Outcome And Cost Of Family Physicians' Care — Pilot Study Of Three Diagnosis-Related Groups In Elderly Inpatients

Duncan S. MacLean, MD

Background: **Health care analysts have speculated whether a primary-care–based delivery system would improve outcomes and costs for the currently preponderantly specialist-directed United States system. This pilot study compares outcomes and costs of family physicians' care with those of other physicians' care for three diagnosis-related groups (DRGs) and shows the feasibility of a larger scale research program using the Pennsylvania MedisGroups data base.**

Methods: **A cross-sectional comparison of outcomes and costs of hospital care given by either family physicians or other physicians was conducted using patients aged 65 years and older who were hospitalized during 1990. The study population was 2420 inpatients: 847 with DRG 174 (gastrointestinal bleeding), 628 with DRG 243 (medical back pain), and 945 with DRG 296 (metabolic disorders). Patients were stratified into five admission severity groups for each DRG. Comparison variables were MedisGroups major morbidity classification, in-hospital mortality, mean length of stay, and total charges.**

Results: **The family physician group delivered inpatient care to the elderly with gastrointestinal bleeding at significantly less cost than other physicians without any compromise of effectiveness. There was a similar but less striking trend for medical back pain. Results were variable for metabolic disorders.**

***Conclusions*:** **The results of this pilot study apply only to the three DRGs in the region at the time investigated. Generalizations should be avoided before additional DRGs are studied throughout the rest of the state. This pilot study demonstrates the feasibility of such research. (J Am Board Fam Pract 1993; 6:588-93.)**

The current debate regarding health care reform has reached every region and sociopolitical segment of our country.[1] Several editorialists have suggested that health care delivery systems should receive the same attention from reformers that financing systems do. They have expressed the belief that care directed by primary care generalists would be more efficient and effective than the care currently provided by our preponderantly specialist-directed system.[2-4] Until recently, this belief has been based on an intuitive opinion supported by only sparse, indirect data.[5] A notable exception are the data from a large cross-sectional study in Los Angeles, Chicago, and Boston, which reported significantly higher medical resource utilization by specialists than by generalists, even when adjusted for severity case mix.[6] These circumstances have prompted calls for further outcome research on the effectiveness of primary care.[7]

Outcome research has been facilitated by the emergence of large public data bases.[8] One such data base was established in Pennsylvania by the Health Care Cost Containment Act of 1987, which requires all hospitals in the state to monitor their care using the MediQual System MedisGroups quality management software program. This data base was the first to be implemented on a statewide basis and was the first to normalize outcomes against admission severity. I used the Pennsylvania MedisGroups data base to measure the comparative efficiency and effectiveness of family physicians in the southcentral region of the state. This study demonstrates the potential use-

Submitted, revised, 3 May 1993.

From the Department of Extended Care and Geriatrics, Polyclinic Medical Center, Harrisburg, Pennsylvania. Address reprint requests to Duncan S. MacLean, MD, Department of Extended Care and Geriatrics, Polyclinic Medical Center, 2601 North Third Street, Harrisburg, PA 17110-2098.

Supported by a grant from the Pennsylvania Academy of Family Physicians.

fulness of this data base for an expanded future program of outcome research.

Methods

A cross-sectional comparison of outcomes of hospital care given by either family physicians or all other specialties was conducted using patients admitted during 1990 to hospitals in Pennsylvania region 5 for any of three diagnosis-related groups (DRGs). DRGs selected for this pilot study were gastrointestinal bleeding (DRG 174), medical back pain (DRG 243), and nutritional, metabolic, and dehydration disorders (DRG 296). These DRGs had shown a large difference in outcomes among various clinical departments in a preliminary quality-assurance study of high-frequency diagnoses carried out in a single hospital (unpublished data).

