NATIONAL DRG VALIDATION STUDY
SPECIAL REPORT ON CODING ACCURACY

OFFICE OF INSPECTOR GENERAL

OFFICE OF ANALYSIS AND INSPECTIONS

FEBRUARY 1988

The Office of Inspector General

The mission of the Office of Inspector General (OIG) is to promote the efficiency, effectiveness and integrity of programs in the United States Department of Health and Human Services (HHS). It does this by developing methods to detect and prevent fraud, waste and abuse. Created by statute in 1976, the Inspector General keeps both the Secretary and the Congress fully and currently informed about program or management problems and recommends corrective action. The OIG performs its mission by conducting audits, investigations and inspections with approximately 1,200 staff strategically located around the country.

Office of Analysis and Inspections

This special report is produced by the Office of Analysis and Inspections (OAI), one of the three major offices within the OIG. The other two are the Office of Audit and the Office of Investigations. The OAI conducts inspections which are typically short-term studies designed to determine program effectiveness, efficiency and vulnerability to fraud and abuse.

This Report

Entitled "National DRG Validation Study - Special Report on Coding Accuracy," this study was conducted to determine the extent of coding errors across a random sample of all hospitals paid under the Prospective Payment System. The report was prepared by the following staff of the Office of Analysis and Inspections:

Headquarters

Cathaleen A. Ahern
David C. Hsia, J.D., M.D., M.P.H.
Wm. Mark Krushat, M.P.H.
Brian M. Rawdon
Barry L. Steeley
Jane C. Tebbutt, M.S.
Nicole Bass
Patti Snyder

Region I

Joyce M. Greenleaf, M.B.A.
Cindy A. Piltch, M.P.H.

Region V

Natalie A. Coen
Suzanne G. Johnson
Margaret Shell

Region IX

Kathy L. Admire
Leonard Czajka
Kaye D. Kidwell, M.A.

NATIONAL DRG VALIDATION STUDY
SPECIAL REPORT ON CODING ACCURACY

RICHARD P. KUSSEROW
INSPECTOR GENERAL

OAI-12-88-01010   FEBRUARY 1988

EXECUTIVE SUMMARY

The Office of Inspector General (OIG) studied the accuracy with which hospitals reported the diagnoses and procedures performed on Medicare beneficiaries under the prospective payment system (PPS). A two-stage cluster design sampled 7,050 medical records from 239 hospitals, stratified by size. Medical records specialists reabstracted the information in the medical record for each case to arrive at the correct diagnosis-related group (DRG), on which payment is based.

The wrong DRG was originally assigned, based on hospital coding, in 20.8 percent of the cases reviewed. Smaller hospitals made significantly more errors than mid-size and large hospitals. A significant percentage of the errors (over 61 percent) favored the hospitals, that is, the hospitals were paid more for the hospital stay than they would have been if the correct codes had been submitted. Previous studies had found the direction of errors to be random, overpaying and underpaying the hospitals about equally.

Most errors (48 percent) derived from physician designations of diagnoses or procedures, which, although incorrect or insupportable, were not corrected before a claim for payment was submitted. Other causes of these errors include incorrect sequencing of diagnoses or codes (27 percent) and incorrect coding (12 percent).

These errors caused an overall 1.3 percent overpayment in the sample, which, if projected to all PPS hospitals, would represent $308 million in excess payments in Fiscal Year 1985.

TABLE OF CONTENTS

INTRODUCTION

From the inception of Medicare in 1965 until 1983, hospitals had been paid on a reasonable cost basis for caring for Medicare beneficiaries. They were reimbursed for their expenditures, based on cost reports submitted annually. The more services a hospital rendered, and the more those services cost, the more the hospital received in payment from Medicare. Third-party payers, including Medicare, tended to isolate both the patient and the provider of care from the effects of steadily increasing hospital expenditures. Many of these increases were due to new techniques and procedures, with life-saving and life-enhancing consequences. Others were due, however, to a system of payment which rewarded inefficiency. As medical expenditures outpaced general inflation rates, various limits were put on the amounts that would be reimbursed, but the cost-based methodology remained the same.

