Do Variations in Disease Prevalence Limit the Usefulness of Population-Based Hospitalization Rates for Studying Variations in Hospital Admissions?

Michael Shwartz, PhD, Erol A. Peköz, PhD,* Arlene S. Ash, PhD,† Michael A. Posner, MS,‡ Joseph D. Restuccia, DrPH,* and Lisa I. Iezzoni, MD§*

Background: Studies of geographic variation in hospitalizations commonly examine age- and gender-adjusted population-based hospitalization rates (ie, the numbers of persons hospitalized relative to what is expected given the age/gender distributions in the area population).

Objective: To determine whether areas identified as extreme using population-based hospitalization rates remain extreme when ranked by disease-based hospitalization rates (the numbers of persons hospitalized relative to what is expected given the amount of disease in the area).

Design: The authors examined 1997 Medicare data on both inpatient admissions and outpatient visits of patients 65 years and older in each of 71 small areas in Massachusetts for 15 medical conditions. For each area, the number of people having each condition was calculated as the sum of those hospitalized plus those treated as outpatients only. The authors used hierarchical Bayesian modeling to estimate area-specific population-based hospitalization rates, disease-based hospitalization rates (DHRs), and disease prevalence.

Main Outcome Measure: The extent to which the same areas were identified as extreme based on population-based hospitalization rates versus DHRs.

Results: Area-specific population-based hospitalization rates, DHRs, and disease prevalence varied substantially. Areas identified as extreme using population-based hospitalization rates often were not extreme when ranked by DHRs. For 11 of the 15 conditions, 5 or more of the 14 areas ranked in top and bottom deciles by population-based hospitalization rates were more likely than not (ie, with probability ≥ 0.50) to be at least 2 deciles less extreme when ranked by DHRs.

Conclusion: Differences in disease prevalence can limit the usefulness of population-based hospitalization rates for studying variations in hospital admissions.

Key Words: small-area variations, hospital utilization, disease prevalence

(*Med Care* 2005;43: 4–11)

Many studies have reported large differences in age- and gender-adjusted rates of hospitalization across small geographic areas.[1–13] Studies of surgery and procedures[14–19] and medical conditions[20] have found that higher rates of inappropriateness do not explain higher hospitalization rates. Despite this, some suggest that savings are possible without sacrificing quality of care by reducing rates in high-rate areas to levels in lower rate areas.[21,22] This suggestion assumes that much of the variation is due to "practice style" differences, either "uncertainty" about best treatment practices[23] or "enthusiasm" for one type of treatment over another,[24] which only minimally affect health care outcomes.

Area-specific hospitalization rates are population based because denominators used in calculating the rates reflect the populations in the areas rather than the numbers of people within areas who have the disease. Differences in age- and gender-adjusted population-based hospitalization rates might be due to differences in the likelihood that patients with the disease are admitted to the hospital. Alternatively, they may reflect differences in disease prevalence.

The most widely used approach to account for differences in disease burden across areas is to adjust area-specific rates further for differences in hospitalization rates for selected marker conditions chosen as proxies for the underlying burden of illness in the population (eg, hospitalizations for hip fracture, colon or lung cancer treated surgically, acute myocardial infarction, and stroke).[25] This approach hypothesizes that hospitalization rates for the marker conditions are highly correlated with the total amount of other diseases in an

From the *School of Management, Boston University; the †Department of General Internal Medicine, Boston University School of Medicine; the ‡Boston University School of Public Health; and the §Harvard Medical School, Boston, Massachusetts.

Supported in part by grant no. RO3 HS09832-01 from the Agency for Healthcare Research and Quality.

Reprints: Michael Shwartz, School of Management, Boston University, 595 Commonwealth Avenue, Boston, MA 02215. E-mail: mshwartz@bu.edu.

ISSN: 0025-7079/05/4301-0004

area. The hypothesis has some merit, as suggested by the finding that variation in such marker conditions explains about 25% of the variation in age-, gender-, and race-adjusted Medicare spending across regions.[22] However, use of marker conditions as a proxy for the prevalence of other diseases has not been validated. Also, when examining variations in hospitalization rates for specific conditions, the conceptual justification for using markers as a measure of prevalence is weaker.

Over the last decade, several groups have developed risk assessment models that use both inpatient and ambulatory claims to identify diseases and predict costs.[26] Their success indicates that although diagnosis codes on claims forms are imperfect indicators of true disease prevalence, they contain useful information about the medical problems that are present. Especially in a Medicare population, in which most beneficiaries are insured for both inpatient and outpatient care, differences across areas in amount of disease as determined from claims forms should reflect, reasonably well, differences in underlying disease prevalence.

