Comparative Effectiveness of Direct Oral Anticoagulants Versus Warfarin in Atrial Fibrillation Using An Instrumental Variable Approach

Nicolae Done, Donglin Li, Adam B. Woolley, Julia C. Prentice

October 18, 2017

CAPER

Center for Access Policy, Evaluation and Research

QUERI
VA Quality Enhancement Research Initiative

Veterans Health Administration
Research & Development
Improving Veterans' Lives www.research.va.gov

VA HEALTH CARE
Defining EXCELLENCE in the 21st Century

Atrial Fibrillation and Stroke

- Atrial Fibrillation (AF) is the **second most common** cardiovascular condition

- AF is associated with a **3- to 5-fold increase in risk of stroke** from cardioembolism

- Oral Anticoagulants (OACs) **drastically reduce the risk of stroke** and are recommended in all AF patients with CHA_2DS_2-VASc score>1 (indicating high risk)

Direct Oral Anticoagulants (DOACs) Approved in 2009 and Entered the VA Formulary in 2011

Warfarin

Available since 1954

- Reduces stroke risk by 60-70% (Hart et al. 2002)
- Numerous food and drug interactions
- Frequent monitoring and dose adjustment
- Increased risk of bleeding, but easier to reverse

DOACs

Dabigatran, Rivaroxaban, Apixaban, Edoxaban

- Non-vitamin K antagonists
- Non-inferior efficacy in large phase III trials
- More expensive
- Reductions concentrated in hemorrhagic stroke

Study Objective

Evaluate the comparative **effectiveness** and **safety** of DOACs using real-world evidence from the VA's large administrative database

Rationale

Efficacy and safety demonstrated in clinical trial settings may not translate to routine practice

- Differences in patient populations

  (VA patients tend to have higher morbidity)
- Patient adherence may be substantially lower (mental health conditions and substance abuse may exacerbate this problem)
- Differences in intensity of follow-up
- Appropriate dosing may be difficult to achieve
- Other variations in care provision

Poll Question #1

What is your professional role?

- VA Research Investigator/Data Manager/Analyst
- VA Project Manager/Coordinator/Assistant
- VA Program Office or Operations Staff
- Clinician
- Non-VA researcher
- Non VA (other)

Previous Studies

	Yao et al. (2016)	Lauffenburger et al. (2015)	Graham et al. (2015)	Villines et al. (2015)	Maura et al. (2015)
Patients	US Privately insured	US Privately insured and Medicare	US Medicare	US Department of Defense	French patients
DOACs studied	DAB, RIV, API	DAB	DAB	DAB	DAB, RIV
Study design	Propensity-score matching	Propensity-score weighting	Propensity-score matching	None	Propensity-score matching
	Results: drug favored				
Stroke/ embolism	**API** DAB, RIV=WAR	**DOAC**	**DOAC**	**DOAC**	No difference
Major Hemorrhage	**API, DAB** RIV=WAR	**DOAC** (hem. stroke)	**DOAC**	**DOAC**	No difference
AMI		**DOAC**	No difference	**DOAC**	
Death	–	–	**DOAC**	**DOAC**	No difference

Real-World Efficacy Studies Face a Selection Bias Concern

- Patients on Warfarin ≠ Patients on DOACs

- Failure to control for unobserved characteristics associated with treatment selection and/or outcomes will bias the effect estimates

Instrumental Variable (IV) Methods Can Address Selection Bias

- The key idea is to find a **plausibly exogenous** source of variation in treatment and use it as an "instrument"
- Most obvious source is a **coin flip** – randomly allocates treatment and controls, usually with 50%/50% probability
- In our case – variation in provider prescribing patterns are **quasi-random**, as they depend on local practices and provider preferences (Prentice *et al.* 2014)

Empirical Approach

2-Stage Residual Inclusion (2SRI) technique (Terza *et al.* 2008)

