

Chart validation of inpatient ICD-9-CM administrative diagnosis codes for ischemic stroke among IGIV users in the Sentinel Distributed Database

Eric M. Ammann, PhD^{a,*}, Enrique C. Leira, MD MS^{a,b}, Scott K. Winiecki, MD^c, Nandakumar Nagaraja, MD^b, Sudeepa Dandapat, MD^b, Ryan M. Carnahan, PharmD, MS^a, Marin L. Schweizer, PhD^{b,d}, James C. Torner, PhD^a, Candace C. Fuller, PhD^e, Charles E. Leonard, PharmD, MSCE^f, Crystal Garcia, MPH^e, Madelyn Pimentel, MSN^e, Elizabeth A. Chrischilles, PhD^a

Abstract

The Sentinel Distributed Database (SDD) is a large database of patient-level medical and prescription records, primarily derived from insurance claims and electronic health records, and is sponsored by the U.S. Food and Drug Administration for drug safety assessments. In this chart validation study, we report on the positive predictive value (PPV) of inpatient ICD-9-CM acute ischemic stroke (AIS) administrative diagnosis codes (433.x1, 434.xx, and 436) in the SDD.

As part of an assessment of the risk of thromboembolic adverse events following treatment with intravenous immune globulin (IGIV), charts were obtained for 131 potential post-IGIV AIS cases. Charts were abstracted by trained nurses and then adjudicated by stroke experts using pre-specified diagnostic criteria.

Case status could be determined for 128 potential AIS cases, of which 34 were confirmed. The PPVs for the inpatient AIS diagnoses recorded in the SDD were 27% overall [95% confidence interval (95% CI): 19–35], 60% (95% CI: 32–84) for principal-position diagnoses, 42% (95% CI: 28–57) for secondary diagnoses, and 6% (95% CI: 2–15) for position-unspecified diagnoses (which in the SDD generally originate from separate physician claims associated with an inpatient stay).

Position-unspecified diagnoses were unlikely to represent true AIS cases. PPVs for principal and secondary inpatient diagnosis codes were higher, but still meaningfully lower than estimates from prior chart validation studies. The low PPVs may be specific to the IGIV user study population. Additional research is needed to assess the validity of AIS administrative diagnosis codes in other study populations within the SDD.

Abbreviations: AIS = acute ischemic stroke, CT = computed tomography, ICD-9-CM = International Classification of Diseases, Ninth Revision, Clinical Modification, IGIV = intravenous immune globulin, MRI = magnetic resonance imaging, PPV = positive predictive value, SDD = Sentinel Distributed Database, TEE = thromboembolic event.

Keywords: administrative data, cerebrovascular disease, chart validation, ischemic stroke, positive predictive value

Editor: Nan Liu.

Funding/support: The results reported herein correspond to the objectives of Mini-Sentinel contract HHSF2232009100061 from the U.S. Food and Drug Administration (FDA) and Department of Health and Human Services (HHS). This work was also supported by the Sentinel Coordinating Center, which is funded by the FDA through HHS contract number HHSF2232014000301.

Reproducibility: The data files and SAS code used to produce the results presented in this paper were reviewed and verified by the programming staff at the Sentinel Coordinating Center. The SAS code is available upon request from the corresponding author. For legal reasons, the individual-level patient data that served as the basis for our results are not available for public distribution.

EMA is now an employee in the Medical Device Epidemiology division at Johnson & Johnson. The analyses for this paper were completed and the manuscript drafted prior to his start in that role. The other authors report no conflicts.

Supplemental Digital Content is available for this article.

^a College of Public Health, University of Iowa, Iowa City, IA, ^b Carver College of Medicine, University of Iowa, Iowa City, IA, ^c Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD, ^d Iowa City VA Health Care System, Iowa City, IA, ^e Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, ^f Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA.

* Correspondence: Eric M. Ammann, CPHB S400, 145 North Riverside Drive, The University of Iowa, Iowa City, IA 52242 (e-mail: eric-ammann@uiowa.edu).

