HHS Public Access
Author manuscript
Epidemiology. Author manuscript; available in PMC 2015 December 02.

Published in final edited form as:
Epidemiology. 2013 March ; 24(2): 334–335. doi:10.1097/EDE.0b013e3182821e75.

Validation of Veterans Health Administration administrative data algorithms for the identification of cardiovascular hospitalization and covariates

Kurt Niesner[1], **Harvey J. Murff**[1,2], **Marie R. Griffin**[1,2,4], **Brian Wasserman**[2], **Robert Greevy**[1,3], **Carlos G. Grijalva**[4], and **Christianne L. Roumie**[1,2]

[1]VA Tennessee Valley Geriatric Research Education Clinical Center (GRECC), HSR&D Center, Nashville, TN

[2]Department of Medicine, Vanderbilt University, Nashville, TN

[3]Department of Biostatistics, Vanderbilt University, Nashville TN

[4]Department of Preventive Medicine, Vanderbilt University, Nashville, TN

Keywords

Data sources; Confounding factors (Epidemiology); Bias (Epidemiology); Methods; Stroke; Acute myocardial infarction

To the editor

We identified a cohort of veterans prescribed first treatment for type 2 diabetes in order to evaluate comparative effectiveness of diabetes treatments on cardiovascular outcomes. We thus aimed to determine the accuracy of algorithms to identify stroke or acute myocardial infarction (AMI) hospitalization among patients with diabetes within Veterans Health Administration (VHA) and to determine validity of pharmacy claims for antidiabetic drugs and covariates. The institutional review board approved the study.

The population included veterans ≥18 years old receiving VHA healthcare between January 1, 2000 and December 31, 2007. Diagnoses were coded according to the International Classification of Diseases, Ninth Revision (ICD9-CM) system.(1) The cohort was limited to patients with gender, birth date and a prescription fill for an antidiabetic drug. We randomly selected a sample of patients with a hospitalization indicating stroke or AMI for medical record review. Definitions of stroke, AMI, exposures and covariates are available as online supplemental table 1. (2-4)

Medical records were abstracted using structured forms, specific for stroke and AMI, and outcomes were adjudicated using the following definitions. A stroke was defined by the

Corresponding Author: Christianne L. Roumie, MD MPH, Nashville VA Medical Center, 1310 24th Ave South GRECC, Nashville TN 37212 phone (615) 873-8013 Fax: (615)873-7981, christianne.roumie@vanderbilt.edu.

Disclosures
None

rapid onset of a persistent neurologic deficit attributed to an obstruction or rupture of the arterial system lasting >24 hours unless death supervened or by a lesion compatible with an acute stroke on CT or MRI scan.(5) An AMI was classified as "definite or probable" using an adaption of the Women's Health Initiative definition based on chest pain, electrocardiogram pattern and abnormal cardiac enzymes(3). We considered the medical record as the referent and calculated the positive and negative predictive values (PPV/ NPV) with 95% confidence intervals (CI) for binomial proportions using Wilson's formula(6).

Table 1 demonstrates the PPVs and 95% CIs for each ICD9-CM code used to identify stroke or AMI. Of the 226 stroke hospitalizations in 194 patients (20 patients with 2 to 5 hospitalizations), 183 (PPV=80.9%, CI 75.3, 85.6) met criteria for an acute stroke. Of the 174 AMI hospitalizations among 158 patients (12 patients with 2 to 4 hospitalizations) there were 148 definite and 8 probable AMIs (PPV=89.7%, CI 84.2, 93.4).

Filled prescriptions were compared to medications recorded in patient charts among the 400 hospitalizations. Of the 94 patients with a metformin prescription fill within 90 days prior to hospitalization, 83 had metformin listed as a medication at hospitalization (PPV=88.3%, CI 80.2, 93.3 Sensitivity=79.0%, CI 70.3, 85.7). Results were similar for sulfonylurea and insulin (Supplemental Table 2). The PPVs for aspirin prescriptions (80.2%, CI 72.0, 86.7) and for indicators of smoking (72.2%, CI 63.4, 79.5) were also high; however, it was clear that administrative data did not capture all such exposures and sensitivity for aspirin and smoking were 79% and 56%, respectively (Supplemental Table 3) .

Supplemental table 4 demonstrates that the age-adjusted probability of current smoking was lower among insulin users compared with those with no fill of an antidiabetic drug (22.3 vs. 44.9%, p=0.007). The age adjusted probability of aspirin use by the administrative algorithm was slightly higher among insulin users than those who did not fill an antidiabetic drug (68.3% vs 45.2%, p=0.002).

