Appendix H.—Alternative Patient Classification Systems[1]

Introduction

In **1983,** when Congress changed the basis for Medicare payment to the prospective payment system (PPS), the Diagnosis Related Groups (DRG) system was the best patient classification system available that could adjust for case-mix differences among hospitals. Other patient classification systems were not ready for use because data were unavailable or there were other administrative and technical problems. Although it is unlikely that the DRG system will be replaced in the near future, alternatives to and refinements of DRGs are under development.

This appendix describes and compares DRGs and five other existing systems for classifying hospital inpatients. Each patient classification system has been constructed from a set of principles and data sources that gives it advantages and disadvantages in various uses. The alternatives to DRGs described could be used as substitutes for or modifications of the DRG system. This appendix lays out criteria for evaluation the usefulness of any patient classification system in per-case hospital payment,

Even if they are not useful for payment, however, alternative patient classification systems could be valuable tools in evaluating the impacts of PPS. The effects of PPS on the quality of care can be measured with greater validity if patients with similar medical problems can be identified in both the pre-PPS and post-PPS periods. Patient classification systems that have very homogeneous groups of patients may help in this regard. Also, the underlying causes of changes in the distribution of hospital revenues and surpluses can be identified with greater precision using refined indexes of illness severity or patient resource needs to separate variations due to a hospital's relative efficiency from those due to intrinsic differences in patient needs,

Patient Classification Systems

A number of alternative patient classification systems have been considered for use in a prospective payment system for hospitals. Six such systems are described below:
- Diagnosis Related Groups (DRGs), **1980** Version;
- Disease Staging;
- Severity of Illness Index;
- Patient Management Categories;
- Medical Illness Severity Grouping System (MEDISGRPS); and
- Acute Physiology and Chronic Health Evaluation, Simplified Version (APACHE II).

Key attributes of each system are summarized in table H-1. More technical details are presented in the discussions of evaluation criteria. (Additional details can be found in refs. **142,305.)**

Diagnosis-Related Groups (DRGs)[2]

Two sets of DRGs have been developed. The original set contained 383 categories, and the revised set on which Medicare's PPS is based contains 467 categories.[3] Both sets of DRGs were partially constructed from records of actual patient stays in a sample of hospitals. Creation of the DRGs relied both on statistical techniques to identify subgroups of patients with similar lengths of stay and on clinical judgment to confirm that patient subgroups also had compatible diagnoses. The second set was developed with greater reliance on the judgment of a panel of clinicians for the identification of clinically meaningful categories, although statistical analysis was still important in the development process.

Assignments of patients to DRGs are based on patient data available on patient discharge abstracts and, now, on Medicare's hospital bills. Principal diagnosis—i.e., the principal reason (after study) for a patient's hospitalization—was used to separate patients into 23 major diagnostic categories corresponding to organ systems. Other important variables are: 1) secondary diagnoses; 2) specific surgical procedures; 3) nature of coexisting conditions and complications; 4) discharge status (including death and "left hospital against medical advice"); 5) age; and 6) other factors for some categories. The data are commonly available in computerized form, and a computer algorithm (referred to as GROUPER) classifies each patient into the appropriate DRG, The diagnosis and procedure coding for DRGs is based on the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM), which is an internationally recognized medical coding system developed for statistical monitoring of hospitalized patients,

---

[1]This appendix is based on S. Williams, G. Kominski, and G. Pickens, "The Evaluation of Alternative Patient Classification Systems, " prepared for the Office of Technology Assessment, U.S. Congress, Washington, DC, December 1984.

[2]See also OTA'S 1983 publication *Diagnosis-Related Groups and the Medicare Program: Implications for Medical Technology* (305).

[3]Medicare uses 470, but the 3 additional categories are ones in which "other" cases are placed before being paid for.

