

# Geographic Variation in Diagnosis Frequency and Risk of Death Among Medicare Beneficiaries

H. Gilbert Welch, MD, MPH

Sandra M. Sharp, SM

Dan J. Gottlieb, MS

Jonathan S. Skinner, PhD

John E. Wennberg, MD, MPH

**D**ISEASE DIAGNOSES ARE CONSIDERED a fundamental input for adjusting health outcomes as well as expenditures to credit systems that care for patients who are sicker than average. Ideally, a diagnosis would be solely an attribute of the patient, unaffected by the process of observation.

A recent investigation demonstrated an association between regional diagnostic practice and the number of diagnoses made in the small proportion of Medicare beneficiaries who changed residence during a 3-year period (approximately 5%).<sup>1</sup> Although patients who moved had a similar number of diagnoses prior to moving, those who moved from low- to high-practice intensity regions accumulated more than twice as many diagnoses than those moving in the opposite direction.

The evidence that diagnoses can be associated with system factors led us to explore how frequency of diagnosis might relate to outcomes in the fee-for-service Medicare population as a whole. Specifically, we hypothesized that patients diagnosed with chronic conditions in regions with high diagnosis frequency would have lower all-cause case-fatality rates than those diag-

**Context** Because diagnosis is typically thought of as purely a patient attribute, it is considered a critical factor in risk-adjustment policies designed to reward efficient and high-quality care.

**Objective** To determine the association between frequency of diagnoses for chronic conditions in geographic areas and case-fatality rate among Medicare beneficiaries.

**Design, Setting, and Participants** Cross-sectional analysis of the mean number of 9 serious chronic conditions (cancer, chronic obstructive pulmonary disease, coronary artery disease, congestive heart failure, peripheral artery disease, severe liver disease, diabetes with end-organ disease, chronic renal failure, and dementia) diagnosed in 306 hospital referral regions (HRRs) in the United States; HRRs were divided into quintiles of diagnosis frequency. Participants were 5 153 877 fee-for-service Medicare beneficiaries in 2007.

**Main Outcome Measures** Age/sex/race-adjusted case-fatality rates.

**Results** Diagnosis frequency ranged across HRRs from 0.58 chronic conditions in Grand Junction, Colorado, to 1.23 in Miami, Florida (mean, 0.90 [95% confidence interval {CI}, 0.89-0.91]; median, 0.87 [interquartile range, 0.80-0.96]). The number of conditions diagnosed was related to risk of death: among patients diagnosed with 0, 1, 2, and 3 conditions the case-fatality rate was 16, 45, 93, and 154 per 1000, respectively. As regional diagnosis frequency increased, however, the case fatality associated with a chronic condition became progressively less. Among patients diagnosed with 1 condition, the case-fatality rate decreased in a stepwise fashion across quintiles of diagnosis frequency, from 51 per 1000 in the lowest quintile to 38 per 1000 in the highest quintile (relative rate, 0.74 [95% CI, 0.72-0.76]). For patients diagnosed with 3 conditions, the corresponding case-fatality rates were 168 and 137 per 1000 (relative rate, 0.81 [95% CI, 0.79-0.84]).

**Conclusion** Among fee-for-service Medicare beneficiaries, there is an inverse relationship between the regional frequency of diagnoses and the case-fatality rate for chronic conditions.

*JAMA. 2011;305(11):1113-1118*

[www.jama.com](http://www.jama.com)

nosed in regions with low diagnosis frequency.

## METHODS

### Overview

To investigate the geographic variation in diagnosis frequency, we measured the frequency with which 9 serious chronic conditions were diag-

**Author Affiliations:** Department of Veterans Affairs Medical Center, White River Junction, Vermont (Dr Welch); Department of Economics, Dartmouth College, Hanover, New Hampshire (Dr Skinner); and Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Lebanon, New Hampshire (Drs Welch, Skinner, and Wennberg and Ms Sharp and Mr Gottlieb).

**Corresponding Author:** H. Gilbert Welch, MD, MPH, Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, 35 Centerra Pkwy, Ste 202, Lebanon, New Hampshire 03766 (h.gilbert.welch@dartmouth.edu).

nosed among the fee-for-service Medicare population in each of 306 hospital referral regions (HRRs) in the United States. We categorized the HRRs into quintiles of diagnosis frequency and measured physician encounters and diagnostic testing. We examined death rates across the quintiles, first for the population as a whole and then within strata of a specified number of chronic conditions (ie, 0, 1, 2, 3).