In summary, using a random sample of 400 VHA cardiovascular hospitalizations, our algorithms identified acute stroke with a PPV of 80.9% and AMI with a PPV of 89.7%. Thus, our algorithms have reasonable accuracy in identifying stroke and AMI hospitalizations in the VHA. Compared to chart abstraction, use of ICD-9 codes for identifying stroke and AMI saves time and resources. Measurements of exposures and important covariates through administrative data were also similar to those determined by chart review. Our study suggests that the use of such algorithms is a reasonable strategy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was funded in part by VHA-Clinical Science Research and Development investigator initiated grant I01CX000570-01 and in part by Contract No. 290-2010-00016I from the Agency for Healthcare Research and Quality, US Department of Health and Human Services as part of the Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) program.

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality, the US Department of Health and Human Services or the Department of Veterans Affairs.

References

1. International Classification of Diseases, Ninth Revision, Clinical Modification. Washington, DC: Public Health Service, US Dept of Health and Human Services; 1988.
2. 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. 2007
3. Curb JD, McTiernan A, Heckbert SR, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. Ann Epidemiol. 2003; 13:S122–128. [PubMed: 14575944]
4. Choma NN, Griffin MR, Huang RL, et al. An algorithm to identify incident myocardial infarction using Medicaid data. Pharmacoepidemiol Drug Saf. 2009
5. Howard VJ, Cushman M, Pulley L, et al. The reasons for geographic and racial differences in stroke study: objectives and design. Neuroepidemiology. 2005; 25:135–143. [PubMed: 15990444]
6. Brown L, Cai TT, Dasgupta A. Interval Estimation for a Binomial Proportion. Statistical Science. 2001; 16:101–133.

Table 1

Positive Predictive Value (PPV) and Confidence Intervals for stroke and acute myocardial infarction ICD9-CM codes

ICD9-CM code and definition	Reviewed N (%)	Confirmed as New Case	PPV (95% CIs)
Ischemic/Thrombotic Strokes Total	**226 (100)**	**183**	**80.9 (75.3, 85.6)**
Cerebral artery occlusion, unspecified with infarction (434.91)	118 (52.2)	107	90.7 (84.1, 94.7)
Acute, but ill-defined, cerebrovascular disease (436)	30 (13.3)	20	66.7 (48.8, 80.8)
Occlusion of carotid artery with infarction (433.11)	22 (9.7)	11	50.0 (30.7, 69.3)
Occlusion of multiple arteries with infarction (433.31)	7 (3.1)	6	85.7 (48.7, 97.4)
Occlusion of cerebral arteries with infarction (434.01)	6 (2.7)	5	83.3 (43.6, 97.0)
Occlusion of pre cerebral arteries with infarction (433.01)	2 (0.9)	2	100 (34.2, 100)
Occlusion of vertebral artery with infarction (433.21)	1 (0.4)	1	100 (20.7, 100)
Occlusion/stenosis of other specified pre cerebral artery with infarction (433.81)	1 (0.4)	1	100 (20.7, 100)
Hemorrhagic Strokes			
Intracerebral Hemorrhage (431)	18 (8.0)	14	77.8 (54.8, 91.0)
Subarachnoid Hemorrhage (430)	6 (2.7)	3	50.0 (18.8, 81.2)
Embolic Strokes			
Cerebral Embolism with infarction (434.11)	14 (6.2)	13	92.9 (68.5, 98.7)
Cerebral Embolism (Embolic Stroke) (434.1)	1 (0.4)	0	0 (0, 79.3)
Acute Myocardial Infarction Total	**174 (100)**	**156**	**89.7 (84.2, 93.4)**
Subendocardial infarction (NSTEMI) (410.7)	124 (71.2)	112	90.3 (84.2, 93.4)
ST elevation MI of other inferior wall (410.4)	13 (7.5)	12	92.3 (66.7, 98.6)
Acute Myocardial infarction unspecified site (410.9)	11 (6.3)	7	63.6 (35.4, 84.8)
ST elevation MI of other anterior wall (410.1)	10 (5.7)	9	90.0 (59.6, 98.2)
ST elevation MI of other specified sites (410.8)	9 (5.2)	9	100 (70.1, 100)
ST elevation MI of anterolateral wall (410.0)	4 (2.3)	4	100 (51.0, 100)
ST elevation MI of inferolateral wall (410.2)	1 (0.6)	1	100 (20.7, 100)
ST elevation MI of inferoposterior wall (410.3)	1 (0.6)	1	100 (20.7, 100)
ST elevation MI of other lateral wall (410.5)	1 (0.6)	1	100 (20.7, 100)