacy prescription fills. Comorbidities of interest were ascertained with published algorithms from inpatient and outpatient claims and include prior stroke/transient ischemic attack(TIA), hemorrhagic stroke, heart failure, myocardial infarction, hypertension, diabetes, peripheral arterial disease, liver disease, kidney disease, chronic pulmonary disease, malignancies (except malignant skin neoplasm), metastatic cancer, history of bleeding, hematological disorders (anemia, coagulation defects), dementia, depression, and alcohol abuse.[267, 268] ICD codes for the comorbidity variables are listed in **Table 14.3** (Supplemental Table 1). We captured prescription fills for the following medication groups: clopidogrel,

angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, beta-blockers, calcium channel blockers, antiarrhythmics, and statins. We calculated the CHA_2DS_2-VASc score[269] at AF date and it consisted of congestive heart failure, hypertension, age (1 point for age 65-74; 2 points for ≥ 75), diabetes, prior stroke or TIA (2 points), vascular disease, and female sex. The HAS-BLED score[122] was calculated using the variables of hypertension, abnormal renal/liver function, stroke, bleeding history or disposition, elderly (age>65) and drugs/alcohol concomitantly; the variable International Normalized Ratio (INR) normally included in the HAS-BLED score was not available for this cohort. We used Medicare carrier claims datasets to identify provider specialty at outpatient visits. Beneficiaries who saw a cardiology provider within a predetermined period (30 days prior to or 90 days after AF diagnosis) were classified as the cardiology group, while patients seen exclusively by internal medicine, family practice, medical doctor, or unspecified multispecialty group were classified as primary care. Patients seen by a cardiologist were included in the cardiology provider group, regardless of a primary care visit.

Statistical analysis

We examined the anticoagulant prescription fill patterns in AF Medicare patients in rural vs. urban areas. Baseline characteristics at the time of AF diagnosis were compared between the 4 rurality groups. The proportion of patients with AF who filled oral anticoagulant prescriptions was evaluated graphically, first overall by year and quarter, and then in each rural/urban category by year. We also determined oral anticoagulant prescriptions within the CHA_2DS_2-VASc score by rural category.

To compare proportions of anticoagulants, we matched beneficiaries based on AF date (± 30 days), age (± 1 years), sex, and CHA_2DS_2-VASc score (± 0). One beneficiary from each of the 3 rural categories was matched with up to 2 beneficiaries in the urban category using a greedy matching algorithm. We used Poisson regression models with robust variance estimates to compute risk ratios (RR) and 95% confidence intervals (CI).[270] The model adjusted for age (continuous), race (white, black, other), sex, CHA_2DS_2-VASc (categorical, 0-9), HAS-BLED score (continuous), specialist care (cardiology: yes/no), and the additional covariates listed above and in Table 1.

We examined effect modification by sex, race, and age (<75, ≥75) by adding a multiplicative interaction term in the model. A sensitivity analysis was limited to AF patients who qualify for oral anticoagulants (CHA_2DS_2-VASc scores ≥ 2); due to the advanced age and poly-comorbidity of Medicare patients, we had to exclude a small percentage (<2%) of

beneficiaries for this analysis. We conducted an additional sensitivity analysis requiring a 180-day run-in time instead of 90 days.

14.4. RESULTS

After exclusion criteria were applied, our study included 447,252 AF patients (mean age 79 ± 8 years), in which 369,357 (83%) lived in an urban area, 38,167 (9%) lived in a large rural area, 21,934 (5%) lived in a small rural area, and 17,794 (4%) lived in an isolated rural area. Characteristics of the total cohort of AF patients are listed in **Table 14.1.** Those in urban areas were slightly older, more likely to be a minority, more likely to have seen a cardiologist, and had a higher CHA_2DS_2-VASc score compared to those in rural areas.

We present several figures to graphically present anticoagulation initiation over the study period for the entire cohort. Overall temporal trends of the anticoagulants by year and quarter are depicted in **Figure 14.4** (Supplemental Figure 1). In 2011, only 34% of beneficiaries used anticoagulants and by 2016, the percentage was 53%. The proportion of warfarin users decreased every year whereas the uptake of DOACs increased every year. By 2016, apixaban was the most commonly used anticoagulant. **Figure 14.2** shows the temporal trends of anticoagulation initiation by urban / rural category. Total anticoagulation increased in a similar manner each of the 4 rural/urban categories over time. Warfarin use is depicted with a clear gradient across rurality, and for every year, the highest percentage of AF patients prescribed warfarin are in the rural areas with those in isolated areas appearing to be the mostly likely to receive warfarin and the least likely to receive a DOAC. This pattern persisted when we stratified anticoagulation initiation by CHA_2DS_2-VASc score, presented in **Figure 14.3.** Overall, anticoagulation frequency was higher in those in rural areas, but those in rural areas were more likely to be receiving warfarin compared to those in urban areas. In **Figure 14.5** (Supplemental Figure 2), we present patterns of anticoagulant initiation by rural / urban category in 4 areas of the US. Total anticoagulation was highest in the Northeast region, and lowest in the Southern region. DOAC use was highest in the Southern region.

To formally test if anticoagulation and DOAC prescription patterns differed by rural / urban areas we matched 1 beneficiary from each of the 3 rural areas with up to 2 urban beneficiaries. The characteristics of patients after the matching are listed in **Table 14.4** (Supplemental Table 2). Statistical comparisons between rural / urban areas using the matched sample are listed in **Table 14.2.** Compared to urban areas, those in isolated areas were 7% more likely to use an anticoagulant, RR (95% CI) = 1.07 (1.04-1.10). However, they were 18% less likely to use a DOAC than those in urban areas, RR (95% CI) = 0.82 (0.78-

0.85). A similar pattern was seen, although to a slightly lesser extent, in the two other rural categories: those in the small rural areas were 5% more likely (95% CI = 2-7%) to be on anticoagulants, but 14% less likely (95% CI = 11-17%) to use a DOAC. Those in large rural areas were 2% more likely (95% CI = 0-4%) to be on anticoagulants, but 10% less likely (95% CI = 8-12%) to use a DOAC. Results were nearly identical when limited to those with a CHA2DS2-VASc score ≥ 2, and are listed in **Table 14.2**. Results were nearly identical when we required a 180 day run-in period instead of the 90 day run-in period (results not shown). We observed lower overall anticoagulation rates in women (8% lower) than in men but these rates did not differ by rural/urban status. We observed lower anticoagulation rates in blacks (16% lower) and other race (13% lower) compared to whites, but these rates did not differ by rural/urban status. We also did not observe a significant interaction for age by rural / urban status.

14.5. DISCUSSION

In this retrospective administrative claims analysis of AF Medicare patients, we found that anticoagulant prescription use in AF remains low, but has increased over time to 53% in 2016, mainly due to the introduction of the DOACs. Overall, total anticoagulation use was higher in those in rural areas compared to urban areas; However, DOAC initiation was lower in rural areas compared to those in urban areas with those in isolated areas least likely to be using a DOAC. This pattern persisted across all CHA_2DS_2-VASC scores. There were modest regional variations in the proportion of beneficiaries using anticoagulants and differences in proportions initiating DOACs.

Past studies from the early 2000's have shown that warfarin is underused among Medicare beneficiaries.[175] Our study using updated Medicare data indicated that anticoagulation is still underutilized in the Medicare population, but the introduction of DOACs to the market has increased the percentage of those on anticoagulants. A study from Hernandez et. al [178] looked at the regional variation in anticoagulant use in Medicare patients in 2013-2014 and found the adjusted probability of receiving any anticoagulant use was lowest in the south, and DOAC use was lowest in the northern US. Our study indicated that a similar regional treatment effect held true through 2016. Similar to previous studies, we found that overall oral anticoagulation, including DOAC initiation, was lower in blacks and other races compared to whites, and was also lower in females.[271, 272] In our study, these race and sex patterns held true across all rural / urban categories.

Our study adds to the literature by showing that DOAC initiation in Medicare patients remains lower in isolated and rural areas compared to those in urban areas. The most recent

European and North American guidelines for the management of AF incorporate recommendations on using DOACs as an alternative to warfarin.[35, 140] Currently, recommendations specifically for rural patients are not mentioned in the guidelines. However, due to the individualized approaches to INR monitoring needed for warfarin patients, along with numerous limitations including distance to coagulation clinics, it has been suggested that rural patients should be considered for DOACs instead of warfarin.[177] One barrier to DOAC initiation might be the higher cost of DOACs vs. warfarin use. However, the evidence suggests that long-term therapy with DOACs may be more cost-effective than warfarin treatment,[273, 274] primarily due to lower monitoring costs and reduced numbers of patients with strokes and systemic embolism.

Reports suggest cardiology providers are more likely to prescribe oral anticoagulants compared with primary care providers,[162-165] and this possibly results in a lower risk of stroke among patients who are managed by cardiology specialists.[163] Our study took into account whether patients had seen a cardiology provider in the time period around AF diagnosis, and those in the most isolated areas were less likely to have seen a cardiologist. We observed that cardiology providers did prescribe DOACs at a higher rate compared to primary care providers. However, most patients in our Medicare cohort had seen a cardiology provider around that time of AF (80%) and thus adjusting for provider specialty did not influence our estimates. Still, due to differences in the initiation of DOACs and the fact that those in isolated areas were less likely to see a cardiology provider, educating providers in rural areas to prescribe DOACs over warfarin may reduce the burdens of stroke and healthcare utilization of older adults in rural areas.

This study has several limitations which should be considered. First, this analysis is limited to fee-for-service Medicare beneficiaries with a stand-alone Part D plan, and this is a subset of all Medicare beneficiaries that is known to have a lower SES and more comorbidities than those with supplemental coverage. Therefore, our results may not be generalizable to the entire Medicare (65+) population. Second, ICD-9 and ICD-10 codes were used to identify AF cases and comorbid conditions and misclassification is possible. Third, unmeasured confounding is a known limitation in observational studies using administrative claims data. Although we attempted to account for many measured patient characteristics in our multivariable model account for differences in rural / urban patients, unmeasured factors (eg, socioeconomic status, distance from a clinic) possibly influenced our findings. Lastly, we only have information on prescriptions filled by the patients, not on the medication prescribed by the provider or compliance with therapy. Despite these limitations, our study has numerous key strengths, including a large sample size of Medicare beneficiaries that allowed us to

detect differences between groups. Medicare data contains individual zip code, which allowed us to compare rural status on a patient level which has not been done in other claims-based datasets. Using this large sample of Medicare data allowed us to identify important differences between rural and urban populations.

In conclusion, in this Medicare population with AF, anticoagulation use remains low but has increased over time due to the introduction of DOACs. However, those in rural areas were less likely to receive a DOAC compared to those in urban areas, with the lowest DOAC use occurring in the most isolated areas. Increasing the use of anticoagulants, in general, and of DOACs in particular, in AF patients living in rural areas may reduce the burdens of stroke and healthcare utilization of older adults in rural areas.

Table 14.1. Characteristics at the time of atrial fibrillation diagnosis by urban / rural classification for the entire cohort, Medicare, 2011-2016

	Urban (n=369,357)	Large Rural (n=38,167)	Small Rural (n=21,934)	Isolated Rural (n=17,794)
Age, years	79.0 ± 8.4	78.6 ± 8.2	78.7 ± 8.1	78.7 ± 8.2
Female, %	55	56	56	54
White race, %	84	92	92	94
Black race, %	7	4	4	2
Other race, %	9	4	3	3
CHA_2DS_2-VASC score	5.0 ± 1.9	4.8 ± 1.8	4.9 ± 1.8	4.8 ± 1.8
HAS-BLED score	3.1 ± 1.2	3.0 ± 1.1	3.0 ± 1.1	2.9 ± 1.1
Cardiology involvement	84	77	74	74
Comorbidities, %				
Hypertension	88	88	88	87
Diabetes	42	39	40	39
Myocardial infarction	12	14	15	14
Heart failure	36	36	38	37
Ischemic stroke/TIA	35	31	31	30
Peripheral artery disease	36	31	32	30
Hemorrhagic stroke	2	2	2	2
Dementia	9	7	7	6
Renal Disease	26	24	24	24
Chronic pulmonary disease	31	35	36	35
Liver disease	10	8	7	7
Hematological disorders	26	23	22	20
Gastrointestinal bleed	35	34	33	32
Other bleed	48	42	43	41
Malignancy	21	19	18	19
Metastatic cancer	4	4	3	4
Depression	21	21	21	20
Alcohol abuse	1	1	1	1
Medications, %				
Digoxin	0.4	0.4	0.4	0.4
Clopidogrel	14	15	15	14
Antiplatelet agents	0.3	0.4	0.3	0.5

Angiotensin-converting enzyme inhibitors	26	29	29	29
Angiotensin receptor blockers	15	14	13	12
Beta-blockers	34	37	38	38
Calcium channel blockers	28	28	28	26
Anti-arrhythmic agents	2	2	2	3
Statins	41	41	41	41
Diabetes medications	6	7	7	7
Oral Anticoagulants, %				
Warfarin	23	26	28	29
Dabigatran	4	4	4	4
Rivaroxaban	10	9	9	8
Apixaban	10	9	9	8

*Values correspond to mean ± standard deviation or percentage.

Table 14.2. Anticoagulation Fill Patterns of Matched Atrial Fibrillation Patients by Rural / Urban Classification, Medicare, 2011-2016

	Total n (%)	Isolated Rural	Matched Urban	RR[a] (95%CI)	RR[b] (95%CI)
Isolated vs. Urban					
All	52,335	17,782	35,553		
Any Anticoagulant	25,620 (48%)	8,907 (50%)	16,713 (47%)	1.07 (1.04-1.10)	1.07 (1.04-1.10)
DOAC	12,256 (23%)	3,687 (21%)	8,569 (24%)	0.82 (0.78-0.85)	0.82 (0.78-0.85)
Small Rural vs. Urban	Total n (%)	Small Rural	Matched Urban	RR[a] (95%CI)	RR[b] (95%CI)
All	65,748	21,920	43,828		
Any Anticoagulant	31,634 (48%)	10,850 (50%)	20,784 (47%)	1.05 (1.02-1.07)	1.05 (1.02-1.07)
DOAC	15,483 (24%)	4,768 (22%)	10,715 (24%)	0.86 (0.83-0.89)	0.87 (0.84-0.90)
Large Rural vs. Urban	Total n (%)	Large Rural	Matched Urban	RR[a] (95%CI)	RR[b] (95%CI)
All	114,417	38,149	76,269		
Any Anticoagulant	54,493 (48%)	18,398 (48%)	36,095 (47%)	1.02 (1.00-1.04)	1.02 (1.00-1.04)
DOAC	26,996 (24%)	8,411 (22%)	18,585 (24%)	0.90 (0.87-0.92)	0.89 (0.87-0.92)

Relative risk of prescription fills for rural status vs. urban (reference).

[a]Adjusted for age, race, sex, CHA2DS2-VASc score, HAS-BLED score, cardiology involvement, hypertension, diabetes, myocardial infarction, heart failure, ischemic stroke/transient ischemic attack, hemorrhagic stroke, peripheral artery disease, dementia, renal disease, chronic pulmonary disease, liver disease, hematological disorders, gastrointestinal bleeding, other bleeding, malignancy, metastatic cancer, depression, alcohol abuse, and use of digoxin, clopidogrel, antiplatelets, angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, beta-blockers, calcium channel blockers, statins and diabetes medications

[b]Limited to those with CHA2DS2-VASc score > 2

CI=confidence interval; DOAC=direct oral anticoagulant; HAS-BLED=hypertension, abnormal renal/liver function, stroke, bleeding history or predisposition, elderly (age >65 years), drugs/alcohol concomitantly; RR=relative risk.

Figure 14.1. Analysis flowchart of the 20% sample of traditional fee-for-service Medicare Beneficiaries, 2011-2016.

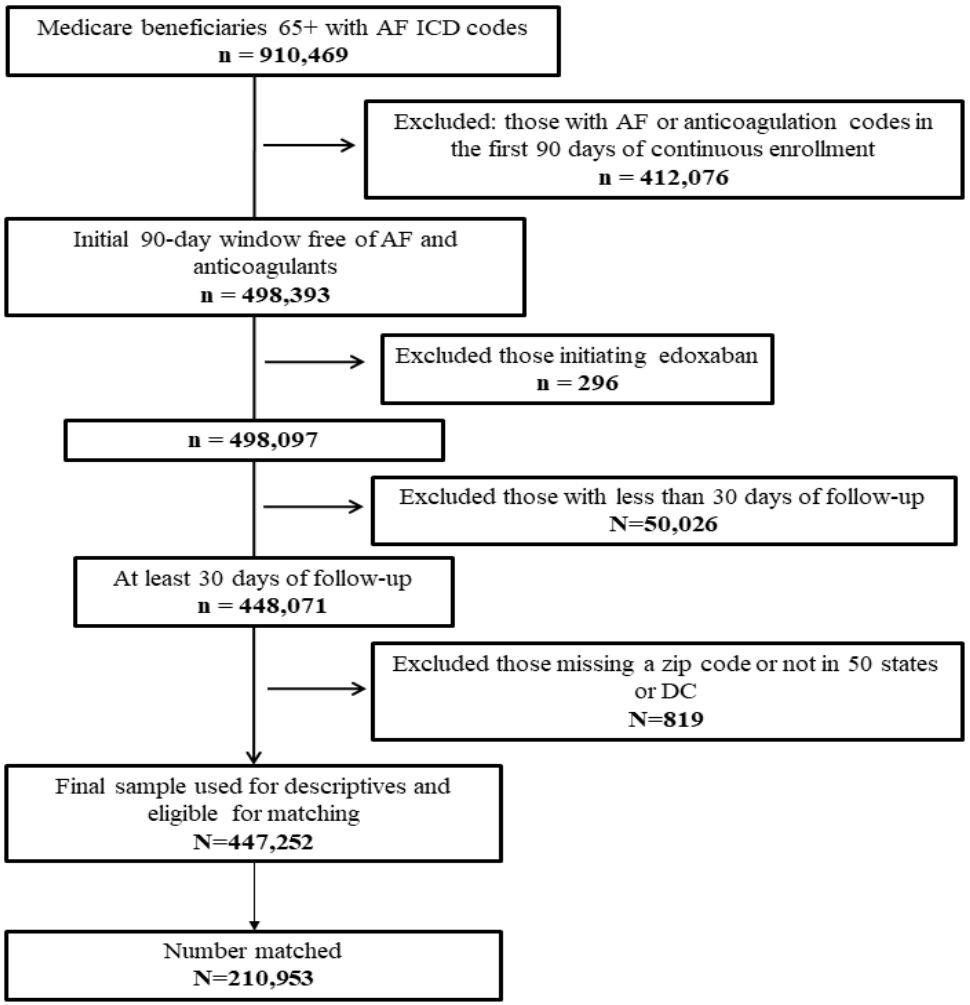


Figure 14.2. Temporal trends of anticoagulant initiation for the treatment of atrial fibrillation in Medicare beneficiaries by Urban / Rural category, 2011-2016

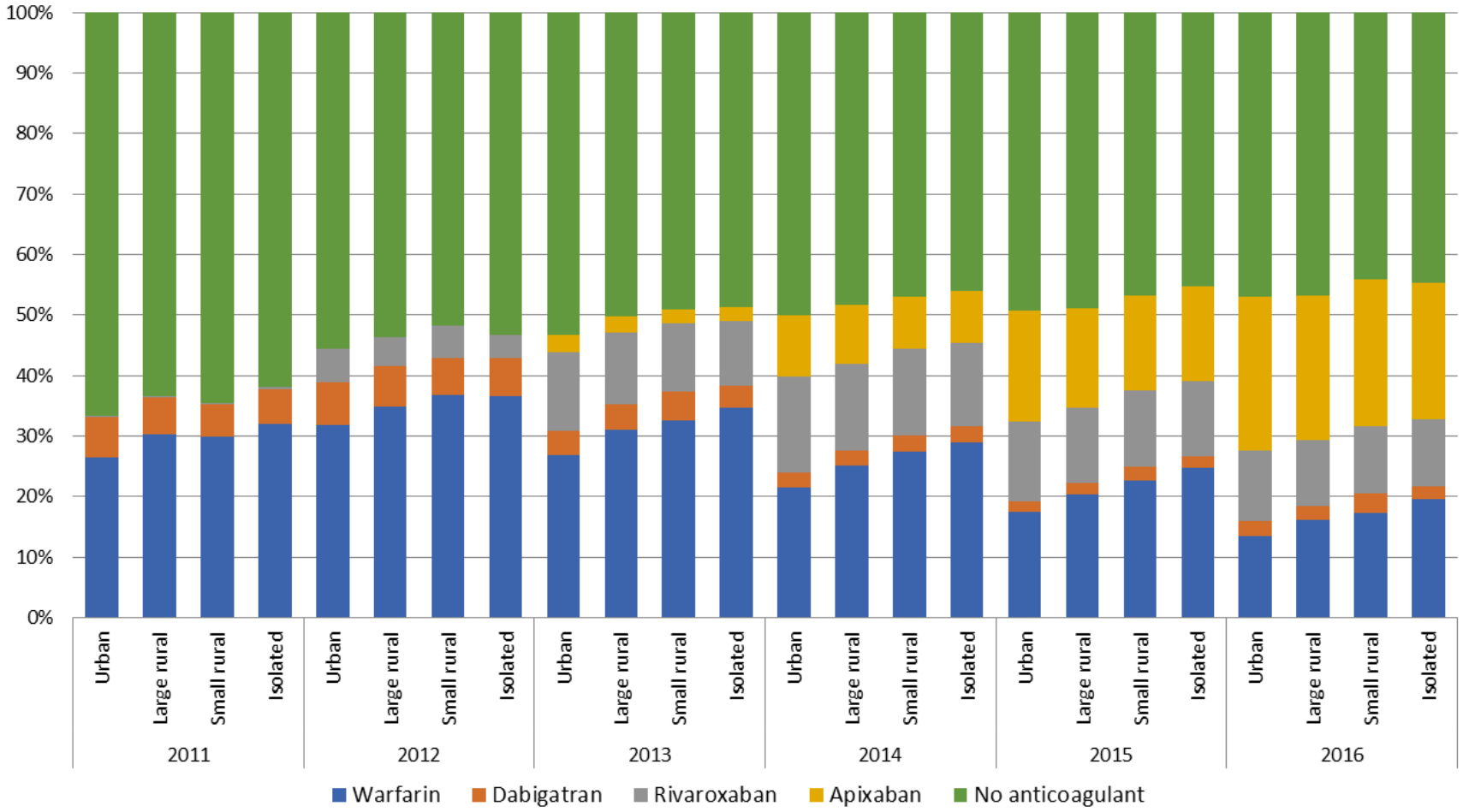


Figure 14.3. Overall temporal trends of oral anticoagulants initiation for the treatment of atrial fibrillation, by CHA_2DS_2-VASc Score, Medicare beneficiaries, pooled 2011-2016. The Y-axis is depicted to 60% and all percentages above that are beneficiaries on no anticoagulants. U=urban, L=large rural, S=small rural, I=isolated rural.