### Data

We analyzed a 20% sample of Medicare beneficiaries in 2007. We restricted the analysis to those beneficiaries who were fully enrolled in Part A and Part B throughout 2007 and who were 65 through 99 years old on December 31, 2007, or who were fully enrolled beginning January 1, 2007, until their death that year and who were 65 through 99 years old at their time of death. Because their claims data are incomplete, beneficiaries enrolled in risk-contract health maintenance organizations were excluded. The final sample totaled 5 153 877 beneficiaries.

### Diagnosis Frequency

We counted the number of chronic conditions documented in the claims for each beneficiary during 2007. We used the 9 major chronic conditions based on the work of Iezzoni et al,<sup>2</sup> as adapted for the 2008 Dartmouth Atlas of Health Care. The conditions were cancer with poor prognosis, chronic obstructive pulmonary disease, coronary artery disease, congestive heart failure, peripheral artery disease, severe liver disease, diabetes with end-organ disease, chronic renal failure, and dementia.

For a beneficiary to be counted as having a chronic condition, the diagnosis had to be either coded on at least 1 hospital discharge abstract following an inpatient stay or on at least 2 claims involving physician contact that were at least 7 days apart. The latter requirement was used to reduce the likelihood of erroneously including “rule out” diagnoses. A beneficiary was

counted as either having or not having each of the 9 conditions, and the total number of conditions for each beneficiary was calculated (range, 0-9).

We calculated the mean number of chronic conditions per beneficiary within each of the 306 HRRs to examine the variation in diagnosis frequency across the Medicare population as a whole. The HRRs were sorted in terms of increasing diagnosis frequency and grouped into quintiles based on population counts (ie, approximately 1 million beneficiaries per quintile). Within each quintile, we examined the distribution of the number of chronic conditions diagnosed. To assess system factors that may be related to the likelihood of diagnosis, we examined 4 measures reflecting physician encounters and diagnostic testing: the number of physician visits, the number of different physicians seen, the number of imaging tests obtained, and the number of laboratory tests obtained.

### Rate of Death

The primary outcome was rate of death in specific groups, measured as either mortality (the population-based rate of death in which the denominator is the entire population in the quintile) or case fatality (the case-based rate of death in which the denominator is the portion of the population in the quintile that has received the diagnosis in question).

The denominator for the population-based mortality rate was the fully entitled Medicare population in the quintile obtained from the Medicare denominator file. The denominator for case-fatality rate was the number of individuals who had the diagnosis in question within that population. For both measures, the numerator was the number of deaths from any cause in calendar year 2007 among those appearing in the denominator. The fact of death was obtained from the Medicare administrative files. To control for population differences across quintiles, all death rates were adjusted for age, sex, and race using the indirect method (which used 20 age/sex/race

cells and obtained standardized rates from the group as a whole).

### Analytic Approach

To consider the relationship of diagnosis frequency and rate of death, we compared age/sex/race-adjusted population-based mortality across the 5 diagnosis frequency quintiles. We repeated the comparison using age/sex/race-adjusted case-fatality in analyses restricted to individuals diagnosed with a specified number of chronic conditions (ie, 0, 1, 2 and 3).

To examine the relationship in specific diagnoses, we compared case fatality across quintiles for each of the 9 chronic conditions. We used logistic regression to adjust for age, sex, and race (using 20 indicator variables for each of the aforementioned age/sex/race cells) and the number of other conditions diagnosed (also using indicator variables to allow for a nonlinear relationship). We estimated 9 different models—one for each condition—with a different (but overlapping) sample for each target condition (eg, only patients with congestive heart failure are included in the congestive heart failure model). Each was evaluated for face validity. Death rates increased with age and the number of conditions and were higher in men and blacks than in women and nonblacks. Each model estimated the expected number of deaths among beneficiaries with a given condition. We constructed the observed to expected ratio (O:E) for each quintile. We used the ratio of O:E in the very high quintile to the O:E in the very low quintile to summarize the association between diagnosis frequency and case fatality.

