THE FINANCIAL IMPACT OF PHYSICIAN PRACTICE STYLE ON HOSPITAL RESOURCE USE

JOE FEINGLASS, Ph.D.
Research Assistant Professor
Northwestern University Medical School, Chicago Il. and
Research Associate,
Center for Health Services and Policy Research,
Northwestern University, Evanston, Il.

GARY J. MARTIN, M.D.
Section Chief, Section of General Internal Medicine, and
Assistant Professor of Medicine,
Northwestern University Medical School
Chicago, Il.

ASHISH SEN, Ph.D.
Professor, School of Urban Planning and Policy,
University of Illinois at Chicago
Chicago, Il.

Address communications to:

JOE FEINGLASS, Ph.D.
Northwestern Memorial Hospital
215 E. Chicago Ave. # 310
Chicago, Il. 60611
w 312-908-8620

ACKNOWLEDGEMENTS

Research for this article was supported by The Pew Charitable Trusts, Grant # 84-1045, "Impact of an Educational Program for House Officers on Containment of Hospital Charges," by Veterans Administration Health Services Research and Development Grant IR # 85-015, "Impact of an Educational Program on House Officers' Generated Costs", and by Northwestern Memorial Hospital, Chicago, Il. The authors would like to thank our two anonymous reviewers for their thoughtful criticisms of earlier drafts of this paper.

THE FINANCIAL IMPACT OF PHYSICIAN PRACTICE STYLE ON HOSPITAL RESOURCE USE

BY

JOE FEINGLASS, Ph.D.
Research Assistant Professor of Medicine
Northwestern University Medical School, Chicago Il. and
Research Associate,
Center for Health Services and Policy Research,
Northwestern University, Evanston, Il.

GARY J. MARTIN, M.D.
Section Chief, Section of General Internal Medicine,
Assistant Professor of Medicine,
Northwestern University Medical School
Chicago, Il.

ASHISH SEN, Ph.D.
Professor, School of Urban Planning and Policy
University of Illinois at Chicago
Chicago, Il.

ABSTRACT

Several specifications of a statistical model were used to measure the impact of internal medicine attending physicians on inpatient charges and length of stay at a large, urban teaching hospital. The study was based on a sample of 1458 patients discharged during 1985-87 with 12 common principal diagnosis clusters. The relationship between 31 physicians' clinical decisions and hospital charges and length of stay was analyzed controlling for patients' health status, as measured by demographic characteristics, diagnostic group, and ratings for the Severity of Illness Index (SOII).

Results indicated that attending physicians were statistically significant predictors of the log of total charges ($p = .0030$) and the log of length of stay ($p < .0001$), although not as significant predictors of untransformed total charges (p =

.1255). Equivalent results were obtained when overall SOII ratings were replaced by SOII subscale ratings for the presenting stage of the principal diagnosis on admission. Examination of individual physician regression coefficients revealed that physicians varied within a 40 percent range of generated per patient charges. No significant differences in mortality, early readmissions or residual impairment on discharge were found between the 10 highest and 10 lowest resource use physicians. The conservatively estimated range of attending physician practice variations observed in this study has serious financial implications for hospitals operating under incentives to minimize operating costs, particularly for teaching hospitals facing reductions in subsidies for graduate medical education.

Physician practice style differences necessarily increase in importance when uncertainty and lack of consensus about appropriate treatment exist. The medical and financial aspects of practice style differences have been the subject of a great deal of recent research. Numerous studies, recently reviewed by Eisenberg (1986), have found that physicians' age, medical specialty, training, and personal characteristics (e.g. values, tastes, and preferences for certain types of clinical decision making), interact with practice setting and financial incentives in influencing both the technical and interpersonal dimensions of care. The medical cost implications of practice style differences are highlighted in recent reviews of literature on geographic variations in the per capita consumption of medical care (Paul Shaheen et al., 1987) and inappropriate hospital utilization (Payne, 1987).

Some investigators believe practice style differences, which evolve into collegial community or institutional standards, are the primary sources of medical practice variations (Wennberg, 1984, 1986). Other studies have reported a more ambiguous relationship between the service intensity of care provided by individual physicians and small area use rates (Stano, 1986; Stano and Folland, 1988). One estimate of the magnitude of physicians' direct impact on hospital costs was revealed in a study controlled for disease stage. It found that physicians accounted for an average of 17.5% of intra-Diagnosis Related Group (DRG) variation in length of stay in an outlier-trimmed sample of high volume DRGs from a teaching hospital, and an average of 29.4% of the intra-DRG variation in length of stay in three other non-teaching hospital samples (McMahon and Newbold, 1986).

Current hospital cost containment initiatives, such as prospective payment, capitation and preferred provider contracts, seek to place providers at risk for the cost of medical care. Because variable hospital costs are largely determined by physician's diagnostic and treatment orders, hospital administrators and third party payors are seeking new methods to reduce costs by altering physician practice styles (Kralewski, 1987; Glandon and Morrisey, 1986). Profiles of physician resource use are now frequently used to identify individual physicians' medical practice patterns.

Created by new management information systems which combine clinical and financial data, physician practice profiles based on Medicare DRGs were supposed to draw attention to variations in physicians' practice style. However, because of wide intra-DRG variation in resource use and the failure of DRGs to incorporate a measure of severity of illness, practice profiles based on DRGs alone have proven inadequate to evaluate differential expenditures (or treatment plans) for patients classified into the same DRG (Hsiao et al., 1986; Horn et al., 1986b).

This study illustrates the financial implications of severity adjusted practice style variations among internal medicine attending physicians at an urban, 700 bed teaching hospital. Despite the fact that the patient population is largely composed of attending physicians' private patients, cost containment educational programs directed at residents had previously produced substantial reductions in charges and length of stay. Yet attending physicians at the study hospital ultimately control (and are legally responsible for) each patients' care. Because most internal medicine attending physicians at the teaching hospital either directed their patients' diagnostic and treatment plans or approved (personally or by phone) residents' plans, it was of interest to determine the cost containment potential of changing attending physician behavior.