1. **First stage:**

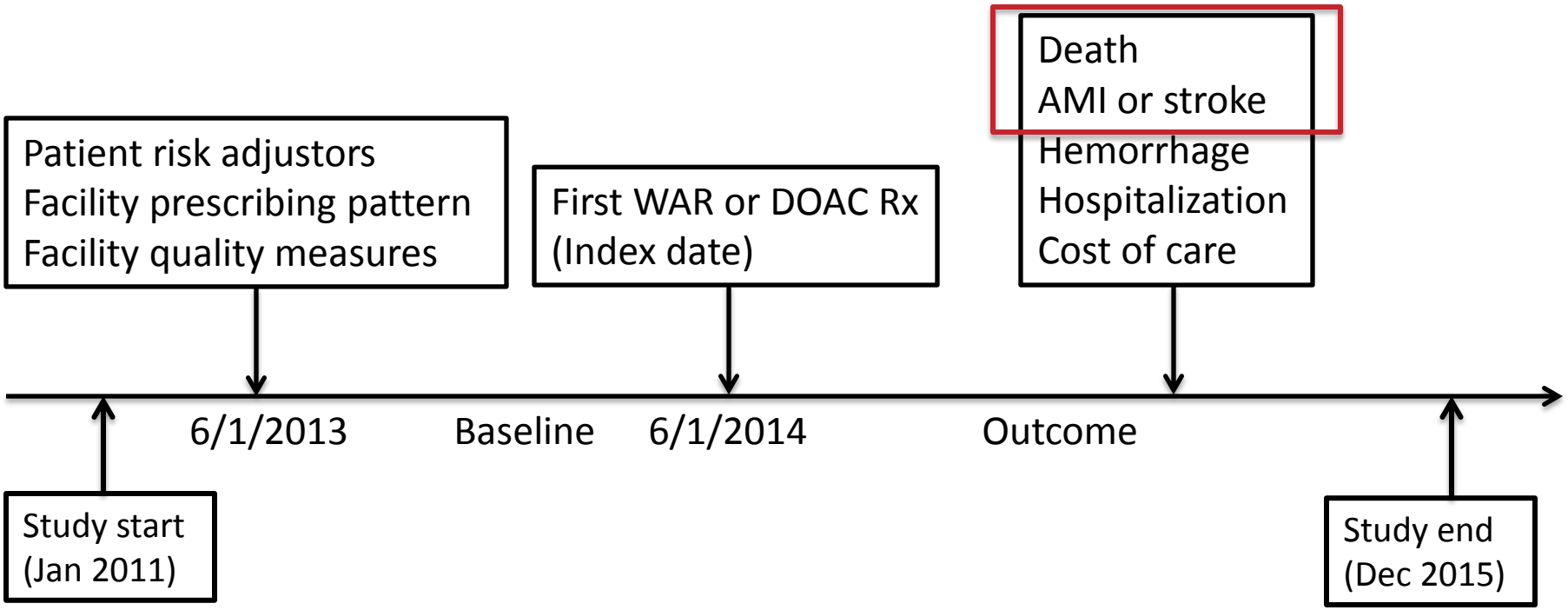
$$DOAC(= P\ \ Facility\ Rx(history + (demographics\ + ((risk + (:provider(quality + (travel(time + ((station + year(r) + \epsilon($$

2. **Second stage:**

$$Outcome(=(F(DOAC(+(\epsilon(+(demographics(+(risk(+(provider(quality : + travel\ time(+(station +((year)$$

- Cox proportional hazards model
- ϵ(controls for unobserved confounders during treatment selection
- Individuals are censored at study end /occurrence of first outcome

Example - Study Design and Timing of Measurements



Data

- VA Corporate Data Warehouse (CDW) (2011-2015)

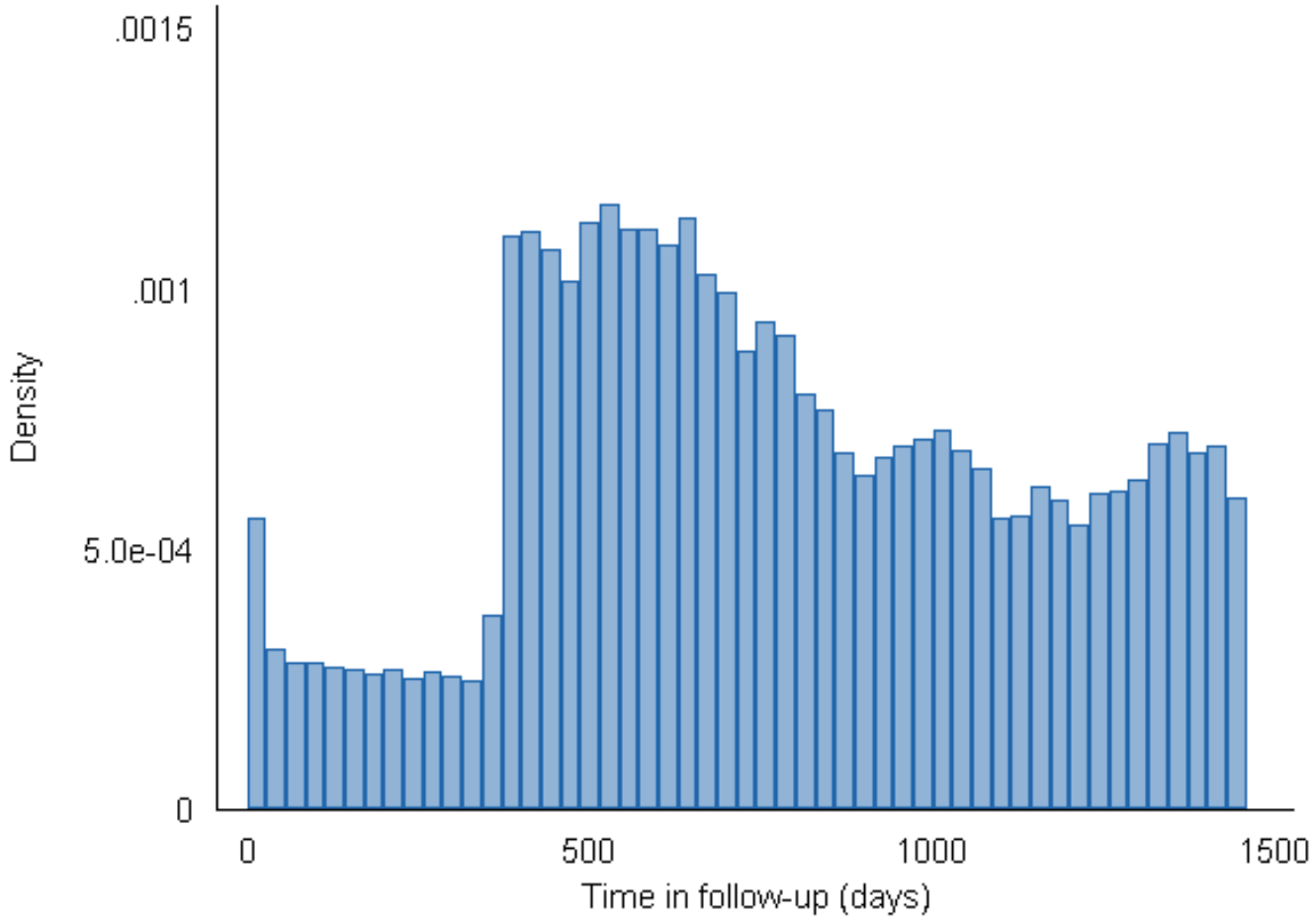
- Medicare Claims (2011-2015)
    - Inpatient and outpatient
    - Part D Rx information
    - Vital statistics file