Copyright © 2017 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

Medicine (2017) 96:52(e9440)

Received: 20 September 2017 / Received in final form: 8 November 2017 / Accepted: 1 December 2017

<http://dx.doi.org/10.1097/MD.00000000000009440>

1. Introduction

In this study, we evaluated the positive predictive value (PPV) of inpatient diagnosis codes for acute ischemic stroke (AIS) in the Sentinel Distributed Database (SDD). The SDD is a large database of longitudinal, patient-level medical and prescription data from a variety of data sources (primarily, billing data from large health insurers and administrative data from integrated healthcare delivery systems) that are converted to a common data format. The SDD and the Sentinel program are sponsored by the U.S. Food and Drug Administration (FDA) for active safety surveillance of marketed medical products. For 2000 to 2016, the SDD has 425 million person-years of longitudinal patient-level data from 223 million health plan members.^[1] Because AIS is a frequent endpoint for studies conducted in the SDD and other administrative databases, it is important that validation studies be conducted on an ongoing basis to establish the PPV of AIS administrative diagnoses.

Prior validation studies conducted outside the SDD indicate that hospital discharge diagnosis codes for AIS generally have high PPVs (80% or higher), with principal-position diagnoses performing somewhat better than secondary diagnoses.^[2] However, to date, no validation studies of AIS diagnosis codes have been performed within the SDD. In addition, medical coding guidelines for AIS were modified in the mid-2000s,^[3] potentially affecting the validity of AIS-related administrative diagnosis codes. To inform the design and interpretation of future studies of AIS based on records from the SDD and other administrative databases, we report on the PPVs associated with inpatient diagnosis codes for AIS recorded during the years 2006 to 2012. Possible cases included in this chart validation study were identified from the SDD as part of a safety assessment of thromboembolic event (TEE) risk following intravenous immune globulin (IGIV).

2. Methods

2.1. Data sources and study population

The administrative health care records and patient medical charts used to identify and validate potential AIS cases came from 13 SDD Data Partners (i.e., large insurers and integrated care delivery systems) who participated in the protocol-based Sentinel assessment of thromboembolic events following immunoglobulin administration.^[4] Potential cases from the years 2006 to 2012 were selected for chart review if an inpatient AIS diagnosis code was recorded in the SDD up to one month following a non-specific (i.e., polyvalent) IGIV treatment episode. A complete description of the criteria used to select potential cases can be found in the Appendix, <http://links.lww.com/MD/C42>. Additional details concerning the design and objectives of the parent study have been described previously.

IGIV is used in the treatment of primary and secondary immunoglobulin deficiencies, as well as inflammatory and autoimmune disorders (e.g., chronic demyelinating polyneuropathy and immune thrombocytopenic purpura).^[5] We provide descriptive information on the patients included in this chart validation study in Table 1, including their possible indications for IGIV use and major cardiovascular risk factors. These health conditions were defined as previously described in the protocol for the parent study.^[4]

2.2. Research ethics and institutional review board review

The data presented in this paper were collected as part of a public health surveillance activity conducted under the auspices of

Table 1

Baseline characteristics of 131 potential acute ischemic stroke (AIS) cases identified from the Sentinel Distributed Database for whom chart retrieval was completed.

Characteristic	N (%)
Demographics	
Age, y	
0–19	14 (11%)
20–39	15 (11%)
40–59	39 (30%)
60–79	46 (35%)
80+	17 (13%)
Female sex	
Female	65 (50%)
Male	66 (50%)
Possible indication for IGIV use	
Autoimmune/inflammatory condition	87 (66%)
Immune deficiency	34 (26%)
Infection	18 (14%)
Bone marrow or hematopoietic stem cell transplant	7 (5%)
Other indication	25 (19%)
Major risk factors for ischemic stroke and atherosclerotic cardiovascular disease*	
Myocardial infarction	13 (10%)
Angina	27 (21%)
Atrial fibrillation or flutter	20 (15%)
Ischemic stroke	20 (15%)
Peripheral vascular disease	12 (9%)
Hypertension, uncomplicated	69 (53%)
Hypertension, complicated	21 (16%)
Diabetes mellitus	40 (31%)
Data Partner type	
Insurer/claims-based	100 (76%)
Integrated health care delivery system	31 (24%)