ORIGIN OF THE PROSPECTIVE PAYMENT SYSTEM

Looking for ways to reverse the trend toward higher health care expenditures, the Health Care Financing Administration (HCFA) funded various studies, one of which resulted in the diagnosis-related group (DRG) system of classification. Robert B. Fetter, John D. Thompson and their colleagues at Yale University's Center for Health Studies developed DRGs as a framework to monitor quality of care and utilization review in institutional settings. Their original goal was to define a "case" in terms of resource consumption. These DRG categories would have to represent clinically coherent groups of patients with similar patterns of resource use. Since treatment, as a rule, is based on diagnosis, this became the logical starting point. More than a third of a million hospital records were abstracted to analyze the patients' condition, age, the presence or absence of complicating illnesses, the length of stay and the treatment rendered. The information was analyzed by statistical means, with length of stay used as a surrogate for resource consumption, but decisions were made by physicians as to the clinical relationships identified. The resulting DRG categories, which were refined several times during development and field testing, represent groups of patients who can be statistically shown to require, on the average, similar levels of care. Not every instance of a given DRG would require the same expenditure of hospital resources as the next of that DRG, but on the average, DRGs can be ranked based on the "weight" given to the intensity of care required by patients in that DRG. In 1983, there were 467 DRGs based on diagnosis, and 1 DRG which represented instances in which the surgical procedure performed was not related to the principal diagnosis (e.g., a patient admitted for respiratory problems who slipped and broke a bone, which was then surgically repaired).

On October 1, 1983, HCFA implemented a wholly new approach to paying hospitals for the inpatient care of Medicare beneficiaries. The Congress had mandated a prospective payment system (PPS) based on DRGs, with rates for each DRG determined prospectively.

The PPS is responsible for reimbursement for Medicare patients at more than 5,400 hospitals nationwide. Another 2,000 hospitals or specialty wings of hospitals are exempt from PPS. These include those specializing in the treatment of psychiatric disorders, physical rehabilitation and, until recently, drug and alcohol treatment.

With the change to PPS, hospitals receive a pre-established payment for each discharge based upon its DRG. As described above, the DRG payment represents an average cost for patients having similar diagnoses. Some discharges consume more services (i.e., cost more) than the payment, while others use less. The hospital retains any surplus from discharges costing less than their DRG payments and absorbs losses on those consuming more services than the payment. An assumption underlying PPS is that a fixed payment per discharge would induce hospitals to implement economies and reduce the unnecessary services previously associated with the retrospective reimbursement system. At the same time, the total payments to the hospital should provide, on the average, the same resources for its patients as the cost-based system.

The PPS also simplified reimbursement administration and program control. For example, the case-mix index (CMI) describes, in a single measure, the complexity of a hospital's discharges, because it represents a weighted average of all the DRGs treated at that hospital. Hospitals that attract a sicker patient population, and, therefore, need to provide a greater degree of care, have higher CMIs. The PPS adjusts reimbursement to compensate for the more complicated patient mix. The CMI also permits administrators to monitor health care delivery, at least to the Medicare population, via previously unavailable comparisons both between hospitals and of a single institution over time.

DRG IMPLEMENTATION

Implementation of PPS depended upon accurate coding using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). This three-volume reference uses a five-digit code to classify approximately 10,000 diseases by anatomic and pathologic taxonomy, and a four-digit code to describe about 1,080 types of diagnostic and therapeutic procedures. The National Center for Health Statistics originally

sponsored the development of ICD-9-CM to automate the collection of morbidity and mortality statistics for epidemiological research.

The PPS radically changed ICD-9-CM's use from research to claims processing. Two areas of confusion were created with this new application. The first is that ICD-9-CM, which was created for research purposes, is used in the construction of DRGs. The second is that DRGs were initially created for reviews of quality and utilization. This has inevitably led to difficulties when used for payment purposes. Formerly, the selection of one diagnosis or code over another would have little effect on the mass of data collected for epidemiological or other research purposes. Now these choices have an impact that can affect not only individual hospitals but the yearly Medicare inpatient expenditure of $27 billion. The Office of Inspector General (OIG) has issued a number of individual reports focusing on coding problems within specific disease entities. This study examines coding errors across a random sample of all hospitals paid under PPS.

