Validity of International Classification of Disease Codes to Identify Ischemic Stroke and Intracranial Hemorrhage Among Individuals With Associated Diagnosis of Atrial Fibrillation

Jonathan L. Thigpen, PharmD; Chrisly Dillon, MD; Kristen B. Forster, MBA; Lori Henault, MPH; Emily K. Quinn, MA; Yorghos Tripodis, PhD; Peter B. Berger, MD; Elaine M. Hylek, MD; Nita A. Limdi, PhD

Background—Because of its association with death and disability, stroke is a focus of outcomes in atrial fibrillation (AF) research. International Classification of Disease-Ninth Revision (ICD-9) edition codes are commonly used to identify stroke in research, particularly in large administrative data. We sought to assess the validity of ICD-9 codes in stroke case ascertainment and for AF across 3 institutions.

Methods and Results—Participating centers included Boston Medical Center (safety net hospital), Geisinger Health System (rural Pennsylvania), and the University of Alabama at Birmingham (academic center in the southeastern stroke belt). ICD-9 codes for ischemic stroke (433–434, 436) and intracranial hemorrhage (430–432) identified 1812 stroke cases with an associated code for AF (427.31) from 2006 to 2010. Cases were vetted through chart review with final adjudication by a stroke neurologist. Review considered 94.2% of ICD-9 identified stroke cases valid with decreased accuracy for concurrent AF diagnosis (82.28%) and stroke attributable to AF (72.8%). Among events with “without infarction” modifiers, 7.2% were valid strokes. ICD-9 stroke code accuracy did not differ by stroke type or site. Stroke code 434 displayed higher accuracy than 433 (94.4% versus 85.2%; P<0.01), and primary stroke codes were more accurate than nonprimary codes (97.2% versus 83.7%; P<0.0001).

Conclusions—Using ICD-9 stroke and AF codes to identify patients with stroke plus AF resulted in inaccuracies. Given the expanded financial and policy implications of patient-oriented research, conclusions derived solely from administrative data without validation of outcome events should be interpreted with caution. (Circ Cardiovasc Qual Outcomes. 2015;8:00-00. DOI: 10.1161/CIRCOUTCOMES.113.000371.)

Key Words: atrial fibrillation ■ International Classification of Disease codes ■ intracranial hemorrhages ■ stroke
WHAT IS KNOWN

- Because of clinical and economic pressures, research strategies often rely on International Classification of Diseases diagnostic codes to capture covariates and outcomes.
- Little is known about the accuracy of stroke diagnostic codes in identifying ischemic and hemorrhagic stroke in the setting of atrial fibrillation.

WHAT THE STUDY ADDS

- Diagnostic codes alone have limited positive predictive values in identifying acute stroke in the setting of active atrial fibrillation.
- This indicates that manual verification is needed to confirm stroke events in the setting of atrial fibrillation so as to reduce biased estimates.
- With increasing reliance on administrative data in research, we must assess the potential implications associated with inaccurate covariate and outcome ascertainment.

considered the major safety net hospital for the city of Boston. Geisinger is a highly integrated healthcare system that serves a predominantly rural population. The University of Alabama is part of the southeastern stroke belt and cares for a diverse patient population. The study was approved by the institutional review board at each respective site.

Stroke events were identified using discharge ICD-9 codes (Table 1) for IS (with infarction 433.434, 436) or ICH (430–432) from hospital admissions during a 5-year period (2006–2010). To ensure the most comprehensive search, we included ICD-9 codes identified in any position—primary (first position) or other (any position other than the first). Among these ICD-9 identified stroke events, all those cases associated with an AF ICD-9 code (427.31) were subject to an in-depth medical record review. To validate the AF diagnosis, we required ECG evidence of AF during the stroke admission or within 6 months of the stroke admission if the AF was not permanent. If neither of these criterion was fulfilled, we sought ECG evidence within 90 days of the stroke discharge. A valid IS was defined as a focal neurological deficit of sudden onset that persisted for >24 hours, corresponded to a vascular territory, and was not explained by an accompanying ICD-9 trauma modifier. In addition, all stroke events were adjudicated by site investigators (P.B.B., E.M.H., N.A.L.) based on review of the medical record. Questionable cases were resolved by consensus of the adjudication committee and a stroke neurologist. All validated IS and ICH cases were assigned a modified Rankin Scale score by the site neurologist. ICD-9 codes were used to identify stroke and AF initially. Covariates, such as hypertension, diabetes mellitus, and heart failure, were abstracted manually, whereas most demographic covariates were automatically populated from electronic records. To assess errors of ICD-9 coding further, we also determined the accuracy of the “without infarction” modifier: 433.00, 433.10, 433.20, 433.30, 434.00, 434.10, and 434.90.