In this study we used Medicare data from both inpatient and outpatient claims to identify people with any of 15 medical conditions. For each condition and geographic area, we calculated both the population-based hospitalization rate (the number of people hospitalized relative to the number expected given the age/gender distribution of the population in the area) and the disease-based hospitalization rate (the number of people hospitalized relative to the number expected given the amount of disease in the area). We then examined the extent to which the same areas were ranked as either particularly high or low using the 2 measures.

METHODS

Database

We studied hospitalizations and outpatient treatment of Medicare patients older than 65 years of age in Massachusetts in 1997. Hospitalization data were obtained from the Center for Medicare and Medicaid Services (CMS, formerly the Health Care Financing Administration) MedPAR file. Outpatient data were obtained from the CMS 1997 Carrier File (claims data for part B physician/supplier services) and Outpatient File (claims data for outpatient facility charges at hospitals and other institutions).

Conditions

Table 1 shows the 15 medical conditions in our study, defined initially by diagnosis-related group (DRG). As described elsewhere,[27] within most DRGs we increased clinical homogeneity by considering only discharges with a principal diagnosis from selected ICD-9-CM codes. All 15 conditions had at least 1000 people hospitalized and 6000 treated as outpatients only in 1997.

Determining Numerators and Denominators

Disease prevalence cannot be inferred from claims for Medicare beneficiaries in health maintenance organizations (HMOs; because their outpatient claims are not submitted to CMS) or those not eligible for outpatient and physician office (part B) reimbursement. Thus, we wanted to include in our analyses only non-HMO part B-eligible enrollees. We did not have individual-level eligibility information. As a proxy for part B eligibility of hospitalized patients, we only counted hospitalized patients with at least 1 part B bill in 1997

TABLE 1. Number of People Hospitalized and Number Treated on an Outpatient-Only Basis, by Condition

DRG	Hospitalized	Outpatient-Only
15 Transient ischemic attack	2001	11,762
88 Chronic bronchitis and emphysema	4842	51,982
89 Bacterial pneumonia	8666	7899
127 Heart failure	11,359	37,187
130 Peripheral vascular disorders	1334	40,936
132 Ischemic heart disease	3432	111,838
138 Cardiac arrhythmia and conduction disorder	3856	74,109
140 Angina pectoris	1596	25,597
141 Syncope and collapse	2046	6002
143 Chest pain	2859	26,282
243 Medical back problems	1453	59,754
277 Cellulitis and abscess	1765	16,339
294 Diabetes	1009	119,713
296 Fluid and electrolyte disorder	4351	6451
320 Kidney and urinary tract infections	3116	24,579

(slightly over 90% of hospitalized Medicare patients). We also eliminated from hospitalization counts the approximately 5% of discharges in the MedPAR file in which an HMO was the payer.

We identified outpatient visits, including office, nursing and rest home, and home visits from CPT codes.[27] Diagnosis coding for outpatient visits is generally less reliable than inpatient coding and is governed by somewhat different rules. To reflect the potential effect of different coding rules, we used 2 methods to assign outpatient visits to conditions: one based on a "series" of outpatient visits and a second based on a single visit. For inpatients, coding guidelines instruct abstractors to code a diagnosis accompanied by such phrases as "rule out," "suspect," or "question" as if the disease had actually occurred. For outpatients, coding guidelines stipulate that only confirmed diagnoses be coded to their highest level of specificity. As a result, a series of outpatient visits may carry multiple codes as the physician attempts to confirm a diagnosis. Because we wished to capture only "final" diagnoses, we proceeded as follows: We considered any outpatient visit within 6 weeks of a previous outpatient visit to be part of the same series of visits. Only diagnostic codes on the last visit in a series were used to identify the conditions being addressed. We required a gap of at least 8 weeks (ie, an additional 2-week buffer) to establish the beginning of a new series.

We examined the sensitivity of conclusions to an approach that ignored visit series and identified the conditions being addressed based on the presence of any relevant diagnosis at any single visit. Because findings concerning the concordance between population-based and disease-based rates were similar for both approaches, we only report analyses using the "visit series" method.

We considered the numbers of people with each condition, either treated in the hospital or as outpatient only, rather than the numbers of admissions or outpatient visits, primarily because the number of people who have a diagnosis is a more direct measure of underlying disease prevalence than service counts, which also reflect practice style.[28] Most variation in overall hospitalization rates is caused by variations in the number of people hospitalized.[27] By focusing on numbers of people rather than numbers of events, we can estimate the observed amount of disease in an area as the total of people hospitalized plus those treated as outpatients only.