Characteristic	Diagnosis-Related Groups	Disease Staging	Severity of Illness Index	Patient Management Categories	MED ISGRPS	APACHE II
Number of classes	467 + 3	1,600 PIUS	4 plus	1,000 plus	Unique score for each patient	Unique score for each patient
Source data	Discharge abstract	Discharge abstract	Medical record	Discharge abstract	Medical record	Medical record
Type of classification	Objective	Objective	Combined subjective and objective	Objective	Objective	Objective
Classification variables	Principal diagnosis Secondary diagnosis Age Principal procedure Secondary procedure Sex Discharge disposition	Principal diagnosis Secondary diagnosis Sex Surgery Discharge disposition	Diagnoses H I story and physical Laboratory tests X-rays Progress and nursing notes	Principal diagnosis Secondary diagnosis Reason for chief problem Elective procedure	Reason for admission Significant history (associated problems) Illness severity (based on key clinical findings)	Principal diagnosis Laboratory tests History and physical
Comorbidity included?	Yes—specific diagnoses	Yes—secondary diagnosis as comorbidity and complications	Yes—an y. factor that affects length of stay	Yes—secondary diagnosis as comorbidity and complications	Yes	Yes — multiple system impacts
Deaths omitted?	No	No	No	No	No	Yes
Outliers omitted?	Yes	No	No	No	No	No
Affected by errors in diagnosis coding on abstracts?	Yes	Yes	No	Yes	No	No
Automated?	Yes	Yes	No	Yes	Yes	No
Mutually exclusive and exhaustive?	Yes	Yes	Yes	Yes	Yes	No
When applied?	After discharge	After discharge	After discharge	After discharge	Day 3 and day 10 after admission	24 hr. after admssion

SOURCES Based on a table in M C Horn brook Techniques for Assessing Hospital Case Mix Ann. Rev. Public Health vol 6 in press 1985 and A C Brewster B G Karlin K A Hyde et al Medical Illness Severity Grouping System (M EDISGR Ps) A Clinically Based Approach to Class ifying Hospital Patients at Admission mimeograph 1984

Disease Staging'

Disease Staging is the product of physicians' judgments about the biological progression of a given disease. Each disease is defined by four required elements: location of the problem, manifestations of the medical problem, cause of the problem, and the severity of the problem. Staging does not depend on observed utilization patterns or expected responses to therapy. Patients may be staged for each disease manifested.

A subset of 420 disease conditions was developed for two purposes: 1) to include the major diseases in each body system, and 2) to cover the majority of typical hospital admissions. Physician experts developed criteria for staging each disease. The criteria assigned patients to one of the following four major stages (substages were also developed for many diseases):

- Stage 1: Conditions with no complications or problems of minimal severity;
- Stage 2: Problems limited to an organ or system, significantly increased risk of complications over Stage 1;
- Stage 3: Multiple site involvement, generalized system involvement, poor prognosis; and
- Stage 4: Death.

Staging patients according to the medical criteria developed by the physician panelists originally required medical records review, but the system now in widespread use assigns patients to categories using information in the Uniform Hospital Discharge Data Set (UHDDS).[5] The computerized system may underestimate the severity in some cases; for example, when an ICD-9-CM code in the UHDDS applies to more than one stage of a disease, the lower stage is selected.

Severity of Illness Index'

The Severity of Illness Index reflects the overall severity of illness of the patient, not just the severity

———. . . ..———

[See also J. S. Gonnella, M. C. Hombrook, and D.Z. Louis, "Staging of Disease: A Case-Mix Measurement, " 1984 (120).

[5]The Uniform Hospital Discharge Data Set (UHDDS) is a minimum basic data set for acute care hospitals developed by a group of experts for the U.S. National Committee on Vital and Health Statistics, based on the work of a conference held in 1969 at Airlie House. The Secretary of the Department of Health, Education, and Welfare (now Health and Human Services) adopted the UHDDS in 1974, but Medicare and Medicaid have resisted using it. However, the Professional Standards Review Organization (now PRO—see ch. 6 and app. G) program adapted UHDDS for its PRO Hospital Discharge Data Set, which includes UHDDS information plus some PRO-specific information (214).

[For additional information and references, see OTA's 1983 publication *Diagnosis-Related Groups and the Medicare Program: Implications for Medical Technology* (305).

of each diagnosis. This index is based on seven dimensions, each of which is divided into four levels, based on criteria developed by researchers at Johns Hopkins University in conjunction with a panel of physicians and nurses, The seven dimensions deemed to be the best indicators of overall illness severity are as follows:

- stage of the principal diagnosis on admission;
- complications of the principal condition;
- concurrent, interacting conditions that affect the course of hospital treatment;
- dependency on the hospital staff;
- extent of non-operating-room life-support procedures;
- rate of response to therapy, or rate of recovery; and
- impairment remaining after therapy for the acute aspect of the hospitalization.

