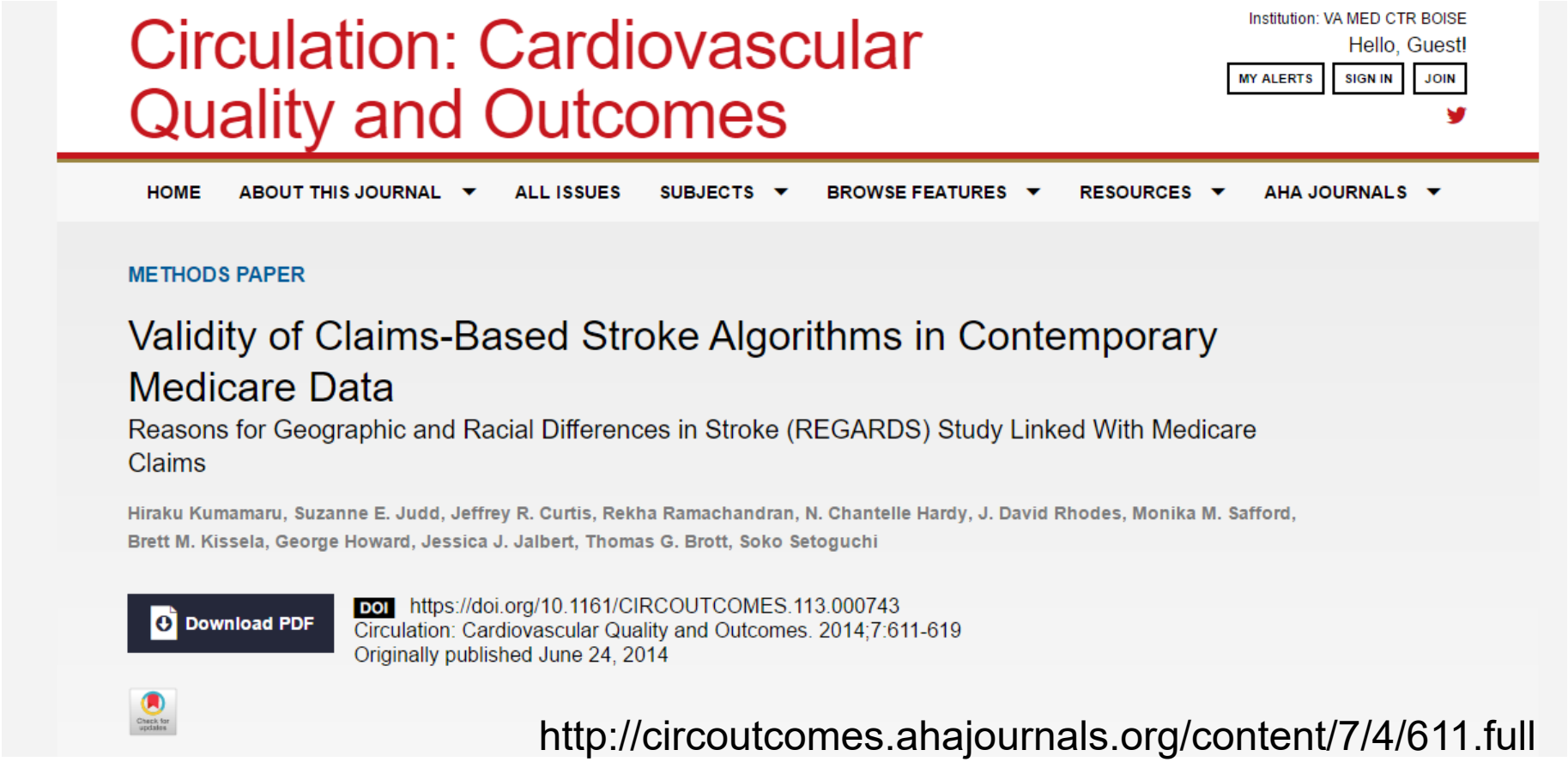
Validity of Stroke Diagnosis in Contemporary Medicare Data: Findings From the REGARDS Study Linked With Medicare Claims

Hiraku Kumamaru, MD ScD

Details of the study are reported in the paper below:



