Published in final edited form as:
Stroke. 2014 March ; 45(3): 815–821. doi:10.1161/STROKEAHA.113.003408.

Comparison of Medicare Claims vs. Physician Adjudication for Identifying Stroke Outcomes in the Women's Health Initiative

Kamakshi Lakshminarayan, MBBS, PhD[1], **Joseph C. Larson, MS**[2], **Beth Virnig, PhD**[3], **Candace Fuller, MPH**[1], **Norrina Bai Allen, PhD, MPH**[4], **Marian Limacher, MD**[5], **Wolfgang C. Winkelmayer, MD, ScD**[6], **Monika M. Safford, MD**[7], and **Dale R. Burwen, MD, MPH**[8]

[1]University of Minnesota, School of Public Health, Division of Epidemiology and Community Health

[2]Women's Health Initiative Clinical Coordinating Center, Fred Hutchinson Cancer Research Center, Seattle, WA

[3]University of Minnesota, School of Public Health

[4]Northwestern University, Department of Preventive Medicine, Feinberg School of Medicine

[5]University of Florida, Division of Cardiovascular Medicine

[6]Stanford University, School of Medicine, Department of Medicine, Division of Nephrology

[7]University of Alabama, Department of Medicine, Division of Preventive Medicine

[8]National Institutes of Health, National Heart, Lung, and Blood Institute

Abstract

Background and Purpose—Many studies use medical record review for ascertaining outcomes. One large, longitudinal study, the Women's Health Initiative (WHI) ascertains strokes using participant self-report and subsequent physician review of medical records. This is resource-intensive. Herein, we assess whether Medicare data can reliably assess stroke events in the WHI.

Methods—Subjects were WHI participants with fee-for-service Medicare. Four stroke definitions were created for Medicare data using discharge diagnoses in hospitalization claims. Definition 1: stroke codes in any position; Definition 2: primary position stroke codes; Definitions 3 & 4: hemorrhagic and ischemic stroke codes respectively. WHI data were randomly split into training (50%) and test sets. A concordance matrix was used to examine agreement between WHI and Medicare stroke diagnosis. A WHI stroke and a Medicare stroke were considered a match if they occurred within +/− 7 days of each other. Refined analyses excluded Medicare events where medical records were unavailable for comparison.

Results—Training data (n=24,428): There were 577 WHI strokes and 557 Medicare strokes using definition 1. Of these, 478 were a match. Algorithm performance: Specificity 99.7%;

Corresponding Author: Kamakshi Lakshminarayan, MD, PhD, University of Minnesota, School of Public Health, Division of Epidemiology and Community Health, 1300 S. Second St., Suite 300, Minneapolis, MN 55454, Tel: 612-624-9492; Fax: 612-624-0315, laksh004@umn.edu.

CONFLICT(S) OF INTEREST/DISCLOSURES
None

Negative Predictive Value 99.7%; Sensitivity 82.8%; Positive Predictive Value 85.8%; kappa 0.84. Performance was similar for test data. While specificity and negative predictive value exceeded 99%, sensitivity ranged from 75 to 88% and positive predictive value ranged from 80 to 90% across stroke definitions.

Conclusion—Medicare data appear useful for population-based stroke research; however the performance characteristics depend on the definition selected.

INTRODUCTION

Medicare claims provide a nationwide data source for individuals over age 65 that is reasonably representative of the population. Ongoing collection of these medical data allows for a variety of secondary uses in public health and health services research, including disease surveillance, tracking of patient outcomes and health care utilization. Randomized clinical trials and major prospective observational cohort studies have traditionally relied on intensive data collection processes including medical record review to ascertain outcomes. These rigorous approaches are generally considered necessary to yield accurate outcome information. The extent to which Medicare administrative data may be useful to ascertain outcomes in this context is unclear, but is important to evaluate because of the cost-efficiencies of secondary data use and the augmented research potential from the comprehensive data collected by Medicare regarding healthcare utilization and expenditures.

The Women's Health Initiative (WHI) is a national longitudinal study of 161,808 women 50–79 years old that evaluated strategies for preventing major causes of morbidity and mortality including cardiovascular disease and stroke.[1] Like many large studies, the WHI outcomes adjudication involves medical record review by physicians. This is resource intensive. Chronic disease outcomes increase substantially with age, necessitating ever escalating resources for outcome ascertainment. Most current WHI participants are enrolled in Medicare at baseline or subsequently during follow-up. To evaluate whether Medicare data may be useful for outcome ascertainment, the WHI program initiated a validation effort to assess the agreement between Medicare and WHI data for cardiovascular disease outcomes. This report focuses on stroke.