Relevant data from patients' medical records are abstracted manually by trained raters at the same time the hospital discharge abstract is prepared. Based on the combined patterns of severity levels within each dimension, the rater judges the overall index of severity of the patient's illness from 1 (least severe) to 4 (most severe). A computerized system is being developed that can be used with modified recordkeeping systems. Clinical data are to be used to modify the present 5-digit ICD-9-CM coding system to record severity considerations into discharge abstract data as a sixth digit for each disease condition.

Patient Management Categories

Patient Management Categories were developed by panels of physician experts. The goal of this system was to specify patient attributes that identify patients with very similar needs for care. The categories are based on patient clinical characteristics and severity of the illness. The development of specific categories did not depend on empirical analysis of actual use of services, but the physician panelists did specify components of effective care in order to provide a basis for subsequently estimating the cost of providing care in each category. Components of care include diagnostic services, treatment procedures, and expected length of stay, which in the view of the physician panels are required for effective patient management (i.e., patient management paths).

Over 750 Patient Management Categories have been defined for patients in acute care hospitals. A computerized algorithm can be used to map cases into the categories from information on discharge abstracts.

——

[See Blue Cross of Western Pennsylvania, *Hospital Case Mix: Development and Implementation, 1983 (38).

The software criteria include combinations of diagnoses (ICD-9-CM codes) and specific surgical procedures when necessary. Sequencing of the codes on the abstract does not affect category assignment. Age and sex are used to categorize patients in only a few instances.

Medical Illness Severity Grouping System (MEDISGRPS) [8]

MEDISGRPS is a patient classification system that groups patients by severity levels on the basis of data acquired upon admission. Key clinical findings were selected by the group of researchers who developed the system. These key clinical findings include the results of laboratory, pathology, or radiology tests or physical examination findings that indicate illness. Test results, not their interpretations, are used. Each finding is assigned to one of the following severity groups:
- Severity group O: No findings;
- Severity group 1: Minimal findings;
- Severity group 2: Severe or acute findings;
- Severity group 3: Severe and acute findings; and
- Severity group 4: Critical findings.

A computer program can be used to assign severity groups to each patient based on data from the medical record. The system is currently being tested in six sites.

Acute Physiology and Chronic Health Evaluation, Simplified Version (APACHE 11) [9]

The APACHE classification system was developed to answer clinical questions about intensive care unit (ICU) patients. Data for the APACHE system are collected within 24 hours of ICU admission, rather than after hospital discharge as is typical of most other patient classification systems. The data used to assign patients are not available on discharge abstracts, and the extent of availability of the necessary data items from the medical records of non-ICU patients is unclear.

The original APACHE system consisted of an Acute Physiology Score based on 34 physiologic variables and a chronic health assessment, which is a four-category scale. The variables and the weights used to incorporate them into a total score for a specific patient were chosen by a group of ICU clinicians.

The simplified version, APACHE II, is based on 12 of the most commonly used physiologic measures of the original 34 variables. Again, clinicians chose the 12 measures based on their judgment of the validity

————..—.——

'See J.A. Brewster, C.M. Jacobs, and R.G. Bradbury, "Classifying Severity of Illness Using Clinical Findings, " 1984 (43).

'See D.P. Wagner and E. Draper, "APACHE II and Medicare Reimbursement, " 1984 (374).

and specificity of the measure, breadth of organ system coverage, and objectivity, reliability, and frequency of measurement. Most of these measurements are available for the majority of ICU patients shortly after hospital admission. The weighting scheme was also slightly modified for APACHE II. Data collection and analysis for payment purposes on large national samples of non-ICU patients have not been done.

Criteria for Evaluating the Usefulness of Patient Classification Systems in Prospective Payment

Six criteria that could be used to evaluate alternative patient classification systems for the purpose of paying hospitals for patient care are discussed below. For each criterion, the following information is presented: the rationale for its selection, how to conduct an evaluation using the criterion, a brief review of any evaluations that have been conducted, and speculations about how candidate patient classifications might perform when evaluated.

