Variation in Hospital Length of Stay: Do Physicians Adapt Their Length of Stay Decisions to What Is Usual in the Hospital Where They Work?

Judith D. de Jong, Gert P. Westert, Ronald Lagoe, and Peter P. Groenewegen

Objective. To test the hypothesis that physicians who work in different hospitals adapt their length of stay decisions to what is usual in the hospital under consideration.

Data Sources. Secondary data were used, originating from the Statewide Planning and Research Cooperative System (SPARCS). SPARCS is a major management tool for assisting hospitals, agencies, and health care organizations with decision making in relation to financial planning and monitoring of inpatient and ambulatory surgery services and costs in New York state.

Study Design. Data on length of stay for surgical interventions and medical conditions (a total of seven diagnosis-related groups [DRGs]) were studied, to find out whether there is more variation between than within hospitals. Data (1999, 2000, and 2001) from all hospitals in New York state were used. The study examined physicians practicing in one hospital and physicians practicing in more than one hospital, to determine whether average length of stay differs according to the hospital of practice. Multilevel models were used to determine variation between and within hospitals. A *t*-test was used to test whether length of stay for patients of each multihospital physician differed from the average length of stay in each of the two hospitals.

Principal Findings. There is significantly ($p < .05$) more variation between than within hospitals in most of the study populations. Physicians working in two hospitals had patient lengths of stay comparable with the usual practice in the hospital where the procedure was performed. The proportion of physicians working in one hospital did not have a consistent effect for all DRGs on the variation within hospitals.

Conclusion. Physicians adapt to their colleagues or to the managerial demands of the particular hospital in which they work. The hospital and broader work environment should be taken into account when developing effective interventions to reduce variation in medical practice.

Key Words. Length of stay, variation, hospitals, multihospital physicians

A persistent finding in health services research is that hospital utilization varies widely (Paul-Shaheen, Clark, and Williams 1987; Ashton et al. 1999; Wennberg 1999). These variations have been observed between geographic areas, hospitals, and physicians. The variation within these units of analysis has been found to be smaller than between the units, for different types of services, numbers of admissions, and length of stay (Wennberg and Gittelsohn 1982; Read et al. 1983; Westert, Nieboer, and Groenewegen 1993; Arndt, Bradbury, and Golec 1995; O'Connor et al. 1999).

Several explanations have been sought for the variation between and similarities within units; a summary can be found in Table 1. Wennberg and Gittelsohn (1975) suggested that an explanation lies with the physicians themselves. First, they theorized that professional uncertainty explains whether the specific procedure or diagnosis will have high or low variation. Second, they hypothesized that the judgment and preferences of groups of physicians give rise to a unique pattern over time, which has been termed a "surgical signature." Chassin (1993) suggested that variation is caused by a difference in the prevalence of physicians who are enthusiastic about certain procedures. The "surgical signature" and enthusiasm hypotheses assume that physicians have a preference for certain procedures, but the behavioral mechanisms that produce different practice styles remain unclear.

Westert and Groenewegen (1999) offered an alternative to this preference-centered approach, emphasizing incentives and environmental conditions that influence the behavior of physicians by providing opportunities and constraints. Westert (1992) applied this approach in a model of local standards that predicts similarities among physicians who share a common work environment and thus a social system and similar constraints. We tested whether variations in medical practice are indeed related to the hospital in which physicians practice. This would imply that variation within hospitals is small compared with variation between hospitals for physicians treating similar patients. A second implication would be that physicians working in more than one hospital conform to the usual practice of each hospital (Westert 1992). This implication cannot be deduced from a preference-centered approach, as

Address correspondence to Judith D. de Jong, M.Sc., NIVEL-Netherlands Institute for Health Services Research, PO Box 1568, 3500 BN Utrecht, The Netherlands. Gert P. Westert, Ph.D., is with the RIVM-National Institute of Public Health and the Environment, BA Bilthoven, The Netherlands. Ronald Lagoe, Ph.D., is with the Hospital Executive Council, Syracuse, NY. Peter P. Groenewegen, Ph.D., is with the NIVEL-Netherlands Institute for Health Services Research, BN Utrecht, The Netherlands.

Table 1: Summary of Explanations for Variation

<i>Type of Explanation</i>	<i>Content</i>	<i>Hypotheses</i>	<i>Study Citations</i>
Preference centered	Practice styles are caused by differences in preference for certain procedures	Practice style remains the same when social context is changed	Wennberg and Gittelsohn (1975) and Chassin (1993)
Constraint centered	Practice styles are caused by differences in characteristics of the (social) context, providing opportunities and constraints	Similarities of practice styles within the same social context, practice style changes with changes in the social context	Westert (1992) and Westert and Groenewegen (1999)

the preferences of an individual would not change when working in another hospital.

Indications were found by Westert, Nieboer, and Groenewegen (1993) that multihospital physicians in the Netherlands have a patient length of stay close to the usual practice in the hospital where the surgery was performed. Griffiths, Waters, and Acheson (1979), using data from the British National Health Service, also found that average postoperative stays were similar among physicians who practice in the same hospitals, while the average between hospitals was significantly different for physicians who practice in more than one hospital.

In the present study, we tested these implications with data from the United States, where it is quite common for physicians to work in more than one hospital. As a result, the implications can be tested more reliably than was carried out by Westert with 23 physicians in five hospitals (only four working in more than one hospital) and Griffiths with nine physicians in eight hospitals (six working in more than one hospital). Furthermore, we used a different methodology. We took length of stay, which is a well-defined and important indicator in hospital management, as the outcome variable.

