Exploration of ICD-9-CM Coding of Chronic Disease within the Elixhauser Comorbidity Measure in Patients with Chronic Heart Failure

by Jennifer Hornung Garvin, PhD, MBA, RHIA, CTR, CPHQ, CCS, FAHIMA; Andrew Redd, PhD; Dan Bolton, MS; Pauline Graham, AS, RHIT; Dominic Roche, AS; Peter Groeneveld, MD, MS; Molly Leecaster, PhD; Shuying Shen, MStat; and Mark G. Weiner, MD

Abstract

Introduction: International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes capture comorbidities that can be used to risk adjust nonrandom patient groups. We explored the accuracy of capturing comorbidities associated with one risk adjustment method, the Elixhauser Comorbidity Measure (ECM), in patients with chronic heart failure (CHF) at one Veterans Affairs (VA) medical center. We explored potential reasons for the differences found between the original codes assigned and conditions found through retrospective review.

Methods: This descriptive, retrospective study used a cohort of patients discharged with a principal diagnosis coded as CHF from one VA medical center in 2003. One admission per patient was used in the study; with multiple admissions, only the first admission was analyzed. We compared the assignment of original codes assigned to conditions found in a retrospective, manual review of the medical record conducted by an investigator with coding expertise as well as by physicians. Members of the team experienced with assigning ICD-9-CM codes and VA coding processes developed themes related to systemic reasons why chronic conditions were not coded in VA records using applied thematic techniques.

Results: In the 181-patient cohort, 388 comorbid conditions were identified; 305 of these were chronic conditions, originally coded at the time of discharge with an average of 1.7 comorbidities related to the ECM per patient. The review by an investigator with coding expertise revealed a total of 937 comorbidities resulting in 618 chronic comorbid conditions with an average of 3.4 per patient; physician review found 872 total comorbidities with 562 chronic conditions (average 3.1 per patient). The agreement between the original and the retrospective *coding* review was 88 percent. The kappa statistic for the original and the retrospective coding review was 0.375 with a 95 percent confidence interval (CI) of 0.352 to 0.398. The kappa statistic for the retrospective coding review and physician review was 0.849 (CI, 0.823–0.875). The kappa statistic for the original coding and the physician review was 0.340 (CI, 0.316–0.364). Several systemic factors were identified, including familiarity with inpatient VA and non-VA guidelines, the quality of documentation, and operational requirements to complete the coding process within short time frames and to identify the reasons for movement within a given facility.

Conclusion: Comorbidities within the ECM representing chronic conditions were significantly underrepresented in the original code assignment. Contributing factors potentially include prioritization of codes related to acute conditions over chronic conditions; coders' professional training, educational level, and experience; and the limited number of codes allowed in initial coding software. This study highlights the need to evaluate systemic causes of underrepresentation of chronic conditions to improve the accuracy of risk adjustment used for health services research, resource allocation, and performance measurement.

Keywords: adverse drug events, comorbidity, complications, heart failure, International Classification of Diseases, risk adjustment, veterans

Introduction

The *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) is primarily used for billing and resource allocation in both Veterans Affairs (VA) and non-VA healthcare organizations[1,2] and is secondarily used to capture specific clinical conditions such as adverse drug events (ADEs),[3] complications of medical and surgical care,[4-11] and risk adjustment of nonrandom patient groups.[12-23] Large-scale administrative databases composed of coded diagnoses and other data associated with the billing process[24] are also used for epidemiology,[25-28] health services research,[29-35] and healthcare policy.[36,37] But these uses of data are secondary because the data are not generated specifically for these purposes. For the purposes of this study. we define administrative databases as those containing data resulting from "administering health care delivery, enrolling members into health insurance plans, and reimbursing for services."[38] ICD-9-CM codes would be part of the database. Missing and incorrectly coded data have been reported in the scientific literature in studies related to research and quality measurement.[39-46] In light of this finding, it is important to evaluate the accuracy of chronic disease capture and the potential reasons for the lack of accuracy because codes used to capture chronic diseases are included in risk-adjustment methods that are increasingly being used in domains relevant to the health information management (HIM) professional. For example, risk adjustment is used to compare organizations' performance measurement data, to predict patient outcomes, and to adjust for severity of illness.[47-51] Many causes of coding errors have been outlined in detail in an article by O'Malley et al., but this process provides only a generic error evaluation, and a more targeted evaluation is needed.[52] The research reported in this study explores the capture of chronic conditions in 100 percent of patient discharges from one VA medical center with the highly prevalent principal diagnosis of chronic heart failure (CHF).

The Elixhauser Comorbidity Measure (ECM) is an important risk-adjustment measure that is frequently used. The ECM was developed using coded data from administrative databases from an all adult, nonmaternal cohort of inpatients discharged from 438 acute care hospitals in California in 1992; software for the ECM is updated on an annual basis by the Agency for Healthcare Research and Quality (AHRQ).[53] In the initial study, Elixhauser et al. found that 31 comorbidities, often chronic conditions, predicted length of stay (LOS), hospital charges, and inpatient mortality.[54] Many studies use the ECM as a method to adjust two or more nonrandom groups so that they can be compared. For example, recently the ECM was used in a nationwide comparison of postsurgical outcomes for patients undergoing laparoscopic roux-en-Y gastric bypass,[55] to examine perioperative and all-cause mortality in patients undergoing radical cystectomy,[56] and to predict outcomes in patients with chronic obstructive pulmonary disease (COPD),[57] among studies in other clinical domains. The ECM is one of several frequently used risk-adjustment methods including the Charlson/Deyo method[58,59] and Johns Hopkins University's Aggregated Diagnosis Groups (ADGs).[60]

Though many of the conditions are chronic, there is limited research about the accuracy of capturing these chronic conditions in this commonly used risk-adjustment measure. The results of an examination of the accuracy of ICD codes within the context of health services research are described in detail by O'Malley et al. They note that accuracy must be assessed within the specific unique situation in which it is used. O'Malley et al. outline the process of coding, examine where error can be introduced, and describe the use of sensitivity, specificity, and positive predictive values of the codes. They also describe using a kappa statistic as a measure of agreement that adjusts for chance agreement if applicable.[61]

Osborn defines a measure as having *sensitivity* if it identifies the property of interest when the property is truly there. Similarly, a measure has *specificity* if it excludes cases when the property of interest is truly absent. Last, positive predictive value (PPV) is defined as the number of cases correctly identified by a measure out of the total number of cases with the property of interest.[62] These metrics have been used in studies determining the accuracy of codes from hypertension,[63] determining the accuracy of heart failure codes,[64] and evaluating accurate coding to predict outcomes in ischemic heart disease.[65] Frogner et al. also found that chronic disease conditions were missing from diagnosis coding, which biased the data used to model reimbursement. Frogner and colleagues used predictive ratios and counts of chronic conditions to compare models to improve reimbursement for the care of patients needing coordinated care.[66]

The research reported in this publication is similar to other studies related to risk adjustment in which the coded data is often several years old. For example, Healthcare Cost and Utilization Project (HCUP) data from 2002–2005 were used in research published in 2012.[67] Similarly, in another study comparing several risk-adjustment methods, coded data from administrative databases from 2004–2008 were used.[68] In comparisons of risk-adjustment methods, the accuracy of predicting the outcomes of death, LOS, and healthcare costs, as well as outcome measures, is often explored. Similarly Frogner et al. use data from 2003 to evaluate risk-adjustment models, with the results published in 2011.[69]

In addition to the quantification of accuracy as described above, qualitative methods such as brainstorming and thematic analysis can be used to answer research questions about factors contributing to missing codes and decreased accuracy.

Brainstorming is a process that a given group undertakes to generate ideas.[70] Themes are defined as recurring issues or ideas that are presented during group interactions and are based on the lived experience of members of the group. The identification and coding of the themes constitutes thematic analysis.[71] Thematic analysis has been used by researchers from many disciplines. For example, thematic analysis has been used to explore responses of midwives to reproductive ethical dilemmas,[72] sexual practices of HIV-positive individuals obtaining antiviral treatment,[73] and patient's judgments about surveillance endoscopy, as well as in many other research studies.[74]

To examine the representation of ICD-9-CM codes for chronic conditions in our data set, we used a combination of quantitative and qualitative techniques, a mixed-methods approach, to explore the accuracy of code assignment of comorbidities associated with the ECM. We undertook quantitative analysis by calculating the sensitivity, specificity, PPV, and kappa statistics on the basis of review by an investigator with domain expertise in coding and by clinician review and comparing those statistics with those for the codes identified retrospectively. After reviewing the quantitative results, the research team used brainstorming to identify potential systemic causes of undercoding of chronic conditions in the inpatient VA setting. We used qualitative data analysis by organizing the brainstorming ideas into themes using thematic analysis.

