PLOS ONE

RESEARCH ARTICLE

Utility of automated data-adaptive propensity score method for confounding by indication in comparative effectiveness study in real world Medicare and registry data

Hiraku Kumamaru[1]*, Jessica J. Jalbert[2], Louis L. Nguyen[3], Lauren A. Williams[4], Hiroaki Miyata[1], Soko Setoguchi[5]

1 Department of Healthcare Quality Assessment, Graduate School of Medicine, University of Tokyo, Tokyo, Japan, **2** Health Economics and Outcomes Research, Regeneron Pharmaceuticals, Inc, Tarrytown, New York, United States of America, **3** Division of Vascular and Endovascular Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **4** Philadelphia College of Osteopathic Medicine, Philadelphia, Pennsylvania, United States of America, **5** Department of Medicine, Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey, United States of America

* hik205@mail.harvard.edu

Check for updates

OPEN ACCESS

Citation: Kumamaru H, Jalbert JJ, Nguyen LL, Williams LA, Miyata H, Setoguchi S (2022) Utility of automated data-adaptive propensity score method for confounding by indication in comparative effectiveness study in real world Medicare and registry data. PLoS ONE 17(8): e0272975. https://doi.org/10.1371/journal.pone.0272975

Editor: Yoshihiro Fukumoto, Kurume University School of Medicine, JAPAN

Received: July 13, 2021

Accepted: July 31, 2022

Published: August 15, 2022

Copyright: © 2022 Kumamaru et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data used for the study are available to researchers through data use agreement with CMS as well as the Society for Vascular Surgery's Vascular Quality Initiative (SVS-VQI). Information on the data access and applications can be found at: CMS: https://www.cms.gov/Research-Statistics-Data-and-Systems/CMS-Information-Technology/AccesstoDataApplication SVS-VR (VQI): https://www.vqi.org/data-analysis/.

Abstract

Background

Confounding by indication is a serious threat to comparative studies using real world data. We assessed the utility of automated data-adaptive analytic approach for confounding adjustment when both claims and clinical registry data are available.

Methods

We used a comparative study example of carotid artery stenting (CAS) vs. carotid endarterectomy (CEA) in 2005–2008 when CAS was only indicated for patients with high surgical risk. We included Medicare beneficiaries linked to the Society for Vascular Surgery's Vascular Registry >65 years old undergoing CAS/CEA. We compared hazard ratios (HRs) for death while adjusting for confounding by combining various 1) Propensity score (PS) modeling strategies (investigator-specified [IS-PS] vs. automated data-adaptive [ada-PS]); 2) data sources (claims-only, registry-only and claims-plus-registry); and 3) PS adjustment approaches (matching vs. quintiles-adjustment with/without trimming). An HR of 1.0 was used as a benchmark effect estimate based on CREST trial.

Results

The cohort included 1,999 CAS and 3,255 CEA patients (mean age 76). CAS patients were more likely symptomatic and at high surgical risk, and experienced higher mortality (crude HR = 1.82 for CAS vs. CEA). HRs from PS-quintile adjustment without trimming were 1.48 and 1.52 for claims-only IS-PS and ada-PS, 1.51 and 1.42 for registry-only IS-PS and ada-PS, and 1.34 and 1.23 for claims-plus-registry IS-PS and ada-PS, respectively. Estimates from other PS adjustment approaches showed similar patterns.

Funding: This project is funded by contract HHSA290-2005-0016-I TO3 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services (DHHS) as part of the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) program; and by IAA Contract 500-2010-00001I TO6 and CEA Contract 500-2010-00001I TO2 from the Centers for Medicare & Medicaid Services (CMS), U.S. DHHS. The funding agency had no role in the design and conduct of the study or in the collection, analysis, and interpretation of the data.

Competing interests: HK and HM are affiliated with the department of Healthcare Quality Assessment, University of Tokyo, which is a social collaboration department supported by the National Clinical Database, Johnson & Johnson KK, and Nipro Corporation. HK has received lecture fees from Chugai Pharmaceutical Co., Ltd., Johnson & Johnson KK, and consultancy fees from Mitsubishi Tanabe Pharm and EP Croit Co., Ltd. JJJ is a full-time employee of Regeneron Pharmaceuticals. This study was conducted prior to her employment at the company. SS has served as a member of the US FDA advisory committee and as a consultant for Pfizer Inc., Medtronic, and Merck. SS has also received research funding from the National Institutes of Health, the Cystic Fibrosis Foundation, Pfizer Inc., and Bristol-Myers Squibb. All other authors have no conflicts to declare. The above do not alter our adherence to Plos One policies on sharing data and materials.

Conclusions

In a comparative effectiveness study of CAS vs. CEA with strong confounding by indication, ada-PS performed better than IS-PS in general, but both claims and registry data were needed to adequately control for bias.

Introduction

Real-world data (RWD) including large electronic databases such as claims databases, electronic health records databases, and patient/procedure registries are rich sources of data for conducting comparative effectiveness and safety studies. While these studies may be used to complement findings from randomized clinical trials for cardiovascular disease, [1–6] confounding is one of the biggest challenges to study validity. Confounding by indication is particularly problematic when comparing the effectiveness of two or more alternative treatment modalities because each modality may be indicated for clinically distinct patient groups. For example, carotid artery stenting (CAS) was originally indicated only for patients with high surgical risk when its use was first approved for reimbursement by CMS in 2004 [7]. It was not until 2011 when a guideline recommended CAS for non-high risk patients based on later trials such as CREST proving the efficacy in the non-high risk population [8, 9]. Thus, comparative effectiveness studies of CAS vs. carotid endarterectomy (CEA) using early phase data would likely be confounded by indication.

Most claims databases lack detailed clinical information to fully account for confounding by indication, and adding clinical information through data linkage with registries has been shown to greatly improve the validity of comparative studies in patients with cardiovascular diseases [2, 3]. At the same time, registries may not be available or linkage may not be possible. Various methodological developments such as propensity score (PS)-based data-adaptive methods with automated variable generation and selection algorithms have been proposed to improve the validity of results using RWD [10–14]. However, in the presence of linked clinical information from a registry, the relative utility of such approaches vs. additional clinical information in CER studies with strong confounding by indication has not been fully elucidated.

The aim of the present study was to evaluate the performance of the automated data-adaptive PS approaches for confounding adjustment in a CER study with strong confounding by indication when claims data only, registry data only, and linked claims plus registry data are available. We used the example of CAS vs. CEA using data from 2005–2008 when CMS coverage indication for CAS were limited to high surgical risk patients with predefined levels of stenosis severity and symptoms [15].

Methods

Data sources

We used the Society for Vascular Surgery's Vascular Registry (SVS-VR), 2005–2008 as well as the denominator, institutional, non-institutional, and vital status Medicare files (2000–2009). The design and protocols of SVS-VR have been described in detail elsewhere [16]. The SVS-VR collects detailed information on medical history, pre-procedural diagnostics (including degree of carotid stenosis, symptomatic status, and high surgical risk status), procedure-related factors, and intra-operative and pre-discharge complications for patients undergoing CAS as well as CEA, allowing for comparison of the two procedures. The SVS-VR data are

audited to ensure that all cases of CAS and CEA are reported and to verify the accuracy and completeness of the data [16].