Diagnostic codes pertaining to stroke in healthcare administrative databases have been reported to have variable performance when compared to clinical definitions of stroke.[2–9] A recent systematic review found that some algorithms for stroke and intracranial bleeds had positive predictive values greater than 80%. Other metrics including sensitivity were less frequently reported.[2] Confirmation criteria for stroke events varied substantially across studies. Although occasional studies involved neurologists to confirm strokes[2, 5] they were typically geographically limited or included a relatively small number of events. Despite the incidence of stroke being the greatest in the elderly, there is limited published information comparing Medicare data with neurologist adjudicated strokes.[6] The WHI uses vascular neurologists to perform stroke adjudication. Medical record review by these specialists who have substantial expertise regarding the diagnostic nuances of stroke and relevant neuroimaging is intended to identify strokes with a high level of accuracy. Also, cases which were not clear cut were discussed by the committee of vascular neurologist adjudicators to develop consensus. The WHI is also geographically diverse, and the number of neurologist-confirmed adult strokes is larger than any previous report used for validation studies of administrative data. Hence the WHI-Medicare linkage is a valuable data source for evaluating the performance of Medicare data.

The goal of this study is to assess whether Medicare data can be reliably used to assess stroke events in the WHI. We developed algorithms that used variables in Medicare claims data to define stroke and examined algorithm performance for the accurate detection of stroke hospitalizations using clinically defined strokes in the WHI. The results presented herein have implications for future large clinical trials as exemplified by the WHI and also for the broader agenda of population science and health services research.

METHODS

Study population

The WHI enrolled 161,808 women 50–79 years old from 1993 until 1998 in a set of randomized clinical trials (68,132 participants) as well as an observational study (93,676 participants). The observational and clinical trial studies ran until 2005 at which time, the women were invited to participate in the WHI extension study through 2010 (1^{st} extension). Longitudinal follow-up continues and current participants have 14–19 years of follow up to date. The WHI was conducted in 24 states across the United States using 40 field centers. Field center locations are indicated at https://cleo.whi.org/about/SitePages/Recruitment.aspx. The WHI database has been linked to Medicare data from the Centers for Medicare & Medicaid Services (CMS).

We used incident stroke events between the start of WHI in 1993 through 2007. We included women from the Observational Study (OS) who either had Medicare Parts A & B, fee-for-service (FFS) coverage at the time of WHI enrollment (1993–1998) or later met age criteria for Medicare enrollment through 2007. Women were excluded if they were enrolled in a Medicare managed care plan at the time of WHI enrollment and were censored at the time of entry to managed care and when they lost Medicare FFS eligibility. Women who experienced a WHI adjudicated stroke outcome prior to Medicare eligibility were excluded since the WHI adjudicated only the first stroke outcome. In the event based analysis, participants were censored from observation 7 days after the WHI stroke allowing for a 7-day match window between WHI and Medicare. The study population was randomly split into a training data set and a test data set (50% each) to replicate the results of the algorithm.

Stroke events in the WHI

The WHI process uses self-report via annual questionnaires completed by participants (or their proxies) who are asked if they were hospitalized overnight since their last report. Using the hospitalization information in the participant self-report, medical records were obtained and subsequently adjudicated by trained vascular neurologists based on available medical record documentation using detailed standards.[10] Details of the record request process are in Supplemental Methods. Medical records were typically inpatient hospitalization records and included admission and discharge notes, emergency room notes, neurology consultations, therapy (physical, occupational, speech) evaluations and imaging results (e.g., brain, cerebrovascular and cardiac studies). The WHI adjudicated events included hospitalizations with overnight inpatient stays. Hence, emergency room visits that did not lead to a hospital admission were excluded. For the WHI, stroke was defined as *"the rapid onset of a persistent neurologic deficit attributed to an obstruction or rupture of the brain arterial system (including stroke occurring during or resulting from a procedure). The deficit was not known to be secondary to brain trauma, tumor, infection, or other cause. The deficit had to last more than 24 hours unless death supervened or there was a demonstrable lesion compatible with an acute stroke on CT or MRI. A stroke was defined as procedure-related if it occurred within 24 hours after any procedure or within 30 days after a cardio-version or invasive cardiovascular procedure."* Venous infarcts, traumatic brain injury, and subdural and epidural hemorrhages were specifically not considered strokes. The WHI collected data on only the first stroke event for each participant – (i.e., once a participant had a confirmed stroke outcome, further stroke events were not adjudicated).