Pearson correlation coefficients ( $r$ ) were calculated to describe the relationship at the HRR level ( $n=306$ ) between diagnosis frequency and the 4 measures reflecting physician encounters and diagnostic testing. Confidence intervals (95% CIs) around the Pearson  $r$  values were calculated using the Fisher  $r$ -to- $z$  transformation; 95% CIs at the population level (eg, quintile, patients with 1 chronic condi-

tion) were calculated empirically using the bootstrap method. Linear regression was performed to test for trends, using 2-sided *P* values and with a 5% significance level. All statistical analyses were performed using SAS version 9.2 (SAS Institute Inc, Cary, North Carolina) or STATA version 11 (Stata-Corp, College Station, Texas). The study was approved by the Dartmouth College institutional review board.

## RESULTS

### Variation in Diagnosis Frequency

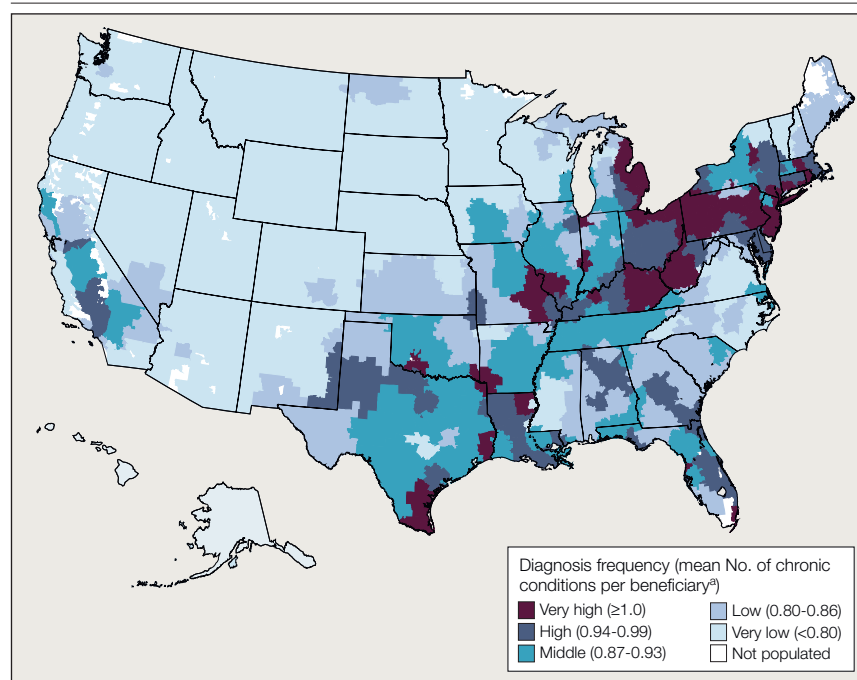
In 2007, the mean number of chronic conditions diagnosed among Medicare beneficiaries across 306 HRRs was 0.90 (95% confidence interval [CI], 0.89-0.91; median, 0.87 [interquartile range, 0.80-0.96]). As shown in FIGURE 1, the frequency of diagnosis varied substantially with geography. The mean number of chronic conditions diagnosed per Medicare beneficiary ranged from 0.58 in Grand Junction, Colorado, and Idaho Falls, Idaho, to 1.23 in Miami, Florida, and McAllen, Texas.

The distribution of the number of chronic conditions varied accordingly. FIGURE 2 depicts this distribution for each of the 5 quintiles of diagnosis frequency. Beneficiaries in the highest quintile were almost twice as likely to have had a diagnosis of 3 or more chronic conditions than those in the lowest quintile (14.3% vs 7.7%; relative rate, 1.86 [95% confidence interval {CI}, 1.84-1.88]). Conversely, beneficiaries in the lowest quintile of diagnosis frequency were more likely to have no chronic condition diagnosis at all (57.3% vs 44.4%; relative rate, 1.29 [95% CI, 1.28-1.30]).