The HCFA calculated "relative weights" for 468 DRGs from historical Medicare cost and charge data. The relative weights came from a 20 percent sample of 1981 Medicare cases from each participating hospital and matching cost reports. Thus Medicare weights represent the actual care rendered to Medicare patients. Weights could be different for other populations because of their varying age, sex, health status, etc. At the inception of PPS, the relative weights theoretically averaged 1.0000 (after adjustments for DRG frequency). Those DRGs consuming more resources have higher relative weights. For instance, a cardiac valve procedure with pump and cardiac catheterization (DRG 104) bore a relative weight of 6.8527, whereas malignancy aftercare (DRG 465) carried a relative weight of 0.2071 in 1983. In this report, we discuss our findings in terms of the weights in use during Fiscal Year (FY) 1985, the period under review. The relative weights have changed each year in response to changes in hospitals' practices. For example, if most simple cases of a given procedure are now performed in an outpatient setting, the relative weight for the remaining (more complex) inpatient cases will increase. Conversely, a very complex procedure may have become less time- or resource-intensive due to advancing technology, and now have a lower relative weight.

A hospital's payment depends on the product of the DRG's relative weight times the hospital's blended payment rate, which takes into account the hospital's specific historical costs. The blended rate in effect in FY 85 provided for a hospital's historic costs to represent 50 percent of its rate, while the federal portion was based 75 percent on regional rates

3

(recognizing varying market-basket costs, area wage levels, etc.) and 25 percent on national rates. The blended payment rate will be phased out as PPS is fully implemented, and most hospitals will be paid based on the same national rate. Adjustments are made to the blended payment rate for area wage levels, urban versus rural locations, indirect medical education and other factors.

Additional payments can be made for discharges which are statistical "outliers," requiring much longer than average lengths of stay or incurring much higher than average costs. Some types of hospitals, such as sole community providers, rural referral centers and those which treat a disproportionate share of Medicare and Medicaid patients, receive supplemental payments of various sorts. Yet other costs are "pass-throughs," at least temporarily. These include direct medical education and capital costs.

The HCFA calculates standardized amounts, which represent the cost of an average Medicare discharge (i.e., the payment for a case having a relative weight of 1.0000). In 1985, the national standardized amount averaged $2,972 in urban areas and $2,358 in rural areas. Congress periodically modifies the standardized amount to reflect changes in the costs of delivery of health care or for other reasons it judges appropriate.

Processing a Medicare claim for payment commences with the patient's discharge from the hospital, so cases under PPS are frequently referred to as "discharges," but also include patients who die, are transferred or leave the hospital against medical advice. At the time of discharge, the attending physician (1) lists the principal diagnosis, secondary diagnoses and any inpatient procedures on the front of the chart; and (2) signs an attestation certifying the correctness of these statements. The hospital then assigns ICD-9-CM codes to all diagnoses and procedures for each discharge, using the rules of the Uniform Hospital Discharge Data Set (UHDDS) and the coding conventions known to Accredited Record Technicians (ARTs) and Registered Record Administrators (RRAs), the professional personnel trained in management of medical records and use of coding systems. These codes are shown on the "face sheet" of the medical record and on the claim for payment from Medicare.

Local fiscal intermediaries (FIs) contract with HCFA to process these codes as part of the hospital claim. The FI applies the Medicare Code Editor, a program which checks for valid codes, coherent data, and coverage by Medicare. Claims that fail these edits are returned to the hospital for correction. Using the hospital-generated codes, the FI processes each record through a computer program, named GROUPER and mandated by HCFA, that

4

returns the appropriate DRG. In addition to codes, the grouping process also considers the patient's age, sex, complicating conditions and status at discharge. A related computer program named PRICER adjusts for the area wage level, hospital-specific historic costs, etc., and assigns a payment amount to the DRG.