BACKGROUND AND HYPOTHESES

In general, people show similar behavior within work environments; colleagues show similarities (Denton 1999); attitudes and beliefs of individuals change toward the group norm (Dambrun, Guimond, and Duarte 2002); "... people both shape and are shaped by social networks" (Pescosolido 1992). Similarities as a

consequence of attitude and belief change is a gradual process. Other explanations for similar behavior are selection and adaptation to the circumstances under which physicians work (de Jong, Groenewegen, and Westert 2003). Similarities as a consequence of the latter mechanisms is much quicker than gradual adaptation. Physicians will adapt to the usual practice in the hospital in which they treat patients, in order to avoid being criticized (Eddy 1984; Westert 1992). Physicians expected to show similar behavior will be selected by the hospital management or by colleagues. If this strategy is followed, local standards of medical care come into being within medical teams that share the same work environment (Westert and Groenewegen 1999). Whether local standards come into being or not, might depend on the extent to which the physicians working in a hospital form a social group, and are able to develop into a normative community. We assume that the chances of developing local standards in a hospital are lower when more physicians in this hospital work in more than one hospital. If there is a small proportion of physicians who practice in this hospital only, average length of stay is expected to be less standardized.

In this study, we first examined whether there is less variation between physicians working in the same hospital than between different hospitals. Second, we examined whether physicians working in more than one hospital practice differently in different hospitals. Third, we tested the effect of the proportion of physicians working in only one hospital on the variation within hospitals. The results of this study have implications for the explanation of variation between hospitals and similarities within hospitals. If physicians working in different hospitals behave according to the local standards in the hospitals where the procedures were performed and thus show intradoctor variation, selection, based on characteristics that relate to similar behavior, is not an explanation for more similarities within hospitals than between them. This is an important test to find out whether selection holds true, because the physicians are the same. If selection based on similarities is an explanation of similar behavior, a physician would not do different things in different hospitals.

If variations between physicians are related to the hospital, as we argued,

Hypothesis 1: The variation between physicians working in the same hospital is less than the variation between hospitals.

Hypothesis 2: Physicians working in different hospitals in which the average lengths of stay differ, will choose lengths of stay similar to the average in the hospital in which the patient is treated.

Normative processes cause similarities within hospitals. The influence of colleagues or of the hospital management will be more important when the average length of stay within a hospital is more standardized; deviation from

the average length of stay is more obvious when there is less variation. The distribution will show less variability for those physicians who only practice in that hospital, as they are more dependent on each other and will avoid being criticized. Therefore:

Hypothesis 3: The higher the proportion of physicians working in only one hospital, the less variation in the length of stay there will be between physicians in that hospital.

DATA AND METHOD

Description of the Data

Data were obtained from the New York Statewide Planning and Research Cooperative System (SPARCS), which is a comprehensive patient data system established as a result of cooperation between the health care industry and the government. SPARCS is a major management tool assisting hospitals, agencies, and health care organizations with decision making in relation to financial planning and monitoring of inpatient and ambulatory surgery services and costs in New York state. Several research articles have been published that are based on SPARCS data (Pasley, Lagoe, and Marshall 1995; Westert and Lagoe 1995; Hainsworth et al. 1997; Lagoe 1998; Murphy, Noetscher, and Lagoe 1999; Noetscher et al. 1999; Lagoe, Noetscher, and Murphy 2001; Noetscher, Morreale, and Lagoe 2001; de Jong et al. 2004).

We used 1999, 2000, and 2001 SPARCS data to study seven diagnosis-related groups (DRGs): two medical (DRGs 88 and 127), one surgical (DRG 209), and four obstetrical (DRGs 358, 359, 370, and 371). Cases for which no physician was known were excluded (1.3 percent of all cases). Only patients above the age of 20 were included and patients with extremely long stays (defined as the average length of stay plus 1.96 times the standard deviation) were excluded. This involved a minimum of 0.86 percent and a maximum of 3.72 percent of cases per procedure. In 2001, 70 percent of the physicians worked in only one hospital and 30 percent worked in more than one hospital. About 93 percent of all physicians work in one or two hospitals, 7 percent work in more than two hospitals. This is about the same for all 3 years of data. The study populations for all 3 years are summarized in Table 2.

Methods

Data for each of the seven diagnoses and procedures contained in the DRGs were separately evaluated with respect to three hypotheses.

Table 2: Study Population: Number of Patients, Physicians, Hospitals and Percentage of Excluded Cases per DRG 1999, 2000, 2001

DRG	Diagnosis/Procedure	Number of Discharges			Number of Physicians			Number of Hospitals			Percentage of Cases Excluded		
		1999	2000	2001	1999	2000	2001	1999	2000	2001	1999	2000	2001
88	Chronic obstructive pulmonary disease	38,424	36,478	34,401	9,125	8,779	8,476	237	234	228	0.86	3.46	2.62
127	Congestive heart failure	62,682	62,599	59,763	11,282	11,285	11,048	233	234	229	3.21	3.72	3.00
209	Major joint procedures	28,426	29,827	32,016	1,267	1,238	1,240	210	207	204	2.58	3.04	3.13
358	Hysterectomy with complications	8,408	8,412	8,137	2,236	2,154	2,104	215	208	207	1.87	3.41	3.58
359	Hysterectomy without complications	21,962	22,926	22,604	2,630	2,608	2,520	216	215	212	2.46	2.09	1.78
370	Cesarean section with complications	11,731	12,125	11,954	2,217	2,240	2,183	164	161	159	2.61	2.54	2.42
371	Cesarean section without complications	39,844	42,980	42,770	2,602	2,654	2,571	163	161	159	1.31	1.25	1.22

DRG, diagnosis-related group.