  Important to observe all outcomes in acute-care settings, death, and Dx in baseline

Patients Data Selection

Sample	N	% prior	% initial
All patients aged 60+ with a CDW Dx of AF in 2012-2014	204,892	N/A	100.0%
Who were prescribed an OAC in 2012-2014	61,050	29.8%	29.8%
Excluding patients in Medicare Advantage in 2011-2015	40,358	66.1%	19.7%
Excluding patients with post-acute care stays in 2011-2015	39,732	98.4%	19.4%
Excluding patients under 66 at date of index prescription	**35,478**	89.3%	17.3%

Time in Follow-Up



Baseline Demographic Characteristics

	Mean or Percent	SD	Min	Max
Age	75.7	7.3	66	100
Male	98.6%			
Race				
White	90.2%			
Black	7.7%			
Other race	2.1%			
Distance to nearest VA (miles)	13.2	13.8	0.1	624.9

Baseline Clinical Characteristics

	Mean or Percent	SD	Min	Max
Number of Elixhauser comorbidities	6.4	3.1	0	22
Congestive Heart Failure	41.9%			
Valvular Disease	28.6%			
Renal Failure	26.4%			
Liver Disease	5.0%			
Depression	24.0%			
Alcohol Abuse	6.8%			
Body Mass Index	30.3	6.3	9.3	84.1
CHA2DS2-VASc score	4.4	1.7	1	9
HAS-BLED score	2.2	1.1	0	7
Average BP<140/90	71.0%			
Average LDL-C<100	59.0%			
Provider quality				
Facility-level BP<140/90	76.4%	3.8	62.8	88.1
Facility-level LDL-C<100	75.2%	7.9	33.3	100.0
Facility-level HbA1c poor control	18.6%	3.2	8.5	29.5

Index Drugs Prescribed

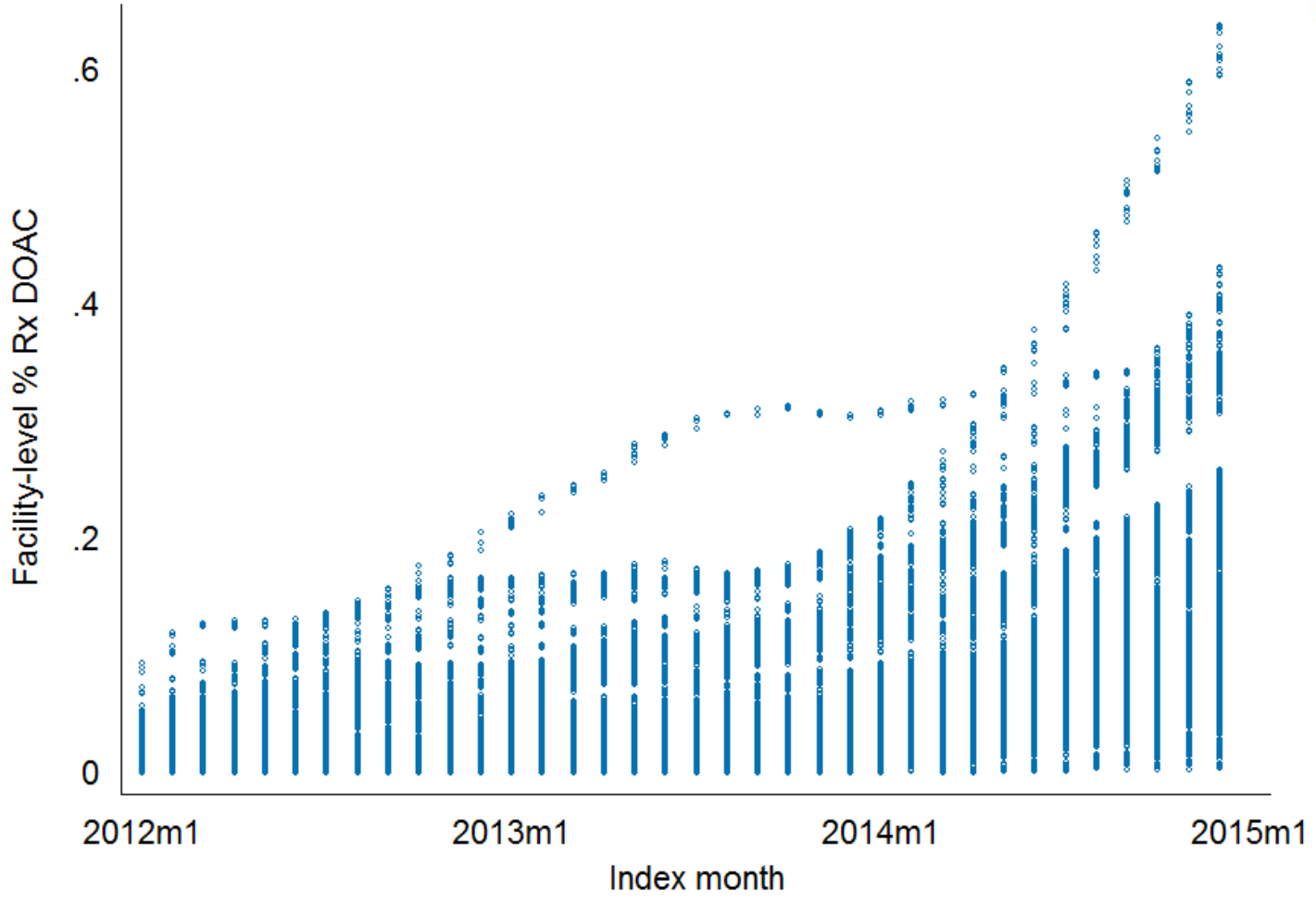
	Percent	SD	Min	Max
Assignment – Index drug prescribed				
Warfarin	80.5%			
DOAC	19.5%			
Dabigatran	11.2%			
Rivaroxaban	5.7%			
Apixaban	2.7%			
Instrument – Facility-level % Rx DOAC in baseline	6.6%	7.2%	0.0%	63.7%

