Department of Health and Human Services

OFFICE OF
INSPECTOR GENERAL

NATIONAL DRG VALIDATION STUDY
UPDATE: SUMMARY REPORT

AUGUST 1992

OFFICE OF INSPECTOR GENERAL

The mission of the Office of Inspector General (OIG), as mandated by Public Law 95-452, as amended, is to protect the integrity of the U.S. Department of Health and Human Services (HHS) programs as well as the health and welfare of beneficiaries served by those programs. This statutory mission is carried out through a nationwide network of audits, investigations, and inspections conducted by three OIG operating components: Office of Audit Services (OAS), Office of Investigations (OI), and Office of Evaluations and Inspections (OEI). The OIG also informs the Secretary of HHS program and management problems and recommends actions to correct them.

OFFICE OF AUDIT SERVICES

The OAS provides all auditing services for HHS, either with its own resources or by overseeing work done by others. Audits examine the performance of HHS programs and/or its grantees and contractors in carrying out their respective responsibilities. Audits provide independent assessments of HHS programs and operations in order to reduce waste, abuse, and mismanagement; and to promote economy and efficiency throughout the Department.

OFFICE OF INVESTIGATIONS

The OI conducts criminal, civil, and administrative investigations of alleged wrongdoing in HHS programs or to HHS beneficiaries, and of unjust enrichment by providers. The investigative efforts of OI lead to criminal convictions, administrative sanctions, or civil money penalties. The OI also oversees State Medicaid fraud control units which investigate and prosecute fraud and patient abuse in the Medicaid program.

OFFICE OF EVALUATION AND INSPECTIONS

The OEI conducts short-term management and program evaluations (called inspections) that focus on issues of concern to the Department, Congress, and public. The findings and recommendations contained in these inspection reports generate rapid, accurate, and up-to-date information on the efficiency, vulnerability, and effectiveness of departmental programs.

The information for this report was collected under contracts with Baxter-Health Data Institute and BOTEC Analysis Corporation. A list of project participants appears in Appendix A.

Department of Health and Human Services

OFFICE OF
INSPECTOR GENERAL

NATIONAL DRG VALIDATION STUDY
UPDATE: SUMMARY REPORT

EXECUTIVE SUMMARY

PURPOSE

This inspection reabstracted the International Classification of Diseases, 9th Edition, Clinical Modification (ICD-9-CM) codes from a sample of Medicare discharges on a blinded basis. It compared the resulting diagnosis-related group (DRG) to the hospital's DRG to determine over-reimbursement or under-reimbursement. The sample was nationally representative and covered all of calendar year 1988, the most recent data available.

This study updates a previous inspection from the Office of Inspector General (OIG). It found that 20.8 percent of 1985 Part A bills contained coding errors that changed the DRG and that 61.7 percent of these errors over-reimbursed the hospitals. This improper DRG "creep" increased total prospective-payment system (PPS) disbursements by 1.9 percent or $308 million.

FINDINGS

DRG coding error reduced: 14.7 percent of 1988 discharges had DRG coding errors. This proportion of coding errors was statistically significantly lower than the 20.8 percent reported for 1985.

DRG creep eliminated: 50.7 percent of DRG errors over-reimbursed the hospital. This proportion differed significantly from the 61.7 percent over-reimbursed reported for 1985.

DRG coding errors, overall, no longer over-reimburse hospitals: Taken together, 1988 DRG errors had the net financial effect of under-reimbursing all hospitals a non-significant $69.8 million, or 0.1 percent of the $52 billion in 1988 PPS disbursements. Projected nationally, over-reimbursements totaled $2,657.8 million and under-reimbursements totaled $2,588.0 million.

Mis-specification errors under-reimbursed the hospitals: Of the 361 DRG errors in the sample, 63.2 percent occurred because the attending physician mis-specified the narrative diagnoses. Of these 227 mis-specification errors, only 43.2 percent over-reimbursed the hospitals. The statutorily-required attestation probably reminded physicians of their obligation to select accurate narrative diagnoses.

Resequencing errors over-reimbursed the hospitals: Of the 361 DRG errors in the sample, another 26.6 percent occurred because the hospital substituted a secondary diagnosis for the (correct, narrative) principal diagnosis. Of these 96 resequencing errors, 66.7 percent over-reimbursed the hospitals. The sentinel effect of the Peer Review Organizations (PRO) surveillance apparently did not fully prevent over-reimbursement due to resequencing.

RECOMMENDATIONS

▶ **The Peer Review Organizations should continue their surveillance of hospital coding for DRG reimbursement accuracy.**

The attestation requirement appears to have deterred over-reimbursement due to mis-specification by attending physicians, causing their hospitals to be under-reimbursed. However, the sentinel effect did not fully prevent over-reimbursement due to resequencing by hospitals. Although these two trends approximately offset each other, this equilibrium may not continue in the future.

The Agency for Health Care Policy Research plans to reabstract the ICD-9-CM codes from a large sample of Medicare charts to determine their diagnostic accuracy for health services research purposes. Using 1991 data, grouping these ICD-9-CM codes to DRGs, and selecting reasons for any differences would provide a third time point with which to track trends in reimbursement accuracy. The OIG supports this effort.

AGENCY COMMENTS

In its May 29, 1992 comments to the draft of this inspection, the Health Care Financing Agency (HCFA) concurred with the OIG recommendation. The HCFA noted that the improvement in DRG coding accuracy may be attributable to increased hospital experience with PPS, PRO review, and HCFA educational efforts.

The HCFA also made a number of technical comments. Based on these comments, the OIG made several changes to this report. The full text of the HCFA comments appear as an appendix.