From CMS, we obtained administrative claims data from January 1st, 2000 to December 31st, 2009 for patients who underwent CAS (ICD-9-CM codes: 00.61, 00.63, 00.64) or CEA (ICD-9-CM codes: 38.12) in an inpatient setting, including information on patient demographics, eligibility, inpatient and outpatient facility services use, physician services use covered under Part B, and date of death.

Linkage of SVS-VR to the Medicare data has been previously described [17]. Briefly, we deterministically linked CAS and CEA procedures in the registry to the Medicare institutional file using date of birth, sex, facility, and the requirement that the procedure date in the registry be between a hospital admission and discharge date for the same procedure in the Medicare data.

Patients

The patient population included in this study has been described [3]. In short, the study sample included Medicare fee-for-service beneficiaries who had undergone CAS or CEA at SVS-VR-participating facilities between 2005 and 2008, who were at least 66 years of age at the time of the procedure, were eligible for Medicare for at least one year prior to the procedure, and for whom Medicare was the primary payer. If patients underwent multiple procedures, the first procedure was selected. We set the day of the procedure as the index date and the preceding 1-year period as the baseline period. We followed patients from the index date until the earliest occurrence of one of the following events: death from any cause, loss of Medicare eligibility, or the end of the study period (December 31st, 2009).

Outcome

The outcome of interest was death from any cause following carotid revascularization over the study period.

Study variables

Medicare claims data. We extracted data on patient age, sex, and race from Medicare denominator files. We also derived measures of healthcare utilization, such as number of past-year hospitalizations, physician visits, and nursing home admissions [18]. Patients were categorized as having undergone elective CAS unless the hospital admission type was defined as urgent, emergent, or of a traumatic nature in the institutional Medicare file. Patients were categorized as having symptomatic carotid stenosis if they had recorded diagnoses of stroke, transient ischemic attack, or amaurosis fugax in the year leading up the procedure. We also defined comorbidities using the diagnoses and procedure codes recorded during the baseline period.

SVS-VR. Information on the ipsilateral and contralateral degrees of carotid stenosis was available from interpreted pre-procedural carotid ultrasound and angiogram exam results. We categorized patients as having mild (<50%), moderate (50–69%), or severe (≥70%) ipsilateral carotid stenosis or contralateral stenosis ≥70% using information from ultrasound results and, when ultrasound data were missing, from angiograms. We considered patients to be at high surgical risk if they met any of the following criteria: age >80, New York Heart Association (NYHA) class III or IV heart failure, left ventricular ejection fraction <30%, unstable angina, myocardial infarction in the past 30 days, recurrent stenosis, prior radical dissection or adiation, contralateral occlusion, contralateral laryngeal nerve palsy or injury, high anatomic lesion, or other physiologic or anatomic surgical risk factors from investigational device

exemption trials of CAS. We classified patients as per the National Coverage Determination's indications for CAS [15]: high surgical risk with symptomatic stenosis 50–69%, high surgical risk with symptomatic stenosis ≥70%, high surgical risk with asymptomatic carotid stenosis ≥80%, or those not matching the conditions above. We also obtained information on comorbidities from the registry.

Propensity score modeling

We employed: 1) investigator-specified approach and 2) automated data-adaptive approach. For investigator-specified PS model, the investigators preselected the list of variables based on clinical knowledge and prior literature. For automated data-adaptive model, we chose high dimensional (hd) PS, one of the most commonly used algorithms. Full lists of variables selected for each model are described in the footnote of Table 2.

The hdPS algorithm empirically generates and selects covariates that collectively act as proxies for the patient's health status from the database for inclusion in the propensity score modeling. The full description of this algorithm is available elsewhere [10, 19]. Briefly, the algorithm consists of three parts: (1) empirical variable identification; (2) variable ranking (prioritization); and (3) variable selection. First, the algorithm goes through each type of data such as inpatient diagnosis, inpatient procedure, namely 'data dimension' and generates covariates that represent the presence or absence of each code meeting a prespecified prevalence threshold (e.g. 1% or more). Second, the generated variables are ranked based on its potential to cause confounding, measured by its association with the outcome and the exposure [17]. Third, top 500 (or the number the investigators select) of these empirically selected variables are included to develop a PS model. In our study, 'data dimensions' considered for hdPS included inpatient, outpatient, non-institutional (carrier), and nursing home diagnosis and procedure codes. We also categorized variables in the registry into registry-based 'data dimensions' including medical history, symptomatic status, diagnostic imaging results, pre-procedural medications, and high-surgical risk status.

We constructed three propensity score models in each of the investigator-specified and hdPS approach using data dimensions in claims data only, with all SVS-VR variables only, and with claims data and SVS-VR variables (Table 2). We also included an additional PS model (4th model in hdPS approach) using both registry and claims data but excluding the information on patients' high surgical risk status to assess the importance of the high surgical risk status in confounding adjustment (Table 2).

Statistical analysis

We evaluated the baseline characteristics of the study cohort undergoing CAS or CEA using both registry and claims data. We report crude cumulative risks and 95% confidence intervals for 3-year mortality, derived from Kaplan-Meier estimators. We used multiple imputation to maximize power and handle missing covariate information from the SVS-VR (less than 5% of data were missing for all variables except contralateral carotid occlusion [5.4%], modified Rankin Scale score [5.5%], creatinine level [9.8%], NYHA class [17.6%], and hyperlipidemia/dyslipidemia [33.2%]).

Hazard for all cause death following CAS compared to CEA were derived using Cox proportional hazards regression models. The propensity score was included in the Cox model in three ways: using 1:1 nearest-neighbor matching with a caliper of 0.2 times the logit of the propensity score, and grouping patients into propensity score quintiles with and without implementing 5% asymmetric trimming [20]. We used the sandwich variance estimator to account for clustering of patients at the physician and hospital levels. The hazard ratios from these

models were compared to the benchmark risk ratio of 1.0 based on the findings reported in randomized clinical trials [9, 21]. All analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC). The study was approved by the Institutional Review Board of Brigham and Women's Hospital.

Results

We successfully linked 2867 SVS-VR CAS records and 4381 SVS-VR CEA records to Medicare claims data [3]. After applying inclusion and exclusion criteria, 1999 CAS patients and 3255 CEA patients were included in the study cohort (Fig 1).

