12-2013

Validation of Diagnostic and Procedural Codes for Identification of Acute Cardiovascular Events in US Veterans with Rheumatoid Arthritis

Lisa A. Davis
Denver Health and Hospital Authority, lisa.davis@ucdenver.edu

Alyse Mann
Denver Veterans Affairs Medical Center, alyse.mann@va.gov

Grant W. Cannon
George E. Wahlen Veterans Affairs Medical Center, grant.cannon@va.gov

Ted R. Mikuls
University of Nebraska Medical Center, tmikuls@unmc.edu

See next pages for additional authors

The Electronic Data Methods (EDM) Forum is supported by the Agency for Healthcare Research and Quality (AHRQ), Grant 1U18HS022789-01. eGEMs publications do not reflect the official views of AHRQ or the United States Department of Health and Human Services.

Validation of Diagnostic and Procedural Codes for Identification of Acute Cardiovascular Events in US Veterans with Rheumatoid Arthritis

Abstract

Objective: To assess the accuracy of International Classification of Diseases, Ninth Revision, and Current Procedural Terminology codes for identifying cardiovascular (CV) events (myocardial infarction [MI], stroke, coronary artery bypass graft [CABG], and percutaneous coronary intervention [PCI]) in enrollees of the Veterans Affairs Rheumatoid Arthritis (VARA) registry.

Design: We performed a validation study from VARA enrollment until 6/1/2010 to compare the accuracy of CV events in those with and without CV-event coding in inpatient and outpatient records to evaluate for CV events +/- 3 months of the coding. The positive predictive value (PPV) was calculated, and codes with a PPV ≥50% were included in a composite coding algorithm.

Results: We evaluated 107 individuals for 21 CV-event codes and 60 individuals without CV-event coding. The PPV varied between 0-100%. Composite coding algorithms' PPV ranged from 70-100%.

Conclusions: Validation of these algorithms allows for identification of acute CV events with known accuracy. The sensitivity and PPV of coding algorithms for CABG and PCI exceed that of stroke and MI.

Key words: Methods, Cohort Identification, Data Reuse

Acknowledgements

This work was supported in part by the Rheumatology Research Foundation's Scientist Development Award, the National Institutes of Health (NIH) Loan Repayment Program, and the NIH's T32 AR007534-24.This material is the result of work supported with resources and the use of facilities at the Denver Veterans Affairs Medical Center.

Keywords

Methods, Cohort Identification, Data Reuse

Disciplines

Health Services Research

Authors

Lisa A Davis, *Denver Health and Hospital Authority*; Alyse Mann, *Denver Veterans Affairs Medical Center*; Grant W Cannon, *George E. Wahlen Veterans Affairs Medical Center*; Ted R Mikuls, *University of Nebraska Medical Center*; Andreas M Reimold, *Dallas Veterans Affairs Medical Center*; Liron Caplan, *Denver Veterans Affairs Medical Center*.

Validation of Diagnostic and Procedural Codes for Identification of Acute Cardiovascular Events in US Veterans with Rheumatoid Arthritis

Lisa A. Davis, MD, MSCS;[i,ii,iii] Alyse Mann, BS;[ii] Grant W. Cannon, MD;[iv] Ted R. Mikuls MD, MPH;[v] Andreas M. Reimold, MD;[vi] Liron Caplan, MD, PhD [ii,iii]

Abstract

Objective: To assess the accuracy of International Classification of Diseases, Ninth Revision, and Current Procedural Terminology codes for identifying cardiovascular (CV) events (myocardial infarction [MI], stroke, coronary artery bypass graft [CABG], and percutaneous coronary intervention [PCI]) in enrollees of the Veterans Affairs Rheumatoid Arthritis (VARA) registry. **Design:** We performed a validation study from VARA enrollment until 6/1/2010 to compare the accuracy of CV events in those with and without CV-event coding in inpatient and outpatient records to evaluate for CV events +/- 3 months of the coding. The positive predictive value (PPV) was calculated, and codes with a PPV ≥50% were included in a composite coding algorithm. Results: We evaluated 107 individuals for 21 CV-event codes and 60 individuals without CV-event coding. The PPV varied between 0-100%. Composite coding algorithms' PPV ranged from 70-100%. **Conclusions:** Validation of these algorithms allows for identification of acute CV events with known accuracy. The sensitivity and PPV of coding algorithms for CABG and PCI exceed that of stroke and MI.

Background

Cardiovascular (CV) disease is a major comorbidity in patients with rheumatoid arthritis (RA). RA patients experience two to three times the risk of CV disease compared with the non-RA population.[1,2]To accurately examine the important association of RA and CV disease, however, validated CV outcomes—such as myocardial infarction (MI), stroke (ischemic or hemorrhagic), coronary artery bypass graft (CABG) surgery, and percutaneous coronary intervention (PCI)—must be established.

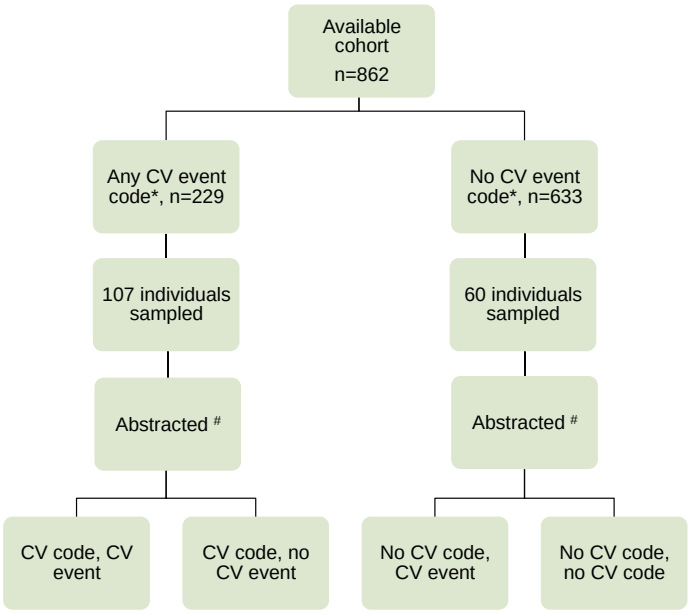
International Classification of Diseases, Clinical Modification, Ninth Revision (ICD-9-CM, hereafter referred to as ICD-9); ICD-9-Procedure, and Current Procedural Terminology (CPT) codes represent a convenient method of developing such CV outcomes. Use of validated ICD-9, ICD-9-Procedure and CPT codes in electronic medical records (EMR) can be convenient and requires fewer hours when compared to manual chart abstraction. However, there are some important gaps in the literature regarding validation of CV coding. First, previous studies report widely divergent Positive Predictive Values (PPVs) for each individual ICD-9 or CPT code. The PPVs for myocardial infarction ICD-9 codes, for example, vary from 2.5 percent to 100 percent.[3] Second, prior literature tends to only offer estimates for the accuracy of administrative data for *prevalent* CV disease.[4,5] When CV events are defined without explicitly defining a window of allowable time for the event around the date the code was assigned, it limits the validity of the administrative data as an outcome measure.[6-8]Without an associated date, such coding algorithms are also unsuitable for time-to-event analysis.

