Supplementary Online Content 1

Yao X, Gersh BJ, Holmes DR, et al. Association of surgical left atrial appendage occlusion with subsequent stroke and mortality among patients undergoing cardiac surgery. *JAMA*. doi:10.1001/jama.2018.6024

Statistical Analysis Plan

STATISTICAL ANALYSIS PLAN

Study Title: Surgical Occlusion of Left Atrial Appendage during Cardiac Surgery: Impact on Stroke, Mortality and Subsequent Atrial Fibrillation

Statistical Analysis Plan Authors: Xiaoxi Yao and Peter A. Noseworthy

Statistical Analysis Plan Date: 10/13/2017

The purpose of this analysis plan is to provide guide to our analyst when conducting the study. Most of the content will be included in the manuscript in order to guide researchers who want to replicate our findings or conduct similar studies. We also provided justifications for our methods and decisions so other researchers can make a choice or adjust their methods accordingly.

TABLE OF CONTENTS

ABBREVIATIONS

AF Atrial Fibrillation

CABG Coronary Artery Bypass

CI Confidence interval

eGFR Estimated Glomerular Filtration Rate, a measure of renal function, calculated using age, sex, race and serum creatinine based on the CKD-EPI equation

HR Hazard Ratio, a measure of relative risk used to compare different treatments; calculated from time-to-event Cox proportional hazards regression

IPTW Inverse Probability Treatment Weighting, a propensity score method used to balance patients' baseline characteristics between treatment groups (ATE weight used in this study)

IQR Interquartile Range

LAAO Surgical Occlusion of Left Atrial Appendage

NOAC Non-vitamin K antagonist oral anticoagulants, including four drugs: apixaban, dabigatran, edoxaban and rivaroxaban

OAC Oral anticoagulant, including five drugs, four NOACs and warfarin

Key Definition

Index Date (variable name index_date) the date patients received cardiac surgery

Baseline Period

Any time before and including the index date used to establish a patient's medical history. Detailed description and justification can be found on page 10.

Study Period

The study period will be January 1st, 2009 and March 31st, 2017. The last day of follow up is March 31st, 2017. Because we will require patients to have at least 1 day of follow up, the last day of index procedure will be March 30th, 2017.

1. BACKGROUND AND OBJECTIVES

Cardiac surgeries are among the most commonly performed procedures, with more than 300,000 coronary artery bypass (CABG) and valve operations performed annually in the United States.[1] Many patients undergoing cardiac surgery have a history of atrial fibrillation (AF), which is associated with a five-fold risk of stroke.[2,3] Because thrombi in the left atrial appendage (LAA) are believed to account for the majority of cardioembolic strokes in AF,[4] surgical occlusion of the LAA (LAAO) is sometimes performed during the surgery to reduce long-term risk of stroke.

There is a paucity of data on the effectiveness of LAAO to guide evidence-based decision making. Although LAA is the dominant source of thrombi only in the setting of AF,[4] in a recent observational study, more than half of the patients undergoing LAAO did not have prior AF.[5] This is likely because surgeons perceived some patients were at a high risk of developing AF[6,7] and chose to close the LAA preemptively. However, little is known whether this approach is justified. Therefore, this study aims to investigate whether LAAO during cardiac surgery is associated with reduced risks of stroke and mortality. We will specifically assess the outcomes stratified by whether patients had a history of AF at the time of surgery.

A secondary aim is to investigate whether LAAO promotes subsequent AF. A previous study found that LAAO may increase the risk of post-operative AF,[5] likely by promoting an atrial arrhythmogenic state due to increased left atrial filling pressures, inflammation, and sympatho-vagal imbalance.[8-10] As such, this study will examine the

association between LAAO and post-operative AF, as well as long-term AF-related health utilization.

2. STUDY DESIGN AND DATA SOURCE

We will conduct a retrospective cohort analysis using OptumLabs Data Warehouse, which contains over 160 million privately insured and Medicare Advantage enrollees of all ages and races from all 50 states.[11,12]

3. STUDY POPULATION

The study population will be adult patients (≥18 years) who underwent their first coronary artery bypass grafting (CABG) or valve surgery (open-heart valve replacement or repair) between January 1st, 2009 and March 31st, 2017. Patients will be required to have at least 6 months of continuous enrollment in health insurance plans prior to the surgery, defined as the baseline used to capture patients' medical history. This requirement is to allow us to have sufficient data for ascertaining patients' medical history. Please calculate the median and interquartile range [IQR] of the baseline period.

4. MEASUREMENTS

4.1 Baseline Characteristics

Baseline characteristics include socio-demographic characteristics, medical history, concurrent medication use, and biomarkers. *Socio-demographic characteristics*, such as age, will be determined at the time of index date. *Medical history* will be determined using patients' physician, facility and pharmacy claims before or on the index date. We will use all data available to us to establish patients' medical history, and the length of baseline period will be included in the propensity score model to avoid any potential bias. Typically, there is no substantial difference in the length of baseline period among different treatment groups, especially after propensity score matching. *Baseline medication* will be captured within 3 months of the index date. *Pre-operative medication*, including, beta blocker, amiodarone, statin and corticosteroid, will be captured within 1 week of the index date. These drugs could potentially prevent post-operative AF.[13-16] *Biomarkers*, such as serum creatinine, serum calcium, serum albumin, hemoglobin, low density lipoprotein cholesterol (LDL-C), and hemoglobin A1c (HbA1c), will be captured using the most recent measurements within 12 months of the index date. About one third of the patients have linked outpatient laboratory data. The availability of laboratory data depends on the contract between laboratory testing facilities and OLDW, rather than individual patient characteristics. Because biomarkers are missing in a proportion of patients, and some biomarkers were only tested in patients with certain conditions, these variables will not be included in the propensity score model, but we will test on the balance of both the values and proportion of missing values after balancing on all other patient characteristics.