TABLE OF CONTENTS

INTRODUCTION

Since the inception of the Medicare program, health care expenditures have grown faster than the rest of the economy, and the Medicare program has grown even faster than general health care expenditures. The need for innovative reimbursement policies and their evaluation has become critical to medical communities, beneficiaries, and payers.[1]

ORIGIN OF THE PRESENT STUDY

Since 1965, Medicare has provided hospital and medical insurance to the nation's elderly and disabled.[2] However, rapidly escalating health care costs coupled with a rise in the proportion of the population eligible for Medicare have burdened its hospital insurance trust fund (Part A). Health care costs accounted for 7.4 percent of the gross national product in 1970, 9.1 percent in 1980, and 10.9 percent in 1986. In 1970, Medicare paid 8 percent of all health care costs, but by 1986 Medicare's proportion increased to 17 percent.

In 1983 Congress changed Medicare inpatient reimbursement from a retrospective, reasonable-cost basis to a prospective-payment system (PPS).[3] Under the PPS, hospital payments depended upon the patient's diseases and procedures as defined by the International Classification of Diseases, 9th Edition, Clinical Modification (ICD-9-CM).[4] The PPS "grouped" various combinations of the approximately 10,000 ICD-9-CM codes into 476 diagnosis-related groups (DRGs).[5] By reimbursing the average cost of care for each DRG, the PPS provided incentives for efficient delivery of health care. The Veterans Administration Medical System Resource Allocation Method emulated this change.

A hospital files a claim for Medicare payment at the time of patient discharge. The attending physician writes the narrative diagnoses and procedures on the face sheet; and attests to the principal diagnosis, secondary diagnoses, and any procedures. The hospital's medical records department then assigns numeric, ICD-9-CM codes to the narrative diagnoses and procedures, using the rules of the Uniform Hospital Discharge Data Set (UHDDS), "a minimum data set used by HHS programs to collect data on individual hospital discharges on a continuing basis."[6] The hospital's billing department enters up to five diagnoses and three procedure codes on the Medicare claim form. A Fiscal Intermediary (FI) for each State receives the hospital bills. It runs GROUPER software to select the correct DRG by assessing diagnosis and procedure codes. It then runs PRICER software that adjusts for geographic location, teaching costs, and other factors to calculate the payment due to the hospital.

Since 1976 the Office of the Inspector General (OIG) has had responsibility for protecting the integrity of the programs and program beneficiaries of the U.S. Department of Health and Human Services (HHS).[7] Following this statutory mandate, the OIG commissioned the 1985 National DRG Validation Study (1985 Study), which examined how the processes that assigned diagnosis and procedure

codes affected the accuracy of PPS reimbursement for Medicare discharges.[8] The 1985 Study analyzed a representative, national sample of medical records to obtain information on the accuracy of the diagnosis and procedure coding, the impact of coding errors on DRG assignment, and the potential net financial impact of coding errors. It also assessed the appropriateness and quality of the services provided to Medicare patients.

The 1985 Study found an overall error proportion of 20.8 percent in assigning DRGs.[9] In 61.7 percent of the errors, the hospitals over-reimbursed themselves. Physician mis-specification of narrative diagnoses and hospital billing department resequencing of diagnoses caused most errors. These errors caused $308 million in overpayments to hospitals, 1.9 percent of PPS disbursements.

The Health Care Financing Administration (HCFA), as the payer of the Medicare reimbursements, made the following points in commenting on the 1985 Study.

▶ The PPS started in 1983. The 1985 Study data covered FY 1985. Physicians, hospital coders, and Peer Review Organizations (PROs) needed a longer learning period to adjust to the new payment and quality monitoring system.

▶ The 1985 Study appeared in 1987, so that its coclusions derived in part from information and coding conventions that had undergone updating by that time.

▶ The HCFA and hospital associations recently had begun to conduct training seminars for hospital personnel. The HCFA therefore expected that coding errors would decrease over time.

To examine these trends, the OIG initiated a follow-up project, implemented through contracts with the Health Data Institute (HDI) of Lexington, Massachusetts; American Medical Records Association (AMRA) of Chicago, Illinois; and BOTEC Analysis Corporation (BOTEC) of Cambridge, Massachusetts using 1988 data. The purpose of this inspection, the DRG Validation Study Update (1988 Study), was to determine how the PPS's coding accuracy had changed over time. The OIG designed this inspection to meet the following objectives.

▶ Replicate the 1985 work on coding accuracy, using 1988 data.

▶ Compare the findings from 1985 and 1988.

▶ Identify patterns of coding errors that PPS changes could modify.

As in the previous study, this inspection assessed how coding errors affected hospital reimbursement and identified the coding problems most subject to error. It evaluated hospital and beneficiary characteristics for their relative contribution to DRG errors.

METHODOLOGY

The OIG randomly selected 2,680 discharges from 1,744 acute care hospitals. The study population consisted of the 10.8 million Medicare funded discharges for calendar year (CY) 1988 from the 6,715 acute care, short-stay hospitals in the United States.[10] The design excluded discharges from specialty institutions such as children's hospitals, tuberculosis units, and psychiatric facilities. It also excluded discharges in Maryland and New Jersey, which the PPS still exempted in 1988. It excluded bills for pediatric, obstetric, and psychiatric DRGs (principally drug and alcohol rehabilitation performed by a general hospital). Unlike its FY 1985 predecessor, it included hospitals established since the advent of the PPS in 1983.

The OIG requested that the hospitals' medical records departments send complete copies of the selected medical records to the OIG's contractor, the HDI. With follow-up, the OIG ultimately obtained 2,451 medical records, 91.4 percent of those selected. The OIG compelled the cooperation of four hospitals by administrative subpoenas.

The AMRA reabstracted the charts by selecting ICD-9-CM codes supported by the record, selected the principal diagnosis, and grouped to select the DRG. To ensure that the original ICD-9-CM codes and DRG codes did not affect the AMRA's reabstraction, the AMRA coders conducted this reabstraction without knowledge of the original ICD-9-CM codes and DRG codes. The coders had instructions not to treat marginal problems or honest differences in judgment about appropriate coding as DRG errors. This standard should have produced a conservative estimate of the proportion of discharges having DRG errors.