The attending physicians studied here know each other and frequently discuss the state of medical knowledge in an academic setting. They share common, state of the art hospital facilities and supervise the same resident teams. The results presented here thus provide a conservative reflection of the extent of practice style variations which would likely exist in a larger universe of community hospital patients and medical specialties.

METHODS

Sample Selection

An initial sample of 4100 discharges was chosen to represent patients with clinical diagnoses most frequently treated by Department of Medicine physicians. Seventy common ICD-9-CM principal diagnoses were grouped into twelve diagnostic clusters designed to be more medically meaningful and etiologically consistent than their associated DRG classifications. These diagnoses include acute myocardial infarction, congestive heart failure, ischemic heart disease, chronic obstructive pulmonary disease, lung cancer, diabetes, asthma, hypertension, gastroenteritis and gastrointestinal hemorrhage. (Acute myocardial infarction patients were divided into two diagnostic groups, one group being those patients admitted to the Coronary Care Unit and the other group being those patients treated exclusively on the medicine floor). The twelve diagnostic clusters studied here fall within the 40 acute and chronic illnesses which account for about 70% of all medical hospital admissions nationally. Over 90% of all listed records with an appropriate ICD-9-CM principal diagnosis were retrieved and abstracted. All patients were discharged between July, 1985 and June, 1987.

Some attending physician cross-coverage exists for patients in the study sample. "Primary" attending physicians were identified for each patient from discharge abstract signatures and review of admitting and transfer notes. Research project medical records coders reviewed all ICD-9-CM principal diagnoses and physician signatures and deleted inappropriately coded records from the study. Because of the relative value unit accounting techniques used at the study hospital, total fixed and variable accounting costs, length of stay, and total charges have been found to be highly correlated in the .85 to .95 range. Only very small (5%) price increases in selected charge codes occurred during the 1985-87 study period. Total charges and length of stay were therefore used as a reasonable proxy for hospital resource use.

From the initial sample of 4100 discharges based on ICD-9-CM codes, a final sample of 1458 teaching hospital discharges managed by 31 internal medicine attending physicians was selected from hospital files on the basis of four further criteria:

1.) Patients with surgical procedures or operating room charges over $150 were deleted. This was done to reduce the effect on charges of multiple attending physicians, including surgeons not in the study sample. Care for the remaining patients was the responsibility of the attending physicians in the sample and the residents directly under their supervision.

2.) Diagnostic cardiac catheterization admissions were deleted to reduce the impact of specialized cardiology procedures on the overall sample. These patients may have been concentrated among a few attending physicians (thus potentially distorting the distribution properties of the overall sample). Two highly unusual cases with lengths of stay in the 70-80 day range were also deleted, as was one stay with under $750 in total charges.

3.) To further reduce potential subspecialty referral pattern distortions related to the "nesting" of certain diagnoses among certain physicians, each physician in the sample was required to have treated patients in at least seven separate clinical diagnoses or patients in at least three separate Major Diagnostic Categories (MDCs). This rule assures that the physician sample better represents attending physicians with a general medical practice, rather than physicians who exclusively practice subspecialty medicine. While several physicians do primarily practice as subspecialists, only three physicians who otherwise had enough patients in the sample were deleted by this rule. The average number of diagnoses per physician was 9.5; 23 of 31 physicians treated patients in all four Major Diagnostic Categories (circulatory, respiratory, digestive, metabolic). The influence of physician subspecialty was separately tested as a predictor of resource use with indicator variables in the regression model described below.

4.) Each attending physician was required to have treated 15 or more patients in each of the two academic years represented in the sample (mean = 47). This rule assures that the sample represents teaching hospital internists with a significant hospital practice. The 31 physicians analyzed here are among the hospital's heaviest admitters.

The number of patients in each principal diagnosis cluster is presented in Table 1; sample descriptive statistics are presented in Table 2. When compared to charges and length of stay for other Department of Medicine patients discharged in the same DRGs, the sample for the present study was found to be a conservative reflection of the resource use variance in the teaching hospital's overall patient population (Feinglass et al., 1988). 27 admissions were Health Care Financing Administration (HCFA) defined high length of stay or high charges DRG outliers.

Severity of Illness

The Severity of Illness Index (SOII), developed over six years at Johns Hopkins University by Horn and associates (Horn et al., 1984; 1986a; Horn and Horn, 1986), attempts to measure patient-related sources of variations in resource use. The SOII defines an ascending one to four scale (with four often a proxy for death) for seven rated dimensions of each inpatient hospital stay: (1) the presenting stage of the principal diagnosis; (2) complications of the principal condition; (3) concurrent interacting illnesses; (4) dependency on hospital staff; (5) the extent of non-operating room procedures; (6) rate of response to therapy; and (7) residual impairment on discharge (related to the acute aspect of hospitalization). Project SOII raters were trained by consultants from Horn's staff, and received four quarterly inter-rater reliability checks. Each check involved a blind review by expert consultants of 25 randomly selected cases for each rater. The average overall disagreement rate for the four checks was under 7%.

The SOII has been criticized for potential circularity in predicting resource use from ratings based, in part, on the level of medical care intervention (Richards et al., 1988). To the extent this criticism is valid, using SOII ratings as a covariate in this analysis will tend to underestimate the true magnitude of physicians' impact on resource use. For instance, avoidable iatrogenic complications, excessive physician test ordering or unnecessary therapy might increase a patients' severity level, and therefore the expected level of resource use for that patient. For this reason, the impact of physicians on resource use was also tested in an alternative model, in which overall SOII ratings are replaced by SOII subscale ratings for the presenting stage of the principal diagnosis. This permits comparisons of physician effects between overall, or "peak" severity during each stay and severity of illness on admission.