For each 5-year age category from 65 years and older, and for each gender, we determined the number of Medicare enrollees in each zip code in Massachusetts from the Annual Zip Code Enrollment File.

Creating Small Geographic Areas

As described elsewhere,[27,29] we used Ward's clustering algorithm to create small geographic areas. Ward's clustering algorithm[30,31] creates areas by combining zip codes based on similarity in the proportion of total hospital discharges from the zip code that were from each hospital. Discharges of patients with the following characteristics were used in clustering: age 65 years or older, Massachusetts resident, and discharged in 1997 from a hospital in Massachusetts paid under Medicare's Prospective Payment System. The results of the clustering were 71 small areas with the following distribution of residents 65 years and older: 20 areas had less than 5000; 21 had 5000 to 9999; 19 had 10,000 to 19,999; and 11 areas had more than 20,000.

Analysis

We considered 3 types of area-specific "relative rates" (referred to more simply as "rates"), defined as observed counts divided by expected.

- Population-based hospitalization rate = number of people hospitalized relative to the number expected.
- Disease prevalence = sum of people hospitalized plus those treated as outpatients only (ie, the number with the disease) relative to the number expected (we use the term prevalence rather than rate to emphasize what this rate is measuring).
- Disease-based hospitalization rate = proportion of people with the disease that were hospitalized relative to the ratio of the expected number hospitalized to the expected number with the disease.

Let O_{ij} = number of people hospitalized (ie, treated as inpatients) in area j

O_{oj} = number of people treated as outpatients only in area j

E_{ij} = expected number of people hospitalized in area j

E_{oj} = expected number of people treated as outpatients only in area j

Both E_{ij} and E_{oj} are adjusted for age and gender distribution in the area using indirect standardization.[32] Observed relative rates are calculated as follows:

Observed population-based hospitalization rate = O_{ij}/E_{ij}

Observed disease rate = $(O_{ij} + O_{oj})/(E_{ij} + E_{oj})$

Observed disease-based hospitalization rate =

$$\frac{O_{ij}/(O_{ij} + O_{oj})}{E_{ij}/(E_{ij} + E_{oj})} = \frac{O_{ij}/E_{ij}}{(O_{ij} + O_{oj})/(E_{ij} + E_{oj})}$$

The rates are centered at 1 because over all areas the observed number of events is equal to the number expected. Because population-based hospitalization rate equals disease-based hospitalization rate multiplied by disease prevalence, population-based hospitalization rates are similar to disease-based hospitalization rates only when disease prevalence varies little across the area from what is expected based on age and gender alone.

We ranked areas from lowest to highest according to their population-based hospitalization rates. Areas in the first decile (ranks 1–7) had the lowest rates and those in the last

decile (ranks 65–71) had the highest. We call areas in these deciles extreme. For each extreme area, we examined its ranking according to its disease-based hospitalization rate. We present detailed results for heart failure (DRG 127), which has over 10,000 people treated as inpatients and nearly 40,000 treated as outpatients only.

Using the observed rates as calculated in the previous equations as estimates of "true" underlying rates does not explicitly take into account random variation of "true" rates across areas.[8] To estimate "true" rates more accurately, we used a hierarchical Bayesian model[33–38]—specifically, a Poisson model with area treated as a random effect. We validated the model by showing that, across the 71 areas, inpatient and outpatient counts generated from the model corresponded well to the observed counts. (Details of the model and validation are available from the first author.)

We estimated "true" rates in each area using Gibbs sampling as implemented in WINBUGS 1.4.[39] These estimates can be thought of as weighted averages of the observed rates in each area (as calculated from the previous formulas) and the average rate over all areas in the state (which is 1, because our rates are relative rates). Thus, the weighting "shrinks" each observed rate toward 1. We refer to these estimates as "shrunken" rates. The areas with the most shrinkage are those with the most extreme observed rates and those with the fewest people. Shrunken estimates are more accurate than traditional estimates in predicting small-area hospitalization rates.[29]

Based on the posterior means from Gibbs sampling, we ranked each area using both population-based and disease-based rates. Ranks, even when based on shrunken estimates, are inherently unstable.[40] To describe the extent to which areas were ranked differently when using population-based versus disease-based rates, we estimated how often (in repeated Gibbs samples from the appropriate posterior distribution) the rank produced by the disease-based rate differed from the rank produced by the population-based rate by at least 2 deciles (14 or more ranks). In summarizing results across conditions, we report the number of the 14 extreme areas in which it was "more likely than not" (ie, probability ≥ 0.50) that the disease-based rank differed by more than 2 deciles from the population-based rank.