H1: Variation between and within hospitals

The initial component of the study attempted to identify the most significant sources of variation in physician practice with respect to lengths of stay. This was accomplished by comparing variations in lengths of stay between hospitals with variations in stay within hospitals, for all physicians. It can be argued that most of the variation in hospital utilization is caused by patient characteristics, like severity of illness and case mix (Powell et al. 2004). Some hospitals might treat patients with a higher severity of illness. It is possible that physicians can choose the hospital where to treat their patients on the basis of the severity of illness, or on other unmeasured characteristics. In our study, we excluded as much as possible that the variation found is related to severity of illness and case mix. The comparison involved the proportion of total variation in stays that did not result from patient age, gender (male, female), race (white, black, native, Asian, other, unknown), or severity of illness (index from one to four). Hospital severity of illness was identified using the All Patients Refined (APR) System developed by the 3M Corporation of Wallingford, Connecticut, USA (Averill et al. 2002). The APR System is used to identify severity in hospitals throughout the United States and Europe. This system identifies severity of illness by assigning a level of illness to each individual International Classification of Diseases, Ninth Revision medical records code. The levels of illness range from 1, minor; 2, moderate; 3, major; and 4, extreme. For each patient, the principal diagnosis and each secondary diagnosis receive one of these levels of severity.

In order to identify a level of severity for an individual patient, the APR System uses a formula to develop a single level based on the levels for each diagnosis. All variables were centered around their means, so the model has an interpretable meaning (Snijders and Bosker 1999).

We also measured the extent to which physicians in the same hospital choose similar lengths of stay. This is carried out by calculating the intraclass correlation (ICC), defined as the variation between hospitals divided by the total variation in length of stay, and it was accomplished using multilevel analysis (Snijders and Bosker 1999). Multilevel analysis is an extension of ordinary least squares regression analysis. Ordinary least squares regression analysis allows us, for example, to estimate the relationship between length of stay and age in our study population, assuming that there is no physician or hospital effect in addition to the characteristics of the patient. With multilevel analysis, however, total variation in length of stay is separated into three parts: a part because of differences between patients, a part because of differences

between physicians, and a part because of differences between hospitals (Leyland and Groenewegen 2003). The model is described in Appendix A.

We did not adjust for hospital differences, because this is part of the context and thus subject of the study: the objective was not to explain the differences in length of stay, but to explain patterns of variation.

H2: Variation in physicians working in two hospitals.

The second component of the analysis evaluated differences in practice patterns in physicians who practiced in more than one hospital. Inclusion criteria were developed for physicians and hospitals.

Physicians were included for the analysis if they discharged patients from either one or two hospitals. These situations accounted for a substantial majority of the physicians in the data (93 percent). All physicians who practiced at more than two hospitals were excluded (7 percent). Those who practiced at only one hospital formed a comparison group; those who practiced at two hospitals were the focus of the study. In order to permit meaningful comparison of physician lengths of stay per DRG, a hospital was included if it had at least two physicians who practiced only at this hospital as well as a third physician who practiced at another hospital as well for a specific DRG.

The analysis for this component of the study involved comparison of physician lengths of stay with a "local standard" or usual practice. The local standard for a hospital was defined as the mean length of stay for physicians working only in this hospital. In order to provide a baseline for subsequent analyses, local standards for hospitals in the study were compared, to determine which of these facilities had lengths of stay that differed by at least 0.5 day.

One half day was chosen since it is associated with savings in expenses, as well as with additional revenue if another patient can be admitted to the vacant bed. In 2001, the difference in mean lengths of stay between the two hospitals where the physician practiced was less than 0.5 days for 48 percent of all physicians who worked in different hospitals. The focus of the analysis was on comparison of the average length of stay for each multihospital physician with the local standard. A *t*-test (Hays 1973) was used to determine whether the mean length of stay for patients of each multihospital physician differed from the local standard at each of the two hospitals. When this test was non-significant, the physician's practice pattern was assumed to be similar to the local standard at that hospital (measured as the mean length of stay generated by the physicians who only work in that hospital). SPSS was used for this *t*-test. The means of the hospital and the multihospital physicians are compared in this procedure, which computes the differences between the multihospital

physicians' average and the hospital's average in each case, and tests whether the average difference is other than 0.

This component of the analysis also compared the quantitative differences between local standard lengths of stay and differences between multi-hospital physician lengths of stay between facilities. This analysis determined whether these differences in stays pointed in the same direction. A sign test was used in this comparison (Wonnacott and Wonnacott 1990). Although a rather weak test, it is appropriate, for we are only interested in the direction of the difference. First, local standards were subtracted and second, physician stays were subtracted for the same pairs of hospitals. If the signs (positive or negative) of the differences were the same, it was concluded that the physician adjusted patients' lengths of stay were in the direction of the usual practice pattern at the hospital.

H3: Influence of the proportion of physicians practicing in only one hospital on total variation in length of stay within a hospital.

The potential impact of lengths of stay generated by physicians working in only one hospital on total practice variation within a hospital was also evaluated. This was measured using a covariance estimated with multilevel analysis, measuring that part of the variation that can be attributed to the proportion of physicians practicing in only one hospital. The exact model is described in Appendix B. The percentage of physicians working in one hospital was centered around 99 percent. It was expected that the variation within hospitals would decrease when the proportion of physicians working in only one hospital increased.

RESULTS

Results of the study for Hypothesis 1 are summarized in Table 3. This information relates to that proportion of the variation in stays that was not associated with age, gender, race, or severity of illness.