Methods

This descriptive, retrospective study used a cohort of patients ($N = 181$) representing 100 percent of patients discharged with a principal diagnosis code of CHF from one VA medical center in 2003 as identified through coded data. One admission per patient was used in the study; with multiple admissions, only the first admission was analyzed. The coding professional who assigned the original codes in the data set used the coding guidelines for FY 2003 for discharges from January 1, 2003, through September 30, 2003, and FY 2004 guidelines for discharges from October 1, 2003, through December 31, 2003. The coding software used for the original code assignment did not adjust for severity or risk of mortality. After the patients were identified, an investigator on the research team, blinded to the original codes, retrospectively reviewed the charts to determine if the conditions within the ECM should have been assigned a code. The assignment was assessed on the basis of FY 2007 national VA coding guidelines that included the definitions from the Uniform Hospital Discharge Data Set (UHDDS) including the assignment of codes for "other diagnoses."[75,76] Five physicians received training about the UHDDS and rules for the assignment of codes for "other diagnoses." These clinicians reviewed the medical records and completed an independent assessment for conditions specified in the ECM that were present in the record. The clinician completed an abstraction form to record the findings. One of the five physicians reviewed each record to determine if a condition within the ECM was present and should therefore be coded.

The data from all independent reviews were collected via an abstract form and entered into a research database. The abstract form also collected the discharge disposition, including inpatient death. Because an inpatient death is significant and the patient may have many acute conditions coded, we analyzed differences in code assignment to see if there were any differences based on discharge disposition including death. The original codes assigned at the time of discharge, the conditions found that should have been coded during retrospective review by the investigator on the research team, and the conditions found by the physicians were compared. As has been done in other studies,[77] the percent agreement and kappa statistic were calculated between the original code assignment and the conditions found at the time of retrospective review to assess the degree of concordance and the kappa statistic accounting for chance agreement. Statistical analysis was performed using R version 2.15.1 and Stata 12.[78,79] The concordance of code assignments and conditions found were evaluated and reviewed by three of the study team members. These three team members were experienced with assigning ICD-9-CM codes and with VA coding processes. All three team members reviewed the results of the study. Two team members conducted a brainstorming session via teleconference and developed a list of potential systemic reasons why chronic conditions were not coded in the VA medical center. The list was circulated to the three aforementioned members of the research team (the two original members and one additional member who could not attend the teleconference) via e-mail for edits, corrections, or additions. This approach is an "applied" use of qualitative analysis, in which the methods are used to solve a problem rather than to allow themes to emerge from the data.[80] The method is a theoretical thematic analysis,[81] or a top-down approach to analysis of the data analysis, in contrast to other qualitative analysis techniques that are inductive. We used these thematic analysis techniques to answer the primary research question "What are systemic causes of undercoding of chronic conditions in the VA medical center?"

Analysis

Descriptive statistics were developed regarding patient demographics, prevalence of conditions, accuracy of coded data, and outcome variables that could be examined by the risk-adjustment process, such as death. LOS and readmission rates were also calculated. Agreement measures were developed using Cohen's kappa. The Landis and Koch kappa interpretations were used as a guide of concordance as poor (<0.20), fair (0.21–0.4), moderate (0.41–0.60), or substantial (0.61–0.80).[82] The association between the number of comorbidities and prediction of the outcome readmission was assessed using a *t*-test between readmitted and non-readmitted groups, defined for 30, 60, and 90 days. The association between the number of comorbidities and prediction of mortality and LOS was assessed using simple linear regression. As has been done in other studies,[83,84] the sensitivity, specificity, and PPV were also calculated, using the results found by the researcher with coding expertise as the reference standard for the majority of the analyses to measure accuracy.

The results of the brainstorming discussion were organized and coded into themes to answer the research question. The themes are presented in the following section.

Results

Patient Characteristics

Within the study cohort, the majority of the patients (98.3 percent) were male, and almost two-thirds were African American (65.8 percent). (See Table 1.) The average age for patients was 66.9 years, and ages ranged from 31 to 96 years.

Table 1
Patient Characteristics

Characteristics		Count	%
Gender			
	Male	178	98.3%
	Female	3	1.7%
Race			
	African American	119	65.8%
	White	54	29.8%
	Other	6	3.3%
	Unknown	2	1.1%
Age			
	Below 55 years	28	15.5%
	55–64 years	56	30.9%
	65–74 years	45	24.9%
	75 years or older	52	28.7%

Conditions Identified

A total of 388 comorbid conditions were identified by the original coding professionals. Of these, 305 were chronic comorbid conditions (average 1.7 per patient), and 67 were comorbid conditions that are both chronic and acute (average 0.37 per

patient). The original coding professional thus identified a total of 372 chronic or chronic and acute conditions (average 2.1 per patient). The researcher with coding expertise identified 937 comorbidities coded in the retrospective expert review with an average of 5.2 per patient. Of these, 618 were chronic comorbid conditions (average 3.4 per patient), and 209 conditions were both chronic and acute (average 1.2 per patient), for a total of 827 chronic or chronic and acute conditions (average 4.6 per patient). Comparing the researcher with coding expertise and the original coding professional on the combination of chronic and chronic and acute conditions, the result of the Wilcoxon signed-rank test was highly significant ($p < .0001$), so we concluded that the researcher with coding expertise and the original coding professional differed in the number of comorbid conditions identified. The overall agreement, across all 31 conditions, between the original and the researcher coding expert was 87 percent. The kappa statistic was 0.375, with a 95 percent confidence interval of 0.352 to 0.398. Similarly, overall agreement between the original coding professional and the researcher with coding expertise was 86 percent for chronic conditions, 89 percent for conditions that were both chronic and acute, and 87 percent for conditions that were chronic or chronic and acute across all 31 conditions. The kappa statistics and 95 percent confidence intervals are 0.396 (0.365–0.426) for chronic conditions, 0.400 (0.357–0.443) for conditions that were both chronic and acute, and 0.398 (0.373–0.423) for conditions that were either chronic or chronic and acute.

The top three chronic conditions that the original coding professionals failed to identify are renal failure (74 misses), heart valve disease (65 misses), and hypertension with complications (63 misses). Other important diagnoses such as depression and diabetes were commonly missed as well.

Discharge Disposition and Readmission

Four inpatient deaths, four deaths within 30 days of discharge, zero deaths in 30 to 60 days, and seven deaths within 60 to 90 days were observed. Readmission within 30 days, 60 days, and 90 days was found to be 8.3 percent, 19.3 percent, and 21.6 percent (see Table 2).

Table 2
Death and Readmission

	Count	%
Inpatient death	4	2.2%
Death occurring after discharge		
Within 30 days	4	2.2%
30 days to 60 days	0	0%
60 days to 90 days	7	3.9%
Readmission		
Within 30 days	15	8.3%
Within 60 days	35	19.3%
Within 90 days	39	21.6%

Using the investigator with coding expertise as the reference standard, for 17 chronic or chronic and acute conditions of the 31 total ECM conditions, the overall PPV of the original codes assigned was 80 percent or better. For low-prevalence conditions, the PPV could reach 100 percent but was lower in more common conditions. (See Appendix A.) This less-than-perfect sensitivity reflects miscoding by the original coding professionals. The sensitivity and PPV are provided for the following conditions: deficiency anemia (sensitivity = 29.3 percent, PPV = 100 percent), chronic obstructive pulmonary disease (COPD) (68.3 percent, 91.5 percent), depression (7.4 percent, 100.0 percent), diabetes mellitus with chronic complications (DM-Comp)

(15.6 percent, 83.3 percent), drug abuse (38.9 percent, 100.0 percent), alcohol abuse (31.0 percent, 100.0 percent), heart valve disorders (14.5 percent, 84.6 percent), hypertension with complications (HTN-Comp) (16.0 percent, 85.7 percent), liver disease (66.7 percent, 85.7 percent), lymphoma (33.3 percent, 100.0 percent), obesity (7.5 percent, 100.0 percent), paralysis (33.3 percent, 83.3 percent), psychosis (56.3 percent, 90.0 percent), pulmonary circulation disorders (14.3 percent, 87.5 percent), peripheral vascular disease (31.0 percent, 81.8 percent), renal failure (14.9 percent, 92.9 percent), and weight loss (16.7 percent, 100.0 percent). Five of these conditions had a prevalence of 10 percent or lower (see Appendix A). In cases where there was a high PPV in combination with a low sensitivity, there are comparatively few false positives and many false negatives. A false positive means that the condition is not present but the original coding professional thought it was present. A false negative means that the condition is present but the original coding professional thought it was not present. The combination of a high PPV and a low sensitivity reflects that the original coding professional did not assign a code for conditions that were not present, but they often missed or omitted conditions identified by the researcher with coding expertise.