### Diagnosis Frequency and Measures of Physician Encounters and Diagnostic Testing

The TABLE shows the relationship between diagnosis frequency and measures of physician encounters and diagnostic testing. Across the 306 HRRs, diagnosis frequency had a strong positive correlation with each of the 4 measures: the number of physician visits

**Figure 1.** Variation in Number of Chronic Conditions Diagnosed Among Medicare Beneficiaries in the United States, 2007



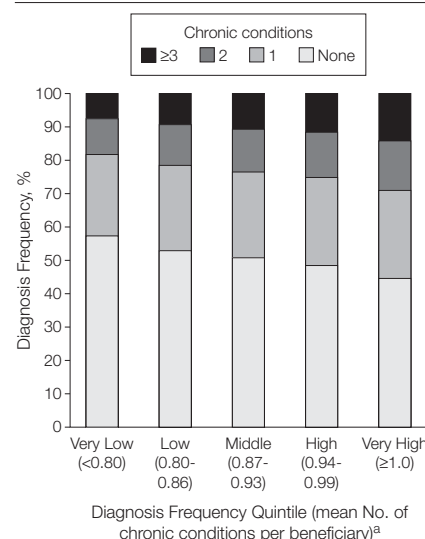
<sup>a</sup>Values in parentheses indicate range for each quintile.

( $r=0.72$ ), number of different physicians seen ( $r=0.59$ ), number of imaging tests ( $r=0.72$ ), and number of laboratory tests ( $r=0.33$ ). The TABLE also shows that the mean number of physician visits increased in a stepwise fashion in progressively higher quintiles of diagnosis frequency ( $P=.002$  for trend). A similar pattern was evident for the 3 other measures of physician encounters and diagnostic testing.

### Diagnosis Frequency and Risk of Death

The relationship between diagnosis frequency and population-based mortality is illustrated in the left-hand portion of FIGURE 3. When considering the entire population within each quintile, there was little relationship between diagnosis frequency and age/sex/race-adjusted all-cause mortality: 49 per 1000 in lowest quintile of diagnostic frequency; 52, 55, and 52 per 1000 in the intervening quintiles; and 53 per 1000 in the highest quintile ( $P=.31$  for trend).

**Figure 2.** Distribution of Chronic Conditions Among Medicare Beneficiaries Across Regions With Varying Diagnosis Frequency



See Table 1 for denominator data used to calculate diagnosis frequency.

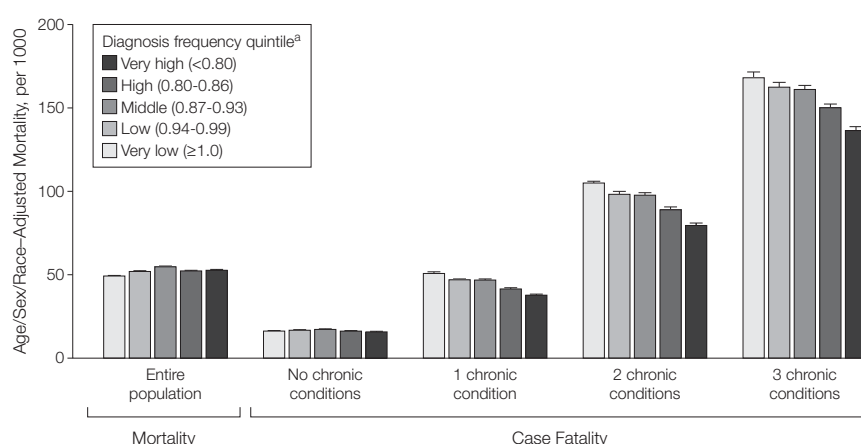
<sup>a</sup>Values in parentheses indicate range for each quintile.

The relationship between diagnosis frequency and case-fatality is illustrated in the right-hand portion of

**Table.** Relationship Between Diagnosis Frequency and Measures of Physician Encounters and Diagnostic Testing in the United States, 2007

	Measure of Physician Encounters and Diagnostic Testing, per Beneficiary			
	Physician Visits, No.	Different Physicians Seen, No.	Imaging Studies, No.	Laboratory Tests, No.
Correlation with diagnosis frequency, <i>r</i> (95% CI) <sup>a</sup>	0.72 (0.66-0.77)	0.59 (0.51-0.66)	0.72 (0.66-0.77)	0.33 (0.23-0.43)
Diagnosis frequency, quintile mean (95% CI)				
Very low (n = 1 019 557)	11.35 (11.32-11.38)	3.80 (3.79-3.81)	3.66 (3.65-3.68)	10.15 (10.11-10.18)
Low (n = 1 038 091)	12.94 (12.90-12.97)	4.26 (4.25-4.26)	4.16 (4.15-4.17)	11.90 (11.87-11.93)
Middle (n = 1 019 701)	13.20 (13.17-13.24)	4.21 (4.21-4.22)	4.36 (4.35-4.37)	11.96 (11.92-11.99)
High (n = 1 039 413)	14.90 (14.85-14.93)	4.68 (4.67-4.69)	4.53 (4.52-4.55)	12.45 (12.41-12.49)
Very high (n = 1 037 115)	16.19 (16.14-16.24)	4.95 (4.94-4.96)	4.69 (4.68-4.71)	13.93 (13.89-13.97)
P value for trend	.002	.008	.008	.01

Abbreviations: CI, confidence interval; HRR, hospital referral region.