RESULTS

Table 1 shows the number of people hospitalized and the number treated on an outpatient-only basis for each condition.

We use heart failure (DRG 127) to illustrate the analyses. Table 2 shows the effect of shrinkage on both the population-based and the disease-based hospitalization rates for those areas in the lowest and highest decile based on their observed rates. For example, although the area with the smallest observed population-based hospitalization rate had an observed (relative) rate of 0.43, its shrunken rate was 0.62. The area ranked 65th according to its observed population-based rate was particularly small. Its shrunken rate was pulled a lot toward 1 (from 1.30–1.08), such that after shrinkage it was ranked only 49th. Although shrunken estimates were less spread out than the observed rates, shrinkage rarely caused large changes in rank. In fact, among the 28 observed-versus-shrunken rank comparisons in Table 2, 22 changed by 3 ranks or less and only 2 comparisons (both relating to the very small area with 909 residents) changed ranks by 14 or more.

TABLE 2. Heart Failure (DRG 127) Example: Effect of Shrinkage on Rates in Areas With the Most Extreme Observed Population-Based and Disease-Based Hospitalization Rates

Population-Based Rate (Rank)			Disease-Based Rate (Rank)		
Population	**Observed**	**Shrunk**	**Population**	**Observed**	**Shrunk**
5183	0.43 (1)	0.62 (1)	5183	0.55 (1)	0.78 (3)
3973	0.50 (2)	0.66 (4)	3384	0.67 (2)	0.83 (5)
1291	0.53 (3)	0.66 (3)	11,474	0.67 (3)	0.74 (1)
3776	0.62 (4)	0.71 (7)	3973	0.72 (4)	0.89 (10)
7943	0.62 (5)	0.65 (2)	33,791	0.76 (5)	0.78 (2)
15,271	0.63 (6)	0.67 (5)	6693	0.79 (6)	0.85 (6)
3383	0.66 (7)	0.71 (6)	30,476	0.81 (7)	0.83 (4)
909	1.30 (65)	1.08 (49)	11,904	1.21 (65)	1.17 (66)
11,904	1.32 (66)	1.26 (66)	909	1.22 (66)	1.05 (42)
23,750	1.32 (67)	1.30 (68)	2178	1.22 (67)	1.09 (57)
2178	1.33 (68)	1.16 (58)	4978	1.23 (68)	1.14 (64)
7751	1.41 (69)	1.32 (70)	7828	1.24 (69)	1.17 (69)
3527	1.44 (70)	1.31 (69)	7751	1.25 (70)	1.19 (70)
2899	1.65 (71)	1.44 (71)	8261	1.36 (71)	1.26 (71)

We were primarily interested in the extent to which areas identified as extreme (ie, in the top and bottom deciles) according to their population-based hospitalization rate were also extreme according to their disease-based hospitalization rate. Table 3 shows for heart failure the shrunken population-based hospitalization rate and rank, and the shrunken disease-based hospitalization rate and rank, for the most extreme 14 areas according to their shrunken population-based rates. The area with the lowest population-based hospitalization rate (62% of expected, rank 1) also had a low disease-based hospitalization rate (78% of expected, rank 3). We call the area with the second lowest population-based hospitalization rate (rank 2) "area A" (we refer to it again later). Although area A's population-based hospitalization rate was 65% of expected, based on its disease-based hospitalization rate (105% of expected), it was ranked 41st. Of the 7 areas ranked in the first decile (ranks 1–7) based on their population-based hospitalization rate, 4 were ranked in the 3rd decile or higher (rank 22 or higher) based on their disease-based hospitaliza-

TABLE 3. Heart Failure (DRG 127) Example: Shrunken Rates and Ranks* of 3 Measures For Areas With the Most Extreme Population-Based Hospitalization Rates

Population-Based Hospitalization		Disease-Based Hospitalization		Disease	
Rate	Rank	Rate	Rank	Prevalence	Rank
0.62	1	0.78	3	0.79	11
0.65	2	1.05	41	0.62	1
0.66	3	0.95	22	0.69	4
0.66	4	0.89	10	0.74	7
0.67	5	0.92	13	0.73	6
0.71	6	1.03	38	0.68	3
0.71	7	0.96	23	0.74	8
1.25	65	1.01	33	1.24	66
1.26	66	1.17	66	1.08	53
1.29	67	0.90	11	1.44	71
1.30	68	1.09	58	1.19	63
1.31	69	1.08	52	1.22	65
1.32	70	1.19	70	1.12	59
1.44	71	1.07	50	1.35	69