Third, prior literature frequently does not give all of the details that might be desirable for different types of analysis. For example, many of the previously published manuscripts do not report the sensitivity or specificity of the codes used to identify incident disease [6,9-12], and PPV has not been reported in the few studies regarding PCI and CABG.[10,12-14]Fourth, all previously published works in this field [4,13,15] rely exclusively upon data from hospital discharges, which may introduce a significant bias by neglecting to account for CV events documented in outpatient records.

To address these deficiencies, we evaluated the validity of ICD-9, ICD-9 Procedure, and CPT codes for CV events by examining a cohort of patients with well-established RA who were enrolled in the Veterans Affairs Rheumatoid Arthritis (VARA) registry. Using electronic medical record (EMRs) of the VARA patients, we determined the validity of CV-event codes within a six-month window using inpatient, outpatient, and inpatient + outpatient records. In contrast to prior investigation, our study focused on acute events (those occurring within months of the billing code, as opposed to CV events occurring at any time), estimated sensitivity and specificity for the coding algorithms, evaluation of the value in using outpatient in addition to inpatient records, and clarification of the acceptable time window necessary to meet our case definition.

[i]Denver Health and Hospital Authority, [ii]Denver Veterans Affairs Medical Center (VAMC), [iii]University of Colorado School of Medicine, [iv]George E Wahlen VAMC, [v]Omaha VAMC and University of Nebraska Medical Center, [vi]Dallas VAMC

Figure 1. Study subject sampling methodology



Notes: *International Classification of Diseases, Clinical Modification, 9th Revision (ICD-9), diagnostic codes limited to the first position for myocardial infarction (MI) and stroke. ICD-9-Procedure and Current Procedural Terminology (CPT) codes were utilized in all positions for coronary artery bypass graft (CABG) and percutaneous coronary intervention (PCI). Codes evaluated for MI: 410.x, 411.x, 412.x, 413.x, 414.x, 429.2, and v45.81; Stroke: 433.11, 433.91, 434.91, 435.x, 436.x, 437.9, 438.x; CABG: 361.x; PCI: 36.06, 36.07, 0.66 (ICD-9-Procedure) and 92973, 92980, 92995 (CPT).

Sampled individuals' entire medical record was abstracted using a structured chart abstraction instrument for MI, stroke, PCI, and CABG. Those with cardiovascular- (CV) event coding were specifically evaluated for CV events within a six-month window of the assignment of the code (three months on either side of the date on which the code was entered into the EMR).

Methods

Patients and Setting: Our study population consisted of subjects enrolled in the VARA prospective registry. The VARA registry is a multicentered observational cohort at 11 VA Medical centers that has been fully described elsewhere.[16,17] In brief, all patients with RA at participating sites are invited to enroll and provide informed consent for the collection of demographic and longitudinal clinical information as well as biologic samples (sera, plasma, DNA). For this study, we relied on a convenience sample of participants from enrollment sites for which we had direct access to medical records: Dallas, Texas; Denver, Colo.; and Omaha, Neb.

Study Design: We performed a validation study and utilized administrative data from inpatient and outpatient national VA databases, based on clinical encounters from time of enrollment into VARA (initiated in 2003) until June 1, 2010 or until the patient's most recent clinical encounter, whichever came first. We examined CV events subsequent to VARA enrollment, as the additional attention of study coordinators, coinvestigators, and site principle investigators is likely to increase the odds of detecting CV events above the surveillance performed by primary care providers during usual care.

Baseline Characteristics: Participants in VARA have baseline demographic information collected upon enrollment into the registry. These data include the following: age at time of registry enrollment; gender; age at time of death; education (years); race/ethnicity; tobacco use at time of enrollment; rheumatoid factor (RF) antibody positivity; anticyclic citrullinated protein (anti-CCP);presence of rheumatoid nodules; and RA disease duration from the time of diagnosis until enrollment in VARA. In addition, 28-joint disease activity score (DAS28) and health assessment questionnaire (HAQ) values were collected over the duration of enrollment in VARA. DAS28 is a clinical measure of active disease and HAQ is a measure of disability used in clinical practice and for RA research.

Procedures: Our study cohort included two populations: those with CV-event coding, which included ICD-9, ICD-9-Procedure, and CPT coding for MI (410.x, 411.x, 412.x, 413.x, 414.x, 429.2, and v45.81), stroke (433.11, 433.91, 434.91, 435.x, 436.x, 437.9x, 438.x), CABG (ICD-9-Procedure 36.1x), or PCI (ICD-9-Procedure 36.06, 36.07, 0.66 or CPT 92973, 92980, 92995); and those without CV-event coding (see Figure 1). These codes were selected based on prior work for MI,[4] stroke,[4] CABG,[13] and PCI.[13] Individuals with and without coding for MI, stroke, CABG, and PCI were randomly chosen based on a randomly assigned numeric identification code from the list of VARA participants for medical chart review. CV-related codes were drawn from the inpatient and outpatient files of the VA's Corporate Franchise Data Center, a national centralized computer-processing center. Samples of 110 and 60 individuals for CV-related and no-CV-related codes were chosen a priori.

Prior work suggests that codes appearing in nonprimary positions (i.e., codes that were not the primary reason for hospitalization or outpatient visits) often represent false positives.[18] For this reason, we limited the coding position to the first position for in-hospital records, hospital discharge records, and outpatient records for ICD-9 diagnostic codes (MI, stroke). For ICD-9-Procedure and CPT codes (CABG, PCI), which were not anticipated to demonstrate high false positive rates, we accepted procedural codes in any position.

The entire medical records of those individuals randomly chosen for medical chart review (both in the CV-event coding and in the no-CV-event coding populations) were abstracted using a structured chart abstraction instrument for MI, stroke, PCI, and CABG. Those with CV-event coding were specifically evaluated for CV events within a six-month window of the assignment of the code (three months on either side of the date on which the code was entered into the EMR). Since many analyses (such as time-to-event) require that the timing of events be known, our case definition included this specified allowable time frame to characterize valid events. Because these events may not be the patient's initial CV event (i.e., events occurring prior to introduction of the EMR may not be captured), it is not entirely accurate to describe them as incident events.[19] Thus, we defined the first event since the introduction of the EMR data as "acute" events—in contrast to prevalent events.

http://repository.academyhealth.org/egems/vol1/iss3/8
DOI: 10.13063/2327-9214.1023

The case definition of a CV event required documentation by a clinician of a MI, stroke, PCI, or CABG in progress notes, discharge summaries, or procedure notes. For MI and strokes, we accepted events described as "likely," but did not accept those characterized as "possible." For those in the no-CV-event population, the patient's entire medical record was reviewed for the presence of a CV event, from the time of enrollment into VARA until June 1, 2010.

The PPV was then determined for each individual diagnostic or procedural code. To assess a more sensitive approach, a composite coding algorithm was then created consisting of the sum of the individual codes with a PPV ≥50 percent for each clinical condition. The negative predictive value (NPV), sensitivity, and specificity of each of the individual codes and composite codes were also calculated. In order to account for the bias that may have resulted from our sampling method (choosing from those who had CV coding and those who did not), we utilized the method described by Weiner et al.[20] to calculate sensitivity and specificity.