Statistical Analyses

Analyses were stratified according to event type (ie, IS or ICH) and event position (ie, primary or other). Validation rates were compared across the 3 clinical sites and across calendar years. For IS, we also assessed the accuracy of specific ICD-9 codes (433, 434, and 436). We subsequently assessed the validity of the AF diagnosis ICD-9 code (427.31) based on ECG. Final validation assessment excluded IS because of a different mechanism (ie, strokes determined by medical record review to be attributable to vascular procedures, tumors, infection, or vasculitis) and ICH secondary to a surgical procedure or major trauma that had not been initially excluded because of lack of an accompanying ICD-9 trauma code. To assess errors of ICD-9 coding further, we also determined the accuracy of the “without infarction” modifier.

Data are displayed as counts and percentages. Bivariate analyses were performed, and differences were assessed with χ² tests for event type, event coding position, clinical site, and IS code. We calculated the positive predictive value (PPV) for each of these categories. All analyses were performed using SAS software (version 9.3; SAS Institute, Cary, NC).

Results

Our electronic search identified a total of 1812 stroke events with AF diagnosis (1358 IS and 454 ICH). Subsequent steps in the case-validation algorithm are displayed in the Figure. Manual review confirmed 1706 (94.2%) events to be valid stroke events. After excluding patients without ECG-confirmed AF (n=217), 1489 patients with manually confirmed stroke and AF remained (1136 IS and 353 ICH). The PPV of ICD-9 stroke codes alone was 94.2%, and when combined with validation of ICD-9 AF codes, the total accuracy dropped to 82.2%. Additional manual review excluded IS events because of a different mechanism (eg, vascular procedure, endocarditis, tumor, sepsis, and vasculitis) or ICH resulting from major head trauma with skull fracture. These additional trauma-related ICH cases were not initially filtered because they lacked the corresponding trauma codes. This final step yielded a total of 1320 strokes (1000 IS and 320 ICH) confirmed in the setting of AF (Table 2). The accuracy of ICD-9 stroke codes after these validation exercises fell to 72.8%. The final cohort included 33 events with no mutually exclusive coding position (ie, had a combination of IS and ICH codes).

The validation rates of ICD-9 stroke codes were similar across event type, IS (94.0%), and ICH (94.7%; P=0.55; Table 3). Validation rates were also similar across the 3 clinical sites: Boston Medical Center (94.9%), Geisinger Health System (93.3%), and University of Alabama (94.3%; P=0.55), and year of stroke admission. Approximately 15% of events designated as 433 were not valid IS. When compared with stroke code 434, this difference was statistically significant (85.2% versus 94.4%; P=0.003). Accuracy of stroke case ascertainment was influenced by the ICD-9 coding position.

Table 1. List of ICD-9 Codes Used to Identify Stroke Events

<table>
<thead>
<tr>
<th>Condition</th>
<th>ICD-9 Code</th>
<th>ICD-9 Code Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ischemic stroke</td>
<td>433</td>
<td>Occlusion and stenosis of precerebral arteries</td>
</tr>
<tr>
<td></td>
<td>434</td>
<td>Occlusion of cerebral arteries</td>
</tr>
<tr>
<td></td>
<td>436</td>
<td>Acute but ill-defined cerebrovascular disease</td>
</tr>
<tr>
<td>Intracranial hemorrhage</td>
<td>430</td>
<td>Subarachnoid hemorrhage</td>
</tr>
<tr>
<td></td>
<td>431</td>
<td>Intracerebral hemorrhage</td>
</tr>
<tr>
<td></td>
<td>432</td>
<td>Other and unspecified intracranial hemorrhage</td>
</tr>
</tbody>
</table>