<sup>a</sup>Across 306 hospital referral regions.**Figure 3.** Population-Based Mortality and Case Fatality Across Regions With Varying Diagnosis Frequency

Error bars indicate 95% confidence intervals. Relative rates of death in regions with very high vs very low diagnosis frequency: relative rate, 1.07 (95% CI, 1.06-1.08 [ $P = .31$  for trend]) for entire population; 0.97 (95% CI, 0.94-1.00 [ $P = .46$  for trend]) for no chronic conditions; 0.74 (95% CI, 0.72-0.76 [ $P = .006$  for trend]) for 1 chronic condition; 0.76 (95% CI, 0.74-0.78 [ $P = .007$  for trend]) for 2 chronic conditions; 0.81 (95% CI, 0.79-0.84 [ $P = .01$  for trend]) for 3 chronic conditions. Entire population,  $N = 5\,153\,877$ ; no chronic conditions,  $n = 2\,611\,646$ ; 1 chronic condition,  $n = 1\,318\,809$ ; 2 chronic conditions,  $n = 665\,093$ ; 3 chronic conditions,  $n = 331\,999$ ; not shown are 226 300 persons with 4 or more conditions.

<sup>a</sup>Quintiles indicate mean number of chronic conditions per beneficiary. Values in parentheses indicate range for each quintile.

Figure 3. In all regions, case-fatality increased as the number of chronic conditions diagnosed increased. Among Medicare beneficiaries diagnosed with 0, 1, 2, and 3 conditions in the United States as a whole, the case-fatality rate was 16, 45, 93, and 154 per 1000, respectively (data not shown).

When focusing on the subsets of Medicare beneficiaries with a fixed number of chronic conditions, increased diagnosis frequency was associated with declining case fatality.

Among patients diagnosed with 1 condition, case fatality was 51 per 1000 for HRRs in the lowest quintile of diagnosis frequency; 47, 47, and 41 per 1000 in the intervening quintiles; and 38 per 1000 in the highest quintile ( $P = .006$  for trend). Among these patients, the relative rate of death in the highest to lowest diagnosis frequency quintile was 0.74 (95% CI, 0.72-0.76).

A similar stepwise relationship was observed across the quintiles among the subsets diagnosed with 2 and 3

chronic conditions ( $P = .007$  and  $P = .01$  for trend, respectively). For both subsets the relative rate of death in the highest vs lowest diagnosis frequency quintile was significantly less than 1: among patients with 2 chronic conditions, the relative rate was 0.76 (95% CI, 0.74-0.78); among those with 3 chronic conditions, the relative rate was 0.81 (95% CI, 0.79-0.84).

FIGURE 4 shows the relative rate of death in the highest to lowest diagnosis frequency quintile among beneficiaries diagnosed with each of the 9 chronic conditions. For each condition, the relative rate was significantly less than 1. In 7 of 9 conditions, a significant stepwise relationship was observed across quintiles ( $P < .05$  for trend for all).

## COMMENT

As the number of diagnoses of chronic conditions increased in individual patients, there was an associated increased risk of death. However, as the number of diagnoses increased among geographically defined populations (ie, across quintiles of diagnosis frequency), there was little relationship with population-based mortality. These apparently paradoxical findings were explained by a third observation: among patient subgroups with a given number of chronic conditions, there was a consistent stepwise decrement in case fatality as diagnosis frequency increased.



As with all observational data, causality cannot be directly tested. Our analysis could be limited by residual confounding—ie, from variables that could explain increased diagnosis frequency as well as declining case fatality. The conventional explanation for our findings would be that the geographic variation in diagnosis frequency reflects underlying differences in disease burden, ie, regions with high diagnosis frequencies must have sicker patients. However, this explanation fails to explain why population-based mortality is stable across quintiles of diagnosis frequency.