Poll Question #2

What is your experience with Instrumental Variables methods?

- Not heard about IV methods at all
- I have learned about them in a class/seminar/conference but not used them
- I have used them in my research, but only linear models
- I have used them extensively, including 2-Stage Residual Inclusion

Instrument Variation Over Time



Covariate Balance by Index Drug and By Instrument Quantiles

	Index drug		
	WAR	**DOAC**	**Std. Diff.**
Age	75.7	76.5	-0.11
Male	98.6%	98.5%	0.01
White	89.1%	93.2%	-0.15
Black	8.5%	4.8%	0.15
Other race	2.4%	2.0%	0.03
Distance to nearest VA	13.2	14.4	-0.08
Number of comorbidities	6.4	6.1	0.10
Congestive Heart Failure	42.1%	39.1%	0.06
Valvular Disease	27.7%	31.1%	-0.08
Renal Failure	27.4%	21.0%	0.15
Liver Disease	5.1%	4.6%	0.02
Depression	24.2%	21.1%	0.07
Alcohol Abuse	6.8%	5.0%	0.08
Body Mass Index	30.3	30.2	0.02
CHA2DS2-VASc score	4.4	4.4	0.02
HAS-BLED score	2.2	2.1	0.09
Average BP <140/90	67.7%	68.7%	-0.02
Average LDL-C<100	56.4%	56.0%	0.01
Observations	**28,354**	**7,124**	

Covariate Balance by Index Drug and By Instrument Quantiles

	Index drug			Facility DOAC proportion, quantile		
	WAR	**DOAC**	**Std. Diff.**	**Below median**	**Above median**	**Std. Diff.**
Age	75.7	76.5	-0.11	75.8	75.9	-0.01
Male	98.6%	98.5%	0.01	98.6%	98.6%	0.00
White	89.1%	93.2%	-0.15	90.1%	89.8%	0.01
Black	8.5%	4.8%	0.15	7.6%	7.8%	-0.01
Other race	2.4%	2.0%	0.03	2.2%	2.4%	-0.01
Distance to nearest VA	13.2	14.4	-0.08	13.3	13.5	-0.01
Number of comorbidities	6.4	6.1	0.10	6.3	6.3	0.01
Congestive Heart Failure	42.1%	39.1%	0.06	42.2%	40.7%	0.03
Valvular Disease	27.7%	31.1%	-0.08	27.4%	29.4%	-0.05
Renal Failure	27.4%	21.0%	0.15	26.3%	26.0%	0.01
Liver Disease	5.1%	4.6%	0.02	4.7%	5.4%	-0.03
Depression	24.2%	21.1%	0.07	23.5%	23.6%	0.00
Alcohol Abuse	6.8%	5.0%	0.08	6.1%	6.7%	-0.02
Body Mass Index	30.3	30.2	0.02	30.3	30.3	0.00
CHA2DS2-VASc score	4.4	4.4	0.02	4.5	4.4	0.06
HAS-BLED score	2.2	2.1	0.09	2.2	2.2	0.03
Average BP <140/90	67.7%	68.7%	-0.02	67.8%	68.1%	0.00
Average LDL-C<100	56.4%	56.0%	0.01	57.2%	55.4%	0.04
Observations	**28,354**	**7,124**		**17,740**	**17,738**	

