SUPPLEMENTAL MATERIAL
Data S1.

STATISTICAL ANALYSIS PLAN

The purpose of this analysis plan is to provide a 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.
### ABBREVIATIONS

| Abbreviation | Description                           |
|--------------|---------------------------------------|
| AF           | Atrial fibrillation                    |
| CI           | Confidence interval                    |
| HR           | Hazard ratio                           |
| IQR          | Interquartile range                    |
| LAAO         | Left atrial appendage occlusion        |
| NOAC         | Non-vitamin K antagonist oral anticoagulant |
1. BACKGROUND AND OBJECTIVES

Lifelong oral anticoagulation is the mainstay for stroke prevention for most patients with atrial fibrillation (AF).\textsuperscript{1-3} However, systemic anticoagulation increase the risk of bleeding and the adherence to medication is poor, leaving many patients under-treated.\textsuperscript{4-6} Percutaneous left atrial appendage occlusion (LAAO) is an alternative to anticoagulation for select patients. In the two pivotal randomized controlled trials (RCTs) that supported the FDA approval, LAAO demonstrated similar risks of stroke or systemic embolism and major bleeding, but a reduction in mortality when compared to warfarin.\textsuperscript{7-9}

However, for most of the 33.5 million patients with AF worldwide, the non-vitamin K antagonist oral anticoagulant (NOACs) are now first-line therapy for stroke prevention. The performance of LAAO relative to NOACs is less certain. The only available evidence is from a recently completed randomized controlled trial (RCT) demonstrated noninferiority between LAAO and NOACs in 402 high-risk patients. It is uncertain whether these results would hold in a larger and broader patient population.

Give the quick uptake in routine practice and the substantial healthcare utilization and cost associated with LAAO implantation, there is a critical need to provide more evidence to guide shared decision making. Therefore, the current study aimed to investigate the risk of stroke, bleeding, and mortality associated with LAAO in comparison to NOACs in a large national cohort of patients managed in routine clinical practice.
2. STUDY DESIGN AND DATA SOURCE

A retrospective cohort analysis will be conducted using OptumLabs Data Warehouse, which contains over 130 million privately insured and Medicare Advantage enrollees of all ages and races from all 50 states.\textsuperscript{10,11} In 2014, this amounted to 19\% of all commercially insured and Medicare Advantage beneficiaries in the U.S.

OptumLabs includes patients of all ages and races from all 50 states. To demonstrate the similarities between the OptumLabs cohort and a nationally representative cohort, we used 4,991,144 people aged \( \geq40 \) years from the OptumLabs cohort in 2013 and compared them with a nationally representative cohort aged \( \geq40 \) years in a previous study using the Medical Expenditure Panel Survey (MEPS) 2012-2013 data.\textsuperscript{12}

Table 1. Comparison of OptumLabs Data and a U.S. Nationally Representative Cohort

|                         | OptumLabs Cohort | Nationally Representative Cohort |
|-------------------------|------------------|----------------------------------|
| Age, mean, y            |                  |                                  |
| Age category            |                  |                                  |
| <65                     | 58.3             | 58.5                             |
| 65-74                   | 69.4             | 69.9                             |
| \( \geq75 \)            | 18.4             | 17.4                             |
| Sex                     |                  |                                  |
| Male                    | 47.6             | 47.8                             |
| Female                  | 52.4             | 52.2                             |
| Race/ethnicity          |                  |                                  |
| Asian                   | 3.9              | 4.9                              |
| Black                   | 10.6             | 10.8                             |
| Hispanic                | 8.4              | 11.4                             |
| White                   | 74.5             | 71.1                             |
| Other/unknown           | 2.7              | 1.8                              |
In addition, the Medicare Fee-for-Service claims data were obtained and linked to the commercial insurance and Medicare Advantage claims at a patient level. Therefore, the study will be the first to create longitudinal datasets linking private insurance, Medicare Advantage, and Medicare Fee-for-Service records of the same patient in millions of patients with AF.

Most administrative datasets are limited to patients within a certain age range (e.g., ≥65 years), geographic regions, or practice settings. The fragmented nature of such data has resulted in a lack of statistical power and generalizability. This study will be able to assemble the largest AF cohort to date with long-term follow up including patients of all ages and races managed at heterogeneous practice settings from all 50 states.
3. STUDY POPULATION

The study population will be adult patients (≥18 years) with AF treated with LAAO or a NOAC between 3/13/2015 and 12/31/2018. For LAAO recipients, the index date will be the discharge date. For the NOAC group, the index date will be the initiation (i.e., the first prescription) of one of the NOACs (apixaban, dabigatran, edoxaban, or rivaroxaban). If patients initiated multiple NOACs during the study period, one of them will be randomly selected as the index date. The NOAC-treated patients could have been treated with a different NOAC or warfarin before the index date. Such NOAC-treated patients will be included in the cohort since many LAAO patients also received an oral anticoagulant drug before LAAO.

