United States General Accounting Office

GAO

Report to Congressional Requesters

October 1994

HEALTH CARE

Employers Urge Hospitals to Battle Costs Using Performance Data Systems

GAO/HEHS-95-1

GAO

United States
General Accounting Office
Washington, D.C. 20548

Health, Education, and
Human Services Division

B-252283

October 3, 1994

The Honorable Michael A. Andrews
House of Representatives

The Honorable Jim Cooper
House of Representatives

The Honorable Charles W. Stenholm
House of Representatives

Many large employers have become increasingly concerned about the wide variation in hospital costs across their communities. They believe that they may be paying for care that is not being delivered in the most efficient manner. At the same time, they contend that they lack the information to assess the value of health care services they are paying for. To meet this need, employers in some communities—such as Walt Disney World in Orlando, Proctor and Gamble in Cincinnati, and LTV Steel in Cleveland—have organized health care coalitions to help them make better purchasing decisions. By developing and sharing comparable information on hospital performance in their communities, they hope to make health care providers more accountable for the services they deliver.

Because of growing interest in obtaining comparative information about health care providers, you asked us to review the experience of communities in which employer coalitions have encouraged hospitals to adopt automated performance measurement systems. Specifically, we examined (1) the purposes for which employer coalitions and hospitals are using comparative performance measurement systems and (2) whether these systems report the information employers and hospitals need to compare outcomes.

Background

In general, performance measurement systems generate large computerized databases that compare providers treating similar patients. They produce data on various indicators of hospital performance—typically how long patients stay in the hospital (length of stay), death rates by category of condition (mortality rates), and treatment charges. Earlier approaches to measuring a provider's performance did not account for the fact that patients of the same age, sex, and general diagnosis could have different outcomes after surgery because of the severity of the patients' conditions. For example, a 55-year-old woman

admitted for colon cancer with an intestinal obstruction from a large colon tumor is not likely to have as favorable an outcome as a patient of the same age and sex with only superficial colon cancer. Hospitals contended that because the systems did not adjust for patients' severity of condition, charge and outcome comparisons were invalid.[1] Hospitals faring poorly in comparison with other hospitals claimed their patients were sicker and thus required more costly care and had an increased likelihood of poor outcomes. To address this concern, a number of private vendors have developed computerized data systems designed to account for differences in patients' severity of illness.

Cincinnati, Cleveland, and Orlando are cities in which employer coalitions selected particular vendors and encouraged the local hospitals to purchase their severity-adjusted performance measurement systems. Cincinnati hospitals use Iameter's Acuity Index Method (AIM). Cleveland hospitals use three instruments: (1) the Cleveland Health Outcome Indicator of Care Evaluation (CHOICE) system for general medical and surgical admissions, designed by Michael Pine and Associates; (2) Apache Medical System's Acute Physiology, Age, Chronic Health Evaluation (APACHE III) system for intensive care admissions; and (3) the Patient Viewpoint Survey, a patient satisfaction survey administered by NCG Research. Orlando hospitals use MediQual's Medical Illness Severity Grouping System (MedisGroups II). Table 1 presents the basic characteristics of the systems used in each community.

[1]In 1986, the Health Care Financing Administration (HCFA) started releasing hospital mortality rates as an indicator of hospitals' quality of care. Hospitals complained, however, that HCFA mortality data were unreliable because they did not accurately adjust for differences in patients' severity of illness. HCFA has continued to refine its methods, but in 1993 delayed release of hospitals' mortality data because of concerns over the accuracy of the severity adjustments.

Table 1: Characteristics of Performance Measurement Systems Used in Three Communities

	Cincinnati	Cleveland	Orlando
Systems used	AIM	CHOICE, APACHE, and Patient Viewpoint Survey	MedisGroups
Type and source of data collected	Administrative data retrieved from patient discharge forms	Clinical data abstracted from medical records and survey of a sample of patients	Clinical data abstracted from medical records
Performance indicators reported	Charges Length of stay Mortality rate	Length of stay Mortality rate Patient satisfaction	Charges Length of stay Morbidity rate[a] Mortality rate
Level of clinical detail reported	Major diagnostic categories and specific high-volume diagnoses	Types of services (intensive care, general medical, and general surgical) and major diagnostic categories	Specific diagnoses
Level of provider detail reported	Hospital and physician	Hospital only	Hospital and physician

[a]MedisGroups defines morbidity as occurring if a second severity rating taken midstay indicates that the patient has not responded to treatment.

Systems based on administrative data, as used in Cincinnati, and clinical data, as used in Orlando and Cleveland, are fundamentally different. To make severity adjustments, administrative systems use information on patient characteristics (age and sex), diagnoses (primary and secondary), and procedures performed. Because hospitals routinely develop administrative data for billing purposes, little additional cost is associated with collecting these data as inputs to performance measurement systems. In contrast, clinically based systems use more medically precise data, such as laboratory test values, radiology reports, and physiology readings (patient temperature, blood pressure, and so on), to make severity adjustments. As a result, clinically based systems require trained staff to review and abstract each medical record and are therefore more expensive than administrative systems.[2]

[2]The burden of collecting clinical data can be simplified if medical records are maintained electronically. The Cleveland Health Quality Choice Program reported that it is beginning to develop an electronic format for collecting laboratory test values and other clinical data, but currently most medical records are manually abstracted.

All of the systems are confined to assessing care for patients admitted to the hospital. None is able to assess care provided in ambulatory settings—such as physicians' offices, hospital outpatient clinics, or freestanding laboratories; such care accounts for more than half of most employers' health care costs. Also, the systems do not identify whether a hospital admission is appropriate, although some hospital admissions may be unnecessary or care could be more effectively provided elsewhere. Finally, none of the systems provides a single summary "grade" for hospitals to highlight that hospitals perform well for some types of care but not for others.[3]

Scope and Methodology

To address our objectives and in consultation with your staffs, we selected Cincinnati, Cleveland, and Orlando as case studies. Although hospitals nationwide have had experience with performance measurement systems, the selection and use of severity-adjusted performance measurement systems in each of these cities resulted from an employer coalition initiative. In addition, each community adopted different systems, reflecting a variety of hospital data systems currently available. Because of the small number of sites analyzed and the unique circumstances of each, the experience of these three cities may not be representative of performance measurement projects elsewhere in the nation.

In each community, we interviewed employer coalition directors, individual employers' health benefit managers, and hospital and physician associations' executives. We also visited several hospitals in each city that represented a range of sizes and types—including community and teaching hospitals—and met with hospital staff, including administrative officers and physicians serving as clinical department chiefs.[4] We contacted the principal developers of the severity adjustment systems and met with other experts in the field of outcomes research.

Finally, to develop trend data, we sought data on costs, charges, and length of stay for specific medical conditions from each hospital participating in the three community projects. However, difficulties with access, completeness, and timing limited our use of this information. We surveyed each hospital for information on costs, charges, and payments

[3]For a review of initiatives to provide summary information comparing health plans' costs and quality of care, see Health Care Reform: "Report Cards" Are Useful But Significant Issues Need to Be Addressed (GAO-HEHS-94-219, Sept. 29, 1994).

[4]To obtain the full cooperation of hospitals involved in these initiatives, we provided pledges of confidentiality. Therefore, we have not identified hospitals by name in this report.

from 1988 to 1993, but we were unable to use these data because of partial and inconsistent responses. We also obtained hospital data on charges and length of stay collected by the Florida and Ohio governments. However, these data were available only through 1992, and in Ohio the data only included privately insured patients (Medicare and Medicaid patients were excluded) for the 100 highest-volume diagnosis related groups (DRG) reported by the hospital. Therefore, we were often restricted to anecdotal information provided during our interviews that could not be verified. We conducted our review between May 1993 and July 1994 in accordance with generally accepted government auditing standards.

Results in Brief

In the communities we reviewed, the introduction of severity-adjusted performance measurement systems has given hospitals an impetus to initiate efficiency improvements. Employers plan to use the information produced by these systems to steer patients to cost-effective providers, while hospitals can use the information to help manage treatment costs. Because employer coalitions have only recently encouraged communitywide use of the systems, the cost-saving potential of the initiatives remains an unresolved issue. Also, in light of data quality concerns, employers should use caution in interpreting systems' results.

Employer coalitions in Cincinnati, Cleveland, and Orlando have made severity-adjusted performance measurement systems an important element of their communities' cost containment strategies. That is, employers will direct or encourage their employees to hospitals that demonstrate superior cost performance without compromising quality of care. By influencing employer health care purchasing decisions, performance measurement systems may increase hospital accountability and competition.

In the communities we examined, hospitals generally regard the systems as one of several useful internal tools for identifying efficiency problems. Many hospitals have attempted to reduce their average charge and length of stay by modifying their patterns of care. These efforts include (1) shifting certain services into ambulatory settings (where costs may be lower); (2) using fewer or less costly supplies and pharmaceuticals; and (3) reducing, substituting, or rescheduling treatments and services. Anecdotal evidence linking the performance measurement systems to cost reductions for certain services implies a potential to generate savings for individual hospitals. However, statistical evidence to date on the effectiveness of these systems as a sustainable, communitywide cost

containment tool is neither complete nor conclusive. This may be explained by (1) the fact that the systems have not been in use long enough to produce sufficient evidence of cost savings and (2) the difficulty in isolating the influence of these systems from other factors contributing to hospital improvements.

Hospitals, physicians, and experts in the field of outcomes research caution employers that the results of these systems should not be the sole guide for health care purchasing decisions. Although the systems allow hospitals to be compared more fairly than previous approaches because they adjust for severity of illness, the systems' reports are incomplete sources of information on costs and quality. In particular, concerns exist about the selection of performance indicators reported to employers and the validity of the severity adjustments applied to the data. Although the systems offer improved comparisons of mortality rates and length of hospital stays, many of the most meaningful patient outcomes for employers, such as functional status and time needed to return to work, are not documented by the systems. Furthermore, certain hospitals may still be disadvantaged in the comparisons because of their unique characteristics and limitations in the severity adjustment methodology applied.

Employers Use Performance Measurement Systems to Stimulate Hospital Efficiency

Vendors have recently begun to market their performance measurement systems to employer coalitions, stressing the systems' capability to highlight opportunities for aggregate hospital cost savings. Employers in Cincinnati, Cleveland, and Orlando have succeeded in prompting most local hospitals to use a performance measurement system selected by the employers. Many hospitals have incorporated the data from these systems in their management of clinical performance, and some hospitals have begun to make practice improvements that they attribute to the employer coalitions' initiatives. Because many hospitals have only recently adopted these systems and few employers have yet to use the comparative information on hospital performance to select providers for their health care networks, it is too early to fully assess the cost containment potential of these employer initiatives.

Coalitions Provide Strong Incentives for Hospitals to Improve Efficiency

Coalitions have sought to have all local hospitals participate in the performance measurement programs with the goal of comparing hospital performance across their communities. Employers in the coalitions represent a substantial share of the communities' patients and either

contract directly with hospitals or make selections through insurers' prearranged network plans. To increase hospital participation, the coalition employers announced that they would be reluctant to deal with hospitals that choose not to use the coalition-sponsored performance measurement system.