The Model of Physicians' Impact on Resource Use

The general linear model of physicians' impact on resource use relies on four distinct (sets of) independent variables: patients' clinical diagnosis cluster; patients' rated severity of illness; patients' demographic status represented by age, sex, and race; and the presence of specific attending physicians who are deemed responsible for the economic resources (estimated by hospital charges and length of stay) generated by physician orders. The model of inpatient resource use per case is expressed as:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \tau_{ij} + \epsilon_{ij}$$

where

Y_{ij} = inpatient resource use (per diem plus ancillary charges) per case for the j th patient treated by the i th physician

μ = a constant, "baseline" resource use per case

α_i = the contribution of attending physician i to resource use per case

β_j = the contribution of the underlying diagnosis, severity of illness, and demographic characteristics of patient j to resource use per case

τ_{ij} = the contribution of the interaction effects of physician i and the diagnosis, illness severity, and demographic characteristics of patient j to resource use per case

ϵ_{ij} = assumed to be independently distributed random errors

All the independent variables are expressed as dummy indicator variables, set equal to one when present and zero when absent. Age (like race and sex) is expressed as a dichotomous variable where age > 69 = 1 and age < 70 = 0. (About one-half the sample was older than 70). Attending physicians are also entered as dummy, indicator variables set equal to one when a particular attending physician manages a patient and zero when absent. The constant term yields the predicted resource use when all indicator variables equal zero, i.e. for the omitted diagnosis, severity level, attending physician and demographic variables "left out" of the equation. Because of the very large number of potential interaction effects between each attending physician and each diagnosis and severity level, the model presented here assumes interaction effects are negligible. The lack of interaction terms will increase the standard error of physician coefficients, diluting the true significance of observed physician effects.

The diagnosis, severity and demographic variables are covariates expressing the expected burden of illness brought to the hospital by each patient, regardless of subsequent clinical decisions about hospital tests, procedures and length of stay. The test statistics associated with the presence of the physician variables are used to estimate the impact of clinical decisions on resource use. The resulting physician regression coefficients rank individual attending physicians in terms of their relative resource-generating behavior. Because it is possible that physicians who generate higher than average resource use are disproportionately subspecialists, the influence of physician subspecialty was tested by inclusion of a dummy, indicator variable for the 19 board certified subspecialist physicians in the sample.

The magnitude of each physician's regression coefficient is determined by which physician in the sample was omitted. For this reason, physicians were initially ranked according to their relative "costliness". The initial rankings were computed by summing each physician's actual charges for each patient treated and then subtracting the sum of the across-sample diagnosis-severity mean charges for each patient treated. The attending physician with the smallest deviation (-$35) from his patients' across-sample diagnosis-severity mean charges was selected to be "left out". The values of the remaining physician regression coefficients thus roughly express the distance of each physician from the physician whose generated charges initially appeared closest to the sample mean for his patients.

Specifying the Model Form

The log transformation of the dependent charges and length of stay variables deserves some discussion. Initially no transformation was made. However, during routine regression diagnostics examining the appropriateness of the model, extreme heteroscedasticity was discovered. Heteroscedasticity, or unequal variance of the dependent variable, violates one of the key assumptions of least squares theory. Its effect is two-fold: it causes variance of estimates to rise and the estimates of the variances to be wrong. Since the level of heteroscedasticity found was very high, very adverse effects on significance testing were expected. Some corrective action was necessary.

There are essentially three courses of action to handle such situations. One is the use of a variance stabilizing transformation. Another is to weight the observations. Both of these alternatives could remove the heteroscedasticity or, at least, would render it low enough to be bearable. A third alternative is to accept the heteroscedascticity, and its attendant increase in variance, but use an estimator for variances (White, 1980). This would make at least the variance estimates reasonably good. The decision was made to use a logarithmic transformation, which turned out to be variance stabilizing.

There were several intuitively appealing reasons for this decision. If the dependent resource use variables are not transformed (whether the model was weighted or not), the implications of positive physician coeffiecients in the model are that the increase in charges would be the same no matter what the overall charge was. For example, there would be an expected $50 increase in a $500 hospital charge as well as a $50,000 hospital charge. On the other hand, a log transformation would imply that the increase would be by a percentage of the overall charge. Because the latter seems so much more plausible, results from the log transformed model are emphasized. Results from untransformed models are also presented to indicate the sensitivity of the underlying assumptions, particularly with respect to physician coefficients and test statistics.

The log-linear model hypothesizes that physicians differ from one another by percentages rather than in constant dollars. Coefficients in a log-linear model will be exponents of e, interpretable as percentage differences from the omitted indicator variables. The log-linear regression coefficients for each physician indicator variable thus reflect each physicians' resource use ranking (the percentage difference in resource use from the omitted physician). Like all other variables in the log-linear model, physician variable coefficients are exponents of e where the highest positive coefficients correspond to physicians with the highest percentage dollar difference from the physician set equal to zero when present. Physician test statistics and individual rankings are compared across logged and untransformed models based on both overall and admission severity.

Inpatient mortality, the incidence of early readmissions and SOII subscale ratings for residual impairment on discharge (unrelated to patients' pre-existing impairments or expected condition on discharge) are also used to determine if quality of care differences exist between high and low resource use physicians. Examining residual impairment ratings is especially important to determine if resource savings are related to low resource use physicians' willingness to risk early readmission by discharging patients "quicker and sicker." Ordinary least squares, run on SPSS PC+ software, was used to determine the strength and significance of each (set) of the independent variables and to test the impact of physicians on resource use.

RESULTS

Physicians' Impact on Resource Use

Table 3 presents partial F-test probabilities for each (set) of the independent variables, given the presence of all other variables in each equation. Results based on overall severity and admission severity are presented seperately. As expected, the indicator variables for clinical diagnosis cluster and severity of illness were highly significant predictors of resource use in all equations, demonstrating the importance of severity adjustments to patients' diagnostic classifications. The demographic variables were weaker, generally non-significant predictors of resource use. Physicians were significant predictors of the log of total charges ($p = .0030$) and the log of length of stay ($p < .0001$). Results from the untransformed model indicate that attending physicians were not as highly predictive of actual total charges (p=.1255), although they remained significant predictors of actual length of stay (p=.0004).