The full characteristics of the population is described in the previous article by Jalbert and colleagues [3]. Briefly, the distribution of age, sex and race of the patients were similar between the two groups (Table 1). CAS patients were more likely to be symptomatic (CAS 38.5% vs. CEA 30.7%) and had a higher prevalence of high surgical risk factors (96.7% vs. 44.5%). Specifically, patients undergoing CAS had higher percentages of patients with class III/IV heart failure (10.7% vs. 3.2%), left ventricular ejection fraction $< 30\%$ (4.9% vs. 1.0%), unstable angina (3.6% vs. 0.5%), recurrent stenosis (27.4% vs. 1.8%), contralateral occlusion (16.7% vs. 7.1%), prior neck radiation (6.3% vs. 0.2%), and high anatomic lesion (6.4% vs. 1.8%) than patients undergoing CEA. CAS patients had a higher prevalence of cardiac comorbidities (e.g., coronary artery disease, 62.1% vs. 52.1%; myocardial infarction, 21.9% vs. 16.2%), and slightly

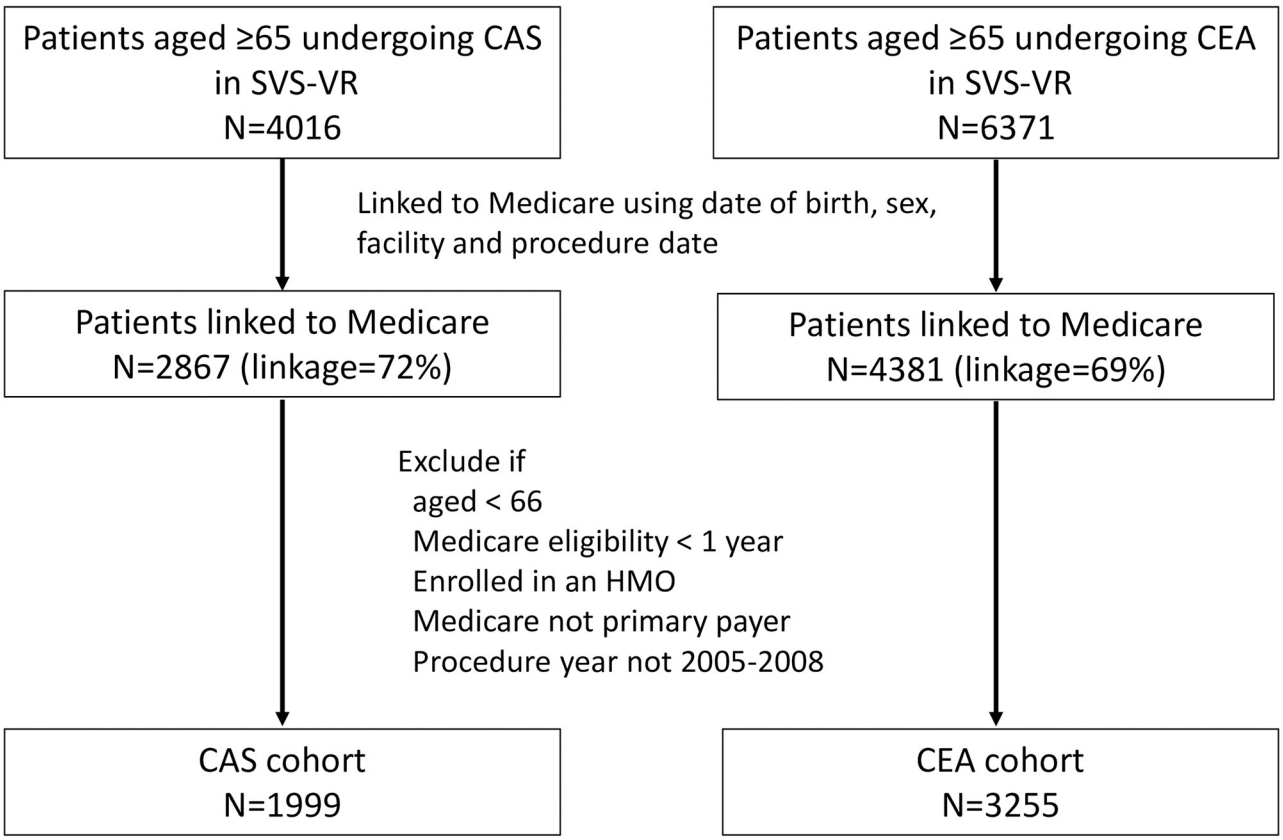

Fig 1. Flow chart for creation of Medicare-linked SVS-VR CAS and CEA study cohorts. Abbreviations: CAS = Carotid stenting; SVS-VR = Society of Vascular Surgeons Vascular Registry; CEA = Carotid endarterectomy; HMO = Health maintenance organization.

https://doi.org/10.1371/journal.pone.0272975.g001

Table 1. Selected patient characteristics at baseline, by procedure.

Patient Characteristics	CAS N (%)	CEA N (%)
N	1999	3255
Data from SVS-VR		
Age		
66–69	362 (18.1)	539 (16.6)
70–74	538 (26.9)	890 (27.3)
75–79	600 (30.0)	976 (30.0)
80-	499 (25.0)	850 (26.1)
Woman	814 (40.7)	1358 (41.7)
White Race	1866 (93.3)	3060 (94.0)
*Carotid Symptoms in Past 12 months***		
Symptomatic	770 (38.5)	1000 (30.7)
Ipsilateral Stroke	250 (12.5)	346 (10.6)
Ipsilateral Transient Ischemic Attack	479 (24.0)	555 (17.1)
Ipsilateral Amaurosis Fugax	171 (8.6)	226 (6.9)
*Carotid Stenosis Level**		
<50%	17 (0.9)	19 (0.6)
50–69%	79 (4.0)	206 (6.6)
≥70%	1899 (95.2)	2882 (92.8)
Contralateral stenosis ≥70%***	457 (24.5)	542 (17.5)
High-Surgical Risk Criteria		
High Surgical Risk	1934 (96.7)	1447 (44.5)
Age >80	520 (26.0)	832 (25.6)
NYHA Class III/IV Heart Failure	214 (10.7)	105 (3.2)
Left Ventricular Ejection Fraction <30%	97 (4.9)	33 (1.0)
Unstable angina	71 (3.6)	16 (0.5)
Recent MI (30 days)	25 (1.3)	8 (0.2)
Other Physiologic High-Surgical Risk from Investigational Device Exemption (IDE) CAS trials	662 (33.1)	269 (8.3)
Other Physiologic High-Surgical Risk	240 (12.0)	84 (2.6)
Recurrent Stenosis	548 (27.4)	57 (1.8)
Radical Neck Dissection	56 (2.8)	3 (0.1)
Contralateral Occlusion	334 (16.7)	231 (7.1)
Prior Neck Radiation	125 (6.3)	5 (0.2)
Contralateral Laryngeal Nerve Injury or Palsy	14 (0.7)	1 (0.0)
High Anatomic Lesion (C2 or higher)	128 (6.4)	60 (1.8)
Other Anatomatic High-Surgical Risk from Investigational Device Exemption (IDE) CAS trials	281 (14.1)	257(7.9)
Other Anatomic High-Surgical Risk	86 (4.3)	50 (1.5)
Cardiac Comorbidities		
Cardiac Arrhythmia	328 (16.4)	478 (14.7)
Coronary Artery Disease	1242 (62.1)	1695 (52.1)
Heart Failure	302 (15.1)	267 (8.2)
Myocardial Infarction	438 (21.9)	528 (16.2)

(*Continued*)

Table 1. (Continued)

Patient Characteristics	CAS N (%)	CEA N (%)
Other Comorbidities		
Cancer	351 (17.6)	457 (14.0)
Chronic Obstructive Pulmonary Disease	385 (19.3)	579 (17.8)
Diabetes Mellitus	645 (32.3)	944 (29.0)
Renal Failure	86 (4.3)	113 (3.5)
Data from claims		
Non-Elective Hospital Admission	478 (23.9)	432 (13.3)
Measures of Healthcare Utilization		
Mean Number of Past-Year Physician Visits (SD)	12.7 (8.8)	11.4 (8.3)
Median (IQR)	11 (7–17)	10 (6–15)
Mean Number of Past-Year Hospitalizations (SD)	1.0 (1.4)	0.6 (1.0)
Median (IQR)	1 (0–1)	0 (0–1)
Past-Year Nursing Home Admission	127 (6.4)	116 (3.6)

Abbreviations: SD = Standard Deviation; IQR = Interquartile Range

*Missing 2.9%;

**Missing <1%;

***Missing 5.4%;

† Missing 33.2%

https://doi.org/10.1371/journal.pone.0272975.t001

higher prevalence of non-cardiac comorbidities including cancer (17.6% vs. 14.0%), COPD (19.3% vs. 17.8%), and renal failure (4.3% vs. 3.5%).