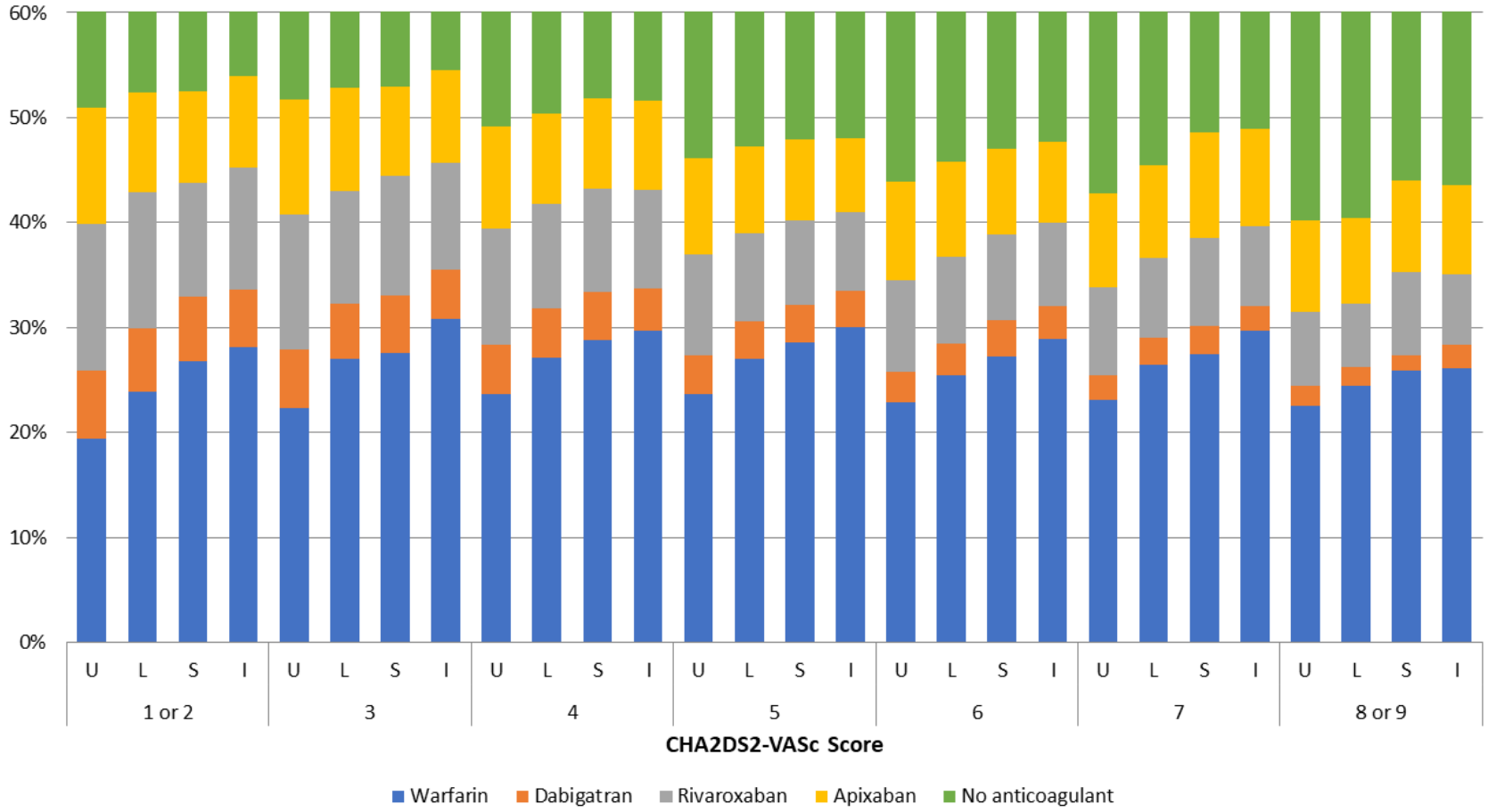


Table 14.3. Supplementary Table 1. ICD codes used to define pre-defined comorbidities

Condition	**ICD-9-CM codes**	**ICD-10-CM codes**
Alcoholism	265.2, 291.1, 291.2, 291.3, 291.5, 291.6, 291.7, 291.8, 291.9, 303.0, 303.9, 305.0, 357.5, 425.5, 535.3, 571.0, 571.1, 571.2, 571.3, 980, V11.3	F10, E52, G62.1, I42.6, K29.2, K70.0, K70.3, K70.9, T51.x, Z50.2, Z71.4, Z72.1
Chronic pulmonary disease	490-492, 494, 496	J40-J44, J47
Dementia	290, 294.1, 331.2	F00.x-F03.x, F05.1, G30, G31.1
Depression	296.2, 296.3, 296.5, 300.4, 309, 311	F20.4, F31.3, F31.5, F32.x, F33.x, F34.1, F41.2, F43.2
Diabetes	250	E10.0-E10.9, E11.0-E11.9, E12.0-E12.9, E13.0-E13.9, E14.0-E14.9
Gastrointestinal bleeding	455.2, 455.5, 455.8, 456.0, 456.20, 530.7, 530.82, 531.0, 531.2, 531.4, 531.6, 532.0, 532.2, 532.4, 532.6, 533.0, 533.2, 533.4, 533.6, 534.0, 534.2, 534.4, 534.6, 535.01, 535.11, 535.21, 535.31, 535.41, 535.51, 535.61, 537.83, 562.02, 562.03, 562.12, 562.13, 568.81, 569.3, 569.85, 578.0, 578.1, 578.9	K25.0,K25.2, K25.4, K25.6, K26.0, K26.2, K26.4, K26.6, K27.0, K27.2, K27.4, K27.6, K28.0, K28.2, K28.4, K28.6, K22.6, K22.8, K92.0, K64.8, K64.4, K64.8, K66.1, K62.5, K92.1, K92.2, K29.01, K29.41, K29.51, K29.61, K29.21, K29.71, K29.91, K29.81, K31.811, I85.01, I85.11, K57.11, K57.13, K57.31, K57.33, K55.21
Heart failure	398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 425.4, 425.9, 428	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5-I42.9, I43.x, I50.x, P29.0
Hematological disorders (Coagulopathy, anemia)	280, 281, 286, 287.1, 287.3, 287.4, 287.5	D65-D68, D69.1, D69.3-D69.6
Hemorrhagic stroke	430, 431, 432	I60-I62
Hypertension	401, 402, 403, 404, 405	I10.x, I11.x-I13.x, I15.x
Ischemic stroke / TIA	362.34, 433-438	G45-G46, I63-I69, H34.0
Kidney disease	403.01, 403.11, 403.91, 404.02, 404.03, 404.12, 404.13, 404.92, 404.93, 582, 583.0, 583.1, 583.2, 583.3, 583.4, 583.5, 583.6, 583.7, 585, 586, 588.0, V42.0, V45.1, V56	I12.0, I13.1, N03.2-N03.7, N05.2-N05.7, N18.x, N19.x, N25.0, Z49.0-Z49.2, Z94.0, Z99.2
Liver disease	070.22, 070.23, 070.32, 070.33, 070.44, 070.54, 070.6, 070.9, 456.0, 456.1, 456.2, 570, 571, 572.2, 572.3,	B18.x, K70.0-K70.3, K70.9, K71.3-K71.5, K71.7, K73.x, K74.x, K76.0, K76.2-K76.4,

	572.4, 572.5, 572.6, 572.7, 572.8, 573.3, 573.4, 573.8, 573.9, V42.7	K76.8, K76.9, Z94.4, I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5, K76.6, K76.7
Malignancy	140-172, 174-195, 200-208, 238.6	C00-C26, C30-C34, C37-C41, C45-C58, C60-C76, C81-C85, C90-C97, C43-C88
Metastatic cancer	196-199	C77-C80
Myocardial infarction	410, 412	I21, I22, I25.2
Other bleeding	423.0, 459.0, 568.81, 593.81, 599.7, 623.8, 626.6, 719.1, 784.7, 784.8, 786.3	N92.0, N92.1, I62.1, I62.0, I62.9, I31.2, K66.1, M25.0, R04.0, R04.1, R04.2, D50.0, D64.9, R79.1, R31, R58, D62
Peripheral artery disease	093.0, 437.3, 440, 441, 443.x, 471, 557.1, 557.9, V434	I70, I71, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9

*ICD-9-CM comorbidity codes were translated to ICD-10-CM codes using cross-walks, with review of face-validity

Table 14.4. Supplemental Table 2. Characteristics at the Time of Atrial Fibrillation Diagnosis by Matched Urban / Rural Classification, Medicare, 2011-2016

	Urban (n=133,102)	Large Rural (n=38,149)	Small Rural (n=21,920)	Isolated Rural (n=17,782)
Age, years	78.7 ± 8.2	78.6 ± 8.2	78.7 ± 8.1	78.7 ± 8.2
Female, %	55	56	56	54
White race, %	84	92	92	94
Black race, %	7	4	4	2
Other race, %	9	4	3	3
CHA_2DS_2-VASC score	4.9 ± 1.8	4.8 ± 1.8	4.9 ± 1.8	4.8 ± 1.8
HAS-BLED score	3.0 ± 1.1	3.0 ± 1.1	3.0 ± 1.1	2.9 ± 1.1
Cardiology involvement	83	77	74	74
Comorbidities, %				
Hypertension	88	88	88	87
Diabetes	41	39	40	39
Myocardial infarction	11	14	15	14
Heart failure	34	36	38	37
Ischemic stroke/TIA	32	31	31	30
Peripheral artery disease	34	31	32	30
Hemorrhagic stroke	2	2	2	2
Dementia	9	7	7	6
Renal Disease	24	24	24	24
Chronic pulmonary disease	30	35	36	35
Liver disease	9	8	7	7
Hematological disorders	25	23	22	20
Gastrointestinal bleed	35	34	33	32
Other bleed	46	42	43	41
Malignancy	21	19	18	19
Metastatic cancer	4	4	3	4
Depression	20	21	21	20
Alcohol abuse	1	1	1	1
Medications, %				
Digoxin	0.4	0.4	0.4	0.4
Clopidogrel	13	15	15	14
Antiplatelets	0.3	0.4	0.3	0.5

Angiotensin-converting enzyme inhibitors	26	29	29	29
Angiotensin receptor blockers	14	14	13	12
Beta-blockers	34	37	38	38
Calcium channel blockers	28	28	28	26
Anti-arrhythmias	2	2	2	3
Statins	41	41	41	41
Diabetes medications	6	7	7	7
Oral Anticoagulants, %				
Warfarin	23	26	28	29
Dabigatran	4	4	4	4
Rivaroxaban	10	9	9	8
Apixaban	10	9	9	8

*Values correspond to mean ± standard deviation or percentage.

Figure 14.4. Supplemental Figure 1. Overall temporal trends of oral anticoagulant initiation for the treatment of atrial fibrillation in Medicare beneficiaries, 2011-2016

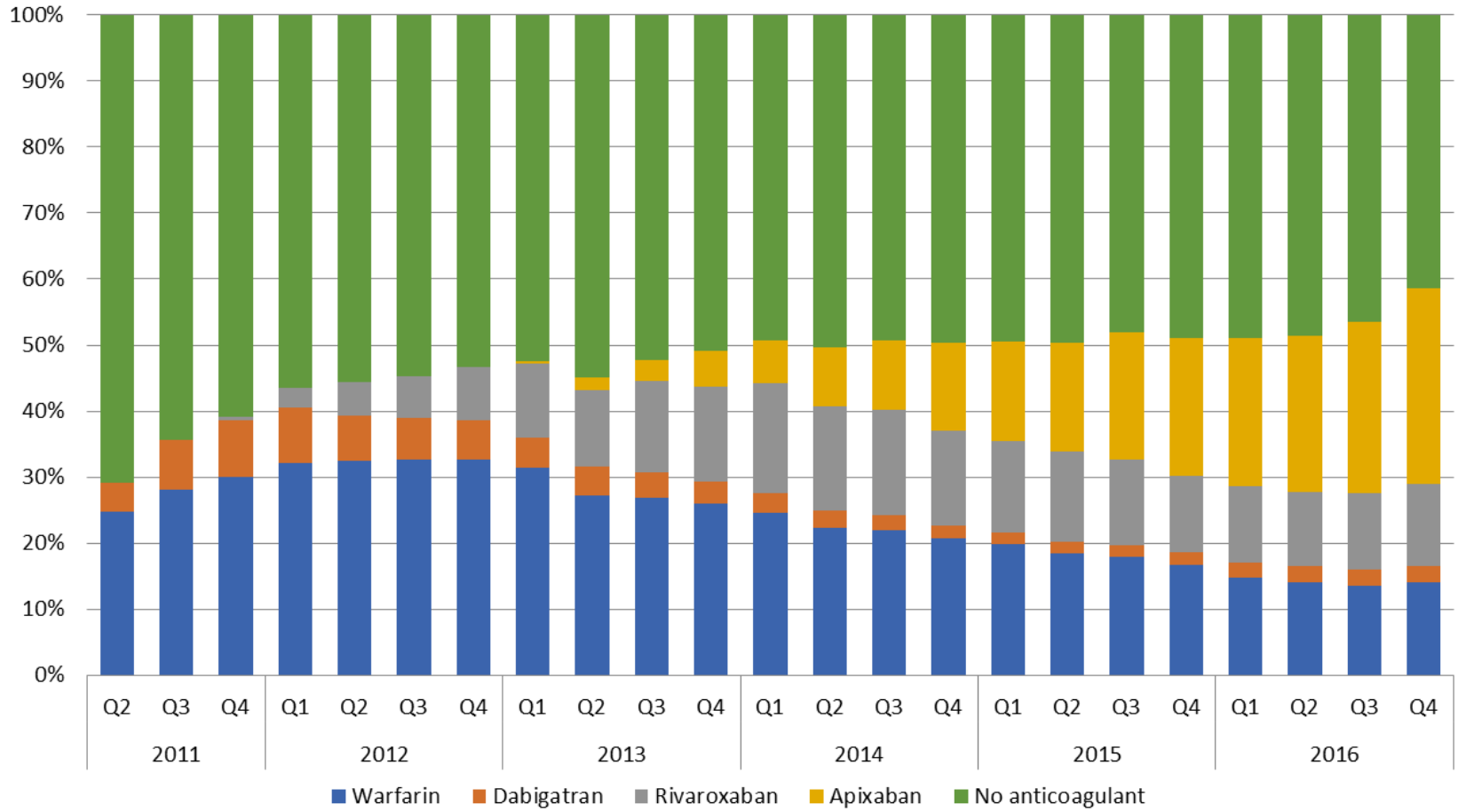
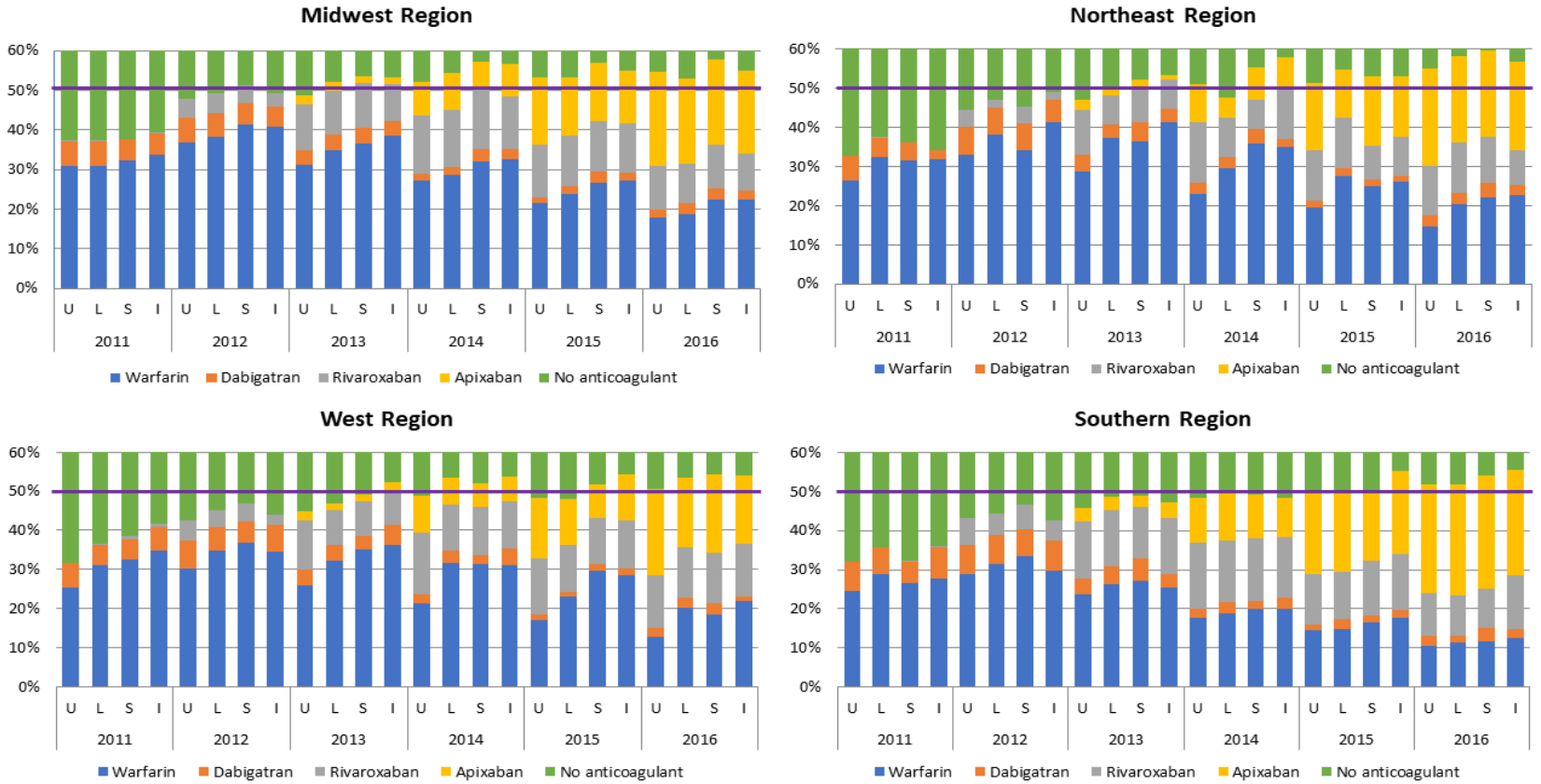


Figure 14.5. Supplemental Figure 2. Temporal trends of oral anticoagulants initiation for the treatment of atrial fibrillation in Medicare beneficiaries by Urban / Rural category in 4 regions of the US, 2011-2016.
The Y-axis is depicted to 60% and all percentages above that are beneficiaries on no anticoagulants. A line has been inserted at 50% for comparison. U=urban, L=large rural, S=small rural, I=isolated rural.



15. Manuscript 4 – Atrial fibrillation outcomes in rural vs. urban Medicare beneficiaries in the United States

15.1. OVERVIEW

Background – Rural areas in the US have higher rates of many cardiovascular outcomes and mortality compared with urban areas; however, it is unknown whether rural-urban disparities exist in atrial fibrillation (AF) outcomes. We used Medicare data to compare the risks of adverse events in newly-diagnosed AF patients in rural versus urban areas.