The threat of losing coalition employers' business has been a powerful incentive, prompting nearly all of the hospitals in the communities we visited to purchase the systems. In Cincinnati, where individuals insured by coalition employers constitute about 13 percent of the population, all the hospitals that the employers asked to participate bought the employer-sponsored system. Similarly, in Cleveland and Orlando, coalition members insure about 20 to 30 percent of the population, and most local hospitals purchased the coalition-sponsored systems.[5]

To obtain the hospitals' initial cooperation, employers in the Cincinnati, Cleveland, and Orlando coalitions agreed to refrain for several years from using comparative data to exclude hospitals from their provider networks. Ultimately, individual employers intend to compare hospitals' performance and select the most cost effective to participate in their managed care networks.

Hospitals Encouraged to Curb Charges and Stays

First-year data received by the employers indicated considerable variation among hospitals in charges and length of stay. For example, average severity-adjusted charges among the 14 Cincinnati hospitals in 1992 varied by about $900 for pregnancy cases, $3,300 for circulatory cases, $4,000 for respiratory cases, $4,300 for digestive-tract cases, and $4,600 for musculoskeletal cases. Similarly, average length of stay varied by about 0.5 day for pregnancy cases, 1.5 days for digestive-tract cases, 2 days for circulatory and respiratory cases, and 2.5 days for musculoskeletal cases.[6] In many cases, hospitals had inconsistent performance profiles, scoring well for some patient groups but poorly for others.

Similarly, the systems' data showed some variation in mortality rates. For example, the spring 1994 Cleveland Health Quality Choice report showed mortality rates for general medical patients indicating that 2 hospitals had

[5]Hospitals in some communities considered the costs of these systems to be substantial. Costs to hospitals varied widely depending on the characteristics of the system, size of the hospital, and amount of data requested.

[6]For respiratory care, for example, these ranges represent variations of 25 percent above and 26 percent below the average charge and 16 percent above and 12 percent below the average length of stay in Cincinnati hospitals.

more deaths than expected, 2 hospitals had fewer deaths than expected, and 25 hospitals were as expected. For intensive care patients, 18 of 29 Cleveland hospitals had the expected mortality rates and the other 11 had fewer deaths than expected.

Many hospitals we visited reported efforts to improve their performance results by modifying treatment patterns.[7] Generally, hospitals use the systems to guide "continuous quality improvement" efforts.[8] The systems help hospitals identify particular conditions or clinical departments needing improvement.[9] In addition, hospital officials noted that the systems are useful for demonstrating to physicians how their performance indicators differ from those of colleagues caring for comparable patients. These comparisons lead hospital administrators and clinical staff to conduct further reviews of medical records and discuss alternative care procedures. Hospitals we visited highlighted various practice changes they attributed in part to the use of the systems. The following are examples:

- **Performing more diagnostic tests and procedures in ambulatory settings**—Some hospitals reported reducing length of stay and inpatient charges by increasing their use of ambulatory services. Physicians are more often ordering diagnostic tests and procedures before hospitalization so that patients may be admitted on the day of surgery, thereby shortening hospital stays by as much as 1 or 2 days. Physicians are also reducing length of stay by increasing their use of ambulatory rehabilitation services, relying more on home health services and discharging patients earlier into other settings, such as skilled nursing facilities. Hospitals also reported renewed efforts to avoid hospital admissions by substituting ambulatory care.[10]

[7]In addition, hospitals and physicians have also made administrative changes to improve their performance results. By revising their documentation of patient care (medical records and discharge summaries), they may receive a higher severity-of-illness classification to justify high charges and length of stay.

[8]In many hospitals we visited, clinical teams are organized to create internal consensus among physicians as to what constitutes optimal care for patients with a given condition. They often develop standard treatment protocols (also referred to as care maps, treatment paths, or practice guidelines) for various conditions.

[9]Hospitals are also focusing attention on improving their care of patients with diagnoses of particular importance to local employers, such as maternity care and back pain. In addition to such focused efforts, several hospital officials noted evidence of a "halo" effect where nontargeted areas also benefit from process improvements initiated as a result of the systems.

[10]In some instances, this substitution increases the average length of stay because patients eligible for ambulatory care would have had relatively short inpatient stays, whereas patients who continue to receive hospital care have relatively long stays. Similarly, many low-cost surgeries may be performed in alternative settings, reducing total costs but raising the average cost of operations performed in the hospital.

- **Changing how physicians schedule and use ancillary personnel and services**—After learning that its actual average length of stay for patients admitted for knee joint replacement surgery was greater than expected, a Cleveland hospital determined that several physicians had been underutilizing physical therapy services. The physicians are now ordering physical therapy for these patients earlier and more often. The hospital increased the utilization of this particular service but expected to reduce costs through shorter hospital stays.
- **Restricting the purchase and use of certain supplies and pharmaceuticals**—In a Cincinnati hospital, all orthopedic surgeons are generally now using a single type of artificial hip device instead of choosing from among several available types. This allows the hospital to take advantage of volume discounts. An Orlando hospital noted that it had convinced physicians to use less expensive pharmaceuticals by pointing out that other physicians used those medications with similarly effective results. Another Orlando hospital estimated that substantial savings resulted from switching to a less expensive pulmonary treatment—using a simple metered dose inhaler rather than extensive therapy—for patients with respiratory problems.

Some changes hospitals have made in response to the performance data initiatives may have already achieved their full cost containment effect. For instance, once hospitals have eliminated presurgery days by performing diagnostic and preoperative procedures in an ambulatory setting, no further reduction in days prior to surgery is possible. Similarly, discounts achieved by purchasing pharmaceuticals and medical equipment from a single supplier may initially reduce prices but may not continue to affect the rate of cost growth. To achieve further savings, hospital officials said they would need to make more profound changes in their operations, such as reducing bed capacity and staff.

Also, some of the cost savings employers attribute to efficiency improvements in inpatient hospital care are partially offset by higher expenditures for ambulatory care. The hospitals we visited acknowledged that efforts to reduce charges and length of stay in inpatient settings have been accompanied by the greater use of ambulatory care (such as outpatient hospital departments, physicians' offices, and home care) for testing, visits, and procedures. Although ambulatory care costs less than hospital care, it is a large and growing expense for employers—but one that the systems do not reveal. Because the severity adjustment systems do not capture ambulatory care costs, when a hospital shifts care from an

inpatient to an ambulatory setting, the change registers as a complete savings instead of a partial savings.

Performance Comparisons Expected to Contribute to Changes in Market Structure

Employer coalitions as well as hospitals predict that the periodic performance ratings will contribute to changes in local hospital markets. Over time, use of comparative ratings could result in individual hospitals choosing to specialize in the services for which they score relatively well. For instance, a hospital that receives high ratings for cardiac services but low ratings for obstetric services could decide to concentrate on cardiac care and drop obstetric care. Such decisions may improve quality of care since hospitals often have better surgical outcomes for procedures performed at high volumes.[11] Some would also argue, however, that patients could lose prompt access to specialized services if fewer local hospitals were able to serve patients needing immediate care.

In addition, hospital officials expect to complement their strengths by forming affiliations. Under a larger integrated system, hospitals would try to achieve economies through volume while offering comprehensive services at convenient locations. They would also expect to be in a better position to negotiate with health care purchasers, which have also been organizing to gain negotiation advantage. This would also allow hospitals receiving lower ratings for some services to maintain their economic viability. Some predict, however, that consolidation taken to the extreme could result in reduced competition and thus less pressure to perform efficiently.

Too Soon to Identify Cost Savings Attributable to Systems

Evidence of the systems' ability to hold down hospital cost growth has been limited because of the short time these systems have been in place and the time it takes to identify and implement changes in practice. Also, it is difficult to isolate the influence of severity-adjusted performance data from other factors contributing to hospital improvements.

To some extent, efficiency gains may reflect a sentinel effect. That is, in anticipation of employers' examination of length of stay and charge data, hospitals began taking action before the systems' data were available. This effect may help explain why some of the changes hospitals attributed to

[11]For example, 11 of 14 studies reviewed by the Office of Technology Assessment found that lower volume of coronary artery bypass graft surgeries in hospitals was associated with higher mortality rates. See The Quality of Medical Care: Information for Consumers, OTA-H-386 (Washington, D.C.: U.S. Government Printing Office, June 1988).

the systems actually began before hospitals received or had time to analyze and act on the systems' results.

For example, one Cincinnati hospital submitted its 1991 annual data to Iameter in October 1992 and submitted its 1992 data in January 1993. The Iameter results showed significant improvement for the hospital between 1991 and 1992. However, during the 4 months between data submissions, the hospital would not have been able to analyze the initial Iameter data, identify and implement improvements based on the data, and have the changes reflected in the subsequent data. Therefore, the identified improvements in charges and length of stay must have resulted from changes made before the system data were available.

Another factor making it difficult to gauge the effect of performance systems on cost containment is that many hospitals already had continuous quality improvement programs designed to improve efficiency. Such programs seek to streamline administrative functions and modify clinical practice patterns. Some hospitals we visited stated that the performance measurement systems would enhance their quality improvement efforts by supplying valuable data. Other hospitals whose performance indicators improved stated that the performance information provided by the systems added little to existing quality improvement efforts.

Furthermore, hospitals and employers noted that increased enrollment in managed care plans may have contributed to slower cost growth. Under such plans, employers and insurers have become increasingly aggressive in negotiating reimbursement rates with hospitals. The advent of statewide reform efforts to stem rising health care costs may have also provided an incentive for hospitals to mitigate rate increases.

Employers Should Use Caution in Interpreting Performance Data

We found that severity-adjusted performance measurement systems are in a relatively early stage of development and may not provide adequate information for accurately comparing hospitals' performance. Limitations exist in the indicators that the systems report and the methods they use to adjust the data for severity of illness. Because of these limitations, additional information and methodological improvements are needed to provide more useful data on which to base purchasing decisions.

Systems' Results Not Sufficient for Measuring Hospital Performance

Used alone, performance indicators measured by these systems are insufficient for comparing cost effectiveness—a function of cost and quality of care. Pressure to reduce charges and length of stay could, in fact, provide an incentive for hospitals to restrict services. As a safeguard against potential adverse effects on quality, the systems generate information to monitor hospitals for signs of compromised care. However, the two performance indicators generated by the systems to measure quality—inpatient mortality rates and average length of stay—are considered too narrow to truly reflect quality. Similarly, the indicators used for assessing costs—charges and length of stay—are poor proxies for measuring resource consumption.

Inpatient mortality rates are not considered a good indicator of the quality of care most patients receive because only a small percentage of hospital stays result in death. For example, admissions of particular interest to employers, such as maternity and orthopedic cases, have extremely low expected mortality rates. Because death and major complications for such patients are rare, systems that measure variation in providers' mortality rates for such services will not yield very useful information about the quality of inpatient hospital care.[12]

Some system vendors contend that average length of stay serves as an indicator of both costs and quality. They assert that longer hospital stays tend to consume more resources and increase patients' risk of contracting hospital-induced infections or other complications. However, hospital representatives we interviewed noted that it is difficult to determine a clinically optimal length of stay. Some patients may benefit from longer-than-average hospitalizations. For example, patients who lack an appropriate environment for home care may benefit from extra days in the hospital. Ever shorter stays may, in fact, result in poorer quality care if patients are discharged before they are ready.