Patients will be required to have at least 12 months of continuous enrollment in health insurance plans before the index date to have sufficient data to capture patients’ baseline medical history. Patients who underwent surgical left atrial appendage closure at any time will be excluded. To ensure the NOAC-treated patients received treatment for AF stroke prevention, they were required to have an AF diagnosis within the past 12 months. Patients who received a NOAC dose for other indications were excluded (i.e., rivaroxaban 2.5 mg and 10 mg; dabigatran 110 mg; edoxaban 15 mg). Patients who had a diagnosis of venous thromboembolism (VTE) within the 12 months were excluded; patients who received a hip replacement within the past 6 weeks or a knee replacement within the past 2 weeks were excluded.
4. MEASUREMENTS

4.1 Baseline Characteristics

Baseline characteristics include socio-demographic characteristics, medical history, and concurrent medication use. Socio-demographic characteristics include age, sex, race/ethnicity, region, and health plan, determined at the time of index date. Race/ethnicity is provided by OptumLabs, classified as non-Hispanic White (White), non-Hispanic Black (Black), Asian, Hispanic, or other/unknown. Self-report was the primary source, and when it was missing, imputation was made by the data provider based on other available administrative data.\textsuperscript{13}

The medical history will be determined using physician, facility, and pharmacy claims before the index date. All data available will be used to establish patients’ medical history, and the length of the baseline period will be included in the propensity score model to avoid any potential bias. In our previous studies, the baseline period was on average 3-4 years, and there was no substantial difference in the length of the baseline period among different treatment groups, especially after propensity score (PS) matching or weighting. Concurrent medication, such as antiplatelet, anti-hypertensive, and anti-diabetic medications, will be captured within 3 months prior to the index date.

4.2 Follow up and Outcomes

The primary outcome will be a composite endpoint of ischemic stroke or systemic embolism (hereafter referred to as stroke), major bleeding, and all-cause mortality. Secondary outcome will be each of the individual outcomes and intracranial bleeding.
Patients will be followed from the index date until the end of the study period (December 31, 2018), the end of enrollment in health insurance plans, or death, whichever happened first. The primary analysis will not censor NOAC-treated patients on the discontinuation or switch of the drugs. Avoiding medication non-adherence is one of the advantages of LAAO, as a result, if the follow-up of NOAC-treated patients is censored, these patients might have shorter follow up and the results might favor NOAC-treated patients.

Stroke and bleeding will be defined as a primary diagnosis on an emergency room visit or a primary diagnosis on the inpatient admission date. Mortality will be identified based on the information provided by the Centers for Medicare and Medicaid Services, Social Security Death Master File, and discharge status.

Table 2. Diagnosis Codes for Outcomes

| Diagnosis Codes | ICD-9-CM | ICD-10-CM |
|-----------------|----------|-----------|
| Intracranial bleeding | 430, 431, 432.x, 852.x, 853.x, 800.2x, 800.3x, 800.7x, 800.8x, 801.2x, 801.3x, 801.7x, 801.8x, 803.2x, 803.3x, 803.7x, 803.8x, 804.2x, 804.3x, 804.7x, 804.8x | I60.x, I61.x, I62.x, S06.34x, S06.35x, S06.36x, S06.37x, S06.38x, S06.4x, S06.5x, S06.6x |
| Gastrointestinal bleeding | 456.0, 456.20, 530.21, 530.7, 530.82, 531.0x, 531.2x, 531.4x, 531.6x, 532.0x, 532.2x, 532.4x, 532.6x, 533.0x, 533.2x, 533.4x, 533.6x, 534.0x, 534.2x, 534.4x, 534.6x, 535.01, 535.11, 535.21, 535.31, 535.41, 535.51, 535.61, 535.71, 537.83, 537.84, 562.02, 562.03, 562.12, 562.13, 569.3, 569.85, 569.86, 578.x | I85.01, I85.11, K22.11, K22.6, 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, K29.x1, K31.811, K31.82, K55.21, K57.x1, K57.x3, K62.5, K63.81, K92.0, K92.1, K92.2 |
| Other bleeding | 423.0, 459.0, 568.81, 596.7, 599.71, 719.1x, 784.8, 786.3 | I31.2, K66.1, M25.0, R04.1, R04.2, R31.0, R58 |
We did not present diagnosis codes for comorbidities due to a large number of codes. The analyst has access to SAS codes for all the comorbidities and outcomes and can provide the codes upon request.