When results from the untransformed model of actual total charges were disaggregated, physicians were statistically significant predictors of routine room charges (p=.0087) and lab charges (p=.0417). (Mean routine room and lab charges account for over 55% of mean total charges in the sample; routine room charges do not include intensive care days). Physicians were also significant predictors of supply charges (p=.0425) and x-rays (p=.0230), while being close to statistically significant predictors of overall drug charges (p=.0596) and overall therapy charges (p=.0902), including respiratory therapy (p=.0752). When the 27 high length of stay or high charges outlier admissions were deleted, physicians were close (p=.0533) to significant predictors of actual total charges in the untransformed regression for the remaining 1433 non-outlier admissions in the sample.

As anticipated, the models using overall, or "peak" SOII ratings explained more variance in the dependent resource use variables than models based on admission severity (R^2=.5394 in the overall severity model of the log of total charges; R^2=.3253 in the admission severity model of the log of total charges). Age greater than 69 becomes more highly significant in the admission severity models, a reflection of how increasing age is at least partly accounted for in higher overall severity ratings.

Physicians become even more significant in the models based on admission severity (p=.0001 for the log of total charges and p < .0001 for the log of length of stay; p=.0768 for untransformed actual total charges and p=.0001 for untransformed actual length of stay). Findings across all model specifications thus tend to confirm the existence of an important attending physician practice style effect on hospital resource use for patients in this sample.

How High and Low Resource Use Physicians Differ

Individual physician regression coefficients reflect each physicians' expected contribution to total charges or length of stay, controlled for their patients' diagnoses, illness severity, and demographic characteristics. Physician rankings were remarkably stable across each specification of the models in Table 3. Physician rankings were very similar in models based on overall severity and admission severity. For instance, the Spearman rank order correlation was .80 ($p < .0001$) between the 31 paired coefficients derived from the admission severity and overall severity log linear models of total charges. Physician rankings were almost identical in the log linear and untransformed total charges equations using overall severity (Spearman $r = .91$, $p < .0001$). Physician subspecialty was not significantly associated with resource use in any of the models; signs on the subspecialty indicator variables consistently indicated lower resource use than generalists.

Table 4 presents the ten highest, the middle eleven and the ten lowest physician coefficients and standard errors, ranked from lowest to highest, derived from the overall severity models predicting both logged and untransformed total charges. The log linear coefficient values can be interpretted as approximate percentage differences from the omitted physician, ID # 39, who was initially selected because the sum of his patients' actual charges was closest to the sum of the expected sample diagnosis-severity mean charges for his patients. (Which physician is omitted does not affect either the rank order or the overall range of differences). The untransformed coefficients in Table 4 reflect unit (actual dollar) differences from the omitted physician.

The full set of physician resource use rankings presented in Table 4 offers some insight into the magnitude of individual physician resource use differences. Percentage differences for total charges, from the highest to lowest resource use physician (based on coefficients from the log linear model of the log of total charges), have a 40% range. Physician length of stay coefficients in the the log linear model ranged from 40% below to 17% above the omitted physician, a 57% range.

The case mix frequencies displayed in Table 5 indicate how the log linear model distinguishes high and low resource use physicians. The distribution of patients in each MDC between the ten highest and ten lowest resource use physicians is quite similar, with lower resource use physicians tending to treat a relatively larger number of patients with respiratory disease.

Turning to the distribution of the SOII ratings displayed in Table 5, it would appear that high resource use physicians are treating a somewhat more severely ill patient population, despite the lack of statistically significant differences in either overall or admission severity. While 39.8% of the 487 patients treated by low resource use physicians were rated as overall severity level one, only 33.3% of the 418 patients treated by high resource use physicians were rated as level one. Similarly, 36.6% of patients treated by high resource use physicians were rated as level 3 for admission severity as opposed to 32.2% for low resource use physicians. These results initially appear to somewhat contradict the assumptions of the statistical model, which by entering overall severity of illness indicator variables, controls for such case mix or referral pattern differences.

However, closer inspection of evidence on differences between physicians for patient charges <u>within</u> overall severity levels suggests that important differences exist between high and low resource use physicians across the full spectrum of severity. Intra-severity level data are displayed in Table 6. Table 6 and Figure 1 present the actual mean charges for all patients treated by the ten highest and ten lowest resource use physicians identified in Table 4. The right-hand columns of Table 6 present separate statistics for the actual mean charges for patients rated as overall SOII levels one and two. Looking only at patients in SOII level one, high resource use physicians generated $646 more per patient than low resource use physicians. The per patient difference within SOII level two was $1345. (The difference in mean total charges within SOII level three was $3374). Differences in charges appear to be uniformly spread across almost all utilization categories. These intra-severity level data confirm the validity of the across-severity level case mix adjustments which originally produced the high and low resource use rankings from physicians' log-linear regression coefficients.

High and low resource use physicians do not appear to differ in the quality of care provided. 15 of 418 patients (3.6%) treated by high resource use physicians died in the hospital; 18 of 487 patients (3.7%) treated by low resource use physicians died in the hospital. (Deaths accounted for exactly one-half of all patients rated as SOII level four in both groups). There was virtually no difference in the two groups' incidence of readmissions at 10 or 30 days; 19 of the 418 patients (4.5%) managed by high resource use physicians were readmitted within 30 days of a previous discharge as compared to 21 of the 487 patients (4.3%) managed by low resource use physicians. A chi-square test of association between high and low resource use physician groups and patients' SOII subscale rating for residual impairment on discharge was non-significant (Chi square=.526, p=.9130). 37 of the 418 patients (8.4%) treated by high resource use physicians had residual impairment on discharge ratings of either three or four; 43 of the 487 patients treated by low resource use physicians (8.8%) had residual impairment on discharge ratings of either three or four. These findings indicate that despite a 1.75 day mean per patient difference in length of stay, low resource use physicians were not discharging patients "quicker and sicker."