CAS patients were followed for a median of 909 days, during which time 446 deaths were identified (incidence of 3-year mortality: 25.5% (95% CI: 23.3–27.6)) (Fig 2). Similarly, we identified 424 deaths during 847 days of median follow-up among the CEA patients (incidence of 3-year mortality: 16.8% (95% CI: 15.2–18.3)). The estimated crude hazard ratio was 1.82 (95% CI: 1.60–2.08).

Fig 3 shows the distribution of the propensity scores from the six models with different combinations of data source and variable selection approach. All PS models that included

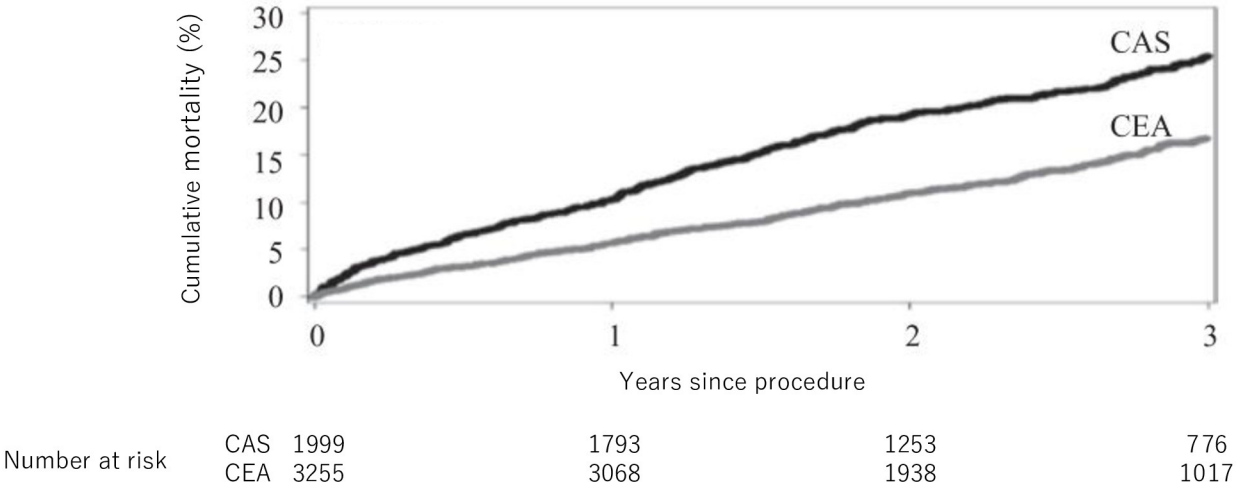

Fig 2. 3-year cumulative incidence of death among the Medicare-linked SVR-VR patients, by procedure. Abbreviations: CAS = Carotid Stenting; CEA = Carotid endarterectomy.

https://doi.org/10.1371/journal.pone.0272975.g002

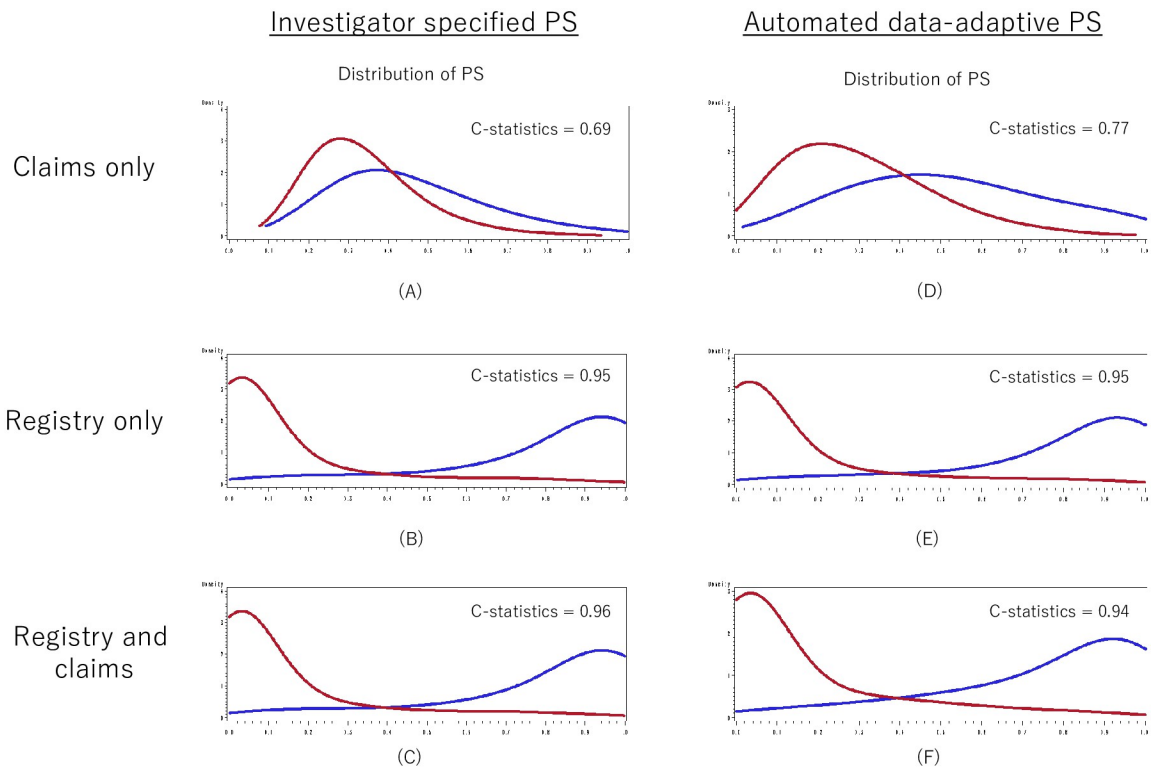

Fig 3. Distribution of the propensity scores estimated from the models with different combinations of data source and variable selection approach. Abbreviations: PS = propensity score.

https://doi.org/10.1371/journal.pone.0272975.g003

registry data (B, C, E, F) clearly separated the distribution between the CAS group and the CEA group. The PS models using claims data only (A, D) resulted in much greater overlap between the two groups, indicating less ability to discriminate between those receiving CAS and those receiving CEA. The c-statistics of the propensity score models ranged from 0.69 to 0.96. The difference in the PS distributions and c-statistics for IS-PS and hdPS were identical or nearly identical except for the models using claims data only (A, D) where the discrimination between the two groups was higher using hdps (c-statistic of 0.77 vs. 0.69). Among the 4 models using either registry data alone or registry plus claims data (B, C, E, F), the c-statistics were similar regardless of whether investigator-specified variables or hdPS-selected variables were used. All models with information on surgical risk had c-statistics of 0.94 and above, while none without had such high c-statistics ($\leq$0.80) (Table 2).