4.3 Missing Data

Studies using administrative claims data generally do not have the problem of missing data, *per se*. 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 will be considered to have a comorbidity, outcome or drug exposure if they have a claim, and will 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 have demonstrated good performance in previous studies.\(^{14-18}\) Our internal validation also suggested good performance of the algorithms. We anticipate that any existing residual misclassification will be non-differential between treatment groups and should not meaningfully impact our findings.

For the demographic data, we typically will delete a very small percentage (<1%) of patients with invalid demographic data during the cohort creation process (e.g., missing residence region or inconsistent birth year). For race/ethnicity, 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.

4.4 Internal Validation of Diagnosis Codes

The codes and algorithms used herein have been commonly used and validated in many previous studies.\textsuperscript{14-20}

We also leveraged the ability to link to laboratory results and electronic health records to validate our diagnosis codes. For example, we compared the ejection fraction documented in electronic health records and the diagnosis codes for H.F. Using a cutoff of LVEF ≤40\% for heart failure with reduced ejection fraction (HFrEF) diagnosis codes and LVEF ≥50\% for heart failure with preserved ejection fraction (HFpEF) codes, we observed the specificity of 91\% and 81\%, respectively, and sensitivity of 81\% and 91\%, respectively.

We also 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 the OptumLabs Data Warehouse.
We have also conducted validation of the major bleeding diagnosis codes based on the International Society on Thrombosis and Haemostasis (ISTH) criteria: (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. 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

A PS, i.e., the probability of undergoing LAAO, will be estimated using logistic regression based on all the baseline characteristics. The overlap weighting method will be used to balance treatment groups. The overlap weight will be calculated as one minus PS for the LAAO patients, and PS for the NOAC patients. Other commonly used PS methods include matching and inverse probability treatment weighting (IPTW). We chose not to use matching because a large number of patients may be dropped during matching. We chose not to use IPTW, since IPTW gave imprecise estimates of treatment effect and undue influence to a small number of observations when substantial confounding was present.\textsuperscript{23} The performance of IPTW often gets worse when the prevalence of treatment is low.\textsuperscript{24}

The overlap weighting was chosen because this approach minimizes the asymptotic variance of the treatment effect, while also possessing a desirable exact balance property.\textsuperscript{25} Unlike IPTW, the overlap weights are bounded between 0 and 1 and thus are less sensitive to extreme weights. Compared to the common practice of truncating weights or discarding patients with extreme weights, the overlap weights avoid this arbitrary choice of a cutoff point for inclusion. The overlap weight also possesses an attractive exact balance property, i.e., the means of all variables
(including the proportions of a binary or categorical variable) will be exactly the same between treatment and control groups after weighting.

The results using the overlap weight should be interpreted as the average treatment effect for the overlap population. The overlap population typically represents a target population of intrinsic substantive interest, i.e., patients who could appear in either treatment groups. In such patients, clinical consensus regarding the treatment choice is often ambiguous, and thus research is most needed to guide decision making.

The balance between treatment groups will be evaluated by comparing standardized mean differences of baseline covariates between two groups. A standardized mean difference of 0.2 or less was deemed to be an acceptable balance.\textsuperscript{26,27} If any substantial imbalance exists, a sensitivity analysis will be performed adjusting for such characteristics in regression models.

Cox proportional hazards regression will be used to compare treatments in the weighted population, with a robust sandwich estimator for variance estimation. The Fine and Gray method will be used to consider death as a competing risk when assessing non-fatal outcomes.\textsuperscript{28} The event rate per 100 person-years will be calculated and the Kaplan-Meier curves were be plotted and used to estimate the cumulative risks for each outcome at different time points. Statistical tests for the proportional hazards assumption are unnecessary because it is expected that the hazard ratio will vary over the follow-up period and tests of the proportional hazards assumption yielding high P values are probably under-powered.\textsuperscript{29} Therefore, a hazard ratio needs to be interpreted
as a weighted average of the true hazard ratios over the entire follow-up period and the absolute risks will be provided at different time points to facilitate the interpretation.

Data management will be performed using SAS Enterprise Guide 7.1; all other analyses will be performed using Stata 15.1.

