

Using diagnosis-related groups for studying variations in hospital admissions

by Noralou P. Roos, John E. Wennberg, and Klim McPherson

The diagnosis-related groups (DRG's) have classically focused on resources consumed during a hospital stay. DRG's can also be considered categories for describing cases admitted to a hospital. In this article, we illustrate how consistent patterns of variations in admission rates can be used to classify DRG categories according to the Index of

Discretionary Admissions. The consistency of variation in admission rates for modified DRG categories across hospital service areas in Iowa, California, Massachusetts, and Maine was high. The proportion of hospital admissions in the DRG's judged to be most discretionary ranged from 22 percent in Iowa to 14 percent in California.

Introduction

The diagnosis-related groups (DRG's) were developed as a method of classifying hospitalized patients into categories having a similar process of care and a predictable range of services (Fetter et al., 1980). DRG's have subsequently been used for prospective payment for hospitalized Medicare patients. Although input from physicians was used to ensure that DRG categories were clinically meaningful, the length of a patient's hospital stay was the primary grouping criterion because length of stay is highly correlated with case complexity and total charges and because data on this measure are available and reliable (Fetter et al., 1980). One purpose of the DRG system is to help hospitals and reimbursement agencies manage the discretionary component in hospital expenditures. If physicians or hospitals routinely utilize more resources than the recognized standard, they are penalized through the reimbursement system.

However, the length of stay in a hospital is only one of the two primary factors determining the per capita numbers of patient days per 1,000 population, the other being the frequency of admission. As it turns out, for most diagnostic categories, the admission rate is much more important than the length of stay as a predictor of the per capita patient days of care among populations living in hospital market areas (Wennberg, McPherson, and Caper, 1984; Knickman and Foltz, 1985; Roos et al., 1986).

Small-area variation studies provide a growing body of evidence on the importance of discretionary physician decisions concerning the necessity of hospitalization. By discretionary decisions we mean

decisions made with uncertainty. When the clinical outcomes of one treatment alternative over another are not well established by outcome studies and physicians disagree among themselves on the appropriate course of care, the decision to hospitalize is considered discretionary.

The argument for the importance of professional discretion in explaining variation is based on characteristics of the pattern of variations and on evidence that rates of illness and other demand-side factors such as income and insurance coverage are not correlated with variations in hospital use among small areas (Wennberg, Barnes, and Zubkoff, 1982; Wennberg, 1987; Roos and Roos, 1982).

Moreover, the degree of variation in per capita admission rates across geographic areas for a given condition seems to be inversely related to the degree of professional consensus about the appropriateness of treatment requiring acute hospitalization for that condition. Little variation in hospital admission rates has been found for myocardial infarction and inguinal hernia repair, conditions for which current standards for care generally mandate treatment in a hospital. However, per capita rates of hospital admissions for certain surgical procedures, such as hysterectomy, prostatectomy, and tonsillectomy, have been shown to vary widely among hospital service areas (Wennberg and Gittelsohn, 1982; McPherson et al., 1982; McCracken, Latessa, and Wennberg, 1982). The medical literature is replete with controversy concerning the value of these procedures (Wennberg and Gittelsohn, 1982; Deane and Ulene, 1977; Bunker, McPherson, and Henneman, 1977; Harper, 1962; Bolande, 1969), and changes in surgical rates have been noted once physicians became aware of the variations (Dyck et al., 1977; Wennberg et al., 1977). The importance of physician practice styles and professional discretion in influencing the number of operations performed and, hence, the number of hospitalizations that take place is clearly shown in these studies.

In contrast to efforts to rationalize hospital length of stay, little attention has been paid to variations in admission rates. In this article, we show the feasibility of using a measure of variation in hospitalization rates to classify DRG categories according to the likelihood that patients in a category have illnesses

This research was supported by Grant No. 04932 from the National Center for Health Services Research. The senior author received a Career Scientist Award (6607-1001-22) and Grant No. 6607-1419-57 from the National Health Research and Development Program. Dr. Wennberg received a grant from the Hartford Foundation. An earlier version of this article was presented at the Conference on Management and Financing of Hospital Services, Dec. 1986, in London, England.

Reprint requests: Noralou P. Roos, Department of Community Health Sciences, S101D, 750 Bannatyne, Winnipeg, Canada R3E0W5.

and conditions for which professional discretion plays an important role in the decision to hospitalize or operate. This index of discretion should allow priorities to be set in the agenda to reduce professional uncertainty through research on outcomes. Moreover, it can serve as a practical guide for physicians, hospital administrators, peer review bodies, and third-party payers to identify hospitals and physicians according to their practice styles with regard to discretionary DRG's.

Based on consistencies in the pattern of variation among DRG's in four States—California, Massachusetts, Maine, and Iowa—we propose to classify DRG's according to the Index of Discretionary Admissions. In the article, we deal with methodologic issues related to determining the consistency of variation measures across States and discuss the use of this information in reducing the discretionary component of hospital admissions.