Methods –We identified 443,703 incident AF events in a 20% sample of fee-for-service Medicare beneficiaries from 2011-2016, and collected beneficiary residential zip code and covariates at the time of AF diagnosis. We categorized beneficiaries into 4 rural/urban areas by linking zip code to rural-urban commuting area codes, resulting in 82% categorized as urban, 9% large rural, 5% small rural and 4% isolated. We matched rural with urban beneficiaries based on characteristics at the time of AF diagnosis, and used Cox proportional hazards models to compare risk of mortality and incident hospitalized cardiovascular outcomes.

Results – Our study included 197,931 (mean age 79 ± 8 years) matched beneficiaries with AF. During a mean follow-up time of 2.1 ± 1.7 years, 2.1% of the cohort had an incident stroke, 4.3% had an incident myocardial infarction (MI), 3.5% had incident heart failure (HF), and 35% died. In multivariable adjusted analysis, those in rural areas had a higher risk of total mortality; the hazard ratio (95% CI) was 1.04 (95% CI=1.01-1.07) in isolated areas, 1.08 (1.04-1.10) in small rural areas, and 1.09 (1.07-1.11) in large rural areas compared to those in urban areas. Additionally, the risk of HF and MI in rural areas were 19% and 14% higher, respectively, compared to those in urban areas. The risk of mortality differed by rural/urban status with regards to sex, age, and race.

Conclusion – In this Medicare population with AF, those in rural areas had modestly higher risk of cardiovascular outcomes and death compared to those in urban areas. Further research is needed to identify ways to intervene to reduce adverse outcomes in AF patients in rural areas.

15.2. INTRODUCTION

The prevalence of atrial fibrillation (AF) increases with age, from 0.1% among people younger than 55 years, a doubling with each successive decade, and exceeding 20% by

age 80 years.[2, 3, 40] AF is associated with increased risk of subsequent major cardiovascular conditions including stroke,[4] myocardial infarction (MI)[5, 106] and heart failure (HF).[80, 107] Furthermore, AF independently increases the risk of mortality, with the relative risk being highest during the first year after AF manifests.[6, 107, 128, 129]

There are nearly 60 million people (19% of the population) living in rural areas according to the US Census Bureau. Those in rural areas have higher rates of adverse cardiovascular risk factors such as cigarette smoking, hypertension, diabetes, and obesity.[26-28] There are known rural vs. urban disparities in cardiovascular disease (CVD) in the US, such as a 40% higher heart disease prevalence in rural areas and higher risk of stroke.[29, 30] However, it is unknown if CVD disparities exist in AF patients living in rural versus urban areas of the US. Using a sample of Medicare beneficiaries, which included patient residential location, we compared the rates of adverse outcomes in AF patients living in rural vs. urban areas of the US from 2011-2016.

15.3. METHODS

Study population

Using a 20% sample of Medicare beneficiaries, we conducted a retrospective longitudinal cohort study using claims data from 2011-2016. We limited the cohort to beneficiaries receiving fee-for-service Medicare who were 65 years or older living in the US, and enrolled in a stand-alone Part D prescription drug plan. We included those continuously enrolled in traditional fee-for-service Medicare Parts A/B/D without supplemental coverage for at least 90 days during 2011-2016. We required at least the first 90 days of a beneficiaries' follow-up time to be free of AF diagnosis codes and anticoagulation codes. This 90 day run-in period was used to capture patient health information and comorbidities. This also allowed us to identify incident AF events and to capture the first anticoagulation prescription following an AF event. If a beneficiary enrolled in supplemental coverage we censored that individual at the time of supplemental enrollment. For this analysis, we required at least a 1-day follow-up period after AF diagnosis in order to assess outcomes that occurred beyond an inpatient hospitalization or outpatient diagnosis. Additionally, we excluded patients on oral anticoagulation prior to their AF event as that could be in indicator this was not an incident AF case, or that these agents were prescribed for other conditions (e.g., venous thromboembolism).

The initial sample included 910,649 AF patients aged 65 to 112 years. The exclusion flow chart is depicted in **Figure 15.1.** We excluded those with an AF diagnosis or prescription fill for an anticoagulant during the first 90 days of enrollment, and those on anticoagulants

prior to their first AF event (n=448,691), those who initiated edoxaban during the study period (due to small numbers; n=296), those with less than 1 day of follow-up (16,901), and those with a missing zip code or those with a zip code in a US territory (n=878). Our sample eligible for matching was 443,703 and for the comparative analyses, 197,931 of those were successfully matched.

Ascertainment of AF Patients

This analysis included patients age 65+ with at least one inpatient claim for AF or 2 outpatient claims for AF 7 to 365 days apart. AF claims were identified using International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis codes 427.3, 427.31, and 427.32, and ICD-10-CM codes I48.x in any position, which is a standard definition used in claims analysis.[45, 212] The validity of ICD-9-CM codes for the identification of AF has been well-established with a systematic review of studies showing a positive predicted value (PPV) of approximately 90% and a sensitivity of approximately 80%.[213] We defined the diagnosis date as the earlier of 1) the earliest discharge date for an inpatient claims, or 2) the earliest service date of the outpatient or physician claim.

Defining rural and urban beneficiaries

We identified beneficiary residential zip code at the time of AF diagnosis. We mapped zip codes to Rural-Urban Communing Area (RUCA) codes, which are approximation codes developed by the University of Washington Research[179] and commonly used to define rural and urban areas.[180] RUCA codes combine standard Census definitions with area commuting behaviors to capture functional and work relationships between regions. We used a 4- category classification to access the rurality of beneficiaries: urban (RUCA codes 1-3, 4.1, 5.1, 7.1, 8.1, 10.1), large rural (RUCA codes 4.0, 4.2, 5.0, 5.2, 6.0, 6.1), small rural (RUCA codes 7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2) and isolated (RUCA codes 10, 10.2, 10.3, 10.4, 10.5, 10.6).

Ascertainment of outcomes

We identified the following incident hospitalized outcomes in AF patients from inpatient claims using validated algorithms:[267, 275, 276] 1) ischemic stroke, 2) intracranial bleeding, 3) MI, 4) HF, and 5) gastrointestinal (GI) bleeding. Although oral anticoagulation in AF reduces the risk of ischemic stroke and systemic thromboembolism, this benefit is accompanied by an increased bleeding risk.[116-120] Therefore GI bleeding is an important and non-trivial outcome to consider. A list of ICD-9-CM and ICD-10-CM codes used to define outcomes is provided in **Table 15.3** (Supplementary Table 1), and consistent with validated algorithms, we identified incident outcomes using codes from the primary position, except for

MI, which could be listed in the primary or secondary position.[267, 275, 276] Additionally, we defined all-cause mortality as an outcome. Medicare links beneficiary records to the National Death Index, and we used the date of death obtained from the death certificate.

Covariates

We identified pre-determined covariates prevalent at the time of AF diagnosis using Medicare inpatient, outpatient, carrier, and pharmacy claims using validated published algorithms. [120, 267, 268] Race was self-reported at time of enrollment and we categorized it into race categories of white, black, and other/unknown (due to small numbers). We included the following comorbidities: prior ischemic stroke/transient ischemic attack(TIA), hemorrhagic stroke, heart failure, myocardial infarction, hypertension, diabetes, peripheral arterial disease, liver disease, kidney disease, chronic pulmonary disease, malignancies (except malignant skin neoplasm), metastatic cancer, history of bleeding, hematological disorders (anemia, coagulation defects), dementia, depression, and alcohol abuse.[267, 268] ICD codes for the comorbidity variables are listed in **Table 15.4** (Supplemental Table 2). We captured the presence of prescription fills for the following medications: clopidogrel, angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, beta-blockers, calcium channel blockers, anti-arrhythmic agents, and statins. We identified filled prescriptions initiated for oral anticoagulation and beneficiaries were assigned to the first anticoagulant filled, if any, following their first AF claim. We included prescriptions for warfarin and the direct oral anticoagulants (DOACs) consisting of dabigatran, rivaroxaban, and apixaban. We excluded initiators of the DOAC edoxaban due to small numbers. We calculated the CHA_2DS_2-VASc score[269] at the date of AF and it consisted of congestive heart failure, hypertension, age (1 point for age 65-74; 2 points for ≥ 75), diabetes, prior stroke or TIA (2 points), vascular disease, and female sex. We calculated the HAS-BLED score[122] using the variables of hypertension, abnormal renal/liver function, stroke, bleeding history or disposition, elderly (age>65) and drugs/alcohol concomitantly; the variable International Normalized Ratio (INR) normally included in the HAS-BLED score was not available for this cohort. We used Medicare carrier claims to identify provider specialty at outpatient visits. We classified beneficiaries who saw a cardiology provider within a predetermined period (30 days prior to or 90 days after AF diagnosis) as the cardiology group, while patients seen exclusively by internal medicine, family practice, medical doctor, or unspecified multispecialty group were classified as primary care. Patients seen by a cardiologist were included in the cardiology provider group, regardless of a primary care visit.

Statistical analysis

We used multivariable logistic regression to predict the probability of living in each rural area (vs the urban area) based on the covariates listed above and created a propensity score. To compare the rates of adverse outcomes, we matched beneficiaries based on AF date (± 30 days), age (± 1 years), sex, CHA_2DS_2-VASc score (± 0), and by propensity score (± 0.004). One beneficiary from each of the 3 rural categories was matched with up to 2 beneficiaries in the urban category using a greedy matching algorithm. One quarter of a standard deviation of the propensity score was used a caliper for matching. Baseline characteristics at the time of AF diagnosis were assessed in the 4 matched groups.

We used Cox proportional hazards models with time to event to assess the risk of each of the 6 outcomes in those in rural areas compared to those in urban areas. Time to event was calculated as the time in days from AF diagnosis date to each incident outcome, health plan disenrollment or enrollment in supplemental insurance, death, or the end of study follow-up (2016), whichever occurred first. Those with prevalent conditions for each outcome were not included in the analysis for the incident outcome. For example, a patient with an ischemic stroke prior to AF was not included in the analysis for incident ischemic stroke, but could be included in the analysis for incident MI, provided they did not have prevalent MI. Thus, outcomes are incident post-AF events. For each outcome, we ran 2 models. Model 1 adjusted for age (continuous), race (white, black, other), sex, CHA_2DS_2-VASc score, HAS-BLED score, specialist care (cardiology: yes/no), and the propensity score. In Model 2, we additionally adjusted for anticoagulant use (none, warfarin, DOAC) as a time-dependent variable.

We examined effect modification of the rural/urban association by sex, race, and age (<79, ≥79) by adding a multiplicative interaction term in the model. We additionally split follow-up time into early vs. late (<90 days, ≥90 days) to determine if the rates of outcomes occurring soon after AF differed by rurality.

We ran several sensitivity analyses. The first was limited to AF patients who qualify for oral anticoagulants according to the guidelines (CHA_2DS_2-VASc scores ≥ 2); due to the advanced age and poly-comorbidity of Medicare patients, we had to exclude a small percentage (<2% of beneficiaries) from this analysis. We conducted an additional analysis requiring a 180-day run-in time instead of 90 days.

15.4. RESULTS

After the initial exclusion criteria were applied, our study included 443,703 AF patients (mean age 79 ± 8 years), of which 82% lived in an urban area, 9% in a large rural area, 5% in a small rural area, and 4% in an isolated rural area. After matching, 197,931 AF

patients remained in our analyses. Characteristics of the matched cohort are listed in **Table 15.1.** Beneficiaries were well-matched with similar characteristics across categories, although those in isolated rural areas were more likely to be white race, less likely to have cardiology involvement, and more likely to be on warfarin compared to those in urban areas.

The number of adverse outcomes, the unadjusted incidence rate, and the associations between rural vs. urban beneficiaries are listed in **Table 15.2.** During a mean follow-up time of 2.1 ± 1.7 years after AF diagnosis, 2.1% of the matched cohort had an incident stroke, 0.41% had an incident intracranial bleed, 4.3% had an incident MI, 3.5% had incident HF, 3.5% had an incident GI bleed, and 35% died. Unadjusted incidence rates for mortality were higher in rural areas compared to urban; those in rural areas had 6 -16 more deaths per 1000 person-years compared to those in urban areas. The unadjusted incidence rates were mostly similar between rural and urban groups for the other outcomes. Compared to those in urban areas, those in isolated areas had increased relative risks of mortality (HR=1.04, 95% CI = 1.01-1.07), MI (HR=1.18, 95% CI = 1.09-1.27), and HF (HR=1.18, 95% CI = 1.11-1.25). Additional adjustment by time-varying anticoagulant use did not change the associations. Those in small rural and large rural areas also had an increased risk of mortality, MI, and HF compared to those in urban areas. The risk of mortality was highest in the large rural group (HR=1.09 (95% CI: 1.07-1.11) compared to the urban group. The risk of stroke was higher in those in small rural areas compared to those in urban areas (HR=1.10, 95% CI =1.02-1.19), and this association remained significant after adjustment for anticoagulant use. The risk of stroke in the other rural areas was not significantly higher than the urban area. We did not detect any rural-urban differences in the risk of intracranial bleeding or GI bleeding.

Results for each outcome were similar across all rural areas and therefore we combined those in the 3 rural areas into 1 group. We then assessed for rural vs. urban interactions for each outcome by sex, race, age, and early vs. late follow-up time. Of the outcomes, the only interactions that were significant were for mortality, and these are depicted in **Figure 15.2**. The hazard ratio of mortality for rural vs. urban areas was higher in men than in women (HR of 1.12 vs. 1.05; p for interaction = 0.0001), higher in those age < 79 compared to those ≥ 79 (HR of 1.12 vs. 1.05; p for interaction = 0.002), and in those in the other race category than in blacks or whites (HR of 1.24 for other, 1.05 for blacks and 1.07 for whites; p for interaction = 0.003).

In sensitivity analyses, associations remained nearly identical when we limited the analysis to those with a CHA_2DS_2-VASc scores ≥ 2, and also when we required a 180 day run-in period instead of 90 days (results are not shown).

15.5. DISCUSSION

In this retrospective longitudinal cohort of AF Medicare patients, we found that disparities in mortality and cardiovascular outcomes exist in rural areas versus urban areas. Specifically, the risks of mortality, MI, and HF were higher in rural areas, and the risk of stroke was higher for those in in small rural areas compared to urban areas. These associations were modest, although across a large population these disparities amounts to a large number of events. There were urban-rural differences in mortality where rural men had a higher risk than urban men, rural individuals aged 65-79 had a higher risk than urban individuals in the same age range, and those with a race other than white or black and living in a rural area had a higher risk than those of other race living in urban areas.

Few research studies have been conducted regarding rural vs. urban disparities in outcomes in AF patients in the US. A recent paper found in-hospital mortality of AF patients is higher in rural hospitals than in urban hospitals, and these results persisted across sex, race, and region.[172] However, this study looked at the hospital location rather than the patient location, and only focused on in-hospital mortality. Our study did not measure in-hospital mortality, but rather we required at least 1 day of follow-up in order to assess outcomes post-AF diagnosis. Nevertheless, our findings were consistent in that there was a higher risk of mortality of AF patients in rural areas compared to urban areas. Our results for the outcome of stroke were similar to a Canadian study that reported patients in rural areas were slightly more likely to fill a prescription for warfarin, but they experienced similar stroke and major bleeding rates to their urban counterparts.[277]

Our study adds to the literature by showing that despite adjustment for anticoagulation treatment, cardiology involvement (vs. primary care), and a plethora of co-morbidities, disparities in outcomes exist in rural compared with urban AF patients. Although disparities of CVD exist in the general US population, those disparities are driven mainly by race and socioeconomic status.[166] We observed those with a race other than white or black and living in a rural area had a higher risk than those of other race living in urban areas, and efforts to reduce the risk in this population should be a priority.

This study has several limitations which should be considered. First, this analysis is limited to fee-for-service Medicare beneficiaries with a stand-alone Part D plan, and this is a subset of all Medicare beneficiaries that is known to have a lower SES and more comorbidities than those with supplemental Part D coverage. Therefore, our results may not

be generalizable to the entire Medicare (65+) population. Second, unmeasured confounding is a known limitation in observational studies using administrative claims data, although we attempted to account for many measured patient characteristics in our multivariable model account for differences in rural / urban patients, unmeasured factors (eg, socioeconomic status, distance from a clinic) possibly influenced our findings. To further account for confounding, we adjusted for many pre-defined variables and created a propensity score to match groups and make them comparable. Therefore, our results only apply to the matched population, which may be different from the entire treated population. Third, misclassification is possible when using ICD codes. Fourth, we did not have cause of death data and therefore deaths specifically from CVD events such as stroke and MI were not counted as CVD events. Lastly, we only have information on prescriptions filled by the patients, not on the medication prescribed by the provider or compliance with therapy, and we did not report medication adherence. Despite these limitations, our study has numerous key strengths. The Medicare data contains individual ZIP code, which allows us to compare rural status on a patient level, which has not been done in other claims-based datasets. Using this large sample of Medicare data allowed us to identify important differences between rural and urban populations.

In conclusion, in this Medicare population with AF, those in rural areas had modestly higher risks of mortality and cardiovascular outcomes compared to those in urban areas. Further research is needed to identify ways to intervene to reduce adverse outcomes in AF patients in rural areas.

Table 15.1. Characteristics at the time of atrial fibrillation diagnosis by urban / rural classification for the entire cohort, Medicare, 2011-2016

	Urban (n=124,322)	Large Rural (n=35,983)	Small Rural (n=20,859)	Isolated Rural (n=16,767)
Age, years	78.8 ± 8.1	78.7 ± 8.2	78.9 ± 8.1	78.8 ± 8.2
Female, %	55	56	56	54
White race, %	84	91	92	94
Black race, %	7	4	4	2
Other race, %	9	4	4	4
CHA_2DS_2-VASC score	4.8 ± 1.8	4.9 ± 1.8	4.9 ± 1.8	4.8 ± 1.8
HAS-BLED score	2.8 ± 1.1	2.8 ± 1.1	2.8 ± 1.1	2.8 ± 1.1
Cardiology involvement	85	77	75	74
Comorbidities, %				
Hypertension	88	88	88	87
Diabetes	40	39	40	39
Myocardial infarction	12	13	14	13
Heart failure	36	36	38	37
Ischemic stroke/TIA	31	31	31	30
Peripheral artery disease	32	32	32	30
Hemorrhagic stroke	2	2	2	2
Dementia	7	8	7	6
Renal Disease	24	24	24	24
Chronic pulmonary disease	34	34	36	34
Liver disease	8	8	7	7
Hematological disorders	22	23	22	20
Gastrointestinal bleed	34	34	33	32
Other bleed	43	43	43	41
Malignancy	20	19	19	19
Metastatic cancer	4	4	4	4
Depression	20	21	21	20
Alcohol abuse	1	1	1	1
Medications, %				
Clopidogrel	14	14	14	14
Angiotensin-converting enzyme inhibitors	27	28	28	28
Angiotensin receptor blockers	14	14	13	12
Beta-blockers	35	35	36	36
Calcium channel blockers	28	28	28	27
Anti-arrhythmic agents	2	2	2	3
Statins	41	40	40	40

Oral Anticoagulants, %				
Warfarin	19	21	22	24
Dabigatran	4	4	4	4
Rivaroxaban	9	8	8	8
Apixaban	10	9	8	8

*Values correspond to mean ± standard deviation or percentage.