*Ranks are based on the mean of the posterior distribution of the relative rates estimated by the Gibbs sampler.

tion rate. Of the 7 areas in the 10th decile (ranks 65–71) based on their population-based hospitalization rate, 3 were in the 7th decile or lower (ranks 50 or lower) based on their disease-based hospitalization rate. Notably, when using the disease-based hospitalization rate, 2 areas in the highest population-based hospitalization rate decile (ranks 65 and 67) were ranked lower (ranks 33 and 11 respectively) than 2 of the areas in the lowest population-based hospitalization rate decile (population-based hospitalization rate ranks 2 and 6, which were ranked 41st and 38th based on disease-based hospitalization rate).

The right 2 columns in Table 3, which show estimates of disease prevalence, demonstrate why population-based hospitalization rates and disease-based hospitalization rates can be so different. Area A, discussed earlier, had 62% of the amount of disease expected (the lowest estimated disease prevalence). Its low population-based hospitalization rate was largely a reflection of the low disease prevalence in the area. After taking disease prevalence into account (by using the disease-based hospitalization rate), the area actually had 5% more hospitalizations than expected. All 7 areas with the lowest population-based hospitalization rates had very low levels of disease and, to a large extent, this explains their very low population-based hospitalization rates. The highest decile demonstrates the same phenomenon. The area ranked 67th in terms of its population-based hospitalization rate had 44% more disease than expected. When this was taken into account, the proportion hospitalized in the area was 90% of expected rather than the 29% more than expected indicated by the population-based hospitalization rate.

Figure 1 is a box plot of shrunken population-based and disease-based hospitalization rates for the 15 conditions. Many of the conditions exhibited more variability in rates than heart failure. For most of the conditions, population-based rates varied from around 50% of expected to 50% more than expected. Disease-based rates also varied considerably, although in most conditions somewhat less than population-based rates.

Table 4 shows for each of the 15 conditions for areas in the top decile of population-based hospitalization rates, what their rank would be according to their disease-based hospitalization rate. To illustrate, for transient ischemic attacks (TIAs, top row) the area ranked highest (71st) based on its population-based hospitalization rate was ranked 62nd based on its disease-based hospitalization rate; the area ranked second highest based on its population-based hospitalization rate (70th) was ranked 49th based on its disease-based hospitalization rate. The numbers in parentheses indicate the fraction of the time in repeated Gibbs samples from the posterior distributions of the ranks that the 2 ranks differed by at least 2 deciles (shown if the fraction was ≥ 0.50). For example, for TIA for the highest ranked area based on its population-based hospitalization rate, there was a 0.50 chance that the disease-based hospitalization rate rank differed from the population-based hospitalization rate rank by at least 2 deciles. For the second highest ranked area based on its

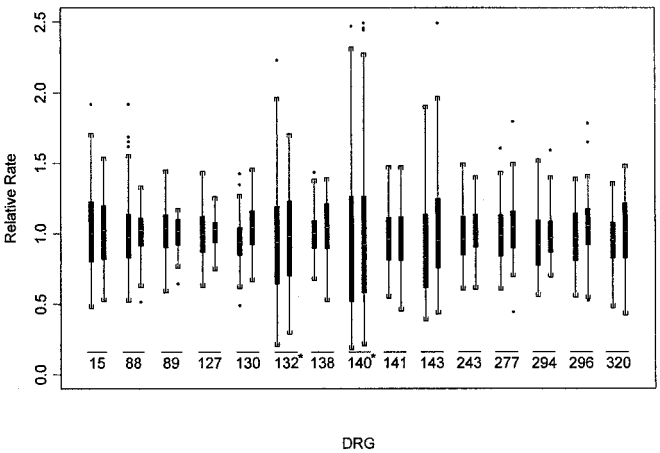

FIGURE 1. Box plot of shrunken population-based hospitalization rates (left) and disease-based hospitalization rate (right) by condition. To retain the graph's scale, we excluded extreme points, as follows:

DRG 132, right plot: 3.2

DRG 140, left plot: 2.7, 2.8, 3.1, 3.2, 3.4, 3.9, 5.4; right plot: 2.8, 4.0, 4.3

TABLE 4. For Areas With the Highest Ranked Shrunken Population-Based Hospitalization Rates, Rank of the Shrunken Disease-Based Hospitalization Rate and Probability That the 2 Ranks Differ by at Least 14*