The AMRA also identified the reasons why a hospital's bill differed from the correct codes. Where multiple reasons applied, the coders had instructions to select the first chronological reason (i.e., mis-specification → miscoding → resequencing → other). Finally, AMRA entered its reabstracted codes into a personal computer database provided by the OIG. A series of reliability checks verified the reproducibility and accuracy of the AMRA coding.

SAMPLE REPRESENTATIVENESS

The sample accurately represented the characteristics of the underlying population. Distributed by hospital demography, it did not differ from the population in bed size, teaching status, location, or control. [Figure 1].

The sample also accurately represented the underlying population by patient age and sex. However, the OIG made remedial efforts to classify unknowns by race. This match to other government files reduced the proportion of sample unknowns in comparison to the underlying population.[11] The volume of cases precluded a similar reclassification of unknowns for the entire population. Without the OIG's reclassification, the sample would also have conformed to racial distribution of the underlying population. [Figure 2].
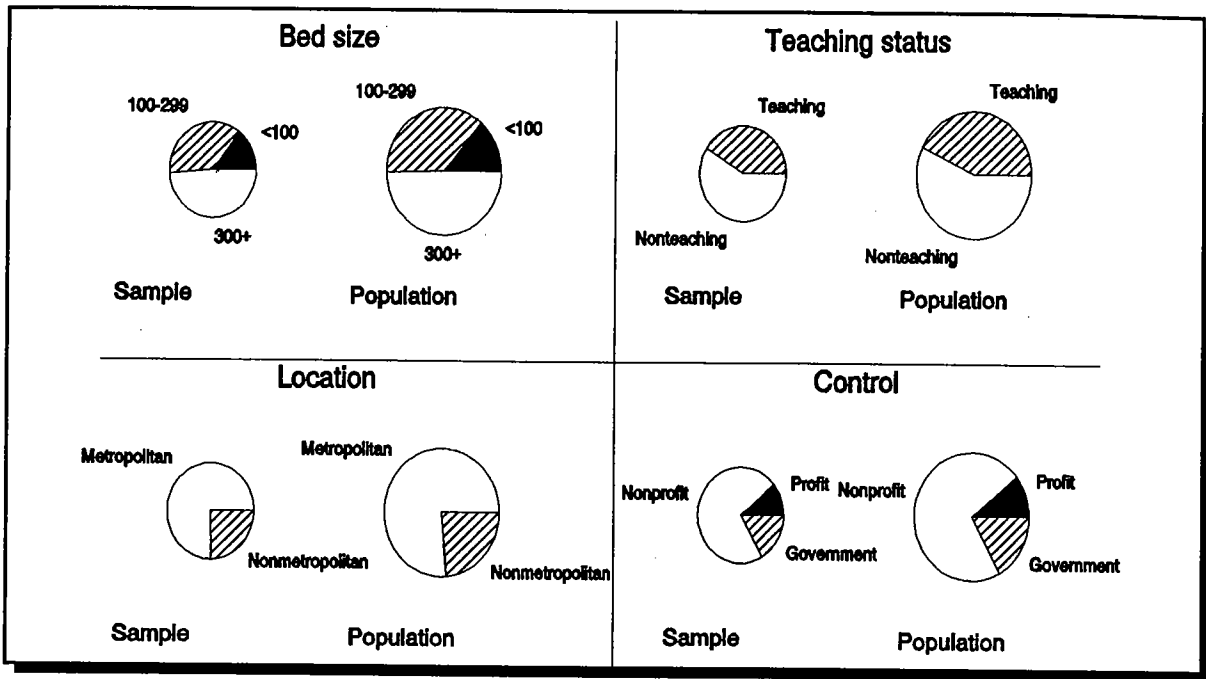

3

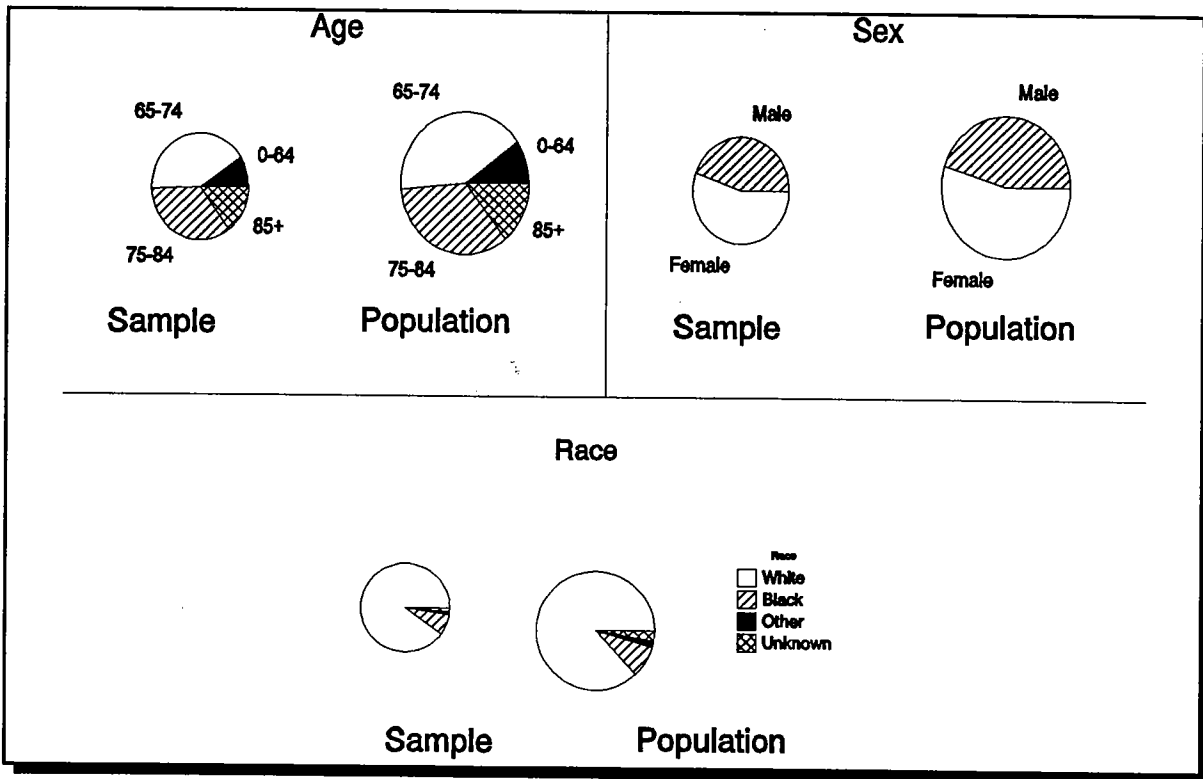**Figure 1:** Sample representativeness by hospital demography, 1988