ICD-9 indicates International Classification of Disease.
Discussion

In this study, we found less than optimal accuracy of ICD-9-coded stroke events when compared with the gold standard of medical record primary data review. For both IS with infarction and ICH, the proportion of events meeting the respective definitions was 94.2%. This was remarkably consistent across 3 sites. The accuracy of case identification was higher (97.2%) for events with the stroke-related ICD-9 code in the primary position when compared with those that were identified with codes in the nonprimary position (83.7%). However, reliance on a search strategy that identifies cases based solely on ICD-9 codes can be a viable screening tool if used correctly but should not be used as the sole method of case-validation. The ramifications of these observations are significant. Stroke research identifies the gaps in care that drive health policy and resource allocation. Given the emphasis on performance measures and public reporting, the financial implications are tangible, particularly for those institutions without sufficient margin to invest in the data systems and personnel training that are requisite to improve accuracy.

For ICD-9 stroke codes, a PPV ≥ 0.85% has been suggested to be adequately accurate for research purposes. Assuming non-differential misclassification, similar high PPVs (ie, >90%) for outcome capture will pose little bias on the relative estimate but should not be used as the sole method of case-validation. The ramifications of these observations are significant. Stroke research identifies the gaps in care that drive health policy and resource allocation. Given the emphasis on performance measures and public reporting, the financial implications are tangible, particularly for those institutions without sufficient margin to invest in the data systems and personnel training that are requisite to improve accuracy.

In this study, we found less than optimal accuracy of ICD-9-coded stroke events when compared with the gold standard of medical record primary data review. For both IS with infarction and ICH, the proportion of events meeting the respective definitions was 94.2%. This was remarkably consistent across 3 sites. The accuracy of case identification was higher (97.2%) for events with the stroke-related ICD-9 code in the primary position when compared with those that were identified with codes in the nonprimary position (83.7%). However, reliance on a search strategy that identifies cases based solely on ICD-9 codes can be a viable screening tool if used correctly but should not be used as the sole method of case-validation. The ramifications of these observations are significant. Stroke research identifies the gaps in care that drive health policy and resource allocation. Given the emphasis on performance measures and public reporting, the financial implications are tangible, particularly for those institutions without sufficient margin to invest in the data systems and personnel training that are requisite to improve accuracy.

For ICD-9 stroke codes, a PPV ≥ 0.85% has been suggested to be adequately accurate for research purposes. Assuming non-differential misclassification, similar high PPVs (ie, >90%) for outcome capture will pose little bias on the relative estimates and may be considered a target threshold for the use of administrative data for comparative effectiveness research for stroke prevention in AF. Bias in the risk estimate is also linked to a test’s specificity more so than sensitivity with decreased specificity resulting in increased estimate bias. Most bias is observed when specificity is ≤ 0.85%. Furthermore, decreasing the PPV will reduce specificity thereby increasing bias in the estimate. We can also assume that even a minimal amount of misclassification leads to a bias in utilization studies and outcome incidence estimates.

The first phase of validation in our study, ICD-9 stroke code accuracy (94.2%), is similar to that previously reported.
similar methodology has led to a highly variable IS event of expertise may be considerably less. Previous research using to surmise that the accuracy of coding found outside centers in this study are all recognized stroke centers, it is reasonable

<table>
<thead>
<tr>
<th>Event coding position, n (%)</th>
<th>Ischemic Stroke (n=1000)</th>
<th>Intracranial Hemorrhage (n=320)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>569 (86.9%)</td>
<td>281 (87.8%)</td>
</tr>
<tr>
<td>Other</td>
<td>109 (10.9%)</td>
<td>28 (8.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study site, n (%)</th>
<th>Ischemic Stroke (n=1000)</th>
<th>Intracranial Hemorrhage (n=320)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston Medical Center</td>
<td>230 (23.0%)</td>
<td>70 (21.9%)</td>
</tr>
<tr>
<td>Geisinger Health System</td>
<td>370 (37.0%)</td>
<td>105 (32.8%)</td>
</tr>
<tr>
<td>University of Alabama</td>
<td>400 (40.0%)</td>
<td>145 (45.3%)</td>
</tr>
</tbody>
</table>

*There were 33 events with combinations of stroke codes recorded, so there was no mutually exclusive stroke coding position.