Defining stroke in Medicare data

Stroke hospitalizations in Medicare data were defined using variables in the MedPAR (Medicare Provider Analysis and Review) files. We used MedPAR since our analysis was focused on hospitalized stroke and MedPAR files have claims data on inpatient hospital stays. Each MedPAR record has claims from a single hospital stay and has up to 10 International Classification of Diseases (9th Revision: ICD-9-CM) discharge diagnosis codes indicating principal diagnosis in the primary position or co-existing conditions in subsequent or secondary positions. Four stroke definitions were created using a preliminary set of these codes. The preliminary set of ICD-9-CM codes was based on a review of published studies which identified acute stroke in administrative data as well as our prior experience with population-based stroke surveillance.[2,5,11] The most comprehensive definition was the *all stroke* definition which included ICD-9 430.xx, 431.xx, 433.x1, 434.x1, 436.xx, 437.1x and 437.9x in any diagnostic position while the *primary position (indicating stroke as the principal diagnosis or reason for hospitalization) stroke* definition included the same codes but limited to the first position. *Ischemic stroke* was defined as ICD-9 433.x1, 434.x1, 436.xx, 437.1x and 437.9x in any diagnostic position *and hemorrhagic stroke* was defined as ICD-9 430.xx and 431.xx in any diagnostic position. Claims not meeting these stroke definitions were classified as non-strokes. Claims with ICD-9 codes pertaining to transient ischemic attacks (TIA) were not considered strokes. Using the training data set, we ultimately refined the definitions by omitting codes 437.1x and 437.9x as described below. All tables use final definitions that omit these codes. We also evaluated whether the coding algorithms could be further optimized by using rehabilitative therapy charges (i.e., non-zero charge amounts for physical, occupational, and/or speech therapy during the same hospital stay) as an added criterion to define stroke. Preliminary analyses suggested that this would likely increase specificity at the expense of a decrease in sensitivity which was non-trivial since 20% of matched events did not have these charges. Therefore, these therapy charges were not used further for defining stroke. We explored the use of codes indicating iatrogenic stroke and this is described in the online supplement.

Analysis

For the main analysis, the analytic unit was any hospitalization event. Included were confirmed WHI strokes after neurologist adjudication and all hospitalization claims from the Medicare MedPAR file including stroke and non-stroke hospitalizations. Women who did not have any hospitalizations recorded in the Medicare data, and who did not have a WHI confirmed stroke, were excluded from this analysis. The event-based analysis was especially intended to inform whether a hospitalization with stroke diagnosis codes in Medicare data was likely to represent a true stroke, and the completeness of ascertainment. For each participant, Medicare stroke events within a 7-day time window (+/− 7 days using admission dates) were regarded as a single stroke hospitalization. The first analytic step matched WHI and Medicare events, populating a 2X2 concordance matrix (Table 1). A Medicare stroke and a WHI stroke were considered a match if they occurred within +/− 7 days of each other. The 7-day match window was selected based on prior work showing that the post-discharge readmission rate due to recurrent stroke at 7 days was low at 0.3% (95% CI 0–0.7). [12] For both Medicare and WHI data, we used the admission date as the event date. We defined a WHI diagnosis of stroke as the reference standard, and WHI vs. Medicare concordance was evaluated using the kappa statistic, as well as by positive and negative predictive value (PPV, NPV) and sensitivity and specificity. The discordant cells were evaluated in greater detail by examining reasons for a non-match. In order to understand the performance of Medicare data when adjudicated hospital medical records were available for comparison, we analyzed concordance-discordance after excluding events which were not informative (e.g., no medical records were received and hence the Medicare event could not be judged to be true or false).

While our main analysis was event-based, we also performed a person-based analysis to evaluate usefulness of Medicare data for ascertainment of incident strokes, because this is often of interest in clinical trials and cohort studies. For the person-based analysis, we assessed whether each participant had a WHI stroke or not and similarly a Medicare stroke or not at any time during the overlapping follow-up period. Concordance was examined on the Medicare vs. WHI stroke status for each participant. If a participant had multiple hospitalizations with stroke codes in the Medicare data, the first hospitalization was used for any analyses requiring dates.

The analysis was first performed on the training set to allow potential fine-tuning of the coding algorithms, and then repeated on the test (validation) data set.

RESULTS

A total of 48,877 WHI-OS participants met inclusion criteria for the stroke validation study. Of these, 24,428 randomly selected participants were used for algorithm development and the remaining 24,422 formed the test/validation set. There were no significant differences in demographic characteristics between the training and validation datasets.

Training results (Tables 1 and 2)

A total of 31,399 Medicare hospitalizations (among 24,428 participants) were analyzed in the event-based analysis (Table 1). Using WHI data, there were 582 strokes. Using Medicare data and the most general stroke definition (stroke diagnosis code in any position), there were 796 strokes. Of these, 478 were found in both the WHI and Medicare databases with a date within +/− 7 days of each other.