Predicting Readmission and Length of Stay

Further, we explored using the sum of comorbid conditions in terms of predicting readmission at 90 days, and no statistically significant association was observed in any of the groups (see Table 3).

Table 3
Predicting Readmission and Length of Stay

Variable	Sum of Chronic Comorbid Condition Codes	
	Expert	**Original**
Readmission within 30 days		
No	4 (3, 6)	2 (1, 3)
Yes	5 (3, 6)	2 (1, 3)
p-value	1.000	1.000
Readmission within 60 days		
No	5 (3, 6)	2 (1, 3)
Yes	4 (3, 6)	2 (2, 2)
p-value	1.000	1.000
Readmission within 90 days		
No	4.5 (3, 6)	2 (1, 3)
Yes	5 (3, 6)	2 (1, 3)
p-value	1.000	1.000
Length of stay (LOS) log regression		
Estimate	0.1787	0.1469
R^2	0.1349	0.0297
p-value	<0.0001	0.012

Note: Values for sums of chronic comorbid conditions are reported as median (interquartile range [IQR]).

LOS ranged from 1 day to 48 days, excluding one patient whose length of stay exceeded a year, with the first quartile at 2 days and the third quartile at 8 days. The log of LOS was modeled using linear regression with the sum of comorbid conditions that were either chronic or chronic and acute as the explanatory variable. One model used the sum of expert codes as predictor, and the other used the sum of original codes. Both predictors were significant ($p < .05$) in their respective models, but the sum of expert codes was a better predictor of LOS than the sum of original codes was (see Table 3). Models based on weighting of comorbidities would likely be more significant because more serious conditions may predict LOS more precisely. Table 3 also presents the analysis of the sum of conditions coded by the researcher with coding expertise and the original coding professionals in relation to readmission to the hospital within 30, 60, or 90 days. Test results are reported with median, interquartile range (IQR), and *p*-value adjusted for multiple comparisons. The test results are not significant.

Agreement

Substantial agreement based on the kappa statistic was achieved in two high-prevalence conditions, COPD and diabetes mellitus, with fair agreement for anemia between the researcher with coding expertise and the original coding professional (Table 4). Clinically significant conditions such as heart valve disease and hypertension (when complications are coded or when the condition is coded without complications) did not have high agreement. The original coder and the physician reviewer had similar results, with substantial agreement on COPD, moderate agreement on diabetes mellitus, and fair agreement on anemia (see Table 5). Table 6 shows much higher agreement between the physician reviewer and the researcher with coding expertise: the agreement between them is either substantial or almost perfect for all high-prevalence conditions.

Table 4
High-Prevalence Chronic Comorbid Condition Codes: Original Coder vs. Researcher with Coding Expertise

Comorbidity Code	Prevalence[a]	Percent Agreement	Probability of Random Agreement	Kappa
Renal failure	48%	59%	52%	0.143
Heart valve disease	42%	63%	57%	0.142
HTN-Comp[b]	41%	64%	57%	0.160
COPD	35%	87%	57%	0.689
HTN	35%	64%	50%	0.267
DM	34%	82%	52%	0.621
Anemia	32%	77%	65%	0.360
Obesity	29%	73%	70%	0.104
Pulm Circ	27%	76%	71%	0.184
DM-Comp[b]	18%	85%	80%	0.220

Abbreviations: HTN, hypertension; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; Pulm Circ, pulmonary circulation disorders; Comp, complications.
[a]Prevalence for this table comes from the expert coder.
[b]Denotes when complications of hypertension or diabetes are present.

Table 5
High-Prevalence Chronic Comorbid Condition Codes: Original Coder vs. Physician Reviewer

Comorbidity Code	Prevalence[a]	Percent Agreement	Probability of Random Agreement	Kappa
Renal failure[b]	48%	61%	54%	0.156
Heart valve disease	42%	65%	61%	0.091

HTN-Comp[b]	41%	66%	59%	0.176
COPD	35%	87%	59%	0.673
HTN	35%	61%	50%	0.210
DM	34%	77%	52%	0.514
Anemia	32%	77%	67%	0.301
Obesity	29%	74%	73%	0.039
Pulm Circ	27%	79%	74%	0.179
DM-Comp[b]	18%	85%	82%	0.129

Abbreviations: HTN, hypertension; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; Pulm Circ, pulmonary circulation disorders; Comp, complications.
[a]Prevalence for this table comes from the expert coder.
[b]Denotes when complications of hypertension or diabetes are present.

Table 6
High-Prevalence Chronic Comorbid Condition Codes: Researcher with Coding Expertise vs. Physician Reviewer

Comorbidity	Prevalence[a]	Percent Agreement	Probability of Random Agreement	Kappa
Renal failure	46%	92%	50%	0.845
Heart valve disease	37%	94%	52%	0.873
HTN-Comp[b]	39%	90%	52%	0.794
COPD	30%	96%	56%	0.900
HTN	30%	86%	56%	0.685
DM	31%	90%	56%	0.760
Anemia	29%	93%	57%	0.831
Obesity	26%	92%	60%	0.807
Pulm Circ	23%	94%	62%	0.839
DM-Comp[b]	15%	92%	72%	0.721

Abbreviations: HTN, hypertension; COPD, chronic obstructive pulmonary disease; DM, diabetes mellitus; Pulm Circ, pulmonary circulation disorders; Comp, complications.
[a]Prevalence for this table comes from the physician coder.
[b]Denotes when complications of hypertension or diabetes are present.

However, when the respective subcategories for diabetes and hypertension are combined (see Table 7), the agreement is much higher between the researcher with coding expertise and the original coding professional. There appears to be agreement that these conditions are present, but less agreement about whether or not there are associated complications.

Table 7
Diabetes Mellitus (DM) and Hypertension (HTN) Category Codes Combined: Original Coder vs. Researcher with Coding Expertise

Comorbidity	Percent Agreement	Probability of Random Agreement	Kappa

HTN COMB	77%	54%	0.512
Compared to:			
HTN	64%	50%	0.267
HTN-Comp[a]	64%	57%	0.160
DM COMB	93%	50%	0.857
Compared to:			
DM	82%	52%	0.621
DM-Comp[a]	85%	80%	0.220

Abbreviations: HTN, hypertension; COMB, all cases with and without complications combined DM, diabetes mellitus; Comp, complications.
[a]Denotes when complications of hypertension or diabetes are present.

The code assignment of two conditions, renal failure (74 misses) and complications of hypertension (63 misses), was likely influenced by a change of coding guidelines between FY 2003 and FY 2006. However, although different conditions were added to the ECM, all of the conditions contained within the record should still have been coded because of the requirements of the official coding guidelines and the UHDDS. Other (additional) diagnoses are required to be coded, and they are defined in the UHDDS as "all conditions that coexist at the time of admission, that develop subsequently, or that affect the treatment received and/or the length of stay." For reporting purposes the definition of "other diagnoses" is interpreted as additional conditions that affect patient care in terms of requiring clinical evaluation, therapeutic treatment, diagnostic procedures, extended length of hospital stay, or increased nursing care and/or monitoring.[85] The only new condition that was captured via a new ICD-9-CM code definition was chronic kidney disease (CKD). The coding of CKD affects the coding of hypertension with complications (HTN-Comp) because the 403.xx code category includes hypertension with CKD. Because of this change, the only conditions in the ECM explored in this study that would have a higher prevalence would be the renal failure and HTN-Comp categories.

The original coder and the physician reviewer have moderate agreement for the combined hypertension category and substantial agreement for the combined diabetes category (see Table 8). The researcher with coding expertise and the physician reviewer have substantial agreement for the combined hypertension category and almost perfect agreement for the combined diabetes category (see Table 9).

Table 8
Diabetes Mellitus (DM) and Hypertension (HTN) Category Codes Combined: Original Coder vs. Physician Reviewer

Comorbidity	Percent Agreement	Probability of Random Agreement	Kappa
HTN COMB	77%	54%	0.512
Compared to:			
HTN	64%	50%	0.267
HTN-Comp[a]	64%	57%	0.160
DM COMB	93%	50%	0.857
Compared to:			
DM	82%	52%	0.621
DM-Comp[a]	85%	80%	0.220

Abbreviations: HTN, hypertension; COMB, all cases with and without complications combined DM, diabetes mellitus; Comp, complications.
[a]Denotes when complications of hypertension or diabetes are present.