Circulation: Cardiovascular Quality and Outcomes

METHODS PAPER

Validity of Claims-Based Stroke Algorithms in Contemporary Medicare Data

Reasons for Geographic and Racial Differences in Stroke (REGARDS) Study Linked With Medicare Claims

Hiraku Kumamaru, Suzanne E. Judd, Jeffrey R. Curtis, Rekha Ramachandran, N. Chantelle Hardy, J. David Rhodes, Monika M. Safford, Brett M. Kissela, George Howard, Jessica J. Jalbert, Thomas G. Brott, Soko Setoguchi

DOI https://doi.org/10.1161/CIRCOUTCOMES.113.000743
Circulation: Cardiovascular Quality and Outcomes. 2014;7:611-619
Originally published June 24, 2014

http://circoutcomes.ahajournals.org/content/7/4/611.full

HOW TO DEAL WITH MISSING GOLD STANDARD?

Background 1

- Multiple studies validated the accuracy of stroke diagnosis in administrative claims
  - No data available for the contemporary population of Medicare enrollees
- Improvements in diagnostic modalities have changed clinical practice over the last 15 years.
  - Validity of stroke ascertainment algorithms in a contemporary population differ from that suggested by older studies?

Background 2

- REasons for Geographic And Racial Differences in Stroke (REGARDS) Study
  - Nationwide epidemiologic study of 30,239 patients
  - Linked to Medicare claims data.
  - Provided a unique opportunity to assess the validity of claims-based algorithms for stroke diagnosis in contemporary Medicare beneficiaries.

Objective

To assess the validity of claims-based algorithms to identify incident and recurrent stroke in the contemporary Medicare population using the linked REGARDS-Medicare database

Data Sources

REGARDS cohort database

Medicare claims data 2003 - 2009

REasons for Geographic And Racial Differences in Stroke (REGARDS) cohort database

30,239 participants

≥45

Randomly sampled/recruited 2003 – 2007

Oversampling blacks, “Stroke Belt”

After enrollment:

Phone interview (demographic, risk factor socioeconomic, and medical history)

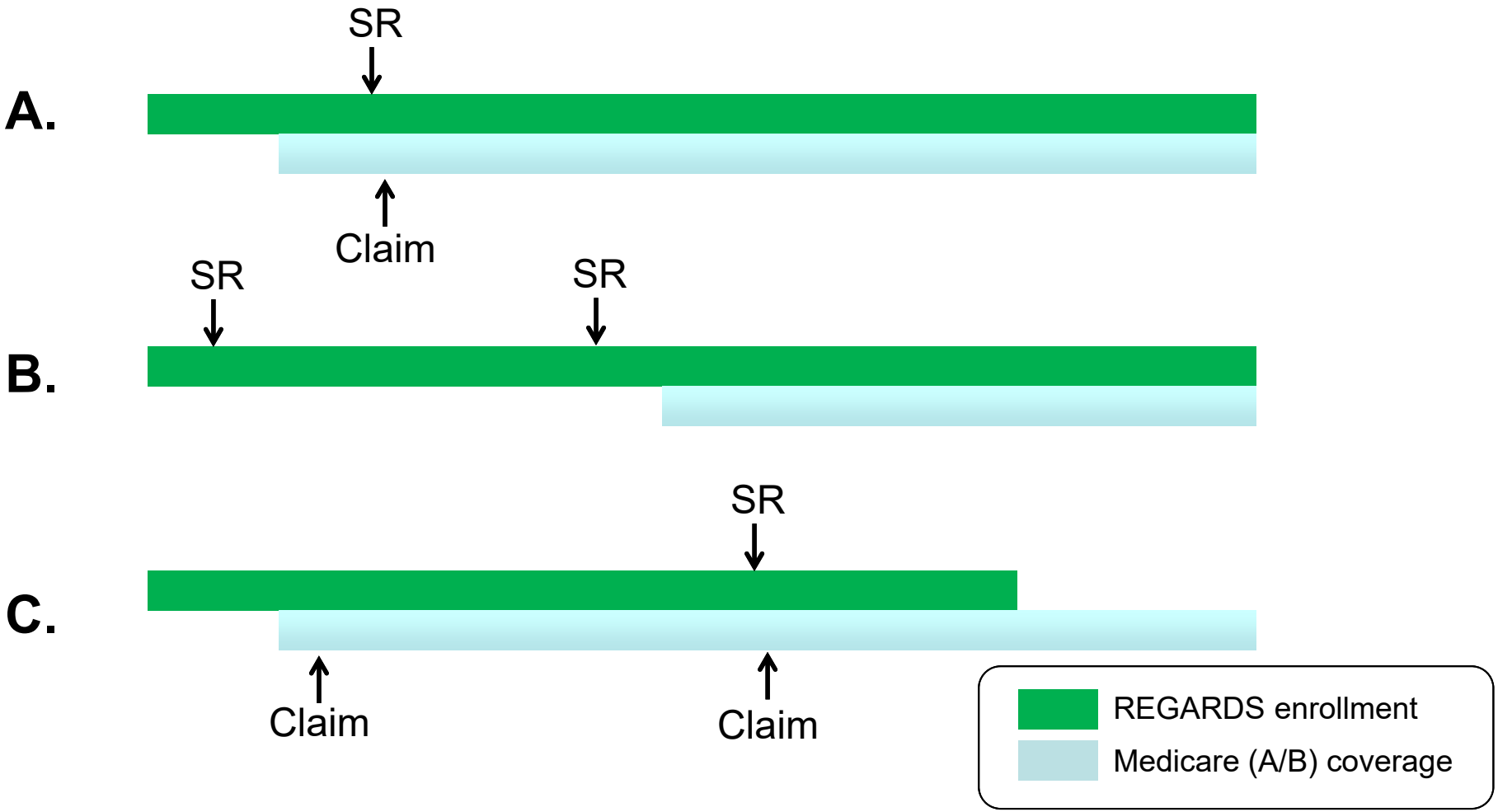
Physical exam (BP, ECG, blood/urine)