Another explanation would suggest that regions with high diagnosis frequencies are more effective in treating sick patients, thereby reducing mortality rates that otherwise would have been substantially increased. While the population of patients in the lowest quintile of diagnosis frequency may have fewer chronic conditions, those who do have fewer also may have poor access to health services (eg, fewer specialists, longer wait times), leading to increased case fatality. Conversely, the population in the highest quintile of diagnosis frequency may have more chronic conditions yet better access to services (and perhaps more experienced physicians), leading to decreased case fatality. However, to pro-

duce the observed pattern of reductions in case fatality in a stepwise relationship, another condition would have to be met: access and the ability to provide effective care must be directly related to diagnosis frequency. Given the required conditions, the conventional explanation is not particularly parsimonious.

An alternative explanation may be that geographic variation in diagnosis frequency substantially reflects the intensity of observation. This is consistent with the associations we report between diagnosis frequency and measures of physician encounters and diagnostic testing. This explanation provides a more parsimonious explanation for our case-fatality findings; ie, if diagnosis frequency reflects intensity of observation, then the pattern of case fatality we observed here would be expected. More testing and more opportunities to make diagnoses may translate into the typical patient given a diagnosis being less sick. The finding of Song et al<sup>1</sup>—that the number of diagnoses accumulated by migrating Medicare beneficiaries is associated with the location to which they moved—provided evidence from a natural experiment that supports this hypothesis.

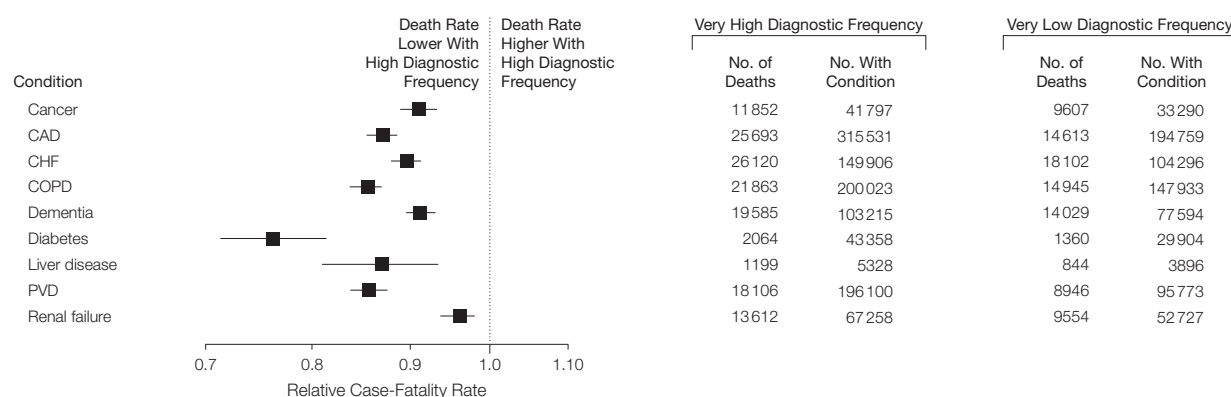
Our analysis has a number of limitations. First, although Medicare claims are

the most complete population-based data available in the United States, they are not entirely complete. Specifically, beneficiaries enrolled in plans outside of fee-for-service (plans that receive capitated payments from Medicare, such as Medicare Advantage) are not included in claims data. Although the proportion of beneficiaries enrolled in these plans is not correlated with diagnosis frequency across HRRs (in fact, enrollment is most common in the very low and very high quintiles), the possibility of differential selection raises the question of whether the relationship we observed would be present in the entire Medicare population. This highlights the importance of developing mechanisms to capture data from capitated plans<sup>3</sup> to foster population-based analyses.

Second, the ability of the logistic regression models to adequately isolate the effect of each individual condition from others may be limited. To address this concern, we repeated the analysis on the subset of patients with no other confounding diagnoses: those diagnosed with only 1 chronic condition. This analysis also showed lower case fatality in the very high quintile of diagnosis frequency for each of the 9 chronic conditions.