Methods

Uniform hospital discharge data sets were used in each of the four States to define hospital market areas according to standard methods (Wennberg and Gittelsohn, 1982). All hospitalizations for all age groups in 1983 (California), 1982 (Massachusetts), 1980 (Iowa), and 1980-82 (Maine) were available. The data on utilization were drawn from the period before the implementation of DRG-based reimbursement, so coding practices and admission policies were not influenced by concerns about DRG-based reimbursement. DRG-specific utilization rates were calculated for each of 30 hospital market areas in Maine, 23 hospital market areas in Iowa, 30 planning areas in California, and 60 areas in Massachusetts. The 22 DRG's pertaining to obstetrical and neonatal conditions were excluded. Utilization rates were adjusted for age and sex by the indirect method.

A modified DRG case-mix classification system was used to simplify the results and reduce variation resulting from minor differences in coding practices. A full listing of the component DRG's contained in each modified group can be found in Figure 1. A total of 111 medical and surgical modified DRG's resulted from the following grouping strategy, developed while working with the Maine data.

- Individual DRG's with unique primary diagnoses were left ungrouped if the case count exceeded 1,500.
- DRG's with the same primary diagnosis were grouped, thus removing the variables of age and presence or absence of a serious complication, both allowable variables under the DRG classification system. However, pediatric DRG's with more than 1,500 cases were left ungrouped.
- For surgical DRG's, related operations were grouped when each represented fewer than 1,500 cases.
- DRG groups that contained fewer than 1,500 cases after the foregoing steps were grouped within their major disease category.

These steps resulted in 77 closely related modified DRG's and 34 residual modified DRG's. These latter were excluded from all analyses unless otherwise noted. All utilization rates were corrected for market-boundary crossing; they are based on population and are specific for the residents of each market whether services were obtained at local or out-of-area hospitals.

For measuring and comparing variation in rates, we used the systematic component of variation (SCV), a measure developed by McPherson et al. (1981). It is important that the index used to measure variation not be greatly affected by either the size of the hospital market area or the prevailing rate under study. The SCV is based on the assumption that observed rates of hospitalization for any category of DRG will vary across small areas because of two factors, random variation and systematic variation. The random component of variation measures differences caused by regional sizes and by the magnitude of actual rates. For instance, small-sized groups may produce excess random variation in rates. The systematic component of the total variability across observed small-area rates is our primary interest. Such systematic variation in rates may be attributed to the underlying characteristics of different regions, including differences in population health characteristics, or to such factors as differences in physician practice style or differences in the availability of health care services across regions. (Differing age and sex composition across regions is accounted for through age and sex standardization.)

The systematic component of variation is estimated using a proportional hazards model by subtracting the random component of variation from the estimate of total variance. This measure is useful for comparing relative variation among DRG's because, unlike the coefficient of variation or chi-square statistic, it is not correlated with the prevailing rate. Thus, the SCV can be used to estimate the relative variation in admission rates for each DRG, even though mean rates differ substantially. Characteristics of small-area rates that produce low and high SCV's are shown in the technical note. Details on the necessary distributional assumptions and calculation of the SCV are described in detail in the original work by McPherson et al. (1981).

First, to determine consistency of variation scores across the four States, within each State, the 77 modified DRG categories were ordered by SCV scores and divided into decile ranks. (The five psychiatric DRG categories were not available for California.) Two standards of consistency in variation scores were applied across either three States or four States, yielding a total of four definitions. According to the stricter standard, the DRG category was judged to have a consistent pattern of variation in admission rates if every State's variation scores, when ranked in deciles, were within two deciles of the scores of the other States. According to the second standard, variation scores were determined to be consistent if every State's decile-ranked score was within three

Figure 1

Modified diagnosis-related group (DRG) categories by whether medical or surgical and degree of variation in hospitalization rate

Medical DRG category and DRG code number inclusion	Surgical DRG category and DRG code number inclusion	Medical DRG category and DRG code number inclusion	Surgical DRG category and DRG code number inclusion
Low variation †Acute myocardial infarction (121-123) †Specific cerebrovascular disorders (14)	Low variation †Inguinal and femoral hernia operations (161-162) †Major small and large bowel operation (148-149) Gallbladder disease with cholecystectomy (195-198) †Hip repair except joint replacement (210-211)	Very high variation I—Con't. Circulatory disorders excluding acute myocardial infarction, with cardiac catheterization (124-125) Red blood cell disorders (395-398) Cellulitis (277-279) Adult diabetes (294)	
Moderate variation †Gastrointestinal hemorrhage (174-175)	Moderate variation Appendicitis with appendectomy (164-167)	Very high variation II Chest pain (143) Deep vein thrombophlebitis (128) †Acute adjustment reaction (425) Adult otitis media and upper respiratory infection (68-69) Respiratory signs and symptoms (99-100) Chronic obstructive lung disease (88) Trauma to skin, subcutaneous tissue, and breast (280-282) Adult bronchitis and asthma (96-97) †Hypertension (134) †Pediatric gastroenteritis (184) †Peptic ulcer (176-178) †Pediatric bronchitis and asthma (98) †Organic mental syndromes (438) †Atherosclerosis (132-133) †Pediatric otitis media and upper respiratory infection (70) †Pediatric pneumonia (91) †Chemotherapy (410)	Very high variation II Tonsillectomy (59-60) Extraocular operations (40-41) Miscellaneous ear, nose, and throat operations (55) †Other tonsillectomy and adenoidectomy operations (57-58) Dilatation and curettage, conization except for malignancy (364) Breast biopsy and local excision for nonmalignancy (262) †Other female laparoscopic operations (361) †Tubal interruptions for nonmalignancy (359) †Dental extractions and restorations (187) †Laparoscopic tubal interruption (362)
High variation †Urinary tract stones (323-324) Digestive malignancy (172-173) †Syncope and collapse (141-142) Heart failure and shock (127) †Respiratory neoplasms (82) Seizures and headaches (24-25) Angina pectoris (140) Toxic effects of drugs (449-450) †Cardiac arrhythmias (138-139) Gastrointestinal obstruction (180-181) †Adult simple pneumonias (89-90) Miscellaneous injuries to extremities (253-254)	High variation Hysterectomy (354-355) Major joint operations (209) Major cardiovascular operations (103-111) Other adult hernia operations (159-160) Anal operations (157-158) Soft tissue operations (226-227) †Lens operations (39) Foot operations (225) Transurethral operations (310-311)		
Very high variation I Kidney and urinary tract infections (320-321) Adult gastroenteritis (182-183) †Medical back problems (243)	Very high variation I Back and neck operations (214-215) Hand operations except ganglion (229) Knee operations (221-222)		