Figure 2: Sample representativeness by patient demography, 1988

Because of its random design, the sample should also accurately represent the distribution of DRGs.

FINDINGS

DRG CODING ERROR REDUCED

Of the sample discharges, 14.7 percent proved to have coding errors that changed their DRGs. This proportion was statistically significantly lower than the 20.8 percent for 1985 (18.4 percent after reweighting to make the 1985 sample comparable). The standard error of 0.7 indicated this point estimate to be quite precise, a secondary effect of the sample size. [Figure 3].

No hospital demographic characteristics had a statistically significant effect upon the proportion of DRG coding errors. Smaller hospitals had higher proportions of miscodes, but this apparent difference did not attain statistical significance.

Figure 3: Proportion of discharges with DRG coding errors, 1985 and 1988

Similarly, no patient demographic characteristics had a statistically significant effect upon the proportion of DRG coding errors. Discharges of younger patients had higher error rates, However, this difference also failed to attain statistical significance.
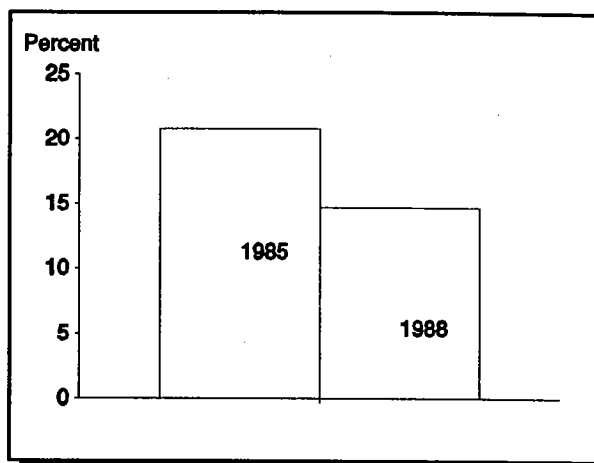
DRG CREEP ELIMINATED

The 1988 DRG errors divided evenly between errors that had over-reimbursed the hospital (50.7 percent) and under-reimbursed the hospital (49.3 percent). This non-directionality differed significantly from the 61.7 percent of 1985 coding errors that over-reimbursed the hospital. [Figure 4].

The equal division between coding errors that over-reimbursed and under-reimbursed the hospital occurred across all hospital characteristics. For-profit hospitals over-reimbursed themselves more than did other types of hospitals, but these differences did not attain
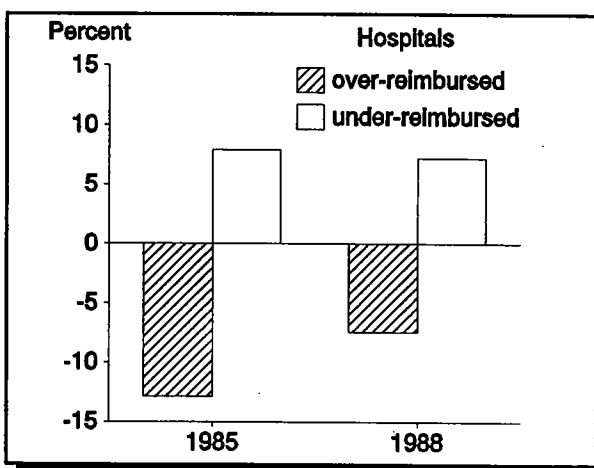
Figure 4: Direction of DRG coding errors, 1985 and 1988

statistical significance. The direction of coding error also exhibited no statistically significant trend by patient characteristics.

The OIG calculated the financial effect of coding errors using the case-mix index (CMI), which quantifies the complexity and resource intensity of Medicare reimbursed discharges. Hospitals with higher mean relative weights serve Medicare patients with more complex conditions and therefore consume more resources. This inspection calculated CMI as (1) submitted to the FIs for reimbursement, (2) reabstracted by the AMRA on a blinded basis, and (3) the difference or mean weight change due to coding inaccuracy. After AMRA recoding, the overall CMI increased 0.0023, a statistically nonsignificant difference. Extrapolating this financial change to all 10.8 million Medicare discharges, DRG coding errors under-reimbursed hospitals by $69.8 million. This difference amounted to only 0.1 percent of the $52 billion in 1988 PPS expenditures.[12] The average hospital billed for slightly less reimbursement for itself than it should have received. [Figure 5].
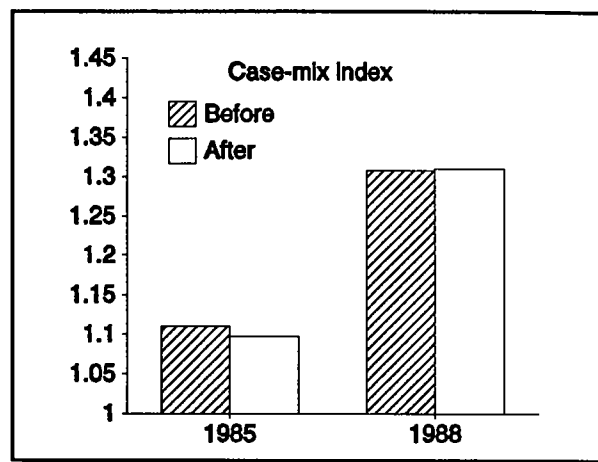