Background

With the change from retrospective reimbursement to prospective payment, the OIG began a number of studies aimed at understanding PPS and its effects on utilization and provider behavior. Among our concerns was that inappropriate provider behaviors, intended to maximize reimbursement, could place Medicare beneficiaries at unnecessary risk and cause inappropriate charges to HCFA.

The OIG anticipated the possibility that hospitals seeking maximum allowable payments might engage in impermissible coding practices. This "gaming" could take the form of coding diseases that the patient did not have or that the attending physician had not attested to, changing the order of the ICD-9-CM codes, choosing a principal diagnosis other than the condition responsible for admission to the hospital, or picking a non-specific diagnosis. A hospital that influences the coding process in an improper manner to increase its revenue engages in a practice known as "DRG creep."

In contrast, choosing among viable, alternate principal diagnoses does not violate ICD-9-CM priority rules. However, this "optimization" still has the potential to increase reimbursement and contribute to overall increased costs to the government. Various attempts have been made to measure the components of the increase in CMI among all PPS hospitals. The weighted average CMI for FY 84 PPS bills nationally was 1.1389; by FY 85, it was 1.1739, and preliminary FY 86 data showed an increase to 1.2066. As mentioned earlier, part of this increase is due to the shift of some services to an outpatient setting, and to more attention being paid to precise and thorough coding. The CMI increases may also be due to the advent of new, more intense treatment regimens for diseases affecting the elderly. This inspection was designed to measure the increase due to inaccurate or improper coding.

Methodology

Based on the National DRG Validation Study sample, the OIG compared the codes paid by Medicare with the codes supported by the underlying medical record. Where inconsistencies occurred, the study investigated their causes, effect on DRG assignment, resulting influence on reimbursement, and changes in the CMI. The study also measured whether systematic coding changes were biased in favor of either the providers or Medicare.

5

This study used a stratified, two-stage, cluster sampling design based on hospitals. The first stage used simple random sampling without replacement to select 80 hospitals from each of three groups based on bed size: small (fewer than 100 beds); mid-size (100 to 299 beds); and large (300 or more beds). The three criteria that defined the hospitals in the sampling frame were: they were acute care, short stay facilities; they were not located in a PPS waiver state (Massachusetts, Maryland, New Jersey, or New York); and, they had participated in the Medicare program since at least 1981 (and therefore had contributed to the establishment of the DRG weights).

SAMPLE DEMOGRAPHICS BY HOSPITAL SIZE

BED SIZE	< 100	100-299	300 +	TOTAL
TOTAL NUMBER OF HOSPITALS	79	80	80	239
Urban	16	56	75	147
Rural	63	24	5	92
Teaching	2	15	45	62
Nonteaching	77	65	35	177
Profit	7	14	2	23
Nonprofit	72	66	78	216
TOTAL NUMBER OF DISCHARGES	2276	2388	2386	7050
Urban	454	1676	2244	4374
Rural	1822	712	142	2676
Teaching	60	449	1316	1825
Nonteaching	2216	1939	1070	5225
Profit	210	419	60	689
Nonprofit	2066	1969	2326	6361

quality control process revealed no significant discrepancies in the reviewers' code assignments.

The results of this recoding process, analyzed for statistical significance, form the basis of this report. (Statistical methods are discussed in the Appendix.) Additional review by nurses and physicians identified concerns with the quality of care rendered under PPS. Subsequent OIG reports will discuss these concerns in detail.

SAMPLING FRAME BY HOSPITAL SIZE

BED SIZE	<100	100-299	300+	TOTAL
NUMBER OF HOSPITALS				
Population	2536	1603	774	4913
Sample	79	80	80	239
NUMBER OF DISCHARGES				
Population	1522000	3105000	3649000	8277000
Sampled Hospitals	18199	59481	144716	222396
Sample	2276	2388	2386	7050

FINDINGS

ERROR RATE OF 20.8 PERCENT

The DRG assigned by the FI to the codes provided by the hospital was found to be incorrect in 1,374 discharges, or 20.8 percent of the discharges reviewed when weighted by the proportion of hospitals in each bed size stratum. The error rate was 16.6 percent in large facilities, 18.4 percent in mid-size facilities and 23.6 percent in small facilities. The small hospitals had

The number of hospitals available for study was 4,913. Of the initial sample of 240 hospitals, one facility terminated its Medicare eligibility before the actual collection of medical records. Therefore, the first stage sample included 239 (4.9 percent) randomly selected hospitals meeting the three criteria mentioned above. The second stage of the design employed systematic random sampling to select 30 Medicare discharges from each of the 239 hospitals for the period October 1, 1984 to March 31, 1985. If the hospital discharged fewer than 30 Medicare patients during that time, all Medicare discharges for the period were selected for review.