†Variation scores within 2 decile ranks across 4 States.
NOTE: All categories shown in this figure have consistency in variation scores across at least 3 States.
SOURCE: Maine Health Information Center, Augusta, Maine; Iowa Foundation for Medical Care, Des Moines, Iowa; Massachusetts Rate Setting Commission, Boston, Massachusetts; California Health Facilities Commission, Sacramento, California: Uniform hospital discharge data sets.

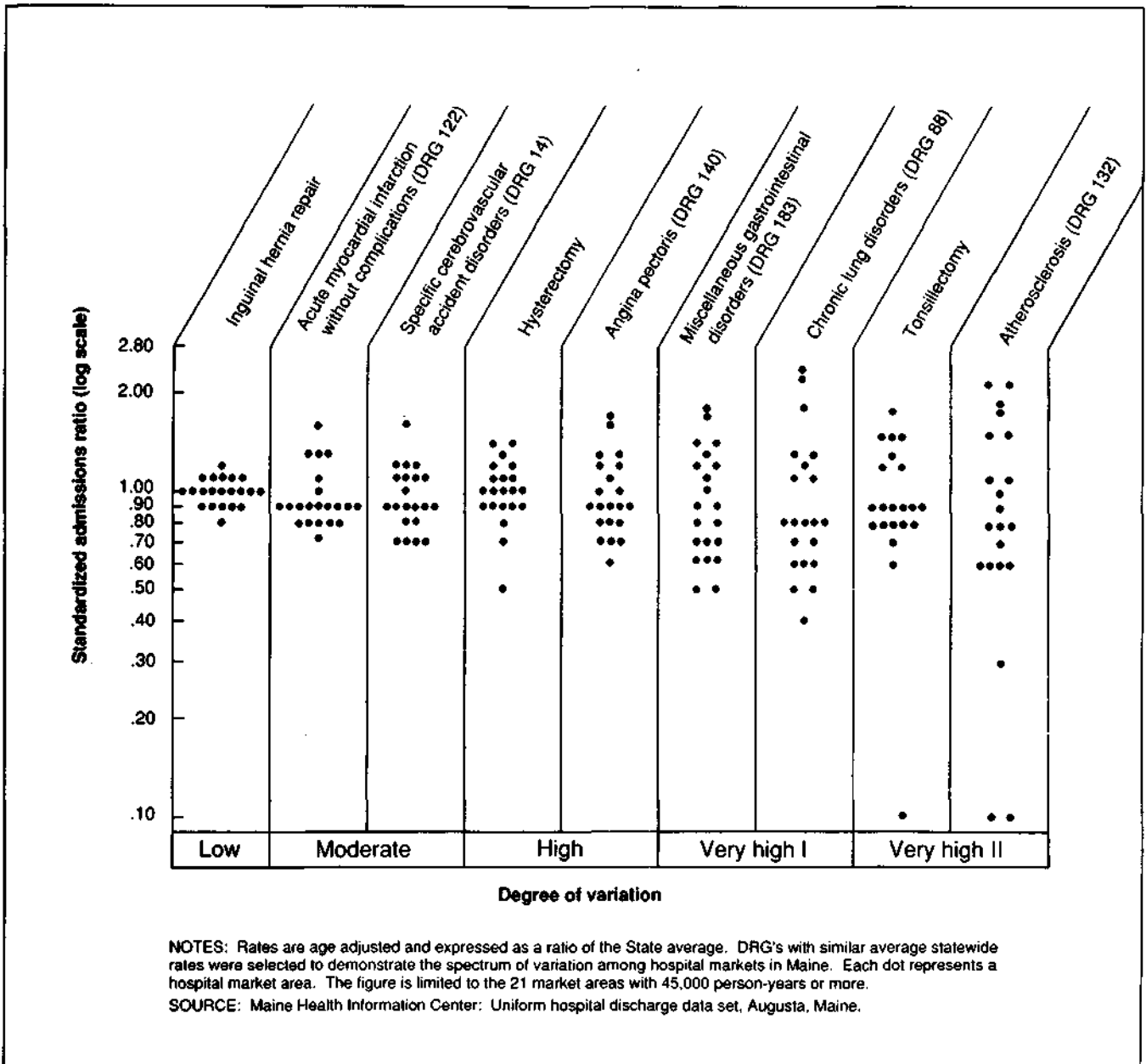
deciles of that of the other States. If all decile-ranked variation scores fell within the range of three to six, there would be agreement according to the second concept but not the first. Four definitions of consistency resulted:

- I. Decile-ranked variation scores were within three deciles for at least three of the States.
- II. Decile-ranked variation scores were within two deciles for at least three of the States.
- III. Decile-ranked variation scores were within three deciles for all four States.