Figure 5: Case-mix index before and after DRG reabstraction, 1985 and 1988

MIS-SPECIFICATION UNDER-REIMBURSES, WHILE RESEQUENCING OVER-REIMBURSES

Each stage of the reimbursement process introduced coding errors for a different reason. The OIG identified three types of errors: mis-specification, miscoding, and resequencing. Mis-specification describes physician errors in attesting to the narrative diagnoses. Miscoding refers to the medical records department selecting the wrong ICD-9-CM code for a correct narrative diagnosis. Resequencing involves the billing department improperly substituting a secondary diagnosis for the correct principal diagnosis. [Figure 6].
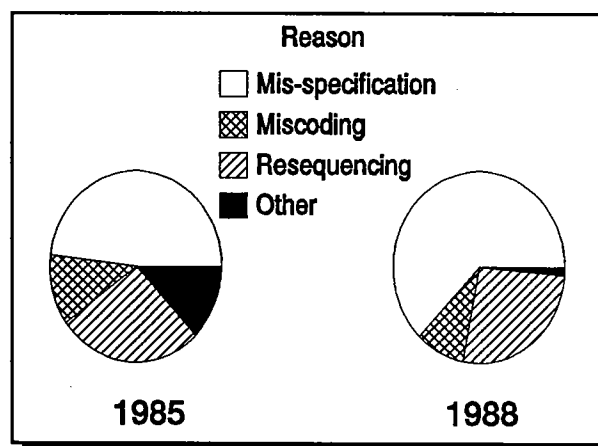

Figure 6: Reasons for DRG coding errors, 1985 and 1988

Narrative changes accounted for 63.2 percent of the 361 DRG errors. Examples of mis-specification include (1) the physician selecting the wrong principal diagnosis and (2) including or excluding a complication, comorbidity, or operating room procedure.

Resequencing cause 26.6 percent of DRG errors. In all of these cases, the AMRA review disagreed with the providers' listing sequence. Usually, a secondary diagnosis was listed as the principal diagnosis. In other cases, the diagnosis needed greater specificity. ICD-9-CM ruling changes accounted for the remaining sequence changes.

Miscoding comprised 9.1 percent of DRG errors. The incorrect numeric code was usually for the principal diagnosis, rather than for complications, comorbidities, or operating room procedures. A few changes represented inaccurate discharge destinations or apparent typographical errors.

Mis-specification significantly under-reimbursed hospitals; while resequencing significantly over-reimbursed them. Thus, the attending physicians selected their patients' diagnoses and procedures with caution, earning their hospitals less reimbursement than they should have received.

The billing departments behaved more aggressively, obtaining greater reimbursement than their employers should have received. Possibly the hospitals had greater influence over employees, or the employees identified closely with the economic interests of the

Figure 7: Reasons for DRG changes by direction of reimbursement change, 1988

hospital. The sentinel effect of the Peer Review Organizations surveillance did not fully prevent over-reimbursement due to resequencing. [Figure 7].

These two trends, under-reimbursement due to physician mis-specification and over-reimbursement due to resequencing, approximately offset each other; so that overall hospitals received correct reimbursement. Whether this equilibrium will continue is unknown.

CERTAIN DRGS ARE MOST SUSCEPTIBLE TO ERROR

In this random sample of discharges, vascular disorders such as heart failure, angina, stroke, and arrhythmias cause Medicare hospitalizations most frequently. Overall, 28 of the 476 DRGs accounted for half of the total bills, and 10 DRGs accounted for 28.1 percent of all bills. Most coding errors fell into these DRGs, but not in direct ratio to the DRGs' volume. [Table 1].

This report identified DRGs with high proportions of coding errors by dividing the number of errors by the frequency of bills for specific DRGs. Many of these DRGs covered vague or nonspecific diagnoses such as atherosclerosis (DRG 132), other circulatory system procedures (DRG 120), respiratory signs and symptoms (DRG 99),
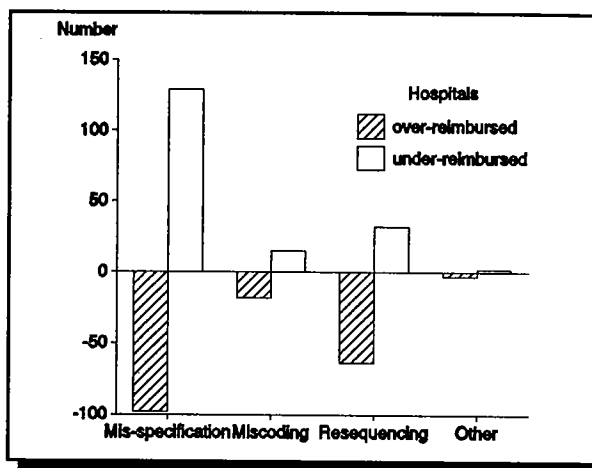
#	DRG description	n [%]
127	Heart failure and shock	133 [5.4]
140	Angina pectoris	89 [3.6]
14	Specific cerebrovascular disorders except TIA	75 [3.1]
89	Simple pneumonia	74 [3.0]
182	Esophagitis, gastrointestinal, and miscellaneous digestive disorders	66 [2.7]
96	Bronchitis and asthma with complications	63 [2.6]
209	Major joint procedure	55 [2.2]
15	Transient ischemic attacks	47 [1.9]
138	Cardiac arrhythmia and conduction disorders	45 [1.8]
296	Nutritional and miscellaneous metabolic disorders with complications	44 [1.8]
	Other	1,760 [71.8]
	Total	2,451 [100.0]

Table 1: DRGs billed most frequently, 1988

and other nervous system disorders (DRG 34). This indeterminateness suggests inherent ambiguities in medical taxonomy. For example, DRG 99 includes apnea, dyspnea, hemoptysis, hypercapnia, pleurodynia, stridor, and ventilatory failure. [Table 2].