Baseline AF will be defined as an AF diagnosis (International Classification of Diseases, 9th Revision, Clinical Modification [ICD-9] diagnosis 427.31 or International Classification of Diseases, 10th Revision [ICD-10] codes I48.0, I48.1, I48.2 and I48.91) on an inpatient or outpatient claim at baseline. These diagnosis codes performed relatively well in previous validation studies, with a median positive predictive value (PPV) of 89%.[17] We note that some previous studies required multiple AF diagnoses when creating cohorts to minimize the impact of rule-out diagnoses and improve the specificity.[18] By contrast, we will require only one AF diagnosis in order to increase the sensitivity of the cohort definition, as the diagnosis codes often have good specificity but sometimes low sensitivity.[17,19] In our previous studies, the specificity of requiring one AF diagnosis is very good, because nearly all of AF patients had more than one diagnosis, and patients on average had 20 AF diagnoses on different dates at baseline.

We will also need to code a few procedure-related characteristics, such as surgery types, types of valve surgery, whether the surgery is an on-pump surgery, and whether patients had pre-operative hemodynamic instability and pre-operative endocarditis, and the year of the surgery. Pre-operative hemodynamic instability includes cardiogenic shock, cardiac arrest and resuscitation within one week prior to the procedure. Pre-operative endocarditis will also be captured within one week prior to the procedure.

We will need to fill out Table 1 and eTables 1-3 shown below. eTable 1 will include baseline characteristics not presented in Table 1, since journals often request us to reduce the size of Table 1. eTable 2 will be patient characteristics after propensity score matching including all characteristics in Table 1 and eTable 1, with a third column

for standardized mean difference. Because the main analysis will be stratified by prior AF, it would be interesting to see the patient characteristics by prior AF in the propensity score cohort. This will be e-Table 3.

Table 1. Baseline Characteristics

	No LAAO (N=)	LAAO (N=)	Total (N=)
Age, yr			
Female			
White			
Surgery types			
CABG			
Mechanical valve replacement			
Bioprosthetic valve replacement			
Valve repair			
CABG+valve surgery			
Types of valve surgery			
Aortic			
Mitral			
Tricuspid or pulmonary			
Medical History			
Atrial fibrillation			
Other supraventricular arrhythmia			
Thromboembolism			
Heart failure			
Diabetes mellitus			
Stage 3-5 CKD			
Myocardial infarction			
Peripheral Artery Disease			
Major bleeding			
Intracranial bleeding			
Hypertension			
Hyperlipidemia			
Falls			
COPD			
Alcoholism			
Obesity			
Smoking			
Obstructive sleep apnea			
Non skin cancer			
CHA_2DS_2-VASc			
0, 1			
2, 3			
≥4			
HAS-BLED≥3			

eTable 1. Baseline Characteristics before Propensity Score Matching Not Presented in Table 1

	No LAAO (N=)	LAAO (N=)	Total (N=)
Race			
Asian			
Black			
Hispanic/Latino			
White			
Other/Unknown			
Geographic Region			
Midwest			
Northeast			
South			
West			
On-pump Surgery			
Pre-operative Hemodynamic Instability			
Pre-operative Endocarditis			
Medical History			
Ischemic stroke or systemic embolism			
TIA			
Ventricular arrhythmia			
Systolic heart failure			
Diabetes requiring insulin			
Dialysis			
Cardioversion			
Ablation			
Pacemaker/ICD			
PCI			
Anemia			
Liver disease			
Depression			
Dementia			
Hypothyroidism			
Thyrotoxicosis			
Ulcer in upper GI tract			
Baseline Medication			
Oral anticoagulant			
Antiplatelet			
Rate control drugs			
Antiarrhythmic drugs			
Other adrenergic blocking agents			
Other calcium channel blockers			
Renin angiotensin system antagonists			

Loop diuretics
Thiazides
Cholesterol lowering drugs
NSAIDs
Diabetes drugs
Antiulcer agents
Pre-operative Medication
Beta blocker
Amiodarone
Statin
Corticosteroid
Year of Index Procedure
2009
2010
2011
2012
2013
2014
2015
2016
2017
Biomarkers, Mean±SD
Serum creatinine, mg/dL
Serum calcium, mg/dL
Serum albumin, g/dL
Hemoglobin, g/dL
LDL-C, mg/dL
HbA1c, %

eTable 2. Baseline Characteristics after Propensity Score Matching

	No LAAO (N=)	LAAO (N=)	Standardized Difference
Age, yr			
18-64			
65-74			
≥75			
Female			
Race			
Asian			
Black			
Hispanic/Latino			
White			
Other/Unknown			
Geographic Region			
Midwest			
Northeast			
South			
West			
Surgery types			
CABG			
Mechanical valve replacement			
Bioprosthetic valve replacement			
Valve repair			
CABG+valve surgery			
Types of Valves			
Aortic			
Mitral			
Tricupid or pulmonary			
Medical History			
Atrial fibrillation			
Other supraventricular arrhythmia			
Thromboembolism			
Heart failure			
Diabetes mellitus			
Stage 3-5 CKD			
Myocardial infarction			
Peripheral artery disease			
Major bleeding			
Intracranial bleeding			
Hypertension			
Hyperlipidemia			
Falls			
Anemia			