IV. Decile-ranked variation scores were within two deciles for all four States.

Second, the modified DRG's were classified with regard to their relative variation; a mean variation score for each DRG category was calculated using the SCV scores for the three States with the closest decile ranks. The five-point Index of Discretionary Admissions was developed based on this mean score (Figure 1). DRG's with values that had ranks ranging from the rank for inguinal hernia up to (but not including) the rank for appendectomy were labeled

Figure 2
Hospitalization rates for selected medical diagnosis-related groups (DRG's) and surgical procedures, by degree of variation: 21 Maine market areas, 1980-81



“low variation”; those from appendectomy up to hysterectomy, “moderate variation”; those from hysterectomy up to back surgery, “high variation”; those from back surgery up to tonsillectomy, “very high I”; and those as high or higher than tonsillectomy, “very high II” (Wennberg, McPherson, and Caper, 1984).

The patterns of admission rates that produce these differing variation scores are illustrated in Figure 2. For the low variation procedure inguinal hernia repair, all 21 areas in Maine had admission rates that clustered in a narrow range. For the very high variation procedures tonsillectomy and atherosclerosis,

the range of variation in admission rates across hospital areas was tenfold.

Results

For a preliminary look at consistency patterns, the five modified DRG categories with the highest variation scores and the five categories with the lowest variation scores were determined for each State. The variation in admission rates across the four States was strikingly consistent. Admissions for laparoscopic tubal interruptions and tubal interruptions for nonmalignancies had variation scores among the five

highest in each of the four States.¹ Admissions for chemotherapy ranked among the highest five categories in three States. (They had the 12th highest variation score in California.) At the other end of the scale, admissions for inguinal and femoral hernia repairs, as well as admissions for acute myocardial infarction, had among the lowest variation scores in each of the four States.

In Figure 1, we list the 66 DRG categories (out of 77 reviewed) that had a consistent pattern of variation in admission rates across hospital areas using Criterion II. If a DRG category's SCV score ranked in the first through third deciles (or in the sixth through eighth, seventh through ninth, etc.) in at least three out of four States, the DRG category was judged to have a consistent pattern of variation in admission rates. DRG categories are listed in ascending order of variation. Categories with variation scores within two deciles across four States (Criterion IV) have been marked with a dagger. A majority of the DRG categories with the highest scores were consistent across all four States.

Because categories based on the three-State criterion might include situations in which one State was widely divergent from the other three States, each category was also examined for signs of marked inconsistency

¹We were concerned that the high variation for the two tubal ligation DRG categories might result from different coding practices, that areas with high rates in one category of tubal ligations would be low in the other. However, this does not appear to be the case. We assessed the relationship of rates for one type of tubal ligation to rates for the other across Maine's 20 hospital areas. The Pearson correlation coefficient was -.26 and the Spearman coefficient -.08. Neither reached a level of statistical significance of .10 or lower.

in variation scores. The scores were judged markedly inconsistent if, in any category, one State or more had an SCV score in the top three deciles (1-3) and simultaneously one State or more had a variation score in the bottom three deciles (8-10). The results are shown in Table 1. Only 4 of the 77 categories had markedly inconsistent variation scores across the four States, and only 1 category (major cardiovascular operations) was included among the 66 judged consistent in Figure 1. The three-State criterion, therefore, provides little overestimation of the degree of consistency in the scores.

A statistical assessment of the consistency in variation scores across the 77 DRG categories, using

Table 1
Diagnosis-related group (DRG) categories with markedly inconsistent systematic component of variation (SCV) scores among States: Maine, 1980-82; Iowa, 1980; Massachusetts, 1982; and California, 1983

DRG category	Deciles of State SCV scores
Major cardiovascular operations	3,3,3,10
Stomach, esophageal, and duodenal operations	3,5,6,8
Uterus and adenexa operations	2,3,7,8
Pediatric hernia operations	2,3,9,10

NOTE: For these DRG categories, at least 1 State had a score in the lowest 3 deciles (1-3) of variation and at least 1 had a score in the highest 3 deciles (8-10).

SOURCE: Maine Health Information Center, Augusta, Maine; Iowa Foundation for Medical Care, Des Moines, Iowa; Massachusetts Rate Setting Commission, Boston, Massachusetts; California Health Facilities Commission, Sacramento, California; Uniform hospital discharge data sets.

Table 2
Pearson correlation coefficients for systematic component of variation (SCV) scores and Spearman rank-order correlation coefficients for rankings of SCV scores, by type of DRG category: Maine, 1980-82; Iowa, 1980; Massachusetts, 1982; and California, 1983

Type of DRG category and State	Pearson coefficients for SCV scores			Spearman coefficients for rankings of SCV scores		
	Maine	Iowa	Massachusetts	Maine	Iowa	Massachusetts
All categories¹						
California	.55	.44	.69	.70	.60	.79
Maine		.84	.80		.67	.73
Iowa			.66			.67
Surgical only²						
California	.66	.60	.69	.74	.56	.82
Maine		.76	.84		.80	.74
Iowa			.66			.60
Medical only³						
California	.70	.50	.84	.68	.67	.74
Maine		.89	.80		.59	.75
Iowa			.72			.75

¹73-77 categories, depending on State.