Criterion 1: How Well Does the Patient Classification System Account for Variations in Patient Needs for Hospital Services?

Prospective payment requires a method to adjust payment rates for differences among patients in their real need for resources. Without such an adjustment, hospitals would receive the same rate of payment for all patients, and strong incentives would exist for hospitals to treat only the less costly patients. Hospitals which, because of their mission, could not discriminate among patients in this way would suffer unfair financial burdens.

A patient classification system can comprise either a fixed number of discrete categories, each with a separate weight, or a scoring system which calculates a unique score for each patient from a formula based on individual patient characteristics. The, objective of the patient classification system is to reduce as much as possible the difference between the amount paid for a given patient and the costs of needed care. If the price paid for a patient is based on the category weight or the patient's score, then the performance of the system depends upon the residual variation between patients' needs and the category weights. If there is high variation because patients with very different needs for care are contained in the same category (and therefore paid at the same rate), then the classification system can be said to perform poorly on this critical dimension. Thus, the primary criterion for evaluating the usefulness of any patient classification system for

per-case prospective payment is the extent to which the system reduces the total variation in resource needs around the mean by forming subgroups, each of which has its own mean value.[19]

Identifying Patient Needs.—Unfortunately, a prescriptive measure of need for services is difficult or impossible to construct for most patients. Also, the complete separation of resource need from actual resource use in the construction of a patient classification system is probably infeasible.

There is an extensive literature on the objective measurement of health status or, conversely, medical needs. The basic measurement issue is the correlation between objective indexes of health status and clinical assessments of needed services. The capability of general health status instruments to measure patient needs is at best unclear. Some investigators have developed general health status measures for use in clinical settings (32,226,256). Although some studies have found general health status measures to be sensitive to the provision of medical care when there are dramatic changes in clinical status (247), other studies have failed to confirm such sensitivity in more common situations **(81).**

Recently, attention has been focused on specific patient groups, because resource needs are more likely to be measured successfully in patients who are similar. Intensive care patients form one such group, and APACHE II has been developed for them (374). The index based on APACHE II is significantly related to observed nursing resource use. Because APACHE II was developed only for patients who require intensive care, it is uncertain whether APACHE II will be useful as a general patient classification system. The MEDISGRPS classification system measures need using clinical laboratory, radiology, pathology, and physical examination findings. In these respects, MEDISGRPS attempts to generalize the approach used by the APACHE II system to nonintensive-care patients.

The developers of the six patient classification systems reviewed in this appendix recognized resource needs indirectly, often by attempting to define groups of patients with similar patterns of resource use. The developers of the DRG system, for example, used a combination of clinical judgment and data on variations among patients in length of stay (one measure of resource use) to develop the classification system. The Disease Staging system recognizes resource need indirectly by assuming that the severity of illness within a disease condition was correlated with need,

but there was no direct reference either to the resource needs of patients or resource use in the construction of the system. Patient Management Categories were defined with clinically acceptable management patterns in mind. To the degree that the physicians who defined categories accurately specified appropriate services for each group, Patient Management Categories appear to be more strongly linked to patient resource need than other systems. The Severity of Illness Index attempts to measure resource need directly (such as the need for nursing services), but need is measured after examining resource use, and it is unclear how the Severity of Illness Index separates need from use.

Although desirable, it is not possible at present to identify a general measure of patient need for hospital resources that is sufficiently accurate. Furthermore, there is little expectation for substantial improvement in the near future. Until measures of patient need are developed, patient use of hospital resources, as measured by length of stay (LOS) and total cost, not the need for hospital resources, will continue to be used as the outcome variable in evaluation of this criterion.

The most commonly used measures of resource use are LOS, total charges, or estimated cost per admission. These measures appear to be highly correlated with one another. For example, the correlation between LOS and variable cost was high when estimated for a sample of patients in 77 New Jersey hospitals (168). The Pearson correlation coefficients were 0.84 for the total sample of patients and 0.78 for Medicare LOS inliers.