MedisGroups Quality Management System

This study used the MedisGroups data base compiled by the Pennsylvania Health Care Cost Containment Council (PHC4). Each hospital in Pennsylvania is by statute required to abstract charts for all admissions using technicians who have been certified as having an accuracy of 95 percent or greater by MediQual, Inc. Abstracters extract objective indicators, called key clinical findings, on admission and at periodic intervals thereafter. Key clinical findings are physical signs, vital signs, and laboratory values correlated with potential for organ system failure and are described in more detail elsewhere.[5] Their validity in predicting mortality has been generally, but not universally, accepted.[9]

The MedisGroups system assigns patients into admission severity groups ranging from 0 (least severe) through 4 (most severe) based on key clinical findings present on admission, using a complex patented algorithm for each DRG. Similarly, the system determines the presence of major morbidity at the first interval review (usually 5 to 7 days after admission) based on persistence of key clinical findings. Thus, according to MediQual, major morbidity is an inverse indicator of the effectiveness of treatment in stabilizing the patient against organ system failure.

For this study, variables measured for the family physician and other physician groups were MedisGroups major morbidity classification, in-hospital mortality, mean length of stay, and total charges from audited hospital financial statements.

Hospitals and Admitting Physicians

This pilot study was carried out in PHC4 region 5, an 11-county area of southcentral Pennsylvania in which the largest standard metropolitan statistical area is Harrisburg-Lebanon-Carlisle, with a population of 583,700 (1987 estimate). The 17 general hospitals in PHC4 region 5 vary in size, proportion of Medicare and Medicaid discharges, and academic category (Table 1).

A list was compiled that was composed of 614 physicians who were active members of the Pennsylvania Academy of Family Physicians or Diplomates of the American Board of Family Practice or both. License numbers could be obtained from the Commonwealth of Pennsylvania Medicine Licensing Board for 437 (74.2 percent); this set of physicians constituted the family physician study group. The 153 family physicians for whom license numbers could not easily be obtained were included in the group of all other physicians admitting patients to region 5 hospitals and together constituted the control group. It was beyond the scope of this study to characterize further the control group by number of general practitioners, general internists, or subspecialists.

Subjects

The study considered all Medicare patients admitted to region 5 general hospitals during the study year and assigned to DRG 174, DRG 243 and DRG 296. The following numbers of patients

Table 1. Profile of Hospitals in Pennsylvania Region 5 (N = 17).

Characteristics	Number of Hospitals
Number of beds	
100–199	7
200–399	6
400–600	4
Percentage of Medicare discharges	
25–29	2
30–39	9
40–50	6
Percentage of Medicaid discharges	
2–9	10
10–14	5
15–18	2
Academic category	
Nonteaching	9
Teaching*	7
University	1

*ABMS-certified or AOA-certified postgraduate training hospital.

were in each group for family physicians and all others, respectively: DRG 174 had 168 patients admitted by family physicians and 679 by all others, DRG 243 had 68 patients admitted by family physicians and 560 by all others, and DRG 296 had 137 patients admitted by family physicians and 808 by all others. The one unusable record in the PHC4 MedisGroups data base was discarded.

Statistical Analysis

Statistical analysis was carried out using the Statistical Package for the Social Sciences, Personal Computer version.[10] The Student 2-tailed t-test, assuming separate variance estimate, was used for continuous variables (total charges, length of stay). The Pearson chi-square test was used for discrete variables (morbidity, mortality). The significance level for all analyses was set at $P < 0.05$.

Sensitivity Analysis

Because 25.8 percent of family physicians' data could not be separated from that of the control group, sensitivity analyses were carried out. Unidentified family physicians' outcomes were alternatively assumed to be the same as those for either family physicians or other physicians. Subtracting these results from other physicians' results and adding them to family physicians' results changed the magnitude of differences observed but did not alter the pattern of the results. In addition, because the university hospital in region 5 admitted no patients to family physicians, comparisons were carried out eliminating its data. This manipulation resulted in small changes in magnitude but not in pattern or statistical significance for other physicians' data.

Results

For DRG 174 (gastrointestinal bleeding), family physicians showed a significantly lower total charges for severity groups 0, 1, and 2 with no difference in mortality (Table 2). Total charges per case for family physicians ranged from $2,950 to $7,963, depending on the admission severity group, with differences favoring the family physician group ranging from $763 to $3,197 per case. Length of stay was lower for family physicians for severity groups 0, 1, 2, and 3 by 0.69 to 1.62 days but reached statistical significance only for severity group 1.