Possible indications and cardiovascular risk factors were assessed using diagnoses and procedures recorded in administrative data during the 183 days before the patient's IGIV treatment date. The indication indicator variables are not mutually exclusive, so indication percentages may sum to greater than 100%.

IGIV = intravenous immune globulin.

the FDA Sentinel Initiative. For this reason, the collection and analysis of these data did not qualify as human subjects research under the Common Rule and were not subject to institutional review board (IRB) review.^[6–8] The administrative data records and medical charts reviewed for these analyses were stored on password-protected secure servers to maintain patient confidentiality.

2.3. Case identification and chart retrieval

The endpoint definition used to identify potential AIS cases included the following International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes originating from an inpatient hospital encounter: 433.x1 (occlusion and stenosis of pre-cerebral arteries with cerebral infarction), 434.xx (occlusion of cerebral arteries), or 436 (acute but ill-defined cerebrovascular disease).

Within the Sentinel Common Data Model, diagnosis codes associated with inpatient encounters are categorized as principal, secondary, or “unable to classify” (i.e., position unspecified). These classifications reflect standard coding practices and the addition of a third category to accommodate heterogeneity across Sentinel Data Partners in how encounters and coding positions are defined. Under Uniform Hospital Discharge Data Set (UHDDS) guidelines used by U.S. hospitals and insurers,^[9] inpatient diagnoses are coded as follows:

- Principal diagnosis: the condition established after study to be chiefly responsible for occasioning the admission of the patient to the hospital
- Secondary diagnosis: a condition also present on admission, which developed during the hospital stay, or that influenced the care of the patient or length of stay

In the SDD, there are also position-unspecified diagnoses that cannot be classified as principal or secondary. These diagnosis codes may represent diagnoses originating from nonfacility claims associated with an inpatient stay, for example, a physician services claim submitted separately from the facility claim. Codes of this type generally come from claims-based Data Partners.

Eligible post-IGIV inpatient encounters with an AIS diagnosis code listed in any position (principal, secondary, or unspecified) were selected for review. For each potential AIS case identified, Sentinel Data Partners were asked to retrieve a medical chart corresponding to the encounter during which the AIS diagnosis was recorded. In this validation report, we restricted the denominator for our PPV calculations to the subsample of potential cases for whom we received a chart that was sufficiently complete to determine whether an AIS occurred (Fig. 1).

2.4. Chart abstraction

A trained nurse and stroke data abstractor (A.N. or E.R) reviewed the medical chart(s) associated with the AIS hospital encounter. The abstractors recorded information concerning symptom onset, relevant clinician notes, brain imaging studies, and other factors relevant for the IGIV-TEE safety assessment.

2.5. Case adjudication

Completed abstraction forms (and the original medical charts if needed) were reviewed by a vascular neurologist (E.L., N.N., or S. D.). The adjudication criteria were based on the 2013 American Heart Association/American Stroke Association definition of ischemic stroke,^[10] and the addition of a “possible” AIS category for cases where chart information was incomplete. Potential cases were adjudicated as a definite, probable, or possible AIS, no AIS, or as status unknown/insufficient information, as described below.

- Definite: To qualify as a definite AIS, a potential case was required to have documentation of an acute focal ischemic infarction of the central nervous system based on imaging, for example, computed tomography (CT) or magnetic resonance imaging (MRI), surgical or pathological findings.
- Probable: A case was counted as a “probable” AIS if there was rapid onset of neurologic deficit documented but CT/MRI was unavailable or done too early. The deficit must not have been secondary to brain hemorrhage, trauma, tumor, infection, or another identifiable mimic, and must have lasted more than 24 hours (unless death supervened).^[10]
- Possible: If neither imaging evidence nor clinical signs and symptoms consistent with AIS were documented in the chart, a physician diagnosis of AIS recorded in the chart was counted as a “possible” event.
- No AIS or AIS status unknown. Cases that fit none of the criteria above were classified as no AIS or status unknown/insufficient information based on the completeness of the patient’s medical chart.