Table 9
Diabetes Mellitus (DM) and Hypertension (HTN) Category Codes Combined: Researcher with Coding Expertise vs. Physician Reviewer

Comorbidity	Percent Agreement	Probability of Random Agreement	Kappa
HTN COMB	88.4%	60%	0.710
Compared to:			
HTN	86%	56%	0.685
HTN-Comp[a]	90%	52%	0.794
DM COMB	92%	49%	0.837
Compared to:			
DM	90%	56%	0.760
DM-Comp[a]	92%	72%	

Abbreviations: HTN, hypertension; COMB, [all cases with and without complications combined]; DM, diabetes mellitus; Comp, complications.
[a]Denotes when complications of hypertension or diabetes are present.

An overall test of validity for the original coding professional, the researcher with coding expertise, and the physician reviewer appears in Appendix B. Results show high validity by both the physician reviewer and the researcher with coding expertise, with sensitivity and PPVs greater than 80 percent. Original coder validity was tested using both the physician reviewer and the researcher with coding expertise as the reference standard. In both of these analyses the original coder did not perform at the same level as the other coders: sensitivity values were approximately 30 percent, while PPVs were 74 percent and 65 percent.

Themes Related to Decreased Chronic Condition and Comorbidity Code Capture

Factors were identified after team review of the study data using brainstorming. Factors were primarily based on the team members' coding experience. Three members of the research team participated. The professional background of those in the brainstorming group was as follows: two members were HIM professionals with certified coding specialist (CCS) credentials, the third was a physician, and all three members were involved in health services research within the VA. The average number of years of coding experience for members of the group was 16.7 years with a range of 5 to 24 years, and the average number of years working with the VA was 10.6 years with a range of 5 to 16 years. The research team identified several factors that influence the capture of chronic conditions, including familiarity with inpatient VA guidelines as well as familiarity with non-VA-specific coding guidelines,[86,87] the quality of documentation from which to discern the presence of chronic conditions so that codes may be accurately assigned, operational requirements to complete the coding process within short time frames and identify the reasons for movement within a given facility with an ICD-9-CM code, coding staff shortages resulting in backlogs of discharge records, lack of knowledge on the behalf of coding staff of secondary use of administrative data (such as for risk adjustment), and limited numbers of codes that could be entered in the coding software at the time of the study. We also identified that many conditions are present in the record and only the conditions that are relevant to the current discharge, including patient movement, are given priority when codes are assigned. The code assignment focuses on making sure that the events that occurred during the patient's stay are captured, rather than focusing on comprehensively identifying chronic diseases. This focus is reflected in the coding guidelines,[88] in which emphasis is placed on acute conditions. Furthermore, in the computer system used during 2003 by the VA,[89] only a specified number of diagnoses could be inserted in the coding software, which may have resulted in a prioritization of acute conditions over chronic

conditions. Further, the VA requires that patient movements among bed sections in the facility have associated diagnoses coded.[90] Clear documentation of complications of hypertension and diabetes is not always present.

The following are examples of factors identified in the brainstorming process:

- We have had a lot of coder education since in the last several years.
- Because there was a backlog in the department, some of the outpatient coders had to code the inpatient records. They were not used to assigning a lot of codes.
- It takes extra steps to add more codes than the 15 allowed in the coding software.
- Coders may not know how the codes are used for research.
- Chronic conditions may be bumped.
- In the VA, we have to consider the patient movement when assigning codes because we are required to show why the patient was transferred to a specific unit. For example, we have to reflect a psych code if the patient is transferred from the medical unit to the psychiatric unit.

Discussion

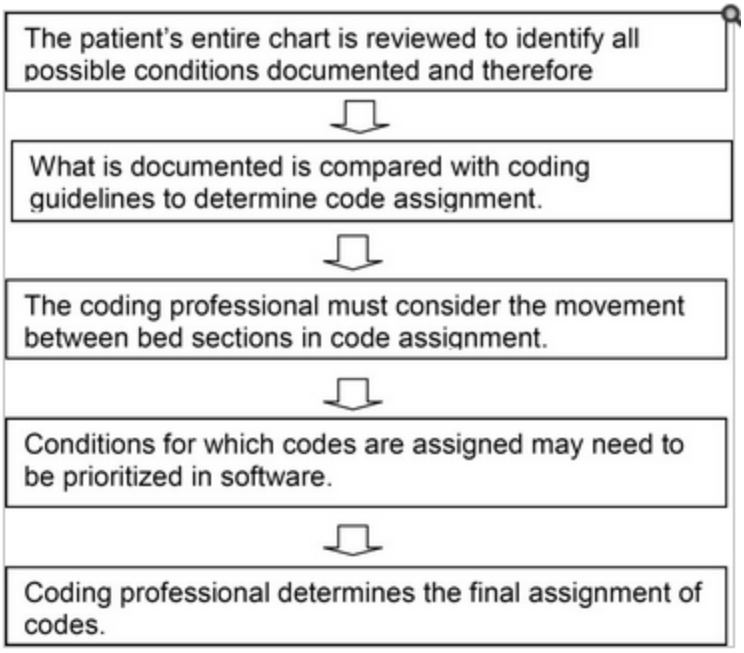
In our research, we found underreporting of chronic conditions. Similarly, recent research found that chronic disease is underreported in administrative data, requiring changes to Medicare risk-adjustment processes.[91] Similarly, two methods have been used to improve the ability of coded data sets to capture patient safety indicators (PSIs), comorbidities, and ADEs accurately. The first is to combine administrative databases with clinical data. LaTour and Maki define clinical data as data that are captured during the diagnosis and treatment process.[92] The clinical data that were added to coded data in this case are laboratory data.[93,94] The process of combining administrative and clinical data has also been piloted in Minnesota with increases in accuracy comparisons of risk-adjusted mortality and risk-adjusted complications.[95] The second method is to add a variable indicating whether the patient's coded diagnoses were present on admission (POA) to differentiate between comorbidities and complications.[96-98]

Limitations of this study include the following: the principal investigator was the individual with coding expertise, this is a single-site study, and the inclusion and exclusion of codes changed between 2003 (data of discharge) and 2007 (date of the coding guidelines used). Other limitations include that we could not calculate the ECM in our small cohort and so we could not measure the effect of underreporting comorbidities based on administrative data directly. However, in terms of the evaluating the impact on LOS, we measured how well the sum (count) of comorbid condition codes predicted this outcome, and the inclusion of a greater number of comorbidity codes resulted in significantly better prediction. A future direction would be to conduct a larger study so that the effect of decreased comorbidity code capture on the outcomes predicted by ECM could be measured. A review of the literature indicates that this appears to be a gap in the current knowledge.

Despite the inability to compute the ECM directly, this study highlights the problem of incomplete data when computing the ECM. Researchers should be aware of this limitation when risk adjusting patients of different populations. When patterns of missing comorbidity codes are different in study populations, differences in the computed risk will reflect differences in coding practices and patterns of missing data rather than the true desired difference in risk. Possible sources of differences in coding practices that we have discovered include the time of coding, but other sources are not unreasonable either. Our population was taken only from the Veterans Health Administration system; other healthcare systems may have different coding practices and training and may have additional sources of differences in patterns of missing codes.

O'Malley et al.[99] identified several factors that impact coding quality, including coding professionals' knowledge, training, and experience. Figure 1 builds upon the foundation provided by O'Malley et al. Our research found additional systemic contributors to error, including operational requirements to complete the coding process within short time frames, the requirement to identify the reasons for movement within a given facility with an ICD-9-CM code, coding staff shortages resulting in backlogs of discharge records, lack of knowledge on the part of the coding staff of secondary uses of administrative data such as for risk adjustment, and limited numbers of codes that could be entered in the coding software at the time of the study.

Figure 1



Process of Assigning ICD-10 Codes within the VA by Coding Staff

We found that the kappa statistic was higher for those who have a doctoral-level degree, even though the investigator with coding expertise had a doctor of philosophy degree whereas the physicians are doctors of medicine. This observation may indicate that higher educational level is a contributing factor to higher recognition of chronic conditions, thus supporting the efforts to move the HIM curriculum from the associate and baccalaureate levels and to strengthen informatics and leadership competencies among HIM professionals.[100] In addition to preparing HIM professionals with master's degrees to take the senior-level positions needed in HIM in the future, a higher educational level may also strengthen core areas of competence in coding-related secondary uses of data such as for risk adjustment and data analytics.