The Instrument Affects Treatment Assignment

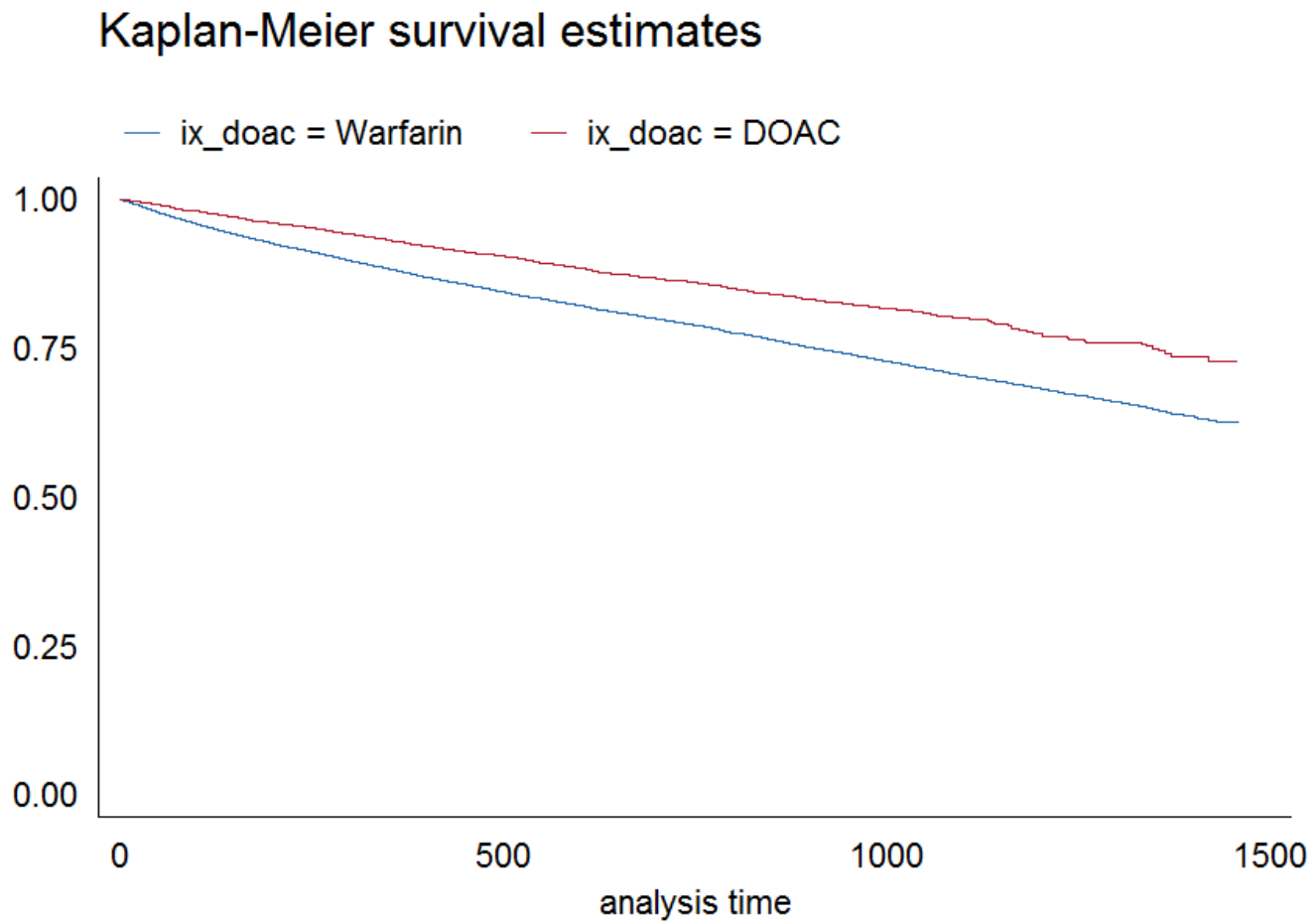
Quartile of Facility DOAC Proportion	Percent assigned to DOAC
1	4.5%
2	14.0%
3	23.5%
4	38.3%
Average	**20.1%**

First Stage Logit – IV Effect on Probability of Index DOAC

Explanatory variables	Odds ratio	95% CI	P<|t|
Instrument - Facility-level % Rx DOAC	6.044***	2.648 - 13.79	0.000
Age	1.011***	1.005 - 1.016	0.000
Male	1.000	0.784 - 1.276	0.998
White	2.113***	1.843 - 2.423	0.000
Other race	1.653***	1.294 - 2.112	0.000
Distance to nearest VA missing	1.039	0.883 - 1.222	0.648
Distance to nearest VA (miles)	1.006***	1.003 - 1.008	0.000
Congestive Heart Failure	0.931**	0.868 - 0.999	0.047
Valvular Disease	1.148***	1.073 - 1.228	0.000
Renal Failure	0.639***	0.574 - 0.713	0.000
Depression	0.899***	0.835 - 0.967	0.0038
Overweight	1.153***	1.054 - 1.263	0.002
Obese	1.106**	1.003 - 1.220	0.044
BMI missing	1.535***	1.362 - 1.729	0.000
Average BP<140/90	1.097***	1.028 - 1.170	0.005
Average LDL-C<100	1.062*	1.000 - 1.128	0.050
Facility-level BP<140/90	0.996	0.982 - 1.011	0.630
Facility-level LDL-C<100	1.036***	1.029 - 1.043	0.000
Facility-level HbA1c poor control	1.016**	1.001 - 1.032	0.039

*** p<0.01, ** p<0.05, * p<0.1

All-Cause Mortality – Survival Curves by Index Drug



All-Cause Mortality – Log-Rank Test for Equality of Survivor Functions

Index Drug	Events observed	Events expected
Warfarin	6721	6274
DOAC	926	1373
Total	7647	7647

chi2(1) = 178.80

Pr>chi2 = 0.0000

Regressions Models

2SRI model

1. **First stage:**

$$DOAC = P\ \ Facility\ Rx + history + demographics + risk + provider + quality + travel + time + station + year + \epsilon$$

2. **Second stage:**

$$Outcome = F(DOAC + \epsilon + demographics + risk + provider + quality + travel + time + station + year)$$