Among the 104 WHI strokes that were not found in Medicare data, five of the women were not hospitalized, i.e. they were managed in the outpatient setting only. Hence, they were not identifiable in the Medicare data of inpatient stays. For the remaining 99 WHI strokes, a key reason for discordance was that WHI picked up many stroke cases which were discharged with non-stroke diagnosis codes in the Medicare claims. We found a total of 198 distinct Medicare non-stroke discharge codes among hospitalizations within +/− 7 days of the 99 WHI stroke events (since there could be multiple discharge codes and multiple non-stroke events in that +/− 7 day period). Review of all diagnosis codes in these hospital claims revealed a wide variety of conditions. The most frequent codes included diagnoses of essential hypertension (ICD-9 CM 401.9, 33 events), atrial fibrillation (ICD-9 CM 427.31, 21 events), diseases of the urinary system (ICD-9 599.0, 14 events), unspecified transient ischemic attacks (ICD-9 435.9, 12 events), and heart failure (428.0, 11 events). We did not find any predominant code that could have been added to our stroke definition to improve overall performance.

Among the 318 events found in Medicare but not WHI data, 182 (57%) had no record of a corresponding hospitalization (for any medical condition) reported to the WHI program. Of these, 62 (34%) participants had died within 365 days after the Medicare hospitalization. (In contrast, only 18.6% (89 of 478) of participants who had matching WHI and Medicare stroke events had died within 365 days of the hospitalization). Of the 136 Medicare events that did have a corresponding WHI hospitalization reported, medical records were adjudicated by the stroke committee for only 79. Reasons for not being adjudicated included 11 events for which no medical records were received due to administrative reasons (e.g., no signed release of records or no documents received), and 46 events for which no stroke adjudication was attempted (e.g., reason reported by the participant for the hospitalization was not suggestive of a WHI outcome of interest). Among the 79 records adjudicated by the WHI stroke committee and not found to be stroke, 27 (34%) were found to be transient ischemic attacks.

Specificity (99.0%) and NPV (99.7%) were high for Medicare ascertainment of stroke. Sensitivity on the other hand was more modest at 82.1%. PPV was low at 60.1% in the initial analysis. When events that were discordant because of outpatient strokes or lack of adjudicated medical records were excluded, the PPV increased to 85.8% while sensitivity remained largely unchanged (82.8%). The exclusion of events for which no WHI adjudication was performed improved the “WHI no, CMS yes” discordant cell (Table 1) and algorithm performance as measured by the kappa statistic increased from 0.69 indicating moderate agreement to 0.84 indicating high agreement. The pattern of results was similar for all stroke definitions.

Compared to using diagnosis codes in any position to define stroke, use of only the principal diagnosis increased the PPV from 60.1% to 64.4% when all records were included, and remained unchanged at 85.8% when only events with adjudicated medical records were included. Sensitivity of ascertaining inpatient strokes decreased from 82.8% to 74.4%.

Among the 582 WHI confirmed strokes, 109 (18.7%) were hemorrhagic, 453 (77.8%) were ischemic, and 20 (3.4%) were unknown type. The hemorrhagic stroke coding algorithm had the best PPV of all stroke definitions (91.1% when limited to events with adjudicated medical records) though the sensitivity of this algorithm was lower (75.9%). In comparison, the ischemic stroke coding algorithm had a PPV of 79.4% and a sensitivity of 82.2%.

Person-based analyses (Table 2) showed modestly improved concordance compared to the event-based results. Using the most general stroke definition (codes in any diagnosis position), 505 participants had a stroke in both WHI and Medicare data during the overlapping follow-up period. For most of these (83%), the date matched exactly in the two data systems; 88% were within +/− 1 day, 91% within +/− 3 days, 95% within +/− 7 days, 96% within +/− 30 days.

We examined the robustness of the person-based results by expanding the match window for discordant events to +/− 30 days (Table I, available online) as done in other WHI-Medicare validation studies. Specifically, among the 240 participants with a Medicare stroke diagnosis without a WHI stroke, 125 had a WHI reported hospitalization within +/− 30 days of the Medicare admission date, and 65 of these had hospital records adjudicated. The other cells in the concordance matrix remained unchanged. The sensitivity was unchanged between the 7-day match window and the 30-day match window at 87.4%. The PPV was 88.6% when using adjudicated WHI hospitalizations in a 30-day time window compared to a PPV of 88.9% when using the +/− 7-day time window.

Optimizing training algorithm performance

We examined the contribution of each stroke related ICD-9 CM code used in the preliminary CMS stroke definition to concordance and discordance between WHI and CMS stroke events. Specifically, we examined the distribution of the codes in the WHI Yes/CMS Yes cells and WHI No/CMS Yes cells. Of all the cerebrovascular disease codes we found that codes 437.1 and 437.9 (defined as generalized and unspecified cerebrovascular disease in the ICD-9 CM manuals) had disproportionately high number of events in the WHI No/CMS Yes cells (14% of events for the most comprehensive stroke definition) and low numbers in the WHI Yes/CMS Yes (<1% events for the most comprehensive stroke definition) cells. Therefore we eliminated these codes from the final stroke definitions and there was no change in sensitivity while the PPV improved slightly. For example, in the event based analysis of the most general CMS stroke definition, PPV improved from 56.4% (data not shown) to 60.1% when the 437.1x and 437.9x codes were eliminated.