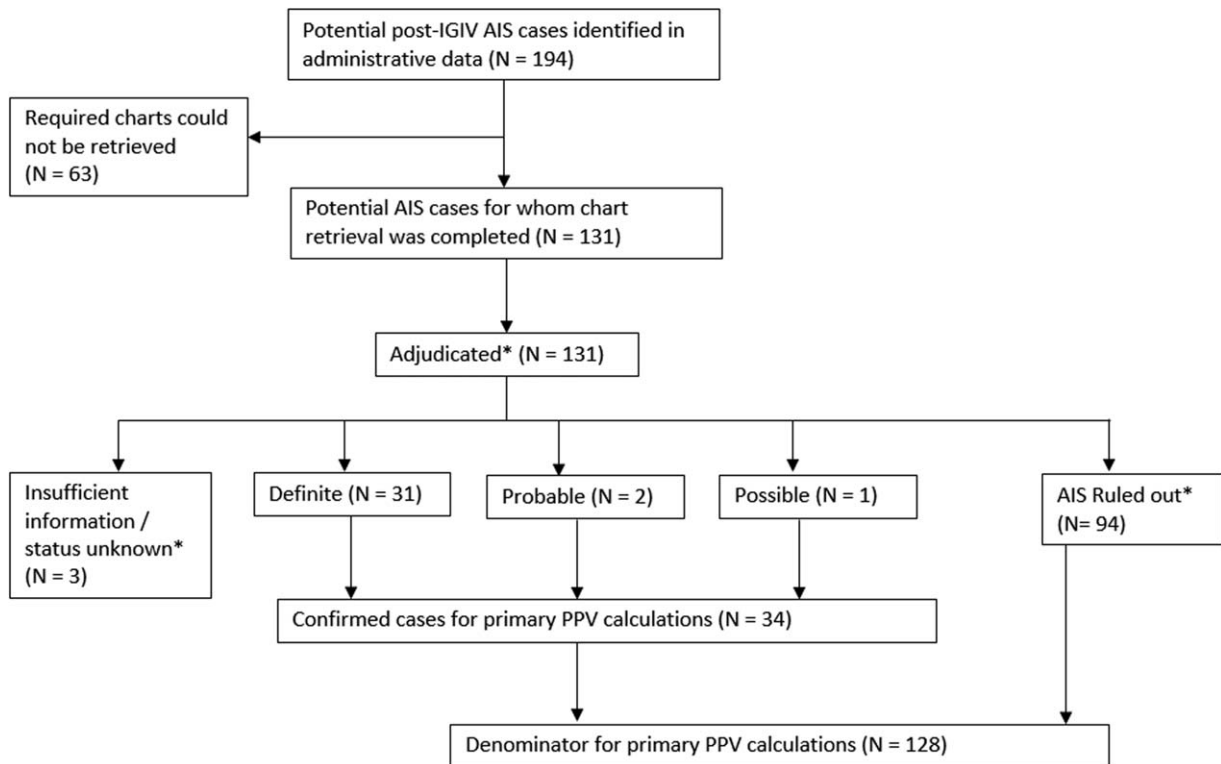


Figure 1. Disposition of potential acute ischemic stroke (AIS) cases identified in the SDD. In limited circumstances (see methods section), a potential case could be ruled out or classified as uninformative based on the judgment of the abstractor and was not physician-adjudicated. For AIS, this included 2 cases evaluated as insufficient information/status unknown and 3 cases as no AIS. * AIS = acute ischemic stroke, IGIV = intravenous immune globulin, PPV = positive predictive value.