In the PS-matched analysis, the largest number of patients were retained when the investigator-specified PS model with claims data only was used (1674 patients each in the CAS and CEA groups). The adjusted HR estimate in the matched cohort was 1.45 (95% CI: 1.24–1.69), significantly different from the benchmark value of 1.0. The HR estimates were similar for the models using either registry data only or claims only, with or without the use of hdPS approach for variable selection. Investigator-specified propensity scores yielded a point estimate slightly closer to the benchmark of 1.0 at 1.36 (1.02–1.81) when both claims and registry data were used, and hdPS from claims plus registry data yielded a smaller HR estimate of 1.22 (0.92–1.63), which was the only estimate without statistical significance across the different approaches compared. The hdPS model with exclusion of high surgical risk status yielded an HR estimate of 1.32 (1.07–1.64) that was further away from the benchmark.

Table 2. Number of events, total cell sizes and hazard ratios for CAS relative to CEA based on different propensity score adjustment techniques across the different propensity score models.

	Propensity score models from investigator-specified variables			propensity score models from automated data-adaptive approach			
	Data from claims only	Data from registry only	Data from both claims and registry	Data from claims only	Data from registry only	Data from both claims and registry excluding high surgical risk status	Data from both claims and registry
Model c-statistic	0.69	0.95	0.96	0.77	0.95	0.80	0.94
PS quintiles Adjustment–no trimming							
Events / Number of CEA Patients	465 / 3255	465 / 3255	465 / 3255	465 / 3255	465 / 3255	465 / 3255	465 / 3255
Events / Number of CAS Patients	528 / 1999	528 / 1999	528 / 1999	528 / 1999	528 / 1999	528 / 1999	528 / 1999
HR	1.48 (1.28–1.70)	1.51 (1.22–1.86)	1.34 (1.07–1.69)	1.52 (1.31–1.76)	1.42 (1.14–1.76)	1.39 (1.20–1.62)	1.23 (1.00–1.50)
PS quintiles adjustment + trimming							
Events / Number of CEA Patients	417 / 2886	175 / 1053	156 / 869	387 / 2650	196 / 951	371 / 2551	224 / 1181
Events / Number of CAS Patients	396 / 1669	154 / 652	133 / 574	323 / 1494	168 / 670	320 / 1462	187 / 828
HR	1.47 (1.25–1.71)	1.36 (1.05–1.76)	1.21 (0.92–1.59)	1.41 (1.19–1.67)	1.38 (1.08–1.76)	1.32 (1.12–1.56)	1.18 (0.94–1.49)
Matching on PS							
Events / Number of CEA Patients	279 / 1,674	108 / 588	95 / 548	230 / 1446	103 / 575	224 / 1,332	112 / 668
Events / Number of CAS Patients	381 / 1,674	144 / 588	127 / 548	320 / 1446	103 / 575	283 / 1,332	154 / 668
HR	1.45 (1.24–1.69)	1.41 (1.09–1.83)	1.36 (1.02–1.81)	1.47 (1.21–1.78)	1.40 (1.05–1.86)	1.32 (1.07–1.64)	1.22 (0.92–1.63)

Age, sex, race, and year of procedure were pre-specified variables included in the high-dimensional propensity score by the investigator.

Abbreviations: HR = hazard ratio; CAS = Carotid stenting; CEA = Carotid endarterectomy

Note: Investigator-specified propensity score using claims data includes: age, sex, race, year of procedure, recent myocardial infarction, heart failure, atrial fibrillation, angina, pulmonary disease, cervical osteoarthritis, hypertension, ischemic heart disease, other atherosclerotic arterial disease, ventricular arrhythmia, peripheral vascular disease, valvular heart disease, chronic kidney disease, anemia, hyperlipidemia, diabetes, chronic obstructive pulmonary disease, coagulopathies, cancer, gastrointestinal bleed, glaucoma, osteoporosis, dementia, delirium, depression, bipolar disorder, anxiety, other psychiatric disorders, smoking history, alcohol abuse, drug abuse, pressure ulcers, cardiomyopathy, flu vaccination, pneumococcal vaccination, mammography, occult fecal blood test, bone mass density test, prostate-specific antigen test, prior carotid endarterectomy, prior CAS, prior radical neck dissection, prior open heart surgery, prior tracheostomy, prior percutaneous coronary intervention, prior coronary artery bypass graft procedure, non-elective hospital admission, Elixhauser comorbidity score, past-year hospitalizations, past-year physician visits, and past-year nursing home stay.

Investigator-specified propensity score using registry data includes: age, sex, race, year of procedure, CAS indication, medical history (coronary artery disease, cardiac arrhythmia, myocardial infarction, valvular heart disease, heart failure, hypertension, diabetes mellitus, stroke, transient ischemic attack, amaurosis fugax, peripheral vascular disease, renal failure, gastrointestinal ulcer or bleed, chronic obstructive pulmonary disease, current or past smoking, cancer, coagulopathy, hyperlipidemia/dyslipidemia, kidney function as per estimated glomerular filtration rate), American Society of Anesthesiology grade, New York Heart Association heart failure class, procedure side, carotid symptoms (carotid stenosis symptomatic in past year, past-year ipsilateral stroke, past-year ipsilateral transient ischemic attack, past-year amaurosis fugax, ipsilateral stroke more than one year prior to the procedure, ipsilateral transient ischemic attack more than one year prior to the procedure, amaurosis fugax more than one year prior to the procedure, modified Rankin Scale score), physiologic high surgical risk factors (age >80, NYHA class III or IV heart failure, left ventricular ejection fraction <30%, unstable angina, myocardial infraction 30 days prior to the procedure, other physiologic high surgical risk criteria from investigational device exemption CAS trials), anatomic high surgical risk factors (restenosis, radical neck dissection, contralateral occlusion, prior neck radiation, contralateral laryngeal nerve palsy or injury, high anatomic lesion, other anatomic high surgical risk criteria from investigational device exemption CAS trials), ipsilateral carotid stenosis (mild, moderate, severe), contralateral carotid stenosis ≥70%, and pre-procedural medication use (none, aspirin, clopidogrel, ticlopidine, other antiplatelet medication).