Test/Validation Results (Tables II and III, Online)

Algorithm performance on the test data set was very similar to the performance on the training data set. In the event-based analyses (Table II, Online) using adjudicated WHI hospitalizations for the most comprehensive stroke definition, the sensitivity was 82.8% for the training data set vs. 82.0% for the test data set and the PPV was 85.8% vs. 84.6% for the test data set. The specificity and NPV were both 99.7% for the training and test data sets. Results for the stroke definition using only the principal diagnosis (definition 2) and the ischemic strokes (definition 3) showed that the algorithm performance on the test data set was similar (and frequently identical) to the performance on the training data set. For hemorrhagic strokes (definition 4), while sensitivity (75.9% training data set vs. 75.3% test data set), specificity (> 99.9 training and test), and NPV (99.9% training and test) were similar, the PPV diminished from 91.1% for the training data set to 84.3% for the test data set but was not statistically significant.

In the person-based analysis, a similar pattern was found (Table III, online). Also, the test set results were robust with expanding the match window to +/− 30 days (Table IV, available online). Specifically, among the 277 participants with a Medicare stroke diagnosis without a WHI stroke, 145 had a WHI reported hospitalization within +/− 30 days of the Medicare admission date and 77 of these had their hospital records adjudicated. The PPV was 86.9% when using adjudicated WHI hospitalizations compared to a PPV of 87.5% when using the +/− 7-day time window.

Comparative performance across all Medicare stroke definitions using the event-based and person-based frameworks in the Test data set for events where medical records were available is shown in Table 3.

DISCUSSION

These analyses of linked WHI-Medicare data showed that a claims-based algorithm for identifying acute strokes had very good sensitivity and excellent specificity and negative predictive values when compared with patient self-report followed by expert adjudication. Since strokes are infrequent overall, the PPV is easily influenced by small changes in specificity. Nevertheless, the PPV was fairly high (85% in the event based and 88% in the person-based analyses in the test set, definition 1) when only Medicare stroke events with WHI adjudicated medical records were included. Comparative performance across all Medicare stroke definitions in the more generalizable test data set comparing events where WHI medical records were available (Table 3) showed that, overall, the person-based approach had higher sensitivity and PPV compared to the event-based approach. The definitions using stroke codes in any position or ischemic stroke codes had the highest sensitivity ranging from 82% in the event-based to 88% in the person-based analyses for both definitions. A more stringent Medicare definition using only principal position stroke codes and the hemorrhagic stroke definition had lower sensitivities in the range of 75% for the event-based approach and 81% for the person-based approach (Table 3). The definition using the principal position stroke code had the best PPV at 88% in the event-based and 90% in the person-based approach. Based on our results, we conclude that Medicare data appear useful for population-based stroke research and that the performance characteristics of the Medicare data depend on the definition selected. The choice of the stroke definition may be guided by the sensitivity and PPV needed for the research purpose. One consideration is that Medicare claims data are typically available multiple months after the calendar year ends due to the time needed for claims accrual and processing, especially for hospitalizations occurring late in the calendar year. E.g., data on Medicare claims for the year 2011 may be available approximately mid-to late 2012. This has to be factored into research planning. There is also some lag in event ascertainment using the WHI approach due to the time needed for obtaining self-report data and requesting and adjudicating medical records.

Our results are consistent with other published reports. A recent systematic review of 26 articles comparing administrative codes vs. medical record abstraction found PPV in the ranges of 80% or higher when stroke specific ICD-9 codes such as 430, 431, and 434 were used. Although the sample sizes ranged from 50–4000, many of the studies had a small number of cases, were from limited geographic areas and did not examine training versus test data validation. [2] Our study is unique in the timespan covering 10–14 years, includes neurologist review, and the cohort of women is drawn from across the U.S. Medicare population, a high risk population for stroke. We recognize that our training and test data samples were both derived from the same larger data sample and were thus not completely independent data sets. Since our training and test samples were non-overlapping, we nevertheless believe that this study provides some validation of algorithm performance beyond that typically reported.

While the terms sensitivity and specificity are used in the traditional sense, we note that our application is not typical in that the reference-standard (WHI) was imperfect. There were some hospitalizations with stroke diagnosis codes recorded in Medicare data for which no hospitalization was reported by the WHI participants or their proxies. More complete ascertainment could potentially be achieved using a multi-pronged approach such as adding Medicare data or for studies that were geographically limited (unlike WHI) using discharge code lists from hospitals within a certain geographic area and surveillance of administrative claims.