For a small number of potential cases, the chart(s) received contained no recorded diagnosis of an acute AIS, no indication that an acute AIS was considered as part of a differential diagnosis, no diagnostic testing, and no symptoms suggestive of a possible AIS. These cases were flagged by the abstractors and not reviewed by the physician adjudicators due to resource constraints. For these cases, if the chart(s) received included the discharge summary for the index AIS hospital encounter, the potential case was considered to have been miscoded and classified as no AIS. Otherwise, the case was classified as having an unknown status due to chart incompleteness.

2.6. Positive predictive value (PPV) calculation

We calculated the PPV of the AIS diagnoses codes identified in the administrative data by dividing the number of confirmed AIS cases (definite, probable, or possible) by the total number of cases for whom a sufficiently complete chart was obtained. Potential cases adjudicated as AIS status unknown/insufficient information were removed from the denominator for the PPV calculation (Fig. 1). Exact binomial 95% confidence intervals (95% CI, Clopper–Pearson) were calculated for the PPV estimates to quantify their precision. Because the study sample was selected for chart review due to the presence of an administrative diagnosis code for AIS, we were unable to calculate sensitivity, specificity, or negative predictive value associated with AIS diagnosis codes.

3. Results

One hundred ninety-four potential post-IGIV AIS cases were identified in the SDD in 2006 to 2012; required charts could be obtained for 131 (68%) of these patients. Common reasons that charts were unavailable included an inability to map the encounter record in the SDD to patient and provider identifiers required for chart requests, an inability to locate the medical chart corresponding to the requested encounter, and refusal by the healthcare provider. (See Appendix Table A1 for a complete list of reasons

that charts were unobtainable, <http://links.lww.com/MD/C42>.) Of the 131 potential AIS cases for which charts were available, 100 were from claims-based Data Partners, and 31 from integrated care delivery systems. The median age of the patients was 65 years; 50% were female. On the basis of administrative diagnoses recorded during the 6 months prior, these patients had a high burden of risk factors for cerebrovascular disease: 15% had a prior ischemic stroke, 15% had atrial fibrillation, 10% had a prior myocardial infarction, and 69% had hypertension. Additional descriptive details on the patient sample are provided in Table 1.

Outcome status could be determined for 128 potential AIS cases, of which 34 were confirmed by physician adjudicators (Fig. 1). The PPVs for the inpatient AIS diagnoses recorded in the administrative data were 27% overall (34/128, 95% CI: 19–35), 60% (9/15, 95% CI: 32–84) for principal-position diagnoses, 42% (21/50, 95% CI: 28–57) for secondary diagnoses, and 6% (4/63, 95% CI: 2–15) for position-unspecified diagnoses. One patient was found to have a venous rather than arterial stroke; in accordance with the study protocol, this patient was counted as a false positive. PPVs were higher for ICD-9-CM diagnosis codes 433.x1 and 434.x1 than codes 434.x0 and 436. Detailed PPV estimates stratified by coding position, ICD-9-CM diagnosis code, Data Partner type, prior AIS diagnosis, and type of indication for IGIV are provided in Table 2.

4. Discussion

In this chart validation study, which relied on data from a protocol-based assessment of the risk of thromboembolic events following IGIV treatment,^[4] we evaluated the validity of inpatient administrative diagnosis codes for AIS within the SDD. PPVs were lower than anticipated: 60% for principal diagnoses, 42% for secondary diagnoses (though sample size limited the precision of this estimate), and only 6% for position-unspecified diagnoses. As discussed in more detail below, these PPV estimates were meaningfully lower than what has been reported in the majority of prior chart validation studies of administrative ICD-9-CM diagnosis codes for AIS.

Table 2

Positive predictive values (PPVs) associated with inpatient administrative diagnosis codes for acute ischemic stroke (AIS) by position.