The data in Table 3 indicate there was significantly more variation in lengths of stay between hospitals than within hospitals in most of the study populations. These populations included patients treated for major joint procedures, and cesarean section with and without complications in all 3 years of the data, as well as those with chronic obstructive pulmonary disease in one of the 3 years of the data. Variation within hospitals was significantly greater than variation between hospitals in the case of patients who received

Table 3: Average Lengths of Stay, Corrected for Age, Sex, Race, and Severity of Illness, 95 Percent Confidence Intervals for the Average per Hospital, ICC, and Variance Components per DRG for 1999, 2000, and 2001

DRG	Diagnosis/Procedure	1999				2000				2001			
		Average (SE)	Hospital (95% CI)	ICC	Variance (Components)	Average (SE)	Hospital (95% CI)	ICC	Variance (Components)	Average (SE)	Hospital (95% CI)	ICC	Variance (Components)
88	Chronic obstructive pulmonary disease	6.01 (0.08)	3.66-8.36	0.09	1.44 (0.15)* 1.09 (0.07)[†] 13.85 (0.11) [‡]	5.48 (0.06)	3.69-7.26	0.08	0.83 (0.09)* 0.69 (0.04) [†] 8.77 (0.07) [‡]	5.37 (0.07)	3.49-7.24	0.09	0.92 (0.10)* 0.74 (0.05) [†] 8.84 (0.07) [‡]
127	Congestive heart failure	5.78 (0.06)	4.04-7.53	0.06	0.79 (0.08) 0.77 (0.04) 11.91 (0.07)	5.65 (0.06)	3.91-7.39	0.06	0.79 (0.08) 0.85 (0.04) 10.82 (0.07)	5.63 (0.06)	3.89-7.38	0.06	0.79 (0.09) 0.86 (0.04) 11.36 (0.07)
209	Major joint procedures	6.42 (0.10)	3.96-8.88	0.23	1.57 (0.18) 0.65 (0.04) 4.60 (0.04)	6.13 (0.08)	3.97-8.28	0.21	1.21 (0.14) 0.59 (0.04) 3.92 (0.03)	5.80 (0.08)	3.75-7.86	0.22	1.10 (0.13) 0.53 (0.04) 3.42 (0.03)
338	Hysterectomy with complications	4.12 (0.04)	3.32-4.91	0.04	0.17 (0.03) 0.34 (0.04) 3.36 (0.06)	3.83 (0.04)	3.13-4.53	0.05	0.13 (0.02) 0.24 (0.03) 2.20 (0.04)	3.74 (0.05)	2.73-4.75	0.10	0.27 (0.04) 0.23 (0.03) 2.26 (0.04)
359	Hysterectomy without complications	2.90 (0.02)	2.34-3.46	0.09	0.08 (0.01) 0.16 (0.01) 0.69 (0.01)	2.82 (0.02)	2.30-3.33	0.08	0.07 (0.01) 0.14 (0.01) 0.67 (0.01)	2.75 (0.02)	2.17-3.34	0.10	0.09 (0.01) 0.14 (0.01) 0.67 (0.01)
370	Cesarean section with complications	4.62 (0.04)	3.87-5.38	0.04	0.15 (0.03) 0.06 (0.02) 3.23 (0.05)	4.53 (0.04)	3.69-5.36	0.06	0.18 (0.03) 0.07 (0.02) 2.87 (0.04)	4.43 (0.05)	3.44-5.42	0.08	0.25 (0.04) 0.05 (0.02) 3.07 (0.04)
371	Cesarean section without complications	3.69 (0.03)	2.98-4.40	0.16	0.13 (0.02) 0.03 (0.00) 0.64 (0.00)	3.66 (0.03)	2.94-4.38	0.17	0.13 (0.02) 0.03 (0.00) 0.62 (0.00)	3.59 (0.03)	2.88-4.31	0.17	0.13 (0.02) 0.03 (0.00) 0.62 (0.00)

*Variance component at the hospital level.

[†]Variance component at the physician level.

[‡]Variance component at the patient level.

Values are bold if the difference between the hospital and physician variance is significant.
DRG, diagnosis-related group; ICC, intraclass correlation; SE, standard error; CI, confidence interval.

hysterectomies without complications in 3 years of data, and for hysterectomies with complications in 2 years of data. The ICC tells us how strong the resemblance is within hospitals (Leyland and Groenewegen 2003).

On the basis of these analyses, Hypothesis 1 was confirmed. The finding that there is more variation between than within hospitals was the basis for the second component of the analysis. Results of the study for Hypothesis 2 are summarized in Tables 3 and 4. This information relates to lengths of stay for physicians working in two hospitals and the relationships between their practice patterns and local standards, defined as the mean length of stay for physicians working in one hospital.

For almost all of the study populations, including both medical and surgical patients, the *t*-tests indicated that lengths of stay for multihospital physicians (between 71 and 1,009 physicians, depending on the DRG) did not differ significantly from the local standard in the 3 years (Table 4). It should be emphasized that these comparisons involved the stays chosen by multihospital physicians and the local standards at each location of practice (hospital A and hospital B). The only significant difference between length of stay chosen by multihospital physicians and local standards relates to a single year (1999) and involves major joint procedures and hysterectomy without complications.

The length of stay for hysterectomy without complications is not related to the hospital, an indication that is based on the observation that variation is lower between hospitals than between physicians. The last component of this part of the analysis compared quantitative differences between local standard hospital lengths of stay and differences between multihospital physicians' lengths of stay for the same hospitals. Results of this sign test are summarized in Table 4.

This information demonstrates that differences between hospital stays pointed in the same direction as differences between physician stays in a majority of the cases for all seven DRGs in each of the 3 years of study data. The proportion of physicians for whom the stays pointed in the same direction varied from 51 to 86 percent.

The combination of analyses which comprised the second component of the study demonstrated that physicians who practiced at more than one hospital produced lengths of stay that were similar to the local standards where they were practicing. The *t*-tests demonstrated that stays for multihospital physicians did not differ significantly from the local standard for five of the DRGs in all 3 years, and for two of these categories in 2 of the 3 years. Finally, the sign tests demonstrated that differences between hospital stays were similar to differences between stays chosen by multihospital physicians in a majority of the cases evaluated.