²31 categories.

³42-46 categories, depending on State.

NOTE: All correlation coefficients in this table are significantly greater than zero at the 0.001 level.

SOURCE: Maine Health Information Center, Augusta, Maine; Iowa Foundation for Medical Care, Des Moines, Iowa; Massachusetts Rate Setting Commission, Boston, Massachusetts; California Health Facilities Commission, Sacramento, California; Uniform hospital discharge data sets.

the Pearson product-moment correlation coefficients for the actual scores and the Spearman rank-order correlation coefficients for the ranked scores, is presented in Table 2. The coefficients for the ranked scores tended to be higher when States had three or four DRG categories with extremely high SCV values. Usually this was because one or two areas had expected or observed values with less than five observations. Recalculating the coefficients after weighting each category by number of admissions increased the size of the coefficients only marginally (for example, from .44 to .49 for the correlation between California and Iowa across all DRG

categories), confirming the robustness of the SCV measure.

Overall, the correlations, whether by actual score or ranked score, demonstrate a high level of consistency in variation scores across the four States. When the coefficients were calculated separately for the surgical categories and medical categories, the coefficients generally increased. The scores of Iowa and California appear to be least similar (showing the lowest correlation coefficients), although even these correlations were highly significant statistically.

In Table 3, we present data on the percentage of DRG groups and percentage of hospitalized cases assigned to DRG groups that were judged to have consistent variation scores using the four definitions of consistency. Using the least stringent criterion (I—three States having SCV scores within three deciles), 93.5 percent of the DRG categories were judged to have consistent variation scores. Using the most stringent criterion (IV—four States having SCV scores within two deciles), 40.0 percent of the DRG categories were judged consistent with regard to variation. The proportion of hospitalizations in each of the four States that were for DRG categories judged to have consistent variation scores also varied markedly depending on the standards applied (Table 3). Requiring four States (rather than three) to have similar variation patterns appears to be more restrictive than narrowing the range of scores defined as similar.

In Table 4, all hospitalizations in each of the four States are classified according to the Index of Discretionary Admissions. Criterion II is used to define consistency in the columns on the left, and Criterion IV is used in the columns on the right. Hospitalizations for DRG categories that had inconsistent variation scores and those that fell among the 34 residual categories are listed separately.

Table 3
Percent of diagnosis-related group (DRG) categories and of hospitalizations judged to have consistent systematic component of variation scores among States using 4 criteria: Maine, 1980-82; Iowa, 1980; Massachusetts, 1982; and California, 1983

Item	Consistency of systematic component of variation scores			
	Across 3 States		Across 4 States	
	I—within 3 deciles	II—within 2 deciles	III—within 3 deciles	IV—within 2 deciles
	Percent			
DRG categories	93.5	85.7	61.0	40.0
Hospitalizations:				
Maine	94.3	89.8	60.6	41.2
Iowa	94.3	90.2	61.4	41.1
Massachusetts	94.7	90.4	61.3	40.4
California	97.4	92.8	58.1	38.6

NOTE: Calculations are based on 73-77 DRG categories that represent 166,517-474,608 cases, depending on State. Residual DRG categories are excluded.

SOURCE: Maine Health Information Center, Augusta, Maine; Iowa Foundation for Medical Care, Des Moines, Iowa; Massachusetts Rate Setting Commission, Boston, Massachusetts; California Health Facilities Commission, Sacramento, California; Uniform hospital discharge data sets.

Table 4
Percent distribution of hospitalizations, by consistency criteria and variation scores: Maine, 1980-82; Iowa, 1980; Massachusetts, 1982; and California, 1983

Variation in hospital admission rates	Consistency criterion II ¹					Consistency criterion IV ²				
	Total	Maine	Iowa	Mass.	Calif.	Total	Maine	Iowa	Mass.	Calif.
Number of hospitalizations	428,056	242,383	669,059	678,012	2,017,510	428,056	242,383	669,059	678,012	2,017,510
	Percent distribution									
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Low variation	8.2	6.2	8.3	8.4	8.1	6.6	4.9	6.9	6.9	6.6
Moderate variation	1.8	1.7	1.7	2.3	1.9	.9	.7	1.0	1.0	.9
High variation	19.4	16.7	20.2	23.0	20.6	6.5	5.5	6.7	7.3	6.7
Very high variation I	11.9	13.6	12.0	12.4	12.3	2.6	3.3	2.5	2.7	2.7
Very high variation II	19.8	22.2	17.2	14.2	17.3	11.3	13.1	9.4	7.2	9.5
Inconsistent categories	5.7	5.3	8.5	4.6	6.2	38.9	38.2	41.4	39.8	39.9
Residual categories	33.2	34.3	32.1	35.1	33.6	33.2	34.3	32.1	35.1	33.6

¹Decile-ranked variation scores within 2 deciles for at least 3 of the States.

²Decile-ranked variation scores within 2 deciles for all 4 States.

SOURCE: Maine Health Information Center, Augusta, Maine; Iowa Foundation for Medical Care, Des Moines, Iowa; Massachusetts Rate Setting Commission, Boston, Massachusetts; California Health Facilities Commission, Sacramento, California; Uniform hospital discharge data sets.