COPD	
Alcoholism	
Obesity	
Smoking	
Obstructive sleep apnea	
Non skin cancer	
Ischemic stroke or systemic embolism	
TIA	
Ventricular arrhythmia	
Systolic heart failure	
Diabetes requiring insulin	
Dialysis	
Cardioversion	
Ablation	
Pacemaker/ICD	
PCI	
Liver disease	
Depression	
Dementia	
Hypothyroidism	
Thyrotoxicosis	
Ulcer in upper GI tract	
On-pump Surgery	
Pre-operative Hemodynamic Instability	
Pre-operative Endocarditis	
CHA_2DS_2-VASc	
0, 1	
2, 3	
≥4	
HAS-BLED≥3	
Baseline Medication	
Oral anticoagulant	
Antiplatelet	
Rate control drugs	
Antiarrhythmic drugs	
Other adrenergic blocking agents	
Other calcium channel blockers	
Renin angiotensin system antagonists	
Loop diuretics	
Thiazides	
Cholesterol lowering drugs	
NSAIDs	
Diabetes drugs	

Antiulcer agents
Periprocedural Medication
Beta blocker
Amiodarone
Statin
Corticosteroid
Year of Index Procedure
2009
2010
2011
2012
2013
2014
2015
2016
2017
Length of Baseline Period, yr
Biomarkers
Serum creatinine, mg/dL
Serum calcium, mg/dL
Serum albumin, g/dL
Hemoglobin, g/dL
LDL-C, mg/dL
HbA1c, %

4.2 Follow up and Outcomes

The primary outcomes will be: (1) ischemic stroke or systemic embolism (hereafter referred to as stroke) and (2) all-cause mortality. Stroke will be defined as a primary diagnosis on an emergency room visit or a primary or secondary diagnosis on an inpatient stay. Mortality will be identified based on the Social Security Death Master File and patient discharge status.

The secondary outcomes will be: (1) post-operative AF, defined as newly diagnosed AF within 30 day after the surgery; and (2) long-term AF-related health utilization, measured by the event rates of outpatient visits and hospitalizations with a diagnosis of AF. When assessing post-operative AF, we will restrict the analysis to patients without prior AF, because in patients with prior AF, a diagnosis code could be used for the previous history rather than a new episode of AF. The long-term AF-related health utilization will be assessed in both patients with and without prior AF from the date after surgery until the end of the study period (March 31st, 2017), the end of enrollment in health insurance plans, or death, whichever happened first. Although the AF-related health utilization may under-estimate the occurrence and frequency of AF episodes, it could be considered as a surrogate measure of the impact of AF on patients' health and quality of life, as well as the burden on the healthcare system.

Table Diagnosis Codes for Outcomes

	Diagnosis Codes	
	ICD-9-CM	**ICD-10-CM**
Ischemic stroke	433.x1, 434.x1, 436	I63.x
Systemic embolism	444.x	I74.x

We did not present diagnosis codes for comorbidities due to the large amount of codes. The analyst has access to SAS codes (with diagnosis and procedure codes inside) for all the comorbidities and outcomes.

4.3 Missing Data

Studies using administrative claims data generally do not have the problem of missing data. We will define the presence of a condition, outcome or drug use by the presence of a claim with eligible diagnosis or procedure codes or prescription fills. Patients would be considered having a comorbidity, outcome or drug exposure if they have a claim, and would be considered not having a comorbidity, outcome or drug exposure if they do not have a claim. Therefore, we do not have missing data in comorbidities, drug use, or outcomes. However, misclassification may exist. This is a limitation of using claims data, but the algorithms used to define our outcomes of interest and important covariates are commonly used and demonstrated good performance in previous studies.[19-23] Our internal validation suggested good performance of the algorithms. We anticipate that any existing residual misclassification is non-differential between treatment groups and should not meaningful impact our findings.

For the demographic data (age and gender), we typically will delete a very small percentage of patients with invalid demographic data during the cohort creation process. For race, the categories in the database are non-Hispanic white, non-Hispanic black, Hispanic, Asian, other and unknown. The other and unknown will be used as a separate category in the propensity score model.

The values of biomarkers all had missing value to some extent. We will not impute the missing values, because some biomarkers were only tested in patients with certain conditions, as such it is difficult to impute the missing values. In our past studies,

both the values and proportion of missing values were well balanced after balancing on all other patient characteristics.