Table 4: χ^2 Test and Sign Test for Multihospital Physicians for 1999, 2000, and 2001

DRG	Diagnosis/Procedure	N (Physicians in Two Hospitals)			t-Test Pair Physician 1 in A and Hospital A			t-Test Pair Physician 1 in B and Hospital B			Sign Difference between Hospitals Equals Difference for the Physicians (%)		
		1999	2000	2001	1999	2000	2001	1999	2000	2001	1999	2000	2001
		537	516	516	nsig	nsig	nsig	nsig	nsig	nsig	56.1	54.7	54.7
88	Chronic obstructive pulmonary disease	877	972	1,009	nsig	nsig	nsig	nsig	nsig	nsig	55.8	57.8	57.6
127	Congestive heart failure	123	135	95	sig*	nsig	nsig	nsig	nsig	nsig	68.3	67.4	68.4
209	Major joint procedures	120	87	103	nsig	nsig	nsig	nsig	nsig	nsig	58.3	60.9	58.3
358	Hysterectomy with complications	100	73	71	nsig	nsig	nsig	sig*	nsig	nsig	51.0	56.2	57.7
359	Hysterectomy without complications	128	130	124	nsig	nsig	nsig	nsig	nsig	nsig	67.2	62.3	62.9
370	Cesarean section with complications	104	74	74	nsig	nsig	nsig	nsig	nsig	nsig	86.5	70.3	73.0
371	Cesarean section without complications												

* $p < .05$.
 DRG, diagnosis-related group; sig, significant; nsig, not significant.

These two analyses provided convincing evidence of the linkage between multihospital physician lengths of stay and local standards. On the basis of this information, Hypothesis 2 was confirmed.

Results of the study of Hypothesis 3 are summarized in Table 5. This information relates to the potential impact on lengths of stay of the proportion of physicians working in only one hospital on total practice variation within a hospital.

The results of the physician variance analysis demonstrated that the relationship between the proportion of the physicians working in only one hospital and the total physician variation was nonsignificant for a majority of the DRGs and years of data evaluated. This means that the variation does not change when the proportion of physicians working in one hospital changes. Significant relationships were identified for chronic obstructive pulmonary disease, congestive heart failure, major joint procedures, and cesarean section with complications in 2 years of data and in two isolated instances, via hysterectomy with complications in 2000 and cesarean section without complications in 2001. Only for chronic obstructive pulmonary disease the variation decreased when the proportion of physicians working in one hospital increased, in all other significant cases the variation increased when the proportion of physicians increased.

On the basis of this analysis, Hypothesis 3 was not confirmed. The conclusion might indicate a connection between length of stay choice and hospital influences, like management. Besides normative processes, regulative processes cause similarities. Regulative processes, or formal rules, will cause these similarities also when there is a small proportion of physicians working in one hospital and not in another also. In the last decades, there has been a change from professional control, through normative processes, to managerial control, which is regulative (Scott et al. 2000).

CONCLUSION AND DISCUSSION

In the classical conception of medicine as a profession, medical practice is supposed to be uniform because of the shared body of (theoretical) knowledge. Variation originates from the necessity to apply this theoretical knowledge to individual patients. The underlying assumption is one of professional autonomy; the decision of physicians is influenced by science and the medical condition of the patient only.

Table 5: Variances of the Physician Level Intercept and the Proportion of Physicians Working in One Hospital within Hospitals, in 1999, 2000, and 2001

DRG	Diagnosis/Procedure	1999		2000		2001	
		Variance	Significance	Variance	Significance	Variance	Significance
88	Chronic obstructive pulmonary disease	1.11 (0.08)		0.60 (0.05)		0.73 (0.05)	
		- 2.22 (0.60)	$p < .01$	1.18 (0.53)	$p = .03$	- 0.36 (0.45)	nsig
127	Congestive heart failure	0.79 (0.04)		0.85 (0.04)		0.88 (0.05)	
		- 1.39 (0.55)	$p = .01$	0.01 (0.51)	nsig	- 1.58 (0.48)	$p < .01$
209	Major joint procedures	0.78 (0.05)		0.71 (0.05)		0.55 (0.04)	
		- 1.00 (0.10)	$p < .01$	- 0.73 (0.13)	$p < .01$	- 0.19 (0.17)	nsig
358	Hysterectomy with complications	0.35 (0.05)		0.28 (0.03)		0.21 (0.03)	
		- 0.35 (0.33)	nsig	- 0.63 (0.19)	$p < .01$	0.22 (0.30)	nsig
359	Hysterectomy without complications	0.17 (0.01)		0.15 (0.01)		0.15 (0.01)	
		- 0.09 (0.06)	nsig	- 0.07 (0.07)	nsig	- 0.07 (0.07)	nsig
370	Cesarean section with complications	0.08 (0.02)		0.08 (0.02)		0.09 (0.02)	
		- 0.33 (0.15)	$p = .03$	- 0.19 (0.19)	nsig	- 0.55 (0.16)	$p < .01$
371	Cesarean section without complications	0.03 (0.00)		0.03 (0.00)		0.03 (0.00)	
		0.01 (0.02)	nsig	- 0.02 (0.02)	nsig	- 0.04 (0.01)	$p = .01$

DRG, diagnosis-related group; nsig, not significant.

Residual variation still remains, however, when clinical variables and patient characteristics are taken into account and this residual variation is not random, but shows clear patterns. The most studied ones are small area variations (Wennberg and Gittelsohn 1975, 1982; Ashton et al. 1999). Researchers differ as to the attribution of this residual variation, some arguing that physicians prefer different practice styles, e.g., as a result of their education and professional socialization. The spatial phenomenon then emerges as a result of (self-)selection of physicians in certain areas or adaptation of preferences to those of others, resulting in a typical "surgical signature."

Other researchers have argued that the circumstances under which physicians work have a profound influence on their decisions, and that these shared circumstances make for homogeneity and differing circumstances result in variation. Freidson (1975) argued that in group practice, physicians cannot make their decisions autonomously, because they depend on colleagues, and are controlled by them. Although individual physicians can be conscientious and ethical, circumstances can cause them to change their performance.