Investigator-specified propensity score using claims and registry data includes all the variables listed in the investigator-specified propensity score using claims only and registry data only.

https://doi.org/10.1371/journal.pone.0272975.t002

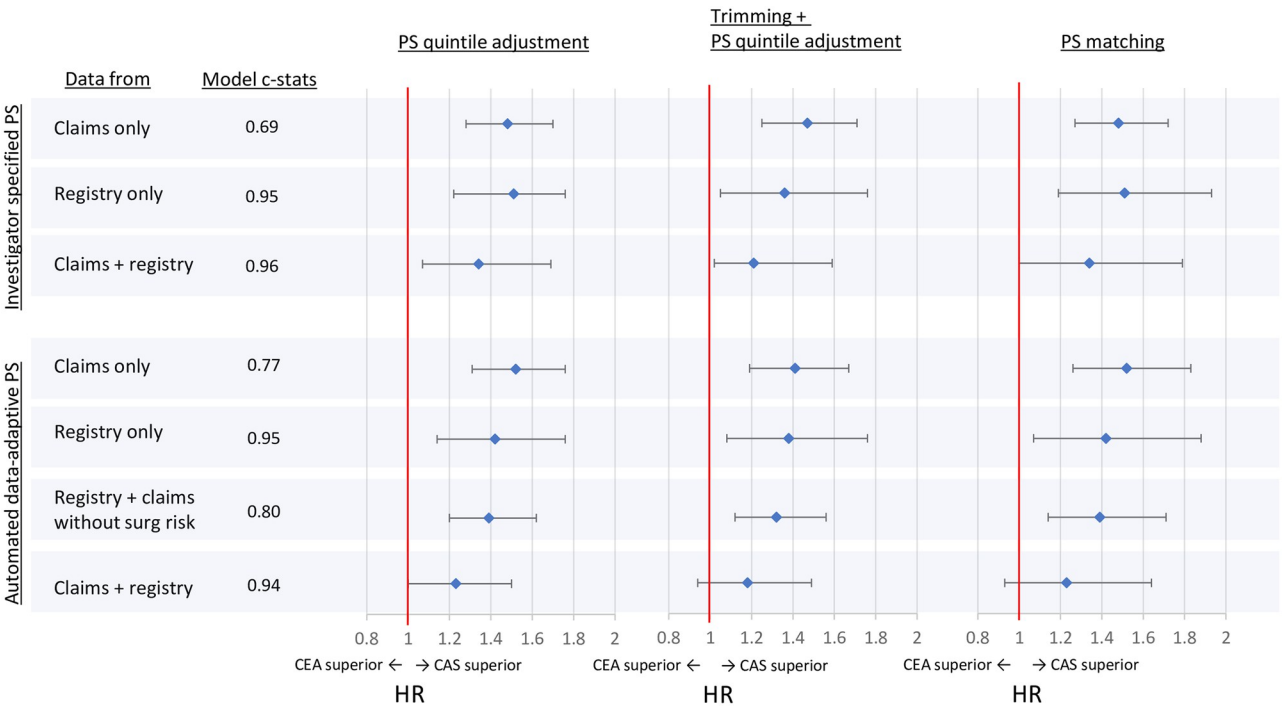

Fig 4. Hazard ratios comparing CAS vs. CEA from outcome models adjusting for investigator specific vs. automated data-adaptive propensity score using covariates from claims only, registry only, or claims+registry. Abbreviations: PS = propensity score, CEA = carotid endoarterectomy, CAS = coronary artery stenting, HR = hazard ratio.

https://doi.org/10.1371/journal.pone.0272975.g004

The findings from the two PS adjustment approaches were similar to those from the matched analysis. PS quintile adjustment without trimming yielded HR estimates slightly larger (i.e., away from the benchmark) compared to the matched analyses, although the difference was small. On the other hand, PS quintile adjustment with asymmetric trimming yielded point estimates that were generally similar or slightly smaller than those from the corresponding matched analyses, despite having more patients in the cohort than the matched analyses. In both adjustment approaches, hdPS from claims plus registry data yielded the smallest estimated HR; hdPS without information on high surgical risk status yielded an HR further from the benchmark. The HRs and 95% confidence intervals estimated from all of the PS model are summarized in Fig 4.

The top 25 variables identified by the hdPS algorithm as having the greatest potential to cause confounding from claims data, registry data and the combination of the two are listed in Table 3. The variables selected from registry data were clinical factors, most of which matched the investigator-specified PS variables. Variables selected from claims data included clinical factors such as diagnoses for heart failure and pneumonia, as well as CPT codes indicating health services use such as hospital care, electrocardiogram, and ambulance transportation. The factors selected via the hdPS algorithm when both registry data and claims data were a mix of these two.

Discussion

In a cardiovascular CER study with strong confounding by indication, we compared the performance of various methods for confounding adjustment using propensity scores (both investigator-specified and automated data-adaptive method) and using both claims and registry

Table 3. Top 25 variables identified by the high-dimensional propensity score algorithm as having the greatest potential to cause confounding.

All claims dimensions	All registry dimensions	All Claims and registry dimensions
4280 (ICD-9-CM; heart failure)	High surgical risk–physiologic (registry)	High surgical risk–physiologic (registry)
4280 (ICD-9-CM; heart failure)	High surgical risk–other physiologic from CAS Investigational Device Exemption trials (registry)	4280 (ICD-9-CM; heart failure)
4280 (ICD-9-CM; heart failure)	Heart failure (registry)	High surgical risk–other physiologic from CAS Investigational Device Exemption trials (registry)
4280 (ICD-9-CM; heart failure)	Left ventricular ejection fraction <30% (registry)	4280 (ICD-9-CM; heart failure)
99232 (CPT-4; subsequent hospital care)	NYHA class IV (registry)	4280 (ICD-9-CM; heart failure)
99232 (CPT-4; subsequent hospital care)	NYHA class III/IV (registry)	4280 (ICD-9-CM; heart failure)
4280 (ICD-9-CM; heart failure)	History of stroke (registry)	99232 (CPT-4; subsequent hospital care)
99231 (CPT-4; subsequent hospital care)	High surgical risk–anatomic (registry)	99232 (CPT-4; subsequent hospital care)
99232 (CPT-4; subsequent hospital care)	Contralateral carotid occlusion (registry)	4280 (ICD-9-CM; heart failure)
99231 (CPT-4; subsequent hospital care)	Carotid restenosis (registry)	99231 (CPT-4; subsequent hospital care)
4280 (ICD-9-CM; heart failure)	High surgical risk–other physiologic (registry)	Heart failure (registry)
93010 (CPT-4; electrocardiogram)	Previous neck radiation (registry)	99232 (CPT-4; subsequent hospital care)
99233 (CPT-4; subsequent hospital care)	Symptomatic carotid stenosis (registry)	99231 (CPT-4; subsequent hospital care)
71010 (ICD-9-CM; systemic sclerosis)	Radical neck dissection (registry)	4280 (ICD-9-CM; heart failure)
4280 (ICD-9-CM; heart failure)	History of myocardial infarction (registry)	93010 (CPT-4; electrocardiogram)
93010 (CPT-4; electrocardiogram)	Contralateral carotid stenosis ≥70% (registry)	Left ventricular ejection fraction <30% (registry)
A0425 (CPT-4; ambulance transportation)	End-stage renal disease (registry)	99233 (CPT-4; subsequent hospital care)
99238 (CPT-4; hospital discharge)	Modified Rankin scare score of 0 (registry)	71010 (ICD-9-CM; systemic sclerosis)
99223 (CPT-4; initial hospital care)	History of stroke (registry)	NYHA class IV (registry)
71010 (ICD-9-CM; systemic sclerosis)	History of cancer (registry)	4280 (ICD-9-CM; heart failure)
A0425 (CPT-4; ambulance transportation)	Transient ischemic attack in past year (registry)	93010 (CPT-4; electrocardiogram)
99231 (CPT-4; subsequent hospital care)	Diabetes mellitus (registry)	A0425 (CPT-4; ambulance transportation)
486 (ICD-9-CM; pneumonia)	Modified Rankin scale score of 4 (registry)	99238 (CPT-4; hospital discharge)
99254 (CPT-4; initial inpatient consultation visit)	Stroke in the past year (registry)	NYHA class III/IV (registry)
486 (ICD-9-CM; pneumonia)	Chronic renal failure (registry)	99223 (CPT-4; initial hospital care)

Note: Codes may appear more than once because codes are assessed for recurrence (once, sporadically, or frequently) in each dimension. For more information, see Schneeweiss et al. [10].