This study validates the concordance between WHI and Medicare stroke events using two approaches. The first is the more stringent event-based analysis which required that the WHI and Medicare strokes had to be within +/− 7 days of each other to be declared a match. The person-based approach relaxed this requirement and this may have increased the matches slightly as well as decreasing discordant events for individuals with multiple stroke hospitalizations and hence, sensitivity and PPV were slightly higher in the person-based analyses. The vast majority of matches however occurred within a 7-day interval. The event-based analyses results presented here are relevant when any event (initial or recurrent) in Medicare hospital data is the focus, while the person-based results apply to incident events in trials and cohort studies.

When the Medicare events were limited to those where medical records were available for review by the stroke adjudication committee, we found generally high though imperfect concordance between Medicare data and medical records. We also note that there were WHI participants who had confirmed strokes based on hospital record adjudication, for whom the Medicare diagnosis codes were not stroke related. These Medicare hospitalizations which corresponded temporally with the WHI hospitalizations did not include stroke related codes but instead included codes for hypertension, atrial fibrillation, diseases of the urinary system, heart failure and transient ischemic attacks among others. Since the hospital record clearly supported a stroke diagnosis this under-coding could reflect a coding of stroke related complications or co-morbid conditions rather than stroke. However, we did not feel that we could add any of the ICD-9 CM codes to our Medicare stroke definition since while the sensitivity would increase, the specificity and PPV would be compromised. Furthermore, there were a variety of ICD-9 CM codes and not just a few predominant ones to account for this discordance. Conversely, there were many cases with a principal diagnosis of stroke in Medicare data which did not pass adjudication in the WHI. About a quarter of these were TIAs by neurologist adjudication in the WHI. Most however appeared to be other conditions. Thus, we conclude that while errors in the assignment of hospital codes result in both over and under-ascertainment of stroke in Medicare data, there was moderate agreement in the initial analysis with an overall kappa of 0.69 which increased to 0.84 when discordance was limited to adjudicated records.

The process of bringing the WHI cohort and the Medicare data in line for an event-based comparison revealed important insights about the strengths and weaknesses of each data source for the purpose of identifying incident strokes. As noted previously, using the Medicare data involves reliance on hospital coding. In some cases, strokes may be considered among the differential diagnoses and ruled out by the patient’s physicians. These cases should not get coded as strokes although they may get coded if the conclusion is not clearly documented by the treating clinicians in the medical record. In other cases, under-coding may occur, e.g., if an actual stroke is one among multiple medical problems and co-morbidities and not identified by the coder.

We note that Medicare identified more stroke events than WHI. Some of these cases were not identified or adjudicated by the WHI and these constituted the majority of the discordant cases –for example, in Table 1 the WHI No, CMS Yes cell declined 75% (from 318 to 79) when the discordance was limited to only those with adjudicated records. The main reason for this is the WHI case ascertainment process which relies on self-report.

Strengths of our work include the large nationwide cohort of WHI participants linked to Medicare, the availability of neurologist adjudicated stroke events in the WHI and the examination of both event-based and person-based frameworks for evaluating concordance. Limitations include the fact that we lacked Medicare data for participants enrolled in managed care plans. These participants were excluded from our analyses. In general, for studies such as the WHI which enroll persons from a broad US population, the extent to which this may cause a problem depends on the degree of managed care enrollment and the extent to which managed care enrollees are dissimilar to those with traditional fee-for-service Medicare. Another limitation is that our analysis includes predominantly persons 65 years or older since that is the age of Medicare eligibility, though younger persons with disability or end stage renal disease are also eligible. A different consideration is that using the adjudicated subset as the best estimate of the PPV (i.e. third row for each stroke definition in the Tables 1, 2, online Tables II and III) implicitly assumes that the non-reported and non-adjudicated events in the WHI No/CMS Yes cell would yield a similar PPV if all their data had been available for adjudication. However, if these events differ systematically from the adjudicated events, and then the true PPV may lie in between PPV values which includes all events (i.e. first row) and those which include only those events where records were available for adjudication (i.e. third row).

SUMMARY/CONCLUSIONS

This study compares two case ascertainment approaches for acute stroke hospitalization events: patient self-report of stroke followed by neurologist adjudication (WHI stroke) vs. ascertainment in Medicare data using four Medicare stroke definitions in an event-based or person-based framework. Both approaches were imperfect and while Medicare data claims helped capture some strokes not identified as such by the WHI, Medicare data also missed some WHI-adjudicated stroke events. We report the following insights into the performance characteristics of Medicare data: all Medicare stroke definitions data had very high specificity and negative predictive values when compared to the WHI approach. Medicare stroke definitions which used stroke codes in any diagnosis position or used ischemic stroke codes had high sensitivities in a person-level framework. The definition using stroke codes in any diagnosis position also had a high PPV. Medicare definitions which used hemorrhagic stroke codes or only principal position stroke codes to identify stroke were less sensitive especially in the event-based framework, though using a stroke code in the principal position had the highest PPV. The results of this study can assist researchers in selecting the best Medicare definition to suit their purpose. Our findings pave the way for an informed use of Medicare data to ascertain strokes, for example during long-term follow up or in large resource efficient clinical trials, with the caveat that outcome ascertainment in managed care enrollees may have to rely on other data sources.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: https://cleo.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf”