In mid-1986, the Office of Inspector General requested copies of the medical records for the sampled discharges. Administrative subpoenas compelled the cooperation of a few institutions. Of the 222,396 records available from the hospital sample, the design selected 7,076. The reviewers ultimately received and reviewed 7,050 (99.6 percent) medical records. Payment records from the FI's files documented the DRG as billed and the amount of payment.

The OIG then contracted with the Health Data Institute of Lexington, MA, to perform DRG validation on the sample records. Clerks taped over the hospital ICD-9-CM codes on the face sheet and physician attestation on all records. An ART reviewed each chart and translated the supportable diagnoses and procedures into ICD-9-CM codes. An RRA supervised the coding team, providing guidance on the application of coding rules and resolving coding issues as they arose. The official GROUPER program processed the codes to determine the correct DRG.

Coders had instructions to ignore marginal problems or honest differences in judgment about appropriate coding. If the newly abstracted DRG nevertheless differed from the DRG paid by HCFA, a member of a panel of physicians evaluated the record. The physician did not know the assigned codes during the review. If the physician felt that the discharge record in question presented complex classification problems, a physician committee resolved the issue. The physicians reviewing coding discrepancies had Board certification in appropriate specialties, experience with ICD-9-CM classification, and recent patient care responsibility. The physician's involvement was not to clarify coding, per se, but to identify principal diagnoses and provide clarification of vague or complex language used by the attesting physician. The physicians were further involved in reviewing instances of poor quality care and unnecessary admissions, which will be addressed in separate reports.

Five percent of the sample underwent blind re-coding by a different ART to verify the accuracy of the process. This

significantly more DRG assignment errors than large and mid-size
hospitals (Wald Chi-square 39.4, 2 degrees of freedom, p<.0001).
Hospitals in the same bed size stratum had similar rates of
errors regardless of location (urban or rural), teaching status
or proprietary versus not-for-profit status. Weighting by the
proportion of discharges in each stratum and projecting these
error rates to the 8.3 million cases paid under PPS in FY 85
suggests 1,500,000 discharges were paid incorrectly.

SIGNIFICANTLY MORE ERRORS IN HOSPITAL'S FAVOR

In 61.7 percent of the discharges with DRG errors, the DRG
assigned on recoding has a lower weight than originally paid.
Thus the hospital's coding gave it a larger payment than it
should have received for the discharge. This differs
significantly from the 50 percent rate expected if errors were
made at random (Wald Chi-square 68.7, 2 d.f., p<.0001). Previous
studies of coding accuracy, performed prior to PPS but using DRG
systems of classification, have reported higher rates of error,
but found the direction of coding error to be random or to
underpay the hospital at nonsignificant rates. Bed size had no
effect on the direction of errors.

DRG CHANGES BY HOSPITAL SIZE

BED SIZE	<100	100-299	300+	TOTAL	PERCENT OF SAMPLE DISCHARGE	PERCENT OF SAMPLE HOSPITAL
NUMBER OF DISCHARGES:						
Total	538	411	395	1374	18.6	20.8
Urban	102	323	363	788	17.6	20.4
Rural	436	118	32	586	20.9	21.3
Teaching	12	94	212	318	17.2	19.6
Nonteaching	526	347	183	1056	19.2	20.9
Profit	50	79	11	140	19.7	21.3
Nonprofit	488	362	384	1234	18.5	20.8