Follow-up phone interview every 6 months

Suspected Events

Events are identified through 2 paths in the study:

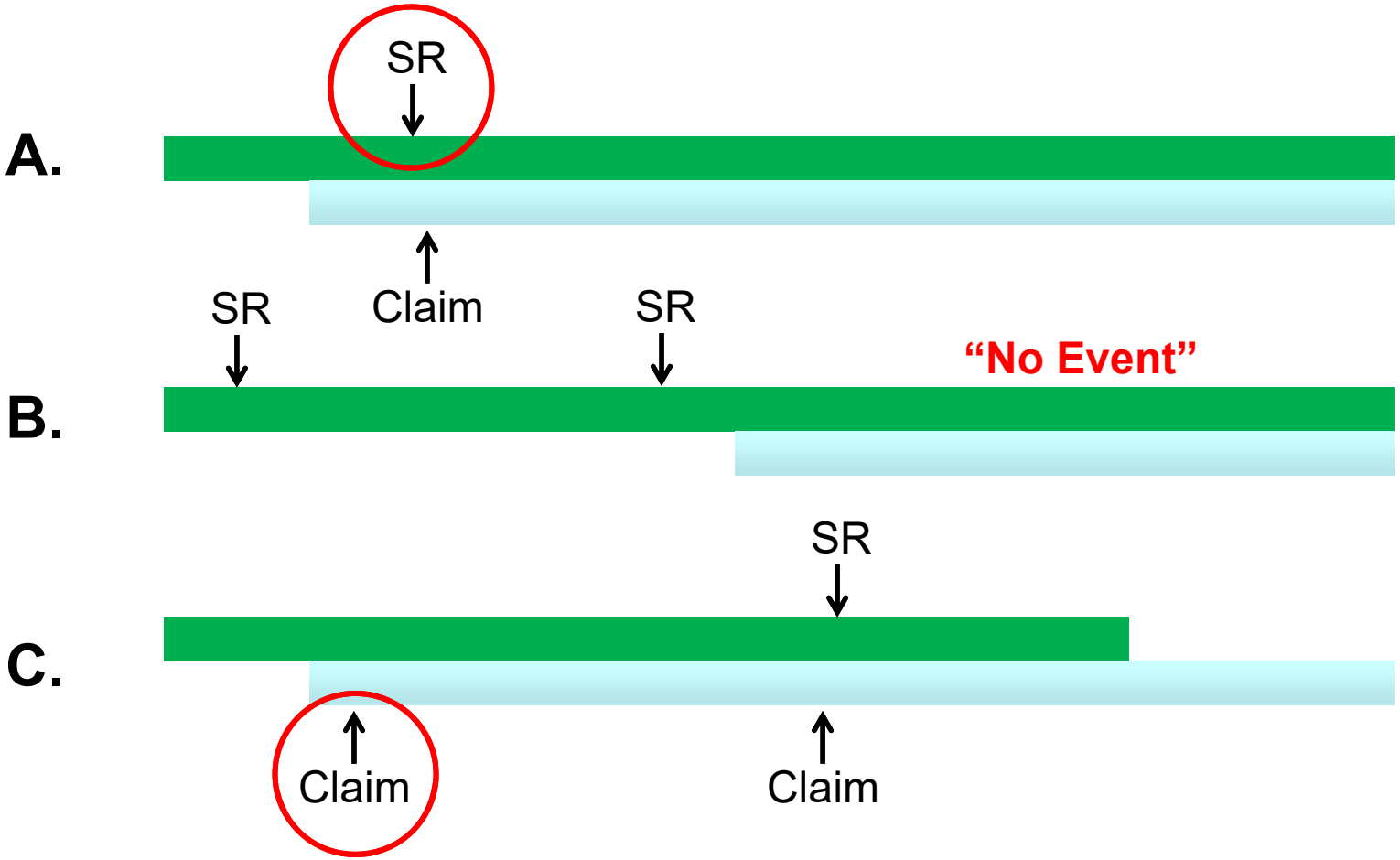
1) Patient self report (SR) at phone interview
2) Medicare Claim algorithm