Statistical Analysis: Students' t-test and chi-square were used to determine differences in baseline variables, as appropriate. A p-value of ≤ 0.05 was considered significant. All analyses were performed using STATA 11.2 (College Station, Texas).

Human Subjects Review: Data were obtained from the U.S. Department of Veterans Affairs. However the funding sources had no role in the study design; in the collection, analysis and interpretation of data; or in the writing of the report. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the U.S. Department of Veterans Affairs. This work was approved by the Internal Review Board of the University of Colorado and the Veterans Affairs Medical Center in Denver, Colorado. Additionally, this project was approved by the Scientific and Ethics Approval Committee for the VARA registry.

Results

Of 862 individuals available for analysis, 229 (27 percent) were coded for a CV-related event in the first position for ICD-9 or in any position for ICD-9-procedure or CPT; 633 (73 percent) had no CV-related event coding (see Figure 1). Of the CV-event coded cohort, the medical records of 107 of the intended 110 individuals were abstracted for acute CV-related coding; of those without CV-event coding, the medical records of 60 individuals were abstracted for CV-related events (for further details of those without CV-event codes, see Supplemental Table 1).

There were significant differences in age, gender, reported race/ethnicity, and RA disease duration between those coded for a CV-related event and those not coded for a CV-related event (see Table 1). RA patients with a coded CV event were older, more often male, more likely to report Caucasian race, and had a longer RA disease duration then those without a coded CV event. Additionally, we examined the baseline differences between individuals at the three sites (Dallas, Texas; Denver, Colo.; and Omaha, Neb.). Significant differences were found in mean education, race/ethnicity, mean disease duration, and RF positivity (see Supplemental Table 2). We also examined how representative our sample was of the entire cohort. There were significant differences in age at enrollment (abstracted individuals 64.76 years versus 62.78 years in those not abstracted, p-value 0.041) and in RF positivity (92 percent versus 86 percent positive, p-value 0.041). See Supplemental Table 3 for full details.

The mean difference between an actual event and the CV-related code being assessed was 9.5 days (SD 20.2 days).

The PPV and its 95 percent confidence interval, NPV, sensitivity and specificity of each individual code for both inpatient and outpatient records are presented in Table 2, while those of the inpatient records only are in Table 3 and those of the outpatient records only are in Table 4. In general, there was a wide range of the PPV of the individual codes ranged (from 0–100%). After this initial PPV of the individual codes were determined, those with a PPV ≥50 percent were included in the composite coding algorithms. A summary of the composite codes for the total cohort, inpatient records only, and outpatient records only, may be found in Table 5.

Supplemental Table 1. Results of individuals with no cardiovascular related coding

Condition	TN	n=60	%
Acute MI	59	60	98.33
Stroke	60	60	100
CABG	60	60	100
PCI	59	60	98.33
Total	58	60	96.67

Notes: TN= True Negative; MI= myocardial infarction; CABG = coronary artery bypass graft; PCI = percutaneous coronary intervention.

Table 1. Characteristics of the cohort

		Entire cohort (n=862)			Event (n=229)			No Event (n=633)			
Variable		**n=**	**%, mean**	**SD**	**n=**	**%, mean**	**SD**	**n=**	**%, mean**	**SD**	**p-value**
Age, years		**862**	**63.17**	**11.22**	**229**	**66.10**	**9.65**	**633**	**62.11**	**11.56**	**<0.0001**
Age of death, years		127	73.29	9.20	55	73.99	8.99	72	72.75	9.37	0.454
Gender, male		**829**	**89.75**		**210**	**95.89**		**534**	**87.54**		**0.001**
Education, years		765	12.69	2.53	203	12.46	2.96	562	12.77	2.36	0.134
Race/Ethnicity	**Caucasian**	**656**	**79.04**		**190**	**86.76**		**466**	**76.27**		**0.005**
	African American	**112**	**13.49**		**23**	**10.50**		**89**	**14.57**		
	Hispanic	**42**	**5.06**		**2**	**0.91**		**40**	**6.55**		
	Native American	**13**	**1.57**		**4**	**1.83**		**9**	**1.47**		
	Asian American	**2**	**0.24**					**2**	**0.33**		
	Other	**5**	**0.60**					**5**	**0.82**		
Tobacco Use	Never	154	18.55		37	16.89		117	19.15		0.227
	Former	410	99.64		120	54.79		290	47.46		
	Current	263	50.24		62	28.31		201	32.90		
Average DAS		807	3.74	1.15	215	3.80		592	3.72	1.15	0.376
Average HAQ		815	1.00	0.52	216	1.06		599	0.99	0.52	0.079
Average disease duration, years		**729**	**17.43**	**12.48**	**184**	**19.67**		**545**	**16.68**	**12.29**	**0.005**
RF positive		827	87.42		191	87.21		532	87.50		0.913
Anti-CCP positive		606	80.59		159	81.96		447	80.11		0.575
Rheumatoid Nodules		463	58.91		127	61.65		336	57.93		0.352

Notes: SD=Standard deviation; Age=age at the time of registry enrollment; Avg. DAS28= average disease activity score, utilizing 28 joint count; Avg. HAQ= average health assessment questionnaire score; Avg. disease duration= average rheumatoid arthritis disease duration at time of registry enrollment; RF= rheumatoid factor; Anti-CCP= anti-citrullinated protein antibody; Rheum. Nodules= rheumatoid nodules. P-value represents test of differences between event and no event populations. Student's t-test was used to evaluate age, age of death, education, average DAS, average HAQ, and average disease duration. Chi-square was used to evaluate race/ethnicity, tobacco use, RF positivity, anti-CCP positivity, and presence of rheumatoid nodules.

http://repository.academyhealth.org/egems/vol1/iss3/8
DOI: 10.13063/2327-9214.1023

Supplemental Table 2. Comparison of characteristics at the three sites contained in the cohort