4.4 Validation of Diagnosis Codes

The codes and algorithms used herein have been commonly used and validated in many previous studies.[19-23] We also leveraged the ability to link to laboratory results and electronic health records to validate our diagnosis codes. For example, we compared eGFR with the presence of a diagnosis code of Stage 3-4 chronic kidney disease (CKD) in those who did not have renal failure. We found 88% of patients who had a diagnosis of Stage 3-4 CKD had eGFR <60 mL/min/1.73m^2, and 90% of those who did not have a diagnosis had eGFR ≥60 mL/min/1.73m^2, which indicates good performance of the diagnosis codes. Moreover, the discrepancy between the diagnosis codes and eGFR could be because some patients may have a temporary decline in eGFR, but later recovered and did not develop to CKD or some patients had serum creatinine tests in facilities that did not submit data to OLDW.

We also compared the ejection fraction documented in electronic health records and the diagnosis codes of heart failure. Using an ejection fraction cutoff of ≤40% for systolic heart failure diagnosis codes and ejection fraction of ≥50% for diastolic heart failure codes; we observed the specificity of 91% and 81%, respectively and sensitivity of 81% and 91%, respectively. We will include systolic heart failure in the propensity score model as a surrogate for reduced ejection fraction, but we acknowledge the inherent limitations in classification of heart failure by ejection fraction.[8]

We conducted internal validation of major bleeding based on the International Society on Thrombosis and Haemostasis (ISTH) criteria[9]: (1) fatal bleeding, and/or, (2) symptomatic bleeding in a critical area or organ, such as intracranial, intraspinal, intraocular, retroperitoneal, intraarticular or pericardial, or intramuscular with compartment syndrome, and/or, (3) bleeding causing a fall in hemoglobin level of 2 g/dL or more, or leading to transfusion of two or more units of whole blood or red cells. We used ICD-9 and CPT procedure codes to identify transfusion, but we were not able to know the units of whole blood or red cells used in the transfusion. We also identified other procedures to control or manage bleeding, such as endoscopic procedures to address gastrointestinal bleeding, neurosurgical decompression for intracranial bleeding, evacuation of hematoma, or vascular embolization procedures to control bleeding. Among all bleeding events, one in four was bleeding in critical areas, and one third required transfusion. This is generally consistent with previous studies that adapted ISTH definition using administrative data.[10] Nearly 80% of patients had a procedure to control or manage bleeding. In patients with hemoglobin test results, we abstracted the most recent test performed within six months prior to the bleeding. The median time from the previous hemoglobin test to the date of bleeding is 29 (IQR 8-66) days. The median hemoglobin level during the bleeding was 8.2 (IQR 7.3-11.2) g/dL, with a median drop of 2.1 (IQR 1.1-3.6) g/dL. Among patients with transfusion, the median hemoglobin level was 7.3 (IQR 6.5-8.1) g/dL with a median drop of 2.7 (IQR 1.1-3.6) g/dL. In patients without transfusion, the median hemoglobin level was 10.4 (IQR 8.2-12.3) g/dL, with a median drop of 2.1 (IQR 1.2-3.6) g/dL. Overall, 95% of patients identified using diagnosis codes had bleeding in critical area, or a transfusion,

or a procedure used to control bleeding, which suggests high specificity of our algorithm. Even in the remaining 5% patients, the hemoglobin level was low, a median of 10.5 (IQR 8.7-12.0), with a median drop of 2.1 (IQR 1.2-3.5) g/dL.

5. STATISTICAL METHODS

5.1 Statistical Analyses

Propensity score matching will be used to balance the differences in baseline characteristics between patients who underwent concurrent LAAO versus those who did not. A propensity score, the probability of undergoing LAAO, was estimated using logistic regression based on socio-demographics, procedure-related characteristics, medical history, concurrent medication use, the year of the surgery, and the length of baseline period (variables in Table 1 and eTable 1). One-to-one nearest neighborhood caliper matching will be used to match patients based on the logit of the propensity score using a caliper equal to 0.2 of the standard deviation of the logit of the propensity score.[27] Patients will be exact matched on whether they had a history of AF and whether they used oral anticoagulants (OAC) within three months before the surgery. This is because the main analysis will be stratified by prior AF, and prior AF and baseline OAC should be the main predictors of follow-up OAC use.

Standardized difference will be used to assess the balance of covariates after matching and a standardized difference within 10% will be considered acceptable.[28] We will also assess balance of baseline characteristics in patients with and without prior AF separately, since the comparisons between LAAO and no LAAO will be stratified by the presence of prior AF.

Cox proportional hazards regression will be used to compare patients undergoing LAAO versus those who did not for stroke, mortality, and post-operative AF in propensity-score matched cohort, with robust sandwich estimates to account for the

clustering within matched sets.[29] The proportional hazards assumption will be tested on the basis of Schoenfeld residuals.[30] Poisson regression will be used to assess AF-related outpatient visits and hospitalizations. We will calculate both the absolute difference in event rates and the rate ratios.

5.2 Subgroup Analyses

We will perform subgroup analyses for stroke and mortality stratified by age, sex, race, surgery types, CHA_2DS_2-VASc score, HAS-BLED score, prior thromboembolism, prior bleeding, heart failure, and stage 3-5 chronic kidney disease (CKD). The subgroup analyses will be performed in patients with and without prior AF separately. The exact cut off points for grouping the CHA_2DS_2-VASc score will depend on the distribution of these variables.