The physician influence on utilization of care has been the subject of several studies (e.g., Burns and Wholey 1991; Geller, Burns, and Brailer 1996). In a study by Burns, Chilingerian, and Wholey (1994) it was concluded that the physician is an important source of variation. Freidson (1975) stated that the social environment in which physicians work is more important for their medical behavior than their formal professional education. For instance physicians working in hospitals are subject to collegial norms, affecting their (medical) behavior. Coser (1957) demonstrated that there are different norms between wards within the same hospital. These differences, informal as opposed to formal, were related to different constraints because of differences in medical tasks. The wards, surgical and medical, were situated on two sides of the same floor.

In the debate in which explanations for variation are sought, the question arises at what level of analysis homogeneity might be expected. Reasoning from shared circumstances as a source of homogeneity, this study focused on hospitals as more appropriate than areas (unless areas are defined on such a scale that they form the market area of a single hospital). Because we studied several DRGs separately, this is equivalent to studying wards of the same specialty. The hypothesis that variation within hospitals was small compared with variation between hospitals was confirmed in this study and persistent over 3 years, in most of the study populations.

This result, however, is still compatible with both the approach based on preferences for a certain practice style and the approach based on work circumstances. We therefore also analyzed the decisions of the same physicians

in different hospitals. Preferences for a certain practice style are supposed to be relatively stable within the same person, but circumstances may clearly differ. Different analyses confirmed that physicians working in two hospitals with different average lengths of stay have a length of stay similar to the usual practice in the hospital where the procedure was performed.

Underlying assumption in our study is that the physician, not the hospital, decides over patients' discharges. The hospital-physician relationship has changed from physician owned in the early twentieth century to one of joint control in the late 1980s (Shortell 1991). However, the lack of common economic incentives makes it difficult for physicians and hospitals to cooperate (Shortell et al. 2000). With a change from professional dominance to managerial-market orientation (Scott et al. 2000), the question of whether the within-hospital similarities are choice or constraint is raised. Are physicians the central actors, or is it management? It might be that in the modern hospital formal management is more important than normative control by physicians. Still, it is the physician who signs the discharge note, and who runs the risk of malpractice suits. For the hospital, although there are incentives to do less, quality is important as well. Therefore, although regulative control might have become more important, physicians can still be considered important actors in the length of stay decisions. The physician is the one who discharges patients, and patients can only be discharged on the day their physician is around.

However, this might change in the future when for instance hospitalists, physicians used by the hospital, sign discharge notes for patients of other physicians. For hospitals this has the advantage of being able to discharge a patient, who is in the condition of being discharged, when the physician is not around.

The hypothesis that variation within a hospital would be smaller when more physicians practice in that hospital alone was not confirmed. This finding could be, however, because of the fact that there are not many hospitals in which there is a low proportion (range 0.16–0.99, mean 0.82) of physicians working only in that hospital. Results in this study are consistent with the results found in the study by Westert, Nieboer, and Groenewegen (1993) and the study by Griffiths, Waters, and Acheson (1979).

This study demonstrated that hospitals are important in studying variation in physicians' practice and explaining length of stay and practice variation. It is important to understand the variation phenomenon, because it will facilitate effective interventions to improve quality of care (Blumenthal 1994). Knowing where variation originates, combined with knowledge about which variation is undesirable, is the key to successful interventions. The existence of variation is often interpreted as a sign of overuse of health care resources and

resources could be saved if all physicians adapted to the lowest utilization rates (Fisher et al. 1994). Although overuse gets more attention, however, underuse could be a problem as well (McNeil 2001). Sources of undesirable variation need to be identified, whether they are indicators of overuse or underuse, with an emphasis on those sources that can be influenced to improve the quality of medical care. If similar patients receive different treatments, or if in some areas it is more likely to receive a certain procedure than in others, some do not get what they need, or get more than they need. Both are undesirable. The former meaning that some people's health could be improved, the latter that there is a waste of resources by providing ineffective, unnecessary care. Part of variation can be legitimate, while another part is not. That part of variation that is not legitimate should be reduced. Although, up to now, there is no evidence that less variation is related to a higher quality of care (Weide et al. 1999; Fertig et al. 1993). Patients should receive the treatment that works best, against acceptable costs. So, besides the question whether variations matter for mortality, morbidity, and quality of life, the question whether some patients receive high quality of care against lower costs is important.

Beyond the importance of understanding the variation phenomenon, it is necessary to study physicians who practice in more than one hospital. Previous research showed that length of stay was longer for the medical patients and shorter for the surgical patients of physicians working in several hospitals, and it was found that inpatient resource use is higher for physicians working in several hospitals (Burns and Wholey 1991; Burns, Geller, and Wholey 1995; Miller, Welch, and Welch 1996).

In this study, we did not find that physicians who practice in two hospitals choose different lengths of stay compared with physicians who practice in one hospital only. We found instead, that multihospital physicians had lengths of stay comparable with the usual practice in the hospital. This implies that physicians adjust to colleagues or circumstances in the hospital where they perform procedures. This implication is useful in our understanding of the variation phenomenon; there are forces within shared work environments that cause physicians to make similar decisions. In this study, we were not able to distinguish between influences exerted by colleagues and by circumstances. This could have been performed if specific characteristics of the hospital had been taken into account. Those characteristics were unknown.

Patient selection could also be an explanation. Multihospital physicians may apply some criterion that is possibly related to length of stay, when choosing the hospital to which they admit patients. This, however, would reinforce the theory of the influence of the hospital on length of stay. The purpose of this

article is not to describe differences between hospitals based on patient selection and therefore the possibility of composition effects should be excluded as much as possible. We did this by adjusting for case mix including severity of illness.

An important question is whether these work environments can be used in interventions to improve the quality of medical care. Several further possible explanations remain for the patterns of variation within work environments. It may be that physicians adapt to the colleagues they work with—that they follow the pack as Eddy (1984) called it—or that there is another factor in the work environment, such as hospital management or the availability of beds and facilities (Westert 1992; Westert and Groenewegen 1999; Kroneman 2001), which influences their medical choices.