Average length-of-stay data are useful but are also limited as an indicator of resource consumption. Shorter stays reduce the cost of care because when patients spend less time in the hospitals, fewer hours of nursing care and fewer routine tests and treatments are generally provided. However, the cost savings are not proportionate to the reductions in days. The most expensive days are generally early in the stay when resource-intensive procedures are performed; the latter days—when charges represent mostly room and board—are less costly. In fact, for most conditions, the

[12]A maternal death during childbirth would be considered a sentinel event that warrants further hospital review regardless of the availability of the performance measurement systems.

final day of a hospital stay costs less than half the average cost for all days.[13] Therefore, if it is assumed that costs per day remain constant, estimates of savings based on length-of-stay reductions may overstate actual cost reductions.

Similarly, data on hospital charges are not precise indicators of efficiency.[14] A hospital's overhead costs, which are included in individual charges, are not distributed across services on the basis of the actual cost of the service.[15] Therefore, the rise or fall of charges for cardiac bypass surgery, for example, may have little or nothing to do with the efficiency with which this procedure is done.

In addition, charges do not necessarily represent what the health care purchaser actually pays. Several employers participating in the coalitions we visited had managed care plans receiving 20- to 40-percent discounts on hospital charges. Other plans negotiate flat rates per day or per case, regardless of the hospital's charge. Thus, charges may not be relevant because often the amount the purchaser actually pays depends on the discount or rate negotiated.

Many of the experts, employers, and providers we spoke with believe that other indicators should be developed for a more complete comparison of quality. To obtain additional information about quality of care for individual hospitals, inpatient data on readmission rates, infection rates, and complication rates could also be monitored. Other postdischarge patient outcomes, such as 30- and 60-day mortality rates, functional status, and time needed to return to work, would also be relevant but are not measured by the systems we examined. If additional indicators of quality were assessed, hospitals would have less of an incentive to unduly focus on improving a single performance indicator, such as length of stay, to the potential neglect of other indicators.

Outcomes researchers have increasingly emphasized the importance of gaining patients' perspectives when monitoring quality. Patient satisfaction surveys may provide additional information on quality for individual hospitals and are relatively easy to conduct. For these reasons, the

[13]See Grace M. Carter and Glenn A. Melnick, How Services and Costs Vary by Day of Stay for Medicare Hospital Stays (Santa Monica, Calif.: RAND Corporation, Mar. 1990).

[14]In our contacts with hospitals, we found that many hospitals do not have cost accounting systems to estimate their actual operating costs for discrete types of services and that little consistency exists in their cost estimates.

[15]Charges are also a poor measure of expenses because they are believed to overvalue procedures, such as surgeries, and undervalue evaluation and management services.

Cleveland Health Quality Choice Program has complemented its severity-adjusted systems with a patient satisfaction survey. However, critics of patient satisfaction surveys note that quality of care as perceived by patients is subjective and often associated with characteristics of the provider that make care more "personal" rather than improve outcomes.

Severity Adjustments Not Adequate for Comparing All Hospitals

Many experts and hospitals noted that the appropriateness of the severity adjustments remain suspect. If the severity adjustment methods are incomplete or flawed, then hospital comparisons may be skewed. For example, some researchers have found that the systems tend to overestimate the expected mortality rate among hospitals with relatively healthy patient populations and underestimate it among hospitals with more severely ill patient populations.[16] Another expert noted that severity adjustment "is a necessary but inherently imperfect tool."[17]

Hospitals we visited were wary that the severity adjustment methodology is often a "black box" because little independent validation occurs. This apprehension about proprietary severity adjustment systems led Cleveland hospitals to participate in designing a new severity-adjusted performance measurement system. Despite their involvement in the design, some Cleveland hospitals continue to criticize the severity adjustment methods for not accounting for important factors that may influence their severity ratings.[18] In particular, the Cleveland-designed system has been criticized for not properly adjusting for seriously ill patients transferred from other hospitals.

Because most of the measurement systems were initially developed for internal hospital use, severity adjustment limitations are compounded when the systems are used to make comparisons across hospitals. Several large inner-city and teaching hospitals we visited contended that the systems do not accurately compare the severity of illness of their patients with that of patients treated in smaller community hospitals. Inner-city

[16]See, for example, Harry P. Selker, "Systems for Comparing Actual and Predicted Mortality Rates: Characteristics to Promote Cooperation in Improving Hospital Care," Annals of Internal Medicine, Vol. 118, No. 10 (May 15, 1993), pp. 820-22.

[17]Lisa I. Iezzoni, "Risk Adjustment for Medical Outcome Studies," Medical Effectiveness Research: Data Methods, Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services, 1992.

[18]Some outcomes researchers noted that systems achieve greater precision by specifying unique severity adjustment factors for specific conditions. For example, adjusting for severity of illness for coronary-artery bypass graft patients may require a very different methodology than for hernia repair patients.

hospitals often treat patients who have socioeconomic characteristics that severity adjustment systems do not capture. For example, a downtown Cincinnati hospital believes that it is put at a disadvantage in the comparisons because it cares for many drug-abusing patients who are not so identified in the data used for severity adjustment. Similarly, teaching hospitals often treat patients with rare conditions that may not be adequately accounted for in severity adjustments, and their costs may also be higher than community hospitals because of their teaching mission.

Hospitals that treat a large number of uninsured individuals may also be disadvantaged in the comparisons. Much of the health care received by the over 37 million Americans who are currently uninsured is not paid for, prompting providers to recover those costs from the insured. Hospitals that provide more services to uninsured patients may charge higher rates to patients with private insurance and thereby show poor performance results. Thus, because performance measurement programs give hospitals an incentive to reduce their charges to compare more favorably on the systems' indicators, the programs also could create an unintended incentive to reduce services to patients without private insurance.

Conclusion

Collecting and sharing data among health care purchasers and providers is the goal of the employer coalitions in our review. The severity-adjusted measurement systems used to meet this goal are designed to produce more precise and relevant information. Because of the systems' limitations, however, it is important that coalitions not overestimate the capabilities they offer.

For a more detailed description of the severity adjustment systems we reviewed and how they are being used in the communities we visited, see appendixes I to III. We asked a representative of each performance measurement initiative to review the appendixes. These representatives generally agreed with the technical content, and we incorporated their comments as appropriate.

As arranged with your office, unless you publicly announce its contents earlier, we plan no further distribution of this report until 30 days after its issue date. At that time, we will send copies to interested parties and make copies available to others on request.

Please call me on (202) 512-7119 if you or your staff have any questions about this report. Other major contributors are listed in appendix IV.

Mark V. Nadel
Associate Director, National and
Public Health Issues

Contents

Abbreviations

AIM	Acuity Index Method
APACHE	Acute Physiology, Age, Chronic Health Evaluation
APR-DRG	All Patient Refined Diagnosis Related Group
CHOICE	Cleveland Health Outcome Indicator of Care Evaluation
DRG	diagnosis related group
HCFA	Health Care Financing Administration
ICD-9-CM	International Classification of Diseases, Ninth Revision, Clinical Modification
MEDPAR	Medicare Provider Analysis and Review

Central Florida Health Care Coalition

Recognizing that the Orlando area's health care costs were rising faster than elsewhere in the nation, a group of local employers developed the Central Florida Health Care Coalition in 1984 with a charter to "limit the increases in health-care expense without reducing the quality of health care." As part of its cost containment efforts, the Coalition began encouraging local hospitals to purchase MediQual's MedisGroups severity-adjusted performance measurement system in 1990.[19] MedisGroups uses information from medical records to adjust for patients' severity of illness and compares providers' charges, length of stay, mortality rates, and morbidity rates to national norms. Several local hospitals reported that they are using MedisGroups data in their continuous quality improvement programs. The Coalition believes that the initiative has helped contain the increase in Orlando's hospital costs.

During our visit to Orlando, we met with the president and the executive director of the Central Florida Health Care Coalition; officers representing the area's two hospital chains and an independent hospital; the vice president of the Florida Hospital Association; and health benefits managers for several large Orlando employers. We also contacted executives with MediQual, the president of the Orange County Medical Society, and representatives of other Orlando-area hospitals.

History of the Coalition's Initiative

The Central Florida Health Care Coalition was formed in 1984 by 10 of the Orlando area's largest employers, including Walt Disney World, the Orange County Public Schools, Martin Marietta, and General Mills Restaurants. The Coalition has expanded to include about 100 local employers, most having fewer than 300 workers. Currently, the Coalition represents about 180,000 employees. Including dependents, the Coalition members insure about 380,000 individuals in the Orlando area, nearly 30 percent of the area's population.[20]

The Coalition encouraged local hospitals to purchase the MedisGroups severity adjustment system to provide comparable performance information. At the time of our visit, the Orlando hospital market was dominated by two hospital chains—Orlando Regional Healthcare System and Florida Hospitals—representing 9 of the 12 local hospitals. Hospitals affiliated with the Orlando Regional Healthcare System contracted with

[19]"MedisGroups" stands for Medical Illness Severity Grouping System.

[20]The Central Florida Health Care Coalition has recently merged with the Space Coast Labor Management group, including employers such as Lockheed. However, because these employers are in a neighboring county, they are not actively involved in the Orlando-area MedisGroups initiative.

MediQual in 1990, and Florida Hospitals contracted with MediQual in late 1991. Because both hospital chains phased in the use of MedisGroups, complete data for these hospitals were not available until early 1993.

Since our visit, most of the remaining independent Orlando-area hospitals have been purchased by Columbia Hospital System to form a third hospital chain. Although at least one of the hospitals in this new chain had been using MedisGroups, other Columbia-affiliated hospitals had decided not to participate in the Coalition's MedisGroups initiative, in part because of the costs of using the MedisGroups system. As of July 1, 1994, Columbia Hospital System had not yet decided whether the chain, as a whole, will purchase MedisGroups.

Some Orlando-area hospitals are also using other severity-adjusted performance measurement systems, including Iameter's Acuity Index Method (AIM). Several hospitals had begun using these severity adjustment systems before the Coalition's request that they use MedisGroups uniformly. Also, some area hospitals participated in a Florida Hospital Association project using 3M's All Patient Refined Diagnosis Related Groups (APR-DRG) system. Several hospitals that use multiple systems reported that the systems are complementary because they have different relative strengths. For example, one hospital noted that MedisGroups is preferred for assessing mortality and morbidity results, whereas AIM more precisely measures length of stay and charges.