The methods used to estimate the cost of resources used can vary widely, Most studies have relied on cost allocation techniques used with hospital accounting data. Other approaches exist, but they require the collection of data that are not provided by routinely available data sources. For example, microcosting, which involves direct observation of resource use in a sample of patients, could be used to measure patient costs more accurately than estimates of per-case costs based on accounting data (397). Because they are more accurate, costs derived by microcosting techniques may be a useful alternative when classification systems are evaluated. Microcosting is, however, expensive and must be repeated periodically to reflect changes in hospital operations. Moreover, a high proportion of hospital costs are fixed and must be allocated to patients indirectly even with microcosting techniques. Despite these drawbacks, many hospitals are implementing microcosting techniques, and enough data might be available from them to be used in comparative evaluations of patient classification systems.

Statistical Issues.—For reasons enumerated above, the central question in comparative studies of patient classification systems has been: "How much reduction

---

[19]When the classification system calculates a unique score for each patient, the performance criterion is the extent of variance of the patient's true need for resources from the calculated score.

in the variance among patients in resource use is achieved by candidate patient classification systems?" Ideally, this comparison should occur on a nationally representative sample of discharges using the same dependent variables and the same statistical techniques. Few such studies have been conducted. For the most part, developers of each classification system have published studies that document the variance reduction capabilities of their own systems, or limited comparisons have been performed on local samples, often restricted to only a few patient categories. This is understandable, because most classification systems have been created recently and their developers must establish the reliability and validity of their products. The most extensively studied system has been the DRG system, in part because reduction in variation measures were used to create it (33,104,230).

Gonnella, Hornbrook, and Louis (120) have evaluated the performance of the Disease Staging system on a national stratified sample of patients who were discharged from 377 hospitals in 1977. In this study, attention was focused on LOS variation in diabetes mellitus cases. Mean LOS by stage, patient characteristics, and hospital characteristics were reported. The assessment was limited to pairwise LOS contrasts for patient and hospital subgroups by stage. No overall measure of variance reduction was provided. This patient classification system performed well in the limited situation in which it was tested.

Ament and colleagues (9) have reported comparative studies of the DRG and Disease Staging systems. Using a database for 50 hospitals from which extreme outlier cases had been removed, they found a 49-percent reduction in variance (measured as R^2)[11] for DRGs and a 35-percent reduction for Disease Stages.

Two studies (62,235) have reported a comparative evaluation of DRGs and Disease Staging in Maryland hospitals. Using a sample of DRGs, they found variance reduction values of about 15 percent for DRGs and 12 percent for Staged Disease Conditions and Disease Stages. Finally, an attempt to combine DRGs, Staged Disease Conditions and Disease Stages resulted in a reduction in variance of about 21.1 percent.

Computerized Patient Management Categories are only now being released, so there is little direct evidence about their variance reduction properties. The precursor to Patient Management Categories was studied by Horn, Sharkey, and Bertram (138) and found to have variance reduction values between 31 percent and 50 percent in a restricted sample.

[11] R^2 is the proportion of the total variance around the mean that is explained by a set of linear variables.

Horn, Sharkey, and Bertram **(138)** have compared reduction of variance measures for DRGs, the Severity of Illness Index, Disease Staging, and a precursor to Patient Management Categories in a sample of four hospitals and for limited groups of disease conditions. DRGs result in R^2 values in the 10- to 20-percent range for total charges in patients with gall bladder disease and congestive heart failure. The other classification systems are comparable, except that the Severity of Illness Index has higher R^2 values, especially when combined with the DRG system. Horn, Horn, and Sharkey (137) also have compared the performance of DRGs and the Severity of Illness Index in eight broad categories of disease in a sample of 14 hospitals using total charges. Percent reduction in variance for DRGs varied between 17 and 31 percent. The range for the Severity of Illness Index was 40 to 95 percent but was lower when outliers (i.e., cases with extremely high or low charges) were excluded **(139).**

In summary, although comparisons of variance reduction have been carried out for some of the classification systems, none of the evaluations has compared all the systems on a nationally representative sample of data. Thus, it is unknown at present how the alternative patient classification systems perform on this critical dimension.

Criterion 2: How Robust Is the Patient Classification System?

"Robustness" refers to the ability of the system to maintain its properties despite changes in the data used to create or operate the system. Roughly speaking, a system is robust if it preserves the same categories when different databases are used for its creation and if it preserves the same level of proportional reduction in variance when applied to different databases. Little is known about the robustness of most classification systems, but some speculation is possible.