#	DRG description	n	Miscoded Number [%]
132	Atherosclerosis	4	3 [75.0]
413	Myeloproliferative	3	4 [66.7]
185	Dental except extractions	3	2 [66.7]
120	Other circulatory system procedures	2	3 [66.7]
99	Respiratory signs and symptoms	8	5 [62.5]
403	Lymphoma and nonacute leukemia	5	3 [60.0]
34	Other nervous system disorders	5	3 [60.0]
	Other	2,420	341 [14.1]
	Total	2,451	361 [14.7]

Table 2: DRGs with high proportions of coding errors, 1988

Over-reimbursement concentrated in selected DRGs and certain types of hospitals. This report identified DRGs with maximum savings potential by multiplying the error frequency by reimbursement change for each DRG. Although this inspection found no overall over-reimbursement to hospitals, certain DRGs still significantly over-reimbursed the hospitals. In particular, 13 DRGs each had over $20 million in projected overpayments. This group consisted primarily of DRGs with operating room procedures, probably because of their high relative weights. [Table 3].

#	DRG description	n	Over-reimbursement	
			$ per discharge	$ million total
104	Cardiac valve procedure with pump & cath	5	4915	81.1
468	Unrelated operating room procedures	21	1128	78.2
475	Respiratory system diagnosis with ventilator	19	791	49.6
110	Major reconstructive vascular procedures	21	682	47.3
191	Major pancreas, liver, & shunt procedures	4	3313	43.7
154	Stomach, esophageal, & duodenal procedures	11	1114	40.5
87	Pulmonary edema & respiratory failure	20	571	37.7
76	Respiratory system operating room procedures	10	764	25.2
82	Respiratory neoplasms	24	306	24.2
121	Myocardial infarction discharged alive	38	189	23.7
415	Operating room procedure for infection	8	858	22.6
148	Major large & small bowel procedures	33	202	22.0
217	Wound debridement & skin graft	2	333	22.0
	Other	2,235	-263	-587.6
	Total	2,451	-29	-69.8

Table 3: DRGs with maximum savings potential, 1988

Additionally, small and for-profit hospitals made more errors that over-reimbursed themselves than did other types of hospitals. This net over-reimbursement totaled $842.2 million.

RECOMMENDATIONS

▶ **The Peer Review Organizations should continue their surveillance of hospital coding for DRG reimbursement accuracy.**

The attestation requirement appears to have deterred over-reimbursement due to mis-specification by attending physicians, causing their hospitals to be under-reimbursed. However, the sentinel effect did not fully prevent over-reimbursement due to resequencing by hospitals. Although these two trends approximately offset each other, this equilibrium may not continue in the future.

The Agency for Health Care Policy Research plans to reabstract the ICD-9-CM codes from a large sample of Medicare charts to determine their diagnostic accuracy for health services research purposes. Using 1991 data, grouping these ICD-9-CM codes to DRGs, and selecting reasons for any differences would provide a third time point with which to track trends in reimbursement accuracy. The OIG supports this effort.

AGENCY COMMENTS

In its May 29, 1992 comments to the draft of this inspection, the HCFA concurred with the OIG recommendation. The HCFA noted that the improvement in DRG coding accuracy may be attributable to increased hospital experience with PPS, PRO review, and HCFA educational efforts.

The HCFA also made a number of technical comments. Based on these comments, the OIG made several changes to this inspection. The full text of the HCFA comments appear as an appendix.

ENDNOTES

1. Schweiker R S. Report to Congress: hospital prospective payment for medicare. Washington, DC: U.S. Department of Health and Human Services, 1982.

2. Social Security Amendments of 1965. Pub. L. No. 89-97, 79 Stat 286 (1965).

3. Social Security Amendments of 1983. Pub. L. No. 98-21, 97 Stat 65 (1983).

4. Commission on Professional and Hospital Activities. International classification of diseases, 9th edition, clinical modifications. Ann Arbor, MI: Edward Brothers, 1980.

5. Fetter R B, Shin Y, Averill R F, & Freedman A T. The new ICD-9-CM diagnostic related groups classification scheme user's manual: final report. New Haven, CT: Yale School of Organization and Management, Health Systems Management Group, 1981.

6. Fed Reg, July 31, 1985; 50 (147): 31038.

7. Shriners' Hospital for Crippled Children -- Colorado Corporation -- Conveyance. Pub. L. No. 94-505; October 15, 1976: Title II, §§ 201-07.

8. Delaney A M & Hsia D C. 1985 national DRG validation study. Lexington, MA: Health Data Institute, 1987. Report prepared for the Office of Inspector General. U.S. Department of Health and Human Services; contract HHS-100-87-0015.

9. Hsia D C, Krushat W M, Fagan A B, Tebbutt J A, & Kusserow R P. Accuracy of diagnostic coding for medicare patients under the prospective-payment system. N Engl J Med, 1988; 318: 352-55.

10. Division of Information Analysis, Office of Statistics and Data Management, Bureau of Data Management and Strategy, Health Care Financing Administration, U.S. Department of Health & Human Services. 1989 HCFA Statistics. Baltimore, MD: Health Care Financing Administration, 1989: 16 & 34.

11. Ruther M, Reilly T W, Silverman H A, & Abbott D B. Medicare and Medicaid data book, 1990. Baltimore, MD: Office of Research and Demonstrations, Health Care Financing Administration, U.S. Department of Health & Human Services, 1991: 15. HCFA pub. no. 03314.