	Rank of Population-Based Hospitalization Rate						
	71	**70**	**69**	**68**	**67**	**66**	**65**
Condition	**Rank of Disease-Based Hospitalization Rate (Probability*)**						
Transient ischemic attack	62 (0.50)	49 (0.80)	66	65	45 (0.74)	59	63
Chronic bronchitis and emphysema	64	68	52 (0.72)	69	34 (0.93)	71	58 (0.53)
Pneumonia	51 (0.61)	50 (0.70)	36 (0.99)	67	65	71	55
Heart failure	50 (0.74)	70	52 (0.66)	58 (0.65)	11 (0.99)	66	33 (0.93)
Peripheral vascular disorders	68	71	4 (0.99)	23 (0.98)	16 (0.99)	57	24 (0.92)
Ischemic heart disease	68	70	67	71	69	60	63
Arrhythmia	70	15 (0.99)	48 (0.62)	47 (0.70)	51 (0.52)	67	44 (0.84)
Angina	71	65	68	64	70	66	69
Syncope	67	71	66	70	55	69	65
Chest pain	55 (0.63)	70	71	69	68	66	50 (0.93)
Back problems	61 (0.52)	71	68	63	39 (0.84)	34 (0.94)	35 (0.83)
Cellulitis	67	69	64	60	31 (0.98)	65	19 (0.99)
Diabetes	69	68	62	13 (0.92)	39 (0.79)	58	52
Fluid and electrolyte disorder	5 (0.99)	59	63	70	65	68	66
Kidney and urinary tract infections	71	56	64	54	40 (0.95)	68	46 (0.61)

*Ranks are based on the means of the posterior distribution of the rates estimated by the Gibbs sampler. Probabilities are the fraction of the times that the 2 ranks, as sampled from the posterior distributions, differed by 14 or more.

population-based hospitalization rate, the probability was 0.80.

For 4 conditions (heart failure, peripheral vascular disorders, arrhythmia, and back problems), 10 or 11 of the 14 extreme areas according to their population-based hospitalization rate were "more likely than not" (ie, with a probability of 0.50 or greater) to differ by at least 2 deciles when ranked according to their disease-based hospitalization rate. For 7 conditions (TIA, chronic obstructive pulmonary disease, pneumonia, chest pain, cellulitis, diabetes, and electrolyte disorder), 5 or 6 of the 14 extreme areas based on the population-based hospitalization rate were "more likely than not" to differ by at least 2 deciles when ranked by their disease-based hospitalization rate. Only for ischemic heart disease, angina, and syncope was it more likely than not that most of the areas (12 or more) were within 2 deciles when ranked by each rate.

DISCUSSION

Population-based hospitalization rates varied widely across small areas in Massachusetts. However, disease prevalence also varied widely. Because of differences in disease prevalence, areas with extreme population-based hospitalization rates were not necessarily areas in which a particularly high or low proportion of those with the disease were admitted to the hospital, relative to what was expected given the age and gender distribution in the area.

Our approach relies on claims data to identify people with specific diseases. Because of concern about the validity of disease identification from outpatient codes, we conducted analyses using both a stringent definition (the series method) and a lenient definition (which considered all visits) for identifying disease from diagnostic codes on outpatient claims. Both analyses showed substantial differences in area ranks when calculated with population-based versus disease-based rates. However, claims data can be imperfect in many ways and are unlikely to capture all conditions noted on the medical record.[41] Our conclusions do not depend on accurately estimating actual disease prevalence, but on estimating relative disease prevalence (ie, how much more or less disease one area has compared with another). The key assumption behind our analysis is that outpatient coding is not systematically biased across areas. However, areas with better access to technology or more specialists may identify more disease than other areas with a similar disease burden, or they may code presentations to justify use of the technology or referral to specialists. We are currently exploring the potential for such bias by examining the relationship, within areas, between specialist physician supply and claims-based disease prevalence.

Another concern when identifying disease from claims is that higher rates of coded disease may reflect better access to primary care or may be a proxy for either physician or hospital bed supply. In a Medicare population eligible for

both hospital and outpatient care, however, financial access differences are at least somewhat muted. Also, by focusing on the number of people with any visit for a condition rather than total numbers of medical encounters by those with the condition, we reduce the effect of practice style on our measure of disease prevalence. If supply were the main factor driving demand, the same areas that have high rates of disease would have high rates of hospitalization among those with the disease. Correlations between disease prevalence and the proportion of those with the disease who were hospitalized were, in fact, statistically significant for 9 of the 15 conditions. However, in all 9, the correlation was negative. Finally, despite concerns about outpatient coding, CMS judges them sufficiently valid that, starting in 2004, it will accept diagnoses from either inpatient or outpatient claims to calculate health-based payments to Medicare+Choice HMOs for the Medicare beneficiaries they enroll.