Suspected Events

i) For each individual, we selected their first claim or their first SR

ii) Those with no SR or claim = “No events”



Gold Standard

6 month phone interview picked up
self reported stroke related events:

stroke hospitalizations, TIA hospitalizations,
stroke symptoms, stroke related deaths
→ medical records were pursued for review.

Triaged by a stroke nurse
Reviewed by stroke neurologists.

Gold Standard

Stroke was defined by:

1) World Health Organization definition:
   Rapidly developing clinical signs of focal, at times global, disturbance of cerebral function, lasting more than 24 hours or leading to death with no apparent cause other than that of vascular origin

OR

2) “clinical strokes” with clinical signs with imaging confirmation

And classified as ischemic or hemorrhagic.

Gold Standard

Additional medical records were retrieved for patients with hospitalization claims (ICD-9-CM codes: 430, 431, 433.x1, 434.x1, 436) as primary Dx. in the linked Medicare database

The retrieved medical charts were used to adjudicate these suspected events using the same methodology as used in REGARDS.

Claims-Based Algorithms

(1) acute ischemic stroke (AIS) algorithm
ICD-9-CM code 433.x1, 434.x1, or 436 in the primary position of the discharge Dx.

(2) intracranial hemorrhage (ICH) algorithm
I*CD-9-CM* code 430 or 431 in the primary position of the discharge Dx.

(3) AIS/ICH algorithm
ICD-9-CM code 430, 431, 433.x1, 434.x1, or 436 in the primary position of the discharge Dx.

Failure to retrieve Medical Records

Failure to retrieve the medical record occurred in about 10% of the suspected stroke cases. The proportion was higher for those cases identified through claims only at about 60%.

AIS / ICH algorithm

Category	Claim	Self report	Missing %	Total Number
Claim + SR	**+**	**+**	**0.10**	**185**
Claim only	**+**	**-**	**0.62**	**97**
SR only	**-**	**+**	**0.08**	**1,748**
No event	**-**	**-**	**-**	**13,059**

Complete Case Analysis

	Disease (+)	Disease (-)		missing
Claim (+)	TP = 184	FP = 20	PPV	78 (28%)
Claim (-)	FN = 153	TN = 14,506	NPV	148 (1%)
	Sensitivity	Specificity		

Complete case analysis assumes missingness completely at random. As missingness was highly correlated with the presence of claims in our study, it would underestimate the sensitivity.