DRG CHANGES TO LOWER WEIGHT BY HOSPITAL SIZE

BED SIZE	<100	100-299	300+	TOTAL	PERCENT OF SAMPLE DISCHARGE	HOSPITAL
NUMBER OF DISCHARGES						
Total	345	263	228	836	11.1	12.9
Urban	55	195	207	457	10.3	11.5
Rural	290	68	21	379	13.3	13.7
Teaching	8	56	120	184	9.9	12.4
Nonteaching	337	207	108	652	11.7	12.9
Profit	34	44	7	85	11.7	13.6
Nonprofit	311	219	221	751	11.0	12.9

DRG CHANGES TO HIGHER WEIGHT BY HOSPITAL SIZE

BED SIZE	<100	100-299	300+	TOTAL	PERCENT OF SAMPLE DISCHARGE	HOSPITAL
NUMBER OF DISCHARGES						
Total	193	178	167	538	7.4	7.9
Urban	47	128	156	331	7.4	8.9
Rural	146	50	11	207	7.6	7.6
Teaching	4	38	92	134	7.3	7.3
Nonteaching	189	140	75	404	7.5	7.9
Profit	16	35	4	55	8.0	7.7
Nonprofit	177	143	163	483	7.4	7.9

10

PROJECTED COST SAVINGS OF $308 MILLION

The effect of these errors was to raise the case-mix index in the hospitals under review from 1.0720 (as corrected) to 1.0878 (as paid). Consequently, the sample hospitals were overpaid by $294,000 (out of $22 million total payments for the sample). This represents a potential cost savings of $308 million (1.3 percent of FY 85's $27 billion in PPS payments) if projected to the population of PPS discharges. These figures do not represent costs associated with unnecessary admissions, which will be discussed in another OIG report.

NET FINANCIAL EFFECT OF CODING CHANGES BY HOSPITAL SIZE

BED SIZE	>100	100-299	300+	TOTAL
$				
Sample	$ 187,000	$ 105,000	$ 52,500	$ 294,000
Population	71,200,00	137,000,000	80,200,000	308,000,000
Average/discharge	60	44	22	
Average/error	253	256	133	

CAUSE OF DRG ERRORS BY HOSPITAL SIZE

BED SIZE	<100	100-299	300+	TOTAL
PHYSICIAN ERRORS				
NARRATIVE CHANGE:				
Principal diagnosis change	225	168	145	538
Other narrative change	43	30	50	123
HOSPITAL ERRORS				
CODING CHANGE:				
Principal diagnosis	48	43	30	121
Secondary diagnosis	4	6	4	14
Procedure	8	21	17	46
RESEQUENCING:				
Incorrect sequence	173	108	90	371
Attestation renumbered	4	11	8	23
OTHER:				
No attestation code	45	34	18	97
Insufficient documentation	147	103	65	315
No principal diagnosis	39	24	19	82
Other	68	66	90	224

A discharge may have multiple reasons

PHYSICIAN'S CHOICE OF DIAGNOSIS OR PROCEDURE CAUSED 48 PERCENT OF ERRORS

The physician's choice of diagnosis or procedure, especially the principal diagnosis, was the single most frequent reason for incorrect payment seen in this study. In 39 percent of the errors, the principal diagnosis designated by the physician was not supported by the medical record. The attending physician described a patient who did not have the disease he later attested to, or who had that disease but was admitted for treatment of another condition. For example, a patient might be

discharged with a principal diagnosis of acute pancreatitis, but a subsequent review may not support that diagnosis. If the record did not describe the patient's complaining of pain, had laboratory results showing normal amylase and had no orders for pain medications, the reviewing physician would question whether the reason for admission was actually some other disease in a patient with, perhaps, a past history of pancreatitis. The reviewing physician would also be alerted to the possibility that the attending physician was not adept in the diagnosis and treatment of pancreatitis, and would have subjected the chart to a review of the quality of care delivered as well as to a review of the principal diagnosis.