		Entire cohort (n=862)			Dallas			Denver			Omaha			
Variable		**n=**	**%, mean**	**SD**	**n=562**	**%, mean**	**SD**	**n=115**	**%, mean**	**SD**	**n=185**	**%, mean**	**SD**	**p-value**
Age, years		862	63.17	11.22	562	63.00	11.59	115	62.63	9.83	185	64.00	10.87	0.365
Age of death, years		127	73.29	9.20	99	73.00	8.88	2	75.34	2.89	26	74.20	10.74	0.453
Gender, male		744	89.75		480	89.05		104	90.43		160	91.43		0.645
Education, years		**765**	**12.69**	**2.53**	**507**	**12.46**	**2.52**	**115**	**13.33**	**2.64**	**143**	**12.98**	**2.37**	**<0.0001**
Race/ Ethnicity	**Caucasian**	**656**	**79.04**		**418**	**77.55**		**84**	**73.04**		**154**	**87.50**		**<0.0001**
	African American	**112**	**13.49**		**93**	**17.25**		**7**	**6.09**		**12**	**6.82**		
	Hispanic	**42**	**5.06**		**20**	**3.71**		**17**	**14.78**		**5**	**2.84**		
	Native American	**13**	**1.57**		**4**	**0.74**		**5**	**4.35**		**4**	**2.27**		
	Asian American	**2**	**0.24**		**2**	**0.37**		**0**	**0.00**		**0**	**0.00**		
	Other	**5**	**0.60**		**2**	**0.37**		**2**	**1.74**		**1**	**0.57**		
Tobacco Use	Never	154	18.55		96	17.81		21	18.26		37	21.02		0.674
	Former	410	49.40		262	48.61		57	49.57		91	51.70		
	Current	263	31.69		178	33.02		37	32.17		48	27.27		
Average DAS		807	3.74	1.15	524	3.68	1.13	110	3.74	1.40	173	3.94	0.98	0.481
Average HAQ		815	1.00	0.52	527	1.02	0.53	113	0.95	0.50	175	1.01	0.52	0.995
Average disease duration, years		**729**	**17.43**	**12.48**	**523**	**17.41**	**12.43**	**114**	**20.66**	**12.87**	**92**	**13.57**	**11.20**	**<0.0001**
RF positive		**723**	**87.42**		**482**	**89.59**		**95**	**83.33**		**146**	**83.43**		**0.037**
Anti-CCP positive		606	80.59		420	80.46		90	81.08		96	80.67		0.988
Rheumatoid Nodules		463	58.91		316	58.74		67	58.77		80	59.70		0.979

Notes: SD=Standard deviation; Age=age at the time of registry enrollment; Avg. DAS28= average disease activity score, utilizing 28 joint count; Avg. HAQ= average health assessment questionnaire score; Avg. disease duration= average rheumatoid arthritis disease duration at time of registry enrollment; RF= rheumatoid factor; Anti-CCP= anti-citrullinated protein antibody; Rheum. Nodules= rheumatoid nodules; Note: p-value represents test of differences between event and no event populations.

Supplemental Table 3. Characteristics abstracted individuals versus those not abstracted

		Entire cohort (n=862)			Abstracted (n=167)			Not Abstracted (n=695)			
Variable		**n=**	**%, mean**	**SD**	**n=**	**%, mean**	**SD**	**n=**	**%, mean**	**SD**	**p-value**
Age, years		**862**	**63.17**	**11.22**	**167**	**64.76**	**10.11**	**695**	**62.78**	**62.78**	**0.041**
Age of death, years		127	73.29	9.20	30	74.20	7.62	97	73.00	9.65	0.535
Gender, male		829	89.75		156	92.95		673	89.00		0.144
Education, years		765	12.69	2.53	148	12.60	2.75	617	12.71	2.48	0.650
Race/Ethnicity	Caucasian	656	79.04		132	84.62		524	77.74		0.216
	African American	112	13.49		14	8.97		98	14.54		
	Hispanic	42	5.06		6	3.85		36	5.34		
	Native American	13	1.57		4	2.56		9	1.34		
	Asian American	2	0.24		0	0.00		2	0.30		
	Other	5	0.60		0	0.00		5	0.74		
Tobacco Use	Never	154	18.55		32	20.51		122	18.10		0.744
	Former	410	49.40		47	49.36		333	49.41		
	Current	263	31.69		77	30.13		216	32.05		
Average DAS		807	3.74	1.15	156	3.75	1.14	651	3.74	1.15	0.886
Average HAQ		815	1.00	0.52	156	1.02	0.55	659	1.00	0.51	0.620
Average disease duration, years		729	17.43	12.48	131	18.86	12.22	598	17.12	12.52	0.147
RF positive		**723**	**87.42**		**144**	**92.31**		**579**	**86.29**		**0.041**
Anti-CCP positive		606	80.59		120	85.71		486	79.41		0.089
Rheumatoid Nodules		463	58.91		92	63.89		371	57.79		0.179

Notes: SD=Standard deviation; Age=age at the time of registry enrollment; Avg. DAS28= average disease activity score, utilizing 28 joint count; Avg. HAQ= average health assessment questionnaire score; Avg. disease duration= average rheumatoid arthritis disease duration at time of registry enrollment; RF= rheumatoid factor; Anti-CCP= anti-citrullinated protein antibody; Rheum. Nodules= rheumatoid nodules; Note: p-value represents test of differences between event and no event populations.

http://repository.academyhealth.org/egems/vol1/iss3/8
DOI: 10.13063/2327-9214.1023

Table 2. Positive predictive value, negative predictive value, sensitivity, and specificity for individual and composite codes for both inpatient and outpatient settings

Event	Code	TP (n)	FP (n)	TN (n)	FN (n)	PPV	95% CI		NPV	Prev	Sens	Spec
MI ICD-9CM												
	410.x	13	2	59	1	0.87	0.60	0.98	0.98	0.23	0.94	0.96
	411.x	1	6	59	1	0.14	0.00	0.58	0.98	0.23	0.72	0.79
	412.x	1	7	59	1	0.13	0.00	0.53	0.98	0.23	0.69	0.79
	413.x	2	8	59	1	0.20	0.03	0.56	0.98	0.23	0.78	0.81
	414.x	3	10	59	1	0.23	0.05	0.54	0.98	0.23	0.80	0.81
	429.2	0	4	59	1	0.00	0.00	0.60	0.98	0.23	0.00	0.77
	v45.81	1	9	59	1	0.10	0.00	0.45	0.98	0.23	0.64	0.79
MI Composite (410.x)		**13**	**2**	**59**	**1**	**0.87**	**0.51**	**0.88**	**0.98**	**0.23**	**0.94**	**0.96**
Stroke ICD-9CM												
	433.11	1	0	60	0	1.00	0.03	0.99	1.00	0.05	1.00	1.00
	433.91	0	1	60	0	0.00	0.00	0.98	1.00	0.05	0.00	0.95
	434.91	9	0	60	0	1.00	0.56	1.00	1.00	0.05	1.00	1.00
	435.x	5	4	60	0	0.56	0.21	0.86	1.00	0.05	1.00	0.98
	436.x	3	10	60	0	0.23	0.05	0.54	1.00	0.05	1.00	0.96
	437.9x	0	2	60	0	0.00	0.00	0.84	1.00	0.05	0.00	0.95
	438.x	6	3	60	0	0.67	0.30	0.93	1.00	0.05	1.00	0.98
Stroke Composite (433.11, 434.91, 435.x, 438.x)		**21**	**7**	**60**	**0**	**0.75**	**0.66**	**1.00**	**1.00**	**0.05**	**1.00**	**0.99**
CABG ICD-9 procedure												
	36.1x	10	0	60	0	1.00	0.69	1.00	1.00	0.01	1.00	1.00
CABG Composite (36.1x)		**10**	**0**	**60**	**0**	**1.00**	**0.69**	**1.00**	**1.00**	**0.01**	**1.00**	**1.00**
PCI												
ICD-9 procedure		20	0	59	1	1.00	0.83	1.00	0.98	0.02	0.53	1.00
	36.06	1	0	59	1	1.00	0.25	1.00	0.98	0.02	0.53	1.00
	36.07	10	0	59	1	1.00	0.69	1.00	0.98	0.02	0.53	1.00
	0.66	9	0	59	1	1.00	0.66	1.00	0.98	0.02	0.53	1.00
CPT		4	0	59	1	1.00	0.40	1.00	0.98	0.02	0.53	1.00
	92973	1	0	59	1	1.00	0.25	1.00	0.98	0.02	0.53	1.00
	92980	2	0	59	1	1.00	0.16	1.00	0.98	0.02	0.53	1.00
	92995	1	0	59	1	1.00	0.25	1.00	0.98	0.02	0.53	1.00
PCI Composite (all codes for PCI)		**24**	**0**	**59**	**1**	**1.00**	**0.85**	**1.00**	**0.98**	**0.02**	**0.53**	**1.00**
Any CV code (composite of all composite codes)		**68**	**9**	**58**	**2**	**0.88**	**0.79**	**0.95**	**0.97**	**0.27**	**0.91**	**0.96**