	PPVs for all potential AIS cases (N=128)	PPVs for principal position AIS diagnoses (N=15)	PPVs for secondary AIS diagnoses (N=50)	PPVs for position-unspecified AIS diagnoses (N=63)
All AIS codes	27% (34/128, 95% CI: 19–35)	60% (9/15, 95% CI: 32–84)	42% (21/50, 95% CI: 28–57)	6% (4/63, 95% CI: 2–15)
By diagnosis code recorded in administrative data				
433.x1	50% (3/6, 95% CI: 12–88)	50% (1/2, 95% CI: 1–99)	100% (1/1, 95% CI: 3–100)	33% (1/3, 95% CI: 1–91)
434.x0	0% (0/9, 95% CI: 0–34)	0% (0/2, 95% CI: 0–84)	–	0% (0/7, 95% CI: 0–41)
434.x1	33% (31/95, 95% CI: 23–43)	73% (8/11, 95% CI: 39–94)	43% (20/47, 95% CI: 28–58)	8% (3/37, 95% CI: 2–22)
436	0% (0/18, 95% CI: 0–19)	–	0% (0/2, 95% CI: 0–84)	0% (0/16, 95% CI: 0–21)
Data Partner type				
Insurer/claims-based	19% (19/98, 95% CI: 12–29)	64% (7/11, 95% CI: 31–89)	33% (8/24, 95% CI: 16–55)	6% (4/63, 95% CI: 2–15)
Integrated care delivery systems	50% (15/30, 95% CI: 31–69)	50% (2/4, 95% CI: 7–93)	50% (13/26, 95% CI: 30–70)	–
By whether an AIS diagnosis code was observed in prior 183 d				
No prior AIS	28% (30/108, 95% CI: 20–37)	64% (9/14, 95% CI: 35–87)	41% (17/41, 95% CI: 26–58)	8% (4/53, 95% CI: 2–18)
Prior AIS	20% (4/20, 95% CI: 6–44)	0% (0/1, 95% CI: 0–98)	44% (4/9, 95% CI: 14–79)	0% (0/10, 95% CI: 0–31)
Possible autoimmune/inflammatory indication for IGIV use*				
Autoimmune/inflammatory indication	29% (25/85, 95% CI: 20–40)	54% (7/13, 95% CI: 25–81)	44% (15/34, 95% CI: 27–62)	8% (3/38, 95% CI: 2–21)
No autoimmune/inflammatory indication	21% (9/43, 95% CI: 10–36)	100% (2/2, 95% CI: 16–100)	38% (6/16, 95% CI: 15–65)	4% (1/25, 95% CI: 0–20)

Statistics reported in this table reflect the PPV of administrative ICD-9-CM AIS diagnosis codes for confirmed (definite, probable, or possible) acute AIS. Patients with a classification of insufficient information/acute TEE status unknown were removed from the denominator for the PPV calculations and not included in this table. AIS=acute ischemic stroke, IGIV=intravenous immune globulin, PPV=positive predictive value.

Consistent with a more recent validation study that included cases from after 2005, we found little use of ICD-9-CM code 436 during our study period (2006–2012) on inpatient facility claims (i.e., principal and secondary inpatient diagnoses).^[11] Stroke and cerebrovascular accident not otherwise specified were removed from the inclusion terms for ICD-9-CM code 436 in 2004, and added as exclusions.^[12] So although this code was included in many previously validated algorithms to identify stroke, and past studies found much higher PPVs for this code, the use and PPV of code 436 likely decreased as a result of this change in coding guidance.^[12]

Position-unspecified AIS diagnosis codes infrequently reflected true AIS events. Within the SDD, position-unspecified codes may represent diagnoses originating from non-facility claims associated with an inpatient stay (i.e., a separate physician/provider claim). If an encounter had a position-unspecified AIS diagnosis without an additional principal or secondary AIS diagnosis, a common explanation was that head imaging had been performed but AIS was ruled out. The exclusion of position-unspecified diagnoses would have improved the overall PPV estimate from 27% to 46%, at a cost of missing four of 34 true cases (12%).