Table 15.2. Associations of outcomes comparing rural / urban classification of matched atrial fibrillation patients, Medicare, 2011-2016

Outcomes	Isolated (n=16,767)		Urban (n=32,909)		Hazard Ratio (95% CI)[a]	Hazard Ratio (95% CI)[b]
	# Events	IR (95% CI)	# Events	IR (95% CI)		
Mortality	5875 (35%)	164 (160-168)	11,018 (33%)	158 (155-161)	1.04 (1.01-1.07)	1.04 (1.01-1.07)
Ischemic stroke	352 (2.2%)	10 (9.1-11)	704 (2.2%)	10 (9.7-11)	1.02 (0.90-1.16)	1.01 (0.89-1.15)
Intracranial bleeding	70 (0.42%)	2.1 (1.6-2.5)	146 (0.44%)	2.1 (1.8-2.5)	0.98 (0.73-1.31)	0.93 (0.70-1.25)
Myocardial Infarction	735 (4.5%)	21 (19-22)	1347 (4.2%)	20 (19-21)	1.12 (1.02-1.23)	1.11 (1.01-1.22)
Heart Failure	607 (3.9%)	18 (17-20)	1005 (3.3%)	16 (14-17)	1.24 (1.12-1.37)	1.20 (1.08-1.33)
GI bleeding	394 (2.4%)	11 (10-13)	771 (2.4%)	11 (10-12)	1.03 (0.91-1.16)	0.99 (0.88-1.13)
	Small Rural (n=20,859)		Urban (n=40,806)		Hazard Ratio (95% CI)[a]	Hazard Ratio (95% CI)[b]
	# Events	IR (95% CI)	# Events	IR (95% CI)		
Mortality	7693 (37%)	176 (172-180)	13,873 (34%)	161 (159-164)	1.08 (1.05-1.11)	1.08 (1.05-1.11)
Ischemic stroke	460 (2.3%)	11 (10-12)	822 (2.1%)	10 (8.9-10)	1.13 (1.01-1.27)	1.13 (1.01-1.27)
Intracranial bleeding	64 (0.31%)	1.5 (1.1-1.9)	164 (0.40%)	1.9 (1.6-2.2)	0.80 (0.60-1.08)	0.80 (0.60-1.07)
Myocardial Infarction	901 (4.4%)	22 (20-23)	1714 (4.3%)	21 (20-22)	1.09 (1.01-1.19)	1.10 (1.01-1.19)
Heart Failure	780 (4.0%)	20 (18-21)	1323 (3.5%)	17 (16-18)	1.21 (1.11-1.33)	1.19 (1.09-1.30)
GI bleeding	488 (2.4%)	12 (11-13)	952 (2.4%)	11 (10-22)	1.05 (0.94-1.17)	1.03 (0.92-1.15)
	Large Rural (n=35,983)		Urban (n=69,633)		Hazard Ratio (95% CI)[a]	Hazard Ratio (95% CI)[b]
	# Events	IR (95% CI)	# Events	IR (95% CI)		
Mortality	13,192 (37%)	177 (174-180)	23,569 (34%)	161 (159-163)	1.09 (1.07-1.11)	1.09 (1.07-1.11)
Ischemic stroke	692 (2.0%)	9.6 (8.9-10)	1417 (2.1%)	9.7 (9.2-10)	0.97 (0.89-1.07)	0.97 (0.89-1.07)
Intracranial bleeding	155 (0.43%)	2.1 (1.8-2.4)	286 (0.41%)	2.0 (1.7-2.2)	1.09 (0.90-1.33)	1.07 (0.88-1.31)
Myocardial Infarction	1655 (4.7%)	23 (22-24)	2887 (4.2%)	21 (20-21)	1.17 (1.11-1.25)	1.18 (1.11-1.25)
Heart Failure	1253 (3.7%)	18 (17-19)	2172 (3.3%)	16 (15-17)	1.17 (1.09-1.26)	1.16 (1.08-1.24)
GI bleeding	840 (2.4%)	12 (11-12)	1666 (2.4%)	12 (11-12)	1.02 (0.94-1.11)	1.01 (0.93-1.09)

IR incidence rate, CI confidence interval

*Incidence rate is per 1000 person-years

[a]Adjusted for age, race, sex, CHA_2DS_2-VASc score, HAS-BLED score, cardiology provider, and propensity score

[b]Additionally adjusted for anticoagulation use as a time-dependent variable

Figure 15.1. Analysis flowchart of the 20% sample of traditional fee-for-service Medicare Beneficiaries, 2011-2016.

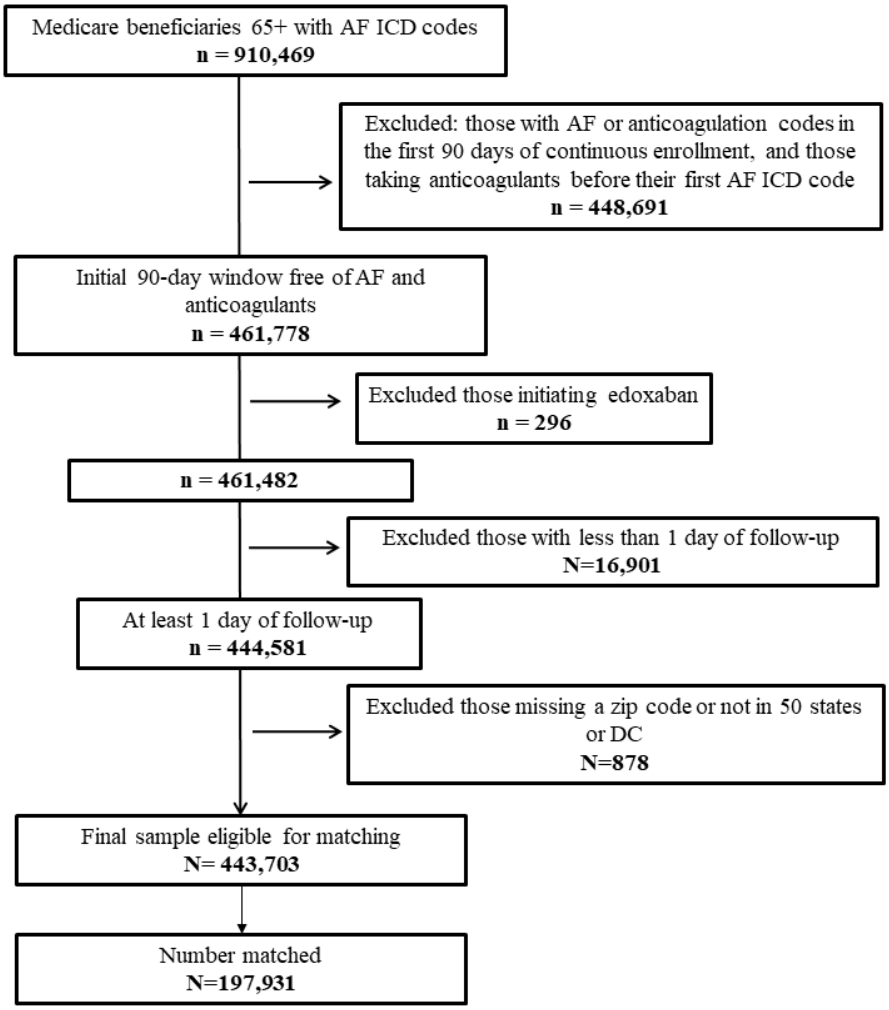


Figure 15.2. Hazard ratio of mortality in rural vs. urban areas, stratified by sex, age, and race, Medicare, 2011-2016

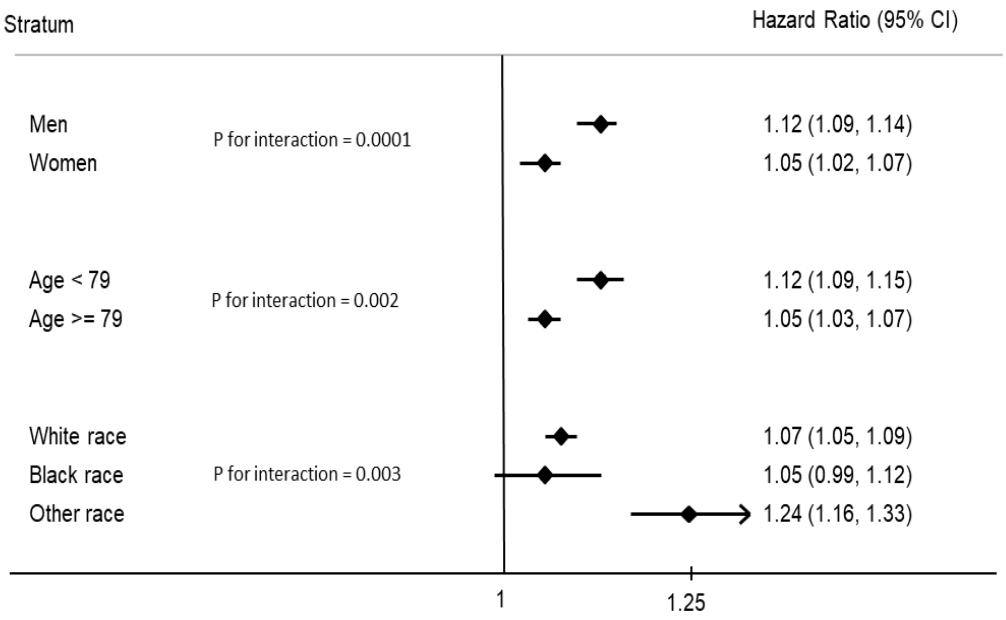


Table 15.3. Supplemental Table 1. ICD codes for outcomes

Condition	**ICD-9-CM codes**	**ICD-10-CM codes**
Ischemic stroke	434, 436	I63, I66, I67.89
Intracranial bleeding	430, 431	I60, I61
Myocardial infarction	410 (except 410.x2)	I21, I22 (except .A1)
Heart Failure	402.x1, 404.x1, 404.x3, 428	I11.0, I13.0, I13.2, I50
Gastrointestinal bleeding	455.2, 455.5, 455.8, 456.0, 456.20, 530.7, 530.82, 531.0, 531.2, 531.4, 531.6, 532.0, 532.2, 532.4, 532.6, 533.0, 533.2, 533.4, 533.6, 534.0, 534.2, 534.4, 534.6, 535.01, 535.11, 535.21, 535.31, 535.41, 535.51, 535.61, 537.83, 562.02, 562.03, 562.12, 562.13, 568.81, 569.3, 569.85, 578.0, 578.1, 578.9	K25.0,K25.2, K25.4, K25.6, K26.0, K26.2, K26.4, K26.6, K27.0, K27.2, K27.4, K27.6, K28.0, K28.2, K28.4, K28.6, K22.6, K22.8, K92.0, K64.8, K64.4, K64.8, K66.1, K62.5, K92.1, K92.2, K29.01, K29.41, K29.51, K29.61, K29.21, K29.71, K29.91, K29.81, K31.811, I85.01, I85.11, K57.11, K57.13, K57.31, K57.33, K55.21

Codes required to be in the 1st position, except for MI, where they were required in 1st or 2nd position. *ICD-9-CM comorbidity codes were translated to ICD-10-CM codes using cross-walks, with review of face-validity

Table 15.4. Supplementary Table 2. ICD codes used to define pre-defined comorbidities.

Condition	**ICD-9-CM codes**	**ICD-10-CM codes**
Alcoholism	265.2, 291.1, 291.2, 291.3, 291.5, 291.6, 291.7, 291.8, 291.9, 303.0, 303.9, 305.0, 357.5, 425.5, 535.3, 571.0, 571.1, 571.2, 571.3, 980, V11.3	F10, E52, G62.1, I42.6, K29.2, K70.0, K70.3, K70.9, T51.x, Z50.2, Z71.4, Z72.1
Chronic pulmonary disease	490-492, 494, 496	J40-J44, J47
Dementia	290, 294.1, 331.2	F00.x-F03.x, F05.1, G30, G31.1
Depression	296.2, 296.3, 296.5, 300.4, 309, 311	F20.4, F31.3, F31.5, F32.x, F33.x, F34.1, F41.2, F43.2
Diabetes	250	E10.0-E10.9, E11.0-E11.9, E12.0-E12.9, E13.0-E13.9, E14.0-E14.9
Gastrointestinal bleeding	455.2, 455.5, 455.8, 456.0, 456.20, 530.7, 530.82, 531.0, 531.2, 531.4, 531.6, 532.0, 532.2, 532.4, 532.6, 533.0, 533.2, 533.4, 533.6, 534.0, 534.2, 534.4, 534.6, 535.01, 535.11, 535.21, 535.31, 535.41, 535.51, 535.61, 537.83, 562.02, 562.03, 562.12, 562.13, 568.81, 569.3, 569.85, 578.0, 578.1, 578.9	K25.0,K25.2, K25.4, K25.6, K26.0, K26.2, K26.4, K26.6, K27.0, K27.2, K27.4, K27.6, K28.0, K28.2, K28.4, K28.6, K22.6, K22.8, K92.0, K64.8, K64.4, K64.8, K66.1, K62.5, K92.1, K92.2, K29.01, K29.41, K29.51, K29.61, K29.21, K29.71, K29.91, K29.81, K31.811, I85.01, I85.11, K57.11, K57.13, K57.31, K57.33, K55.21
Heart failure	398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 425.4, 425.9, 428	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5-I42.9, I43.x, I50.x, P29.0
Hematological disorders (Coagulopathy, anemia)	280, 281, 286, 287.1, 287.3, 287.4, 287.5	D65-D68, D69.1, D69.3-D69.6
Hemorrhagic stroke	430, 431, 432	I60-I62
Hypertension	401, 402, 403, 404, 405	I10.x, I11.x-I13.x, I15.x
Ischemic stroke / TIA	362.34, 433-438	G45-G46, I63-I69, H34.0
Kidney disease	403.01, 403.11, 403.91, 404.02, 404.03, 404.12, 404.13, 404.92, 404.93, 582, 583.0, 583.1, 583.2, 583.3, 583.4, 583.5, 583.6, 583.7, 585, 586, 588.0, V42.0, V45.1, V56	I12.0, I13.1, N03.2-N03.7, N05.2-N05.7, N18.x, N19.x, N25.0, Z49.0-Z49.2, Z94.0, Z99.2

Liver disease	070.22, 070.23, 070.32, 070.33, 070.44, 070.54, 070.6, 070.9, 456.0, 456.1, 456.2, 570, 571, 572.2, 572.3, 572.4, 572.5, 572.6, 572.7, 572.8, 573.3, 573.4, 573.8, 573.9, V42.7	B18.x, K70.0-K70.3, K70.9, K71.3-K71.5, K71.7, K73.x, K74.x, K76.0, K76.2-K76.4, K76.8, K76.9, Z94.4, I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5, K76.6, K76.7
Malignancy	140-172, 174-195, 200-208, 238.6	C00-C26, C30-C34, C37-C41, C45-C58, C60-C76, C81-C85, C90-C97, C43-C88
Metastatic cancer	196-199	C77-C80
Myocardial infarction	410, 412	I21, I22, I25.2
Other bleeding	423.0, 459.0, 568.81, 593.81, 599.7, 623.8, 626.6, 719.1, 784.7, 784.8, 786.3	N92.0, N92.1, I62.1, I62.0, I62.9, I31.2, K66.1, M25.0, R04.0, R04.1, R04.2, D50.0, D64.9, R79.1, R31, R58, D62
Peripheral artery disease	093.0, 437.3, 440, 441, 443.x, 471, 557.1, 557.9, V434	I70, I71, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9

*ICD-9-CM comorbidity codes were translated to ICD-10-CM codes using cross-walks, with review of face-validity

16. SUMMARY

The overarching aims of this dissertation were to assess the relation of proteomics and incident AF, to develop a series of risk prediction scores better calibrated to older adults, and to determine if disparities exist in AF treatment and outcomes in rural versus urban areas of the US.

Each of these manuscripts contributed to the AF literature by addressing an important knowledge gap of public health significance. In the first manuscript, our objective was to make use of the newly-measured proteomic markers in ARIC and assess their relationship to incident AF. In this population-based sample of older adults, we found NT-proBNP was the protein most strongly associated with incident AF risk. For every doubling of NT-proBNP the risk of incident AF was 1.82 times higher (95%CI: 1.68-1.98). In addition to NT-proBNP, after further adjustment for eGFR and medication use, we found 16 other proteins remained significantly associated with incident AF. Through pathway analysis, we explored mechanistic pathways of AF development. Our results offer new observations into the biological changes that may precede AF onset and provide insight into mechanistic pathways of AF development. If replicated further, these novel proteins might be worth evaluating for possible pharmacologic targets in AF.

In the second manuscript, our objectives were to improve the discrimination and calibration of AF risk prediction models, and in the process, consider novel markers such as proteomics for inclusion. Using a population-based sample of older adults, we developed a series of models from simple to involved that selected variables predicting incident AF within a 5 year period. Our final simple prediction model included 8 clinical variables and had moderate discrimination (c-statistic 0.697; 95% CI 0.671-0.723). The addition of blood biomarkers plus 16 proteins from proteomic analysis greatly improved the discrimination (c-statistic 0.795; 95% CI 0.773-0.816) while still showing excellent calibration ($\chi^2 = 7.6$; P = 0.58). Addition of abnormal P wave axis, left atrial diameter, and septal E/e prime moderately increased the c-statistic to 0.806 (95% CI: 0.785-0.827) in the full-developed model that contained 30 variables total. Our series of developed AF prediction models are better targeted and calibrated to older populations. The addition of biomarkers, including proteomics data, improved prediction, suggesting it may be worthwhile to explore developing cost-effective and time-efficient ways to quantify the predictive protein biomarkers.

Despite rural disparities of CVD, there is little known regarding the anticoagulation rates of AF patients in rural areas, and if there are differences in the rates of adverse outcomes of AF patients in rural areas versus urban areas. The objectives of the third and fourth

manuscripts were to use Medicare data to fill in these knowledge gaps. In the third manuscript, we examined the initiation of anticoagulation use and found the overall percentage of AF patients who initiated an anticoagulant rose from 34% in 2011 to 53% in 2016, driven by the uptake of DOACs. There were clear gradients of anticoagulant use by rurality. In a multivariable-adjusted analysis of beneficiaries matched by rural / urban category, those in rural areas were more likely to initiate an anticoagulant compared to those in urban areas. However, those in rural areas were less likely to receive a DOAC; those in isolated areas were 18% less likely (95% CI = 15 to 22%) to initiate a DOAC compared to those in urban areas. In the fourth manuscript, we compared the risks of adverse events in newly-diagnosed AF patients. We found that those in rural areas had modestly higher risk of MI, HF and mortality compared to those in urban areas. Collectively, increasing the percentages of DOAC use in AF patients living in rural areas may reduce the burdens of stroke and healthcare utilization of older adults in rural areas. Further research is needed to identify ways to intervene to reduce adverse outcomes in AF patients in rural areas.

17. REFERENCES

References

1. Mou L, Norby FL, Chen LY, O'Neal WT, Lewis TT, Loehr LR, Soliman EZ and Alonso A. Lifetime Risk of Atrial Fibrillation by Race and Socioeconomic Status: ARIC Study (Atherosclerosis Risk in Communities). *Circulation Arrhythmia and electrophysiology*. 2018;11:e006350.
2. Heeringa J, van der Kuip DA, Hofman A, Kors JA, van Herpen G, Stricker BH, Stijnen T, Lip GY and Witteman JC. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. *European heart journal*. 2006;27:949-53.
3. Lloyd-Jones DM, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, D'Agostino RB, Massaro JM, Beiser A, Wolf PA and Benjamin EJ. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. *Circulation*. 2004;110:1042-6.
4. Wolf PA, Abbott RD and Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke; a journal of cerebral circulation*. 1991;22:983-8.
5. Soliman EZ, Lopez F, O'Neal WT, Chen LY, Bengtson L, Zhang ZM, Loehr L, Cushman M and Alonso A. Atrial Fibrillation and Risk of ST-Segment-Elevation Versus Non-ST-Segment-Elevation Myocardial Infarction: The Atherosclerosis Risk in Communities (ARIC) Study. *Circulation*. 2015;131:1843-50.
6. Miyasaka Y, Barnes ME, Bailey KR, Cha SS, Gersh BJ, Seward JB and Tsang TS. Mortality trends in patients diagnosed with first atrial fibrillation: a 21-year community-based study. *Journal of the American College of Cardiology*. 2007;49:986-92.
7. Kim MH, Johnston SS, Chu BC, Dalal MR and Schulman KL. Estimation of total incremental health care costs in patients with atrial fibrillation in the United States. *Circulation Cardiovascular quality and outcomes*. 2011;4:313-20.
8. Katsnelson M, Koch S and Rundek T. Stroke Prevention in Atrial Fibrillation. *Journal of atrial fibrillation*. 2010;3:279.
9. Alonso A, Agarwal SK, Soliman EZ, Ambrose M, Chamberlain AM, Prineas RJ and Folsom AR. Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. *American heart journal*. 2009;158:111-7.
10. Alonso A, Krijthe BP, Aspelund T, Stepas KA, Pencina MJ, Moser CB, Sinner MF, Sotoodehnia N, Fontes JD, Janssens AC, Kronmal RA, Magnani JW, Witteman JC, Chamberlain AM, Lubitz SA, Schnabel RB, Agarwal SK, McManus DD, Ellinor PT, Larson MG, Burke GL, Launer LJ, Hofman A, Levy D, Gottdiener JS, Kaab S, Couper D, Harris TB, Soliman EZ, Stricker BH, Gudnason V, Heckbert SR and Benjamin EJ. Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *Journal of the American Heart Association*. 2013;2:e000102.
11. Rienstra M, Yin X, Larson MG, Fontes JD, Magnani JW, McManus DD, McCabe EL, Coglianese EE, Amponsah M, Ho JE, Januzzi JL, Jr., Wollert KC, Fradley MG, Vasan RS, Ellinor PT, Wang TJ and Benjamin EJ. Relation between soluble ST2, growth differentiation factor-15, and high-sensitivity troponin I and incident atrial fibrillation. *American heart journal*. 2014;167:109-115 e2.
12. Willeit K, Pechlaner R, Willeit P, Skroblin P, Paulweber B, Schernthaner C, Toell T, Egger G, Weger S, Oberhollenzer M, Kedenko L, Iglseder B, Bonora E, Schett G, Mayr M, Willeit J and Kiechl S. Association Between Vascular Cell Adhesion Molecule 1 and Atrial Fibrillation. *JAMA cardiology*. 2017;2:516-523.
13. Guo Y, Lip GY and Apostolakis S. Inflammation in atrial fibrillation. *Journal of the American College of Cardiology*. 2012;60:2263-70.