Naïve model

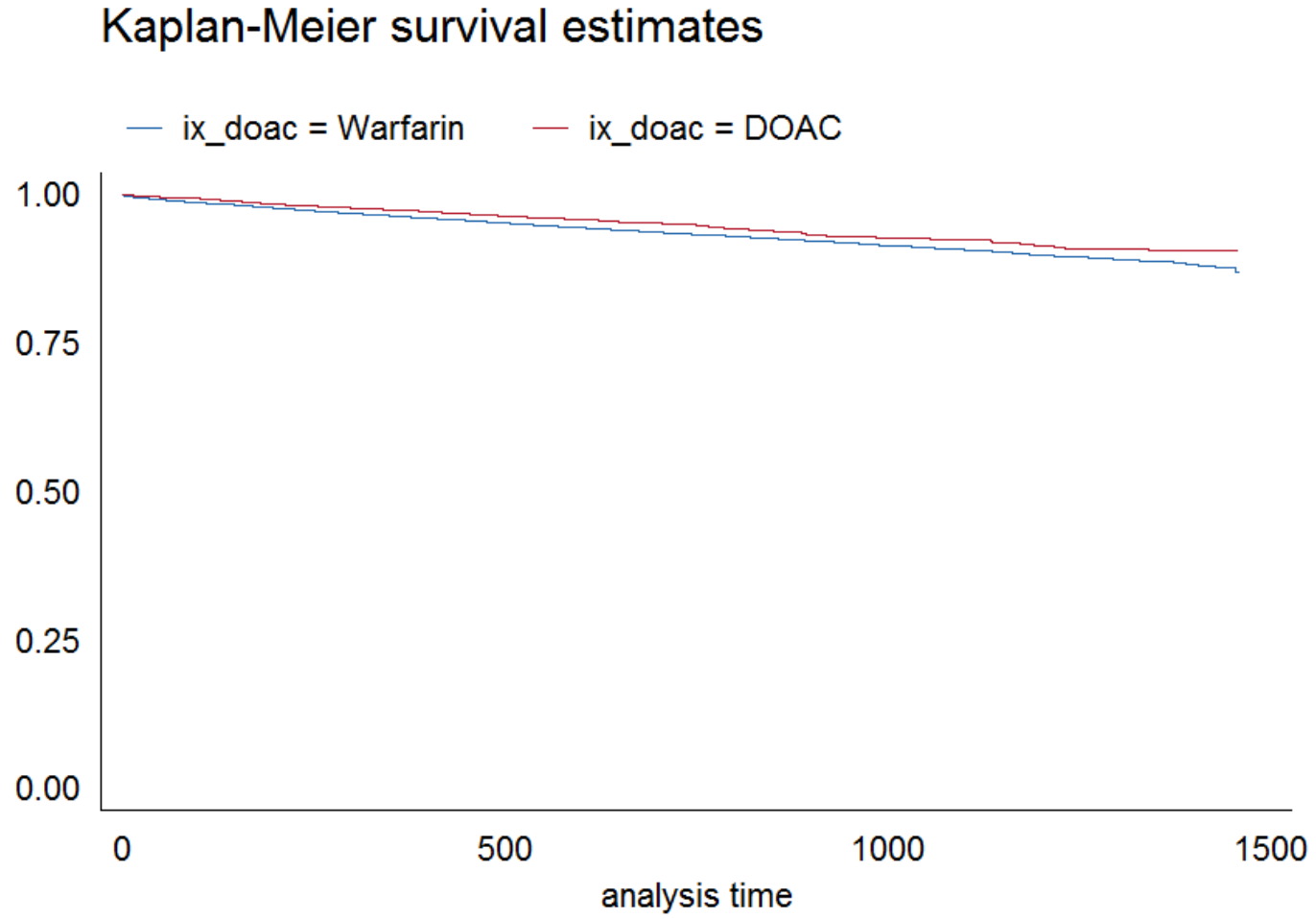
$$Outcome = F(DOAC + demographics + risk + provider + quality + travel + time + station + year)$$

IV and Naïve Estimates – Effects on All-Cause Mortality

	2SRI Estimate			Naïve Estimate		
Explanatory variables	**Hazard ratio**	**95% CI**	**P<|t|**	**Hazard ratio**	**95% CI**	**P<|t|**
DOAC	0.343***	0.223 - 0.527	0.000	0.661***	0.614 - 0.711	0.000
Stage I residuals	1.965***	1.271 - 3.039	0.002			

*** p<0.01, ** p<0.05, * p<0.1. Models also control for patient demographics and clinical risk variables, provider quality measures, and provider and year fixed effects.

Stroke or AMI – Survival Curves by Index Drug



Stroke/AMI – Log-Rank Test for Equality of Survivor Functions

Index Drug	Events observed	Events expected
Warfarin	1932	1855
DOAC	333	410
Total	2265	2265

chi2(1) = 17.83

Pr>chi2 = 0.00001

IV and Naïve Estimates – Effects on Stroke/AMI

	2SRI Estimate			Naïve Estimate		
Explanatory variables	**Hazard ratio**	**95% CI**	**P<|t|**	**Hazard ratio**	**95% CI**	**P<|t|**
DOAC	0.573	0.271 - 1.209	0.144	0.883*	0.777 - 1.003	0.0565
Stage I residuals	1.561	0.730 - 3.339	0.251			

* p<0.1. Models also control for patient demographics and clinical risk variables, provider quality measures, and provider and year fixed effects.