Some errors can be particularly advantageous for the hospital, such as a chronic condition shown as the principal diagnosis, when in reality a patient with that chronic condition was admitted for treatment of an acute condition. Chronic conditions such as cirrhosis, emphysema and arteriosclerosis are weighted relatively highly (1.1841, 1.0304 and 0.9392) when compared to those conditions secondary to a principal diagnosis of, for example, acute gastroenteritis (0.6121 and 0.5593, depending on whether the chronic disorder is classified as a complicating condition). This represents a difference in payment of $1,000-$1,300.

The UHDDS definition of principal diagnosis is "the condition established after study to be chiefly responsible for occasioning the admission of the patient to the hospital for care." Errors result when physicians are not cognizant of this rule, or are inconsistent in applying it. In such instances, it is the responsibility of the medical record professional, after an analysis of the entire chart, to inform the physician that the principal diagnosis he or she selected does not conform to UHDDS guidelines. Thus, although the original error is the physician's, it is compounded by a coder who is poorly trained and unaware of the coder's responsibility, aware of it but afraid or unwilling to confront the physician, or under such pressure to complete coding so that a bill may be issued that the decision is made not to follow proper procedure. In some hospitals, coding is done from the face sheet only, not from the entire chart. This can ensure that errors in the identification of principal diagnoses will result in erroneous claims for payment, since no independent review is given to the underlying medical record.

The addition or deletion of a secondary diagnosis can affect a hospital's reimbursement if the secondary diagnosis is a complicating condition. These conditions were determined during the development of DRGs to have a strong likelihood of increasing the patient's length of stay by at least one day. Thus, an increase in payment acknowledges an increase in hospital costs.

(In a few DRGs, a particular secondary code can lead to lower reimbursement.) Similarly, the addition or deletion of an operative procedure can cause the DRG to change. As a rule, surgical DRGs are weighted more highly than medical DRGs to reflect the additional resources involved in performing surgery.

INCORRECT SEQUENCING OF CODES CAUSED 27 PERCENT OF ERRORS

More than a quarter of the errors were due to showing the disease entity that was actually responsible for the admission in the list of narrative diagnoses, but not as the principal diagnosis. In other instances, the codes, which correctly described correct narratives, were renumbered to create an incorrect claim.

To illustrate these errors of both coding and sequencing, assume a patient with chronic lung disease entered the hospital for treatment of acute bronchitis. This would be correctly shown as:

Acute bronchitis   466.0
COPD   496

and would group to DRG 96, with a weight of 0.7996.

Incorrect sequencing by the physician would show the acute condition which actually necessitated the admission as a secondary diagnosis:

COPD   496
Acute bronchitis   466.0

which would group to DRG 88, with a weight of 1.0412.

Leaving the narrative diagnoses in correct order, but submitting codes in the wrong order would also result in incorrect payment of DRG 88:

Acute bronchitis   ② 466.0
COPD   ① 496

Coding acute bronchitis as chronic would also have the effect of causing this example to group to DRG 88:

Acute bronchitis   491.2
COPD   496

IN 12 PERCENT OF ERRORS, AN INCORRECT CODE WAS SELECTED

The narrative diagnoses and procedures written or dictated by the physician are coded using ICD-9-CM, the three-volume work described earlier. The coder should use Volume 2, the diseases

14

index, as the first step in assigning a code, and Volume 1, the tabular list, as the second step. Frequently, however, a coder feels so familiar with the coding system that he or she assigns codes from memory or by turning to what is assumed to be the correct section of Volume 1, which contains over 1,000 pages of codes in numerical order. This can lead the coder to assign an incorrect code, either through failing to properly identify the disease or the stage of the disease referred to by the physician, or by assuming that the familiar code is the most descriptive in a given situation. Sometimes a physician may use language which is unfamiliar to the coder, who assumes an incorrect code describes the disease. Because of the intricacies of medical nomenclature and the nature of handwritten notes, errors such as reading "emphysema" for "empyema" can occur.

Some of these errors were made because coding rules or conventions were not followed. These are technical procedures which are shown in ICD-9-CM as "excludes" notes or as specific instructions, such as "do not code when 'x' is present," or rules of coding such as not coding symptoms when the disease responsible for them has been diagnosed, coding of diseases that have or have not been "ruled out," etc. Although some issues of classification will emerge from time to time, for the most part these coding conventions are well known.