5.3 Sensitivity Analyses

First, we will assess falsification endpoints to test for residual confounding. Treatment effects estimated in observational studies are prone to unmeasured confounding. In recent years, falsification end point, also called control outcome, has become a popular method to assess for unmeasured confounding.[31-33] A falsification endpoint is a health outcome that researchers believe is highly unlikely to be casually related to the treatment in question. If a significant relationship is found between the treatment and a falsification endpoint, it may indicate the treatment groups are different in some unmeasured ways, i.e. the existence of unmeasured confounding. This method is similar to a negative control, a routine precaution taken in the design of biologic laboratory experiments, and is recommended to be used to detect confounding and bias

in observational studies.[32,34,35] This method is particularly useful in observational studies comparing different treatment options, because the unmeasured confounding in these studies tend to make one group systematically healthier or less susceptible to adverse outcomes than the other group.

We selected three endpoints that that are unlikely to be a result of undergoing LAAO – emergency room visit or hospitalization related to chronic obstructive pulmonary disease (COPD), pneumonia, and fracture. If a significant relationship were to be found between LAAO and any of these endpoints, it would indicate the existence of residual confounding.[31]

There are many other methods to test for residual confounding, however, as noted in previous studies, the assumption of no unmeasured confounding cannot be formally tested; instead, subject matter knowledge is required in designing the study so that all confounders are accounted for.[36,37] We will include all potential confounders that we can measure, and some unmeasured clinical characteristics (e.g. left atrial size) are highly correlated with the characteristics we measure (e.g., hypertension and valvular heart disease), and thus it is likely that they will be balanced too.

Second, we will use inverse probability weighting instead of propensity score matching. The main analyses will be conducted using propensity score matching, since the prevalence of LAAO treatment is likely low, and the two treatment groups will be very difference. Propensity score matching may outperform matching in this case.[38] A previous study has shown that covariate adjustment and propensity score matching performed well in all of their examples, and propensity score weighting gave imprecise

estimates of treatment effect and undue influence to a small number of observations when substantial confounding was present.[39]

Third, we will use the method of Fine and Gray, considering death as a competing risk for stroke and post-operative AF.[40]

Fourth, because the use of OAC during follow up may affect the risk of stroke and mortality, we will assess whether the OAC use during follow up differed between patients treated with or without LAAO in the propensity score matched patients. The OAC use will be examined as the proportion of patients who received OAC at different time points during follow up. We will also assess the risks of stroke and mortality by whether patients received OAC during follow up. We will try a few different ways to define OAC during follow-up, such as a prescription within the first 14 days of surgery, a prescription within 30 day of the surgery, a prescription at any time during follow up, or proportion of days covered (PDC)≥80%.

Sixth, in patients without prior AF, some developed AF during follow up, and thus, we will assess stroke and mortality risks associated with LAAO stratified by whether or not patients developed AF during follow up

Seventh, previous studies found post-operative AF was associated with developing AF (or so called “late AF”).[41] We will look at AF-related health utilization excluding the first 30 days, stratified by whether patients had post-operative AF.

All analyses will be conducted using SAS 9.4 (SAS Institute Inc., Cary, North Carolina) and Stata 14.1 (Stata Corp, College Station, TX).

6. LIMITATIONS

First, despite careful adjustment, our study may still be subject to confounding. However, groups that are almost identical in 76 dimensions are not likely to differ dramatically in other unmeasured confounders. Some of the baseline characteristics that we will measure, e.g., prior treatment with ablation, cardioversion, anti-arrhythmic drugs, and medications for other chronic conditions, etc., could be proxies for unmeasured aspects of the underlying diseases. Furthermore, the test of falsification endpoints will provide some reassurance that there is no evidence for residual confounding (will need to confirm this statement).

Second, our study relies on administrative data to ascertain baseline characteristics and outcomes, which could be subject to misclassification. However, it is unlikely there is any systemic difference in the ascertainment of comorbidities and outcomes between different treatment groups, and thus, the misclassification should not meaningfully impact our comparisons between drugs. The diagnosis and procedure codes used in this study have been commonly used in previous studies, and demonstrated good performance in our internal validation using linked laboratory results and electronic health records as well as other validation studies with positive predictive value around 90%.[17,20,42-47]

Third, we are unable to distinguish between excision or exclusion by sutures or stapling since these procedures are all described by a single code. Nor do we have data on the apparent success of closure as gauged by intra-operative transesophageal echocardiography.

Last, our study includes only privately insured and Medicare Advantage patients. The prevalence of risk factors could be different in the Medicaid, Medicare Fee-for-Service, and uninsured populations, and lower socio-economic status in these populations could lead to poorer outcomes in patients with similar risk profile. However, the distribution of age, sex and minorities in the OLDW population is largely consistent with the general U.S. population. Furthermore, the insurance coverage rates are high in older Americans. Over 90% of Americans aged 50-64 have health insurance and over 75% had private health insurance.[48] One in three Medicare patients is enrolled in Medicare Advantage.[49] Although traditionally Medicare Advantage attracted healthier people, after the risk adjustment system was phased in from 2004-2007, the favorable risk selection has been largely reduced.[50] Our cohort is also less selective than most registries, because registries often focus on cardiology practices for recruitment and patients have to sign informed consent and agree to participate and to be actively followed. Therefore, findings from this database, which includes people of all ages and races and from diverse treatment settings, are more generalizable for the majority of older adults.