Falsification Test – CAD Cohort

- If our instrument is valid, then it should affect risk of an outcome **only** by affecting the treatment assignment (Pizer 2016)

- Coronary Artery Disease (CAD) patients – at higher **risk of stroke** but are **not prescribed OACs** (so no treatment assignment)

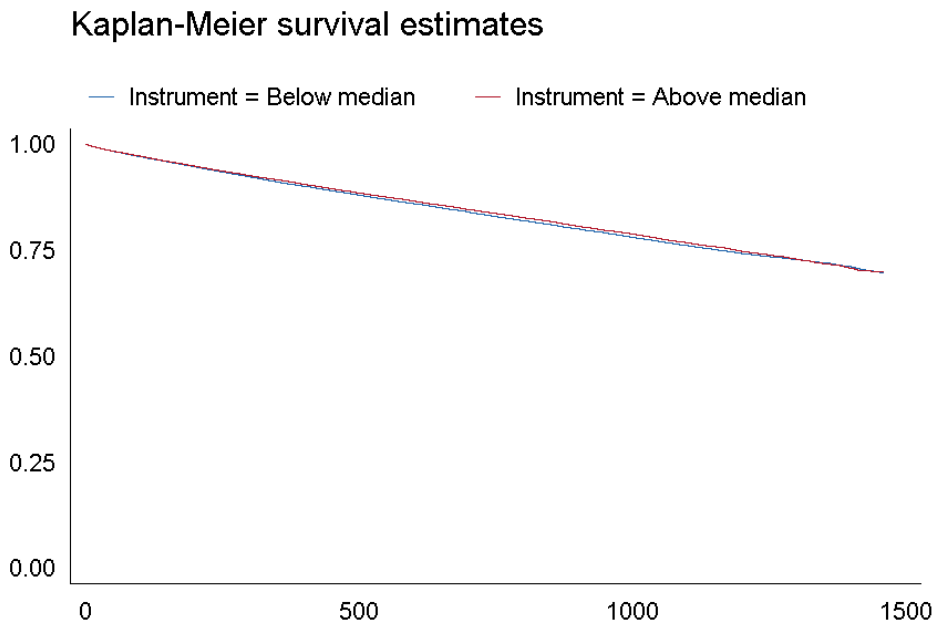
- Cox proportional hazards models using the instrument as an explanatory variable should **not predict outcomes**

Falsification Test – All-Cause Mortality in the CAD Cohort

Facility DOAC Proportion	Events observed	Events expected
Below median	13427	13240
Above Median	11365	11552
Total	24792	24792

chi2(1) = 5.74

Pr>chi2 = 0.0166

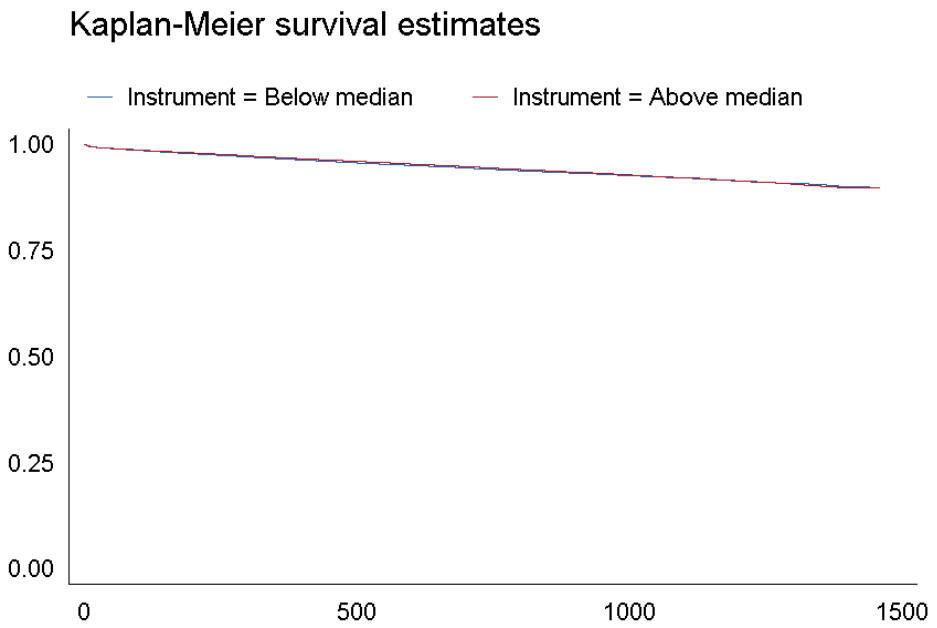


Falsification Test – Stroke and AMI in the CAD Cohort

Facility DOAC Proportion	Events observed	Events expected
Below median	4268	4206
Above Median	3670	3732
Total	7938	7938

chi2(1) = 1.93

Pr>chi2 = 0.1649



Falsification Test – Adjusted Estimates

	Mortality			Stroke or AMI		
Explanatory variables	**Hazard ratio**	**95% CI**	**P<|t|**	**Hazard ratio**	**95% CI**	**P<|t|**
Facility DOAC proportion	0.930	0.746 - 1.159	0.518	0.922	0.633 - 1.341	0.669
N	130,404			130,404		

Models also control for patient demographics and clinical risk variables, provider quality measures, and provider and year fixed effects.