Notes: TP= true positive; FP = false positive; TN= true negative; FN= false negative; PPV = positive predictive value; CI= confidence interval; NPV= negative predictive value; Prev= prevalence; Sens= sensitivity; Spec= specificity; MI= myocardial infarction; CABG= coronary artery bypass graft; PCI= percutaneous coronary intervention; ICD-9= International Classification of Diseases, Clinical Modification, Ninth Revision; CPT= Current Procedural Terminology; CV= cardiovascular.

Table 3. Positive predictive value, negative predictive value, sensitivity, and specificity for individual and composite codes for inpatient setting

Event	Code	TP (n)	FP (n)	TN (n)	FN (n)	PPV	95% CI		NPV	Prev	Sens	Spec
MI ICD-9CM												
	410.x	7	0	59	1	1.00	0.59	1.00	0.98	0.12	0.89	1.00
	411.x	0	3	59	1	0.00	0.00	0.71	0.98	0.12	0.00	0.87
	412.x	0	0	59	1	0.00	0.00	0.00	0.98	0.12	0.00	0.87
	413.x	0	0	59	1	0.00	0.00	0.00	0.98	0.12	0.00	0.87
	414.x	2	4	59	1	0.33	0.04	0.78	0.98	0.12	0.74	0.91
	429.2	0	0	59	1	0.00	0.00	0.00	0.98	0.12	0.00	0.87
	v45.81	0	0	59	1	0.00	0.00	0.00	0.98	0.12	0.00	0.87
MI Composite (410.x)		**7**	**0**	**59**	**1**	**1.00**	**0.51**	**0.88**	**0.98**	**0.12**	**0.89**	**1.00**
Stroke ICD-9CM												
	433.11	0	0	60	0	0.00	0.00	0.00	1.00	0.05	0.00	0.95
	433.91	0	0	60	0	0.00	0.00	0.00	1.00	0.05	0.00	0.95
	434.91	9	0	60	0	1.00	0.66	1.00	1.00	0.05	1.00	1.00
	435.x	2	0	60	0	1.00	0.16	1.00	1.00	0.05	1.00	1.00
	436.x	1	0	60	0	1.00	0.03	1.00	1.00	0.05	1.00	1.00
	437.9x	0	0	60	0	0.00	0.00	0.00	1.00	0.05	0.00	0.95
	438.x	1	0	60	0	1.00	0.03	1.00	1.00	0.05	1.00	1.00
Stroke Composite (434.91, 435.x, 436.x, 438.x)		**13**	**0**	**60**	**0**	**1.00**	**0.66**	**1.00**	**1.00**	**0.05**	**1.00**	**1.00**
CABG ICD-9 procedure												
	36.1x	10	0	60	0	1.00	0.69	1.00	1.00	0.04	1.00	1.00
CABG composite (36.1x)		**10**	**0**	**60**	**0**	**1.00**	**0.69**	**1.00**	**1.00**	**0.04**	**1.00**	**1.00**
PCI												
ICD-9 procedure		20	0	59	1	1.00	0.83	1.00	0.98	0.05	0.76	1.00
	36.06	1	0	59	1	1.00	0.25	1.00	0.98	0.05	0.76	1.00
	36.07	10	0	59	1	1.00	0.69	1.00	0.98	0.05	0.76	1.00
	0.66	9	0	59	1	1.00	0.66	1.00	0.98	0.05	0.76	1.00
CPT		0	0	59	1	0.00	0.00	0.00	0.98	0.05	0.00	0.95
	92973	0	0	59	1	0.00	0.00	0.00	0.98	0.05	0.00	0.95
	92980	0	0	59	1	0.00	0.00	0.00	0.98	0.05	0.00	0.95
	92995	0	0	59	1	0.00	0.00	0.00	0.98	0.05	0.00	0.95
Composite (all codes for PCI)		**20**	**0**	**59**	**1**	**1.00**	**0.85**	**1.00**	**0.98**	**0.05**	**0.76**	**1.00**
Any CV code (composite of all composite codes)		**50**	**0**	**58**	**2**	**1.00**	**0.93**	**1.00**	**0.97**	**0.18**	**0.87**	**1.00**

Notes: TP= true positive; FP = false positive; TN= true negative; FN= false negative; PPV = positive predictive value; CI= confidence interval; NPV= negative predictive value; Prev= prevalence; Sens= sensitivity; Spec= specificity; MI= myocardial infarction; CABG= coronary artery bypass graft; PCI= percutaneous coronary intervention; ICD-9= International Classification of Diseases, Clinical Modification, Ninth Revision; CPT= Current Procedural Terminology; CV= cardiovascular.

http://repository.academyhealth.org/egems/vol1/iss3/8
DOI: 10.13063/2327-9214.1023

Table 4. Positive predictive value, negative predictive value, sensitivity, and specificity for individual and composite codes for outpatient setting