5.2 Subgroup Analyses

Subgroup analyses will be performed stratified by age (<75 years and ≥75 years), sex (male and female), race (white and non-white), and potential contraindications to anticoagulation. Contraindications to anticoagulation include (1) intracranial bleeding, (2) major extracranial bleeding, (3) end-stage kidney disease requiring dialysis, (4) a history of fall, (5) coagulation defects, (6) gastrointestinal lesions, (7) end-stage liver disease, (8) cerebral amyloid angiopathy, (9) cerebral aneurysms, (10) pericarditis/pericardial effusions. Based on the prevalence of each contraindication, we plan to perform subgroup analysis grouping patients to: no contraindication, prior intracranial bleeding, prior extracranial bleeding, and all other contraindications.

Since an increasing number of subgroup analyses could increase the chance of false positive results, we pre-specified the above subgroups. For all analyses performed in this study, adjustment for multiple testing will not be performed. The sample size will be large, and thus, even with the conservative Bonferroni adjustment, many tests will still be statistically significant. All the analyses except those related to the primary outcome will be considered exploratory.
5.3 Residual Confounding

First, falsification endpoints will be assessed to test for residual confounding. Treatment effects estimated in observational studies are prone to unmeasured confounding. In recent years, the falsification endpoint, also called control outcome, has become a popular method to assess for unmeasured confounding. A falsification endpoint is a health outcome that researchers believe is highly unlikely to be causally 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 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. We selected two endpoints that are unlikely to be a result of undergoing LAAO – emergency room visit or hospitalization related to pneumonia and fracture.

Second, e-value will be used to assess how strong an unmeasured confounder would have to be to explain away an observed treatment–outcome relationship. This is a sensitivity analysis technique that is easy to use and does not itself make strong assumptions. The E-value is the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and outcome, conditional on the measured covariates, to fully explain away a specific treatment–outcome association.
5.4 Sensitivity Analyses

First, whether or not and which antiplatelet and anticoagulant drugs LAAO patients received after the device implantation will be descriptively assessed. There are many different post-procedural antithrombotic regimens regarding the choice, combination, and duration of the drugs. Which regimen is best for which patient is a question to be studied within a LAAO population (preferably via a pragmatic trial) and is out of the scope of this study.

Second, apixaban and rivaroxaban will be compared to LAAO, respectively. Apixaban and rivaroxaban are the two most commonly used NOACs. PS will be recalculated to balance patients treated with apixaban or LAAO, and another PS will be calculated to balance patients treated with rivaroxaban or LAAO. Regression analyses will be rerun in each of the two weighted populations.

6. Limitations

First, as with any observational studies, the results are subject to residual confounding even after careful adjustment. However, the groups identical on 87 dimensions were unlikely to substantially differ in other aspects, since many of the measured characteristics are highly correlated with unmeasured ones. For example, age, valvular heart disease, hypertension, previous AADs, and cardioversion are associated with unmeasured characteristics, such as left atrial diameter, AF pattern or burden.

Second, certain variables are not available in our dataset, e.g., aspirin, device-related thrombosis (DRT), and implant success rate. There are many different post-
procedural antithrombotic regimens for Watchman patients, but all would include low-dose aspirin for a minimum of 12 months.36 There is no billing code for DRT but the major consequence of DRT, i.e., stroke, will be assessed.37,38 Implant success rate is generally very high, with a reported 98.3% in a previous study using registry data.39

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 systematic 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 (described in Section 4.4) as well as other validation studies with positive predictive value around 90%.14,40-43
References

1. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *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*. 2016;18(11):1609-1678.

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. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. *Circulation*. 2014;129(8):837-847.

4. Hsu JC, Maddox TM, Kennedy KF, et al. Oral anticoagulant therapy prescription in patients with atrial fibrillation across the spectrum of stroke risk: insights from the NCDR PINNACLE registry. *JAMA cardiology*. 2016;1(1):55-62.

5. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation*. 2007;115(21):2689-2696.
6. Yao X, Abraham NS, Alexander GC, et al. Effect of adherence to oral anticoagulants on risk of stroke and major bleeding among patients with atrial fibrillation. *Journal of the American Heart Association*. 2016;5(2):e003074.

7. Reddy VY, Sievert H, Halperin J, et al. Percutaneous left atrial appendage closure vs warfarin for atrial fibrillation: a randomized clinical trial. *Jama*. 2014;312(19):1988-1998.

8. Holmes DR, Jr., Kar S, Price MJ, et al. Prospective randomized evaluation of the Watchman Left Atrial Appendage Closure device in patients with atrial fibrillation versus long-term warfarin therapy: the PREVAIL trial. *J Am Coll Cardiol*. 2014;64(1):1-12.

9. Reddy VY, Doshi SK, Kar S, et al. 5-Year Outcomes After Left Atrial Appendage Closure: From the PREVAIL and PROTECT AF Trials. *J Am Coll Cardiol*. 2017;70(24):2964-2975.