I. Complete case:
   Sensitivity = 54.6%

II. Assuming same proportion of true cases in missing:
   Sensitivity = 62.2%

Adjudication Results for Each Algorithm (AIS / ICH)
ICD-9-CM: 430, 431, 433.x1, 434.x1, 436

Category	Total No.	Adjudication Results No. (%)			Prop. of True Cases (A/A+B)	Prop. of Missing (C/A+B+C)
		Stroke (A)	No Stroke (B)	Missing (C)		
Claim + SR	185	150	17	18	0.90	0.10
Claim only	97	34	3	60	0.92	0.62
SR only	1,748	153	1,447	148	0.10	0.08
No event	13,059	—	13,059	—	—	—
Total	15,089	337	14,526	226	—	—

Accuracy of ICD-9-CM Codes in the Primary Dx. Position and 95% confidence intervals

Algorithm	PPV (%)	Sensitivity (%)	Specificity (%)	NPV (%)
AIS	91.1 (86.6-95.5)	58.6 (52.4-64.7)	99.9 (99.8-100)	99.2 (99.1-99.4)
ICH	84.8 (71.8-97.9)	59.9 (44.9-74.9)	100.0 (99.9-100.0)	99.9 (99.8-99.9)
AIS/ICH	92.6 (88.8-96.4)	59.5 (53.8-65.1)	99.9 (99.8-100)	99.1 (98.9-99.3)

Accuracy of AIS/ICH algorithm by Subgroup

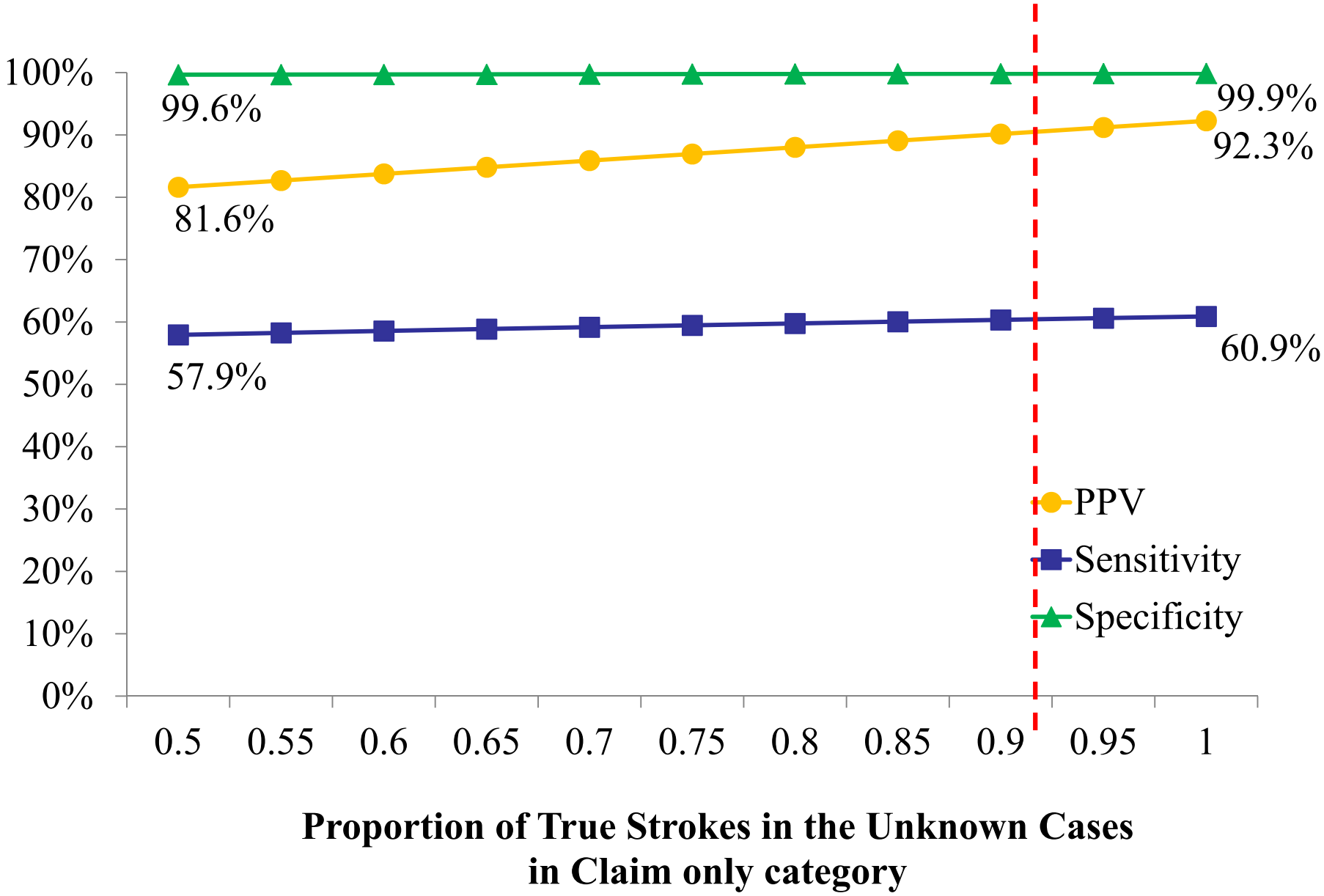
	Sub-group	No.	PPV (%)	Sensitivity (%)	Specificity (%)	NPV (%)
Age	< 65	1,551	—*	—*	—*	—*
	65-74	8,702	96.2 (92.3-100)	56.4 (48.6-64.2)	100.0 (99.9-100)	99.2 (99.0-99.4)
	≥ 75	2,843	92.4 (86.7-98.1)	67.4 (58.8-76.0)	99.8 (99.6-99.9)	98.7 (98.2-99.1)
Gender	Female	6,782	96.3 (92.4-100)	64.6 (56.5-72.6)	99.9 (99.9-100)	99.3 (99.1-99.5)
	Male	6,314	89.9 (83.8-96.0)	55.2 (47.3-63.1)	99.8 (99.7-99.9)	98.9 (98.6-99.2)
Race	Black	4,689	92.8 (87.2-98.4)	64.3 (55.6-72.9)	99.9 (99.8-100)	99.1 (98.8-99.4)
	White	8,407	93.3 (88.5-98.1)	56.5 (49.0-63.9)	99.9 (99.9-100)	99.1 (98.9-99.3)