For DRG 243 (medical back pain), family physicians showed a less striking trend toward lower total charges, with variable differences in mortality and length of stay (Table 3). There was 1 mortality for family physicians' patients among the 4 in severity group 3 compared with none of 16 for other physicians, a statistically significant difference.

For DRG 296 (nutritional, metabolic, and dehydration disorders) family physicians showed a shorter length of stay and total charges for severity groups 2 and 3 but higher in length of stay and total charges for severity groups 0 and 1, with only the total charges for severity group 3 reaching statistical significance favoring family physicians (Table 4). There was a trend toward higher mortality for family physicians' patients in severity groups 3 and 4.

There were no significant differences in percentage of major morbidity classification between family physicians and other physicians for any DRGs or severity groups, with the exception of severity group 0 for DRG 296 (metabolic disor-

Table 2. Comparison of Outcome and Cost of Care by Family Physicians and Other Physicians for Patients in DRG 174 (Gastrointestinal Bleeding).*

	Family Physicians (n = 105)				All Other Physicians (n = 289)			
Admission Severity	No. Patients	Mortality (%)	Length of Stay (days ± S.D.)	Total Charges ($± S.D.)	No. Patients	Mortality (%)	Length of Stay (days ± S.D.)	Total Charges ($ ± S.D.)
0	6	0	4.67 ± 1.03	2,950† ± 755	25	0	5.36 ± 3.53	4,178 ± 2,320
1	36	3	4.64† ± 2.75	3,321† ± 1,965	103	2	6.26 ± 4.34	4,539 ± 4,126
2	77	3	6.60 ± 4.08	5,124† ± 3,798	319	4	7.65 ± 4.89	6,324 ± 4,326
3	44	9	8.55 ± 5.46	7,963 ± 6,062	206	16	9.35 ± 6.53	8,726 ± 6,456
4	5	40	6.00 ± 4.53	6,055 ± 5,336	26	50	5.85 ± 5.65	9,252 ± 9,256
Total	168				679			

*Admission severity = MedisGroups admission severity group. S.D. = standard deviation.
†$P < 0.05$.

Table 3. Comparison of Outcome and Cost of Care by Family Physicians and Other Physicians for Patients in DRG 243 (Medical Back Pain).*

	Family Physicians (n = 47)				All Other Physicians (n = 260)			
Admission Severity	No. Patients	Mortality (%)	Length of Stay (days ± S.D.)	Total Charges ($± S.D.)	No. Patients	Mortality (%)	Length of Stay (days ± S.D.)	Total Charges ($ ± S.D.)
0	30	0	7.17 ± 4.00	3,563 ±2,257	298	0.3	5.42† ± 4.64	3,550 ± 2,944
1	24	0	8.58 ± 5.39	4,498 ± 3,273	173	0	8.36 ± 6.26	5,222 ± 5,858
2	10	10	8.20 ± 6.00	5,357 ± 3,394	72	1	8.22 ± 5.34	5,427 ± 3,344
3	4	25	8.50 ± 6.25	5,599 ± 5,136	16	0†	10.63 ± 8.17	7,471 ± 6,038
4	0	—	—	—	1	0	11.0	8,055
Total	68				560			

*Admission severity = MedisGroups admission severity group. S.D. = standard deviation.
†$P < 0.05$.

ders). Family physicians had 11 percent of their patients in the major morbidity category (1 out of 9 patients) compared with 0 percent out of 55 patients for other physicians, which reached statistical significance.

Conclusion

Family physicians in region 5 treated elderly inpatients with DRG 174 (gastrointestinal bleeding) for probably a lower cost than did other physicians in 1990 without any apparent compromise of effectiveness. There was a similar but less striking trend for DRG 243 (back pain), which was not statistically significant but was financially significant.

Discussion

Reduced utilization of health care resources by family physicians was observed for two of the three DRGs studied. These results are consistent with the hypothesis that the practice style of these family physicians, not just their patients' severity of illness, accounts for their comparative efficiency.

This study is properly understood as a system analysis that applies only to the region and time studied. Characteristics could fluctuate between regions and even over time within the same region, based on changes in the many factors influencing hospital care. Thus, results do not reflect immutable laws of nature, such as investigated by epidemiology and clinical sciences.