Abbreviations: ICD-9-CM = International Classification of Diseases, Clinical Modification; CPT-4 = Current Procedural Terminology; NYHA = New York Heart Association

https://doi.org/10.1371/journal.pone.0272975.t003

data. The c-statistics for the propensity score models were high and were similar for models from registry data with or without claims. However, propensity scores from claims data were not sufficient to achieve adequate confounding adjustment, even with the use of automated data-adaptive PS method. Clinical information on high surgical risk status from the registry was needed to produce estimates similar to what was expected based on results from RCTs [9, 21].

Using the data-adaptive PS method, variables with high potential for confounding were automatically identified based on their association with the exposure and the outcome. When using the registry as the only data source, the top 25 potential confounders identified were mostly clinical factors that constitute high surgical risk. These factors were strongly associated with CAS, since CAS procedures were reimbursed by the CMS only for those with high surgical risk. On the other hand, when we used only claims data, the top 25 potential confounders identified by data-adaptive PS method included heart failure, past hospitalizations, ambulance transportations, and electrocardiograms. Data-adaptive PS method did not identify all of the past diagnoses or treatments that constitute high surgical risk, most likely because they are under- or mis-captured in claims data. For example, conditions such as contralateral occlusion and contralateral nerve injury or palsy may not be captured accurately using the claims data as the information on laterality is oftentimes missing. Factors such as high anatomic lesion (C2 or higher) and the level of stenosis cannot be captured using claims data, because the billing codes are not designed to identify such detailed clinical information.

Confounding adjustment using registry information only also did not produce estimates that were closer to the estimates from RCT and variables from claims data were needed to achieve maximum confounding adjustment. This supports that claims databases may provide important information on confounders not captured in disease or device registries. First, registries for specific diseases or devices tend to have limited information on comorbidities that are not directly related to the main devices or diseases of interest (e.g., details on cancer or chronic medication use for CAS/CEA registry) while claims data capture all clinical encounters of patients regardless of the targeted diseases/devices. Second, claims data can provide proxies of non-clinical health determinants, including use of services and therefore patients' tendency to seek health care, their socioeconomic status, or frailty [18]. Automated data-adaptive approach may be especially useful in this context, as it can identify these proxies without investigators' pre-specification. It is important to reiterate that despite the above, using claims data only did not sufficiently adjust for confounding by indication in the case of CAS vs. CEA.

It is also worth noting that without detailed specification of variables by the researcher, automated data-adaptive PS was able to identify and include potential confounders in the model for the analysis for both registry-only and claims plus registry databases. This finding is consistent with prior studies of hdPS using claims databases [10, 22–26] and underscores the usefulness of the approach, especially in the absence of detailed pre-specification of confounders in situations such as conducting a large number of comparisons (e.g. monitoring and signal detection [27]), or when no previous assessment of potential confounders is available. Yet, at the same time, automated data-adaptive approach was not able to overcome the lack of information on patients' high surgical risk status which was the most important confounding by indication in the current study.

Limitations of our study should be noted. First, to assess the performance of different adjustment approaches, we used the reported relative risk estimates from randomized clinical trials as a benchmark. However, results from randomized studies may differ from comparative effectiveness estimates obtained using routine care data due to differences in patient populations and possible effect measure modification. Second, the precision of our estimates was poor, due to the limited number of patients and events in the linked cohort, and we were thus unable to conduct meaningful statistical tests to compare the estimates. Third, the example study we used is for a particular comparative analysis and may have limited generalizability. While we expect combinations of claims database and registries with common characteristics to show similar tendencies in terms of the usefulness of agsPS for other comparative effectiveness research questions, additional studies with such combinations are needed to confirm our findings.

In summary, in a cardiovascular CER study with a high likelihood of confounding by indication, an automated data-adaptive approach for PS modeling performed well, with similar results seen for individual claims, registry, and claims plus registry databases. Although the c-statistics of propensity score models were similar across registry-based models with and without additional claims data components, neither was sufficient to adequately control confounding by indication by itself; both claims and registry variables were needed for better control. As expected, among the different data dimensions in the registry and claims data, the most important to control for confounding by indication was high surgical risk status data in the registry. Despite its potential to control for unmeasured confounding, the automated data adaptive approach itself was not sufficient to control strong confounding by indication when the information is lacking in the data source.

Acknowledgments

Disclaimer: The views expressed in this article are those of the authors and do not represent policies of the AHRQ, CMS or the U.S. DHHS.

Author Contributions

Conceptualization: Jessica J. Jalbert, Soko Setoguchi.

Data curation: Jessica J. Jalbert, Lauren A. Williams.

Formal analysis: Hiraku Kumamaru, Jessica J. Jalbert.

Funding acquisition: Soko Setoguchi.

Investigation: Louis L. Nguyen, Hiroaki Miyata, Soko Setoguchi.

Methodology: Hiraku Kumamaru, Jessica J. Jalbert, Louis L. Nguyen, Soko Setoguchi.

Project administration: Lauren A. Williams.

Supervision: Louis L. Nguyen, Hiroaki Miyata, Soko Setoguchi.

Visualization: Hiraku Kumamaru.

Writing – original draft: Hiraku Kumamaru, Jessica J. Jalbert, Soko Setoguchi.

Writing – review & editing: Hiraku Kumamaru, Jessica J. Jalbert, Louis L. Nguyen, Lauren A. Williams, Hiroaki Miyata, Soko Setoguchi.