Description of MedisGroups System

After reviewing various severity adjustment systems, the Coalition selected MedisGroups, in part because it preferred a clinically based system. The database tracks length of stay, hospital charges, morbidity rates, and mortality rates, and compares each hospital's performance with national standards.[21] MedisGroups generates data on these indicators by hospital and physician for individual conditions and compares them with the expected rates given the patient's severity of illness. MedisGroups assesses severity of illness once within the first 2 days of admission and then again during days 3 to 7 of the stay. To make the severity adjustment, trained nurses and administrative staff abstract "key clinical findings," including patients' vital signs, laboratory test values, and radiology reports, from the medical record. These data are used to categorize the patients by their likelihood of suffering a major organ failure—"severity 0"

[21]MedisGroups defines morbidity as occurring if a second severity rating taken midstay indicates that the patient has not responded to treatment.

indicates patients are not expected to have major organ failure, whereas "severity 4" indicates the presence of organ failure.[22]

MedisGroups omits some pertinent information that may affect the cost of treatment during the hospital stay. Although the system tracks severity of illness, it does not take into account the patient's medical history, such as whether it is the patient's first or second heart attack. The system is also unable to distinguish between an emergency coronary bypass surgery, a repeat coronary bypass surgery, and a coronary bypass surgery following angioplasty.

The costs of MedisGroups are paid by the hospitals, with the amount varying by size of the hospital. The hospitals we visited paid initial purchase fees of $230,000 to $330,000 for licensing, software, and equipment, and ongoing annual fees of $28,000 to $70,000. Another major expense of using MedisGroups is the staff time required to abstract clinical findings from medical records and enter the data in the MedisGroups system.[23] A relatively small hospital with 300 acute-care beds dedicates 3 full-time staff to MedisGroups data collection at an annual cost of $72,000. A large hospital chain has 15 full-time staff abstracting MedisGroups records for 1,500 hospital beds at an annual cost of $500,000. The Coalition estimates that MedisGroups costs average about $8 to $11 per hospital admission.

Hospitals' Use of MedisGroups Results

Hospital administrators can analyze the MedisGroups-adjusted data to identify the performance of individual physicians for particular conditions. In addition, they can compare their hospital's performance with that of others in the MedisGroups database. Table I.1 provides an example of the MedisGroups data provided to a hospital. It indicates the hospital's actual charges compared with their expected charges, estimated by MedisGroups on the basis of the patients' severity of illness. The ratio of the actual value divided by the expected value is defined as the standard ratio. The system compares whether the standard ratio is significantly different from 1: a standard ratio significantly greater than 1 is worse than average, whereas a standard ratio less than 1 is better than average.

[22]For a more thorough technical discussion of MedisGroups, see Lisa I. Iezzoni and Mark A. Moskowitz, "A Clinical Assessment of MedisGroups," Journal of the American Medical Association, Vol. 260, No. 21, (Dec. 2, 1988), pp. 3159-63.

[23]The training to become a MedisGroups' abstractor is extensive, with abstractors required to meet a 95-percent consistency standard before becoming qualified by MediQual. One hospital reported that abstractor training required 3 to 6 months.

Table I.1: GAO Illustration of Typical Medisgroups Report Information

DRG	Description	Number of patients	Average total charges		
			Actual average	Expected average	Standard ratio
391	Normal newborn	276	$796	$861	.92[a]
371	Cesarean section without complications	64	$6,635	$5,188	1.28[a]
209	Major joint/limb operation of lower extremities	57	$18,204	$19,645	.93[b]
98	Bronchitis, asthma, ages 0 to 17	48	$3,335	$3,334	1.00[c]
127	Heart failure and shock	48	$7,239	$8,919	.81[b]

Note: "Standard ratio" is an index that is created by dividing the actual value by the expected value.

[a]Standard ratio is statistically significant at the .01 level.

[b]Standard ratio is statistically significant at the .05 level.

[c]Standard ratio is not significantly different from 1 at the .05 level.

Several local hospitals have incorporated the MedisGroups data in their quality improvement processes. By using MedisGroups data to identify specific procedures for which performance is significantly worse than expected, the hospital can target its improvement efforts, as the following examples illustrate:

- One hospital established physician task forces to review cesarean sections, pneumonia, and back surgery. The hospital believes that the task force recommendations, such as switching to a less expensive pharmaceutical following back surgeries, have saved the hospital money.
- Another hospital has established committees to focus on acute myocardial infarctions, congestive heart failure, cesarean sections, and respiratory care. One committee recommended using a metered dose inhaler rather than aerosol treatment for patients receiving respiratory therapy. The hospital estimates that changing to the inhaler will save about $300,000 per year, largely because the inhaler requires only about 3 minutes of a respiratory therapist's time, whereas the aerosol treatment requires about 20 minutes.

Several hospitals noted that their quality improvement programs had been active before they received the MedisGroups or other severity adjustment systems' data. One hospital cited an aggressive quality improvement program as a factor in decreasing inpatient hospital charges. Even without

the MedisGroups data, it might have had a similar quality improvement program. But the hospital believes that the employers' interest in the data has provided additional motivation to the hospital's quality improvement process.

Employers' Use of MedisGroups Results

Hospitals summarize the MedisGroups results and their continuous quality improvement efforts in regular reports to the Central Florida Health Care Coalition.[24] For the most part, employers have encouraged hospitals to continue using the MedisGroups data for reviewing their medical practices rather than directly using the data to select the better performing providers. However, several employers have made hospital participation in the MedisGroups initiative a condition for inclusion in their health networks:

- One large employer intends to drop a hospital from its managed care network because it has not purchased MedisGroups.
- A purchasing alliance established by members of the Central Florida Health Care Coalition and other Florida employer coalitions has also made participation in the MedisGroups project a criterion for hospitals to be included in their health plan.
- Another large company has used MedisGroups and other risk-adjusted data in establishing a subset of "most preferred" providers within its existing managed care network. Employees receive financial incentives to use the subset of providers.

Claimed Savings

Several hospitals and employers have identified hospital cost savings that have occurred since purchasing MedisGroups:

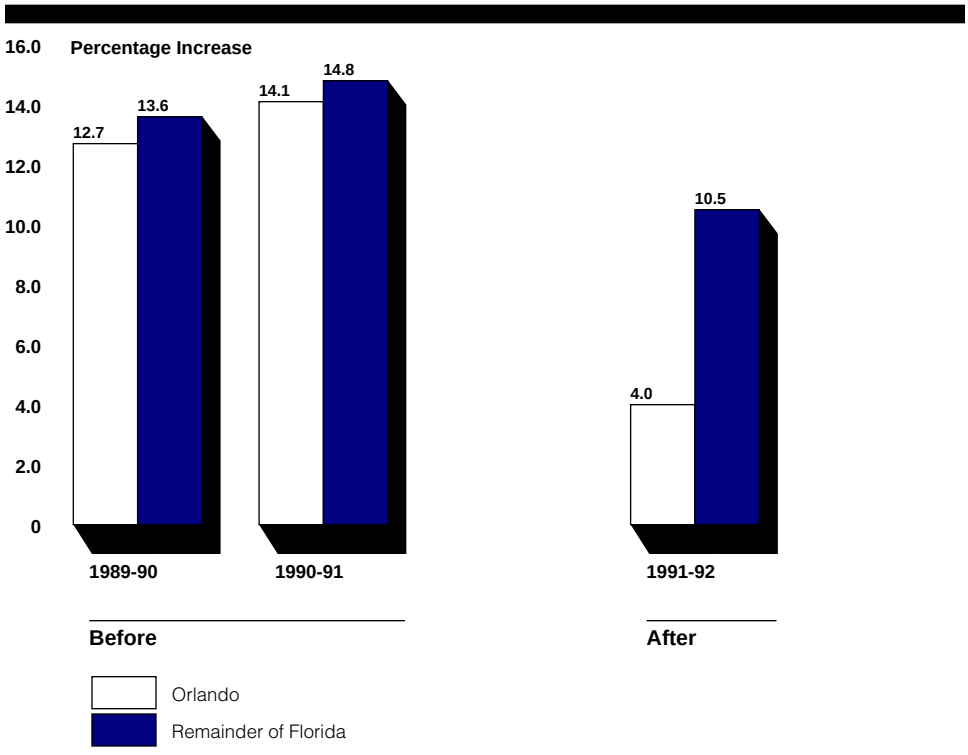
- One Orlando hospital reduced its average length of stay by 13.8 percent in 1992, leading to a 1.7-percent reduction in average costs per patient. This compares with a statewide average length-of-stay decline of about 2.7 percent during the same period.
- Another hospital has identified savings for specific conditions and procedures. For the individual conditions that hospital task forces have assessed, costs per case declined by between 3 and 43 percent, charges by between 8 and 28 percent, and length of stay by between 3 and 32 percent.
- One employer, large enough to have independent clout, also cited benefits from its participation in the coalition. From 1989 to 1991, its premium

[24]At least one hospital also reanalyzes the MedisGroups data to develop customized reports for specific large employers.

increases averaged 11 percent per year; since 1991, the premium growth has been less than half that rate. The employer estimated a savings of $1 million in anticipated inpatient hospital costs due to the MedisGroups project.

Our analysis of data collected by the Florida Agency for Health Care Administration indicates that Orlando's hospital charges have grown more slowly since the implementation of the MedisGroups program (in 1991 and 1992) and relative to other hospitals in Florida. (See fig. I.1.) Similarly, Orlando's average hospital length of stay has declined more rapidly compared with prior years and with other Florida hospitals. (See fig. I.2.)

Figure I.1: Growth in Hospital Charges in Orlando and the Remainder of Florida, 1989 to 1992

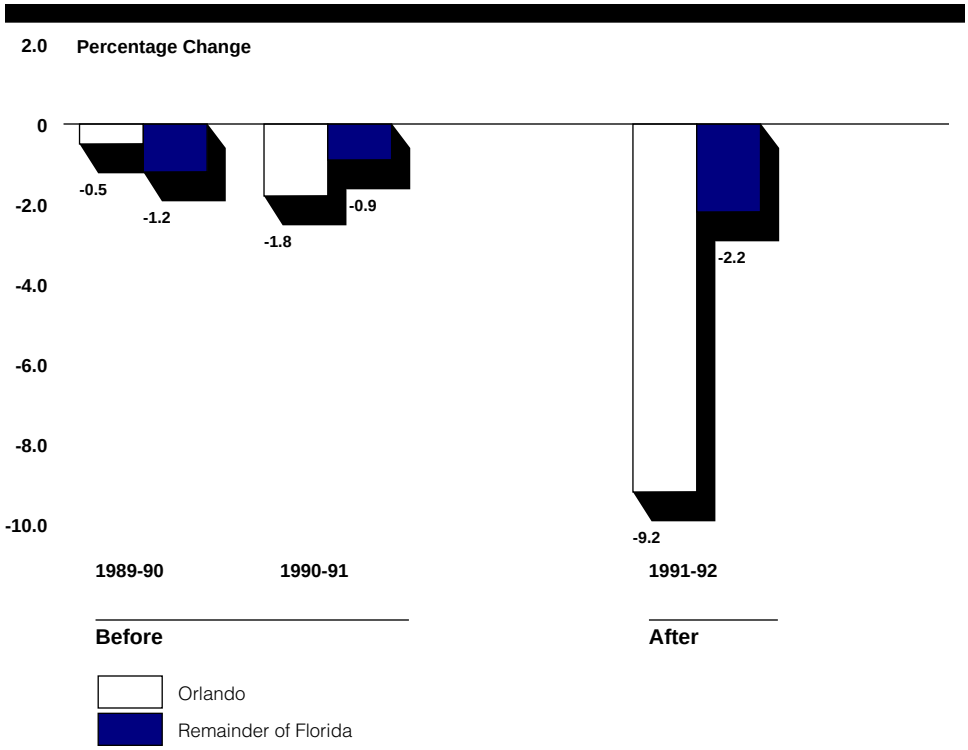


Notes: The Orlando data exclude three hospitals that were not using the MedisGroups system during this period. These hospitals' data are included in the "remainder of Florida" category.