Third, the accuracy of coding diagnostic data are open to question. If in-

**Figure 4.** Relative Rates of Death in Regions With Very High vs Very Low Diagnosis Frequency Among Beneficiaries Diagnosed With Each of 9 Chronic Conditions



Relative rates were calculated using logistic regression adjusted for age/sex/race and number of coexisting chronic conditions and are not calculable from raw counts of numbers of deaths and numbers with each condition. CAD indicates coronary artery disease; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; PVD, peripheral vascular disease.

accuracies are random and not associated with region, they would not affect our findings. Coding practices, however, could vary across regions. It is possible that coding is relatively incomplete in the lowest quintile and relatively complete in the highest quintile. Alternatively, these differences could be purposeful—that is, the result of “gaming” efforts to increase reimbursements (or improve apparent quality) in the highest quintile.<sup>4,5</sup> Although we have no evidence that this is the case, such practices could explain our results.

The frequency of diagnoses reported in claims data are routinely used in methods for risk adjustment in comparative effectiveness research,<sup>6,7</sup> the evaluation of readmissions following hospitalization,<sup>8,9</sup> and in paying insurance plans under the Medicare Advan-

tage program.<sup>10</sup> If diagnosis is not solely an attribute of underlying disease burden, adjustments based on frequency of diagnosis may introduce bias into efforts to compare outcomes, pay for health care, and assess the extent of geographic variation in health care delivery. On the other hand, if more diagnoses (and more frequent encounters and diagnostic testing as well as greater spending) improve outcomes, then standard methods of risk adjustment may provide a more accurate comparison of effectiveness and efficiency. Future research must further evaluate the contribution of the process of observation to diagnosis frequency and explore mechanisms to better measure disease burden.

**Author Contributions:** Dr Welch had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Welch, Wennberg.  
**Acquisition of data:** Gottlieb.  
**Analysis and interpretation of data:** Welch, Sharp, Gottlieb, Skinner, Wennberg.  
**Drafting of the manuscript:** Welch, Sharp, Wennberg.  
**Critical revision of the manuscript for important intellectual content:** Gottlieb, Skinner, Wennberg.  
**Statistical analysis:** Welch, Sharp, Gottlieb, Skinner, Wennberg.  
**Obtained funding:** Skinner.  
**Administrative, technical, or material support:** Wennberg.  
**Study supervision:** Welch.

**Conflict of Interest Disclosures:** All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Wennberg reported serving as a paid consultant to the Foundation for Informed Medical Decision Making and Health Dialog and receiving royalties from Health Dialog. No other authors reported disclosures.

**Funding/Support:** This study was partially supported by grant PO1-AG19783 from the National Institute on Aging.

**Role of the Sponsor:** The National Institute on Aging had no role in the design and conduct of the study; the collection, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

**Disclaimer:** The views expressed herein do not necessarily represent the views of the Department of Veterans Affairs or the United States Government.

## REFERENCES

1. Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES. Regional variations in diagnostic practices. *N Engl J Med*. 2010;363(1):45-53.
2. Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J, Coffman GA. Chronic conditions and risk of in-hospital death. *Health Serv Res*. 1994;29(4):435-460.
3. Welch WP, Welch HG. Fee-for-data: a strategy to open the HMO black box. *Health Aff (Millwood)*. 1995;14(4):104-116.
4. Morreim EH. Gaming the system: dodging the rules, ruling the dodgers. *Arch Intern Med*. 1991;151(3):443-447.
5. Pine M, Jordan HS, Elixhauser A, et al. Modifying ICD-9-CM coding of secondary diagnoses to improve risk-adjustment of inpatient mortality rates. *Med Decis Making*. 2009;29(1):69-81.
6. Ong MK, Mangione CM, Romano PS, et al. Looking forward, looking back: assessing variations in hospital resource use and outcomes for elderly patients with heart failure. *Circ Cardiovasc Qual Outcomes*. 2009;2(6):548-557.
7. Medicare Payment Advisory Commission (MedPAC). *Regional Variation in Medicare Service Use*. Washington, DC: MedPAC; 2011.
8. Krumholz HM, Merrill AR, Schone EM, et al. Patterns of hospital performance in acute myocardial infarction and heart failure 30-day mortality and readmission. *Circ Cardiovasc Qual Outcomes*. 2009;2(5):407-413.
9. Ghaferi AA, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. *N Engl J Med*. 2009;361(14):1368-1375.
10. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev*. 2004;25(4):119-141.