SOURCES OF FUNDING

The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

References

1. [Accessed April 10, 2013] The Women’s Health Initative. Available at: www.whi.org
2. Andrade SE, Harrold LR, Tjia J, Cutrona SL, Saczynski JS, Dodd KS, 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):100–128. [PubMed: 22262598]
3. Tirschwell DL, Longstreth WT Jr. Validating administrative data in stroke research. Stroke. 2002; 33:2465–2470. [PubMed: 12364739]

4. Tirschwell D, Kukull WA, Longstreth WT Jr. Inaccuracy of the ICD-9-CM in identifying the diagnosis of ischemic cerebrovascular disease. Neurology. 1998; 51:921. author reply 922. [PubMed: 9748071]
5. Wahl PM, Rodgers K, Schneeweiss S, Gage BF, Butler J, Wilmer C, 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. [PubMed: 20140892]
6. Goldstein LB. Accuracy of ICD-9-CM coding for the identification of patients with acute ischemic stroke: effect of modifier codes. Stroke. 1998; 29:1602–1604. [PubMed: 9707200]
7. Reker DM, Hamilton BB, Duncan PW, Yeh SC, Rosen A. Stroke: who's counting what? J Rehabil Res Dev. 2001; 38:281–289. [PubMed: 11392661]
8. Leibson CL, Naessens JM, Brown RD, Whisnant JP. Accuracy of hospital discharge abstracts for identifying stroke. Stroke. 1994; 25:2348–2355. [PubMed: 7974572]
9. Benesch C, Witter DM Jr, Wilder AL, Duncan PW, Samsa GP, Matchar DB. Inaccuracy of the International Classification of Diseases (ICD-9-CM) in identifying the diagnosis of ischemic cerebrovascular disease. Neurology. 1997; 49:660–664. [PubMed: 9305319]
10. Curb JD, McTiernan A, Heckbert SR, Kooperberg C, Stanford J, Nevitt M, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol. 2003; 13:S122–S128. [PubMed: 14575944]
11. Lakshminarayan K, Anderson DC, Jacobs DR Jr, Barber CA, Luepker RV. Stroke Rates: 1980–2000: The Minnesota Stroke Survey. Am J Epidemiol. 2009; 169:1070–1078. [PubMed: 19318614]
12. Lakshminarayan K, Schissel C, Anderson DC, Vazquez G, Jacobs DR, Ezzeddine M, et al. Five-year rehospitalization outcomes in a cohort of patients with acute ischemic stroke: Medicare linkage study. Stroke. 2011; 42:1556–1562. [PubMed: 21474798]

NIH-PA Author Manuscript
NIH-PA Author Manuscript

TABLE 1

Event Level Analysis of Training Data Set. Comparison of the Performance of Different Medicare Stroke Definition Algorithms vs. Physician Adjudicated WHI Stroke Events Using a 7-day Match Window.*

	Total	WHI Yes CMS Yes	WHI No CMS Yes	WHI Yes CMS No	WHI No CMS No	SN	SP	PPV	NPV	Kappa
Stroke (any diagnosis code; definition 1)										
All	31399	478	318	104	30499	82.1	99.0	60.1	99.7	0.69
WHI Hospitalization reported [1]	31212	478	136	99	30499	82.8	99.6	77.9	99.7	0.80
WHI Hospitalization adjudicated [2]	31115	478	79	99	30499	82.8	99.7	85.8	99.7	0.84
Stroke (principal diagnosis; definition 2)										
All	31448	429	237	153	30629	73.7	99.2	64.4	99.5	0.68
WHI Hospitalization reported [1]	31311	429	105	148	30629	74.4	99.7	80.3	99.5	0.77
WHI Hospitalization adjudicated [2]	31277	429	71	148	30629	74.4	99.8	85.8	99.5	0.79
Hemorrhagic Stroke (definition 3)										
All	31323	82	49	27	31165	75.2	99.8	62.6	99.9	0.68
WHI Hospitalization reported [1]	31296	82	23	26	31165	75.9	99.9	78.1	99.9	0.77
WHI Hospitalization adjudicated [2]	31281	82	8	26	31165	75.9	>99.9	91.1	99.9	0.83
Ischemic Stroke (definition 4)										
All	31379	369	296	84	30630	81.5	99.0	55.5	99.7	0.65
WHI Hospitalization reported [1]	31218	369	139	80	30630	82.2	99.5	72.6	99.7	0.77
WHI Hospitalization adjudicated [2]	31175	369	96	80	30630	82.2	99.7	79.4	99.7	0.80

Abbreviations: CMS, Centers for Medicare & Medicaid Services data, i.e., Medicare data; NPV, negative predictive value; PPV, positive predictive value; SN, Sensitivity; SP, Specificity; WHI, Women's Health Initiative data.