Some of the classification systems relied on specific databases for their development. The DRG system is the best example. The procedure used to create the DRG system was partly based on reduction of variance in LOS, which is sensitive to the presence of outlier cases. Although regression techniques exist for the analysis of data containing outliers (149), these techniques were not used to develop the DRG system and apparently have not been used to develop any other system, probably because they are computer intensive and may not be suitable for large data sets. The expense and effort required probably prohibit a study that would examine DRG robustness by trying to reproduce the system with an alternative database (406).

Also, since DRGs were developed using expert advice, robustness of DRGs would need to be tested using a sample of expert panels.

Robustness also is an issue for systems that are created from expert judgment. For these systems, which include Disease Stages, the Severity of Illness Index, Patient Management Categories, and MEDISGRPS, robustness could be tested by determining if different expert panels would reproduce the original categories using the same principles. This evaluation also is probably unrealistic, considering the expense and effort that would be required.

Another aspect of robustness is the ability of a given system to classify patients reliably when different databases are used, for example, databases containing patients in different age groups or from different geographical regions. The study by Pettengill and Vertrees **(230)** suggests that the DRG system is sufficiently robust across geographic regions for it to be used in Medicare's PPS, but there are no published studies about the geographic robustness of the other patient classification systems.

Because of concerns about the DRG system's ability to classify children appropriately, most pediatric patients are excluded from Medicare's PPS. The National Association of Children's Hospitals and Related Institutions is conducting a study to determine if the DRG system is robust enough to be used for pediatric patients **(217),** and a special version of Disease Staging has been developed for pediatric patients.

One aspect of robustness that can be studied with existing data is the ability of alternative patient classification systems to preserve reduction in variance when applied to databases that include or exclude outliers, as defined by Medicare, The inclusion of outliers greatly increases overall variance. If outliers are to be paid outside the system, this means that alternative systems must be evaluated for their robustness for inliers only.

Criterion 3: How **Reliable Are Patient Assignments to Category?**

Previous studies have reported large random error rates in hospital abstracts **(123),** and there are many random errors in hospital charge data as well. These errors probably are less serious than they seem because data reporting will improve now that both hospitals and Medicare are monitoring reports to ensure accurate payments. Because of their tendency to cancel each other in relatively large data sets, random errors are less important than systematic errors **(230), al-**though selective correction of these errors could allow hospitals to report more of their patients in categories with high payment rates, which has been called "DRG creep" (267). Even if we assume that random errors are not a problem, however, data reporting could lead to problems of reliability and there could be important differences among the classification systems.

If patient assignments are reliable, each patient will be assigned to the same category when assignments are made more than once, even if assignments are made by different people under different circumstances. Reliability should be assessed at several levels.

One such level concerns the primary data. The more objective are the data used to assign patients to classifications, the more likely the assignments are to be reliable. For example, to the extent that a system, like MEDISGRPS, uses data directly from the clinical laboratory, it is more likely to be reliable than a system, like the Severity of Illness Index, that requires subjective judgments about the patient's illness. Some of the possible data elements are listed below in decreasing order of perceived objectivity (with few data available to establish exact order):

- sex, age, and the values for vital signs and laboratory test results;
- reports of imaging procedures, clinical laboratory examinations, and the findings at surgery and other procedures;
- primary or principal diagnoses;
- secondary diagnoses;
- symptoms reported by the patient and signs observed by clinicians; and
- judgments and interpretations, for example, about the relative severity of illness or the response to therapy.

Another level at which reliability should be assessed involves the integration of data into category assignment. Assignments based on specific criteria that can be incorporated into computer algorithms are more likely to be reliable than assignments that require subjective judgment.

Criterion 4: **How Does the Patient Classification System Affect Economic Incentives Regarding Technology Use?**

In general, patient classification systems can be separated into two types. The first type distinguishes between the clinical condition of the patient and the resources used to treat the condition by excluding all measures of resource use in the assignment of patients to categories. Disease Staging, MEDIGRPS, and APACHE II are three examples of this type. The second type combines measures of resource use with clinical information to assign patients to categories. DRGs and the Severity of Illness Index are examples of this type.