The greatest strength of this pilot study lies in its capture of data for 100 percent of admissions in each DRG through its use of the Pennsylvania public data base. Another strength is stratification of PHC4 data by admission severity. The MedisGroups average weighted severity score is a reliable and valid indicator of mortality risk according to most, but not all, investigators[11] and thus served as a crude but satisfactory control for severity.

By contrast, the MedisGroups methodology has not been externally validated for its in-hospital major morbidity classification. It is not known how this measure correlates with other more widely accepted outcome measures, such as functional status at discharge, subjective well-being, and readmissions within 30 days.

Caution must be used in interpreting total charge data. Charges published by many hospitals correlate poorly with true costs. Total charge data can be used reliably to compare costs within a single hospital, but aggregating such data from different hospitals, such as in the present study, introduces the possibility of misleading results. Large differences in aggregate charge data, however, such as those observed in this study for DRG 174 (gastrointestinal bleeding), are probably valid. Hospital-by-hospital comparisons that should be carried out before drawing final conclusions were beyond the scope of this study.

Statistical calculations carried out as part of systems analyses of quality, such as this study, serve a purpose different from those conducted as part of clinical epidemiology investigations. In the latter types of studies, statistical analysis determines the likelihood that a random sampling error accounts for observed differences among groups. No such error was possible in this study, because it measured the entire data universe within the region studied; differences in length of stay and total charges measured in this study are absolute. Another difference is that statistical

Table 4. Comparison of Outcome and Cost of Care by Family Physicians and Other Physicians for Patients in DRG 296 (Nutritional, Metabolic, Dehydration Disorders).*

	Family Physicians (n = 81)				All Other Physicians (n = 338)			
Admission Severity	No. Patients	Mortality (%)	Length of Stay (days ± S.D.)	Total Charges ($± S.D.)	No. Patients	Mortality (%)	Length of Stay (days ± S.D.)	Total Charges ($ ± S.D.)
0	9	11	9.56 ± 7.18	6,015 ±4,600	55	2	8.42 ± 7.00	5,651 ± 4,934
1	31	0	8.61 ± 8.81	5,346 ± 7,245	176	5	7.59 ± 6.12	4,841 ± 3,719
2	53	11	8.60 ± 6.10	5,364 ± 4,242	330	11	9.01 ± 8.65	5,864 ± 6,185
3	37	22	8.62 ± 6.09	5,523† ± 4,336	209	19	11.14 ± 15.00	8,747 ± 15,488
4	7	71	7.43 ± 6.13	8,247 ± 8,333	38	53	9.79 ± 8.48	8,287 ± 6,331
Total	137				808			

*Admission severity = MedisGroups admission severity group. S.D. = standard deviation.
†$P < 0.05$.

methods of clinical epidemiology assume that study populations are relatively homogenous and similar except with regard to the variables being measured. By contrast, the group of hospital admissions in this study were variegated and complex. For example, under each DRG were included many different diagnoses. Outcomes and charges could have been influenced by many other factors, such as hospital systems, clustering of higher severity cases by specialty or hospital, interplay between outpatient and inpatient management, and presence of chronic disease. It would be fallacious to impute universal validity to the results of this study based on statistical significance of observed differences, because the systems studied did not necessarily conform to the assumptions of statistical methods.[12]

Descriptive statistical methods are increasingly being applied in total quality management both in ambulatory care settings[13] and in hospitals.[14] In quality management, attention is focused as much on variations from the mean within groups as on the difference in means between groups. Shewhart[15] has pointed out that as quality improves in industrial processes, larger assignable variations (attributable to special causes of poor quality) are eliminated and only smaller predictable random variations remain. In this study, family physicians had smaller standard deviations in length of stay and in total charges for both gastrointestinal bleeding (DRG 174) and back pain (DRG 243), and this tighter clustering near the means might correlate with lower assignable variations in quality. The limitations of this pilot study did not permit a full analysis of random and assignable variations, however.