Appendix

Target Journals for this manuscript include JAMA, JACC and Circulation. Our recent relevant studies are listed below. Nearly all the methods used in this study have been used in our previous studies as well as other high-impact studies in our field.

1. Yao X, Tangri N, Gersh BJ, Sangaralingham LR, Shah ND, Nath KA, Noseworthy PA. Renal Outcomes in Anticoagulated Patients With Atrial Fibrillation. J Am Coll Cardiol. 2017 Nov 28; 70 (21):2621-2632 PMID: 29169468 DOI: 10.1016/j.jacc.2017.09.1087

2. Yao X, Shah ND, Sangaralingham LR, Gersh BJ, Noseworthy PA. Non-Vitamin K Antagonist Oral Anticoagulant Dosing in Patients With Atrial Fibrillation and Renal Dysfunction. J Am Coll Cardiol. 2017 Jun 13; 69 (23):2779-2790 PMID: 28595692 DOI: 10.1016/j.jacc.2017.03.600

3. Yao X, Shah ND, Sangaralingham LR, Gersh BJ, Noseworthy PA. Reply: NOAC Dosing in Patients With Atrial Fibrillation and Renal Dysfunction. J Am Coll Cardiol.2017 Nov 28;70:(21):2734-2735.

4. Del-Carpio Munoz F, Yao X, Abraham NS, Bellolio MF, Rabinstein AA, Asirvatham SJ, McBane RD, Gersh BJ, Shah ND, Noseworthy PA. Dabigatran Versus Warfarin in Relation to Renal Function in Patients With Atrial Fibrillation. J Am Coll Cardiol.2016 Jul 5;68:(1):129-31. PMID: 27364057 DOI: 10.1016/j.jacc.2016.04.031

5. Noseworthy PA, Yao X, Shah ND, Gersh BJ. Comparative effectiveness and safety of non-vitamin K antagonist oral anticoagulants versus warfarin in patients with atrial fibrillation and valvular heart disease. Int J Cardiol 2016 Apr 15; 209:181-3 Epub 2016 Feb 02 PMID: 26896618 DOI: 10.1016/j.ijcard.2016.02.005

6. Yao X, Abraham NS, Sangaralingham LR, Bellolio MF, McBane RD, Shah ND, Noseworthy PA. Effectiveness and Safety of Dabigatran, Rivaroxaban, and Apixaban Versus Warfarin in Nonvalvular Atrial Fibrillation. J Am Heart Assoc. 2016 Jun 13; 5: (6). PMID: 27412905 PMCID: 4937291 DOI: 10.1161/JAHA.116.003725

7. Abraham NS, Noseworthy PA, Yao X, Sangaralingham LR, Shah ND. Gastrointestinal Safety of Direct Oral Anticoagulants: A Large Population-Based Study. Gastroenterology. 2017 Apr; 152: (5)1014-1022.e1. PMID: 28043907 DOI: 10.1053/j.gastro.2016.12.018

8. Noseworthy PA, Yao X, Gersh BJ, Hargraves I, Shah ND, Montori VM. Baseline characteristics and event rates among anticoagulated patients with atrial fibrillation in practice and pivotal NOAC trials. Data Brief. 2017 Oct; 14:563-565 Epub 2017 Aug 09 PMID: 28861455 PMCID: 5568872 DOI: 10.1016/j.dib.2017.08.010

9. Noseworthy PA, Yao X, Gersh BJ, Hargraves I, Shah ND, Montori VM. Long-term stroke and bleeding risk in patients with atrial fibrillation treated with oral anticoagulants in contemporary practice: Providing evidence for shared decision-making. Int J Cardiol. 2017 Oct 15; 245:174-177 Epub 2017 July 14 PMID: 28733071 DOI: 10.1016/j.ijcard.2017.07.043

10. Yao X, Gersh BJ, Sangaralingham LR, Kent DM, Shah ND, Abraham NS, Noseworthy PA. Comparison of the CHA2DS2-VASc, CHADS2, HAS-BLED, ORBIT, and ATRIA Risk Scores in Predicting Non-Vitamin K Antagonist Oral Anticoagulants-Associated Bleeding in Patients With Atrial Fibrillation. Am J Cardiol. 2017 Nov 1; 120 (9):1549-1556 Epub 2017 July 31 PMID: 28844514 DOI: 10.1016/j.amjcard.2017.07.051

11. Yao X, Abraham NS, Alexander GC, Crown W, Montori VM, Sangaralingham LR, Gersh BJ, Shah ND, Noseworthy PA. Effect of Adherence to Oral Anticoagulants on Risk of Stroke and Major Bleeding Among Patients With Atrial Fibrillation. J Am Heart Assoc. 2016 Feb 23; 5: (2). PMID: 26908412 PMCID: 4802483 DOI: 10.1161/JAHA.115.003074

12. Noseworthy PA, Yao X, Shah ND, Gersh BJ. Stroke and Bleeding Risks in NOAC- and Warfarin-Treated Patients With Hypertrophic Cardiomyopathy and Atrial Fibrillation. J Am Coll Cardiol.2016 Jun 28;67:(25):3020-1. PMID: 27339501 DOI: 10.1016/j.jacc.2016.04.026