The use of MedisGroups was phased in by hospitals. By 1992, hospitals representing about 85 percent of Orlando hospital beds were implementing MedisGroups.

Source: GAO calculation of data from Florida Agency for Health Care Administration.

Figure I.2: Decline in Hospital Patient Length of Stay in Orlando and the Remainder of Florida, 1989 to 1992



Notes: The Orlando data exclude three hospitals that were not using the MedisGroups system during this period. These hospitals' data are included in the "remainder of Florida" category.

The use of MedisGroups was phased in by hospitals. By 1992, hospitals representing about 85 percent of Orlando hospital beds were implementing MedisGroups.

Source: GAO calculation of data from Florida Agency for Health Care Administration.

Some hospitals and employers recognize that it is difficult to isolate the impact of MedisGroups on Orlando's hospital cost trends. For example, one hospital noted that its ongoing quality improvement program and lower payments from Medicare and managed care plans also contributed to its shorter stays and slower cost growth. In addition, some members of the Central Florida Health Care Coalition are participating in a statewide employers' insurance purchasing alliance. The Coalition attributes cost savings in the area to this purchasing alliance's leverage in achieving hospital discounts as well as to the MedisGroups initiative.

Employers also indicated that recent reductions in their inpatient hospital expenditures have been matched by increases in ambulatory care costs, with their total health care expenditures continuing to increase at about the same rate as in prior years. For example, between 1991 and 1992 one large Orlando employer's inpatient hospital payments per employee declined by about 14 percent, whereas its payments for hospital outpatient care, physician care, and pharmaceuticals rose by nearly 17 percent. Overall, the employer's health care payments per employee still grew by about 4 percent. Similarly, another large Orlando employer's average charge per case for inpatient hospital care declined by about 9 percent from 1990 to 1992, but because of increasing expenditures for ambulatory care overall health plan costs increased 5 percent.[25]

Future Developments in Orlando's Hospital Market

As previously discussed, the Orlando hospital market has recently become more consolidated with the development of a third hospital chain. In the future, specialization is expected to occur within these hospital groups, with hospitals concentrating on providing the services at which they excel. For example, a hospital administrator noted that two affiliated hospitals may decide that one hospital will provide pediatric care while the other hospital will provide cardiac care. In this way, each hospital can reduce hospital beds for some services while the hospital chain can continue to offer a comprehensive package of hospital services to employers.

As part of a broader health reform initiative to make outcomes data from Florida hospitals available to the public, Florida's law requires the use of a uniform severity-adjusted performance measurement system. In 1993, the state selected an administratively based system, APR-DRG, which had been piloted by the Florida Hospital Association. The hospital association advocated this system because of its relatively low cost—it estimated that APR-DRG cost hospitals less than $10,000 per year. Thus, many of the Orlando-area hospitals participated in the Florida Hospital Association pilot and are using APR-DRG in addition to MedisGroups. Because the Central Florida Health Care Coalition prefers the advantages of a clinically based system, it is continuing to encourage Orlando-area hospitals to use MedisGroups despite the state's mandated use of APR-DRG.

[25]This compares to an average 8 percent increase in health plan costs reported for over 300 large Florida employers between 1991 and 1992. See William M. Mercer, Inc., 1992 Florida Health Care Costs and Benefits: Survey Results (Tampa, Fla.: 1992).

Cincinnati Payer Initiative

Frustrated by annual double-digit health plan cost increases, four large Cincinnati employers collaborated to launch the Cincinnati Payer Initiative. The employers view this 3-year initiative as a mechanism to combat hospital cost increases by using severity-adjusted data to identify less efficient hospitals. The potential market clout of the 4 large employers encouraged 14 local hospitals to purchase Iameter's AIM severity adjustment system.[26] AIM provides a method to make severity adjustments for providers' patient populations and compare inpatient charges, length of stay, and mortality rates on a communitywide basis. Hospital officials reported that the data assist in identifying clinical areas on which to focus their cost containment efforts.

We contacted each of the 14 Cincinnati hospitals that have purchased AIM and conducted more detailed visits with 4 major hospitals. We also contacted several of the employers that began the Cincinnati Payer Initiative; the Greater Cincinnati Employer Health Care Alliance; the Greater Cincinnati Hospital Council; the Academy of Medicine of Cincinnati; and representatives of Iameter.

History of the Initiative

In 1984, the Greater Cincinnati Employer Health Care Alliance was established as a forum to exchange information and collaborate on special projects with the goal of promoting cost containment and quality care. Currently, the Alliance represents nearly 80 large and small employers. In 1990, the Alliance began examining various types of severity-adjusted performance measurement systems. Although several experts in the field of performance measurement provided guidance on selecting severity adjustment systems, the project never progressed to the point of actually comparing systems on the market.

Meanwhile, health care costs continued to escalate. From 1989 to 1991, Cincinnati employers experienced annual health care inflation rates averaging 10 to 14 percent. Disappointed by these cost increases and the slow pace of the Alliance's project, four large employers initiated an independent examination of severity-adjusted performance measurement systems in 1991. The four employers—Cincinnati Bell, General Electric Aircraft Engines, Kroger, and Proctor and Gamble—represent about 168,000 employees and dependents, constituting over 13 percent of the local health care market.

[26]All of the hospitals that the employers asked to participate agreed to purchase AIM. The initiative did not include several Cincinnati-area hospitals serving specific types of patients, including a children's hospital, a veterans' hospital, a burn center, and a psychiatric hospital.

After examining several severity adjustment systems, the employers selected Iameter's AIM. Iameter performed an initial comparison of hospital performance based on publicly available Medicare Provider Analysis and Review (MEDPAR) data.[27] The employers shared these preliminary comparisons with representatives from 14 local hospitals in March 1992. After this meeting, each hospital agreed to contract with Iameter and submitted discharge data for all of their patients for 1991.

Iameter's AIM

AIM reports a comparison of charges, length of stay, and mortality rates for two levels of specificity: (1) for five major disease categories, such as circulatory care, and (2) several dozen high-volume diagnosis related groups (DRG), such as coronary-artery bypass grafts. AIM is one of several administratively based systems that uses routinely collected discharge data to adjust for severity. At discharge, the attending physician writes each patient's diagnosis and procedures on an attestation statement. For billing purposes, the hospital's medical records department translates these narratives into numeric codes (DRGs and the codes specified in the International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM]). To make its severity adjustments, AIM collects patient data on age, sex, length of stay, total charges, discharge status, principal diagnosis, secondary diagnosis, and procedures performed. Using this information, AIM classifies patients with the same diagnosis into one of five severity levels.

The employers selected AIM largely because of its relatively low cost and ease of implementation. Several hospitals we visited reported that data for AIM's analysis are readily retrievable from their internal management information systems. The hospitals required little or no additional staff, staff training, or equipment to develop the data for AIM's analysis, so minimal additional costs were incurred. Hospital officials reported that Iameter offers several packages of services ranging in cost from $20,000 to over $80,000. The low-cost package consists only of data comparing hospitals at broad levels of services, while the high-cost package includes condition-specific and physician-specific comparisons and Iameter's consulting services.[28]

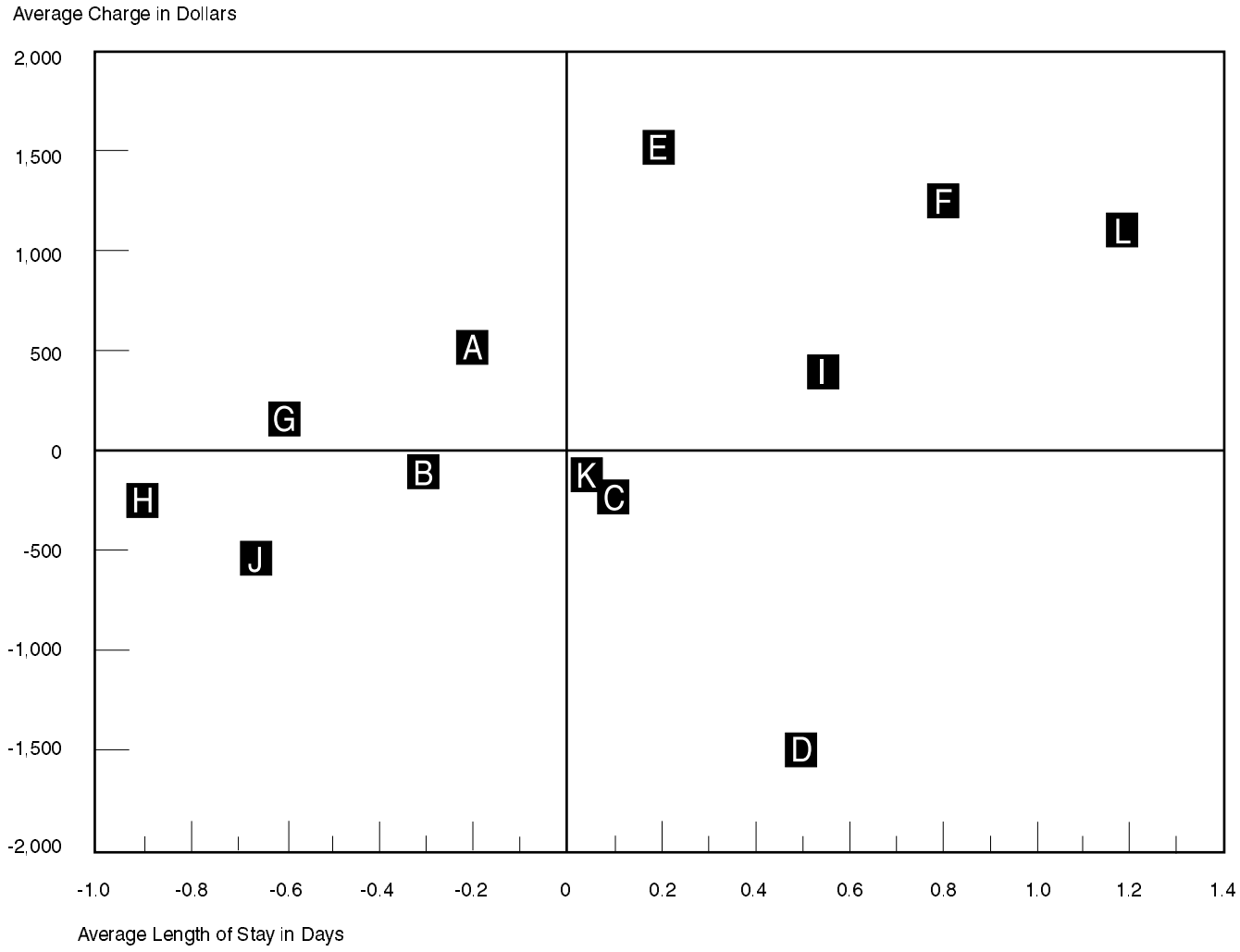
[27]MEDPAR data account only for Medicare hospital patients and do not include patients with Medicaid or private coverage. Medicare patients represented nearly half of hospital discharges nationally in 1991.