In the four States, the percentage of admissions judged to be most discretionary (those with the highest variation scores) depends on the consistency criteria applied. As seen in the columns on the left side of Table 4, Iowa had the greatest portion of admissions in the very high II category (22.2 percent), with only 14.2 percent of admissions to California hospitals in this category. However, even in California, 96,278 hospital days (0.142 times 678,012) were consumed in 1983 by patients for whom physicians applied admission standards that were extremely inconsistent in at least three States.

Discussion

The contribution of this article to the burgeoning but separate literatures on small-area variation and DRG's is our examination of the consistency of the pattern of variation in admission rates for categories of patients defined by the DRG system of classification. The observed consistency was used to label the relative discretion of hospital admissions on a five-point scale ranging from low to very high. Low variation, or low discretion, conditions, such as hip fractures, inguinal hernia repair, and major bowel surgery (which is usually done for cancer), are conditions for which the natural variation in rates of occurrence of disease among populations is the principal factor in explaining the variation. For these conditions, patient reluctance to seek medical attention and physician differences in diagnosis and decisionmaking about treatment and/or hospitalization are not important contributors to the variation in hospitalization rates (Wennberg, 1987).

Closer examination of the kinds of admissions that have the highest discretionary scores clearly shows the probable dynamics of the variation process. Of the 13 categories of surgical admissions producing very high discretionary scores, 10 are procedures routinely performed in many hospitals on an outpatient basis (Carter and Ginsburg, 1985). Only inpatient admissions were counted in this data base. In all four States, many, if not most, of these procedures probably continue to be done on an inpatient basis in some hospitals and on an outpatient basis in hospitals serving other areas. Such differing admission patterns would produce wide variation in admission rates across small areas and high SCV scores. Laparoscopic tubal interruption, the cause of surgical admission with the highest variation score, is a treatment for which there are varying hospital policies and surgeon preferences on whether it should occur on an inpatient or outpatient basis. The extremely high variation in rates of dental extractions does not reflect the underlying rate of dental disease but the fact that some oral surgeons routinely use the hospital for the extraction of teeth and other dentists use their offices.

For major operations, in contrast, the component reflecting professional discretion results at least partially from uncertainty about the natural history of the disease and the effects of alternative treatments.

Examples include tonsillectomy for hypertrophy of the tonsils, back surgery, and major cardiovascular operations. Although some randomized trials have been performed, the practice options that physicians select to treat these disorders are not tightly defined. Therefore, the practices of individual doctors can vary substantially based on their own subjective experience.

Variations in admission patterns for minor surgery reflect disagreements on both the value and the preferred place for surgery. Admissions for D&C (dilatation and curettage of the uterus) undoubtedly vary because of discretion in whether a procedure is performed on an inpatient or outpatient basis and in the timing of the procedure (that is, whether a D&C is recommended following a single irregular menstrual period or after several). Moreover, the value of the treatment as an intervention in the natural history of the menopause is in dispute. We plan to use claims from the ambulatory as well as the inpatient setting to distinguish these two components of professional discretion.

Physician discretion also likely determines whether patients with specific medical illnesses are hospitalized or handled on an outpatient basis. Here the uncertainty concerns not the appropriateness of specific treatments—whether, for example, an antibiotic should or should not be used for treating pneumonia—but rather whether patients with a given level of severity of illness need to be hospitalized or can be safely treated at home or in the outpatient department. There is virtually no research on these matters, and medical textbooks contain few guidelines for practicing physicians. The strong correlations between number of beds per 1,000 population and the admission rates for high variation medical conditions indicate that such professional uncertainty has been solved by pragmatic adjustment of local decision rules to the point at which available inpatient beds are consumed (Wennberg, 1987; Wennberg, Freeman, and Culp, 1987). Because such behavior is evident in populations where most inpatient services are provided in teaching hospitals, it is clear that even current academic standards governing admission policies are vulnerable to the influence of available supply (Wennberg, Freeman, and Culp, 1987).

The consistent variation in admission rates across specific categories of admissions in States as diverse as California, Iowa, Massachusetts, and Maine is, from one perspective, quite remarkable. The medical care systems in the Eastern and Western regions of the United States differ markedly in hospital length of stay, admission rates, and overall days of hospital care (Gornick, 1975; Knickman and Foltz, 1985). These consistencies in variation patterns are no less unexpected than those previously reported for several procedures performed in Norway, the United Kingdom, and New England (McPherson et al., 1982). Although it remains to be seen if consistency in international variation patterns for medical conditions will be observed, the consistency of the pattern among these diverse States suggests that the degree of variation of medical admissions, as well as of surgical

procedures, is probably a characteristic of the medical service itself rather than the way health care is organized or financed.

Categorizing admissions by discretionary levels has several important policy implications. The categories can be used to set priorities for research on outcomes. Fewer than 40 causes of medical admission, several of them related, account for most of the differences in per capita costs and utilization among small areas. For example, in a recent study, it was shown that hospital beds and costs are much greater for residents of Boston than of New Haven, even though the two populations are quite similar demographically (Wennberg, Freeman, and Culp, 1987). In 1982, Medicare Part A reimbursements in the Boston area were \$1,894; in the New Haven area, they were \$1,088. Most of the difference in hospital costs was attributable to differences in hospitalization rates for some of the high variation medical conditions identified in this article. These conditions should have priority in any outcomes research agenda.