When the category **to** which **a** patient is assigned depends on whether or not a technology is used, hospital managers' and physicians' incentives for technology adoption and use can be very different from the incentives under systems in which assignment is unrelated to technology use. The DRG system, for example, classifies patients according to whether specific surgical procedures are performed. Depending on the relative payment rates associated with medical and surgical categories, the DRG system can encourage the use of some surgical procedures. Also, the DRG assignment of patients can change as new technologies are applied to patients. In patient classification systems in which assignment is neutral with respect to the kinds of technologies used, only technologies that reduce the cost per case are encouraged by the financial incentives of the system.

A more subtle example of technology-dependent assignment is the recent finding that the homogeneity of several surgical DRGs with respect to resource use could be substantially improved if they were differentiated on the basis of admission status (i. e., emergency vs. non-emergency). Because emergency admissions have significantly higher costs and lengths of stay, incorporating admission status into the definition of these DRGs would create an incentive for hospitals to classify as many patients as possible as emergency admissions. Unless more objective measures of emergency status can be found, improving the homogeneity with respect to resource use of the DRGs would encourage hospitals to label more admissions as emergency.

Criterion 5: **To What Extent Is the Patient Classification System Meaningful to Clinicians?**

In any classification system, patient categories should be consistent with generally accepted clinical practice. There are several reasons. Clinicians must understand the categories if they are to respond rationally to the inherent incentives of per-case payment. Clinical diagnoses (and levels of severity within diagnoses) dictate hospital management of resources. Clinical diagnoses supplemented with measures of prognosis and severity represent the most highly developed methods for determining a patient's need for hospital care. Finally, classification systems might affect clinical practice adversely if they do not recognize acceptable clinical practice in their category assignments. Being consistent with clinical practice, however, does not mean that the classification system must institutionalize current practice. When current practice and payment policy differ, users of the classification system should be able to relate the system's categories to the categories used for patient management.

Because alternative classification systems have been developed at different times over the last several years, and because development has relied on different physician experts whose qualifications for representing a national consensus often are unspecified, there likely is considerable variation in how well different systems reflect current clinical practice. The **1980** version of the DRG system and APACHE II, for example, reflected the consensus of panels of clinical experts drawn from national samples. The other systems relied on less representative authorities who may or may not have reflected the wide diversity in clinical practice that characterizes the nation's hospitals.

Criterion 6: **How Feasible and Expensive Is the System To Use in Prospective Payment?**

One effect of Medicare's PPS has been to accelerate hospital adoption of computer-processing capabilities. This widespread capability has made it feasible to consider classification systems that use more detailed data than do current ones. Therefore, differences in feasibility and expense among classification systems are probably limited to differences in the type of data required for category assignment and the process used for revision and recalibration, but not to differences in the number of categories in and of themselves. The required data elements should be available on existing hospital reporting systems. The following reporting systems can be used for category assignment, in descending order of availability, with availability determined largely by the extent to which these reports are managed with computer-based methods:

- the patient's itemized bill, perhaps modified by accepted accounting practices, such as Medicare's cost-to-charge ratio or more detailed microcosting methods;
- the Uniform Hospital Discharge Data Set (UHDDS);
- results of diagnostic tests, records of vital signs and other monitored indicators of clinical status, pathology reports, operative reports and other data that are processed by computer in some hospitals;
- the patient's medical record, which is available only on paper;
- patient characteristics that are not gathered as a routine part of the record, for example, the level of social support available to the patient after discharge; and
- the intensity and amount of hospital services that are not commonly measured in most hospitals, for example, nursing services.

Regardless of how data are collected, assignment of patients to categories should be capable of being done by computer to minimize cost and enhance feasibility y.

Conclusions

The DRG system was not selected for use in Medicare's PPS by accident. It was (and remains at present) the classification system that most fully met two critical criteria: reduction in variation among patients in resource use, and feasibility. Though other systems may perform better by creating more homogeneous patient groups, they are not easily implemented or they lack objectivity in coding.

The relative cost and feasibility of alternative methods are changing as hospitals become more sophisticated in automated information processing. In the future, then, other systems that rely on more detailed data may become viable alternatives or refinements of DRGs.