14. Schnabel RB, Larson MG, Yamamoto JF, Kathiresan S, Rong J, Levy D, Keaney JF, Jr., Wang TJ, Vasan RS and Benjamin EJ. Relation of multiple inflammatory biomarkers to incident atrial fibrillation. *The American journal of cardiology*. 2009;104:92-6.
15. Samman Tahhan A, Sandesara PB, Hayek SS, Alkhoder A, Chivukula K, Hammadah M, Mohamed-Kelli H, O'Neal WT, Topel M, Ghasemzadeh N, Ko YA, Aida H, Gafeer M, Sperling L, Vaccarino V, Liang Y, Jones DP and Quyyumi AA. Association between oxidative stress and atrial fibrillation. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2017;14:1849-1855.
16. Filion KB, Agarwal SK, Ballantyne CM, Eberg M, Hoogeveen RC, Huxley RR, Loehr LR, Nambi V, Soliman EZ and Alonso A. High-sensitivity cardiac troponin T and the risk of incident atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study. *American heart journal*. 2015;169:31-8 e3.
17. Hussein AA, Bartz TM, Gottdiener JS, Sotoodehnia N, Heckbert SR, Lloyd-Jones D, Kizer JR, Christenson R, Wazni O and deFilippi C. Serial measures of cardiac troponin T levels by a highly sensitive assay and incident atrial fibrillation in a prospective cohort of ambulatory older adults. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2015;12:879-85.
18. Sinner MF, Stepas KA, Moser CB, Krijthe BP, Aspelund T, Sotoodehnia N, Fontes JD, Janssens AC, Kronmal RA, Magnani JW, Witteman JC, Chamberlain AM, Lubitz SA, Schnabel RB, Vasan RS, Wang TJ, Agarwal SK, McManus DD, Franco OH, Yin X, Larson MG, Burke GL, Launer LJ, Hofman A, Levy D, Gottdiener JS, Kaab S, Couper D, Harris TB, Astor BC, Ballantyne CM, Hoogeveen RC, Arai AE, Soliman EZ, Ellinor PT, Stricker BH, Gudnason V, Heckbert SR, Pencina MJ, Benjamin EJ and Alonso A. B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*. 2014;16:1426-33.
19. Schnabel RB, Larson MG, Yamamoto JF, Sullivan LM, Pencina MJ, Meigs JB, Tofler GH, Selhub J, Jacques PF, Wolf PA, Magnani JW, Ellinor PT, Wang TJ, Levy D, Vasan RS and Benjamin EJ. Relations of biomarkers of distinct pathophysiological pathways and atrial fibrillation incidence in the community. *Circulation*. 2010;121:200-7.
20. Patton KK, Ellinor PT, Heckbert SR, Christenson RH, DeFilippi C, Gottdiener JS and Kronmal RA. N-terminal pro-B-type natriuretic peptide is a major predictor of the development of atrial fibrillation: the Cardiovascular Health Study. *Circulation*. 2009;120:1768-74.
21. Patton KK, Heckbert SR, Alonso A, Bahrami H, Lima JA, Burke G and Kronmal RA. N-terminal pro-B-type natriuretic peptide as a predictor of incident atrial fibrillation in the Multi-Ethnic Study of Atherosclerosis: the effects of age, sex and ethnicity. *Heart*. 2013;99:1832-6.
22. Kara K, Geisel MH, Mohlenkamp S, Lehmann N, Kalsch H, Bauer M, Neumann T, Dragano N, Moebus S, Jockel KH, Erbel R and Mahabadi AA. B-type natriuretic peptide for incident atrial fibrillation-The Heinz Nixdorf Recall Study. *Journal of cardiology*. 2015;65:453-8.
23. Lind L, Sundstrom J, Stenemo M, Hagstrom E and Arnlov J. Discovery of new biomarkers for atrial fibrillation using a custom-made proteomics chip. *Heart*. 2017;103:377-382.
24. Mathew JS, Sachs MC, Katz R, Patton KK, Heckbert SR, Hoofnagle AN, Alonso A, Chonchol M, Deo R, Ix JH, Siscovick DS, Kestenbaum B and de Boer IH. Fibroblast growth factor-23 and incident atrial fibrillation: the Multi-Ethnic Study of Atherosclerosis (MESA) and the Cardiovascular Health Study (CHS). *Circulation*. 2014;130:298-307.

25. Alonso A, Misialek JR, Eckfeldt JH, Selvin E, Coresh J, Chen LY, Soliman EZ, Agarwal SK and Lutsey PL. Circulating fibroblast growth factor-23 and the incidence of atrial fibrillation: the Atherosclerosis Risk in Communities study. *Journal of the American Heart Association*. 2014;3:e001082.
26. O'Connor A and Wellenius G. Rural-urban disparities in the prevalence of diabetes and coronary heart disease. *Public health*. 2012;126:813-20.
27. Moy E GM, Bastian B, et al. Leading Causes of Death in Nonmetropolitan and Metropolitan Areas — United States, 1999–2014. . *MMWR Surveill Summ*. 2017;66(No. SS-1):1–8. .
28. Mainous AG, 3rd, King DE, Garr DR and Pearson WS. Race, rural residence, and control of diabetes and hypertension. *Annals of family medicine*. 2004;2:563-8.
29. Singh GK, Daus GP, Allender M, Ramey CT, Martin EK, Perry C, Reyes AAL and Vedamuthu IP. Social Determinants of Health in the United States: Addressing Major Health Inequality Trends for the Nation, 1935-2016. *International journal of MCH and AIDS*. 2017;6:139-164.
30. Howard G, Kleindorfer DO, Cushman M, Long DL, Jasne A, Judd SE, Higginbotham JC and Howard VJ. Contributors to the Excess Stroke Mortality in Rural Areas in the United States. *Stroke; a journal of cerebral circulation*. 2017;48:1773-1778.
31. Fye WB. Tracing atrial fibrillation--100 years. *The New England journal of medicine*. 2006;355:1412-4.
32. Gallagher MM and Camm J. Classification of atrial fibrillation. *The American journal of cardiology*. 1998;82:18N-28N.
33. Roukoz HL, F; Sakguchi, S. *Catheter Ablation of Cardiac Arrythmias.* 3rd edition ed. New York, NY: Springer; 2015.
34. Gutierrez C and Blanchard DG. Diagnosis and Treatment of Atrial Fibrillation. *American family physician*. 2016;94:442-52.
35. January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Jr., Conti JB, Ellinor PT, Ezekowitz MD, Field ME, Murray KT, Sacco RL, Stevenson WG, Tchou PJ, Tracy CM and Yancy CW. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on practice guidelines and the Heart Rhythm Society. *Circulation*. 2014;130:e199-267.
36. Allessie M, Ausma J and Schotten U. Electrical, contractile and structural remodeling during atrial fibrillation. *Cardiovascular research*. 2002;54:230-46.
37. He B, Scherlag BJ, Nakagawa H, Lazzara R and Po SS. The intrinsic autonomic nervous system in atrial fibrillation: a review. *ISRN cardiology*. 2012;2012:490674.
38. Nattel S, Burstein B and Dobrev D. Atrial remodeling and atrial fibrillation: mechanisms and implications. *Circulation Arrhythmia and electrophysiology*. 2008;1:62-73.
39. Cooke G, Doust J and Sanders S. Is pulse palpation helpful in detecting atrial fibrillation? A systematic review. *The Journal of family practice*. 2006;55:130-4.
40. Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, Selby JV and Singer DE. 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;285:2370-5.
41. Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, Seward JB and Tsang TS. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation*. 2006;114:119-25.
42. Colilla S, Crow A, Petkun W, Singer DE, Simon T and Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. *The American journal of cardiology*. 2013;112:1142-7.

43. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ and Wolf PA. Independent risk factors for atrial fibrillation in a population-based cohort. The Framingham Heart Study. *Jama*. 1994;271:840-4.
44. Psaty BM, Manolio TA, Kuller LH, Kronmal RA, Cushman M, Fried LP, White R, Furberg CD and Rautaharju PM. Incidence of and risk factors for atrial fibrillation in older adults. *Circulation*. 1997;96:2455-61.
45. Piccini JP, Hammill BG, Sinner MF, Jensen PN, Hernandez AF, Heckbert SR, Benjamin EJ and Curtis LH. Incidence and prevalence of atrial fibrillation and associated mortality among Medicare beneficiaries, 1993-2007. *Circulation Cardiovascular quality and outcomes*. 2012;5:85-93.
46. Kannel WB, Wolf PA, Benjamin EJ and Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *The American journal of cardiology*. 1998;82:2N-9N.
47. Feinberg WM, Blackshear JL, Laupacis A, Kronmal R and Hart RG. Prevalence, age distribution, and gender of patients with atrial fibrillation. Analysis and implications. *Archives of internal medicine*. 1995;155:469-73.
48. Rich MW. Epidemiology of atrial fibrillation. *Journal of interventional cardiac electrophysiology : an international journal of arrhythmias and pacing*. 2009;25:3-8.
49. Furberg CD, Psaty BM, Manolio TA, Gardin JM, Smith VE and Rautaharju PM. Prevalence of atrial fibrillation in elderly subjects (the Cardiovascular Health Study). *The American journal of cardiology*. 1994;74:236-41.
50. Chamberlain AM, Agarwal SK, Ambrose M, Folsom AR, Soliman EZ and Alonso A. Metabolic syndrome and incidence of atrial fibrillation among blacks and whites in the Atherosclerosis Risk in Communities (ARIC) Study. *American heart journal*. 2010;159:850-6.
51. Chamberlain AM, Agarwal SK, Folsom AR, Soliman EZ, Chambless LE, Crow R, Ambrose M and Alonso A. A clinical risk score for atrial fibrillation in a biracial prospective cohort (from the Atherosclerosis Risk in Communities [ARIC] study). *The American journal of cardiology*. 2011;107:85-91.
52. Dewland TA, Olgin JE, Vittinghoff E and Marcus GM. Incident atrial fibrillation among Asians, Hispanics, blacks, and whites. *Circulation*. 2013;128:2470-7.
53. Rodriguez CJ, Soliman EZ, Alonso A, Swett K, Okin PM, Goff DC, Jr. and Heckbert SR. Atrial fibrillation incidence and risk factors in relation to race-ethnicity and the population attributable fraction of atrial fibrillation risk factors: the Multi-Ethnic Study of Atherosclerosis. *Annals of epidemiology*. 2015;25:71-6, 76 e1.
54. Ettinger PO WC, De La Cruz C, Jr. et al. . Arrhythmias and the "Holiday Heart": alcohol-associated cardiac rhythm disorders. *American heart journal*. 1978;95:555-562.
55. Djousse L, Levy D, Benjamin EJ, Blease SJ, Russ A, Larson MG, Massaro JM, D'Agostino RB, Wolf PA and Ellison RC. Long-term alcohol consumption and the risk of atrial fibrillation in the Framingham Study. *The American journal of cardiology*. 2004;93:710-3.
56. Mukamal KJ, Psaty BM, Rautaharju PM, Furberg CD, Kuller LH, Mittleman MA, Gottdiener JS and Siscovick DS. Alcohol consumption and risk and prognosis of atrial fibrillation among older adults: the Cardiovascular Health Study. *American heart journal*. 2007;153:260-6.
57. Kodama S, Saito K, Tanaka S, Horikawa C, Saito A, Heianza Y, Anasako Y, Nishigaki Y, Yachi Y, Iida KT, Ohashi Y, Yamada N and Sone H. Alcohol consumption and risk of atrial fibrillation: a meta-analysis. *Journal of the American College of Cardiology*. 2011;57:427-36.
58. Aizer A, Gaziano JM, Cook NR, Manson JE, Buring JE and Albert CM. Relation of vigorous exercise to risk of atrial fibrillation. *The American journal of cardiology*. 2009;103:1572-7.

59. Hoogsteen J, Schep G, Van Hemel NM and Van Der Wall EE. Paroxysmal atrial fibrillation in male endurance athletes. A 9-year follow up. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*. 2004;6:222-8.
60. Furlanello F, Bertoldi A, Dallago M, Galassi A, Fernando F, Biffi A, Mazzone P, Pappone C and Chierchia S. Atrial fibrillation in elite athletes. *Journal of cardiovascular electrophysiology*. 1998;9:S63-8.
61. Karjalainen J, Kujala UM, Kaprio J, Sarna S and Viitasalo M. Lone atrial fibrillation in vigorously exercising middle aged men: case-control study. *BMJ*. 1998;316:1784-5.
62. Frost L, Frost P and Vestergaard P. Work related physical activity and risk of a hospital discharge diagnosis of atrial fibrillation or flutter: the Danish Diet, Cancer, and Health Study. *Occupational and environmental medicine*. 2005;62:49-53.
63. Everett BM, Conen D, Buring JE, Moorthy MV, Lee IM and Albert CM. Physical activity and the risk of incident atrial fibrillation in women. *Circulation Cardiovascular quality and outcomes*. 2011;4:321-7.
64. Mozaffarian D, Furberg CD, Psaty BM and Siscovick D. Physical activity and incidence of atrial fibrillation in older adults: the cardiovascular health study. *Circulation*. 2008;118:800-7.
65. Huxley RR, Lopez FL, Folsom AR, Agarwal SK, Loehr LR, Soliman EZ, Maclehose R, Konety S and Alonso A. Absolute and attributable risks of atrial fibrillation in relation to optimal and borderline risk factors: the Atherosclerosis Risk in Communities (ARIC) study. *Circulation*. 2011;123:1501-8.
66. Chamberlain AM, Agarwal SK, Folsom AR, Duval S, Soliman EZ, Ambrose M, Eberly LE and Alonso A. Smoking and incidence of atrial fibrillation: results from the Atherosclerosis Risk in Communities (ARIC) study. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2011;8:1160-6.
67. Heeringa J, Kors JA, Hofman A, van Rooij FJ and Witteman JC. Cigarette smoking and risk of atrial fibrillation: the Rotterdam Study. *American heart journal*. 2008;156:1163-9.
68. Krahn AD, Manfreda J, Tate RB, Mathewson FA and Cuddy TE. The natural history of atrial fibrillation: incidence, risk factors, and prognosis in the Manitoba Follow-Up Study. *The American journal of medicine*. 1995;98:476-84.
69. Huxley RR, Alonso A, Lopez FL, Filion KB, Agarwal SK, Loehr LR, Soliman EZ, Pankow JS and Selvin E. Type 2 diabetes, glucose homeostasis and incident atrial fibrillation: the Atherosclerosis Risk in Communities study. *Heart*. 2012;98:133-8.
70. Huxley RR, Filion KB, Konety S and Alonso A. Meta-analysis of cohort and case-control studies of type 2 diabetes mellitus and risk of atrial fibrillation. *The American journal of cardiology*. 2011;108:56-62.
71. Wanahita N, Messerli FH, Bangalore S, Gami AS, Somers VK and Steinberg JS. Atrial fibrillation and obesity--results of a meta-analysis. *American heart journal*. 2008;155:310-5.
72. Alonso A, Lopez FL, Matsushita K, Loehr LR, Agarwal SK, Chen LY, Soliman EZ, Astor BC and Coresh J. Chronic kidney disease is associated with the incidence of atrial fibrillation: the Atherosclerosis Risk in Communities (ARIC) study. *Circulation*. 2011;123:2946-53.
73. Alonso A, Yin X, Roetker NS, Magnani JW, Kronmal RA, Ellinor PT, Chen LY, Lubitz SA, McClelland RL, McManus DD, Soliman EZ, Huxley RR, Nazarian S, Szklo M, Heckbert SR and Benjamin EJ. Blood lipids and the incidence of atrial fibrillation: the multi-ethnic study of atherosclerosis and the framingham heart study. *Journal of the American Heart Association*. 2014;3.
74. Lopez FL, Agarwal SK, Maclehose RF, Soliman EZ, Sharrett AR, Huxley RR, Konety S, Ballantyne CM and Alonso A. Blood lipid levels, lipid-lowering medications, and the incidence

of atrial fibrillation: the atherosclerosis risk in communities study. *Circulation Arrhythmia and electrophysiology*. 2012;5:155-62.
75. Agarwal SK, Norby FL, Whitsel EA, Soliman EZ, Chen LY, Loehr LR, Fuster V, Heiss G, Coresh J and Alonso A. Cardiac Autonomic Dysfunction and Incidence of Atrial Fibrillation: Results From 20 Years Follow-Up. *Journal of the American College of Cardiology*. 2017;69:291-299.
76. Wang W, Alonso A, Soliman EZ, O'Neal WT, Calkins H, Chen LY, Diener-West M and Szklo M. Relation of Resting Heart Rate to Incident Atrial Fibrillation (From ARIC [Atherosclerosis Risk in Communities] Study). *The American journal of cardiology*. 2018;121:1169-1176.
77. O'Neal WT, Almahmoud MF and Soliman EZ. Resting heart rate and incident atrial fibrillation in the elderly. *Pacing and clinical electrophysiology : PACE*. 2015;38:591-7.
78. Misialek JR, Lopez FL, Lutsey PL, Huxley RR, Peacock JM, Chen LY, Soliman EZ, Agarwal SK and Alonso A. Serum and dietary magnesium and incidence of atrial fibrillation in whites and in African Americans--Atherosclerosis Risk in Communities (ARIC) study. *Circulation journal : official journal of the Japanese Circulation Society*. 2013;77:323-9.
79. Lopez FL, Agarwal SK, Grams ME, Loehr LR, Soliman EZ, Lutsey PL, Chen LY, Huxley RR and Alonso A. Relation of serum phosphorus levels to the incidence of atrial fibrillation (from the Atherosclerosis Risk In Communities [ARIC] study). *The American journal of cardiology*. 2013;111:857-62.
80. Norby FL, Soliman EZ, Chen LY, Bengtson LG, Loehr LR, Agarwal SK and Alonso A. Trajectories of Cardiovascular Risk Factors and Incidence of Atrial Fibrillation Over a 25-Year Follow-Up: The ARIC Study (Atherosclerosis Risk in Communities). *Circulation*. 2016;134:599-610.
81. Schnabel RB, Wild PS, Wilde S, Ojeda FM, Schulz A, Zeller T, Sinning CR, Kunde J, Lackner KJ, Munzel T and Blankenberg S. Multiple biomarkers and atrial fibrillation in the general population. *PloS one*. 2014;9:e112486.
82. Smith JG, Newton-Cheh C, Almgren P, Struck J, Morgenthaler NG, Bergmann A, Platonov PG, Hedblad B, Engstrom G, Wang TJ and Melander O. Assessment of conventional cardiovascular risk factors and multiple biomarkers for the prediction of incident heart failure and atrial fibrillation. *Journal of the American College of Cardiology*. 2010;56:1712-9.
83. Ho JE, Yin X, Levy D, Vasan RS, Magnani JW, Ellinor PT, McManus DD, Lubitz SA, Larson MG and Benjamin EJ. Galectin 3 and incident atrial fibrillation in the community. *American heart journal*. 2014;167:729-34 e1.
84. Fashanu OE, Norby FL, Aguilar D, Ballantyne CM, Hoogeveen RC, Chen LY, Soliman EZ, Alonso A and Folsom AR. Galectin-3 and incidence of atrial fibrillation: The Atherosclerosis Risk in Communities (ARIC) study. *American heart journal*. 2017;192:19-25.
85. Schnabel RB, Sullivan LM, Levy D, Pencina MJ, Massaro JM, D'Agostino RB, Sr., Newton-Cheh C, Yamamoto JF, Magnani JW, Tadros TM, Kannel WB, Wang TJ, Ellinor PT, Wolf PA, Vasan RS and Benjamin EJ. Development of a risk score for atrial fibrillation (Framingham Heart Study): a community-based cohort study. *Lancet*. 2009;373:739-45.
86. Aeschbacher S, O'Neal WT, Krisai P, Loehr L, Chen LY, Alonso A, Soliman EZ and Conen D. Relationship between QRS duration and incident atrial fibrillation. *International journal of cardiology*. 2018;266:84-88.
87. Mandyam MC, Soliman EZ, Alonso A, Dewland TA, Heckbert SR, Vittinghoff E, Cummings SR, Ellinor PT, Chaitman BR, Stocke K, Applegate WB, Arking DE, Butler J, Loehr LR, Magnani JW, Murphy RA, Satterfield S, Newman AB and Marcus GM. The QT interval and risk of incident atrial fibrillation. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2013;10:1562-8.