Conclusions

- After adjusting for unobserved confounding, we find that DOACs reduce the risk of death by ~66% compared to Warfarin
  - Larger reduction than in other studies
    - Graham et al. 2015: HR = 0.86 (0.77-0.96)
    - Villines et al. 2015: HR = 0.64 (0.55-0.74)

- DOACs also reduce the risk of stroke or AMI by ~43%, but this effect is not statistically significant at our level of precision
  - Also larger reduction compared to other studies

Next Steps

- Incorporate 2016 Medicare claims data → increase sample size and follow-up time
- Compare DOACs to Warfarin individually
- Add measures of patient drug adherence
- Analyze effect on incidence of hospital stays and hemorrhage
- Calculate total cost of care → cost-effectiveness analysis
- Compare with propensity score matching

Selected Limitations

- Medicare data lag leads to **suboptimal sample size**
- Significant **missing data** for some measures (e.g., BMI, ZIP code)
- **Intent-to-treat** analysis (about 10% of patients switch drugs after initial assignment)
- **Unobserved quality** dimensions may still be an issue
  - But any unobserved measures would have to be highly correlated with DOAC prescribing and uncorrelated with our measured quality indicators
- IV estimates **Local Average Treatment Effect** (LATE)
  - In the presence of "essential heterogeneity", even 2SRI methods could lead to LATE that is significantly different from the population Average Treatment Effect (ATE) (Chapman and Brooks, 2016)
- Findings in a sample of Veterans **may not generalize** to other populations

Acknowledgments and Author Affiliations

- **Nicolae Done, PhD - Boston University School of Medicine and CAPER**
- Donglin Li, MS - CAPER
- Adam B. Woolley, PharmD, MEd, BCPS, RPh – VA Boston Healthcare System
- Julia C. Prentice, Boston University School of Medicine, School of Public Health, and CAPER

- Funded by grant number **IIR 15-139** from the Health Services Research and Development Service of the Department of Veterans Affairs
- All opinions expressed are those of the authors and do not necessarily reflect the official position of the U.S. Department of Veterans Affairs, Boston University

- **Correspondence**: nicolae.done@va.gov

References

- Hart, R. G., Benavente, O., McBride, R., & Pearce, L. A. (1999). Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. *Annals Of Internal Medicine*, *131*(7), 492–501.
- Graham, D. J., Reichman, M. E., Wernecke, M., Zhang, R., Southworth, M. R., Levenson, M., … Kelman, J. A. (2015). Cardiovascular, Bleeding, and Mortality Risks in Elderly Medicare Patients Treated With Dabigatran or Warfarin for Nonvalvular Atrial Fibrillation. *Circulation*, *131*(2), 157–164. https://doi.org/10.1161/CIRCULATIONAHA.114.012061
- Lauffenburger, J. C., Farley, J. F., Gehi, A. K., Rhoney, D. H., Brookhart, M. A., & Fang, G. (2015). Effectiveness and safety of dabigatran and warfarin in real-world US patients with non-valvular atrial fibrillation: a retrospective cohort study. *Journal of the American Heart Association*, *4*(4), e001798.
- Maura, G., Blotière, P.-O., Bouillon, K., Billionnet, C., Ricordeau, P., Alla, F., & Zureik, M. (2015). Comparison of the Short-Term Risk of Bleeding and Arterial Thromboembolic Events in Nonvalvular Atrial Fibrillation Patients Newly Treated With Dabigatran or Rivaroxaban Versus Vitamin K Antagonists: A French Nationwide Propensity-Matched Cohort Study. *Circulation*, *132*(13), 1252–1260. https://doi.org/10.1161/CIRCULATIONAHA.115.015710
- Villines, T. C., Schnee, J., Fraeman, K., Siu, K., Reynolds, M. W., Collins, J., & Schwartzman, E. (2015). A comparison of the safety and effectiveness of dabigatran and warfarin in non-valvular atrial fibrillation patients in a large healthcare system. *Thrombosis and Haemostasis*, *114*(6), 1290–1298. https://doi.org/10.1160/TH15-06-0453
- Yao, X., Abraham, N. S., Sangaralingham, L. R., Bellolio, M. F., McBane, R. D., Shah, N. D., & Noseworthy, P. A. (2016). Effectiveness and Safety of Dabigatran, Rivaroxaban, and Apixaban Versus Warfarin in Nonvalvular Atrial Fibrillation. *Journal of the American Heart Association*, *5*(6), e003725. https://doi.org/10.1161/JAHA.116.003725
- Prentice, J. C., Conlin, P. R., Gellad, W. F., Edelman, D., Lee, T. A., & Pizer, S. D. (2014). Capitalizing on Prescribing Pattern Variation to Compare Medications for Type 2 Diabetes. Value in Health, 17(8), 854–862. https://doi.org/10.1016/j.jval.2014.08.2674Elixhauser, A., Steiner, C., Harris, D. R., & Coffey, R. M. (1998). Comorbidity measures for use with administrative data. *Medical Care*, *36*, 8–27.
- Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531–543. https://doi.org/10.1016/j.jhealeco.2007.09.009
- Pizer, S. D. (2016). Falsification Testing of Instrumental Variables Methods for Comparative Effectiveness Research. *Health Services Research*, *51*(2), 790–811. https://doi.org/10.1111/1475-6773.12355
- Chapman, C. G., & Brooks, J. M. (2016). Treatment Effect Estimation Using Nonlinear Two-Stage Instrumental Variable Estimators: Another Cautionary Note. *Health Services Research*, 51(6), 2375–2394. https://doi.org/10.1111/1475-6773.12463