This is not necessarily surprising given that our study had characteristics associated with high ICD-9 accuracy, including using hospital discharge codes, having a majority of events coded in the primary position, having an overwhelming majority of IS (93%) coded with the more accurate 434 stroke code as opposed to the 433 code, and excluding without infarct modifier codes. In addition, given that the 3 participating centers in this study are all recognized stroke centers, it is reasonable to surmise that the accuracy of coding found outside centers of expertise may be considerably less. Previous research using similar methodology has led to a highly variable IS event capture accuracy, varying from 62% to 85%. In addition, informed interpretation of previous IS research is challenged by the inclusion of transient ischemic attack or the nonspecific ICD-9 code 438 (late effects of cerebrovascular disease).

The second phase of our validation sought confirmation of the AF diagnosis. Despite AF being listed among the associated diagnoses, medical record documentation of AF or ECG confirmation of AF was lacking in 12% of stroke cases, indicating that AF ICD-9 codes had a relatively high PPV of 88% in patients with confirmed stroke. Possible explanations for this finding include error in ECG interpretation, propagation of the AF code after a remote transient episode, or physician use of this code to rule out AF and to justify an ECG in the setting of an irregular pulse. A recent systematic review described the PPV of AF ICD-9 codes that ranged from 70 to 96%. Accuracies of AF validation methods differed according to factors such as the number of documented AF codes required, the period of time searched for AF codes, and whether ECGs were used to confirm AF. In general, more strict validation methods resulted in lower PPVs of AF ICD-9 codes. Although not validated in the literature, a validation algorithm incorporating inpatient and outpatient AF diagnosis codes with ECG confirmation may be best to confirm incident or prevalent AF. Strategies based on multiple diagnoses or ECGs potentially could miss the heightened risk of new onset AF.

Accuracy of both stroke and AF codes is critically important for performance measurement and medication postmarketing surveillance, eg, recently introduced novel oral anticoagulants (apixaban, rivaroxaban, and dabigatran). Rigorously conducted comparative effectiveness studies of these agents in clinical practice will be contingent on validation of AF, IS, and ICH. In our population, the 72.8% PPV for ICD-9 codes describing stroke associated with AF indicates modest usefulness of ICD-9 codes alone; however, a higher PPV (and increased usefulness) may be expected when assessing these codes in a population entirely comprised of patients with AF on oral anticoagulants with previous instances of ICD-9 confirmed AF. Interpretation of results across different patient populations will require adjustment for key covariates, which raises additional challenges. A previous study reports additional stroke covariates, such as hypertension (85%), diabetes mellitus (97%), and coronary artery disease (88%) have high predictive values, whereas more difficult to obtain information such as history of cerebrovascular accident (59%) or tobacco use (58%) have lower accuracy. Similarly, another study reported high accuracies of coronary artery disease (96%), diabetes mellitus (98%), and hypertension (97%), but a low accuracy of deep vein thrombosis (72%). Similar to stroke, the accuracy of ICD-9 codes for stroke covariates seems to be highly variable, which confirms our belief that as of currently, a more conservative approach using manual review in addition to ICD-9 codes in retrospective research is necessary to ensure more accurate stroke covariate ascertainment.

Furthermore, manual verification is especially prudent when identifying fluctuating covariates, such as blood pressure or renal function, and key confounding factors that are not reliably captured in electronic health records, such as nonprescription aspirin use, medication interruption for procedures, or medication adherence.

Newer ICD-10 codes are not currently used in hospital discharges in the United States. These codes are considered more specific and provide a more intuitive coding method compared with currently used ICD-9 codes. About stroke, ICD-10 codes specify the hemorrhage location and source in ICH, differentiate between thrombotic and embolic IS, and include codes for intraoperative and postprocedural strokes. Data comparing the accuracy of ICD-9 and ICD-10 stroke codes are limited, and research indicates the same accuracy for IS ICD-9 (85%)
and ICD-10 (85%) codes and similar accuracies of ICH ICD-9 (97%) and ICD-10 (98%) and subarachnoid hemorrhage ICD-9 (98%) and ICD-10 (91%) codes. Another study shows ICD-10 codes for ICH and subarachnoid hemorrhage having PPVs of ≈96%. Until ICD-10 stroke codes accuracies are further determined and compared with ICD-9 stroke codes, manual review of all or a subset of events seems to be warranted based on the results of our study.