This study raises many questions. For instance, do similar differences exist in the other eight regions of Pennsylvania, in small compared with large hospitals, or in urban compared with rural hospitals? For which other DRGs do family physicians show lower charges, and for which, higher? When greater numbers are aggregated over time for metabolic disorders (DRG 296), do family physicians have a worse rate of mortality and morbidity than other physicians, as suggested by the trends in this small study, or do the rates regress toward the mean? Which hospitals, group practices, and health maintenance organizations show the lowest charges and which, the highest? Do any differences in charges correlate with poorer mortality, morbidity, or length of stay? Which hospitals, subspecialties, and subgroups show best results for DRGs requiring more specialized technological care? What are the results for younger, non-Medicare patients? What are the results for 1991 and subsequent years?

This pilot study answers none of these questions because of its limited scope. Nevertheless, it shows that information about efficiency and effectiveness can be obtained using the Pennsylvania MedisGroups data base. Other public data bases are being developed with even greater degrees of sophistication and can serve to validate conclusions.

Large data bases would be most useful for internal quality management by institutions and health delivery systems. They could also be used to track the performance of recent graduates of residency training programs as a measure of their programs' educational effectiveness. The data bases could help determine which specialties or other physician subgroups achieve the most effective, efficient care and thereby could help establish objective standards of quality.

Only if similar trends are found among many data bases can any generalizations be justified. For example, Billings and Hasselblad,[16] using other data bases, have described a set of "ambulatory-care sensitive" hospital admission diagnoses that occur at rates inversely proportional to the availability of primary medical care services. It is probable that certain DRGs will eventually be shown to be geriatric sensitive among many data bases, that is, most efficiently managed using a biopsychosocial-functional approach. DRG 174 (gastrointestinal bleeding) might be one such geriatric-sensitive DRG.

Information about many such questions lies in the Pennsylvania MedisGroups data base. It should be used to conduct further outcome research.

This study was supported by a grant from the Pennsylvania Academy of Family Physicians. J. Marvin Bentley, PhD, assisted with research design. Jill Boyd, Barbara Dobie, RN, Regina Luto, and Michael Markowski assisted with data processing. A version of this paper was presented to the Eleventh Annual Northeast Regional Meeting of the Society of Teachers of Family Medicine, 2 October 1992.

References

1. Lundberg GD. National health care reform. An aura of inevitability is upon us. JAMA 1991; 265:2566-7.
2. Brody H. A policy imperative for primary care: reflections on Keystone II. Fam Med 1990; 22:42-5.
3. Relman AS. Reforming the health care system. N Engl J Med 1990; 323:991-2.
4. Petersdorf RG. Primary care applicants — they get no respect. N Engl J Med 1992; 326:408-9.
5. McGann KP, Bowman MA. A comparison of morbidity and mortality for family physicians' and internists' admissions. J Fam Pract 1990; 31:541-5.
6. Greenfield S, Nelson EC, Zubkoff M, Manning W, Rogers W, Kravitz RL, et al. Variations in resource utilization among medical specialties and systems of care. Results from the medical outcomes study. JAMA 1992; 267:1624-30.
7. Politzer RM, Harris DL, Gaston MH, Mullan F. Primary care physician supply and the medically underserved. A status report and recommendations. JAMA 1991; 266:104-9.
8. Caper P. The epidemiologic surveillance of medical care. Am J Public Health 1987; 77:669-70.
9. Iezzoni LI, Meskowitz MA. A clinical assessment of MedisGroups. JAMA 1988; 260:3159-63.
10. Norusis MJ. Manual for SPSS/PC: machine-readable data file for the IBM-XT. Chicago: SPSS, Inc., 1984.
11. Iezzoni LI, Ash AS, Cobb JL, Moskowitz MA. Admission MedisGroups score and the cost of hospitalizations. Med Care 1988; 26:1068-80.
12. Weinberg GM. An introduction to general systems thinking. New York: John Wiley & Sons, 1975:19-26.
13. Kibbe DC, Bentz E, McLaughlin CP. Continuous quality improvement for continuity of care. J Fam Pract 1993; 36:304-8.
14. Finison LJ, Finison KS, Bliersbach CM. The use of control charts to improve healthcare quality. J Healthcare Quality 1993; 15:9-23.
15. Shewhart WA. Economic control of quality of manufactured product. New York: D. Van Nostrand, 1931.
16. Billings J, Hasselblad V. A preliminary study: use of small area analysis to assess the performance of the outpatient delivery system in New York City. Lebanon, NH: The Codman Research Group, 1989.