Conclusions

Codes for comorbidities representing chronic conditions within the ECM were significantly underrepresented in the original code assignment. Contributing factors potentially include prioritization of codes related to acute conditions over chronic conditions; the coding professionals' training, educational level, and experience; and the limited number of codes allowed in initial coding software. This study highlights the need to evaluate systemic causes of underrepresentation of chronic conditions to improve the accuracy of risk adjustment used for health services research, resource allocation, and performance measurement.

Acknowledgments

The authors would like to acknowledge and thank the Salt Lake City VA Health System Informatics, Decision Enhancement, and Surveillance (IDEAS) Center staff members Linsey Ben-Ami, MPH, program analyst, and Marianne Madsen, MS, senior academic writer, for their assistance in reviewing and editing our manuscript. The authors would also like to thank the Coding and Compliance Department at the Philadelphia VA Medical Center. The Jennifer Garvin would also like to thank the staff of the Center for Health Equity Research and Promotion (CHERP) for their assistance with this research during her fellowship.

This publication is based upon work supported by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, HSR&D, Grant #TMI 95-642. Dr. Garvin was a VA HSR&D medical informatics post-doctoral fellow at the CHERP at the Philadelphia VA Medical Center the time this research was conducted and is currently a research health scientist and core investigator at the research center at the Salt Lake City VA Health System. The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs, the University of Utah School of Medicine, the University of Utah Department of Biomedical Informatics, or the University of Pennsylvania School of Medicine.

Appendix A

Summary of Condition Codes and Associated Accuracy Based on Expert Coding

Elixhauser Comorbidity Measure (ECM) Based on Version 3.4/2009					
Comorbidity or Classification	**ICD-9-CM Code Range**	**Prevalence**	**Sensitivity**	**Specificity**	**Positive Predictive Value**
Adverse Drug Events (ADE)[a]	960–979, E930–E949	13.3%	4.2%	99.4%	99.4%
Anemia (Deficiency) including acute blood loss anemia[b]	280.1–280.9 285.21–285.29 285.9	32.0%	29.3%	100.0%	100.0%
Blood Loss Anemia–chronic blood loss and anemia of pregnancy[bc]	280.0 648.20–648.24	0.6%	0.0%	100.0%	NA
CA Primary-Solid Tumor without Metastasis[d]	140.0–172.9 174.0–175.9 179–195.8, 209.00-209.24, 209.25-209.3, 258.01-258.03	10.5%	26.3%	98.8%	71.4%
CA Secondary (Metastatic Cancer)[cd]	196.0–199.1, 789.51	0.6%	0.0%	100.0%	NA
Chronic Peptic Ulcer Disease (includes bleeding only if obstruction is also present)[cd]	531.41, 531.51, 531.61, 531.70, 531.71, 531.91, 532.41, 532.51, 532.61, 532.70, 532.71, 532.91, 533.41, 533.51, 533.61, 533.70, 533.71, 533.91, 534.41, 534.51, 534.61, 534.70, 534.71, 534.91	1.7%	0.0%	98.9%	0.0%
Coagulation deficiency[cd]	286.0–286.9, 287.1–287.3–287.5, 289.84, 649.30–649.34	1.7%	0.0%	100.0%	NA
Complications[ce]	996–999	7.1%	0.0%	100.0%	0%
Congestive Heart Failure[df]	398.91, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93 428.0–428.9	100%	NA	NA	NA
Chronic Obstructive Pulmonary Disease (COPD)[d]	490–492.8 493.00–493.92 494–494.1495.0–505 506.4	34.8%	68.3%	96.6%	91.5%
Depression[d]	300.4, 301.12 309.00, 309.1 311	14.9%	7.4%	100.0%	100.0%
Diabetes Mellitus without Chronic Complications[d]	249.00–249.31, 250.00–250.33 648.00–648.04	33.7%	88.5%	78.3%	67.5%
Diabetes Mellitus with Chronic Complications[d]	249.40–249.91, 250.40–250.93, 775.1	17.7%	15.6%	99.3%	83.3%
Drug Dependence and Abuse[b]	292.0292.82–292.89 292.9304.00–304.93305.20–	9.9%	38.9%	100.0%	100.0%

	305.93648.30–648.34				
ETOH (Alcohol)Abuse[b]	291.0–291.3291.5 291.8291.81, 291.82 291.89291.9303.00–303.93305.00–305.03	16.0%	31.0%	100.0%	100.0%
Fluid and Electrolyte Disorders[a]	276.0–276.9	40.3%	17.8%	99.1%	92.9%
Heart Valve (Valvular) Disease[d]	093.20–093.24394.0–397.1 397.9424.0–424.99746.3–726.6 V42.2V43.3	42.0%	14.5%	98.1%	84.6%
HIV and AIDS (Acquired Immunodeficiency Syndrome)[cd]	042–044.9	1.1%	0.0%	100.0%	NA
HTN-Hypertension[d]	401.1, 401.9 642.00–642.24	34.8%	68.3%	61.0%	48.3%
HTN-Hypertension (complicated)[d]	401.0, 402.00–405.99, 437.2, 642.10–624.24, 642.70–642.94	41.4%	16.0%	98.1%	85.7%
Hypothyroidism[d]	243–244.2, 244.8 244.9	5.0%	66.7%	98.3%	66.7%
Liver Disease, Including Acute or Chronic Viral[b]	070.22, 070.23,070.33, 070.44, 070.54, 456.0,456.1, 456.20, 456.21, 571.0,571.2, 571.3,571.40–571.49,571.5, 571.6, 571.8, 571.9, 572.3, 572.8, V42.7	9.9%	66.7%	98.8%	85.7%
Lymphoma[d]	200.00–202.38202.50–203.01, 203.02–203.82,203.8–203.81 238.6,(plasmacytoma) 273.3(macroglobulinemia)	1.7%	33.3%	100.0%	100.0%
Obesity[d]	278.0, 278.00, 278.01,649.10–649.14,793.91, V85.30–V85.4, V85.54	29.3%	7.5%	100.0%	100.0%
Other Neuro Disorders[c]	330.1–331.9, 332.0, 333.4, 333.5, 333.71–333.79, 333.85, 333.94, 443.0–335.9, 338.0, 340, 341.1–341.9, 345.00–345.11, 345.2–345.3, 345.40–345.91, 347.00–347.01, 347.10–347.11, 649.40–649.44, 768.7, 780.3, 780.31, 780.32, 780.39,780.97 784.3, 784.3	3.9%	0.0%	99.4%	0.0%
Paralysis[d]	342.0–344.9 438.20–438.53, 780.72	8.3%	33.3%	99.4%	83.3%
Psychoses[d]	295.00–298.9 299.10 299.11	8.8%	56.3%	99.4%	90.0%
Pulmonary Circulation Disorders[b]	415.11–415.19, 416.0–416.9 417.9	27.1%	14.3%	99.2%	87.5%
Peripheral Vascular Disease[b]	440.0–440.9, 441.00–441.9, 442.0–442.9, 443.1–443.9,	16.0%	31.0%	98.7%	81.8%

	447.1, 57.1,557.9, V43.4				
Renal Failure[d]	403.01, 403.11, 403.91, 404.02,404.03, 404.12, 404.13, 404.92, 404.93, 585.3, 585.4, 585.5, 585.6, 585.9, 586, V42.0, V45.1, V45.11, V45.12, V56.0–V56.32,V56.8	48.1%	14.9%	98.9%	92.9%
Rheumatoid arthritis/collagen vascular diseases[cd] connective tissue disorders	701.0–710.0–710.9, 714.0–714.9–720.0–720.9, 725	2.2%	0.0%	100.0%	NA
Weight Loss[d]	260–263.9, 783.21–783.22	3.3%	16.7%	100.0%	100.0%

[a]Acute condition.
[b]Acute and chronic conditions.
[c]These conditions were never coded in our corpus by the original coding professionals (indicated by "NA").
[d]Chronic condition.
[e]Not in original ECM.
[f]Note that this condition is one of the ECM comorbid conditions, but in this case the cohort was determined based on these codes so the accuracy measures are not provided.

Appendix B

This table reports sensitivity, specificity, and positive predictive values. Values for the researcher with coding expertise use the physician reviewer as the reference standard; values for the physician reviewer use the researcher with coding expertise as the reference standard; and values for the original coding professional are reported using both the researcher with coding expertise and the physician reviewer as reference standards. Values are reported with the 95% confidence interval in parentheses.