Similarly, the high variation surgical conditions warrant assessment. Among these, a small number of procedures are conspicuous (Figure 1). Research on outcomes is needed in precisely these areas.

Use of modified DRG categories identified by the Index of Discretionary Admissions can also facilitate utilization review. For example, hospitals interested in rationalizing operations could undertake a review of admissions over the past 3 months. What number of bed days were used by patients admitted with high variation diagnoses? Hospital utilization committees could pull a random sample of charts and assess the degree to which these admissions were necessary. What procedures could have been performed on an outpatient basis? Meetings with medical staff might focus on the general problem of discretionary hospitalizations as well as developing specific guidelines for admission for high variation categories. When such guidelines were developed in Vermont and Pittsburgh for the high variation procedure tonsillectomy, tonsillectomy rates at the two demonstration hospitals fell markedly (Wennberg et al., 1977; Paradise and Bluestone, 1976). People conducting similar research on hysterectomy rates in Saskatchewan, Canada, found that a reduction in admissions for hysterectomy (a high variation procedure) was associated with a drop in the proportion of cases judged "unjustified" by a professional committee (Dyck et al., 1977).

More recently, in the Maine Medical Assessment Program, it has been shown that physicians within a region can respond to information on variations in discretionary admissions by developing an ongoing program to monitor variations and to investigate the outcome significance of differences in practice style. The program has resulted in substantial reduction in rates for common operations, including hysterectomy, tonsillectomy, prostatectomy, and back surgery, and for pediatric hospitalizations (Department of Health Care Review, 1986).

All the methods we have suggested to reduce

admissions of patients in high variation categories have focused on the medical profession and those regulating the profession rather than on the patient. There are two reasons for this.

First, as we highlight in this article, areas of professional uncertainty exist. Some physicians clearly believe strongly in the effectiveness of hospital treatment for a given condition; others, no doubt equally strongly, believe it inappropriate or unnecessary to hospitalize or surgically treat patients with those conditions. The question of who is right in this standoff should be addressed by the profession. For cases in which consensus is difficult to achieve, the profession needs to collaborate in clinical trials, cohort studies, decision analysis, and other forms of research to resolve these critical questions.

A second reason for focusing on the profession, rather than on the patient, as a means of controlling health care costs has been highlighted in the Rand Health Insurance Experiment. In this high-profile project, it has been demonstrated that the programs designed to reduce unnecessary health care use by targeting the consumer-patient, i.e., copayment schemes, work in pernicious ways. Cost sharing resulted in much greater reduction in use of health care services by the poor, especially poor children, than by more economically advantaged groups (Lohr et al., 1986). Perhaps more importantly, individuals facing cost sharing reduced their use of all kinds of medical care, not just that considered unnecessary or inappropriate. Clearly, efforts to ensure that medical care is delivered in the most cost-effective way must focus on the medical profession. The identification of causes of hospital admission that are highly discretionary across widely divergent geographic areas is one step toward helping the profession to focus these efforts.

Finally, DRG groups have some limitations in categorizing variations in hospitalization rates. Because the focus in such studies is on discretion in clinical decisionmaking, the clinical issue must be well described. For medical admissions, we found the DRG system quite useful; with the modified groups used here (Figure 1), we could conveniently categorize at least 70 percent of hospitalized cases using only 40 closely related groups. For surgical operations, the situation was less satisfactory. With some DRG's, congruent identification of clinically meaningful procedures cannot be made. For example, we could not measure prostatectomy rates because of the way DRG's are constructed. For surgical categories, the Professional Activity Study classification convention sometimes provides better descriptions.

Further work will be focused on whether the consistency of variation in admissions patterns observed to date holds when variation scores are obtained from additional States, on the grouping of residual categories, and on analyzing hospital-specific patterns of admission using the Index of Discretionary Admissions. This last analysis will make it possible to determine if there are marked differences across

hospitals with regard to admitting practices for highly discretionary cases.

Technical note

A listing for one State of hospital area discharge rates that result in a low SCV score for the DRG category "specific cerebrovascular disorders" and discharge rates that result in a high SCV score for the DRG category "pediatric gastroenteritis" are shown in Table 5.

Table 5

Person-years and age-adjusted discharge rates for specific cerebrovascular disorders and pediatric gastroenteritis, by hospital service area: Maine, 1980-82

Hospital service area	Person-years	Specific cerebrovascular disorders ¹	Pediatric gastroenteritis ²
		Age-adjusted discharge rates	
All areas	3,264,794	16.3	15.6
Area 1	581,543	12.7	6.9
Area 2	302,173	18.7	13.9
Area 3	300,905	16.7	19.8
Area 4	216,115	11.8	13.2
Area 5	184,141	13.4	36.7
Area 6	163,749	17.1	10.9
Area 7	98,421	14.0	9.1
Area 8	116,234	18.1	11.8
Area 9	96,103	17.0	7.7
Area 10	83,398	17.0	21.9
Area 11	75,956	17.4	12.9
Area 12	74,394	19.8	33.8
Area 13	59,440	14.3	16.3
Area 14	74,093	14.8	7.9
Area 15	41,964	14.5	10.6
Area 16	55,019	27.0	23.5
Area 17	37,068	16.9	24.2
Area 18	51,333	20.6	33.7
Area 19	51,138	19.4	15.0
Area 20	40,841	19.5	9.9
Area 21	45,353	21.4	27.5
Area 22	46,228	21.4	11.8
Area 23	45,154	18.7	28.9
Area 24	40,854	31.1	19.1
Area 25	42,634	18.6	19.8
Area 26	44,609	15.1	5.9
Area 27	39,969	25.5	17.6
Area 28	51,970	13.4	7.5
Area 29	41,828	19.2	11.7
All other	162,167	16.7	17.2
Number of cases	—	5,309	5,090
Systematic component of variation	—	44.3	238.6
Coefficient of variation	—	.239	.524
Ratio of highest to lowest rate	—	2.64	6.21