[1] Excludes strokes in the outpatient setting. Also, excludes events for which no WHI hospitalization was reported by the participant or a proxy with admission date within +/− 7 days of the Medicare hospitalization.

[2] Excludes events for which no medical records were received due to administrative reasons (e.g., no record release or no documents) or when reasons reported for hospitalization did not trigger adjudication (e.g. bunionectomy).

[*] Analysis based on a nationwide dataset.

TABLE 2

Person Level Analysis of Training Data Set. Comparison of the Performance of Different Medicare Stroke Definition Algorithms vs. Physician Adjudicated WHI Stroke Events Using a 7-day Match Window. *

	Total	WHI Yes CMS Yes	WHI No CMS Yes	WHI Yes CMS No	WHI No CMS No	SN	SP	PPV	NPV	Kappa
Stroke (any diagnosis code; definition 1)										
All	24428	505	240	77	23606	86.8	99.0	67.8	99.7	0.75
WHI Hospitalization reported [1]	24294	505	110	73	23606	87.4	99.5	82.1	99.7	0.84
WHI Hospitalization adjudicated [2]	24247	505	63	73	23606	87.4	99.7	88.9	99.7	0.88
Stroke (principal diagnosis; definition 2)										
All	24428	458	182	124	23664	78.7	99.2	71.6	99.5	0.74
WHI Hospitalization reported [1]	24325	458	83	120	23664	79.2	99.7	84.7	99.5	0.81
WHI Hospitalization adjudicated [2]	24298	458	56	120	23664	79.2	99.8	89.1	99.5	0.84
Hemorrhagic Stroke (definition 3)										
All	24428	85	42	24	24277	78.0	99.8	66.9	99.9	0.72
WHI Hospitalization reported [1]	24406	85	20	24	24277	78.0	99.9	81.0	99.9	0.79
WHI Hospitalization adjudicated [2]	24393	85	7	24	24277	78.0	>99.9	92.4	99.9	0.85
Ischemic Stroke (definition 4)										
All	24428	388	233	65	23742	85.7	99.0	62.5	99.7	0.72
WHI Hospitalization reported [1]	24309	388	118	61	23742	86.4	99.5	76.7	99.7	0.81
WHI Hospitalization adjudicated [2]	24273	388	82	61	23742	86.4	99.7	82.6	99.7	0.84

Abbreviations: CMS, Centers for Medicare & Medicaid Services data, i.e., Medicare data; NPV, negative predictive value; PPV, positive predictive value; SN, Sensitivity; SP, Specificity; WHI, Women's Health Initiative data.

[1] Excludes strokes in the outpatient setting. Also, excludes events for which no WHI hospitalization was reported by the participant or a proxy with admission date within +/− 7 days of the Medicare hospitalization.

[2] Excludes events for which no medical records were received due to administrative reasons (e.g., no record release or no documents) or when reasons reported for hospitalization did not trigger adjudication (e.g. bunionectomy).

[*] Analysis based on a nationwide dataset.

TABLE 3

Summary of Medicare Algorithm Performance on Test Data across Various Stroke Definitions. Results are for Events Where WHI Medical Records Were Available for Adjudication. Analyses Use a 7-day Match Window.

	SN	SP	PPV	NPV	Kappa
Event-Based Analysis					
Stroke (any diagnosis code; definition 1)	82.0	99.7	84.6	99.7	0.83
Stroke (principal diagnosis; definition 2)	75.3	99.8	87.5	99.5	0.81
Hemorrhagic Stroke (definition 3)	75.3	>99.9	84.3	99.9	0.79
Ischemic Stroke (definition 4)	82.2	99.7	79.5	99.7	0.81
Person-Level Analysis					
Stroke (any diagnosis code; definition 1)	88.0	99.7	87.5	99.7	0.87
Stroke (principal diagnosis; definition 2)	81.0	99.8	89.6	99.5	0.85
Hemorrhagic Stroke (definition 3)	80.6	>99.9	87.2	99.9	0.84
Ischemic Stroke (definition 4)	88.4	99.6	82.7	99.8	0.85

Abbreviations: CMS, Centers for Medicare & Medicaid Services data, i.e., Medicare data; NPV, negative predictive value; PPV, positive predictive value; SN, Sensitivity; SP, Specificity; WHI, Women's Health Initiative data.