Event	Code	TP (n)	FP (n)	TN (n)	FN (n)	PPV	95% CI		NPV	Prev	Sens	Spec
MI ICD-9CM												
	410.x	6	2	59	1	0.75	0.35	0.97	0.98	0.23	0.93	0.93
	411.x	1	3	59	1	0.25	0.01	0.81	0.98	0.23	0.81	0.82
	412.x	1	7	59	1	0.13	0.00	0.53	0.98	0.23	0.69	0.79
	413.x	2	9	59	1	0.18	0.02	0.52	0.98	0.23	0.76	0.80
	414.x	1	6	59	1	0.14	0.00	0.58	0.98	0.23	0.72	0.80
	429.2	0	4	59	1	0.00	0.00	0.60	0.98	0.23	0.00	0.77
	v45.81	1	9	59	1	0.10	0.00	0.45	0.98	0.23	0.64	0.79
MI composite (410.x)		**6**	**2**	**59**	**1**	**0.75**	**0.51**	**0.88**	**0.98**	**0.23**	**0.93**	**0.93**
Stroke ICD-9CM												
	433.11	1	0	60	0	1.00	0.03	1.00	1.00	0.05	1.00	1.00
	433.91	0	1	60	0	0.00	0.00	0.98	1.00	0.05	0.00	0.95
	434.91	0	0	60	0	0.00	0.00	0.00	1.00	0.05	0.00	0.95
	435.x	3	5	60	0	0.38	0.09	0.76	1.00	0.05	1.00	0.97
	436.x	2	11	60	0	0.15	0.02	0.45	1.00	0.05	1.00	0.96
	437.9x	0	2	60	0	0.00	0.00	0.84	1.00	0.05	0.00	0.95
	438.x	5	5	60	0	0.50	0.19	0.81	1.00	0.05	1.00	0.97
Stroke composite (433.11, 438.x)		**6**	**5**	**60**	**0**	**0.55**	**0.66**	**1.00**	**1.00**	**0.05**	**1.00**	**0.98**
CABG ICD-9 procedure												
	36.1x	0	0	60	0	0.00	0.00	0.00	1.00	0.00	0.00	1.00
CABG composite (36.1x)		**0**	**0**	**60**	**0**	**0.00**	**0.00**	**0.00**	**1.00**	**0.00**	**0.00**	**1.00**
PCI												
ICD-9 procedure		**0**	**0**	**59**	**1**	**0.00**	**0.00**	**1.00**	**0.98**	**0.00**	**0.00**	**1.00**
	36.06	0	0	59	1	0.00	0.00	0.00	0.98	0.00	0.00	1.00
	36.07	0	0	59	1	0.00	0.00	0.00	0.98	0.00	0.00	1.00
	0.66	0	0	59	1	0.00	0.00	0.00	0.98	0.00	0.00	1.00
CPT		**4**	**0**	**59**	**1**	**0.00**	**0.29**	**1.00**	**0.98**	**0.00**	**0.00**	**1.00**
	92973	1	0	59	1	1.00	0.25	1.00	0.98	0.00	0.22	1.00
	92980	2	0	59	1	1.00	0.16	1.00	0.98	0.00	0.22	1.00
	92995	1	0	59	1	1.00	0.25	1.00	0.98	0.00	0.22	1.00
PCI composite (all codes for PCI)		**4**	**0**	**59**	**1**	**1.00**	**0.85**	**1.00**	**0.98**	**0.00**	**0.22**	**1.00**
Any CV code (composite of all composite codes)		**16**	**7**	**58**	**2**	**0.70**	**0.47**	**0.87**	**0.97**	**0.26**	**0.88**	**0.90**

Notes: TP= true positive; FP = false positive; TN= true negative; FN= false negative; PPV = positive predictive value; CI= confidence interval; NPV= negative predictive value; Prev= prevalence; Sens= sensitivity; Spec= specificity; MI= myocardial infarction; CABG= coronary artery bypass graft; PCI= percutaneous coronary intervention; ICD-9= International Classification of Diseases, Clinical Modification, Ninth Revision; CPT= Current Procedural Terminology; CV= cardiovascular.

Table 5: Summary of total cohort, inpatient, and outpatient composite codings for myocardial infarction, stroke, coronary artery bypass graft, percutaneous coronary intervention, and any cardiovascular event

Event	Code	TP (n)	FP (n)	TN (n)	FN (n)	PPV	95% CI		NPV	Prev	Sens	Spec
Total Cohort												
MI Composite (410.x)		13	2	59	1	0.87	0.60	0.98	0.98	0.23	0.94	0.96
Stroke Composite (433.11, 434.91, 435.x, 438.x)		21	7	60	0	0.75	0.66	1.00	1.00	0.05	1.00	0.99
CABG Composite (36.1x)		10	0	60	0	1.00	0.69	1.00	1.00	0.01	1.00	1.00
PCI Composite (ICD: 36.06, 36.07, 0.66; CPT: 92973,92980, 925995)		24	0	59	1	1.00	0.85	1.00	0.98	0.02	0.51	1.00
Any CV code (composite of all composite codes)		68	9	58	2	0.88	0.79	0.95	0.97	0.27	0.91	0.96
Inpatient												
MI Composite (410.x)		7	0	59	1	1.00	0.51	0.88	0.98	0.12	0.89	1.00
Stroke Composite (434.91, 435.x, 436.x, 438.x)		13	0	60	0	1.00	0.66	1.00	1.00	0.05	1.00	1.00
CABG composite (36.1x)		10	0	60	0	1.00	0.69	1.00	1.00	0.04	1.00	1.00
PCI Composite (ICD: 36.06, 36.07, 0.66; CPT: 92973,92980, 925995)		20	0	59	1	1.00	0.85	1.00	0.98	0.05	0.76	1.00
Any CV code (composite of all composite codes)		50	0	58	2	1.00	0.93	1.00	0.97	0.18	0.87	1.00
Outpatient												
MI composite (410.x)		6	2	59	1	0.75	0.51	0.88	0.98	0.23	0.93	0.93
Stroke composite (433.11, 438.x)		6	5	60	0	0.55	0.66	1.00	1.00	0.05	1.00	0.98
CABG composite (36.1x)		0	0	60	0	0.00	0.00	0.00	1.00	0.00	0.00	1.00
PCI Composite (ICD: 36.06, 36.07, 0.66; CPT: 92973,92980, 925995)		4	0	59	1	1.00	0.85	1.00	0.98	0.00	0.22	1.00
Any CV code (composite of all composite codes)		16	7	58	2	0.70	0.47	0.87	0.97	0.26	0.88	0.90

Notes: TP= true positive; FP = false positive; TN= true negative; FN= false negative; PPV = positive predictive value; CI= confidence interval; NPV= negative predictive value; Prev= prevalence; Sens= sensitivity; Spec= specificity; MI= myocardial infarction; CABG= coronary artery bypass graft; PCI= percutaneous coronary intervention; ICD-9= International Classification of Diseases, Clinical Modification, Ninth Revision; CPT= Current Procedural Terminology; CV= cardiovascular.

In comparing the composite codes between inpatient + outpatient records, inpatient only, and outpatient only, the inpatient only records had the highest PPV and specificity. However, the inpatient records in general had a lower sensitivity than that of either inpatient + outpatient or outpatient alone records. For example, the MI composite coding algorithm (410.x) had a PPV of 0.87 for the inpatient + outpatient, 1.0 for the inpatient alone, and 0.75 for outpatient alone, while the sensitivity was 0.94, 0.89, and 0.93, respectively. With regards to the "any CV code composite," which consists of the MI composite, stroke composite, CABG composite, and PCI composite, the inpatient + outpatient had a PPV of 0.88, NPV of 0.97, sensitivity of 0.91 and specificity of 0.96. The values for the inpatient and outpatient records, respectively, were the following: PPV 1.0 and 0.70, NPV 0.97 and 0.97, sensitivity 0.87 and 0.88, and specificity 1.0 and 0.90.

In general, the procedure codes (ICD-9 procedure and CPT) exceeded the ICD-9 diagnostic codes for PPV.