References

1. Kosiborod M, Lam CSP, Kohsaka S, Kim DJ, Karasik A, Shaw J, et al. Cardiovascular Events Associated With SGLT-2 Inhibitors Versus Other Glucose-Lowering Drugs: The CVD-REAL 2 Study. J Am Coll Cardiol. 2018; 71: 2628–2639. https://doi.org/10.1016/j.jacc.2018.03.009 PMID: 29540325
2. Chen CY, Stevenson LW, Stewart GC, Seeger JD, Williams L, Jalbert JJ, et al. Impact of baseline heart failure burden on post-implantable cardioverter-defibrillator mortality among medicare beneficiaries. J Am Coll Cardiol. 2013; 61: 2142–2150. https://doi.org/10.1016/j.jacc.2013.02.043 PMID: 23541973
3. Jalbert JJ, Nguyen LL, Gerhard-Herman MD, Kumamaru H, Chen C-Y, Williams LA, et al. Comparative effectiveness of carotid artery stenting versus carotid endarterectomy among medicare beneficiaries. Circ Cardiovasc Qual Outcomes. 2016; 9, 1–12. https://doi.org/10.1161/CIRCOUTCOMES.115.002336 PMID: 27116974
4. Patorno E, Goldfine AB, Schneeweiss S, Everett BM, Glynn RJ, Liu J, et al. Cardiovascular outcomes associated with canagliflozin versus other non-gliflozin antidiabetic drugs: Population based cohort study. BMJ. 2018; 360. https://doi.org/10.1136/bmj.k119 PMID: 29437648
5. Blin P, Fauchier L, Dureau-Pournin C, Sacher F, Dallongeville J, Bernard MA, et al. Effectiveness and Safety of Rivaroxaban 15 or 20 mg Versus Vitamin K Antagonists in Nonvalvular Atrial Fibrillation: A

Population-Based New Users High-Dimensional Propensity Score Matched Cohorts Study. Stroke. 2019; 50: 2469–2476. https://doi.org/10.1161/STROKEAHA.119.025824 PMID: 31390972

6. Graham DJ, Reichman ME, Wernecke M, Zhang R, Southworth MR, Levenson M, et al. Cardiovascular, Bleeding, and Mortality Risks in Elderly Medicare Patients Treated With Dabigatran or Warfarin for Nonvalvular Atrial Fibrillation. Circulation. 2014; 131: 157–164. https://doi.org/10.1161/CIRCULATIONAHA.114.012061 PMID: 25359164

7. Centers for Medicare & Medicaid Services Dicision Memo for Carotid Artery Stenting (CAG-00085R). [cited 11 Aug 2020]. https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=157

8. Macc CJW. Carotid stenting vs. endarterectomy. [cited 25 Feb 2020]. https://www.acc.org/latest-in-cardiology/articles/2016/04/27/09/38/carotid-stenting-vs-endarterectomy

9. Brott T, Hobson R, Howart G, Roubin G, Clark WM, Brooks W, et al. Stenting versus Endarterctomy for Treatment of Carotid-Artery Stenosis. N Engl J Med. 2010; 363: 11–23. https://doi.org/10.1056/NEJMoa0912321 PMID: 20505173

10. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology. 2009; 20: 512–522. https://doi.org/10.1097/EDE.0b013e3181a663cc PMID: 19487948

11. Rassen JA, Schneeweiss S. Using high-dimensional propensity scores to automate confounding control in a distributed medical product safety surveillance system. 2012; 21: 41–49. https://doi.org/10.1002/pds

12. Patorno E, Glynn RJ, Hernández-Díaz S, Liu J, Schneeweiss S. Studies with many covariates and few outcomes: selecting covariates and implementing propensity-score-based confounding adjustments. Epidemiology. 2014; 25: 268–78. https://doi.org/10.1097/EDE.0000000000000069 PMID: 24487209

13. Kumamaru H, Gagne JJ, Glynn RJ, Setoguchi S, Schneeweiss S. Comparison of high-dimensional confounder summary scores in comparative studies of newly marketed medications. J Clin Epidemiol. 2016; 76. https://doi.org/10.1016/j.jclinepi.2016.02.011 PMID: 26931292

14. Schneeweiss S. Automated data-adaptive analytics for electronic healthcare data to study causal treatment effects. Clin Epidemiol. 2018; 10: 771–788. https://doi.org/10.2147/CLEP.S166545 PMID: 30013400

15. Centers for Medicare and Medicaid Survices Medicare Coverage Database. Decision Memo for Carotid Artery Stending. 2005 [cited 23 Sep 2020]. https://www.cms.gov/medicare-coverage-database/details/nca-decision-memo.aspx?NCAId=157&ver=29&NcaName=Carotid+Artery+Stenting+(1st+Recon)

16. Sidawy AN, Zwolak RM, White RA, Siami FS, Schermerhorn ML, Sicard GA. Risk-adjusted 30-day outcomes of carotid stenting and endarterectomy: Results from the SVS Vascular Registry. J Vasc Surg. 2009; 49: 71–79. https://doi.org/10.1016/j.jvs.2008.08.039 PMID: 19028045

17. Setoguchi S, Zhu Y, Jalbert JJ, Williams LA, Chen CY. Validity of deterministic record linkage using multiple indirect personal identifiers linking a large registry to claims data. Circ Cardiovasc Qual Outcomes. 2014; 7: 475–480. https://doi.org/10.1161/CIRCOUTCOMES.113.000294 PMID: 24755909

18. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Robert J, et al. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol. 2001; 154: 854–64. Available: http://www.ncbi.nlm.nih.gov/pubmed/11682368 PMID: 11682368

19. Bross IDJ. Spurious effects from an extraneous variable. J Chronic Dis. 1966; 19: 637–647. https://doi.org/10.1016/0021-9681(66)90062-2 PMID: 5966011

20. Stürmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution—a simulation study. Am J Epidemiol. 2010; 172: 843–54. https://doi.org/10.1093/aje/kwq198 PMID: 20716704

21. Yadav J, Wholey M, Kuntz R, Fayad P, Katzen B, Mishkel GJ, et al. Protected Carotid-Artery Stenting versus Endarterectomy in High-Risk Patients. N Engl J Med. 2004; 351: 1493–1501. https://doi.org/10.1056/NEJMoa040127 PMID: 15470212

22. Rassen J, Choudhry N, Avorn J, Schneeweiss S. Cardiovascular outcomes and mortality in patients using clopidogrel with proton pump inhibitors after percutaneous coronary intervention. Circulation. 2009; 120: 2322–9. https://doi.org/10.1161/CIRCULATIONAHA.109.873497 PMID: 19933932

23. Le H V, Poole C, Brookhart MA, Schoenbach VJ, Beach KJ, Layton JB, et al. Effects of aggregation of drug and diagnostic codes on the performance of the high-dimensional propensity score algorithm: an empirical example. BMC Med Res Methodol. 2013; 13: 142. https://doi.org/10.1186/1471-2288-13-142 PMID: 24245772

24. Garbe E, Kloss S, Suling M, Pigeot I, Schneeweiss S. High-dimensional versus conventional propensity scores in a comparative effectiveness study of coxibs and reduced upper gastrointestinal complications. Eur J Clin Pharmacol. 2013; 69: 549–557. https://doi.org/10.1007/s00228-012-1334-2 PMID: 22763756

25. Schneeweiss S, Patrick AR, Solomon DH, Dormuth CR, Miller M, Mehta J, et al. Comparative safety of antidepressant agents for children and adolescents regarding suicidal acts. Pediatrics. 2010; 125: 876–888. https://doi.org/10.1542/peds.2009-2317 PMID: 20385637

26. Patorno E, Rhonda L, Wahl P, Avorn J, Patrick A, Liu J, et al. Anticonvulsant medications and the risk of suicide, attempted suicide, or violent death. JAMA. 2010; 303: 1401–9. https://doi.org/10.1001/jama.2010.410 PMID: 20388896

27. Gagne JJ, Rassen J a, Walker AM, Glynn RJ, Schneeweiss S. Active safety monitoring of new medical products using electronic healthcare data: selecting alerting rules. Epidemiology. 2012; 23: 238–46. https://doi.org/10.1097/EDE.0b013e3182459d7d PMID: 22266893