There is probably not one single explanation, although one might be more important than the other. If variation is very much related to the influence of colleagues, consensus conferences are a useful instrument for effective interventions. Two strategies could be applied, via broad conferences trying to influence physicians from many different hospitals to adopt the same evidence-based standards and guidelines, or hospital-based conferences trying to influence all physicians in a certain group to change their “local standards” toward more global, evidence-based standards. It is, however, questionable whether conferences are effective in changing physician behavior, or that stronger interventions should be developed. Thom (2000) showed in a study on the effects of a training course for physicians to improve their behavior and to increase patients’ trust, that the intervention was not strong enough. Available evidence suggests that physician profiling can be efficient (Evans, Hwang, and Nagarajan 1995). Physician profiling is a technique used to change hospital length of stay choice of physicians, by comparing their individual average to a benchmark figure, adjusted for severity of illness. The physicians are confronted with their average, without knowing the average of their direct colleagues. This individual-based, managerial approach avoids the informal standards within a hospital, which could be a barrier when trying to influence length of stay choices in a group approach. Managerial interventions might be more effective when applied to individuals, while interventions focusing on professional content might be more effective when the approach is based on teams of physicians.

REFERENCES

- Arndt, M., R. C. Bradbury, and J. Golec. 1995. “Indications for Hysterectomy: Variation within and across Hospitals.” *Medical Care Research and Reviews* 52: 342–63.

- Ashton, C. M., N. J. Petersen, J. Soucek, T. J. Menke, H. Yu, K. Pietz, M. L. Eigenbrodt, G. Barbour, K. W. Kizer, and N. P. Wray. 1999. "Geographic Variations in Utilization Rates in Veterans Affairs Hospitals and Clinics." *New England Journal of Medicine* 340: 32-9.
- Averill, R. F., N. Goldfield, B. A. Steinbeck, and T. Grant. 2002. *All Patients Refined Diagnosis Related Groups (APDRGs)*. Wallingford, CT: 3M Health Information Systems.
- Blumenthal, D. 1994. "The Variation Phenomenon in 1994." *New England Journal of Medicine* 331: 1017-8.
- Burns, L. R., J. A. Chilingerian, and D. R. Wholey. 1994. "The Effect of Physician Practice Organization on Efficient Utilization of Hospital Resources." *Health Services Research* 29 (5): 583-603.
- Burns, L. R., S. E. Geller, and D. R. Wholey. 1995. "The Effect of Physician Factors on the Cesarean Section Decision." *Medical Care* 33: 365-82.
- Burns, L. R., and D. R. Wholey. 1991. "The Effects of Patient, Hospital, and Physician Characteristics on Length of Stay and Mortality." *Medical Care* 29: 251-71.
- Chassin, M. R. 1993. "Explaining Geographic Variations: The Enthusiasm Hypothesis." *Medical Care* 31 (5, suppl): YS37-44.
- Coser, R. L. 1957. "Authority and Decision-Making in a Hospital: A Comparative Analysis." *American Sociological Review* 23 (1): 56-63.
- Dambrun, M., S. Guimond, and S. Duarte. 2002. "The Impact of Hierarchy-Enhancing vs. Attenuating Academic Major on Stereotyping: The Mediating Role of Perceived Social Norm." *Crisp* 7 (8): 114-35.
- Denton, D. W. 1999. "The Attraction-Selection-Attrition Model of Organizational Behavior and the Homogeneity of Managerial Personality." *Crisp* 4 (8): 146-59.
- Eddy, D. M. 1984. "Variations in Physician Practice: The Role of Uncertainty." *Health Affairs* 3 (2): 74-89.
- Evans, J. H. III, Y. Hwang, and N. Nagarajan. 1995. "Physicians' Response to Length-of-Stay Profiling." *Medical Care* 33: 1106-19.
- Fertig, A., M. Roland, H. King, and T. Moore. 1993. "Understanding Variation in Rates of Referral among General Practitioners: Are Inappropriate Referrals Important and Would Guidelines Help to Reduce Rates?" *British Medical Journal* 307: 1467-70.
- Fisher, E. S., J. E. Wennberg, Th. A. Stukel, and S. M. Sharp. 1994. "Hospital Readmission Rates for Cohorts of Medicare Beneficiaries in Boston and New Haven." *New England Journal of Medicine* 331: 989-95.
- Freidson, E. 1975. *Profession of Medicine. A Study of the Sociology of Applied Knowledge*. New York: Dodd, Mead & Company.
- Geller, S. E., L. R. Burns, and D. J. Brailer. 1996. "The Impact of Nonclinical Factors on Practice Variations: The Case of Hysterectomies." *Health Services Research* 30 (6): 729-50.
- Griffiths, M., W. E. Waters, and E. D. Acheson. 1979. "Variation in Hospital Stay after Inguinal Herniorrhaphy." *British Medical Journal* 6166 (1): 787-9.
- Hainsworth, D. S., E. Lockwood-Cook, M. Pond, and R. J. Lagoe. 1997. "Development and Implementation of Clinical Pathways for Stroke on a Multihospital Basis." *Journal of Neuroscience Nursing* 29 (3): 156-62.