Discussion

Using the administrative records of RA patients enrolled in the VARA registry, we evaluated the accuracy of ICD-9, ICD-9-Procedure, and CPT codes for identifying acute CV events within a six-month window in inpatient, outpatient and inpatient + outpatient records. When the composite coding for any CV-related event is utilized for both inpatient and outpatient records, the PPV is 0.88 with an NPV of 0.97.There were significant baseline differences between the event and no-event groups in age, gender, reported race/ethnicity, and average RA disease duration. This is not unexpected, as classic risk factors include age and gender,[1] and longer RA disease duration has been associated with increased risk of CV events.[2] As expected, the CV composite codes for the combined inpatient and outpatient records had a PPV between that of the inpatient and outpatient only records, as did sensitivity and specificity. Qualitative review of the medical records revealed that the relatively low sensitivity for the composite CV code for inpatient records is likely due to the number of MI events not being captured by ICD-9 codes in the VA medical records; frequently, this is because patients were evaluated and

http://repository.academyhealth.org/egems/vol1/iss3/8
DOI: 10.13063/2327-9214.1023

treated at outside facilities, and follow-up occurred outside of the six-month window. This "dual care" issue (care received under two or more health care systems) is likely more problematic for emergent conditions, such as MI. We attempted to evaluate this potential limitation by performing an outpatient and outpatient + inpatient records analysis, where a patient would likely report an event to that patient's primary care provider.

Stroke composite coding for outpatient records had a PPV of 0.55, but a sensitivity of 1.0. We hypothesize that the relatively low PPV is due to frequent assignment of "stroke" as either a working diagnosis during emergent events (such as the emergency room visit, which does not create an inpatient record unless the patient is actually admitted) or for follow up for neurologic sequelae of a stroke. For example, a patient presenting to the emergency room with altered mental status may be initially coded as stroke, but subsequent evaluations, including outpatient evaluations, may render a different diagnosis.

Composite coding for any CV event achieved acceptable levels for PPV for inpatient + outpatient (0.88), inpatient (1.0), and outpatient settings (0.70). While certain circumstances may require the identification of specific types of events (MI, stroke, CABG, PCI) as an outcome event, there are many situations that warrant the use of composite coding (which would produce a higher event rate and allow researchers greater power in detecting differences in CV events between therapeutic approaches, for example). However, the specific types of records (inpatient + outpatient, inpatient only or outpatient only) should be determined by the researchers' needs with regards to sensitivity and specificity.

A number of investigators have evaluated the accuracy of administrative coding for CV conditions, including the following: MI;[3,9,18,21-24] stroke and transient ischemic attack;[6-8,11,25-28] CABG surgery;[10,12-14] and various procedures within the spectrum of PCI, including percutaneous transluminal coronary angioplasty (PTCA) and cardiac catheterization.[10,12-14]

Due to the wide range of differences in populations and methodologies used, these studies demonstrate an extraordinarily wide range of reported PPV for each individual ICD-9 or CPT code. For example, the PPV for myocardial infarction ICD-9 codes range from 2.5 to 100 percent,[3] and there is a similarly broad range for incident stroke (<1 to 94 percent).[8,11] Furthermore, very little has been published regarding the sensitivity and specificity of these coding algorithms for the identification of CV disease; and, of the studies that have reported the sensitivity and specificity, there is also a wide range.

While some of these studies examine incident events[6-8] most of these do not specify a specific window of allowable time for the event, making these results less useful for time-to-event analyses. Also of importance, all previously mentioned articles rely exclusively upon data from hospital discharges, which may introduce a significant bias by neglecting to account for CV events documented in outpatient records. This may be particularly relevant to comprehensive health care systems or accountable care organizations, where a substantial proportion of patients live in rural locations. These patients may be managed acutely at a facility that is not within the health care system, and thereby avert detection in administrative data generated solely by the inpatient record. It is therefore vitally important in this setting to capture these events in the outpatient record, which may contain evidence of subacute care for the selected conditions delivered by the patients' primary care physicians or other outpatient staff.

Depending upon the focus of the research, prevalent, incident, or acute coding algorithms may be the most appropriate for an individual researcher's use. We assessed the PPV, NPV, sensitivity, and specificity of acute CV event coding within a specified window of time in inpatient, outpatient, and inpatient + outpatient settings to evaluate our particular codings for time-to-event and other analyses where timing is an essential element.

Administrative data are used for a variety of activities outside of their original billing-related purpose, including assessing quality of care, monitoring health care utilization, and performing health services research (HSR).[14] While these are all important applications, the value of administrative data for these uses depends upon the validity of the codes for the particular application. A number of factors may influence the validity of administrative data including the particular disease state under examination, whether prevalent versus incident disease is being assessed, whether the data are restricted to inpatient sources or whether outpatient data are incorporated, and the health care system in which the data are gathered (a self-contained health care system versus fee-for-service insurance).

As is the case with all studies, our investigation includes a number of strengths and limitations. Notably, our study was conducted within the VA using a well-defined RA population, which will allow application to other patient groups with RA. The VA Computerized Patient Record System (CPRS) is an integrated systemwide EMR that allows comprehensive capture of health care events for patients receiving care within the VA system. However, these results may not be generalizable to other health care settings, and these coding algorithms will need to be validated in other systems. Supporting the generalization or our results, however, is work demonstrating similar clinical characteristics for RA patients from our study cohort compared to more traditional female-predominant community-based cohorts.[29]

Limitations include the fact that our study was performed in a cohort of subjects with RA. While clinicians or coders may exhibit a bias in assigning CV codes to RA patients when compared to non-RA populations, we believe that such a scenario is unlikely. As our subjects are registry participants, they may be systematically different from patients who do not participate in registries. Additionally, we did not independently survey patients directly to determine whether they had experienced a CV event. Such an approach might have identified more CV events (and thus, diminished the observed sensitivity further); however, patient surveys

might have also resulted in substantially more misclassification of events (e.g., classification of angina as MI or of TIAs as stroke), when compared with clinician documentation, and resulted in a lower specificity. Additionally, we chose to assess only the primary coding positions for both inpatient and outpatient records. We did not perform a sample size calculation, but rather chose our numbers to abstract a priori. These statistical choices may be improved upon when this study is repeated in other health care systems. Also, it is important to note that because the VA health care system pays for veterans' care that is delivered outside the VA system, policies are in place to reduce cost by aggressively seeking out veterans hospitalized outside the VA and transferring them to VA facilities. This likely improves the capture of non-VA delivered care and, thus, is anticipated to improve the sensitivity of our approach.

Conclusion

In conclusion, we report the PPV, NPV, sensitivity, and specificity of CV-event coding algorithms in a well-defined RA population of U.S. veterans. Our work has evaluated the proposed composite algorithms for identifying acute CV events in inpatient, outpatient, and inpatient + outpatient settings. The ability to detect incident CV events using reliable codes may further researchers' efforts in health services research, guideline implementation, and health care utilization. These data will enable researchers to make a decision about whether the coding algorithms fit their requirements for use in their research. By providing PPV, NPV, sensitivity, and specificity for different patient settings, our study provides investigators with reliable composite coding for MI, stroke, PCI, and CABG in inpatient, outpatient, and inpatient + outpatient settings. Future avenues of research include reproduction of our study in other health care settings, reproduction of our study in other patient populations, and evaluation of "hybrid" models for identification of CV events, including laboratory values and medications.