13. Noseworthy PA, Yao X, Abraham NS, Sangaralingham LR, McBane RD, Shah ND. Direct Comparison of Dabigatran, Rivaroxaban, and Apixaban for Effectiveness and Safety in Nonvalvular Atrial Fibrillation. Chest. 2016 Dec; 150: (6)1302-1312. PMID: 27938741 DOI: 10.1016/j.chest.2016.07.013

14. Noseworthy PA, Yao X, Deshmukh AJ, Van Houten H, Sangaralingham LR, Siontis KC, Piccini JP Sr, Asirvatham SJ, Friedman PA, Packer DL, Gersh BJ, Shah ND. Patterns of Anticoagulation Use and Cardioembolic Risk After Catheter Ablation for Atrial Fibrillation. J Am Heart Assoc. 2015 Nov; 4: (11). PMID: 26541393 PMCID: 4845220 DOI: 10.1161/JAHA.115.002597

15. Madhavan M, Yao X, Sangaralingham LR, Asirvatham SJ, Friedman PA, McLeod CJ, Sugrue AM, DeSimone CV, Noseworthy PA. Ischemic Stroke or Systemic Embolism After Transseptal Ablation of Arrhythmias in Patients With Cardiac Implantable

Electronic Devices. J Am Heart Assoc. 2016 Apr 13; 5: (4)e003163. PMID: 27076562
PMCID: 4843539 DOI: 10.1161/JAHA.115.003163

References

1. Fingar K, Stocks C, Weiss A, Steiner C. Most frequent operating room procedures performed in US hospitals, 2003–2012. In: Agency for Healthcare Research and Quality, ed. *HCUP Statistical Brief# 186* 2014.
2. Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of atrial fibrillation on the risk of death the Framingham Heart Study. *Circulation.* 1998;98(10):946-952.
3. Kannel WB, Wolf PA, Benjamin EJ, Levy D. Prevalence, incidence, prognosis, and predisposing conditions for atrial fibrillation: population-based estimates. *Am J Cardiol.* 1998;82(7):2N-9N.
4. Blackshear JL, Odell JA. Appendage obliteration to reduce stroke in cardiac surgical patients with atrial fibrillation. *Ann Thorac Surg.* 1996;61(2):755-759.
5. Melduni RM, Schaff HV, Lee HC, et al. Impact of Left Atrial Appendage Closure During Cardiac Surgery on the Occurrence of Early Postoperative Atrial Fibrillation, Stroke, and Mortality: A Propensity Score-Matched Analysis of 10 633 Patients. *Circulation.* 2017;135(4):366-378.
6. Maisel WH, Rawn JD, Stevenson WG. ATrial fibrillation after cardiac surgery. *Annals of Internal Medicine.* 2001;135(12):1061-1073.
7. Mathew JP, Fontes ML, Tudor IC, et al. A multicenter risk index for atrial fibrillation after cardiac surgery. *JAMA.* 2004;291(14):1720-1729.
8. Hoit BD, Shao Y, Tsai LM, Patel R, Gabel M, Walsh RA. Altered left atrial compliance after atrial appendectomy. Influence on left atrial and ventricular filling. *Circ Res.* 1993;72(1):167-175.
9. Hoit BD, Walsh RA. Regional atrial distensibility. *Am J Physiol.* 1992;262(5 Pt 2):H1356-1360.
10. Davis CA, 3rd, Rembert JC, Greenfield JC, Jr. Compliance of left atrium with and without left atrium appendage. *Am J Physiol.* 1990;259(4 Pt 2):H1006-1008.
11. Wallace PJ, Shah ND, Dennen T, Bleicher PA, Crown WH. Optum Labs: building a novel node in the learning health care system. *Health Aff (Millwood).* 2014;33(7):1187-1194.
12. Optum. Optum Research Data Assets. 2015; https://www.optum.com/content/dam/optum/resources/productSheets/5302_Data_Assets_Chart_Sheet_ISPOR.pdf. Accessed 6/22/2015.
13. Mitchell LB, Exner DV, Wyse DG, et al. Prophylactic Oral Amiodarone for the Prevention of Arrhythmias that Begin Early After Revascularization, Valve Replacement, or Repair: PAPABEAR: a randomized controlled trial. *Jama.* 2005;294(24):3093-3100.
14. Fauchier L, Pierre B, de Labriolle A, Grimard C, Zannad N, Babuty D. Antiarrhythmic effect of statin therapy and atrial fibrillation a meta-analysis of randomized controlled trials. *J Am Coll Cardiol.* 2008;51(8):828-835.
15. Brinkman W, Herbert MA, O'Brien S, et al. Preoperative beta-blocker use in coronary artery bypass grafting surgery: national database analysis. *JAMA Intern Med.* 2014;174(8):1320-1327.