[28]In the comparisons, data from the other hospitals are coded to protect confidentiality.

Cincinnati hospitals we visited reported that since the Cincinnati Payer Initiative began they have paid more attention to the codes documenting patient services. Because AIM relies on these codes to make severity adjustments, differences in coding practices by the hospitals may significantly influence the patients' severity classifications and hence the hospitals' relative standing. Because caring for more severely ill patients requires more resources and entails greater risk, hospitals that obtain high severity ratings appear justified in having higher charges, length of stay, and mortality rates. As a result, several Cincinnati hospitals noted that they have conducted seminars and other efforts to help physicians and other staff code more "effectively." Other hospitals have changed administrative coding processes by introducing concurrent coding in which services and procedures are coded on billing forms as they are delivered instead of after the patient has been discharged.

After the information is collected and analyzed, Iameter provides participating employers and hospitals with the comparative results. Figure II.1 is our adaption of an Iameter chart comparing hospitals' performance for circulatory care. The hospitals in the upper right quadrant had shorter patient stays and lower charges, whereas the hospitals in the lower left quadrant had longer patient stays and higher charges for equivalent cases. Positive numbers on the chart reflect performance better than the community norm.

Figure II.1: GAO Illustration of a Typical Chart in an Iameter Report (Length of Stay and Charges for Circulatory Care)



Note: Average charges and average length of stay are stated as the difference from the community norm, which is shown as the "0" point on each axis.

Standardized for severity, Iameter's 1992 comparisons demonstrated wide variations in hospital performance in Cincinnati. For example, average

charges across Cincinnati's 14 hospitals varied by about $900 for pregnancy cases, $3,300 for circulatory cases, $4,000 for respiratory cases, $4,300 for digestive-tract cases, and $4,600 for musculoskeletal cases. Similarly, average length of stay varied by about 0.5 day for pregnancy cases, 1.5 days for digestive-tract cases, 2 days for circulatory and respiratory cases, and 2.5 days for musculoskeletal cases.[29] In many cases hospitals had variable performance profiles, scoring well for some patient groups but poorly for others.

Hospitals' and Employers' Use of AIM Data

Many of the hospitals we visited are using AIM's data in their ongoing continuous quality improvement programs. Because AIM's data report variations among physicians, hospital officials said the information helps them identify where more detailed examination of processes, procedures, and physician practices is needed. Many hospitals believe that AIM does not provide enough data on how to improve performance. However, information on variation is viewed as a useful tool to help physicians identify problems and to target hospital quality improvement efforts. Several local hospitals contract with Iameter for additional consulting services to identify opportunities for improvement.

Hospitals and physicians have begun to make changes to improve efficiency as these examples show:

- AIM reported that physicians at one hospital had average stays for pneumonia patients ranging from 7 days longer to 6 days shorter than the community norm and average charges differed by nearly $12,000. In an effort to reduce this variation by standardizing care, a team of physicians determined that treating patients with a less expensive antibiotic and involving respiratory therapists to collect sputum samples earlier in treatment could decrease charges and length of stay. The team developed treatment protocols based on the practices of the best-performing physicians that are used to guide treatment for all pneumonia patients.
- Orthopedists at one hospital agreed to use just one or two brands of artificial hips in hip replacement operations rather than choosing from among the many versions on the market. This allows the hospital to save money with volume purchases from a single vendor and helps physicians standardize treatment protocols, which should reduce variation in charges and length of stay.

[29]For respiratory care, for example, these ranges represent variation of 25 percent above and 26 percent below the average charge and 16 percent above and 12 percent below the average length of stay in Cincinnati hospitals.

Whereas a hospital sees only how it compares with other unnamed hospitals, the employers' reports identify each of the hospitals in the communitywide comparison. The employers expect that trends in hospital performance will begin to appear after 3 years of data are compiled, and they will then start using the AIM information in making their purchasing decisions. One employer reported that it has shared the AIM data with managed care plans that bid for a contract, asking that the plans consider the AIM data in selecting hospitals and physicians for their networks. However, the employer reported that other factors are more important in selecting a managed care plan, including the plan's general management structure. For this reason, the employer has maintained contracts with its existing managed care organization although some hospitals that are highly rated by AIM are not in the plan's network.

Claimed Savings

On May 24, 1994, the hospitals received their most recent AIM-adjusted reports reflecting 1993 performance. Iameter reported that between 1992 and 1993, hospital charges decreased by about 1 percent and length of stay decreased by about 10 percent.[30] Iameter estimated that these changes have led to savings among the hospitals of $200 million between 1991 and 1993.

Although Iameter and employer coalition leaders have claimed cost savings, we found little pattern to the growth rate in charges over the period. We asked hospitals and employers about their cost trends since using the system and were given the following examples:

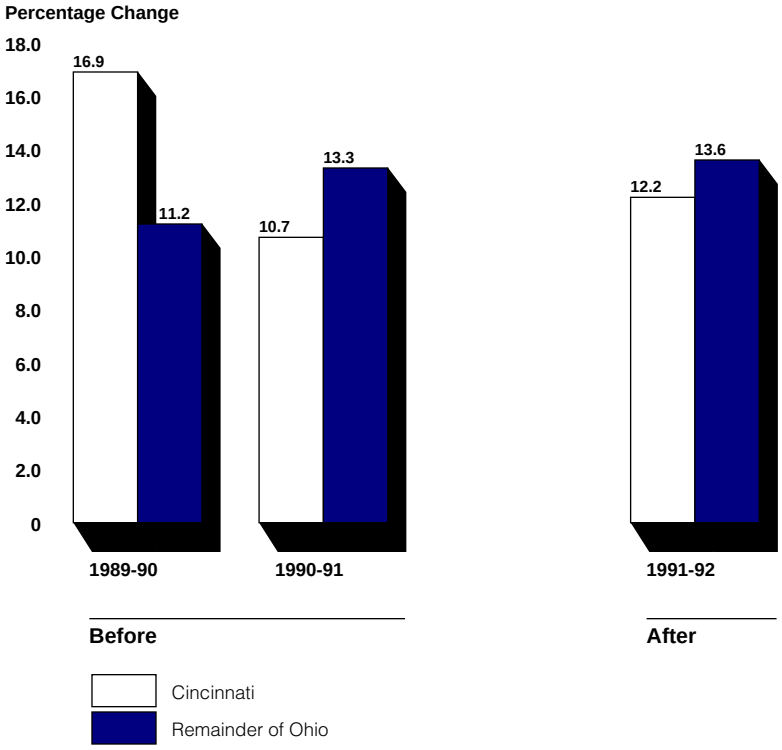
- A Cincinnati hospital estimated that a 0.6-day average annual reduction in its average length of stay for all patients between 1991 and May 1993 could have saved as much as $5 million in nursing care, food, drugs, and supply costs in 1993 (5 percent of the hospital's operating costs). By comparison, the statewide average length of stay in Ohio declined by only 0.2 day between 1991 and 1992.
- Another Cincinnati hospital estimated that reducing its average length of stay by about 0.5 day since the Cincinnati Payer Initiative began led to cost savings of between $1.3 million and $4.8 million during a 15-month period (about 1.3 to 4.8 percent of the hospital's inpatient revenues).[31]

[30]Iameter did not share the most recent detailed performance results with us for 1993.

[31]The hospital estimated the range in savings based on varying assumptions of how much a shorter average length of stay reduces the average cost of a day of hospital care, ranging from saving the daily average cost ($946) to saving only a fraction of the average cost ($250).

Comparing charges and length of stay before and after the employer coalition's initiative began indicates mixed results. Data reported by the Ohio Department of Health indicate that Cincinnati's hospital charges grew at a faster rate the year after hospitals initially purchased AIM than the year before. However, both years had significantly lower rates of growth than 2 years before. The 1992 growth in charges in Cincinnati was somewhat lower than the rate experienced by other hospitals in Ohio, but continued at double-digit rates. (See fig. II.2.) Length of stay declined by 4.2 percent between 1991 and 1992, a greater decline than reported for the 2 years before AIM was used. A similar rate of decline (3.7 percent) was experienced by other Ohio hospitals. (See fig. II.3.)

Figure II.2: Growth in Hospital Charges in Cincinnati and the Remainder of Ohio, 1989 to 1992



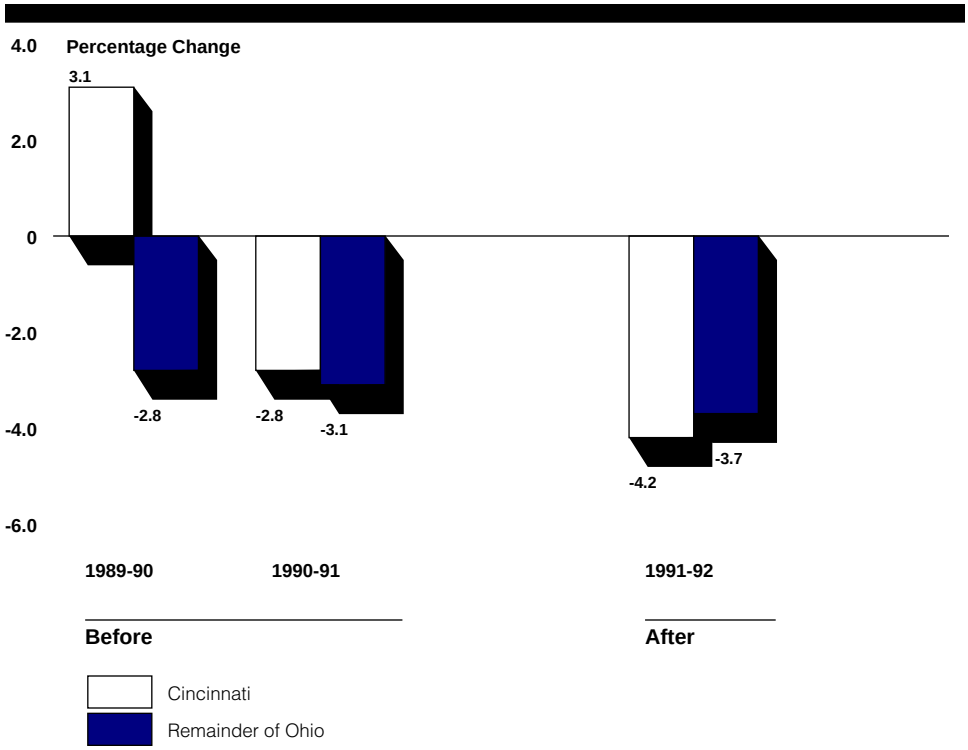
Notes: Ohio Department of Health data exclude Medicare and Medicaid patients.