	Sensitivity	**Specificity**	**Positive Predictive Value**
Researcher with coding expertise	0.91 (0.88–0.92)	0.97 (0.96–0.97)	0.84 (0.82–0.87).
Physician reviewer	0.84 (0.82–0.87)	0.98 (0.98–0.99)	0.91 (0.88–0.92)
Original coding professional vs. researcher with coding expertise	0.31 (0.28–0.34)	0.98 (0.97–0.98)	0.74 (0.70–0.79)
Original coding professional vs. physician reviewer	0.29 (0.26–0.32)	0.97 (0.97–0.98)	0.65 (0.60–0.70)

Contributor Information

Jennifer Hornung Garvin, PhD, MBA, RHIA, CTR, CPHQ, CCS, FAHIMA, is a core research investigator in the Informatics, Decision Enhancement, and Surveillance (IDEAS) Center at the Salt Lake City VA Healthcare System and an Associate Professor in the Department of Biomedical Informatics at the University of Utah School of Medicine in Salt Lake City, UT.

Andrew Redd, PhD, is a statistician at the Informatics, Decision Enhancement, and Surveillance (IDEAS) Center at the Salt Lake City VA Healthcare System and an Assistant Professor at the University of Utah Schoolof Medicine in Salt Lake City,

UT.

Dan Bolton, MS, is a statistician at the Informatics, Decision Enhancement, and Surveillance (IDEAS) Center at the Salt Lake City VA Healthcare System and the University of Utah School of Medicine in Salt Lake City, UT.

Pauline Graham, AS, RHIT, is the acting assistant chief of health administration services at the Philadelphia VA Medical Center in Philadelphia, PA.

Dominic Roche, AS, is a software engineer for the Center for Medicare and Medicaid Services and a former data analyst in the Center for Health Equity Research and Promotion at the Philadelphia VA Medical Center in Philadelphia, PA.

Peter Groeneveld, MD, MS, is a core researcher in the Center for Health Equity Research and Promotion at the Philadelphia VA Medical Center and is an Associate Professor of Medicine at the University of Pennsylvania in Philadelphia, PA.

Molly Leecaster, PhD, is a statistician at the Informatics, Decision Enhancement, and Surveillance (IDEAS) Center at the Salt Lake City VA Healthcare System and an Assistant Professor at the University of Utah School of Medicine in Salt Lake City, UT.

Shuying Shen, MStat, is a statistician at the Informatics, Decision Enhancement, and Surveillance (IDEAS) Center at the Salt Lake City VA Healthcare System and the University of Utah School of Medicine in Salt Lake City, UT.

Mark G. Weiner, MD, is a researcher at AstraZeneca Pharmaceuticals, Research and Development Information formerly at the Center for Health Equity Research and Promotion at the Philadelphia VA Medical Center and an Associate Professor of Medicine at the University of Pennsylvania in Philadelphia, PA.

Notes

[1] Veterans Health Administration. *Coding Handbook*. Washington, DC, February 2008.

[2] American Hospital Association (AHA). *Coding Clinic for ICD-9-CM*. Chicago, IL: AHA, 2008.

[3] Moriomoto, T., T. K. Ghandi, A. C. Seger, T. C. Hsieh, and D. W. Bates. "Adverse Drug Events and Medication Errors: Detection and Classification Methods." *Quality and Safety in Health Care* 13, no. 4 (2004): 306–14.

[4] Jollis, J. G., M. Ancukiewicz, E. R. Delong, D. B. Pryor, L. H. Muhlbaier, and D. B. Mark. "Discordance of Databases Designed for Claims Payment Versus Clinical Information Systems: Implications for Outcomes Research." *Annals of Internal Medicine* 119, no. 8 (1993): 844–50.

[5] Houchens, R. L., A. Elixhauser, and P. S. Romano. "How Often Are Potential Patient Safety Events Present on Admission?" *Joint Commission Journal on Quality and Patient Safety* 34, no. 3 (2008): 154–63.

[6] Naessens, J. M., C. R. Campbell, B. Berg, A. R. Williams, and R. Culbertson. "Impact of Diagnosis-Timing Indicators on Measures of Safety, Comorbidity, and Case Mix Groupings from Administrative Data Sources." *Medical Care* 45, no. 8 (2007): 781–88.

[7] Polancich, S., E. Restrepo, and J. Prosser. "Cautious Use of Adminstrative Data for Decubitus Ulcer Outcome Reporting." *American Journal of Medical Quality* 21, no. 4 (2006): 262–68.

[8] Bahl, V., M. A. Thompson, T. Y. Kau, H. M. Hu, and D. A. Campbell Jr. "Do the AHRQ Patient Safety Indicators Flag Conditions That Are Present at the Time of Hospital Admission?" *Medical Care* 46, no. 5 (2008): 516–22.

[9] Agency for Healthcare Research and Quality. *Guide to Patient Safety Indicators*. 2006. Available at http://www.qualityindicators.ahrq.gov/Downloads/Software/SAS/V30a/psi_guide_v30a.pdf.

[10] Pine, M., B. Jones, and Y. B. Lou. "Laboratory Values Improve Predictions of Hospital Mortality." *International Journal for Quality in Health Care* 10, no. 6 (1998): 491–501.

[11] Borzecki, A. M., A. T. Wong, E. C. Hickey, A. S. Ash, and D. R. Berlowitz. "Identifying Hypertension-Related Comorbidities from Administrative Data: What's the Optimal Approach?" *American Journal of Medical Quality* 19, no. 5 (2004): 201–6.

[12] Selim, A. J., L. E. Kazis, W. Rogers, S. Qian, J. A. Rothendler, A. Lee, X. S. Ren, S. C. Haffer, R. Mardon, D. Millter, A. Spiro III, B. J. Selim, and B. G. Fincke. "Risk-adjusted Mortality as an Indicator of Outcomes: Comparison of the Medicare Advantage Program with the Veterans' Health Administration." *Medical Care* 44, no. 4 (2006): 359–65.

[13] Rumsfeld, J. S., P. G. Jones, M. A. Whooley, M. D. Sullivan, B. Pitt, W. S. Weintraub, and J. A. Spertus. "Depression Predicts Mortality and Hospitalization in Patients with Myocardial Infarction Complicated by Heart Failure." *American Heart Jounal* 150, no. 5 (2005): 961–67.

[14] Desai, R. A., E. A. Stefanovics, and R. A. Rosenheck. "The Role of Psychiatric Diagnosis in Satisfaction with Primary Care: Data from the Department of Veterans Affairs." *Medical Care* 43, no. 12 (2005): 1208–16.

[15] Mortensen, E. M., M. I. Restrepo, A. Anzueto, and J. Pugh. "The Impact of Prior Outpatient ACE Inhibitor Use on 30-day Mortality for Patients Hospitalized with Community-acquired Pneumonia." *BMC Pulmonary Medicine* 5, no. 12 (2005).

[16] Weeks, W. B., D. A. Bazos, D. M. Bott, R. Lombardo, M. J., Racz, E. L. Hannan, and E. S. Fisher. "New York's Statistical Model Accurately Predicts Mortality Risk for Veterans Who Obtain Private Sector CABG." *Health Services Research* 40, no. 4 (2005): 1186–96.

[17] Maciejewski, M. L., C. F. Liu, A. Derleth, M. McDonell, S. Anderson, and S. D. Fihn. "The Performance of Administrative and Self-Reported Measures for Risk Adjustment of Veterans Affairs Expenditures." *Health Services Research* 40, no. 3 (2005): 887–904.

[18] Kaafarani, H. M., K. M. Itani, L. A. Petersen, J. Thornby, and D. H. Berger. "Does Resident Hours Reduction Have an Impact on Surgical Outcomes?" *Journal of Surgical Research* 126, no. 2 (2005): 167–71.

[19] Johnston, J. A., D. P. Wagner, S. Timmons, D. Welsh, J. Tsevat, and M. L. Render. "Impact of Different Measures of Comorbid Disease on Predicted Mortality of Intensive Care Unit Patients." *Medical Care* 40, no. 10 (2002): 929–40.

[20] Rosenthal, G. E., P. J. Kaboli, and M. J. Barnett. "Differences in Length of Stay in Veterans Health Administration and Other United States Hospitals: Is the Gap Closing?" *Medical Care* 41, no. 8 (2003): 882–94.