88. Magnani JW, Zhu L, Lopez F, Pencina MJ, Agarwal SK, Soliman EZ, Benjamin EJ and Alonso A. P-wave indices and atrial fibrillation: cross-cohort assessments from the Framingham Heart Study (FHS) and Atherosclerosis Risk in Communities (ARIC) study. *American heart journal*. 2015;169:53-61 e1.
89. Maheshwari A, Norby FL, Roetker NS, Soliman EZ, Koene RJ, Rooney MR, O'Neal WT, Shah AM, Claggett BL, Solomon SD, Alonso A, Gottesman RF, Heckbert SR and Chen LY. Refining Prediction of Atrial Fibrillation-Related Stroke Using the P2-CHA2DS2-VASc Score. *Circulation*. 2019;139:180-191.
90. Dewland TA, Vittinghoff E, Mandyam MC, Heckbert SR, Siscovick DS, Stein PK, Psaty BM, Sotoodehnia N, Gottdiener JS and Marcus GM. Atrial ectopy as a predictor of incident atrial fibrillation: a cohort study. *Annals of internal medicine*. 2013;159:721-8.
91. Bekwelem W, Misialek JR, Konety S, Solomon SD, Soliman EZ, Loehr LR, Lopez FL, Fox ER, Mosley TH and Alonso A. Echocardiographic measures of cardiac structure and function are associated with risk of atrial fibrillation in blacks: the Atherosclerosis Risk in Communities (ARIC) study. *PloS one*. 2014;9:e110111.
92. Rosenberg MA, Gottdiener JS, Heckbert SR and Mukamal KJ. Echocardiographic diastolic parameters and risk of atrial fibrillation: the Cardiovascular Health Study. *European heart journal*. 2012;33:904-12.
93. Vaziri SM, Larson MG, Benjamin EJ and Levy D. Echocardiographic predictors of nonrheumatic atrial fibrillation. The Framingham Heart Study. *Circulation*. 1994;89:724-30.
94. Tsang TS, Gersh BJ, Appleton CP, Tajik AJ, Barnes ME, Bailey KR, Oh JK, Leibson C, Montgomery SC and Seward JB. Left ventricular diastolic dysfunction as a predictor of the first diagnosed nonvalvular atrial fibrillation in 840 elderly men and women. *Journal of the American College of Cardiology*. 2002;40:1636-44.
95. Ellinor PT, Lunetta KL, Albert CM, Glazer NL, Ritchie MD, Smith AV, Arking DE, Muller-Nurasyid M, Krijthe BP, Lubitz SA, Bis JC, Chung MK, Dorr M, Ozaki K, Roberts JD, Smith JG, Pfeufer A, Sinner MF, Lohman K, Ding J, Smith NL, Smith JD, Rienstra M, Rice KM, Van Wagoner DR, Magnani JW, Wakili R, Clauss S, Rotter JI, Steinbeck G, Launer LJ, Davies RW, Borkovich M, Harris TB, Lin H, Volker U, Volzke H, Milan DJ, Hofman A, Boerwinkle E, Chen LY, Soliman EZ, Voight BF, Li G, Chakravarti A, Kubo M, Tedrow UB, Rose LM, Ridker PM, Conen D, Tsunoda T, Furukawa T, Sotoodehnia N, Xu S, Kamatani N, Levy D, Nakamura Y, Parvez B, Mahida S, Furie KL, Rosand J, Muhammad R, Psaty BM, Meitinger T, Perz S, Wichmann HE, Witteman JC, Kao WH, Kathiresan S, Roden DM, Uitterlinden AG, Rivadeneira F, McKnight B, Sjogren M, Newman AB, Liu Y, Gollob MH, Melander O, Tanaka T, Stricker BH, Felix SB, Alonso A, Darbar D, Barnard J, Chasman DI, Heckbert SR, Benjamin EJ, Gudnason V and Kaab S. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nature genetics*. 2012;44:670-5.
96. Benjamin EJ, Rice KM, Arking DE, Pfeufer A, van Noord C, Smith AV, Schnabel RB, Bis JC, Boerwinkle E, Sinner MF, Dehghan A, Lubitz SA, D'Agostino RB, Sr., Lumley T, Ehret GB, Heeringa J, Aspelund T, Newton-Cheh C, Larson MG, Marciante KD, Soliman EZ, Rivadeneira F, Wang TJ, Eiriksdottir G, Levy D, Psaty BM, Li M, Chamberlain AM, Hofman A, Vasan RS, Harris TB, Rotter JI, Kao WH, Agarwal SK, Stricker BH, Wang K, Launer LJ, Smith NL, Chakravarti A, Uitterlinden AG, Wolf PA, Sotoodehnia N, Kottgen A, van Duijn CM, Meitinger T, Mueller M, Perz S, Steinbeck G, Wichmann HE, Lunetta KL, Heckbert SR, Gudnason V, Alonso A, Kaab S, Ellinor PT and Witteman JC. Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nature genetics*. 2009;41:879-81.
97. Ellinor PT, Lunetta KL, Glazer NL, Pfeufer A, Alonso A, Chung MK, Sinner MF, de Bakker PI, Mueller M, Lubitz SA, Fox E, Darbar D, Smith NL, Smith JD, Schnabel RB, Soliman EZ, Rice KM, Van Wagoner DR, Beckmann BM, van Noord C, Wang K, Ehret GB, Rotter JI, Hazen SL, Steinbeck G, Smith AV, Launer LJ, Harris TB, Makino S, Nelis M, Milan DJ, Perz S,

Esko T, Kottgen A, Moebus S, Newton-Cheh C, Li M, Mohlenkamp S, Wang TJ, Kao WH, Vasan RS, Nothen MM, MacRae CA, Stricker BH, Hofman A, Uitterlinden AG, Levy D, Boerwinkle E, Metspalu A, Topol EJ, Chakravarti A, Gudnason V, Psaty BM, Roden DM, Meitinger T, Wichmann HE, Witteman JC, Barnard J, Arking DE, Benjamin EJ, Heckbert SR and Kaab S. Common variants in KCNN3 are associated with lone atrial fibrillation. *Nature genetics*. 2010;42:240-4.
98. Gudbjartsson DF, Arnar DO, Helgadottir A, Gretarsdottir S, Holm H, Sigurdsson A, Jonasdottir A, Baker A, Thorleifsson G, Kristjansson K, Palsson A, Blondal T, Sulem P, Backman VM, Hardarson GA, Palsdottir E, Helgason A, Sigurjonsdottir R, Sverrisson JT, Kostulas K, Ng MC, Baum L, So WY, Wong KS, Chan JC, Furie KL, Greenberg SM, Sale M, Kelly P, MacRae CA, Smith EE, Rosand J, Hillert J, Ma RC, Ellinor PT, Thorgeirsson G, Gulcher JR, Kong A, Thorsteinsdottir U and Stefansson K. Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature*. 2007;448:353-7.
99. Pfeufer A, van Noord C, Marciante KD, Arking DE, Larson MG, Smith AV, Tarasov KV, Muller M, Sotoodehnia N, Sinner MF, Verwoert GC, Li M, Kao WH, Kottgen A, Coresh J, Bis JC, Psaty BM, Rice K, Rotter JI, Rivadeneira F, Hofman A, Kors JA, Stricker BH, Uitterlinden AG, van Duijn CM, Beckmann BM, Sauter W, Gieger C, Lubitz SA, Newton-Cheh C, Wang TJ, Magnani JW, Schnabel RB, Chung MK, Barnard J, Smith JD, Van Wagoner DR, Vasan RS, Aspelund T, Eiriksdottir G, Harris TB, Launer LJ, Najjar SS, Lakatta E, Schlessinger D, Uda M, Abecasis GR, Muller-Myhsok B, Ehret GB, Boerwinkle E, Chakravarti A, Soliman EZ, Lunetta KL, Perz S, Wichmann HE, Meitinger T, Levy D, Gudnason V, Ellinor PT, Sanna S, Kaab S, Witteman JC, Alonso A, Benjamin EJ and Heckbert SR. Genome-wide association study of PR interval. *Nature genetics*. 2010;42:153-9.
100. Holm H, Gudbjartsson DF, Arnar DO, Thorleifsson G, Thorgeirsson G, Stefansdottir H, Gudjonsson SA, Jonasdottir A, Mathiesen EB, Njolstad I, Nyrnes A, Wilsgaard T, Hald EM, Hveem K, Stoltenberg C, Lochen ML, Kong A, Thorsteinsdottir U and Stefansson K. Several common variants modulate heart rate, PR interval and QRS duration. *Nature genetics*. 2010;42:117-22.
101. Hart RG and Halperin JL. Atrial fibrillation and stroke : concepts and controversies. *Stroke; a journal of cerebral circulation*. 2001;32:803-8.
102. Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Seward JB, Bailey KR, Iwasaka T and Tsang TS. Time trends of ischemic stroke incidence and mortality in patients diagnosed with first atrial fibrillation in 1980 to 2000: report of a community-based study. *Stroke; a journal of cerebral circulation*. 2005;36:2362-6.
103. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation. Analysis of pooled data from five randomized controlled trials. *Archives of internal medicine*. 1994;154:1449-57.
104. Miller VT, Rothrock JF, Pearce LA, Feinberg WM, Hart RG and Anderson DC. Ischemic stroke in patients with atrial fibrillation: effect of aspirin according to stroke mechanism. Stroke Prevention in Atrial Fibrillation Investigators. *Neurology*. 1993;43:32-6.
105. Bogousslavsky J, Van Melle G, Regli F and Kappenberger L. Pathogenesis of anterior circulation stroke in patients with nonvalvular atrial fibrillation: the Lausanne Stroke Registry. *Neurology*. 1990;40:1046-50.
106. Soliman EZ, Safford MM, Muntner P, Khodneva Y, Dawood FZ, Zakai NA, Thacker EL, Judd S, Howard VJ, Howard G, Herrington DM and Cushman M. Atrial fibrillation and the risk of myocardial infarction. *JAMA internal medicine*. 2014;174:107-14.
107. Stewart S, Hart CL, Hole DJ and McMurray JJ. A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. *The American journal of medicine*. 2002;113:359-64.

108. Schnabel RB, Rienstra M, Sullivan LM, Sun JX, Moser CB, Levy D, Pencina MJ, Fontes JD, Magnani JW, McManus DD, Lubitz SA, Tadros TM, Wang TJ, Ellinor PT, Vasan RS and Benjamin EJ. Risk assessment for incident heart failure in individuals with atrial fibrillation. *European journal of heart failure*. 2013;15:843-9.
109. Wang TJ, Larson MG, Levy D, Vasan RS, Leip EP, Wolf PA, D'Agostino RB, Murabito JM, Kannel WB and Benjamin EJ. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. *Circulation*. 2003;107:2920-5.
110. Odutayo A, Wong CX, Hsiao AJ, Hopewell S, Altman DG and Emdin CA. Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. *BMJ*. 2016;354:i4482.
111. Kwok CS, Loke YK, Hale R, Potter JF and Myint PK. Atrial fibrillation and incidence of dementia: a systematic review and meta-analysis. *Neurology*. 2011;76:914-22.
112. Miyasaka Y, Barnes ME, Petersen RC, Cha SS, Bailey KR, Gersh BJ, Casaclang-Verzosa G, Abhayaratna WP, Seward JB, Iwasaka T and Tsang TS. Risk of dementia in stroke-free patients diagnosed with atrial fibrillation: data from a community-based cohort. *European heart journal*. 2007;28:1962-7.
113. Magnani JW, Norby FL, Agarwal SK, Soliman EZ, Chen LY, Loehr LR and Alonso A. Racial Differences in Atrial Fibrillation-Related Cardiovascular Disease and Mortality: The Atherosclerosis Risk in Communities (ARIC) Study. *JAMA cardiology*. 2016;1:433-41.
114. Chen LY, Norby FL, Gottesman RF, Mosley TH, Soliman EZ, Agarwal SK, Loehr LR, Folsom AR, Coresh J and Alonso A. Association of Atrial Fibrillation With Cognitive Decline and Dementia Over 20 Years: The ARIC-NCS (Atherosclerosis Risk in Communities Neurocognitive Study). *Journal of the American Heart Association*. 2018;7.
115. Chen LY, Lopez FL, Gottesman RF, Huxley RR, Agarwal SK, Loehr L, Mosley T and Alonso A. Atrial fibrillation and cognitive decline-the role of subclinical cerebral infarcts: the atherosclerosis risk in communities study. *Stroke; a journal of cerebral circulation*. 2014;45:2568-74.
116. Potpara TS and Lip GY. Oral Anticoagulant Therapy in Atrial Fibrillation Patients at High Stroke and Bleeding Risk. *Progress in cardiovascular diseases*. 2015;58:177-94.
117. Alonso A and Norby FL. Predicting Atrial Fibrillation and Its Complications. *Circulation journal : official journal of the Japanese Circulation Society*. 2016;80:1061-6.
118. Lutsey PL, Norby FL, Zakai NA, MacLehose RF, Chen LY, Shah S, Datta YH and Alonso A. Oral anticoagulation therapy and subsequent risk of venous thromboembolism in atrial fibrillation patients. *Current medical research and opinion*. 2019;35:837-845.
119. Norby FL and Alonso A. Comparative effectiveness of rivaroxaban in the treatment of nonvalvular atrial fibrillation. *Journal of comparative effectiveness research*. 2017;6:549-560.
120. Norby FL, Bengtson LGS, Lutsey PL, Chen LY, MacLehose RF, Chamberlain AM, Rapson I and Alonso A. Comparative effectiveness of rivaroxaban versus warfarin or dabigatran for the treatment of patients with non-valvular atrial fibrillation. *BMC cardiovascular disorders*. 2017;17:238.
121. Gage BF, Yan Y, Milligan PE, Waterman AD, Culverhouse R, Rich MW and Radford MJ. Clinical classification schemes for predicting hemorrhage: results from the National Registry of Atrial Fibrillation (NRAF). *American heart journal*. 2006;151:713-9.
122. Pisters R, Lane DA, Nieuwlaat R, de Vos CB, Crijns HJ and Lip GY. A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey. *Chest*. 2010;138:1093-100.
123. Fang MC, Go AS, Chang Y, Borowsky LH, Pomernacki NK, Udaltsova N and Singer DE. A new risk scheme to predict warfarin-associated hemorrhage: The ATRIA (Anticoagulation

and Risk Factors in Atrial Fibrillation) Study. *Journal of the American College of Cardiology*. 2011;58:395-401.
124. O'Brien EC, Simon DN, Thomas LE, Hylek EM, Gersh BJ, Ansell JE, Kowey PR, Mahaffey KW, Chang P, Fonarow GC, Pencina MJ, Piccini JP and Peterson ED. The ORBIT bleeding score: a simple bedside score to assess bleeding risk in atrial fibrillation. *European heart journal*. 2015;36:3258-64.
125. Barnes GD, Gu X, Haymart B, Kline-Rogers E, Almany S, Kozlowski J, Besley D, Krol GD, Froehlich JB and Kaatz S. The predictive ability of the CHADS2 and CHA2DS2-VASc scores for bleeding risk in atrial fibrillation: the MAQI(2) experience. *Thrombosis research*. 2014;134:294-9.
126. Senoo K, Proietti M, Lane DA and Lip GY. Evaluation of the HAS-BLED, ATRIA, and ORBIT Bleeding Risk Scores in Patients with Atrial Fibrillation Taking Warfarin. *The American journal of medicine*. 2016;129:600-7.
127. Zhu W, He W, Guo L, Wang X and Hong K. The HAS-BLED Score for Predicting Major Bleeding Risk in Anticoagulated Patients With Atrial Fibrillation: A Systematic Review and Meta-analysis. *Clinical cardiology*. 2015;38:555-61.
128. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB and Levy D. Impact of atrial fibrillation on the risk of death: the Framingham Heart Study. *Circulation*. 1998;98:946-52.
129. Vidaillet H, Granada JF, Chyou P, Maassen K, Ortiz M, Pulido JN, Sharma P, Smith PN and Hayes J. A population-based study of mortality among patients with atrial fibrillation or flutter. *The American journal of medicine*. 2002;113:365-70.
130. Gajewski J and Singer RB. Mortality in an insured population with atrial fibrillation. *Jama*. 1981;245:1540-4.
131. Emdin CA, Wong CX, Hsiao AJ, Altman DG, Peters SA, Woodward M and Odutayo AA. Atrial fibrillation as risk factor for cardiovascular disease and death in women compared with men: systematic review and meta-analysis of cohort studies. *BMJ*. 2016;532:h7013.
132. Conen D, Chae CU, Glynn RJ, Tedrow UB, Everett BM, Buring JE and Albert CM. Risk of death and cardiovascular events in initially healthy women with new-onset atrial fibrillation. *Jama*. 2011;305:2080-7.
133. Kabra R, Cram P, Girotra S and Vaughan Sarrazin M. Effect of race on outcomes (stroke and death) in patients >65 years with atrial fibrillation. *The American journal of cardiology*. 2015;116:230-5.
134. Turagam MK, Velagapudi P, Visotcky A, Szabo A and Kocheril AG. African Americans have the highest risk of in-hospital mortality with atrial fibrillation related hospitalizations among all racial/ethnic groups: a nationwide analysis. *International journal of cardiology*. 2012;158:165-6.
135. Marijon E, Le Heuzey JY, Connolly S, Yang S, Pogue J, Brueckmann M, Eikelboom J, Themeles E, Ezekowitz M, Wallentin L and Yusuf S. Causes of death and influencing factors in patients with atrial fibrillation: a competing-risk analysis from the randomized evaluation of long-term anticoagulant therapy study. *Circulation*. 2013;128:2192-201.
136. Chen LY, Sotoodehnia N, Buzkova P, Lopez FL, Yee LM, Heckbert SR, Prineas R, Soliman EZ, Adabag S, Konety S, Folsom AR, Siscovick D and Alonso A. Atrial fibrillation and the risk of sudden cardiac death: the atherosclerosis risk in communities study and cardiovascular health study. *JAMA internal medicine*. 2013;173:29-35.
137. Van Gelder IC and Groenveld HF. Rate Control in Atrial Fibrillation. *Cardiac electrophysiology clinics*. 2010;2:419-427.
138. Olshansky B, Rosenfeld LE, Warner AL, Solomon AJ, O'Neill G, Sharma A, Platia E, Feld GK, Akiyama T, Brodsky MA and Greene HL. The Atrial Fibrillation Follow-up

Investigation of Rhythm Management (AFFIRM) study: approaches to control rate in atrial fibrillation. *Journal of the American College of Cardiology*. 2004;43:1201-8.
139. Van Gelder IC, Groenveld HF, Crijns HJ, Tuininga YS, Tijssen JG, Alings AM, Hillege HL, Bergsma-Kadijk JA, Cornel JH, Kamp O, Tukkie R, Bosker HA, Van Veldhuisen DJ and Van den Berg MP. Lenient versus strict rate control in patients with atrial fibrillation. *The New England journal of medicine*. 2010;362:1363-73.
140. Camm AJ, Lip GY, De Caterina R, Savelieva I, Atar D, Hohnloser SH, Hindricks G and Kirchhof P. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation: an update of the 2010 ESC Guidelines for the management of atrial fibrillation. Developed with the special contribution of the European Heart Rhythm Association. *European heart journal*. 2012;33:2719-47.
141. Arbelo E, Brugada J, Hindricks G, Maggioni A, Tavazzi L, Vardas P, Anselme F, Inama G, Jais P, Kalarus Z, Kautzner J, Lewalter T, Mairesse G, Perez-Villacastin J, Riahi S, Taborsky M, Theodorakis G and Trines S. ESC-EURObservational Research Programme: the Atrial Fibrillation Ablation Pilot Study, conducted by the European Heart Rhythm Association. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*. 2012;14:1094-103.
142. Cappato R, Calkins H, Chen SA, Davies W, Iesaka Y, Kalman J, Kim YH, Klein G, Natale A, Packer D, Skanes A, Ambrogi F and Biganzoli E. Updated worldwide survey on the methods, efficacy, and safety of catheter ablation for human atrial fibrillation. *Circulation Arrhythmia and electrophysiology*. 2010;3:32-8.
143. Damiano RJ, Jr., Gaynor SL, Bailey M, Prasad S, Cox JL, Boineau JP and Schuessler RP. The long-term outcome of patients with coronary disease and atrial fibrillation undergoing the Cox maze procedure. *The Journal of thoracic and cardiovascular surgery*. 2003;126:2016-21.
144. Holmes DR, Jr., Doshi SK, Kar S, Price MJ, Sanchez JM, Sievert H, Valderrabano M and Reddy VY. Left Atrial Appendage Closure as an Alternative to Warfarin for Stroke Prevention in Atrial Fibrillation: A Patient-Level Meta-Analysis. *Journal of the American College of Cardiology*. 2015;65:2614-2623.
145. Holmes DR, Reddy VY, Turi ZG, Doshi SK, Sievert H, Buchbinder M, Mullin CM and Sick P. Percutaneous closure of the left atrial appendage versus warfarin therapy for prevention of stroke in patients with atrial fibrillation: a randomised non-inferiority trial. *Lancet*. 2009;374:534-42.
146. Alli O, Asirvatham S and Holmes DR, Jr. Strategies to incorporate left atrial appendage occlusion into clinical practice. *Journal of the American College of Cardiology*. 2015;65:2337-44.
147. Bartus K, Han FT, Bednarek J, Myc J, Kapelak B, Sadowski J, Lelakowski J, Bartus S, Yakubov SJ and Lee RJ. Percutaneous left atrial appendage suture ligation using the LARIAT device in patients with atrial fibrillation: initial clinical experience. *Journal of the American College of Cardiology*. 2013;62:108-118.
148. Sick PB, Schuler G, Hauptmann KE, Grube E, Yakubov S, Turi ZG, Mishkel G, Almany S and Holmes DR. Initial worldwide experience with the WATCHMAN left atrial appendage system for stroke prevention in atrial fibrillation. *Journal of the American College of Cardiology*. 2007;49:1490-5.
149. Rietbrock S, Heeley E, Plumb J and van Staa T. Chronic atrial fibrillation: Incidence, prevalence, and prediction of stroke using the Congestive heart failure, Hypertension, Age >75, Diabetes mellitus, and prior Stroke or transient ischemic attack (CHADS2) risk stratification scheme. *American heart journal*. 2008;156:57-64.