- Hays, W. L. 1973. *Statistics for the Social Sciences*, 2th edition. London: Holt, Rinehart and Winston.
- de Jong, J. D., P. P. Groenewegen, and G. P. Westert. 2003. "Mutual Influences of General Practitioners in Partnerships." *Social Science and Medicine* 57 (8): 1515–24.
- de Jong, J. D., G. P. Westert, C. M. Noetscher, and P. P. Groenewegen. 2004. "Does Managed Care Make a Difference? Physicians' Length of Stay Decisions under Managed and Non-Managed Care." *BMC Health Services Research* 4: 3.
- Kroneman, M. 2001. *Healthcare Systems and Hospital Bed Use*. Utrecht: Nivel.
- Lagoe, R. J. 1998. "Benchmarking Hospital Admission Rates by Ages." *Healthcare Benchmarks* 5 (7): 104–12.
- Lagoe, R. J., C. M. Noetscher, and M. E. Murphy. 2001. "Hospital Readmission: Predicting the Risk." *Journal of Nursing Care Quality* 15 (4): 69–83.
- Leyland, A. H., and P. P. Groenewegen. 2003. "Multilevel Modelling and Public Health Policy." *Scandinavian Journal of Public Health* 31: 267–74.
- McNeil, B. J. 2001. "Hidden Barriers to Improvement in the Quality of Care." *New England Journal of Medicine* 345: 1612–20.
- Miller, M. E., W. P. Welch, and H. G. Welch. 1996. "The Impact of Practicing in Multiple Hospitals on Physician Profiles." *Medical Care* 34: 455–62.
- Murphy, M., C. M. Noetscher, and R. J. Lagoe. 1999. "A Multihospital Effort to Reduce Inpatient Lengths of Stay for Pneumonia." *Journal of Nursing Care Quality* 13 (5): 11–23.
- Noetscher, C. M., R. J. Lagoe, V. K. Hohner, and G. M. Schmidt. 1999. "Analyzing Hospital Readmissions Using Statewide Data Bases." *Journal of Nursing Care Quality* 13 (6): 57–67.
- Noetscher, C. M., G. F. Morreale, and R. J. Lagoe. 2001. "Length of Stay Reduction: Two Innovative Hospital Approaches." *Journal of Nursing Care Quality* 16 (1): 1–14.
- O'Connor, G. T., H. B. Quinton, N. D. Traven, L. D. Ramunno, T. A. Dodds, T. A. Marciniak, and J. E. Wennberg. 1999. "Geographic Variation in the Treatment of Acute Myocardial Infarction: The Cooperative Cardiovascular Project." *Journal of the American Medical Association* 281: 627–33.
- Pasley, B. H., R. J. Lagoe, and N. O. Marshall. 1995. "Excess Acute Care Bed Capacity and Its Causes: The Experience of New York State." *Health Services Research* 30 (1): 115–31.
- Paul-Shaheen, P., J. D. Clark, and D. Williams. 1987. "Small Area Analysis: A Review and Analysis of North American Literature." *Journal of Health Politics Policy and Law* 12: 741–809.
- Pescosolido, B. A. 1992. "Beyond Rational Choice: The Social Dynamics of How People Seek Help." *American Journal of Sociology* 97 (4): 1096–138.
- Powell, A., H. Davies, and R. Thomson. 2004. "Using Routine Comparative Data to Assess the Quality of Health Care: Understanding and Avoiding Common Pitfalls." In: *Quality Improvement Research. Understanding the Science of Change in Health Care*, edited by R. Grol, R. Baker, and F. Moss, pp. 29–50. London: BMJ Publishing Group.

- Read, J. L., R. S. Stern, L. A. Thibodeau, D. E. Jr. Greer, and H. Klapholz. 1983. "Variation in Antenatal Testing over Time and between Clinic Settings." *Journal of the American Medical Association* 249: 1605–9.
- Scott, W. R., M. Ruef, P. J. Mendel, and C. A. Caronna. 2000. *Institutional Change and Healthcare Organizations. From Professional Dominance to Managed Care*. London: The University of Chicago Press, Ltd.
- Shortell, S. M. 1991. *Effective Hospital-Physician Relationships*. Ann Arbor, MI: Health Administration Press Perspectives.
- Shortell, S. M., R. R. Gillies, D. A. Anderson, K. M. Erickson, and J. B. Mitchell. 2000. *Remaking Health Care in America*. San Francisco: Jossey-Bass.
- Snijders, T. A. B., and R. J. Bosker. 1999. *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. London: SAGE Publications Ltd.
- Thom, D. H. 2000. "Training Physicians to Increase Patient Trust." *Journal of Evaluation in Clinical Practice* 6 (3): 245–53.
- Weide, M. G., G. P. Westert, J. S. A. G. Schouten, and P. P. Groenewegen. 1992. *Praktijkvariatie en gepaste zorg (Practice Variation and Appropriate Care)*. Utrecht: NIVEL/RIVM.
- Wennberg, J. E. 1999. "Understanding Geographic Variations in Health Care Delivery." *New England Journal of Medicine* 340: 52–3.
- Wennberg, J. E., and A. Gittelsohn. 1975. "Health Care Delivery in Maine I: Patterns of Use of Common Surgical Procedures." *The Journal of the Maine Medical Association* 66 (5): 123–49.
- . 1982. "Variations in Medical Care among Small Areas." *Scientific American* 246: 120–34.
- Westert, G. P. 1992. *Variation in Use of Hospital Care*. Assen: Van Gorcum.
- Westert, G. P., and P. P. Groenewegen. 1999. "Medical Practice Variations: Changing the Theoretical Approach." *Scandinavian Journal of Public Health* 27: 173–80.
- Westert, G. P., and R. J. Lagoe. 1995. "The Evaluation of Stays for Total Hip Replacement." *Quality Management in Health Care* 3 (3): 62–71.
- Westert, G. P., A. P. Nieboer, and P. P. Groenewegen. 1993. "Variation in Duration in Hospital Stay between Hospitals and between Doctors within Hospitals." *Social Science & Medicine* 37 (6): 833–9.
- Wonnacott, T. H., and R. J. Wonnacott. 1990. *Introductory Statistics*, 5th edition. New York: Wiley.

SUPPLEMENTARY MATERIAL

The following supplementary material for this article is available online:

APPENDIX A. The Multilevel Model.

APPENDIX B. Multilevel Model with Covariance Estimation.