[21] Silber, J. H., P. R. Rosenbaum, M. E. Trudeau, W. Chen, X. Zhang, R. R. Kelz, R. E. Mosher, and O. Even-Shoshan. "Changes in Prognosis after the First Postoperative Complication." *Medical Care* 45, no. 2 (2005): 122–31.

[22] Silber, J. H., P. R. Rosenbaum, M. E. Trudeau, W. Chen, X. Zhang, S. A. Lorch, R. R. Kelz, R. E. Mosher, and O. Even-Shoshan. "Preoperative Antibiotics and Mortality in the Elderly." *Annals of Surgery* 242, no. 1 (2005): 107–14.

[23] Southern, D. A., H. Quan, and W. A. Ghali. "Comparison of the Elixhauser and Charlson/Deyo Methods of Comorbidity Measurement in Administrative Data." *Medical Care* 42, no. 4 (2004): 355–60.

[24] Iezzoni, L. I. "Assessing Quality Using Administrative Data." *Medical Care* 127, no. 8, pt. 2 (1997): 666–74.

[25] Goff, D. C., Jr., D. K. Pandey, F. A. Chan, C. Ortiz, and M. Z. Nichaman. "Congestive Heart Failure in the United States: Is There More Than Meets the I(CD Code)? The Corpus Christi Heart Project." *Archives of Internal Medicine* 160, no. 2 (2000): 197–202.

[26] Ho, K. K., K. M. Anderson, W. B. Kannel, W. Grossman, and D. Levy. "Survival after the Onset of Congestive Heart Failure in Framingham Heart Study Subjects." *Circulation* 88, no. 1 (1993): 107–15.

[27] Kannel, W. B., and A. J. Belanger. "Epidemiology of Heart Failure." *American Heart Journal* 121, no. 3, pt. 1 (1991): 951–57.

[28] Webster, G., J. Zhang, and D. Rosenthal. "Comparison of the Epidemiology and Co-morbidities of Heart Failure in the Pediatric and Adult Populations: A Retrospective, Cross-Sectional Study." *BMC Cardiovascular Disorders* 6, no. 23 (2006).

[29] Johnston, J., D. P. Wagner, S. Timmons, D. Welsh, J. Tsevat, and M. L. Render. "Impact of Different Measures of Comorbid Disease on Predicted Mortality of Intensive Care Unit Patients."

[30] Ashton, C. M., N. J. Petersen, J. Souchek, T. J. Menke, K. Pietz, H. J. Yu, and N. P. Wray. "Rates of Health Services Utilization and Survival in Patients with Heart Failure in the Department of Veterans Affairs Medical Care System." *American Journal of Medical Quality* 14, no. 1 (1999): 55–63.

[31] Maciejewski, M. L., C. F. Liu, A. Derleth, M. McDonell, S. Anderson, and S. D. Fihn. "The Performance of Administrative and Self-Reported Measures for Risk Adjustment of Veterans Affairs Expenditures."

[32] Kaafarani, H. M., K. M. Itani, L. A. Petersen, J. Thornby, and D. H. Berger. "Does Resident Hours Reduction Have an Impact on Surgical Outcomes?"

[33] Rosenthal, G. E., P. J. Kaboli, and M. Barnett. "Differences in Length of Stay in Veterans Health Administration and Other United States Hospitals: Is the Gap Closing?"

[34] Silber, J. H., P. R. Rosenbaum, M. E. Trudeau, W. Chen, X. Zhang, R. R. Kelz, R. A. Mosher, and O. Even-Shoshan. "Changes in Prognosis after the First Postoperative Complication."

[35] Silber, J. H., P. R. Rosenbaum, M. E. Trudeau, W. Chen, X. Zhang, S. A. Lorch, R. R. Kelz, R. E. Mosher, and O. Even-Shoshan. "Preoperative Antibiotics and Mortality in the Elderly."

[36] Glickman, S.W. and E. D. Peterson, Innovative Health Reform Models: Pay-for-Performance Initiatives, *American Journal of Managed Care*. Vol. 15, No. 10 (2009), S300-S305.

[37] Centers for Medicare and Medicaid Services. 2006. Available at http://www.cms.hhs.gov/PVRP/.

[38] Iezzoni, L. I. "Assessing Quality Using Administrative Data."

[39] Moriomoto, T., T. K. Ghandi, A. C. Seger, T. C. Hsieh, and D. W. Bates. "Adverse Drug Events and Medication Errors: Detection and Classification Methods."

[40] Jollis, J. G., M. Ancukiewicz, E. R. Delong, D. B. Pryor, L. H. Muhlbaier, and D. B. Mark. "Discordance of Databases Designed for Claims Payment Versus Clinical Information Systems: Implications for Outcomes Research."

[41] Houchens, R. L., A. Elixhauser, and P. S. Romano. "How Often Are Potential Patient Safety Events Present on Admission?"

[42] Naessens, J. M., C. R. Campbell, B. Berg, A. R. Williams, and R. Culbertson. "Impact of Diagnosis-Timing Indicators on Measures of Safety, Comorbidity, and Case Mix Groupings from Administrative Data Sources."

[43] Polancich, S., E. Restrepo, and J. Prosser. "Cautious Use of Adminstrative Data for Decubitus Ulcer Outcome Reporting."

[44] Bahl, V., M. A. Thompson, T. Y. Kau, H. M. Hu, and D. A. Campbell Jr. "Do the AHRQ Patient Safety Indicators Flag Conditions That Are Present at the Time of Hospital Admission?"

[45] Pine, M., B. Jones, and Y. B. Lou. "Laboratory Values Improve Predictions of Hospital Mortality."

[46] Borzecki, A. M., A. T. Wong, E. C. Hickey, A. S. Ash, and D. R. Berlowitz. "Identifying Hypertension-Related Comorbidities from Administrative Data: What's the Optimal Approach?"

[47] Hsia, D. C., W. M. Krushat, A. B. Fagan, J. A. Tebbutt, and R. P. Kusserow. "Accuracy of Diagnostic Coding for Medicare Patients Under the Prospective-Payment System." *New England Journal of Medicine* 318, no. 6 (1988): 352–55.

[48] O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy." *Health Services Research* 40, no. 5, pt. 2 (2005): 1620–39.

[49] Zhan, C., A. Elixhauser, C. L. Richards, Jr., Y. Wang, W. B. Baine, M. Pineau, N. Verzier, R. Kliman, and D. Hunt. "Identification of Hospital-acquired Catheter-associated Urinary Tract Infections from Medicare Claims: Sensitivity and Positive Predictive Value." *Medical Care* 47, no. 3 (2009): 364–69.

[50] Pine, M., H. S. Jordan, A. Elixhauser, D. E. Fry, D. C. Hoaglin, B. Jones, R. Meimban, D. Warner, and J. Gonzales. "Modifying ICD-9-CM Coding of Secondary Diagnoses to Improve Risk-Adjustment of Inpatient Mortality Rates." *Medical Decision Making* 29, no. 1 (2009): 69–81.

[51] Glance, L. G., T. M. Osler, D. B. Mukamel, and A. W. Dick. "Impact of the Present-on-Admission Indicator on Hospital Quality Measurement: Experience with the Agency for Healthcare Research and Quality (AHRQ) Inpatient Quality Indicators." *Medical Care* 46, no. 2 (2008): 112–19.

[52] O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."

[53] Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. "Comorbidity Measures for Use with Administrative Data." *Medical Care* 36, no. 1 (1998): 8–27.

[54] Ibid.

[55] Shin, J. H., M. Worni, A. W. Castleberry, R. Pietrobon, P. A. Omotosho, M. Silberberg, and T. Ostbye. "The Application of Comorbidity Indices to Predict Early Postoperative Outcomes After Laparoscopic Roux-en-Y Gastric Bypass: A Nationwide Comparative Analysis of Over 70,000 Cases." *Obesity Surgery* 23, no. 5 (2013): 638–49.

[56] Boorjian, S. A., S. P. Kim, M. K. Tollefson, A. Carrasco, J. C. Cheville, R. H. Thompson, P. Thapa, and I. Frank. "Comparative Performance of Comorbidity Indices for Estimating

Perioperative and 5-Year All-Cause Mortality Following Radical Cystectomy for Bladder Cancer." *Journal of Urology* 190, no. 1 (2013): 55–60.

[57] Austin, P. C., M. B. Stanbrook, G. M. Anderson, A. Newman, and A. S. Gershon. "Comparative Ability of Comorbidity Classification Methods for Administrative Data to Predict Outcomes in Patients with Chronic Obstructive Pulmonary Disease." *Annals of Epidemiology* 22, no. 12 (2012): 881–87.

[58] Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKenzie. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation."