150. Hart RG, Pearce LA and Aguilar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Annals of internal medicine*. 2007;146:857-67.
151. Go AS, Hylek EM, Borowsky LH, Phillips KA, Selby JV and Singer DE. Warfarin use among ambulatory patients with nonvalvular atrial fibrillation: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study. *Annals of internal medicine*. 1999;131:927-34.
152. Patel MR, Mahaffey KW, Garg J, Pan G, Singer DE, Hacke W, Breithardt G, Halperin JL, Hankey GJ, Piccini JP, Becker RC, Nessel CC, Paolini JF, Berkowitz SD, Fox KA and Califf RM. Rivaroxaban versus warfarin in nonvalvular atrial fibrillation. *The New England journal of medicine*. 2011;365:883-91.
153. Connolly SJ, Ezekowitz MD, Yusuf S, Eikelboom J, Oldgren J, Parekh A, Pogue J, Reilly PA, Themeles E, Varrone J, Wang S, Alings M, Xavier D, Zhu J, Diaz R, Lewis BS, Darius H, Diener HC, Joyner CD and Wallentin L. Dabigatran versus warfarin in patients with atrial fibrillation. *The New England journal of medicine*. 2009;361:1139-51.
154. Granger CB, Alexander JH, McMurray JJ, Lopes RD, Hylek EM, Hanna M, Al-Khalidi HR, Ansell J, Atar D, Avezum A, Bahit MC, Diaz R, Easton JD, Ezekowitz JA, Flaker G, Garcia D, Geraldes M, Gersh BJ, Golitsyn S, Goto S, Hermosillo AG, Hohnloser SH, Horowitz J, Mohan P, Jansky P, Lewis BS, Lopez-Sendon JL, Pais P, Parkhomenko A, Verheugt FW, Zhu J and Wallentin L. Apixaban versus warfarin in patients with atrial fibrillation. *The New England journal of medicine*. 2011;365:981-92.
155. Giugliano RP, Ruff CT, Braunwald E, Murphy SA, Wiviott SD, Halperin JL, Waldo AL, Ezekowitz MD, Weitz JI, Spinar J, Ruzyllo W, Ruda M, Koretsune Y, Betcher J, Shi M, Grip LT, Patel SP, Patel I, Hanyok JJ, Mercuri M and Antman EM. Edoxaban versus warfarin in patients with atrial fibrillation. *The New England journal of medicine*. 2013;369:2093-104.
156. Bengtson LG, Lutsey PL, Chen LY, MacLehose RF and Alonso A. Comparative effectiveness of dabigatran and rivaroxaban versus warfarin for the treatment of non-valvular atrial fibrillation. *Journal of cardiology*. 2016.
157. Abraham NS, Singh S, Alexander GC, Heien H, Haas LR, Crown W and Shah ND. Comparative risk of gastrointestinal bleeding with dabigatran, rivaroxaban, and warfarin: population based cohort study. *BMJ*. 2015;350:h1857.
158. Alonso A, MacLehose RF, Chen LY, Bengtson LG, Chamberlain AM, Norby FL and Lutsey PL. Prospective study of oral anticoagulants and risk of liver injury in patients with atrial fibrillation. *Heart*. 2017;103:834-839.
159. Shah S, Norby FL, Datta YH, Lutsey PL, MacLehose RF, Chen LY and Alonso A. Comparative effectiveness of direct oral anticoagulants and warfarin in patients with cancer and atrial fibrillation. *Blood advances*. 2018;2:200-209.
160. Sankaranarayanan R, Kirkwood G, Dibb K and Garratt CJ. Comparison of Atrial Fibrillation in the Young versus That in the Elderly: A Review. *Cardiology research and practice*. 2013;2013:976976.
161. Bhave PD, Lu X, Girotra S, Kamel H and Vaughan Sarrazin MS. Race- and sex-related differences in care for patients newly diagnosed with atrial fibrillation. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2015;12:1406-12.
162. Turakhia MP, Hoang DD, Xu X, Frayne S, Schmitt S, Yang F, Phibbs CS, Than CT, Wang PJ and Heidenreich PA. Differences and trends in stroke prevention anticoagulation in primary care vs cardiology specialty management of new atrial fibrillation: The Retrospective Evaluation and Assessment of Therapies in AF (TREAT-AF) study. *American heart journal*. 2013;165:93-101 e1.
163. Perino AC, Fan J, Schmitt SK, Askari M, Kaiser DW, Deshmukh A, Heidenreich PA, Swan C, Narayan SM, Wang PJ and Turakhia MP. Treating Specialty and Outcomes in Newly

Diagnosed Atrial Fibrillation: From the TREAT-AF Study. *Journal of the American College of Cardiology*. 2017;70:78-86.
164. Fosbol EL, Holmes DN, Piccini JP, Thomas L, Reiffel JA, Mills RM, Kowey P, Mahaffey K, Gersh BJ and Peterson ED. Provider specialty and atrial fibrillation treatment strategies in United States community practice: findings from the ORBIT-AF registry. *Journal of the American Heart Association*. 2013;2:e000110.
165. O'Neal WT, Sandesara PB, Claxton JS, MacLehose RF, Chen LY, Bengtson LGS, Chamberlain AM, Norby FL, Lutsey PL and Alonso A. Provider Specialty, Anticoagulation Prescription Patterns, and Stroke Risk in Atrial Fibrillation. *Journal of the American Heart Association*. 2018;7.
166. Singh GK, Siahpush M, Azuine RE and Williams SD. Widening Socioeconomic and Racial Disparities in Cardiovascular Disease Mortality in the United States, 1969-2013. *International journal of MCH and AIDS*. 2015;3:106-18.
167. Singh GK and Siahpush M. Widening rural-urban disparities in life expectancy, U.S., 1969-2009. *American journal of preventive medicine*. 2014;46:e19-29.
168. Douthit N, Kiv S, Dwolatzky T and Biswas S. Exposing some important barriers to health care access in the rural USA. *Public health*. 2015;129:611-20.
169. Snyder JE, Jensen M, Nguyen NX, Filice CE and Joynt KE. Defining Rurality in Medicare Administrative Data. *Medical care*. 2017;55:e164-e169.
170. Buchmueller TC, Jacobson M and Wold C. How far to the hospital? The effect of hospital closures on access to care. *Journal of health economics*. 2006;25:740-61.
171. Burwell SM. Setting value-based payment goals--HHS efforts to improve U.S. health care. *The New England journal of medicine*. 2015;372:897-9.
172. O'Neal WT, Sandesara PB, Kelli HM, Venkatesh S and Soliman EZ. Urban-rural differences in mortality for atrial fibrillation hospitalizations in the United States. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2017.
173. Barnes GD, Lucas E, Alexander GC and Goldberger ZD. National Trends in Ambulatory Oral Anticoagulant Use. *The American journal of medicine*. 2015;128:1300-5 e2.
174. Desai NR, Krumme AA, Schneeweiss S, Shrank WH, Brill G, Pezalla EJ, Spettell CM, Brennan TA, Matlin OS, Avorn J and Choudhry NK. Patterns of initiation of oral anticoagulants in patients with atrial fibrillation- quality and cost implications. *The American journal of medicine*. 2014;127:1075-82 e1.
175. Gage BF, Boechler M, Doggette AL, Fortune G, Flaker GC, Rich MW and Radford MJ. Adverse outcomes and predictors of underuse of antithrombotic therapy in medicare beneficiaries with chronic atrial fibrillation. *Stroke; a journal of cerebral circulation*. 2000;31:822-7.
176. Flaker GC, McGowan DJ, Boechler M, Fortune G and Gage B. Underutilization of antithrombotic therapy in elderly rural patients with atrial fibrillation. *American heart journal*. 1999;137:307-12.
177. Hayden K and Schweinle W. Demographic considerations in anticoagulation therapy for atrial fibrillation. *South Dakota medicine : the journal of the South Dakota State Medical Association*. 2015;68:116-9.
178. Hernandez I, Saba S and Zhang Y. Geographic Variation in the Use of Oral Anticoagulation Therapy in Stroke Prevention in Atrial Fibrillation. *Stroke; a journal of cerebral circulation*. 2017;48:2289-2291.
179. Rural-Urban Commuting Area Codes (RUCAs): Rural Health Research Center. 2019.
180. USDA. Rural-Urban Communting Area Codes. 2020.
181. Toth M, Holmes M, Van Houtven C, Toles M, Weinberger M and Silberman P. Rural Medicare Beneficiaries Have Fewer Follow-up Visits and Greater Emergency Department Use Postdischarge. *Medical care*. 2015;53:800-8.

182. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Despres JP, Fullerton HJ, Howard VJ, Huffman MD, Isasi CR, Jimenez MC, Judd SE, Kissela BM, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Magid DJ, McGuire DK, Mohler ER, 3rd, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Rosamond W, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Woo D, Yeh RW and Turner MB. Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. *Circulation*. 2016;133:e38-e360.
183. Ping P, Vondriska TM, Creighton CJ, Gandhi TK, Yang Z, Menon R, Kwon MS, Cho SY, Drwal G, Kellmann M, Peri S, Suresh S, Gronborg M, Molina H, Chaerkady R, Rekha B, Shet AS, Gerszten RE, Wu H, Raftery M, Wasinger V, Schulz-Knappe P, Hanash SM, Paik YK, Hancock WS, States DJ, Omenn GS and Pandey A. A functional annotation of subproteomes in human plasma. *Proteomics*. 2005;5:3506-19.
184. Ko D, Benson MD, Ngo D, Yang Q, Larson MG, Wang TJ, Trinquart L, McManus DD, Lubitz SA, Ellinor PT, Vasan RS, Gerszten RE, Benjamin EJ and Lin H. Proteomics Profiling and Risk of New-Onset Atrial Fibrillation: Framingham Heart Study. *Journal of the American Heart Association*. 2019;8:e010976.
185. Uemura T, Kaikita K, Yamabe H, Soejima K, Matsukawa M, Fuchigami S, Tanaka Y, Morihisa K, Enomoto K, Sumida H, Sugiyama S and Ogawa H. Changes in plasma von Willebrand factor and ADAMTS13 levels associated with left atrial remodeling in atrial fibrillation. *Thrombosis research*. 2009;124:28-32.
186. Freynhofer MK, Bruno V, Jarai R, Gruber S, Hochtl T, Brozovic I, Farhan S, Wojta J and Huber K. Levels of von Willebrand factor and ADAMTS13 determine clinical outcome after cardioversion for atrial fibrillation. *Thrombosis and haemostasis*. 2011;105:435-43.
187. Sattlecker M, Kiddle SJ, Newhouse S, Proitsi P, Nelson S, Williams S, Johnston C, Killick R, Simmons A, Westman E, Hodges A, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S and Dobson RJ. Alzheimer's disease biomarker discovery using SOMAscan multiplexed protein technology. *Alzheimer's & dementia : the journal of the Alzheimer's Association*. 2014;10:724-34.
188. Hathout Y, Brody E, Clemens PR, Cripe L, DeLisle RK, Furlong P, Gordish-Dressman H, Hache L, Henricson E, Hoffman EP, Kobayashi YM, Lorts A, Mah JK, McDonald C, Mehler B, Nelson S, Nikrad M, Singer B, Steele F, Sterling D, Sweeney HL, Williams S and Gold L. Large-scale serum protein biomarker discovery in Duchenne muscular dystrophy. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112:7153-8.
189. Marion T, Elbahesh H, Thomas PG, DeVincenzo JP, Webby R and Schughart K. Respiratory Mucosal Proteome Quantification in Human Influenza Infections. *PloS one*. 2016;11:e0153674.
190. De Groote MA, Nahid P, Jarlsberg L, Johnson JL, Weiner M, Muzanyi G, Janjic N, Sterling DG and Ochsner UA. Elucidating novel serum biomarkers associated with pulmonary tuberculosis treatment. *PloS one*. 2013;8:e61002.
191. Smith JG and Gerszten RE. Emerging Affinity-Based Proteomic Technologies for Large-Scale Plasma Profiling in Cardiovascular Disease. *Circulation*. 2017;135:1651-1664.
192. Ngo D, Sinha S, Shen D, Kuhn EW, Keyes MJ, Shi X, Benson MD, O'Sullivan JF, Keshishian H, Farrell LA, Fifer MA, Vasan RS, Sabatine MS, Larson MG, Carr SA, Wang TJ and Gerszten RE. Aptamer-Based Proteomic Profiling Reveals Novel Candidate Biomarkers and Pathways in Cardiovascular Disease. *Circulation*. 2016;134:270-85.
193. Menni C, Kiddle SJ, Mangino M, Vinuela A, Psatha M, Steves C, Sattlecker M, Buil A, Newhouse S, Nelson S, Williams S, Voyle N, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Spector TD, Dobson R and Valdes AM. Circulating Proteomic Signatures of Chronological Age. *The journals of gerontology Series A, Biological sciences and medical sciences*. 2015;70:809-16.

194. Scriba TJ, Penn-Nicholson A, Shankar S, Hraha T, Thompson EG, Sterling D, Nemes E, Darboe F, Suliman S, Amon LM, Mahomed H, Erasmus M, Whatney W, Johnson JL, Boom WH, Hatherill M, Valvo J, De Groote MA, Ochsner UA, Aderem A, Hanekom WA and Zak DE. Sequential inflammatory processes define human progression from M. tuberculosis infection to tuberculosis disease. *PLoS pathogens*. 2017;13:e1006687.
195. Alonso A and Bengtson LG. A rising tide: the global epidemic of atrial fibrillation. *Circulation*. 2014;129:829-30.
196. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ and Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128-38.
197. Schnabel RB, Aspelund T, Li G, Sullivan LM, Suchy-Dicey A, Harris TB, Pencina MJ, D'Agostino RB, Sr., Levy D, Kannel WB, Wang TJ, Kronmal RA, Wolf PA, Burke GL, Launer LJ, Vasan RS, Psaty BM, Benjamin EJ, Gudnason V and Heckbert SR. Validation of an atrial fibrillation risk algorithm in whites and African Americans. *Archives of internal medicine*. 2010;170:1909-17.
198. Alonso A, Roetker NS, Soliman EZ, Chen LY, Greenland P and Heckbert SR. Prediction of Atrial Fibrillation in a Racially Diverse Cohort: The Multi-Ethnic Study of Atherosclerosis (MESA). *Journal of the American Heart Association*. 2016;5.
199. Everett BM, Cook NR, Conen D, Chasman DI, Ridker PM and Albert CM. Novel genetic markers improve measures of atrial fibrillation risk prediction. *European heart journal*. 2013;34:2243-51.
200. Pfister R, Bragelmann J, Michels G, Wareham NJ, Luben R and Khaw KT. Performance of the CHARGE-AF risk model for incident atrial fibrillation in the EPIC Norfolk cohort. *European journal of preventive cardiology*. 2015;22:932-9.
201. Fauchier L, Clementy N, Pelade C, Collignon C, Nicolle E and Lip GY. Patients With Ischemic Stroke and Incident Atrial Fibrillation: A Nationwide Cohort Study. *Stroke; a journal of cerebral circulation*. 2015;46:2432-7.
202. Zuo ML, Liu S, Chan KH, Lau KK, Chong BH, Lam KF, Chan YH, Lau YF, Lip GY, Lau CP, Tse HF and Siu CW. The CHADS2 and CHA 2DS 2-VASc scores predict new occurrence of atrial fibrillation and ischemic stroke. *Journal of interventional cardiac electrophysiology : an international journal of arrhythmias and pacing*. 2013;37:47-54.
203. Cameli M, Lisi M, Focardi M, Reccia R, Natali BM, Sparla S and Mondillo S. Left atrial deformation analysis by speckle tracking echocardiography for prediction of cardiovascular outcomes. *The American journal of cardiology*. 2012;110:264-9.
204. Yuasa T and Imoto Y. Usefulness of Tissue Doppler Imaging-Derived Atrial Conduction Time for Prediction of Atrial Fibrillation. *Circulation journal : official journal of the Japanese Circulation Society*. 2016;80:58-9.
205. Nakanishi K, Fukuda S, Tanaka A, Otsuka K, Sakamoto M, Taguchi H, Yoshikawa J, Shimada K and Yoshiyama M. Peri-atrial epicardial adipose tissue is associated with new-onset nonvalvular atrial fibrillation. *Circulation journal : official journal of the Japanese Circulation Society*. 2012;76:2748-54.
206. Tada H, Shiffman D, Smith JG, Sjogren M, Lubitz SA, Ellinor PT, Louie JZ, Catanese JJ, Engstrom G, Devlin JJ, Kathiresan S and Melander O. Twelve-single nucleotide polymorphism genetic risk score identifies individuals at increased risk for future atrial fibrillation and stroke. *Stroke; a journal of cerebral circulation*. 2014;45:2856-2862.
207. Ben Freedman S and Lowres N. Asymptomatic Atrial Fibrillation: The Case for Screening to Prevent Stroke. *Jama*. 2015;314:1911-2.
208. Lin HJ, Wolf PA, Benjamin EJ, Belanger AJ and D'Agostino RB. Newly diagnosed atrial fibrillation and acute stroke. The Framingham Study. *Stroke; a journal of cerebral circulation*. 1995;26:1527-30.

209. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *American journal of epidemiology*. 1989;129:687-702.
210. White AD, Folsom AR, Chambless LE, Sharret AR, Yang K, Conwill D, Higgins M, Williams OD and Tyroler HA. Community surveillance of coronary heart disease in the Atherosclerosis Risk in Communities (ARIC) Study: methods and initial two years' experience. *Journal of clinical epidemiology*. 1996;49:223-33.
211. Soliman EZ, Prineas RJ, Case LD, Zhang ZM and Goff DC, Jr. Ethnic distribution of ECG predictors of atrial fibrillation and its impact on understanding the ethnic distribution of ischemic stroke in the Atherosclerosis Risk in Communities (ARIC) study. *Stroke; a journal of cerebral circulation*. 2009;40:1204-11.
212. Alonso A, MacLehose RF, Chen LY, Bengtson LG, Chamberlain AM, Norby FL and Lutsey PL. Prospective study of oral anticoagulants and risk of liver injury in patients with atrial fibrillation. *Heart*. 2017.
213. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R and Dublin S. A systematic review of validated methods for identifying atrial fibrillation using administrative data. *Pharmacoepidemiology and drug safety*. 2012;21 Suppl 1:141-7.
214. Molvin J, Jujic A, Melander O, Pareek M, Rastam L, Lindblad U, Daka B, Leosdottir M, Nilsson P, Olsen M and Magnusson M. Exploration of pathophysiological pathways for incident atrial fibrillation using a multiplex proteomic chip. *Open Heart*. 2020;7:e001190.
215. Staerk L, Preis SR, Lin H, Lubitz SA, Ellinor PT, Levy D, Benjamin EJ and Trinquart L. Protein Biomarkers and Risk of Atrial Fibrillation: The FHS. *Circulation Arrhythmia and electrophysiology*. 2020;13:e007607.
216. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S, Katilius E, Keeney TR, Kim N, Koch TH, Kraemer S, Kroiss L, Le N, Levine D, Lindsey W, Lollo B, Mayfield W, Mehan M, Mehler R, Nelson SK, Nelson M, Nieuwlandt D, Nikrad M, Ochsner U, Ostroff RM, Otis M, Parker T, Pietrasiewicz S, Resnicow DI, Rohloff J, Sanders G, Sattin S, Schneider D, Singer B, Stanton M, Sterkel A, Stewart A, Stratford S, Vaught JD, Vrkljan M, Walker JJ, Watrobka M, Waugh S, Weiss A, Wilcox SK, Wolfson A, Wolk SK, Zhang C and Zichi D. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PloS one*. 2010;5:e15004.
217. Gold L, Walker JJ, Wilcox SK and Williams S. Advances in human proteomics at high scale with the SOMAscan proteomics platform. *New biotechnology*. 2012;29:543-9.
218. Kim CH, Tworoger SS, Stampfer MJ, Dillon ST, Gu X, Sawyer SJ, Chan AT, Libermann TA and Eliassen AH. Stability and reproducibility of proteomic profiles measured with an aptamer-based platform. *Sci Rep*. 2018;8:8382.
219. Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, Huang J, Stuccio S, Zingone A, Ryan BM, Tsang JS and Biancotto A. Assessment of Variability in the SOMAscan Assay. *Sci Rep*. 2017;7:14248.
220. Kramer A, Green J, Pollard J, Jr. and Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*. 2014;30:523-30.
221. Patel JV, Lim HS, Varughese GI, Hughes EA and Lip GY. Angiopoietin-2 levels as a biomarker of cardiovascular risk in patients with hypertension. *Ann Med*. 2008;40:215-22.
222. Chong AY, Caine GJ, Freestone B, Blann AD and Lip GY. Plasma angiopoietin-1, angiopoietin-2, and angiopoietin receptor tie-2 levels in congestive heart failure. *Journal of the American College of Cardiology*. 2004;43:423-8.
223. Morrell NW, Bloch DB, ten Dijke P, Goumans MJ, Hata A, Smith J, Yu PB and Bloch KD. Targeting BMP signalling in cardiovascular disease and anaemia. *Nat Rev Cardiol*. 2016;13:106-20.