DISCUSSION

The patients and the attending physicians analyzed in this study both represent relatively homogenous populations. The patients were selected from relatively common diagnostic categories; their signs and symptoms are in the "mainstream" of hospital care, and their diseases pose frequently seen and discussed medical management issues. Their care did not involve surgery or other specialty services. Because they did not experience interdepartmental transfers, their care (including care in the intensive care units) was largely managed by a single attending physician. The attending physicians studied here all practice at a prestigious teaching hospital and have academic appointments; as internists with a large hospital practice they meet and interact about medical issues frequently, often know each other personally, and have access to the same facilities and medical technologies. The ordering decisions of residents, who rotate randomly across patients treated on the medical service, would tend to "wash out" some of the directly observable effect of attending physicians in this model. It is thus of considerable general interest that these attending physicians were nevertheless found to have a significant impact on hospital resource use that is controlled for their specific referral patterns and is independent of residents' orders.

The most important difference between physicians appears to be in discretionary decisions on how long to keep patients in the hospital. The teaching hospital in this study has an aggressive, well-staffed utilization review program, and is subject to frequent Professional Review Organization (PRO) and third party payor utilization audits. Nevertheless, a previous utilization review coordinator, interviewed for this study, indicated her impression that at least 10% of hospital days were still inappropriate in 1985-87, and that most of those days were due to "traditional attending physician practice patterns." Medical staff interviews revealed wide disparities in the ability of some patients and some physicians to fully utilize outpatient preadmission or followup diagnostic facilities, as well as the influence of different third party insurance coverage options on decisions about length of stay.

Other differences in the service intensity of physicians' orders, such as those involving use of laboratory tests, may be related to differences in the supervison of resident orders, in teaching philosophy, or a result of distinctive practice patterns developed in response to the specific referral patterns of a subspecialist in a group practice. Some of the ten high resource use physicians profiled in Table 4 may be more accustomed to consulting for or managing more severely ill, hospitalized patients, despite the relative similarity of SOII ratings between patients treated by high and low resource use physicians. Given an extensive hospital-based, tertiary care practice experience, higher resource use physicians may be "overreacting" to their less severely ill patients by employing more intensive workups requiring longer lengths of stay. Conversely, the lower resource use physicians may be more adept at shifting care to the outpatient setting, tolerating diagnostic ambiguity, or initiating early discharge planning. These differences reflect physicians' varying test/treatment thresholds, or systematic biases about the costs, risks, and benefits of medical services, and the corresponding propensity to order services in the face of ambiguous or conflicting concerns about patient care (Pauker and Kassirer, 1980).

These findings are similar to other studies based on hospitalization rates, outpatient utilization data, and clinical simulations, which have found that high resource use physicians generate a markedly disproportionate share of clinical costs (Linn et al., 1984; White et al., 1984; Roos et al., 1988). Perhaps because of the similar tertiary care setting, the results on the effects of internal medicine subspecialty training reported here are quite similar to those recently reported by Bernard et al. (1990). The literature on the role of medical specialty in determining hospital resource use for diagnostically similar patients remains contradictory (Eisenberg, 1986; Cherkin et al., 1987; Franks and Dickenson, 1986; Strauss et al., 1986; Garg et al., 1979). Although not significantly different, subspecialists were associated with lower resource use in the diagnosis-severity controlled models analyzed here.

The potential impact of inter-physician differences is suggested by the following arithmetic. The 1985-87 mean charge for all patients in this sample was about $6600. A crude estimate of the per-patient dollar differences between the highest resource use physician (#27, 17.4% above the omitted physician) and the lowest resource use physician (#34, 22.7% below the omitted physician) would therefore be over $2600. Given the high variation in total charges and the large standard errors around the physician parameter estimates, a very conservative estimate of potential savings from reducing practice style variation is suggested by the rankings in Table 4, which divide the physicians studied into three about equal groups (on the basis of their rankings in the log-linear model of total charges). The low resource use group average (-.133) was about 13% below the medium resource use group average (-.001) and the high resource use group average (.094) was about 9% above the medium resource use group. An estimate of the per-patient savings if the ten high resource use physicians were to change their practice styles to conform to the average total charges of the medium resource use group is over $600. (This difference is understated insofar as differences in physician fees are not included in hospital charges). If we note that the ten high resource use physicians treated 418 patients in the sample, the resulting average difference in charges is over $250,000, <u>just</u>

for the patients in this sample. This figure is about 3.2% of the $7.6 million in charges in the sample. Although such savings would impact marginal, rather than average resource use, the real financial implication of practice style differences is illustrated by the fact that hospital operating margins dropped from 3.6% in 1985 to 3.0% in 1986, and are continuing to decline nationally (Hospitals, 1987).

Despite the fact that some amount of practice style variation is both inevitable and desirable, given different patient preferences and the state of medical knowledge, practice variations may nevertheless pose a serious challenge to hospitals competing in an environment of growing incentives to minimize operating costs. The relatively less efficient practice style of residents, or the team practice of medicine characteristic of academic medical centers, has been implicated in studies documenting the higher costs of teaching hospitals (Frick et al., 1985; Cameron, 1985; Jones, 1985). Studies comparing faculty and community physicians within the same hospital have also found that faculty attending physicians were responsible for significantly higher costs (Garber et al., 1984; Jones, 1984). Given recent attempts to cut reimbursement for the indirect costs of medical education, reducing costs generated by high resource use physicians could be an important future component of teaching hospital cost containment efforts.

The aging of the U.S. population will make large increases in medical care expenditures inevitable (Schneider et al., 1983).