Coding rules exist as well to govern when a cancelled procedure may be coded as performed. For example, if the abdomen was opened to perform gallbladder surgery, but the patient's heartbeat became so irregular that the wound was closed without the procedure's being done, the code for exploratory laparotomy is appropriately used. If, however, the irregularity caused the surgery to be cancelled while the patient was still in the holding area, no procedure may be coded.

OTHER ERRORS RESPONSIBLE FOR 13 PERCENT OF CHANGES

Among other causes of errors was physicians' use of vague or non-specific diagnostic language. This may also represent an inadequate diagnostic workup, where not enough is determined about the patient's condition to allow a more precise diagnosis.

A few of the erroneous discharges represent non-documentation of the patient's condition and treatment while in the hospital. That is, if the physician had adequately described the patient's status and treatment, the reviewing physician might have agreed with the principal diagnosis.

Transcription and transposition errors, including key entry errors, also caused errors in DRG assignment, as did coding records with multiple attestations, coding from incomplete

records (e.g., those lacking definitive laboratory results, received after the patient was discharged), and showing an incorrect patient age or discharge destination.

CONCLUSIONS

Under PPS, assignment of diagnoses and codes assumes new importance. Previously, payment had very little to do with any factor but the patient's length of stay and/or the costs associated with it. Hospitals may have responded to these new reimbursement factors by devoting greater attention to the accuracy of coding, which would explain the lower overall rate of error found in this study compared to earlier studies.

The direction of the error, however, may confirm that some coding choices are being made with the intent to increase reimbursement. The opportunities to do this are many, as discussed above. It was not within the scope of review of this study to determine a hospital's "motive" in coding a higher-weighted DRG than supported by the medical record. However, the results of this study appear to confirm that hospitals are coding records in a way that results in a higher payment than is justified by the medical record. This up-coding resulted in a 1.3 percent increase in the CMI in this study.

The finding that small hospitals code in error 23.6 percent of their discharges may indicate the relative lack of resources at these hospitals. These resources would include experienced, well-trained coders, consulting physicians who could provide more specific diagnoses and sophisticated software which would allow monitoring of coding accuracy and adequacy of record documentation. This suggests that further review and educational efforts should be directed towards these hospitals.

The pressures on hospitals under PPS should provide an incentive to provide cost-effective care. At the same time, remaining competitive in the marketplace requires a reputation for providing high quality care. Thus a balance should be maintained in which quality of care would not be sacrificed for decreased costs. If this is true, hospitals would look to other means for increasing operating revenues. Among the measures hospitals could take would be more streamlined scheduling of admissions and surgery; cost-benefit analyses of new purchases, particularly of "state-of-the-art" equipment; and tracking of the costs incurred by individual patients, compared to the anticipated DRG of that patient. Sophisticated software packages have been developed which allow precisely this level of integration of financial and clinical data. The results of this study suggest that in addition to legitimate optimization of codes and thus DRGs, "DRG creep" may be a by-product of this new attention to coding and costs.

16

APPENDIX

Statistical Analysis

Because of the two-stage sample design, this report evaluated its data by hospitals rather than by discharges. It calculated proportions of events as the number of events over the total number of discharges reviewed within each bed size group. Estimates for the total sample (weighted proportions) were weighted by the number of hospitals in each group. Projections to the universe of discharges derived from the inverse of the sampling fractions with the estimates and variances calculated accordingly.

Post-stratification analysis followed HCFA practices for classifying hospitals by their demographic characteristics. Urban versus rural status depended on whether the hospital's location fell within the boundaries of a standard metropolitan area as defined by the Census Bureau. The HCFA considered a hospital to have teaching status if it has an accredited residency program. These two characteristics warranted logit analysis because they affected the hospital's payment under PPS.

This survey employed a generalized logit analysis using weighted least squares for categorical data to test the effects of these variables on the various events measured in this study. This procedure reweighted data to properly represent the effect of independent variables not controlled by the appropriate degrees of freedom and provided a test of the significance of the model parameters and the goodness of fit of the models.