In previous validation studies of inpatient diagnosis codes for AIS that were conducted in U.S. adult populations and used similar code ranges to identify AIS cases, reported PPVs have ranged from 73% to 94%, with principal diagnoses associated with somewhat higher PPVs than secondary diagnoses.^[2,13–22] These estimates are higher than the PPVs we found for principal and secondary diagnosis codes. With the exclusion of position-unspecified diagnosis codes, code 434.x0 (cerebral arterial occlusion without mention of infarction) and 436 (acute but ill-defined cerebrovascular disease), our principal and secondary diagnosis PPV estimates would be somewhat higher (69% and 44%, respectively), but still lower than previously reported PPVs.

There are a number of explanations for the lower PPVs reported in our study. As our sample was comprised of potential post-IGIV stroke cases, specific characteristics of that population may have contributed to lower PPVs. We initially hypothesized that false positives might be more common among patients whose indication for IGIV was a neurologic autoimmune or inflammatory condition, which could cause symptoms resembling those associated with a stroke. However, when we stratified on IGIV indication (autoimmune/inflammatory condition or other), we did not find lower PPVs among those patients (Table 2). Second, it is possible that ICD-9-CM PPVs were lower in our study period (2006–2012) than during earlier case identification periods evaluated in prior studies. In the updated stroke validation study conducted in the Atherosclerosis Risk In Communities (ARIC) cohort, PPVs associated with inpatient administrative diagnosis codes for AIS were somewhat lower during 2003 to 2006 and 2007 to 2010 (74% and 72%, respectively), compared with earlier periods (78% in 1991–1994, 79% in 1995–1998, and 85% in 1999–2002).^[13] Our results may be explained by some combination of the factors described above, sampling variability, and/or factors specific to the administrative data records in the SDD.

Our study had a number of important limitations. First, because the study sample was selected for chart review due to the presence of an administrative diagnosis code for AIS, we were unable to calculate sensitivity, specificity, or negative predictive value associated with AIS diagnosis codes. Second, we did not systematically collect data on AIS etiology during the chart validation process, and are unable to provide a breakdown on the number of confirmed stroke cases by etiology. Third, as discussed above, the generalizability of our AIS PPV estimates—

based on a sample of IGIV users—to the broader population of SDD health plan members is unclear.

Our results underscore the fact that the PPVs of administrative diagnosis codes can vary meaningfully due to factors such as the patient population, the nature of the administrative database and its relationship to the underlying claims data and health records, and differences in diagnostic and billing practices by time and place. Future research is needed to assess the validity of ICD-9-CM and ICD-10-CM administrative diagnosis codes for AIS in other patient populations within the SDD.

Acknowledgments

We thank the following individuals for their contributions to this research: Angela Overton, Erin Rindels, Cole Haskins, Michael Mueller, and Nicholas Rudzianski of the University of Iowa, Bruce Fireman of Kaiser Permanente Division of Research, and Meghan Baker and Casey Covarrubias of the Harvard Pilgrim Health Care Institute.