¹Diagnosis-related group number 14.

²Diagnosis-related group number 184.

SOURCE: Maine Health Information Center, Augusta, Maine: Uniform hospital discharge data set.

References

- Bolande, R. P.: Ritualistic surgery—circumcision and tonsillectomy. *New England Journal of Medicine* 280(11):591-596, 1969.
- Bunker, J. P., McPherson, K., and Henneman, P. L.: Elective hysterectomy. In Bunker, J. P., Barnes, B. A., and Mosteller, F., eds. *Costs, Risks and Benefits of Surgery*. New York. Oxford University Press, 1977.
- Carter, G. M., and Ginsburg, P. B.: *The Medicare Case Mix Index Increase: Medical Practice Changes, Aging, and DRG Creep*. Santa Monica, Calif. The Rand Corporation, 1985.
- Deane, R. T., and Ulene, A.: Hysterectomy or tubal ligation for sterilization: A cost-effectiveness analysis. *Inquiry* 14(1):73-86, 1977.
- Department of Health Care Review: *Confronting Regional Variations: The Maine Approach*. Chicago. The American Medical Association, 1986.
- Dyck, F. J., Murphy, F. A., Murphy, J. K., et al.: Effect of surveillance on the number of hysterectomies in the province of Saskatchewan. *New England Journal of Medicine* 296(23):1326-1328, 1977.
- Fetter, R. B., Youngsoo, S., Freeman, J. L., et al.: Case mix definition by diagnosis-related groups. *Medical Care* 18(2)Suppl.:1-53, 1980.
- Gornick, M.: *Medicare Patients: Regional Differences in Length of Hospital Stays, 1969-71*. DHEW Pub. No. (SSA) 75-11700. Social Security Administration. Washington. U.S. Government Printing Office, July 1975.
- Harper, P.: *Preventive Pediatrics: Child Health and Development*. New York. Appleton-Century-Crofts, 1962.
- Knickman, J. R., and Foltz, A. M.: A statistical analysis of reasons for East-West differences in hospital use. *Inquiry* 22(1):45-58, 1985.
- Lohr, K. N., Brook, R. H., Kamberg, C. J., et al.: Use of medical care in the Rand Health Insurance Experiment: Diagnosis- and service-specific analyses in a randomized controlled trial. *Medical Care* 24(9)Suppl.:S1-S87, 1986.
- McCracken, S., Latessa, P., and Wennberg, J. E.: *A Study of Hospital Utilization in Iowa in 1980*. Des Moines, Ia. Servi-Share of Iowa, 1982.
- McPherson, K., Strong, P. M., Epstein, A., et al.: Regional variations in the use of common surgical procedures: Within and between England and Wales, Canada and the United States of America. *Social Science and Medicine* 15(3):273-288, 1981.
- McPherson, K., Wennberg, J. E., Hovind, O. B., et al.: Small-area variations in the use of common surgical procedures: An international comparison of New England, England and Norway. *New England Journal of Medicine* 307(21):1310-1314, 1982.
- Paradise, J. A., and Bluestone, C. D.: Toward rational indications for tonsil and adenoid surgery. *Hospital Practice* 11(2):79-87, 1976.
- Roos, N. P., and Roos, L. L.: Surgical rate variations: Do they reflect the health or socio-economic characteristics of the population? *Medical Care* 20:945-958, 1982.

Roos, N. P., Flowerdew, G., Wajda, A., et al.: Variations in physicians' hospitalization practices: A population-based study in Manitoba, Canada. *American Journal of Public Health* 76(1):45-51, 1986.

Wennberg, J. E.: Population illness rates do not explain population hospitalization rates. *Medical Care* 25(4):354-359, 1987.

Wennberg, J. E., and Gittelsohn, A.: Variations in medical care among small areas. *Scientific American* 246(4):120-134, 1982.

Wennberg, J. E., Barnes, B. A., and Zubkoff, M.: Professional uncertainty and the problem of supplier induced demand. *Social Science and Medicine* 16(7):811-824, 1982.

Wennberg, J. E., Blowers, L., Parker, R., et al.: Changes in tonsillectomy rates associated with feedback and review. *Pediatrics* 59(6):821-826, 1977.

Wennberg, J. E., Freeman, J. L., and Culp, W. J.: Are hospital services rationed in New Haven or over-utilized in Boston? *Lancet* 1(8543):1185-1188, 1987.

Wennberg, J. E., McPherson, K., and Caper, P.: Will payment based on diagnosis-related groups control hospital costs? *New England Journal of Medicine* 311(5):295-299, 1984.