As health care costs continue to outpace general inflation, pressure is mounting to revitalize certificate-of-need programs. For example, the *Wall Street Journal* recently wrote: "the Big Three [auto companies] have lobbied aggressively to keep certificate-of-need programs in states such as Missouri and have fought ardently for the establishment of programs in Ohio and Indiana."[42] Vermont recently issued a request for proposal for consultation services to develop a health resource allocation plan to guide health facility planning and capital expenditures. Hospital capacity is a major focus of such programs. Motivating decisions about hospital capacity are area-specific hospitalization rates. Our analyses suggest that disease prevalence rather than population counts provide a more appropriate denominator for such rates. Unfortunately, outpatient claims are not widely available. At a minimum, our study indicates that databases with outpatient as well as inpatient claims are needed to address better the research and policy questions raised by geographic variations.

Many current quality improvement efforts examine process measures for patients with specific conditions (eg, whether specifically identified patients with heart disease receive appropriate medications or diabetic patients receive appropriate assessment and preventive therapy). Whether patients with specific conditions are hospitalized is an important process measure, both because of iatrogenic events and costs. Although we have focused on variations across geographic areas, our approach applies as well to examining variations in hospitalizations across physician practices. Because of small samples, shrinkage estimators are even more important in this setting.

Our study examined only 15 medical conditions using 1 year of data from 1 state. Also, some Medicare beneficiaries use both Veteran Administration (VA) and non-VA facilities and we have not included diagnostic or utilization data from VA facilities.[43,44] Nonetheless, the large differences in disease prevalence across small areas that remain after adjusting for age and gender raise concerns about the value of population-based hospitalization rates for studying hospital utilization and drawing inferences about physician practice styles.

ACKNOWLEDGMENT

The manuscript has benefited greatly from suggestions from the anonymous reviewers and the editor.

REFERENCES

1. Barnes BA, O'Brien EZ, Comstock C, et al. Report on variation in rates of utilization of surgical services in the Commonwealth of Massachusetts. *JAMA*. 1985;254:371–375.
2. Carlisle DM, Valdez RB, Shapiro MF, et al. Geographic variation in rates of selected surgical procedures within Los Angeles County. *Health Serv Res*. 1995;30:27–42.
3. Connell FA, Day RW, LoGerfo JP. Hospitalization of Medicaid children: analysis of small area variations in admission rates. *Am J Public Health*. 1981;71:606–613.
4. Folland S, Stano M. Small area variations: a critical review of propositions, methods and evidence. *Med Care Rev*. 1990;47:419–465.
5. Gittelsohn AM, Powe NR. Small area variations in health care delivery in Maryland. *Health Serv Res*. 1995;30:295–317.
6. Lewis CE. Variations in the incidence of surgery. *N Engl J Med*. 1969;281:880–884.
7. McMahon LF, Wolfe RA, Tedeschi PJ. Variation in hospital admissions among small areas: a comparison of Maine and Michigan. *Med Care*. 1989;27:623–631.
8. McPherson S, Wennberg JE, Hovind OB, et al. Small-area variations in the use of common surgical procedures: an international comparison of New England, England and Norway. *N Engl J Med*. 1982;307:1310–1314.
9. Paul–Shaheen P, Clark JD, Williams D. Small area analysis: a review and analysis of the North American literature. *J Health Polit Policy Law*. 1987;12:741–809.
10. Roos NP, Roos LL Jr. Surgical rate variations: do they reflect the health or socioeconomic characteristics of the population. *Med Care*. 1982;20:945–958.
11. Roos NP. Hysterectomy: variations in rates across small areas and across physicians' practices. *Am J Public Health*. 1984;74:327–335.
12. Wennberg JE, Gittelsohn A. Health care delivery in Maine I: patterns of use of common surgical procedures. *J Maine Med Assoc*. 1975;66:123–130.
13. Wennberg JE, Gittelsohn A. Variations in medical care among small areas. *Sci Am*. 1982;246:120–133.
14. Berstein SJ, Hilborne LH, Leape LL, et al. The appropriateness of use of coronary angiography in New York State. *JAMA*. 1993;269:766–769.
15. Chassin MR, Kosecoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA*. 1987;258:2533–2537.
16. Hilborne LH, Leape LL, Bernstein SJ, et al. The appropriateness in use of percutaneous transluminal coronary angioplasty in New York state. *JAMA*. 1993;269:761–765.
17. Leape LL, Park RE, Solomon DH, et al. Does inappropriate use explain small-area variations in the use of health care services? *JAMA*. 1990;263:669–672.
18. Leape LL, Hilborne LH, Park RE, et al. The appropriateness of use of coronary artery bypass graft surgery in New York state. *JAMA*. 1993;269:753–760.
19. Restuccia JD, Shwartz M, Kreger BE, et al. Does more "appropriateness" explain higher rates of cardiac procedures among patients hospitalized with coronary heart disease? *Med Care*. 2002;40:500–509.
20. Restuccia JD, Shwartz M, Ash A, et al. High hospital admission rates and inappropriate care. *Health Aff*. 1996;15:156–163.
21. Fisher ES, Welch HG, Wennberg JE. Prioritizing Oregon's hospital resources: an example based on variations in discretionary medical utilization. *JAMA*. 1992;267:1925–1931.

22. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Aff* 13 February 2002. Available at www.healthaffairs.org/WebExclusives/Wennberg_Web_Excl_021302.htm.
23. Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problem of supplier-induced demand. *Soc Sci Med*. 1982;16:811–824.
24. Chassin MR. Explaining geographic variations: the enthusiasm hypothesis. *Med Care* 1993;31:YS37–YS44.
25. Fisher ES, Wennberg JE, Stukel TA, et al. Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. *Health Serv Res*. 2000;34:1351–1362.
26. Cumming RB, Cameron BA. *A Comparative Analysis of Claims-Based Methods of Health Risk Assessment for Commercial Populations*. Available at www.soa.org/sections/riskadjfinalreport1.pdf.
27. Shwartz M, Iezzoni LI, Ash AS, et al. *More Disease: How Major a Factor in Higher Utilization?* Final report submitted to the Agency for Healthcare Research and Quality, August 2001 (Grant No. R03 HS09832-01).
28. Wennberg JE. Population illness rates do not explain population hospitalization rates: a comment on Mark Blumberg's thesis that morbidity adjusters are needed to interpret small area variations. *Med Care* 1987; 25:339.
29. Shwartz M, Ash AS, Anderson J, et al. Small area variations in hospitalization rates: how much you see depends on how you look. *Med Care*. 1994;32:189–201.
30. Ward JH Jr. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc*. 1963;58:236–244.
31. Romesburg CH. *Cluster Analysis for Researchers*. London: Wadsworth; 1984.
32. Rothman KJ. *Modern Epidemiology*. Boston: Little, Brown; 1986.
33. Gelman A, Carlin JP, Stern HS, et al. *Bayesian Data Analysis.* 2nd edition. Boca Raton, FL: Chapman & Hall/CRC; 2003.
34. Carlin BP, Louis TA. *Bayes and Empirical Bayes Methods for Data Analysis.* 2nd ed. Boca Raton, FL: Chapman & Hall/CRC; 2000.
35. Bryk AS, Raudenbush SW. *Hierarchical Linear Models: Applications and Data Analysis Methods. Advanced Quantitative Techniques in the Social Sciences*. Newbury Park: Sage Publications; 1992.
36. Gatsonis CA, Epstein AM, Newhouse JP, et al. Variations in the utilization of coronary angiography for elderly patients with an acute myocardial infarction: an analysis using hierarchical logistic regression. *Med Care*. 1995;33:625–642.
37. Christiansen CL, Morris CN. Improving the statistical approach to health care provider profiling. *Ann Intern Med*. 1997;127:764–768.
38. Normand S-LT, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: issues and applications. *J Am Stat Assoc*. 1997;92:803–814.
39. The software and documentation are available at www.mrc-bsu.cam.ac.uk/bugs/.
40. Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance [with discussion]. *J R Stat Soc A*. 1996;159:385–443.
41. Fowles JB, Lawthers AG, Weiner JP, et al. Agreement between physicians' office records and Medicare part B claims data. *Health Care Financ Rev*. 1995;16:189–199.
42. Hawkins L Jr. Detroit auto makers fight local hospitals' building plans. *The Wall Street Journal*. January 20, 2004. Available at: http://online.wsj.com.
43. Fleming C, Fisher ES, Chang CH, et al. Studying outcomes and hospital utilization in the elderly: the advantages of a merged data base for Medicare and Veterans Affairs hospitals. *Med Care*. 1992;30:377–391.
44. Ashton CM, Souchek J, Petersen NJ, et al. Hospital use and survival among Veterans Affairs beneficiaries. *N Engl J Med*. 2003;349:1637–1646.