Journal of Chronic Diseases 40, no. 5 (1987): 373–83.

[59] Deyo, R. A., D. C. Cherkin, and M. A. Ciol. "Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Databases." *Journal of Clinical Epidemiology* 45, no. 6 (1992): 613–19.

[60] Johns Hopkins University. "Johns Hopkins ACG Case-Mix Adjustment System." 2011. Available at http://www.acg.jhsph.edu.

[61] O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."

[62] Osborn, C. E. *Statistical Application for Health Information Management*. Gaithersburg, MD: Aspen, 2000.

[63] Borzecki, A. M., A. T. Wong, E. C. Hickey, A. S. Ash, and D. R. Berlowitz. "Identifying Hypertension-Related Comorbidities from Administrative Data: What's the Optimal Approach?"

[64] Lee, D. S., L. Donovan, P. C. Austin, Y. Gong, P. P. Liu, J. L. Rouleau, and J. V. Tu. "Comparison of Coding of Heart Failure and Comorbidities in Administrative and Clinical Data for Use in Outcomes Research." *Medical Care* 43, no. 2 (2005): 182–88.

[65] Jollis, J. G., M. Ancukiewicz, E. R. Delong, D. B. Pryor, L. H. Muhlbaier, and D. B. Mark. "Discordance of Databases Designed for Claims Payment Versus Clinical Information Systems: Implications for Outcomes Research."

[66] Frogner, B. K., G. F. Anderson, R. A. Cohen, and C. Abrams. "Incorporating New Research into Medicare Risk Adjustment." *Medical Care* 49, no. 3 (2011): 295–300.

[67] Myers, R. P., J. N. Hubbard, A. A. Shaheen, E. Dixon, and G. G. Kaplan. "Hospital Performance Reports based on Severity Adjusted Mortality Rates in Patients with Cirrhosis Depend on the Method of Risk Adjustment." *Annals of Hepatology* 11, no. 4 (2012): 526–35.

[68] Austin, P. C., M. B. Stanbrook, G. M. Anderson, A. Newman, and A. S. Gershon. Comparative Ability of Comorbidity Classification Methods for Administrative Data to Predict Outcomes in Patients with Chronic Obstructive Pulmonary Disease."

[69] Frogner, B. K., G. F. Anderson, R. A. Cohen, and C. Abrams. "Incorporating New Research into Medicare Risk Adjustment."

[70] White, S. V. *Q Solutions: Essential Resources for the Healthcare Quality Professional, Quality and Performance Improvement*. 3rd ed. Glenview, IL: National Association for Healthcare Quality, 2012.

[71] Gibbs, G. *Analyzing Qualitative Data: The Sage Qualitative Research Kit*. London: Sage, 2007.

[72] Church, S., and M. Ekberg. "Student Midwives' Responses to Reproductive Ethics: A Qualitative Focus Group Approach Using Case Scenarios." *Midwifery* 29, no. 8 (2013) 895–901.

[73] Dessie, Y., and M. Deresa. "Sexual Practices of HIV-Positive Individuals Attending Antiretroviral Treatment (ART) in Addis Ababa Public Hospitals: Findings from In-Depth Interview." *Pan African Medical Journal* 13, no. 80 (2012.

[74] Hinojosa-Lindsey, M., J. Arney, S. Heberlig, J. R. Kramer, R. L. Street Jr., H. B. El-Serag, and A. D. Naik. "Patients' Intuitive Judgments about Surveillance Endoscopy in Barrett's Esophagus: A Review and Application to Models of Decision-Making." *Diseases of the Esophagus*. Published online February 5, 2013. Available at doi:10.1111/dote.12028.

[75] Brown, F. *ICD-9-CM Coding Handbook 2007, With Answers*. Chicago, IL: American Hospital Association, 2007.

[76] Veterans Health Administration. Coding Handbook. Washington, DC, 2011.

[77] Jollis, J. G., M. Ancukiewicz, E. R. Delong, D. B. Pryor, L. H. Muhlbaier, and D. B. Mark. "Discordance of Databases Designed for Claims Payment Versus Clinical Information Systems: Implications for Outcomes Research."

[78] The R Foundation for Statistical Computing. Available at http://www.r-project.org/foundation/ (accessed January 7, 2013).

[79] Stata version 12. College Station, TX: StataCorp.

[80] Guest, Greg, Kathleen M. MacQueen, and Emily E. Namey. *Applied Thematic Analysis*. Thousand Oaks, CA: Sage, 2012.

[81] Braun, V., and V. Clarke. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2006): 77–101.

[82] Landis, J. R., and C. G. Koch. "The Measurement of Observer Agreement for Categorical Data. Biometrics." *Biometrics* 33, no. 1 (1977): 159–74.

[83] Borzecki, A. M., A. T. Wong, E. C. Hickey, A. S. Ash, and D. R. Berlowitz. "Identifying Hypertension-Related Comorbidities from Administrative Data: What's the Optimal Approach?"

[84] Lee, D. S., L. Donovan, P. C. Austin, Y. Gong, P. P. Liu, J. L. Rouleau, and J. V. Tu. "Comparison of Coding of Heart Failure and Comorbidities in Administrative and Clinical Data for Use in Outcomes Research."

[85] Veterans Health Administration. Coding Handbook. Washington, DC, 2011.

[86] Veterans Health Administration. *Coding Handbook*.

[87] American Hospital Association (AHA). *Coding Clinic for ICD-9-CM*.

[88] Ibid.

[89] Veterans Health Administration. *Coding Handbook*.

[90] Ibid.

[91] Frogner, B. K., G. F. Anderson, R. A. Cohen, and C. Abrams. "Incorporating New Research into Medicare Risk Adjustment."

[92] LaTour, Kathleen M., and Shirley Eichenwald Maki. *Health Information Management: Concepts, Principles, and Practice.* 3rd ed. Chicago, IL: AHIMA, 2010.

[93] Pine, M., B. Jones, and Y. B. Lou. "Laboratory Values Improve Predictions of Hospital Mortality."

[94] Pine, M., H. S. Jordan, A. Elixhauser, D. E. Fry, D. C. Hoaglin, B. Jones, R. Meimban, D. Warner, and J. Gonzales. "Enhancement of Claims Data to Improve Risk Adjustment of Hospital Mortality." *Journal of the American Medical Association* 297, no. 1 (2007): 71–76.

[95] Pine, M., M. Sonneborn, J. Schindler, M. Stanek, J. L. Maeda, and C. Hanlon. "Harnessing the Power of Enhanced Data for Healthcare Quality Improvement: Lessons from a

Minnesota Hospital Association Pilot Project." *Journal of Healthcare Management* 57, no. 6 (2012): 406–20.

[96] Zhan, C., A. Elixhauser, C. L. Richards, Jr., Y. Wang, W. B. Baine, M. Pineau, N. Verzier, R. Kliman, and D. Hunt. "Identification of Hospital-acquired Catheter-associated Urinary Tract Infections from Medicare Claims: Sensitivity and Positive Predictive Value." *Medical Care* 47, no. 3 (2009): 364–69.

[97] Pine, M., H. S. Jordan, A. Elixhauser, D. E. Fry, D. C. Hoaglin, B. Jones, R. Meimban, D. Warner, and J. Gonzales. "Modifying ICD-9-CM Coding of Secondary Diagnoses to Improve Risk-Adjustment of Inpatient Mortality Rates." *Medical Decision Making* 29, no. 1 (2009): 69–81.

[98] Glance, L. G., T. M. Osler, D. B. Mukamel, and A. W. Dick. "Impact of the Present-on-Admission Indicator on Hospital Quality Measurement: Experience with the Agency for Healthcare Research and Quality (AHRQ) Inpatient Quality Indicators." *Medical Care* 46, no. 2 (2008): 112–19.

[99] O'Malley, K. J., K. F. Cook, M. D. Price, K. R. Wildes, J. F. Hurdle, and C. M. Ashton. "Measuring Diagnoses: ICD Code Accuracy."

[100] American Health Information Management Association. "Vision 2016: A Blueprint for Quality Education in Health Information Management." September 24, 2007.

Article citation:
Garvin, Jennifer Hornung; Redd, Andrew; Bolton, Dan; Graham, Pauline; Roche, Dominic; Groeneveld, Peter; Leecaster, Molly; Shen, Shuying; Weiner, Mark G. "Exploration of ICD-9-CM Coding of Chronic Disease within the Elixhauser Comorbidity Measure in Patients with Chronic Heart Failure" *Perspectives in Health Information Management* (Fall, October 2013).