References

- [1] Snapshot of Database Statistics. Sentinel Coordinating Center, 2017. Available at: <https://www.sentinelinitiative.org/sentinel/snapshot-data-base-statistics>. Accessed September 10, 2017.
- [2] Andrade SE, Harrold LR, Tjia J, et al. A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data. *Pharmacoepidemiol Drug Saf* 2012;21 (Suppl 1):100–28.
- [3] ICD-9-CM Coding Guidelines: DRG 014—Intracranial Hemorrhage or Cerebral Infarction, DRG 559—Acute Ischemic Stroke with Use of Thrombolytic Agent. *Primaris*, 2006. Available at: http://www.primaris.org/sites/default/files/resources/HPMP/coding%20guidelines%20booklet_DRG_014_559.pdf. Accessed January 22, 2016.
- [4] The Mini-Sentinel Thromboembolic Events after Immunoglobulin Administration Workgroup. Mini-Sentinel Assessment Protocol: Thromboembolic Events After Immunoglobulin Administration: Version 3.0. Mini-Sentinel Coordinating Center, 2015. Available at: https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_Thromboembolic-Events-After-Immunoglobulin-Administration-Protocol_0.pdf. Accessed September 10, 2017.
- [5] Orange JS, Hossny EM, Weiler CR, et al. Use of intravenous immunoglobulin in human disease: a review of evidence by members of the Primary Immunodeficiency Committee of the American Academy of Allergy, Asthma and Immunology. *J Allergy Clin Immunol* 2006;117: S525–53.
- [6] McGraw D, Rosati K, Evans B. A policy framework for public health uses of electronic health data. *Pharmacoepidemiol Drug Saf* 2012;21 (Suppl 1):18–22.
- [7] Rosati K, Evans B, McGraw D. HIPAA and Common Rule Compliance in the Mini-Sentinel Pilot. Unpublished White Paper. Sentinel Operations Center, Boston, MA:2012.
- [8] Farrow S, Campion DM, Herrinton LJ, et al. The organizational structure and governing principles of the Food and Drug Administration's Mini-Sentinel pilot program. *Pharmacoepidemiol Drug Saf* 2012;21(Suppl 1):12–7.
- [9] Health Information Policy Council. 1984 revision of the Uniform Hospital Discharge Data Set: HHS. Notice. *Fed Regist* 1985; 50:31038–31040.
- [10] Sacco RL, Kasner SE, Broderick JP, et al. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2013;44:2064–89.
- [11] Olson KL, Wood MD, Delate T, et al. Positive predictive values of ICD-9 codes to identify patients with stroke or TIA. *Am J Manag Care* 2014;20: e27–34.
- [12] ICD-9-CM Coordination and Maintenance Committee. ICD-9-CM Coordination and Maintenance Committee Meeting Agenda, April 1–2, 2004. 2004. Available at: <https://www.cdc.gov/nchs/data/icd/agendaapril04-revised.pdf>. Accessed November 23, 2016.
- [13] Jones SA, Gottesman RF, Shahar E, et al. Validity of hospital discharge diagnosis codes for stroke: the Atherosclerosis Risk in Communities Study. *Stroke* 2014;45:3219–25.

- [14] Woodfield R, Grant I, Sudlow CL. Group UKBSO, Follow-Up UKB, Outcomes Working Group. Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: a systematic review from the UK Biobank Stroke Outcomes Group. *PLoS One* 2015;10:e0140533.
- [15] Wahl PM, Rodgers K, Schneeweiss S, et al. Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population. *Pharmacoepidemiol Drug Saf* 2010;19:596–603.
- [16] Benesch C, Witter DM Jr, Wilder AL, et al. Inaccuracy of the International Classification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. *Neurology* 1997;49:660–4.
- [17] Goldstein LB. Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke: effect of modifier codes. *Stroke* 1998;29:1602–4.
- [18] Thigpen JL, Dillon C, Forster KB, et al. Validity of international classification of disease codes to identify ischemic stroke and intracranial hemorrhage among individuals with associated diagnosis of atrial fibrillation. *Circ Cardiovasc Qual Outcomes* 2015;8:8–14.
- [19] Rosamond WD, Folsom AR, Chambless LE, et al. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. *Stroke* 1999;30:736–43.
- [20] Roumie CL, Mitchel E, Gideon PS, et al. Validation of ICD-9 codes with a high positive predictive value for incident strokes resulting in hospitalization using Medicaid health data. *Pharmacoepidemiol Drug Saf* 2008;17:20–6.
- [21] Lichtman JH, Leifheit-Limson EC, Goldstein LB. Centers for Medicare and Medicaid services medicare data and stroke research goldmine or landmine? *Stroke* 2015;46:598–604.
- [22] Lakshminarayan K, Larson JC, Virnig B, et al. Comparison of Medicare claims versus physician adjudication for identifying stroke outcomes in the women's health initiative. *Stroke* 2014;45:815–21.