224. Sedaghat S, de Vries PS, Boender J, Sonneveld MA, Hoorn EJ, Hofman A, de Maat MP, Franco OH, Ikram MA, Leebeek FW and Dehghan A. von Willebrand Factor, ADAMTS13 Activity, and Decline in Kidney Function: A Population-Based Cohort Study. *Am J Kidney Dis*. 2016;68:726-732.
225. Sonneveld MA, Franco OH, Ikram MA, Hofman A, Kavousi M, de Maat MP and Leebeek FW. Von Willebrand Factor, ADAMTS13, and the Risk of Mortality: The Rotterdam Study. *Arterioscler Thromb Vasc Biol*. 2016;36:2446-2451.
226. Bourron O, Le Bouc Y, Berard L, Kotti S, Brunel N, Ritz B, Leclercq F, Tabone X, Drouet E, Mulak G, Danchin N and Simon T. Impact of age-adjusted insulin-like growth factor 1 on major cardiovascular events after acute myocardial infarction: results from the fast-MI registry. *J Clin Endocrinol Metab*. 2015;100:1879-86.
227. Busch M, Kruger A, Gross S, Ittermann T, Friedrich N, Nauck M, Dorr M and Felix SB. Relation of IGF-1 and IGFBP-3 with prevalent and incident atrial fibrillation in a population-based study. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2019;16:1314-1319.
228. Huang L, Li L, Yang T, Li W, Song L, Meng X, Gu Q, Xiong C and He J. Transgelin as a potential target in the reversibility of pulmonary arterial hypertension secondary to congenital heart disease. *J Cell Mol Med*. 2018;22:6249-6261.
229. Nakada TA, Russell JA, Boyd JH, Thair SA and Walley KR. Identification of a nonsynonymous polymorphism in the SVEP1 gene associated with altered clinical outcomes in septic shock. *Crit Care Med*. 2015;43:101-8.
230. Sato-Nishiuchi R, Nakano I, Ozawa A, Sato Y, Takeichi M, Kiyozumi D, Yamazaki K, Yasunaga T, Futaki S and Sekiguchi K. Polydom/SVEP1 is a ligand for integrin alpha9beta1. *J Biol Chem*. 2012;287:25615-30.
231. Myocardial Infarction G, Investigators CAEC, Stitziel NO, Stirrups KE, Masca NG, Erdmann J, Ferrario PG, Konig IR, Weeke PE, Webb TR, Auer PL, Schick UM, Lu Y, Zhang H, Dube MP, Goel A, Farrall M, Peloso GM, Won HH, Do R, van Iperen E, Kanoni S, Kruppa J, Mahajan A, Scott RA, Willenberg C, Braund PS, van Capelleveen JC, Doney AS, Donnelly LA, Asselta R, Merlini PA, Duga S, Marziliano N, Denny JC, Shaffer CM, El-Mokhtari NE, Franke A, Gottesman O, Heilmann S, Hengstenberg C, Hoffman P, Holmen OL, Hveem K, Jansson JH, Jockel KH, Kessler T, Kriebel J, Laugwitz KL, Marouli E, Martinelli N, McCarthy MI, Van Zuydam NR, Meisinger C, Esko T, Mihailov E, Escher SA, Alver M, Moebus S, Morris AD, Muller-Nurasyid M, Nikpay M, Olivieri O, Lemieux Perreault LP, AlQarawi A, Robertson NR, Akinsanya KO, Reilly DF, Vogt TF, Yin W, Asselbergs FW, Kooperberg C, Jackson RD, Stahl E, Strauch K, Varga TV, Waldenberger M, Zeng L, Kraja AT, Liu C, Ehret GB, Newton-Cheh C, Chasman DI, Chowdhury R, Ferrario M, Ford I, Jukema JW, Kee F, Kuulasmaa K, Nordestgaard BG, Perola M, Saleheen D, Sattar N, Surendran P, Tregouet D, Young R, Howson JM, Butterworth AS, Danesh J, Ardissino D, Bottinger EP, Erbel R, Franks PW, Girelli D, Hall AS, Hovingh GK, Kastrati A, Lieb W, Meitinger T, Kraus WE, Shah SH, McPherson R, Orho-Melander M, Melander O, Metspalu A, Palmer CN, Peters A, Rader D, Reilly MP, Loos RJ, Reiner AP, Roden DM, Tardif JC, Thompson JR, Wareham NJ, Watkins H, Willer CJ, Kathiresan S, Deloukas P, Samani NJ and Schunkert H. Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. *The New England journal of medicine*. 2016;374:1134-44.
232. Mabuchi M, Kataoka H, Miura Y, Kim TS, Kawaguchi M, Ebi M, Tanaka M, Mori Y, Kubota E, Mizushima T, Shimura T, Mizoshita T, Tanida S, Kamiya T, Asai K and Joh T. Tumor suppressor, AT motif binding factor 1 (ATBF1), translocates to the nucleus with runt domain transcription factor 3 (RUNX3) in response to TGF-beta signal transduction. *Biochem Biophys Res Commun*. 2010;398:321-5.

233. Li J, Solus J, Chen Q, Rho YH, Milne G, Stein CM and Darbar D. Role of inflammation and oxidative stress in atrial fibrillation. *Heart rhythm : the official journal of the Heart Rhythm Society*. 2010;7:438-44.
234. Huxley RR, Lopez FL, MacLehose RF, Eckfeldt JH, Couper D, Leiendecker-Foster C, Hoogeveen RC, Chen LY, Soliman EZ, Agarwal SK and Alonso A. Novel association between plasma matrix metalloproteinase-9 and risk of incident atrial fibrillation in a case-cohort study: the Atherosclerosis Risk in Communities study. *PloS one*. 2013;8:e59052.
235. Nakano Y, Niida S, Dote K, Takenaka S, Hirao H, Miura F, Ishida M, Shingu T, Sueda T, Yoshizumi M and Chayama K. Matrix metalloproteinase-9 contributes to human atrial remodeling during atrial fibrillation. *Journal of the American College of Cardiology*. 2004;43:818-25.
236. Liu Y, Xu B, Wu N, Xiang Y, Wu L, Zhang M, Wang J, Chen X, Li Y and Zhong L. Association of MMPs and TIMPs With the Occurrence of Atrial Fibrillation: A Systematic Review and Meta-analysis. *Can J Cardiol*. 2016;32:803-13.
237. Wakula P, Neumann B, Kienemund J, Thon-Gutschi E, Stojakovic T, Manninger M, Scherr D, Scharnagl H, Kapl M, Pieske B and Heinzel FR. CHA2DS2-VASc score and blood biomarkers to identify patients with atrial high-rate episodes and paroxysmal atrial fibrillation. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*. 2017;19:544-551.
238. Wang W, Zhang HT and Yang XL. Effect of matrix metalloproteinase and their inhibitors on atrial myocardial structural remodeling. *J Cardiovasc Med (Hagerstown)*. 2013;14:265-9.
239. Mukherjee R, Herron AR, Lowry AS, Stroud RE, Stroud MR, Wharton JM, Ikonomidis JS, Crumbley AJ, 3rd, Spinale FG and Gold MR. Selective induction of matrix metalloproteinases and tissue inhibitor of metalloproteinases in atrial and ventricular myocardium in patients with atrial fibrillation. *The American journal of cardiology*. 2006;97:532-7.
240. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J, Lin H, Arking DE, Smith AV, Albert CM, Chaffin M, Tucker NR, Li M, Klarin D, Bihlmeyer NA, Low SK, Weeke PE, Muller-Nurasyid M, Smith JG, Brody JA, Niemeijer MN, Dorr M, Trompet S, Huffman J, Gustafsson S, Schurmann C, Kleber ME, Lyytikainen LP, Seppala I, Malik R, Horimoto A, Perez M, Sinisalo J, Aeschbacher S, Theriault S, Yao J, Radmanesh F, Weiss S, Teumer A, Choi SH, Weng LC, Clauss S, Deo R, Rader DJ, Shah SH, Sun A, Hopewell JC, Debette S, Chauhan G, Yang Q, Worrall BB, Pare G, Kamatani Y, Hagemeijer YP, Verweij N, Siland JE, Kubo M, Smith JD, Van Wagoner DR, Bis JC, Perz S, Psaty BM, Ridker PM, Magnani JW, Harris TB, Launer LJ, Shoemaker MB, Padmanabhan S, Haessler J, Bartz TM, Waldenberger M, Lichtner P, Arendt M, Krieger JE, Kahonen M, Risch L, Mansur AJ, Peters A, Smith BH, Lind L, Scott SA, Lu Y, Bottinger EB, Hernesniemi J, Lindgren CM, Wong JA, Huang J, Eskola M, Morris AP, Ford I, Reiner AP, Delgado G, Chen LY, Chen YI, Sandhu RK, Li M, Boerwinkle E, Eisele L, Lannfelt L, Rost N, Anderson CD, Taylor KD, Campbell A, Magnusson PK, Porteous D, Hocking LJ, Vlachopoulou E, Pedersen NL, Nikus K, Orho-Melander M, Hamsten A, Heeringa J, Denny JC, Kriebel J, Darbar D, Newton-Cheh C, Shaffer C, Macfarlane PW, Heilmann-Heimbach S, Almgren P, Huang PL, Sotoodehnia N, Soliman EZ, Uitterlinden AG, Hofman A, Franco OH, Volker U, Jockel KH, Sinner MF, Lin HJ, Guo X, ISGC MCot, Neurology Working Group of the CC, Dichgans M, Ingelsson E, Kooperberg C, Melander O, Loos RJF, Laurikka J, Conen D, Rosand J, van der Harst P, Lokki ML, Kathiresan S, Pereira A, Jukema JW, Hayward C, Rotter JI, Marz W, Lehtimaki T, Stricker BH, Chung MK, Felix SB, Gudnason V, Alonso A, Roden DM, Kaab S, Chasman DI, Heckbert SR, Benjamin EJ, Tanaka T, Lunetta KL, Lubitz SA, Ellinor PT

and Consortium AF. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nature genetics*. 2017;49:946-952.
241. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, Almgren P, Alonso A, Anderson CD, Aragam KG, Arking DE, Barnard J, Bartz TM, Benjamin EJ, Bihlmeyer NA, Bis JC, Bloom HL, Boerwinkle E, Bottinger EB, Brody JA, Calkins H, Campbell A, Cappola TP, Carlquist J, Chasman DI, Chen LY, Chen YI, Choi EK, Choi SH, Christophersen IE, Chung MK, Cole JW, Conen D, Cook J, Crijns HJ, Cutler MJ, Damrauer SM, Daniels BR, Darbar D, Delgado G, Denny JC, Dichgans M, Dorr M, Dudink EA, Dudley SC, Esa N, Esko T, Eskola M, Fatkin D, Felix SB, Ford I, Franco OH, Geelhoed B, Grewal RP, Gudnason V, Guo X, Gupta N, Gustafsson S, Gutmann R, Hamsten A, Harris TB, Hayward C, Heckbert SR, Hernesniemi J, Hocking LJ, Hofman A, Horimoto A, Huang J, Huang PL, Huffman J, Ingelsson E, Ipek EG, Ito K, Jimenez-Conde J, Johnson R, Jukema JW, Kaab S, Kahonen M, Kamatani Y, Kane JP, Kastrati A, Kathiresan S, Katschnig-Winter P, Kavousi M, Kessler T, Kietselaer BL, Kirchhof P, Kleber ME, Knight S, Krieger JE, Kubo M, Launer LJ, Laurikka J, Lehtimaki T, Leineweber K, Lemaitre RN, Li M, Lim HE, Lin HJ, Lin H, Lind L, Lindgren CM, Lokki ML, London B, Loos RJF, Low SK, Lu Y, Lyytikainen LP, Macfarlane PW, Magnusson PK, Mahajan A, Malik R, Mansur AJ, Marcus GM, Margolin L, Margulies KB, Marz W, McManus DD, Melander O, Mohanty S, Montgomery JA, Morley MP, Morris AP, Muller-Nurasyid M, Natale A, Nazarian S, Neumann B, Newton-Cheh C, Niemeijer MN, Nikus K, Nilsson P, Noordam R, Oellers H, Olesen MS, Orho-Melander M, Padmanabhan S, Pak HN, Pare G, Pedersen NL, Pera J, Pereira A, Porteous D, Psaty BM, Pulit SL, Pullinger CR, Rader DJ, Refsgaard L, Ribases M, Ridker PM, Rienstra M, Risch L, Roden DM, Rosand J, Rosenberg MA, Rost N, Rotter JI, Saba S, Sandhu RK, Schnabel RB, Schramm K, Schunkert H, Schurman C, Scott SA, Seppala I, Shaffer C, Shah S, Shalaby AA, Shim J, Shoemaker MB, Siland JE, Sinisalo J, Sinner MF, Slowik A, Smith AV, Smith BH, Smith JG, Smith JD, Smith NL, Soliman EZ, Sotoodehnia N, Stricker BH, Sun A, Sun H, Svendsen JH, Tanaka T, Tanriverdi K, Taylor KD, Teder-Laving M, Teumer A, Theriault S, Trompet S, Tucker NR, Tveit A, Uitterlinden AG, Van Der Harst P, Van Gelder IC, Van Wagoner DR, Verweij N, Vlachopoulou E, Volker U, Wang B, Weeke PE, Weijs B, Weiss R, Weiss S, Wells QS, Wiggins KL, Wong JA, Woo D, Worrall BB, Yang PS, Yao J, Yoneda ZT, Zeller T, Zeng L, Lubitz SA, Lunetta KL and Ellinor PT. Multi-ethnic genome-wide association study for atrial fibrillation. *Nature genetics*. 2018;50:1225-1233.
242. Zhou M, Liao Y and Tu X. The role of transcription factors in atrial fibrillation. *J Thorac Dis*. 2015;7:152-8.
243. Tin A, Yu B, Ma J, Masushita K, Daya N, Hoogeveen RC, Ballantyne CM, Couper D, Rebholz CM, Grams ME, Alonso A, Mosley T, Heiss G, Ganz P, Selvin E, Boerwinkle E and Coresh J. Reproducibility and Variability of Protein Analytes Measured Using a Multiplexed Modified Aptamer Assay. *J Appl Lab Med*. 2019;4:30-39.
244. Han Z, Xiao Z, Kalantar-Zadeh K, Moradi H, Shafi T, Waikar SS, Quarles LD, Yu Z, Tin A, Coresh J and Kovesdy CP. Validation of a Novel Modified Aptamer-Based Array Proteomic Platform in Patients with End-Stage Renal Disease. *Diagnostics (Basel)*. 2018;8.
245. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, Sterling DG and Williams SA. Development and Validation of a Protein-Based Risk Score for Cardiovascular Outcomes Among Patients With Stable Coronary Heart Disease. *Jama*. 2016;315:2532-41.
246. Williams SA, Kivimaki M, Langenberg C, Hingorani AD, Casas JP, Bouchard C, Jonasson C, Sarzynski MA, Shipley MJ, Alexander L, Ash J, Bauer T, Chadwick J, Datta G, DeLisle RK, Hagar Y, Hinterberg M, Ostroff R, Weiss S, Ganz P and Wareham NJ. Plasma protein patterns as comprehensive indicators of health. *Nat Med*. 2019;25:1851-1857.
247. SomaLogic. Somascan® data standardization. *https://somalogiccom/wp-content/uploads/2017/06/SSM-071-Rev-0-Technical-Note-SOMAscan-Data-Standardizationpdf*.

248. Eriksson H, Caidahl K, Larsson B, Ohlson LO, Welin L, Wilhelmsen L and Svardsudd K. Cardiac and pulmonary causes of dyspnoea--validation of a scoring test for clinical-epidemiological use: the Study of Men Born in 1913. *European heart journal*. 1987;8:1007-14.
249. Loehr LR, Rosamond WD, Chang PP, Folsom AR and Chambless LE. Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study). *The American journal of cardiology*. 2008;101:1016-22.
250. Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, Kusek JW, Manzi J, Van Lente F, Zhang YL, Coresh J, Levey AS and Investigators C-E. Estimating glomerular filtration rate from serum creatinine and cystatin C. *The New England journal of medicine*. 2012;367:20-9.
251. Maheshwari A, Norby FL, Soliman EZ, Koene R, Rooney M, O'Neal WT, Alonso A and Chen LY. Refining Prediction of Atrial Fibrillation Risk in the General Population With Analysis of P-Wave Axis (from the Atherosclerosis Risk in Communities Study). *The American journal of cardiology*. 2017;120:1980-1984.
252. Olsen FJ, Mogelvang R, Jensen GB, Jensen JS and Biering-Sorensen T. Relationship Between Left Atrial Functional Measures and Incident Atrial Fibrillation in the General Population: The Copenhagen City Heart Study. *JACC Cardiovascular imaging*. 2018.
253. Pencina MJ and D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in medicine*. 2004;23:2109-23.
254. D'Agostino R. B. NB. Evaluation of the performance of survival analysis models: discrimination and calibration measures. *Amsterdam: Elsevier*. 2004;23:1-25.
255. Collins GS, Reitsma JB, Altman DG and Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Annals of internal medicine*. 2015;162:55-63.
256. Steyerberg EW, Harrell FE, Jr., Borsboom GJ, Eijkemans MJ, Vergouwe Y and Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of clinical epidemiology*. 2001;54:774-81.
257. Austin PCT, J.V. Bootstrap Methods for Developing Predictive Models. *The American Statistician*. 2004;58:131-137.
258. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr. and Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine*. 2008;27:157-72; discussion 207-12.
259. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC and Greenland P. Interpreting incremental value of markers added to risk prediction models. *American journal of epidemiology*. 2012;176:473-81.
260. Harrell FE, Jr., Lee KL and Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in medicine*. 1996;15:361-87.
261. Thompson RC and Ohlsson K. Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor, a potent inhibitor of leukocyte elastase. *Proceedings of the National Academy of Sciences of the United States of America*. 1986;83:6692-6.
262. Hannila SS. Secretory Leukocyte Protease Inhibitor (SLPI): Emerging Roles in CNS Trauma and Repair. *Neuroscientist*. 2015;21:630-6.
263. Costabel JP, Galve E, Terricabras M, Ametrano C, Ronderos R, Baranchuk A, Evangelista A and Avegliano G. E/e' ratio and left atrial area are predictors of atrial fibrillation in patients with hypertrophic cardiomyopathy. *Echocardiography*. 2018;35:935-940.

264. Shah AM, Claggett B, Kitzman D, Biering-Sorensen T, Jensen JS, Cheng S, Matsushita K, Konety S, Folsom AR, Mosley TH, Wright JD, Heiss G and Solomon SD. Contemporary Assessment of Left Ventricular Diastolic Function in Older Adults: The Atherosclerosis Risk in Communities Study. *Circulation*. 2017;135:426-439.
265. Shah AM, Cheng S, Skali H, Wu J, Mangion JR, Kitzman D, Matsushita K, Konety S, Butler KR, Fox ER, Cook N, Ni H, Coresh J, Mosley TH, Heiss G, Folsom AR and Solomon SD. Rationale and design of a multicenter echocardiographic study to assess the relationship between cardiac structure and function and heart failure risk in a biracial cohort of community-dwelling elderly persons: the Atherosclerosis Risk in Communities study. *Circulation Cardiovascular imaging*. 2014;7:173-81.
266. Garg RK, Glazer NL, Wiggins KL, Newton KM, Thacker EL, Smith NL, Siscovick DS, Psaty BM and Heckbert SR. Ascertainment of warfarin and aspirin use by medical record review compared with automated pharmacy data. *Pharmacoepidemiology and drug safety*. 2011;20:313-6.
267. Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE and Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. *Pharmacoepidemiology and drug safety*. 2011;20:560-6.
268. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE and Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical care*. 2005;43:1130-9.
269. Lip GY, Nieuwlaat R, Pisters R, Lane DA and Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest*. 2010;137:263-72.
270. Zou G. A modified poisson regression approach to prospective studies with binary data. *American journal of epidemiology*. 2004;159:702-6.
271. Waddy SP, Solomon AJ, Becerra AZ, Ward JB, Chan KE, Fwu CW, Norton JM, Eggers PW, Abbott KC and Kimmel PL. Racial/Ethnic Disparities in Atrial Fibrillation Treatment and Outcomes among Dialysis Patients in the United States. *J Am Soc Nephrol*. 2020;31:637-649.
272. Essien UR, Magnani JW, Chen N, Gellad WF, Fine MJ and Hernandez I. Race/Ethnicity and Sex-Related Differences in Direct Oral Anticoagulant Initiation in Newly Diagnosed Atrial Fibrillation: A Retrospective Study of Medicare Data. *J Natl Med Assoc*. 2020;112:103-108.
273. Shields GE, Bates AE and Chapman AM. Implementing Guidelines: The Cost and Clinical Impact of Anticoagulants in the UK Atrial Fibrillation Population. *Appl Health Econ Health Policy*. 2015;13:543-51.
274. Vestergaard AS and Ehlers LH. A Health Economic Evaluation of Stroke Prevention in Atrial Fibrillation: Guideline Adherence Versus the Observed Treatment Strategy Prior to 2012 in Denmark. *Pharmacoeconomics*. 2015;33:967-79.
275. Andrade SE, Harrold LR, Tjia J, Cutrona SL, Saczynski JS, Dodd KS, Goldberg RJ and Gurwitz JH. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiology and drug safety*. 2012;21 Suppl 1:100-28.
276. Saczynski JS, Andrade SE, Harrold LR, Tjia J, Cutrona SL, Dodd KS, Goldberg RJ and Gurwitz JH. A systematic review of validated methods for identifying heart failure using administrative data. *Pharmacoepidemiology and drug safety*. 2012;21 Suppl 1:129-40.
277. Avgil Tsadok M, Jackevicius CA, Essebag V, Eisenberg MJ, Rahme E and Pilote L. Warfarin Treatment and Outcomes of Patients With Atrial Fibrillation in Rural and Urban Settings. *J Rural Health*. 2015;31:310-5.