* Calculations could not be completed as some cells had zero observations

Adjudication Results for Each Algorithm (AIS / ICH)
ICD-9-CM: 430, 431, 433.x1, 434.x1, 436

Category	Total No.	Adjudication Results No. (%)			Prop. of True Cases (A/A+B)	Prop. of Missing (C/A+B+C)
		Stroke (A)	No Stroke (B)	Missing (C)		
Claim + SR	185	150	17	18	0.90	0.10
Claim only	97	34	3	60	0.92	0.62
SR only	1,748	153	1,447	148	0.10	0.08
No event	13,059	—	13,059	—	—	—
Total	15,089	337	14,526	226	—	—

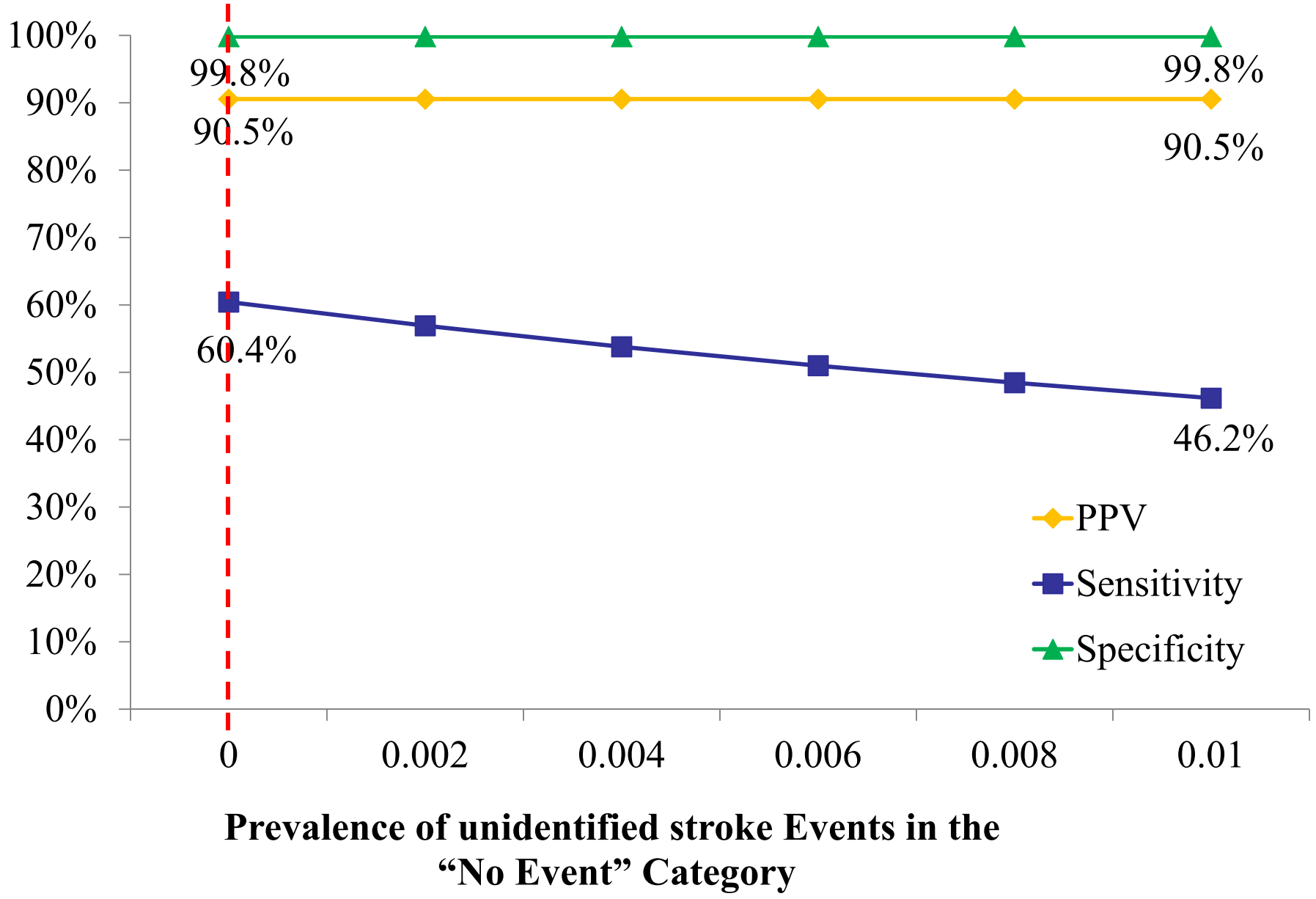
Sensitivity Analysis 1: Effect of altering the assumed proportion of true cases in unknown “claim only” for AIS/ICH



Adjudication Results for Each Algorithm (AIS / ICH)
ICD-9-CM: 430, 431, 433.x1, 434.x1, 436

Category	Total No.	Adjudication Results No. (%)			Prop. of True Cases (A/A+B)	Prop. of Missing (C/A+B+C)
		Stroke (A)	No Stroke (B)	Missing (C)		
Claim + SR	185	150	17	18	0.90	0.10
Claim only	97	34	3	60	0.92	0.62
SR only	1,748	153	1,447	148	0.10	0.08
No event	13,059	—	13,059	—	—	—
Total	15,089	337	14,526	226	—	—

Sensitivity Analysis 2: Effect of altering the assumed proportion of unidentified cases in the “no event” category



Limitations

- Algorithms were limited to those using primary diagnoses.
- Algorithms did not include TIAs
- Differential missingness by stroke triggering events
- Underestimation of confidence interval due to the use of imputed means

Summary

- The claims-based algorithms using primary discharge diagnoses captured true stroke events among the contemporary Medicare enrollees with high positive predictive value and high specificity
- Differences in age, sex, and race had limited influence on specificity and negative predictive values, but sensitivity and positive predictive values varied by these factors, although not statistically significant.

Implications

- High specificity and positive predictive value of the claim-based algorithms will maintain the validity of etiologic studies using administrative data as long as the misclassification of stroke is non-differential
- The low sensitivity of the claim-based algorithms will compromise the power of the study identifying strokes as an outcome

Discussion points

- What other ways are available to treat the differential missingness of the adjudication results for those with failed medical record retrieval?
  - Complete case analyses
  - Multiple imputation