16. Lee JZ, Singh N, Howe CL, et al. Colchicine for prevention of post-operative atrial fibrillation: a meta-analysis. *JACC: Clinical Electrophysiology.* 2016;2(1):78-85.
17. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R, Dublin S. Identifying atrial fibrillation from electronic medical data: a systematic review. *Pharmacoepidemiol Drug Saf.* 2012;21(0 1):141-147.
18. Piccini JP, Hammill BG, Sinner MF, et al. Incidence and prevalence of atrial fibrillation and associated mortality among Medicare beneficiaries, 1993–2007. *Circ Cardiovasc Qual Outcomes.* 2012;5(1):85-93.
19. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol.* 2004;57(2):131-141.
20. Tirschwell DL, Longstreth W. Validating administrative data in stroke research. *Stroke.* 2002;33(10):2465-2470.
21. Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE, Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. *Pharmacoepidemiol Drug Saf.* 2011;20(6):560-566.
22. Arnason T, Wells P, Van Walraven C, Forster A. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. *Thromb Res.* 2006;118(2):253-262.
23. Hwang YJ, Shariff SZ, Gandhi S, et al. Validity of the International Classification of Diseases, Tenth Revision code for acute kidney injury in elderly patients at presentation to the emergency department and at hospital admission. *BMJ Open.* 2012;2(6):e001821.
24. Shah KS, Xu H, Matsouaka RA, et al. Heart Failure With Preserved, Borderline, and Reduced Ejection Fraction: 5-Year Outcomes. *J Am Coll Cardiol.* 2017.
25. Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. *Journal of Thrombosis and Haemostasis.* 2005;3(4):692-694.
26. Jasuja GK, Reisman JI, Miller DR, et al. Identifying major hemorrhage with automated data: results of the Veterans Affairs study to improve anticoagulation (VARIA). *Thrombosis research.* 2013;131(1):31-36.
27. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat.* 2011;10(2):150-161.
28. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med.* 2009;28(25):3083-3107.
29. Gayat E, Resche-Rigon M, Mary JY, Porcher R. Propensity score applied to survival data analysis through proportional hazards models: a Monte Carlo study. *Pharm Stat.* 2012;11(3):222-229.
30. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994;81(3):515-526.
31. Prasad V, Jena AB. Prespecified falsification end points: can they validate true observational associations? *JAMA.* 2013;309(3):241-242.

32. Dusetzina SB, Brookhart MA, Maciejewski ML. Control Outcomes and Exposures for Improving Internal Validity of Nonrandomized Studies. *Health Serv Res.* 2015;50(5):1432-1451.
33. Ioannidis JP. Are mortality differences detected by administrative data reliable and actionable? *Jama.* 2013;309(13):1410-1411.
34. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology.* 2010;21(3):383-388.
35. Wimmer NJ, Resnic FS, Mauri L, Matheny ME, Yeh RW. Comparison of transradial versus transfemoral percutaneous coronary intervention in routine practice: evidence for the importance of "falsification hypotheses" in observational studies of comparative effectiveness. *J Am Coll Cardiol.* 2013;62(22):2147-2148.
36. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* 2015;34(28):3661-3679.
37. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008;168(6):656-664.
38. Austin PC, Schuster T. The performance of different propensity score methods for estimating absolute effects of treatments on survival outcomes: a simulation study. *Statistical methods in medical research.* 2014:0962280213519716.
39. Elze MC, Gregson J, Baber U, et al. Comparison of Propensity Score Methods and Covariate Adjustment: Evaluation in 4 Cardiovascular Studies. *J Am Coll Cardiol.* 2017;69(3):345-357.
40. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association.* 1999;94(446):496-509.
41. Melduni RM, Schaff HV, Bailey KR, et al. Implications of new-onset atrial fibrillation after cardiac surgery on long-term prognosis: a community-based study. *Am Heart J.* 2015;170(4):659-668.
42. Kumamaru H, Judd SE, Curtis JR, et al. Validity of claims-based stroke algorithms in contemporary medicare data reasons for geographic and racial differences in stroke (REGARDS) study linked with medicare claims. *Circulation: Cardiovascular Quality and Outcomes.* 2014;7(4):611-619.
43. Kokotailo RA, Hill MD. Coding of stroke and stroke risk factors using international classification of diseases, revisions 9 and 10. *Stroke.* 2005;36(8):1776-1781.
44. Noseworthy PA, Yao X, Abraham NS, Sangaralingham LR, McBane RD, Shah ND. Direct comparison of dabigatran, rivaroxaban, and apixaban for effectiveness and safety in nonvalvular atrial fibrillation. *Chest.* 2016;150(6):1302-1312.
45. Yao X, Shah ND, Sangaralingham LR, Gersh BJ, Noseworthy PA. Non-vitamin K antagonist oral anticoagulant dosing in patients with atrial fibrillation and renal dysfunction. *J Am Coll Cardiol.* 2017;69(23):2779-2790.

46. Fan J, Arruda-Olson AM, Leibson CL, et al. Billing code algorithms to identify cases of peripheral artery disease from administrative data. *J Am Med Inform Assoc.* 2013;20(e2):e349-354.
47. Yao X, Tangri N, Gersh BJ, et al. Renal Outcomes in Anticoagulated Patients With Atrial Fibrillation. *J Am Coll Cardiol.* 2017;70(21):2621-2632.
48. Barnett JC, Vornovitsky MS. Health insurance coverage in the United States: 2015. *US Census Bureau, Current Population Reports, Report.* 2016(P60-257).
49. Jacobson G, Casillas G, Damico A, Neuman T, Gold M. Medicare Advantage 2016 spotlight: enrollment market update. 2016; http://www.kff.org/medicare/issue-brief/medicare-advantage-2016-spotlight-enrollment-market-update/.
50. McWilliams JM, Hsu J, Newhouse JP. New risk-adjustment system was associated with reduced favorable selection in medicare advantage. *Health Aff (Millwood).* 2012;31(12):2630-2640.