Acknowledgements

This work was supported in part by the Rheumatology Research Foundation's Scientist Development Award, the National Institutes of Health (NIH) Loan Repayment Program, and the NIH's T32 AR007534-24.This material is the result of work supported with resources and the use of facilities at the Denver Veterans Affairs Medical Center.

References

1. del Rincon ID, Williams K, Stern MP, Freeman GL, and Escalante A. High incidence of cardiovascular events in a rheumatoid arthritis cohort not explained by traditional cardiac risk factors. Arthritis Rheum 2001;44:2737-2745.
2. Peters MJ, Symmons DP, McCarey D, Dijkmans BA, Nicola P, Kvien TK, McInnes IB, Haentzschel H, Gonzalez-Gay MA, Provan S, Semb A, Sidiropoulos P, Kitas G, Smulders YM, Soubrier M, Szekanecz Z, Sattar N, and Nurmohamed MT. EULAR evidence-based recommendations for cardiovascular risk management in patients with rheumatoid arthritis and other forms of inflammatory arthritis. Ann Rheum Dis 2010;69:325-331.
3. McAlpine R, Pringle S, Pringle T, Lorimer R, and MacDonald TM. A study to determine the sensitivity and specificity of hospital discharge diagnosis data used in the MICA study. Pharmacoepidemiol Drug Saf 1998;7:311-318.
4. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, and Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care 2005;43:480-485.
5. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, and Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130-1139.
6. Ellekjaer H, Holmen J, Kruger O, and Terent A. Identification of incident stroke in Norway: hospital discharge data compared with a population-based stroke register. Stroke 1999;30:56-60.
7. Roumie CL, Mitchel E, Gideon PS, Varas-Lorenzo C, Castellsague J, and Griffin MR. 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;17:20-26.
8. Tirschwell DL and Longstreth WT, Jr. Validating administrative data in stroke research. Stroke 2002;33:2465-2470.
9. Choma NN, Griffin MR, Huang RL, Mitchel EF, Jr., Kaltenbach LA, Gideon P, Stratton SM, and Roumie CL. An algorithm to identify incident myocardial infarction using Medicaid data. Pharmacoepidemiol Drug Saf 2009;18:1064-1071.
10. Jolly S, Kao C, Bindman AB, and Korenbrot C. Cardiac procedures among American Indians and Alaska Natives compared to non-Hispanic whites hospitalized with ischemic heart disease in California. J Gen Intern Med 2010;25:430-434.
11. Rosamond WD, Folsom AR, Chambless LE, Wang CH, McGovern PG, Howard G, Copper LS, and Shahar E. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. Stroke 1999;30:736-743.
12. Vaughan-Sarrazin MS, Wakefield B, and Rosenthal GE. Mortality of Department of Veterans Affairs patients undergoing coronary revascularization in private sector hospitals. Health Serv Res 2007;42:1802-1821.
13. Nallamothu BK, Lu X, Vaughan-Sarrazin MS, and Cram P. Coronary revascularization at specialty cardiac hospitals and peer general hospitals in black Medicare beneficiaries. Circ Cardiovasc Qual Outcomes 2008;1:116-122.
14. Petersen LA, Wright S, Normand SL, and Daley J. Positive predictive value of the diagnosis of acute myocardial infarction in an administrative database. J Gen Intern Med 1999;14:555-558.
15. Egorova N, Vouyouka AG, Quin J, Guillerme S, Moskowitz A, Marin M, and Faries PL. Analysis of gender-related differences in lower extremity peripheral arterial disease. J Vasc Surg 2010;51:372-378.
16. Curtis JR, Jain A, Askling J, Bridges SL, Jr., Carmona L, Dixon W, Finckh A, Hyrich K, Greenberg JD, Kremer J, Listing J, Michaud K, Mikuls T, Shadick N, Solomon DH, Weinblatt ME, Wolfe F, and Zink A. A comparison of patient characteristics and outcomes in selected European and U.S. rheumatoid

http://repository.academyhealth.org/egems/vol1/iss3/8
DOI: 10.13063/2327-9214.1023

arthritis registries. Semin Arthritis Rheum 2010;40:2-14.
17. Mikuls TR, Kazi S, Cipher D, Hooker R, Kerr GS, Richards JS, and Cannon GW. The association of race and ethnicity with disease expression in male US veterans with rheumatoid arthritis. J Rheumatol 2007;34:1480-1484.
18. Austin PC, Daly PA, and Tu JV. A multicenter study of the coding accuracy of hospital discharge administrative data for patients admitted to cardiac care units in Ontario. Am Heart J 2002;144:290-296.
19. Gordis L. Epidemiology. 2009;4th:
20. Weiner MG, Garvin JH, and Ten Have TR. Assessing the accuracy of diagnostic codes in administrative databases: the impact of the sampling frame on sensitivity and specificity. AMIA Annu Symp Proc 2006;1140-
21. Kiyota Y, Schneeweiss S, Glynn RJ, Cannuscio CC, Avorn J, and Solomon DH. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. Am Heart J 2004;148:99-104.
22. McGovern PG, Pankow JS, Shahar E, Doliszny KM, Folsom AR, Blackburn H, and Luepker RV. Recent trends in acute coronary heart disease--mortality, morbidity, medical care, and risk factors. The Minnesota Heart Survey Investigators. N Engl J Med 4-4-1996;334:884-890.
23. Rosamond WD, Chambless LE, Sorlie PD, Bell EM, Weitzman S, Smith JC, and Folsom AR. Trends in the sensitivity, positive predictive value, false-positive rate, and comparability ratio of hospital discharge diagnosis codes for acute myocardial infarction in four US communities, 1987-2000. Am J Epidemiol 12-15-2004;160:1137-1146.
24. Varas-Lorenzo C, Castellsague J, Stang MR, Tomas L, Aguado J, and Perez-Gutthann S. Positive predictive value of ICD-9 codes 410 and 411 in the identification of cases of acute coronary syndromes in the Saskatchewan Hospital automated database. Pharmacoepidemiol Drug Saf 2008;17:842-852.
25. Benesch C, Witter DM, Jr., Wilder AL, Duncan PW, Samsa GP, and 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.
26. Broderick J, Brott T, Kothari R, Miller R, Khoury J, Pancioli A, Gebel J, Mills D, Minneci L, and Shukla R. The Greater Cincinnati/Northern Kentucky Stroke Study: preliminary first-ever and total incidence rates of stroke among blacks. Stroke 1998;29:415-421.
27. Leibson CL, Naessens JM, Brown RD, and Whisnant JP. Accuracy of hospital discharge abstracts for identifying stroke. Stroke 1994;25:2348-2355.
28. Williams GR, Jiang JG, Matchar DB, and Samsa GP. Incidence and occurrence of total (first-ever and recurrent) stroke. Stroke 1999;30:2523-2528.
29. Shahouri SH, Michaud K, Mikuls TR, Caplan L, Shaver TS, Anderson JD, Weidensaul DN, Busch RE, Wang S, and Wolfe F. Remission of rheumatoid arthritis in clinical practice: application of the American College of Rheumatology/European League Against Rheumatism 2011 remission criteria. Arthritis Rheum 2011;63:3204-3215.