Advocates of new rationing mechanisms have pointed out that increasingly difficult medical resource allocation decisions are upon us. In an editorial in the Journal of the American Medical Association, one of the foremost investigators of medical practice variations discussed the implications of this reality:

> The nation's growing demand for improved quality, efficiency, and equity in its health care system is thus hostage to unresolved theories about correct practice. The unsettled nature of contemporary medical opinion on correct practice, the high prevalence of the underlying conditions that reasonably fit theory, and the steady growth in the numbers of specialists trained in invasive technologies merge to ensure the continued increase in the per capita cost of care. (Wennberg, 1987).

By identifying significant hospital resource use variations among academically distinguished internists in a single teaching hospital, this study has provided a very conservative illustration of the economic magnitude of practice style variations. The relatively primitive nature of the severity and outcome adjustments now available points towards the unresolved methodological issues facing those who would seek to restrict medical care expeditures through more restrictive clinical policies.

REFERENCES

Bernard, A. M., Shapiro, L.R., and McMahon L.F. "The Influence of Attending Physician Subspecialization on Hospital Length of Stay." Medical Care, 28 (1990):170-174.

Cameron, J.M. "Indirect Costs of Graduate Medical Education." New England Journal of Medicine 312 (1985):1233-1238.

Chassin, M.R., Brook, R.H., Park, R.E., Keesey, J., Fink, A., Kosecoff, J., Kahn, K., Merrick, N., and Solomon, D.H. "Variation in the Use of Medical and Surgical Services by the Medicare Population." The New England Journal of Medicine 314 (1986):285-290.

Cherkin, D.C., Rosenblatt, R.A., Hart, L.G., Schneeweiss, R., and Logerfo, J. "The Use of Medical Resources by Residency-Trained Family Physicians and General Internists." Medical Care 25 (1987):455-469.

Eisenberg, J.M. Doctors' Decisions and the Cost of Medical Care: the Reasons for Doctors' Practice Patterns and Ways to Change Them. Ann Arbor, Michigan: Health Administration Press Perspectives, 1986.

Feinglass, J., Scherubel, J., Swartz, J. "Comparing How Two Measures of Illness Severity Predict Hospital Costs."

Journal of the American Medical Records Association 59, no. 7 (1988):34-38.

Franks, P. and Dickinson, J.C. "Comparisons of Family Physicians and Internists: Process and Outcome in Adult Patients at a Community Hospital." *Medical Care* 24 (1986):941-948.

Frick, A.P., Martin, S.G., and Shwartz, M. "Case Mix and Cost Differences between Teaching and Non-teaching Hospitals." *Medical Care* 23 (1985):283-295.

Garber, A.M., Fuchs, V.R., and Silverman, J.F. "Case Mix, Costs and Outcomes: Differences between Faculty and Community Services in a University Hospital." *The New England Journal of Medicine* 310 (1984):1231-1237.

Garg, M.L., Mulligan, J.L., Gliebe, W.A., and Parekh, R.R. "Physician Specialty, Quality and Cost of Inpatient Care." *Social Science and Medicine* 13C (1979):187-190.

Glandon, G.L. and Morrisey, M.A. "Redefining the Hospital-Physician Relationship under Prospective Payment." *Inquiry* 23 (1986): 166-175.

Gonnella, J.S., Hornbrook, M.C., and Louis, D.Z. "Staging of Disease: a Case Mix Measurement." Journal of the American Medical Association 251 (1984):637-644.

Horn, S.D., Horn, R.A., and Sharkey, P.D. "The Severity of Illness Index as a Severity Adjustment to Diagnosis-Related Groups." Health Care Financing Review Annual Supplement (1984):33-45.

Horn, S.D., Horn, R.A. "Reliability and Validity of the Severity of Illness Index." Medical Care 24 (1986):9-178.

Horn, S.D., Horn, R.A., Sharkey, P.D., and Chambers, A.F. "Severity of Illness within DRGs: Homogeneity Study." Medical Care 24 (1986a):225-235.

Horn, S.D., Horn, R.A., and Moses, H. "Profiles of Physician Practice and Patient Severity of Illness." The American Journal of Public Health 76 (1986b):532-535.

Hospitals, "Finance," October 20, 1987, p. 2.

Hsiao, W.C., Sapolsky, H.M., Dunn, D.L., and Weiner, S.L. "Lessons of the New Jersey DRG Payment System." Health Affairs 5, no. 2 (1986):32-45.

Jones, K.R. "The Influence of the Attending Physician on Indirect Graduate Medical Education Costs." Journal of Medical Education 59 (1984):789-798.

Jones, K.R. "Predicting Hospital Charge and Stay Variation: the Role of Teaching Status, Controlling for Diagnosis-Related Group, Demographic Characteristics, and Severity of Illness." Medical Care 23 (1985):220-235.

Kralewski, J.E., Dowd, B., Feldman, R., and Shapiro, J. "The Physician Rebellion." The New England Journal of Medicine 316 (1987):339-341.

Linn, L.S., Yager, J., Leake, B.D., Gastaldo, G., and Palkowski, C. "Differences in the Numbers and Costs of Tests Ordered by Internists, Family Physicians, and Psychiatrists." Inquiry 21 (1984):266-275.

McMahon, L.F. and Newbold, R. "Variation in Resource Use within Diagnosis-Related Groups: the Effect of Severity of Illness and Physician Practice." Medical Care 24 (1986):388-397.

Pauker, S.G. and Kassirer, J.P. "The Threshold Approach to Clinical Decision Making." The New England Journal of Medicine 302 (1980):1109-1117.

Paul Shaheen, P., Clark, J.D., and Williams, D. "Small Area Analysis: a Review and Analysis of the North American Literature." Journal of Health Politics, Policy and Law 12 (1987):741-809.

Payne, S.M.C. "Identifying and Managing Inappropriate Hospital Utilization: a Policy Synthesis." Health Services Research 22 (1987):709-769.

Richards, T., Lurie, N., Rogers, W.H., and Brook, R.H. "Case-mix Adjustment: a Quantitative Evaluation of Alternative Measures at the Patient Level." Medical Care 26, Supplement (1988):S94-S121.