Our study has several strengths, including its multisite nature, diverse patient population, and the comprehensiveness of our manual medical record reviews. Stroke events were adjudicated by the site principle investigator with modified Rankin Scale score assignment by the site neurologist. Questionable stroke cases were vetted through our adjudication committee with ultimate assignment by a stroke neurologist. We conducted comprehensive searches of outpatient, inpatient, and transferring hospital ECGs to validate the AF code. Additional strengths include the comprehensive assessment of commonly used stroke codes and coding position, and chart reviews of the “without infarction” group. Furthermore, we reviewed and assessed ICD-9 accuracies for a large number of ICH (n=454) cases that significantly contributes to the current limited literature.

As mentioned previously, although we included 3 different and complementary sites, our findings may not reflect the coding patterns or coding accuracy of other settings. Although we searched all medical records within 6 months of stroke admission and 90 days after stroke discharge for ECG evidence of AF, subclinical AF cannot be ruled out as the underlying mechanism. Our patient cohort includes patients with concurrent stroke and AF ICD-9 codes; therefore, we cannot calculate PPV for stroke and AF codes independent from each other. In addition, we sought to include the first valid stroke for patients with multiple events. Although not likely impactful, this may overestimate our PPV. We cannot calculate the sensitivity of ICD-9 stroke codes because we did not review patients with diagnosis that may contribute additional cases of false-negatives (ie, TIA). Similarly, we are unable to report meaningful specificities of stroke codes because our estimates of nonvalid stroke cases were restricted to patients assigned with stroke-related codes (430–432, 433–434, and 436). Finally, we only assess PPV; therefore, we cannot calculate the prevalence of stroke in our cohort. This limits our ability to infer how changes in disease prevalence may alter our PPV.

In conclusion, screening ICD-9 codes is common in stroke-related research to identify outcome events. We identified ICD-9–coded IS and ICH events across 3 different stroke centers, and then implemented a thorough manual chart review to

<p>| Table 3. Percentage of Valid Stroke Events Using International Classification of Disease-Ninth Revision Coding Alone by Event Type, Coding Position, Clinical Site, and Ischemic Stroke Code |
|---------------------------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Event type</th>
<th>Valid (%)</th>
<th>Not Valid (%)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event coding position</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>1368 (77.4%)</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Other</td>
<td>399 (22.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston Medical Center</td>
<td>433 (23.9%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geisinger Health System</td>
<td>538 (29.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Alabama</td>
<td>841 (46.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischemic stroke code*</td>
<td>433</td>
<td></td>
<td>0.0034</td>
</tr>
<tr>
<td></td>
<td>61 (4.6%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1265 (95.4%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI indicates confidence interval; and ICH, intracranial hemorrhage.

*There were 2 events coded with International Classification of Disease-Ninth Revision ischemic stroke code 436 (60% valid); there were 30 valid ischemic stroke events coded with combinations of ischemic stroke (IS) codes or intracranial hemorrhage and IS codes. Comparing codes 433, 434, and 436 resulted in a Fisher exact test P value of 0.0103. We present the P value for code 433 vs code 434.
consider whether the stroke events were valid. The accuracy of combined ICD-9 codes for stroke and AF to identify valid stroke accompanied by AF was 72.8%. Furthermore, review of “without infarction” ICD-9-coded events considered 7.2% to be a valid acute stroke. These findings highlight the need for more rigorous methods to validate outcome events identified by ICD-9 codes in large administrative databases.

Sources of Funding
This study was supported, in part, by grants from the National Institute of Neurological Disorders and Stroke (R01 NS070307) and National Heart Lung and Blood Institute (R01 HL092173).

Disclosures
None.

References
7._maps.2011.11.011.


Validity of International Classification of Disease Codes to Identify Ischemic Stroke and Intracranial Hemorrhage Among Individuals With Associated Diagnosis of Atrial Fibrillation

Jonathan L. Thigpen, Chrisly Dillon, Kristen B. Forster, Lori Henault, Emily K. Quinn, Yorghos Tripodis, Peter B. Berger, Elaine M. Hylek and Nita A. Limdi

Circ Cardiovasc Qual Outcomes. published online January 13, 2015;
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/early/2015/01/13/CIRCOUTCOMES.113.000371

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org/subscriptions/