10. Wallace PJ, Shah ND, Dennen T, Bleicher PA, Crown WH. Optum Labs: building a novel node in the learning health care system. *Health affairs (Project Hope)*. 2014;33(7):1187-1194.

11. Optum. Optum Research Data Assets. 
   
   [https://www.optum.com/content/dam/optum/resources/productSheets/5302_Data Assets_Chart_Sheet_ISPOR.pdf](https://www.optum.com/content/dam/optum/resources/productSheets/5302_Data Assets_Chart_Sheet_ISPOR.pdf). Published 2015. Accessed 6/22/2015.
12. Salami JA, Warraich H, Valero-Elizondo J, et al. National Trends in Statin Use and Expenditures in the US Adult Population From 2002 to 2013: Insights From the Medical Expenditure Panel Survey. *JAMA cardiology.* 2017;2(1):56-65.

13. Hershman DL, Tsui J, Wright JD, Coromilas EJ, Tsai WY, Neugut Al. Household net worth, racial disparities, and hormonal therapy adherence among women with early-stage breast cancer. *Journal of Clinical Oncology.* 2015;33(9):1053-1059.

14. Tirschwell DL, Longstreth W. Validating administrative data in stroke research. *Stroke.* 2002;33(10):2465-2470.

15. 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. *Pharmacoepidemiology and drug safety.* 2011;20(6):560-566.

16. Arnason T, Wells P, Van Walraven C, Forster A. Accuracy of coding for possible warfarin complications in hospital discharge abstracts. *Thrombosis research.* 2006;118(2):253-262.

17. 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.
18. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *Journal of clinical epidemiology.* 2004;57(2):131-141.

19. 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.

20. 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.

21. 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.

22. 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.

23. 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.
24. Austin PC, Schuster T. The performance of different propensity score methods for estimating absolute effects of treatments on survival outcomes: A simulation study. *Stat Methods Med Res.* 2016;25(5):2214-2237.

25. Li F, Morgan KL, Zaslavsky AM. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association.* 2017:1-11.

26. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in medicine.* 2013;32(19):3388-3414.

27. Goldstone AB, Chiu P, Baiocch M, et al. Mechanical or Biologic Prostheses for Aortic-Valve and Mitral-Valve Replacement. *The New England journal of medicine.* 2017;377(19):1847-1857.

28. 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.

29. Stensrud MJ, Hernan MA. Why Test for Proportional Hazards? *Jama.* 2020.

30. Prasad V, Jena AB. Prespecified falsification end points: can they validate true observational associations? *Jama.* 2013;309(3):241-242.
31. Dusetzina SB, Brookhart MA, Maciejewski ML. Control Outcomes and Exposures for Improving Internal Validity of Nonrandomized Studies. *Health services research*. 2015;50(5):1432-1451.

32. Ioannidis JP. Are mortality differences detected by administrative data reliable and actionable? *Jama*. 2013;309(13):1410-1411.

33. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology (Cambridge, Mass)*. 2010;21(3):383-388.

34. 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.

35. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of internal medicine*. 2017;167(4):268-274.

36. Saw J, Nielsen-Kudsk JE, Bergmann M, et al. Antithrombotic Therapy and Device-Related Thrombosis Following Endovascular Left Atrial Appendage Closure. *JACC Cardiovascular interventions*. 2019;12(11):1067-1076.
37. Dukkipati SR, Kar S, Holmes DR, Jr., et al. Device-Related Thrombus After Left Atrial Appendage Closure: Incidence, Predictors, and Outcomes. *Circulation*. 2018.

38. Fauchier L, Cinaud A, Brigadeau F, et al. Device-Related Thrombosis After Percutaneous Left Atrial Appendage Occlusion for Atrial Fibrillation. *J Am Coll Cardiol.* 2018;71(14):1528-1536.

39. Freeman JV, Varosy P, Price MJ, et al. The NCDR Left Atrial Appendage Occlusion Registry. *J Am Coll Cardiol.* 2020;75(13):1503-1518.

40. 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.

41. 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.

42. Jensen PN, Johnson K, Floyd J, Heckbert SR, Carnahan R, Dublin S. Identifying atrial fibrillation from electronic medical data: a systematic review. *Pharmacoepidemiology and drug safety*. 2012;21(1):141-147.
Fan J, Arruda-Olson AM, Leibson CL, et al. Billing code algorithms to identify cases of peripheral artery disease from administrative data. *Journal of the American Medical Informatics Association: JAMIA*. 2013;20(e2):e349-354.