Roos, N.P., Flowerdew, G., Wajda, A., and Tate, R. "Variations in Physicians' Hospitalization Practices: a Population-Based Study in Manitoba, Canada." American Journal of Public Health 76 (1986):45-51.

Schneider, E.L. and Brody, J.A. "Aging, Natural Death, and the Compression of Morbidity: Another View." The New England Journal of Medicine 309 (1983):854-855.

Schumacher, D.N., Parker, B., Kofie, V., and Munns, J.M. "Severity of Illness Index and the Adverse Patient

Occurrence Index: a Reliability Study and Policy Implications." Medical Care 25 (1987):965-703.

Stano, M. "A Further Analysis of the 'Variations in Practice Style' Phenomenon." Inquiry 23 (1986):176-181.

Stano, M. and Folland, S. "Variations in the Use of Physician Services by Medicare Beneficiaries." Health Care Financing Review 9 (1988):51-58.

Strauss, M.J., Conrad, D., Logerfo, J.P., Hudson, L.D., and Bergner, M. "Cost and Outcome of Care for Patients with Chronic Obstructive Lung Disease." Medical Care 24 (1986):915-924.

Wennberg, J. "Dealing with Medical Practice Variations: a Proposal for Action." Health Affairs 3, no. 2 (1984):6-32.

Wennberg, J. "Which Rate is Right?" The New England Journal of Medicine 314 (1986):310-311.

Wennberg, J.E. "The Paradox of Appropriate Care." The Journal of the American Medical Association 258 (1987):2568-2569.

White, H. "A Heteroscedascticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedascticity." Econometrics 48 (1980):817-838.

White, R.E., Skipper, B.J, Applegate, W.B., Bennett, M.D., and Chilton, L.A. "Ordering Decision and Clinic Cost Variation among Resident Physicians." The Western Journal of Medicine 141 (1984): 117-122.

TABLE 1

SAMPLE DIAGNOSIS FREQUENCIES (n =1458)

	DX FREQ.	DX PERC.	MDC FREQ.	MDC PERC.
MDC 4 Diseases and Disorders of the Respiratory System				
Chronic Obstructive Pulmonary Disease	48	3.3		
Lung Cancer	43	2.9	427	29.3
Pneumonia & Respiratory Infection	240	16.5		
Asthma	96	6.6		
MDC 5 Diseases and Disorders of the Circulatory System				
Ischemic Heart Disease	257	17.6		
Congestive Heart Failure	273	18.7		
Acute Myocardial Infarction (without CCU admission)	62	4.3	782	53.6
Hypertension	31	2.1		
Acute Myocardial Infarction (with CCU admission)	159	10.9		
MDC 6 Diseases and Disorders of the Digestive System				
Gastroenteritis	53	3.6	178	12.2
GI Hemorrhage	125	8.6		
MDC 10 Endorcrine, Nutritional, and Metabolic Diseases and Disorders				
Diabetes	71	4.9	71	4.9
Total	1458	100.0	1458	100.0

TABLE 2
SAMPLE CHARACTERISTICS (n = 1458)

SEVERITY OF ILLNESS FREQUENCIES (PERCENT)

LEVEL	OVERALL SOII FREQUENCY		STAGE OF THE PRINCIPAL DIAGNOSIS ON ADMISSION	
1	541	(37.1)	23	(1.5)
2	737	(50.5)	912	(62.6)
3	86	(5.9)	485	(33.3)
4	94	(6.4)	38	(2.6)

	MEAN	SD
LENGTH OF STAY	7.08	5.13
TOTAL CHARGES ($)	6625	5776
ROOM CHARGES ($)	2161	1782
LABORATORY CHARGES ($)	1502	1393
INTENSIVE CARE CHARGES($)	880	1744
RADIOLOGY CHARGES ($)	345	449
PHARMACY CHARGES ($)	671	1049
SURGICAL CHARGES ($)	1	3
THERAPY CHARGES ($)	295	762
SUPPLY CHARGES ($)	327	613
MISC.INPATIENT CHARGES($)	333	616

RACE	WHITE		OTHER		BLACK	
	FREQ.	PERC.	FREQ.	PERC.	FREQ.	PERC.
	1199	82.2	54	3.7	205	14.1

AGE	< 60		60-69		> 70	
	FREQ.	PERC.	FREQ.	PERC.	FREQ.	PERC.
	381	26	312	21	765	52

SEX	MALE		FEMALE	
	FREQ.	PERC.	FREQ.	PERC.
	757	51.9	701	48.1

TABLE 3

F-TEST OR T-TEST PROBABILITIES FOR (SETS OF) INDEPENDENT VARIABLES
n = 1458

OVERALL SEVERITY MODEL

	Log of Total Charges	Total Charges	Log of LOS	LOS
Clinical Diagnosis Cluster	<.0001	<.0001	<.0001	<.0001
Overall SOII Rating	<.0001	<.0001	<.0001	<.0001
Race	.5564	.2797	.3347	.1335
Sex	.1701	.2315	.2714	.5269
Age > 69	.9976	.7907	.0728	.1196
Physicians	.0030	.1255	<.0001	.0004
	R^2=.5394	R^2=.4138	R^2=.5040	R^2=.4200

ADMISSION SEVERITY MODEL

	Log of Total Charges	Total Charges	Log of LOS	LOS
Clinical Diagnosis Cluster	<.0001	<.0001	<.0001	<.0001
Admission SOII Rating	<.0001	<.0001	<.0001	<.0001
Race	.9336	.5946	.8530	.4640
Sex	.5878	.6722	.1268	.2020
Age > 69	<.0001	.0012	<.0001	.0033
Physicians	.0001	.0768	< .0001	.0001
	R^2=.3253	R^2=.2493	R^2=.2554	R^2=.2260