12. Division of Information Analysis, Office of Statistics and Data Management, Bureau of Data Management and Strategy, Health Care Financing Administration, U.S. Department of Health and Human Services. 1989 HCFA statistics. Baltimore, MD: Health Care Financing Administration, 1989: 25. HCFA publication no. 03294.

APPENDIX A

PROJECT PARTICIPANTS[1]

<u>OIG</u>
Cathaleen A. Ahern, B.A.
Evan J. Buckingham, B.A.
David C. Hsia, J.D., M.D., M.P.H.
Thomas F. Komaniecki, M.P.A.
W. Mark Krushat, M.P.H.
Linda M. Moscoe, B.A.
Brian P. Ritchie, B.A.
Barry L. Steeley[2]
John M. Traczyk, B.A.

<u>HCFA</u>
Timothy F. Greene, M.A., M.B.A.
Stephen F. Jencks, M.D.
Michael R. McMullan, M.B.A.
Harry L. Savitt, Ph.D.
Jeanette M. Smith, M.D., M.P.H.[3]
Malcolm A. Sneen, B.S.

<u>RAND Corporation</u>
Haya P. Rubin, M.D., Ph.D[4]

<u>Baxter-Health Data Institute</u>[5]
Patricia J. Baxter, R.N.
Patricia Cassidy-Tsnosas, R.N.
Annette M. Delaney, R.N., M.A.
Ellen B. Inghilleri, R.N.
Janet Mathews, A.R.T.
Laurie H. Moore, R.R.A.
Claire Shannon, A.R.T.
Michele A. Wiese, B.A.

<u>AMRA</u>
Margret K. Amatayakul, M.B.A., R.R.A.
Mary Converse, R.R.A.

---

1. Institutions appear in chronological order of involvement.
2. Now at Health Audit Services, Ellicott City, MD.
3. Now at the Journal of the American Medical Association, Chicago, IL.
4. Now at Johns Hopkins Medical Institutions.
5. Ceased operations February 16, 1990.

Nicholas J. Cotsonas, M.D.[6]
Linda Ertl, R.R.A.
Rita M. Finnegan, R.R.A.
Desla Mancilla, A.R.T.
Barbara Manny, R.R.A.
Sonia Martyniuk, R.R.A.
Toula Nicholas, A.R.T.
Charlotte Razor, R.R.A.
LouAnn Schraffenberger, R.R.A.
Lynn Smetko, R.R.A.
Dawn Smith, A.R.T.
Joan Zacharias, A.R.T.

BOTEC Analysis Corporation
Geraldine M. Berenholz, R.R.A.
Andrew H. Chalsma, B.A.
David P. Cavanagh, M.A., Ph.D.
Janet W. Knight, R.N., Ph.D.
Amy L. Lockwood, B.A.

---

6. Outside contractor.

APPENDIX B

HEALTH CARE FINANCING ADMINISTRATION COMMENTS

Memorandum

Date
. MAY 29 1992

From
William Toby, Jr.
Acting Administrator

Subject
OIG Draft Reports: "National DRG Validation Study Update: Summary
Report" (OEI-12-89-00190) and "National DRG Validation Study Update:

To
Technical Report" (OEI-12-89-00191)

Inspector General
Office of the Secretary

We have reviewed the above-referenced draft reports which are
updates to a previous OIG study which re-abstracted the International
Classification of Diseases, 9th Edition, Clinical Modification codes from a
sample of 1985 Medicare discharges. These reports update the previous study
by: replicating the 1985 work on coding accuracy using 1988 data, comparing
the findings from 1985 and 1988, and identifying patterns of coding errors that
the Prospective Payment System (PPS) changes could modify. The sample
was nationally representative and covered all of calendar year 1988, the most
recent year for which data were available.

OIG found a significant improvement in hospital coding of PPS cases
between 1985 and 1988. In the 1985 study, OIG found that 20.8 percent of
1985 Part A bills contained coding errors that changed the diagnostic related
group (DRG) and that 61.7 percent of these errors over-reimbursed the
hospitals. This improper DRG coding increased total PPS disbursements by
1.9 percent or $308 million. During the 1988 study, OIG found that
14.7 percent of the discharges had DRG coding errors, 51 percent of DRG
errors over-reimbursed the hospitals, and 49 percent under-reimbursed the
hospitals. Taken together, the 1988 DRG errors resulted in no net
overpayment of hospitals.

We believe these reports are a useful addition to the continuing
research and evaluation of case mix change among hospitals paid under PPS.
We agree with OIG's recommendation that Peer Review Organizations
continue their surveillance of hospital coding of DRG reimbursement
accuracy. Our detailed comments on these two reports are attached for your
consideration.

Page 2 - Inspector General

Thank you for the opportunity to review and comment on these draft reports. Please advise us if you agree with our position on the report's recommendation at your earliest convenience.

Attachment

Comments of the Health Care Financing Administration (HCFA)
on OIG's Draft Reports: "National DRG Validation Study Update:
Summary Report" OEI-12-89-00190 and "National DRG Validation
Study Update: Technical Report" OEI-12-89-00191

OIG Recommendation

The Peer Review Organizations (PROs) should continue their surveillance of hospital coding for DRG reimbursement accuracy.

HCFA Response

We agree. The PROs will continue their surveillance of hospital coding of diagnostic related group (DRG) reimbursement accuracy.

We are pleased to note the decrease in DRG coding errors identified by the 1988 study, as compared to the 1985 study. We believe this improvement may be attributed to a combination of factors including increased hospital experience under the prospective payment system (PPS), the impact of PRO review and the effect of HCFA educational efforts to enhance coding accuracy.

General Comments

Summary Report, OEI-12-89-00190

This report includes a brief but excellent description of the payment process, including the role of coding. Three types of errors are defined at each stage of the payment process. However, no statement is made as to whether or not the errors are mutually exclusive, and if the 361 records with errors contained only 1 of each type. We believe this information should be included in the final report.