Hospitals report only the 100 highest-volume diagnosis related groups.

Hospitals in the "remainder of Ohio" include those in Cleveland, which were developing their severity-adjusted performance measurement project during this time.

Source: GAO calculations based on data obtained from Ohio Department of Health.

Figure II.3: Change in Hospital Patient Length of Stay in Cincinnati and the Remainder of Ohio, 1989 to 1992



Notes: Ohio Department of Health data exclude Medicare and Medicaid patients.

Hospitals report only the 100 highest-volume diagnosis related groups.

Hospitals in the "remainder of Ohio" include those in Cleveland, which were developing their severity-adjusted performance measurement project during this time.

Source: GAO calculations based on data obtained from Ohio Department of Health.

Hospitals and employers noted that other factors could have contributed to recent changes in Cincinnati hospitals' cost trends. Employers acknowledged that their general, increased attention to health care costs, including an increased use of managed care plans and other cost containment strategies, also influenced the Cincinnati health care market. According to a representative of the Greater Cincinnati Employer Health Care Alliance, a local health maintenance organization that is not involved in the Cincinnati Payer Initiative also claimed credit for the declining length of stay and slowed growth in hospital charges.

In addition, the timing of the initial AIM reports suggests that hospitals were making changes before receiving their AIM results. Iameter's analysis of the data hospitals provided for 1991 was made available in fall 1992. The hospitals then had only a few months to make performance improvements before submitting their 1992 data in January 1993. Thus, although the AIM data indicated improvement by many hospitals between 1991 and 1992, changes in hospital performance may have resulted from a sentinel effect—that is, from hospitals' anticipation of the employers scrutinizing their performance. The hospitals' ongoing quality improvement efforts, initiated before the receipt of AIM data, may also have contributed to the differences in the 1991 and 1992 comparisons.

Cleveland Health Quality Choice Program

The Cleveland Health Quality Choice Program has several unique characteristics. In contrast to the Cincinnati and Orlando initiatives, the program has pursued a slower but a more collaborative approach among the local physician and hospital associations and several employer coalitions. Rather than choosing a single off-the-shelf system, the program custom-designed a three-part system. It compares local hospitals' performance in mortality, length of stay, and patient satisfaction, but omits hospital charges. The severity-adjusted data do not directly make cross-hospital comparisons, but rather report only whether a hospital performs as expected for the type of patient it treats.

During our site visit, we interviewed the executive director of the Cleveland Health Quality Choice Program as well as representatives of several local hospitals, employer associations, and insurers. We also discussed the program with representatives of the Greater Cleveland Hospital Association; the Academy of Medicine of Cleveland; and Michael Pine and Associates, which developed the Cleveland Health Outcome Indicator of Care Evaluation (CHOICE) system. Ohio Department of Health data for 1993 will not be available until fall 1994. Because the first hospital comparisons from the Cleveland system were issued in April 1993, we were unable to compare the changes in Cleveland hospital charges and length of stay before and after the implementation of the system as we did with Orlando and Cincinnati.

History of the Program

During the late 1980s, Cleveland employer coalitions began to negotiate more aggressively with local hospitals, in part as a result of analyses and anecdotes indicating Cleveland's health care costs were relatively high. A Foster Higgins study reported that the cost of care in Cleveland was the fourth highest among metropolitan areas in the United States. This study noted that hospital services costing $100 in Chicago would cost about $140 in Cleveland. Cleveland employers also began telling hospitals that for some high-cost services, it would be less expensive to fly an employee to the Mayo Clinic in Rochester, Minnesota, for treatment than to have the employee treated at a local hospital.

The employers told the hospitals that they would begin using Ohio Department of Health charge data and Medicare mortality reports to select low-cost hospitals that had satisfactory mortality outcomes. The hospitals responded that these comparisons would be unfair. They echoed the argument made by hospitals elsewhere that by not adequately adjusting for differences in the average severity of illness among the hospitals' patients,

hospitals with "sicker" patients would be disadvantaged in the comparisons. Consequently, the hospitals, physicians, and employers collaborated to develop the Cleveland Health Quality Choice Program in 1989.

The directors of the program include representatives from the Academy of Medicine of Cleveland, the Greater Cleveland Hospital Association, and three employer groups. These employer groups are Cleveland Tomorrow, a group of chief executive officers from the 50 largest Cleveland corporations; the Health Action Council of Northeast Ohio, a coalition of about 100 employers; and the Council of Smaller Enterprises, a buying cooperative of 12,000 small businesses. These employer groups provide health coverage to about 350,000 individuals in the greater Cleveland area, about 20 percent of the population. All 29 of the major hospitals in the Cleveland area (except for a veterans' hospital) are participating in the program.

As part of its collaborative approach, the Cleveland Health Quality Choice Program includes the local hospital and physician associations as members. Both groups participated in the development of the program's three-part system, including the design of the CHOICE system and the selection of the Acute Physiology, Age, Chronic Health Evaluation (APACHE III) system and the Patient Viewpoint Survey. As a voting member, the Academy of Medicine of Cleveland successfully opposed including physician-specific data in the program. In addition, the health care provider groups have ensured that published reports meet high standards of statistical validity.

Performance Measurement Systems Used by Cleveland Program

The members of the Cleveland Health Quality Choice Program decided in 1990 that for comparing inpatient hospital services, the severity-adjusted performance measurement systems on the market were not suitable. The program members were concerned that the methodology of many of the systems was a "black box" because little rigorous independent evaluation of its validity had been conducted. Thus, the program opted for a three-part approach to evaluate hospital performance, including the new, custom-designed severity-adjusted performance measurement system developed with the participation of Cleveland hospitals and physicians. For intensive care outcomes, the program's representatives selected the APACHE III system, which has been widely used and independently evaluated. The program also contracted for the Patient Viewpoint Survey to assess patient satisfaction with hospital services.

Most of the costs of the Cleveland Health Quality Choice Program have been borne by the participating hospitals. The development and implementation of the new program cost the hospitals about $8 million over 3 years, with the participating employer groups contributing an additional $1 million.[32] The Cleveland Health Quality Choice Program estimates that the hospitals' overall costs average about $8 per discharge.

CHOICE

The CHOICE system was designed by Michael Pine and Associates in collaboration with Cleveland-area health care providers. The system underwent a 3-year start-up phase beginning in 1990; the first report was issued in spring 1993. A clinically based system, CHOICE estimates patient severity of illness by abstracting data from medical records. In addition to a patient's age, sex, race, and diagnoses, the system retrieves information such as vital signs, radiology results, and laboratory test values.

Using the CHOICE model, the program reports both the expected and the actual mortality rates and average length of stay for each hospital. The CHOICE results are reported for medical and surgical patients admitted for 14 specific diagnoses or surgical procedures.[33] In spring 1994, the program included summary performance for two general categories (medical patients and surgery patients) and five narrower clinical areas (cardiovascular, gastrointestinal, neurologic, respiratory, and coronary-artery bypass graft patients). In future reports, the program hopes to also show comparisons of additional specific services, including obstetrics.

The CHOICE system has been criticized for making some hospitals appear worse than their actual performance warrants. For example, one hospital has criticized the CHOICE model for neglecting to distinguish between patients who are transferred from another hospital and patients who initially seek care at the hospital. This may result in worse scores and unfair comparisons. The hospital contends that the severity of illness of a patient transferred from another hospital may be understated because the patient's clinical status is stabilized before the transfer. However, the patient's underlying conditions remain and may deteriorate soon after

[32]The program has recently contracted with Apache Medical Systems, Inc., to market the CHOICE system to other hospitals; Apache will pay further system development costs.

[33]The medical diagnoses include patients admitted with acute myocardial infarction, congestive heart failure, stroke, gastrointestinal hemorrhage, pneumonia, and chronic obstructive pulmonary disease. Surgical procedures include coronary artery bypass, peripheral vascular repair or bypass, lung resection, lower bowel resection, laminectomy, reduction of hip fracture, prostatectomy, and hysterectomy.

admission. An analysis conducted by Lewin-VHI indicates that the hospital's mortality rates would improve if transferred patients were separately accounted for in the data.[34] The hospital also notes that its share of patients transferred from local hospitals increased sharply in 1992.

APACHE III

APACHE III compares Cleveland hospitals with national norms for intensive care mortality rates and length of stay. APACHE III develops a severity score for intensive care patients, ranging from 0 to 200, from which it computes expected mortality rates and length of stay. To estimate the patient's severity of illness, APACHE uses medical records data, including 16 clinical values (such as temperature, heart rate, and blood and urine test results) and information regarding the presence of chronic health problems (such as acquired immunodeficiency syndrome and cancer).

The Cleveland Health Quality Choice Program reports both the predicted and actual mortality rates and average length of stay for each intensive care unit. In future reports, after a sufficient number of patients are included in the program's database to meet statistical validity standards, the program intends to report mortality and length of stay results for specific categories of intensive care patients, such as cardiovascular, gastrointestinal, respiratory, and neurologic patients.[35]

Patient Viewpoint Survey

In addition to the severity-adjusted hospital performance comparisons described in the previous two sections, the Cleveland Health Quality Choice Program reports patient satisfaction comparisons among the hospitals. The Patient Viewpoint Survey is mailed to 600 randomly selected patients from each hospital for each of the program's semiannual reports. In addition to such global satisfaction questions as "Would you recommend the hospital to a friend?" the survey also asks about 11 particular components of care, such as the availability of doctors, the responsiveness of the nursing staff, the patient's experience with the admitting and billing departments, and the patient's satisfaction with food and housekeeping services.

[34]The Lewin-VHI study acknowledges that accounting for transfer patients may not change the predicted mortality rates for all hospitals, but could affect predicted mortality rates for some hospitals that treat many transferred patients. See Robert J. Rubin and William A. Gold, "An Assessment of the Cleveland Health Quality Choice Mortality Models" (Fairfax, Va.: Lewin-VHI, Apr. 27, 1993).