Outcome

TABLE 4

COMPARATIVE PHYSICIAN REGRESSION COEFFICIENTS
FROM THE OVERALL SEVERITY MODELS OF TOTAL CHARGES

PERCENTAGE OR DOLLAR DIFFERENCE FROM OMITTED PHYSICIAN

ID #	SPECIALTY	LOG LINEAR COEFFICIENTS	SE	UNTRANSFORMED LINEAR COEFFICIENTS	SE
		LOWEST TEN			
34	Cardiology	-.2272	.0960	-1025.46	848.11
22	General	-.2135	.0884	-860.15	781.48
26	Endocrine	-.2024	.0933	-1155.70	824.44
31	Oncology	-.1526	.1108	-1158.54	978.69
23	General	-.1459	.0927	-165.02	819.08
7	Cardiology	-.1421	.0899	-146.73	794.05
18	Cardiology	-.0887	.0954	-343.65	843.16
1	Endocrine	-.0728	.0971	-413.79	857.98
13	Hematology	-.0431	.0821	-150.81	724.89
8	General	-.0429	.0882	-241.83	779.04
		MIDDLE ELEVEN			
21	Cardiology	-.0415	.1142	-396.36	1010.48
4	Cardiology	-.0219	.0923	186.55	815.81
35	Cardiology	-.0188	.1049	101.64	926.74
39	Cardiology	.0000	.0000	000.00	000.00
11	General	.0037	.0982	389.37	867.80
6	General	.0039	.0973	435.65	860.09
29	General	.0064	.1073	232.41	947.99
2	Cardiology	.0075	.0785	639.23	694.01
28	General	.0095	.1061	623.60	937.84
9	GI	.0162	.0949	811.92	838.85
37	Cardiology	.0192	.0787	302.35	695.52
		HIGHEST TEN			
3	Pulmonary	0.0234	.1064	87.87	940.48
30	Endocrine	0.0259	.0972	734.46	859.04
40	Endocrine	0.0335	.0997	158.40	881.31
25	General	0.0918	.0845	1201.36	747.07
14	General	0.0941	.1103	1555.58	974.29
32	Endocrine	0.1002	.0984	1076.38	869.80
10	General	0.1164	.0903	1954.08	797.87
5	General	0.1202	.0984	1014.99	869.86
15	Cardiology	0.1595	.0958	998.97	846.60
27	General	0.1743	.1001	1321.23	884.34

TABLE 5

CASE MIX DIFFERENCES BETWEEN HIGH AND LOW RESOURCE USE PHYSICIANS

	TEN HIGH RESOURCE PHYSICIANS 418 PATIENTS		TEN LOW RESOURCE PHYSICIANS 487 PATIENTS	
	FREQ.	PERCENT	FREQ.	PERCENT
MAJOR DIAGNOSTIC CATEGORIES				
Respiratory	129	30.9	172	35.3
Circulatory	215	51.4	230	47.2
Digestive	50	12.0	55	11.3
Metabolic (Diabetes)	24	5.7	30	6.2
Total	418	100.0	487	100.0

(Chi square =2.305 p=.5114)

	TEN HIGH RESOURCE PHYSICIANS 418 PATIENTS		TEN LOW RESOURCE PHYSICIANS 487 PATIENTS	
	FREQ.	PERCENT	FREQ.	PERCENT
OVERALL SOII LEVEL				
1	139	33.3	194	39.8
2	222	53.1	231	47.4
3	27	6.5	26	5.3
4	30	7.2	36	7.4
Total	418	100.0	487	100.0

(Chi sqaure=4.593 p=.2041)

	TEN HIGH RESOURCE PHYSICIANS 418 PATIENTS		TEN LOW RESOURCE PHYSICIANS 487 PATIENTS	
	FREQ.	PERCENT	FREQ.	PERCENT
ADMISSION SEVERITY LEVEL				
1	9	2.1	5	1.0
2	247	59.1	306	62.9
3	153	36.6	157	32.2
4	9	2.1	19	3.9
Total	418	100.0	487	100.0

(Chi square=6.386 p=.1721)

TABLE 6
DIFFERENCES IN MEAN CHARGES BETWEEN HIGH AND LOW PHYSICIANS FOR ANCILLARY SERVICES AND SELECTED COST CENTERS

	TEN HIGH M.D.s	TEN LOW M.D.s	HIGH M.D.s SOII LEVEL 1	LOW M.D.s SOII LEVEL 1	HIGH M.D.s SOII LEVEL 2	LOW M.D.s SOII LEVEL 2
Patients	418	487	139	194	222	231
Total charges**	$7690	5780	3383	2737	8029	6684
Length of stay**	7.99	6.24	3.94	3.26	9.01	7.75
Room charges**	2425	1995	1278	1136	2778	2402
All Labs**	1712	1260	891	664	1779	1426
Cardiology	425	261	292	101	481	346
Hematology	262	212	139	119	273	233
Chemistry	657	507	293	263	670	563
Microbiology	155	110	70	88	156	105
Blood Gas	64	48	11	15	30	27
Blood Flow	7	8	4	2	8	13
Pharmacology	45	36	18	25	46	38
Pulmonary	8	17	7	12	11	24
Pathology	34	28	20	16	39	38
GI-Lab	39	24	26	20	49	27
All Radiology*	398	332	217	173	404	364
X-Rays	210	162	120	96	192	165
Soft Tissue	55	36	32	14	57	36
Nuclear Medicine	80	83	37	39	96	105
CT Scan	47	35	27	7	50	45
MRI	3	6	0	10	5	5
Supplies**	419	285	124	88	342	269
Respiratory Therapy**	294	172	66	76	274	168
All Drugs**	844	621	243	280	869	671
IV Admixtures	160	132	52	58	192	165
Oral and Topical	684	489	191	222	677	506
All Intensive Care**	984	670	329	197	852	752
MICU-CCU	811	497	221	82	691	549
Emergency Room	87	96	67	91	90	89

* Overall differences significant at $p < .05$
** Overall differences significant $p < .005$