Although the representativeness of the sample is thoroughly addressed, there is no discussion as to how adequately each DRG is represented. We would like to know how the distribution by DRG of the sample relates to the universe of DRGs, and if any correlation exists between high volume DRGs and errors in the sample. Also, we note that most hospitals contributed only one discharge to the sample. We question whether it is statistically valid for 1 record in 1 hospital to represent 3,300 other records in the total population of records.

It is not made clear in the report how payment projections are calculated. The amount used to estimate payment is not stated, although mention is made

of using a rate of $3,118 for metropolitan hospitals and $2,637 for nonmetropolitan hospitals (page 29, Technical Report). It is stated that these are "current dollars," but there is no discussion as to the actual figures used to project payment amounts. While the calculation for case-mix index (CMI) is clearly defined, as is the DRG difference before and after coding, it is not clear how the payment amount is determined. The final report should include a more complete description of payment calculation.

Executive Summary, page iii - "DRG coding errors, overall, no longer over-reimburse hospitals." The figure for under-reimbursements is stated to be "$2,588.0 billion." The correct figure is $2,588 million (also Executive Summary, page iii, Technical Report).

Findings, pages 5, 9 - While it is unwise to ignore magnitudes and only pay attention to statistical significance, we do not believe effects that are not statistically significant should be highlighted. For example, the report notes that for-profit hospitals "over-reimburse themselves" but that "these results did not attain statistical significance" (also CMI discussion, pages 13, 14, Technical Report).

Findings, page 6 - The narrative incorrectly states that the overall CMI decreased after the study's recoding. For hospitals to have underpaid themselves, the CMI would have to increase after recoding, as is correctly stated on page 13 of the Technical Report.

Findings, page 6, Figure 6 - The 1988 pie chart does not match the numbers in the text immediately below the chart. The resequenced and miscoded pieces of the pie do not represent the 27 percent and 9 percent respective figures and may have been reversed in the pie chart.

Technical Report

Introduction, page 1, paragraph 3 - The original number of DRGs in 1983 was 470, not 476.

Introduction, page 1, paragraph 4 - The narrative implies that there is a fiscal intermediary for each State. The sentence should be rewritten to state: "A fiscal intermediary receives the hospital bills for each State."

Page 2, paragraph 3, under the heading "The HCFA . . . made the following points in commenting on the 1985 Study" - It was HCFA's intent to point out that the 1985 study was published in November 1987, and conclusions set forth in that report were based in part on information and coding conventions which had been updated by the time the report was released. The phrasing ". . . needed a longer learning period to adjust to the new payment and quality monitoring system" does not encompass the concept that changes have been made to the whole International Classification of Diseases, 9th Edition, Clinical Modification/Prospective Payment System/DRG system since its conception. Such changes and improvements continue to be made.

Page 4 - We do not believe the data support some of the assertions and language which the report uses. For example, the report describes the findings of the earlier OIG report as showing "intentional 'gaming' or manipulating [of the] coding process" by hospitals. The note cites unrelated sources while the earlier report used much more restrained language.

The report consistently characterizes coding disagreements as provider "errors." A recent Rand study used charts re-abstracted at SuperPRO to examine the 1987-88 increase in the Medicare CMI. Rand estimated that changes in coding common to SuperPRO and hospitals accounted for almost one-third of the 1987-88 CMI increase. Since SuperPRO coders have no incentive to upcode, this probably reflects changes in explicit and implicit coding rules between original coding (1987) and recoding (1988). While the CMI may increase, payments to hospitals will not. However, disagreements which may reflect the evolution of coding practice over time are not strictly provider "errors." Rand and OIG studied slightly different problems, and the Rand finding may be unique to 1988. However, since American Medical Records Association staff should be able to provide insight on this matter, we recommend that OIG explore and report on the question of whether the increase in the CMI is attributable to provider errors or the evolution of coding practices.

Page 4, paragraph 3 - "Each year, the relative weights change to reflect alterations in resource consumption, DRG title, coding and . . . ." The change in relative weights is not related to the title of the DRG.

Page 10 - The report does not find net overpayment due to coding error. Inclusion of a table of DRGs with "maximum savings potential [to Medicare]" due to upcoding, without a parallel table of DRGs with maximum potential payment increase, gives an impression of lack of balance. Financial impact (savings) estimates could be presented in the context of the argument that, while ". . . trends to over-pay and under-pay approximately offset each other, this equilibrium may not continue in the future."

Page 12 - OIG reports that coding errors (incidence of coding error capable of influencing DRG assignment) declined significantly between 1985 and 1988. While true, the report overstates the change. The 1985 report used a two-stage sampling design and reported error rates for the nation of 20.8 percent (hospital-weighted) and 18.6 percent (case-weighted). The update report finds a 1988 error rate of 14.7 percent but compares it to the 1985 hospital-weighted statistic. The 1988 value, based on a simple random sample of cases, is equivalent to a case-weighted statistic and should not be compared with a hospital-weighted value.

Page 15 - Some PPS details in the draft report are obscured. For example, the report discusses "base payment amounts" when it means standardized amounts. Discussion of reimbursement effects states that financial impacts "paralleled CMI changes" when they are simply calculated from and are logically equivalent to such changes.

Appendix 5 - There appears to be an error for hospitals with 300 or more beds. The response rate of 66.8 percent should be changed to 48.9 percent (1,199 responses of 2,451).

There is some disagreement between the narrative and the appendices. Table 9 illustrates the number of records per DRG and the number miscoded. Appendix 21 lists each DRG with the number of records selected and the error frequency and the proportion of errors. While the number of records selected is consistent between the table and the appendix, the number of errors is not. There is no indication why these numbers differ, the source of this data, or which is correct. However, the errors reported in appendix 21 do total 361, the correct total for the study.