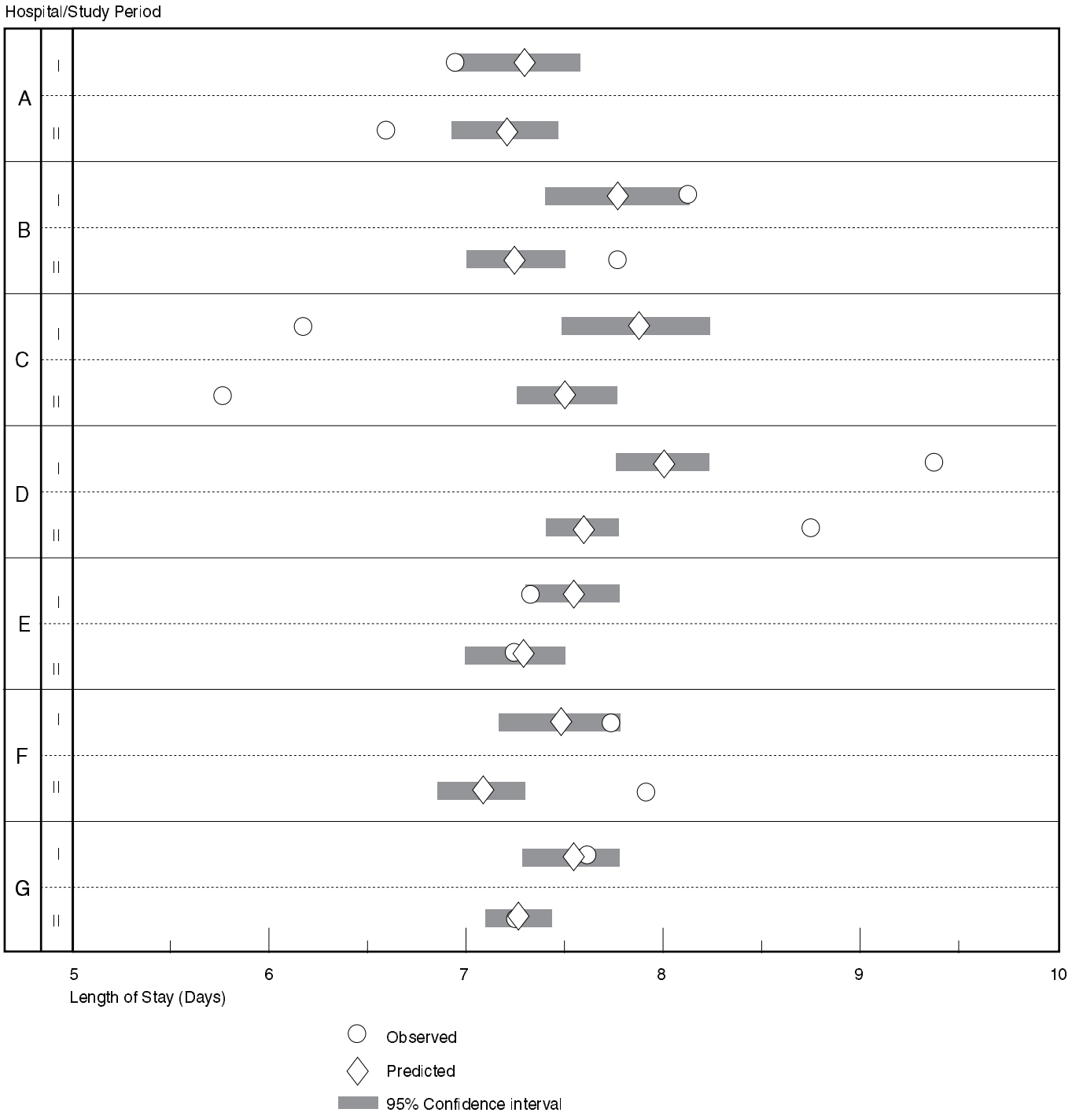
[35]The Cleveland Health Quality Choice Program recently recalled APACHE data for these specific services that had been printed in the June 1994 report. The program attributed the mistaken inclusion of these comparisons to a computer programming error.

Patient satisfaction information has been criticized as being subjective and often associated with characteristics of the hospital that makes care more "personal" rather than improving outcomes of care. However, experts in quality assessment are increasingly finding patients' evaluations of their care to be meaningful.

Reporting Results

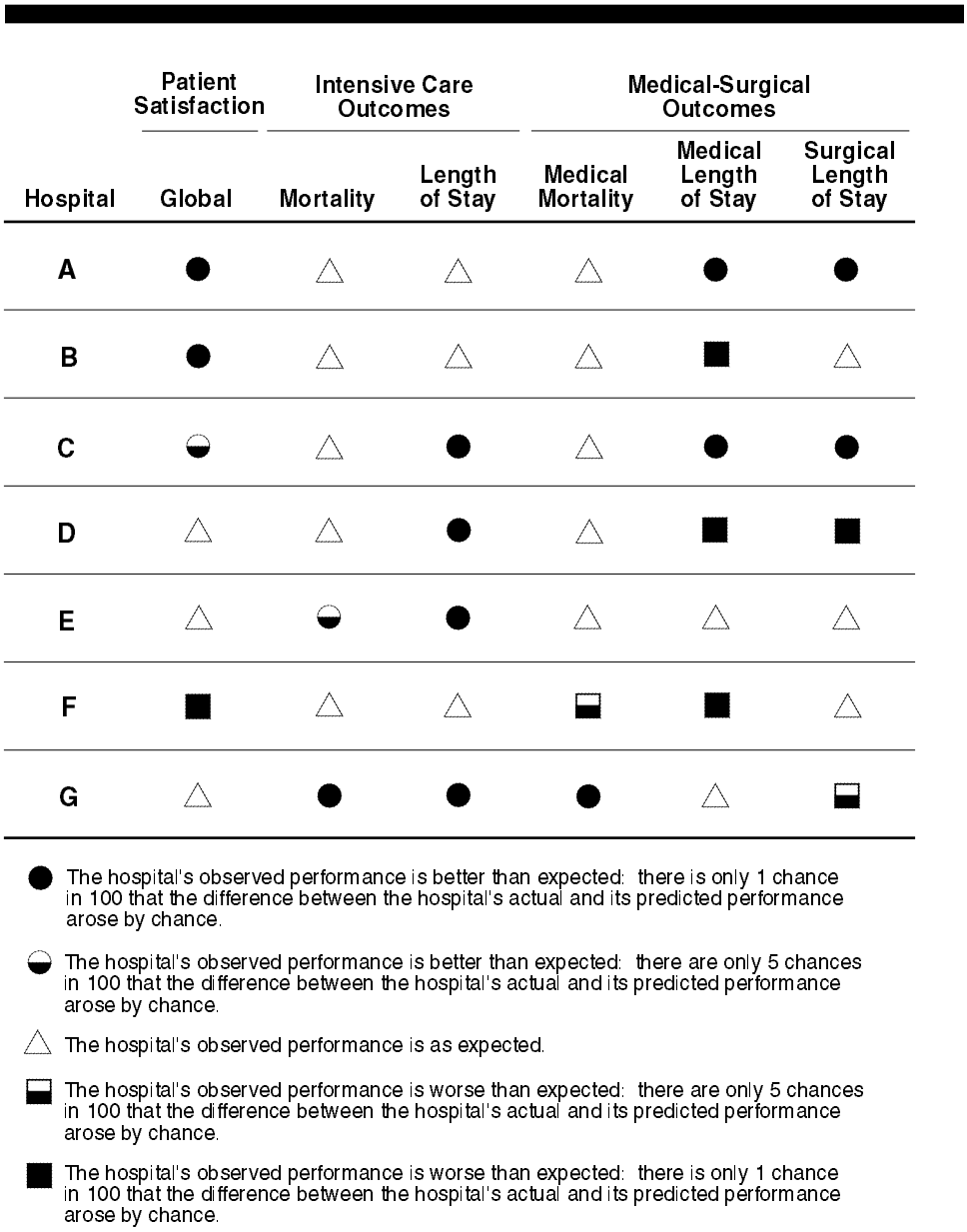
Figure III.1 illustrates a Cleveland Health Quality Choice Program comparison for surgical patients' average length of stay. (In the figure, the hospital names are coded for anonymity, but the program's reports include hospital names.) As shown, the program assesses hospital performance as above, at, or below the expected levels for each performance measure (for example, surgical length of stay). The program cautions that hospital results should not be compared with those at other hospitals, but only with its expected performance.

Figure III.1: GAO Illustration of a Typical Chart in a Cleveland Health Quality Choice Report (Length-of-Stay Data Over Time for Surgical Patients)



The program also summarizes hospitals' performance across indicators, as illustrated in figure III.2.

Figure III.2: GAO Illustration of a Typical Chart in a Cleveland Health Quality Choice Report (Summary Data on Satisfaction and Outcomes)



In contrast to the Cincinnati and Orlando systems, the Cleveland program does not report comparisons of hospital charges. Although such data are of interest to Cleveland employers, hospital and physician representatives have resisted the inclusion of information on charges in the program. A task force is considering whether hospital charges should be compared in the future.

Of the employer-sponsored systems we reviewed, the Cleveland coalition is the only one to have formal safeguards to prevent hospitals from manipulating their data. The Cleveland program routinely audits a sample of records from each participating hospital. If a hospital is found to have overestimated its patients' severity of illness, its medical records are reabstracted.

Hospitals' Use of Cleveland Health Quality Choice Data

Several hospitals provided us with examples of changes they have made in areas identified for improvement with the Cleveland Health Quality Choice Program information:

- One hospital noted that it has focused on developing treatment protocols for ordering laboratory tests and prescribing medications. For example, the hospital has begun more aggressively treating pneumonia patients with antibiotics. Although this change may not directly reduce costs, the hospital believes that it will improve the quality of care for these patients.
- Another hospital recognized that it had not been using physical therapists efficiently for knee joint replacement patients. The hospital is attempting to shorten stays for knee joint replacement patients by extending the hours physical therapists are available, changing the time of the surgery, and involving discharge planners earlier in the admission.
- A Cleveland hospital is establishing clinical teams to develop treatment protocols for four conditions—pneumonia, chronic obstructive pulmonary disease, strokes, and congestive heart failure—for which the hospital had longer than expected stays.

As in the other communities we visited, several Cleveland participants said they expect that hospitals will become more specialized and many will affiliate. Several hospitals and employers noted that the Cleveland hospital market has too many hospital beds, but did not expect widespread hospital closings. Instead, hospitals and employers expect centers of excellence to be established for specific clinical services. Also, hospitals are expected to affiliate so that a patient can stay within a single hospital

system for a range of health services—from primary care through tertiary-level care.

Employers' Use of Cleveland Health Quality Choice Data

In most cases, employers are waiting until several Cleveland Health Quality Choice reports are available before identifying trends in performance and changing their purchasing decisions. In a few cases, payers have already used the Cleveland Health Quality Choice information:

- Aetna Health Plans of Ohio is using data from the Cleveland Health Quality Choice Program as a factor in establishing a more restrictive managed care network than it currently offers. The insurer's existing contracts with hospitals, claims data, and geographic location of hospitals were also important factors in selecting hospitals for this managed care network.
- One large employer indicated that the Cleveland Health Quality Choice information confirmed a decision it had previously made to remove a hospital from its managed care network.

It is difficult to attribute any cost savings to the Cleveland Health Quality Choice Program at this point because the first report was issued in April 1993, probably too recently for hospitals to have changed their practices in response to the program. However, the program's executive director estimated that participating hospitals have saved nearly $20 million per year because patient stays have been reduced by a total of 21,500 days. But he acknowledged that many factors contributed to this reduction. The Greater Cleveland Hospital Association noted that—even before changes associated with the program occurred—hospital costs rose more slowly in Cleveland than elsewhere in Ohio. The hospital association credits the increasingly aggressive health plan negotiations by business groups for this slower growth. Furthermore, a local insurer contends that it has begun negotiating more effectively with hospitals in its managed care networks.

Major Contributors to This Report

Rosamond Katz, Assistant Director, (202) 512-7148
John Dicken, Evaluator-in-Charge, (202) 512-7135
Paula Bonin
Hannah Fein
Trisha Kurtz

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (301) 258-4097 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Mail
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested