# How Much Change in the Case Mix Index Is DRG Creep?

Grace M. Carter, Joseph P. Newhouse,
Daniel A. Relles

RAND

The research described in this report was supported by the Health Care Financing Administration, U.S. Department of Health and Human Services, Grant No. 99-C-98489/9-05.

**Library of Congress Cataloging in Publication Data**

Carter, Grace M.
How much change in the case mix index is DRG creep? / Grace M. Carter, Joseph P. Newhouse, Daniel A. Relles.
p. cm.
"Supported by the Health Care Financing Administration, U.S. Department of Health and Human Services."
"April 1990."
"R-3826-HCFA."
Includes bibliographical references.
ISBN 0-8330-0996-6
1. Diagnosis related groups—United States. 2. Hospital patients—United States—Classification. I. Newhouse, Joseph P.
II. Relles, Daniel A., 1943- . III. RAND Corporation.
IV. United States. Health Care Financing Administration. V. Title.
RA971.32.C37 1989
362.1'1'0973—dc20 89-35962
CIP

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

R-3826-HCFA

# How Much Change in the Case Mix Index Is DRG Creep?

Grace M. Carter, Joseph P. Newhouse,
Daniel A. Relles

April 1990

Supported by the
Health Care Financing Administration,
U.S. Department of Health and Human Services

RAND

RAND/UCLA/Harvard
Center for Health Care
Financing Policy Research

# PREFACE

This report describes the results of a project to measure how much the changes in the Medicare Prospective Payment System's case mix index reflect the resource needs of patients versus how much they reflect coding practices. The project was performed within the RAND/UCLA/Harvard Center for Health Care Financing Policy Research which is supported by cooperative agreement between the center and the Health Care Financing Administration (HCFA). The Prospective Payment Commission provided HCFA with a portion of the funds to support the project.

# SUMMARY

The Medicare case mix index (CMI)—the average diagnosis-related group (DRG) relative payment weight assigned to all discharges paid under the Prospective Payment System (PPS)—has continually increased since the introduction of the PPS. Because the CMI determines the amount hospitals are paid under PPS, this increase has been the subject of considerable debate. On the one hand, hospitals argue that most of this change is truly related to the resources required to treat Medicare patients; less-complex, lower-weighted cases—such as cataract operations—have increasingly been taken care of outside the hospital and technological change has typically been among more-complex, high-weighted DRGs. On the other hand, many observers—both inside and outside the government—are concerned that much of the change may really be an upcoding of cases, with little change in resource needs.

Distinguishing between true change and upcoding is important because a lot of money is at stake. Each 1 percent change in the CMI results in an additional $400 million in revenue for hospitals and in expenditure for the federal government. If the change is primarily upcoding (and if the original amount paid was correct), payments to hospitals should not increase much; if the change is primarily true change, however, they should.

This report presents the results of using a new method to partition the CMI increase between 1986 and 1987 into true change and coding change. The new method exploits the random sample of records gathered by the so-called SuperPRO, the private firm that independently recodes a sample of charts that hospitals and professional review organizations (PROs) have assigned DRGs.

By having SuperPRO assign DRGs to random charts drawn from different time periods, an unbiased measurement of true CMI change is possible. Because coding standards may have changed over time, we had the SuperPRO recode records from the past at approximately the same time it was coding more current records.

In addition to true change and upcoding, the measured case mix change could be affected by changes in the Grouper program that assigns DRGs. Although the weights for a new Grouper are normed on a prior year's case mix in an effort to make the change neutral, if the case distribution changes among DRGs, a new Grouper will produce a different CMI than will an old Grouper.

## METHODOLOGY

Our data come from a subset of 9234 cases in the SuperPRO database and cover roughly two years of PRO reviews. The cases we used constitute a stratified random sample of Medicare PPS discharges, with varying sample proportions in each state and fiscal year of discharge. The sample omits outlier cases, but we show that omission of this small fraction of cases is unlikely to have had a serious impact on our findings.

We decompose the CMI change into four parts, each of which is measured as follows:

1. True change, measured as the CMI change from 1986 to 1987 based on SuperPRO coding using the same, recent coding standards;
2. Coding change common to SuperPRO and the hospital, measured as the CMI change for 1986 cases between SuperPRO's latest coding (that is, using recent standards) and the codes SuperPRO originally assigned to these same cases;
3. Hospital-specific coding change, measured as the difference between the SuperPRO and hospital codings of a given case using time-of-discharge standards;
4. Grouper effect, measured by applying different Groupers to the same cases.

We compute these components of case mix change from the SuperPRO sample by determining case-level estimates of these components, then computing weighted means over the entire sample. These components add up to total change for the sample, which is an unbiased estimate of the total CMI change found in the Medicare data. Since we have just a sample of cases, the (weighted) sample mean differs from the more precise estimate of the population mean available from Medicare data. We exploit this more precise estimate of total change to adjust simultaneously the estimates of the components using a generalized least-squares regression method. Our estimate assumes that reporting of events on the medical record did not become more complete between 1986 and 1987. Although we believe this assumption to be approximately correct, to the degree the medical record became more complete our estimate of true change is overstated.

## RESULTS

Our principal finding is that between 1986 and 1987, some two-thirds of the 2.4 percent CMI change was true; the remainder was attributable to changes in coding practices and to the effect of Grouper program changes.

During 1986 and 1987, hospital coding practices were very similar to coding practices found at SuperPRO. In each year, approximately 85 percent of the cases received the same DRG from the SuperPRO coding as they had received from the hospital's coding. The cases for which SuperPRO disagreed with the hospital were almost equally split between those for which SuperPRO had raised the hospital's CMI and those for which it had lowered the CMI.

The coding practices of both SuperPRO and hospitals have changed since 1986. Many more secondary diagnoses were coded by SuperPRO during the recent reabstraction than were coded the first time SuperPRO abstracted these cases. When we applied the 1986 Grouper to the recoded 1986 cases, it measured a CMI 0.55 percent higher than it had measured on the original coding of these same cases. The 1988 Grouper is more sensitive to coding change than either of the earlier two Groupers we studied. Using the 1988 Grouper on the same recoded 1986 cases increased the CMI by 1.02 percent over the CMI produced by the same Grouper on the original coding.

The 1987 Grouper applied to the 1986 Medicare payments data measures the 1986 CMI at 0.17 percent higher than does the 1986 Grouper. This is the sum total of the Grouper effect on the 1986–1987 CMI increase, because using either Grouper to measure both years gives a very similar result for the CMI change between 1986 and 1987: 2.22 percent or 2.23 percent. The 0.17 percent Grouper effect on the 1986–1987 CMI increase results from recalibration—purely an administrative procedure—and is not the result of true change.

We recommend that there be continued monitoring and analysis of the type we have carried out. If modest changes could be made in data collection, studies such as this one and others involving PPS monitoring would be considerably easier. First, routine identification of the random sample of cases that comes to the SuperPRO should occur. Second, estimates of national statistics would be more precise if each state's sample were closer to a proportionate sample of the number of that state's discharges. Third, the random sample of cases collected at SuperPRO should include outlier cases. Finally, for other purposes, having the random sample include cases from exempt hospitals and units would be useful.

# ACKNOWLEDGMENTS

This research would not have been possible without the efforts of the SuperPRO staff at SysteMetrics, Inc., in Santa Barbara, California. We would like to thank their medical records technicians, physicians, data coders, and programmers. We also thank Cathleen Barnes, who organized and supervised the work of the SysteMetrics staff and was able to meet very tight deadlines for this work.

We thank Timothy Greene of the Health Care Financing Administration and Laura Dummit and Bruce Steinwald of the Prospective Payment Assessment Commission for providing very useful comments on an earlier version of this report. The report also benefited from very helpful reviews by RAND staff members Susan Marquis and Bill Rogers. Rosalyn Dennis patiently and carefully typed the report's many revisions. We are grateful to the Health Care Financing Administration and to the Prospective Payment Assessment Commission for supporting this work.

# CONTENTS

PREFACE ..... iii

SUMMARY ..... v

ACKNOWLEDGMENTS ..... ix

FIGURE AND TABLES ..... xiii

Section

I. INTRODUCTION ..... 1

II. METHODS ..... 6
- PRO and SuperPRO Review Process ..... 6
- SuperPRO Data ..... 7
- Selecting the Reabstraction Sample ..... 9
- Estimation Issues ..... 12
- Patbill Data ..... 14
- Estimating the Components of Case Mix Change ..... 15
- Grouper ..... 16

III. RESULTS ..... 19
- Comparison of SuperPRO and Hospital Coding ..... 19
- Changes in SuperPRO Coding over Time ..... 26
- True Change ..... 27
- Reconciliation with Patbill Data ..... 28
- Sensitivity Analysis ..... 30
- Effect of Grouper ..... 31
- Summary of the Components of Case Mix Change ..... 33

IV. FISCAL YEAR 1988 GROUPER ..... 35

V. CONCLUSION ..... 37

Appendix: ICD-9-CM CODE TRANSLATIONS FOR USE WITH ALTERNATIVE GROUPERS ..... 39

REFERENCES ..... 43

# FIGURE

3.1. Comparison of two SuperPRO codings of the same cases . . . . . 26

# TABLES

1.1. Percentage change in case mix index, by year . . . . . 1
2.1. SuperPRO random sample, by cycle and fiscal year of discharge . . . . . 8
2.2. Sizes of various samples, by fiscal year . . . . . 8
2.3. Population and sample sizes, by state . . . . . 10
3.1. Comparison of weights of DRGs assigned by different coders, current year's coding . . . . . 20
3.2. Unweighted CMIs using 1986 Grouper, current year's coding . . . . . 21
3.3. Analysis of covariance in the logarithm of the ratio of the hospital-assigned DRG to the SuperPRO-assigned DRG, 1986 Grouper . . . . . 21
3.4. Selected coefficients in analysis of covariance, 1986 Grouper . . . . . 22
3.5. Effect of exclusion of outlier cases on difference between hospital/PRO CMI and SuperPRO CMI, 1986 Grouper . . . . . 24
3.6. Comparison of weighted and unweighted estimates of the difference in CMIs assigned by different coders, 1986 Grouper . . . . . 25
3.7. Estimate of true change in CMI of nonoutlier cases . . . . 28
3.8. Preliminary and improved estimates of the components of case mix change, 1986 Grouper . . . . . 29
3.9. Preliminary and improved estimates of the components of case mix change, 1987 Grouper . . . . . 30
3.10. CMI for nonoutlier cases and for total cases—Patbill data, 1986 Grouper . . . . . 31
3.11. CMI from Medicare data, by Grouper and fiscal year . . . 32

3.12. Percentage increase in CMI, by Grouper and fiscal year . . . . . 32
3.13. Components of CMI change, 1986 Grouper . . . . . 34
4.1. CMI produced by each Grouper under old and new SuperPRO codings of the same cases . . . . . 35
4.2. Comparison of the Grouper's sensitivity to coding change . . . . . 36
5.1. Summary of components of percentage CMI change, 1986–1987 . . . . . 37

# I. INTRODUCTION

Under Medicare's Prospective Payment System (PPS), the Health Care Financing Administration (HCFA) pays the hospital an amount proportional to the case's diagnosis related group (DRG) weight for the vast majority of cases. The case mix index (CMI) is the average value of the DRG weights for all Medicare cases. Since the PPS was introduced, the CMI has continuously increased (see Table 1.1). Because each 1 percent increase is an additional $400 million in revenue for hospitals and expenditure for the government, this increase has been controversial.

On the one hand, hospitals have argued that most of this change is true change—that the CMI increase reflects a mix of more-complex cases in the hospital. This mix has occurred, hospitals argue, because less-complex, lower-weighted cases—such as cataract operations—have increasingly been taken care of outside the hospital. Furthermore, technological change has tended to increase the number of cases in the higher-weighted DRGs (for example, Medicare first paid for cardiac defibrillators in fiscal year 1987; a steady increase in coronary artery bypass grafts in the population older than 75 has occurred). If most of the change is true, the increased payments—now some $7 billion more than the 1984 level—made by the government because of the CMI change have a strong rationale: Increased patient resource requirements should be reflected in higher payments.

Table 1.1

PERCENTAGE CHANGE IN CASE MIX INDEX, BY YEAR

| Year | CMI Change |
|---|---|
| 1984–1985 | 4.2 |
| 1985–1986 | 2.6 |
| 1986–1987 | 2.4 |
| 1987–1988 | 3.7 |

NOTE: We calculated the figures for 1984 through 1987 using only PPS cases. The 1987–1988 figure is from the Health Care Financing Administration (1989).

On the other hand, many people both inside and outside the government have been concerned that much of the CMI change is DRG "creep" (Simborg, 1981)—upcoding of cases with no change in resource needs. This view holds that much of the increase has occurred because hospitals now have an incentive to code more completely and, in cases of ambiguity, to assign the most highly weighted diagnosis as the principal diagnosis. To the extent that most of the change is upcoding, the rationale for the increased payments is weakened. This report's main objective is to decompose the 2.4 percent change between 1986 and 1987 into the amount that is true change and the amount that is upcoding.

Earlier work decomposing the index's change was undertaken by Carter and Ginsburg (1985). They showed that in the first year of the PPS, little of the CMI change was true. However, their estimate of true change was largely based on a regression of 1981–1983 (pre-PPS) CMI change using a sample of Commission on Professional and Hospital Activities hospitals. Although extrapolating the rate of change from 1981–1983 to 1984 was reasonable, extrapolating it to future years becomes increasingly unreasonable—in part because the PPS itself could be expected to have a lagged effect on the rate of true change. Without such an extrapolation, the Carter-Ginsburg method breaks down. Hence, a new method for decomposing the observed CMI change into true change and upcoding was necessary.

This report presents the results of using such a new method. The new method exploits the random sample of records gathered by the so-called SuperPRO, the private firm that independently recodes a sample of charts that hospitals and professional review organizations (PROs) have assigned DRGs. The HCFA contracts with a PRO in each state to review the quality and appropriateness of care given to Medicare beneficiaries. As part of its duties, the PRO recodes a random sample of each hospital's cases to verify the accuracy of the hospital's coding. The intent of the SuperPRO's recoding is to check the PRO's accuracy in reviewing hospital coding. An expert coder at the SuperPRO reads the medical chart independently of the hospital and PRO to determine the correctness of the DRG assigned by the hospital and PRO. By having the expert coder assign DRGs to random charts drawn from different time periods, a "gold standard" for true CMI change is possible.

Of course, coding standards can change over time. To test for this possibility and correct for it if present, we had the SuperPRO recode records from the past at approximately the same time it was coding more current records. Specifically, we used cycles 3 through 6 of the SuperPRO's reviewing cycle; cycles 3 and 4 contain charts primarily from 1986, and cycles 5 and 6 contain charts primarily from 1987.

Immediately after coding the cycle 6 cases, we had the SuperPRO recode cycle 3 and 4 charts. Almost all these cases were recoded in the same 12-month period as the cycle 5 and 6 cases. To estimate true change, we then computed the CMI change as determined from contemporaneous SuperPRO codings. The remaining actual CMI change results from a combination of various types of coding effects that we also estimate.

Use of the SuperPRO random sample as a gold standard is not without problems. First, the sample omits outlier cases, and we must estimate the effect of that omission. Second, the random sample contains varying fractions of each state's beneficiaries—it is not a simple random sample. The sample is sufficiently disproportionate that employing weights equal to the reciprocal of the sampling fractions notably increases standard errors. Third, the quality of the medical records kept at some hospitals may be improving over time. We spoke to several experts at the HCFA and elsewhere; all believed that although the quality of medical records had greatly improved since the PPS's beginning in October 1983, little improvement occurred after 1986. If they are wrong, and if the recording of information in the medical record improved between 1986 and 1987 in ways that increased the DRG weight but did not reflect changes in real resource use, then we would incorrectly assign a portion of coding change to true change.

In addition to true change and upcoding, the measured case mix change is affected by changes in the Grouper. The Grouper is the computer program that assigns a DRG to each case based on diagnoses, procedures, and demographic information about the patient. The Grouper is modified annually to provide coverage for new technology and to change the classification rules in ways that will improve the relationship between DRG assignment and expected resource use. New DRG weights are assigned by the HCFA with each new Grouper. Although the weights for a new Grouper are normalized so that the CMI using the new Grouper is equal to that using the old Grouper on the latest available data, a change in the distribution of cases among DRGs will produce a different case mix under the new Grouper than it will under an old Grouper. For the sake of orderly exposition, we initially calculate our results using the fiscal year 1986 Grouper on the records from all years; we then show the effect of changing the Grouper.

If we choose the 1986 Grouper, we can decompose the overall change into true change and coding change at the hospital level. In turn, we can decompose the coding change into that element that was shared between SuperPRO and the hospitals, and the change that was unique to the hospital. Written in equations, we have the following five

relationships and our method of estimating them. (We explain the symbols used after the set of equations.)

1. Measured change = CMI87(H,C) − CMI86(H,C), or the change in the CMI from 1986 to 1987 as coded by hospitals (H) using the standards of the current year (C). Table 1.1 gives this value translated into a percentage.
2. True change = CMI87(S,R) − CMI86(S,R), or the change in the CMI from 1986 to 1987 as recoded (R) by SuperPRO (S) using the same coding standards. This is what Medicare would ideally reimburse.
3. Coding change common to SuperPRO and hospital = CMI86(S,R) − CMI86(S,C), or the change in the CMI for 1986 cases when SuperPRO recoded using later standards.
4. Hospital-specific coding change = [CMI87(H,C) − CMI87(S,C)] − [CMI86(H,C) − CMI86(S,C)], or the change in the difference between hospital coding and SuperPRO coding. This is the change resulting from changes in hospital coding standards that did not occur at SuperPRO.
5. By an identity CMI87(S,C) − CMI87(S,R) = 0, because the SuperPRO coded the 1987 cases contemporaneously (approximately); that is, no recoding was done for 1987.

The symbols used in the equations are as follows: CMI87 and CMI86 are the case mix index in 1987 and 1986, respectively; H represents the combined coding of the hospital and the PRO; S represents SuperPRO coding; C represents current-year coding (either 1986 or 1987); and R represents the recoded records.

If we add together equations 2 through 5, we obtain the expression for measured change in equation 1; one of our principal tasks in the report's remainder is to decompose the change measured in equation 1 into the components described in equations 2 through 4, and especially that of equation 2 versus that of equations 3 and 4.

Note again, however, that this decomposition is conditional upon a given year's Grouper; we shall explore the sensitivity of the decomposition to each year's Grouper. The hospital, of course, used the 1986 Grouper for the 1986 cases and the 1987 Grouper for the 1987 cases. To determine the effect of using the 1987 Grouper on the 1986 cases and conversely, we used the information from the face sheet of the 1986 charts to determine what DRG would have been assigned by the 1987 Grouper on the basis of the hospital's coding, and similarly for the 1987 charts and the 1986 Grouper.

We have a sample of SuperPRO cases from each year, which yields both an unbiased estimate of total case mix change and the decomposition of the total. However, from the Patbill file, a 20 percent random sample of all Medicare cases, we have a much more precise estimate of the total case mix change. We shall use this more precise estimate to adjust our decomposition, as the next section explains.

# II. METHODS

Our study's purpose was to decompose the CMI change between 1986 and 1987 into true change and coding change. As we just discussed, our methodology requires that we separately measure two components of coding change: hospital-specific coding change and coding change common to both SuperPRO and hospitals. This section begins with a description of the SuperPRO review process. Then we describe the SuperPRO sample used in this study and how we adjusted for three problems with the data: (1) stratification in the sample, (2) the possibility of censoring, and (3) the absence of data on outlier cases. We briefly describe the Medicare data we used and how we reconciled our estimated components of case mix change with the CMI's actual paid change. Finally, we describe how we estimated different Groupers' effects.

## PRO AND SUPERPRO REVIEW PROCESS

The PROs monitor the quality of inpatient care, the necessity of admissions, and the accuracy of DRG assignments. They select cases to review for many reasons, but among the cases they review are a 3 percent random sample of all admissions that are not outlier cases.[1] The PROs review the complete medical record for each selected case. The SuperPRO selects a random sample of each PRO's cases and obtains copies of medical records from the PRO. Medical record technicians at the SuperPRO, assisted by physicians, reabstract each medical record, assigning a primary diagnosis and applicable secondary diagnoses and recoding all procedures the medical record describes.[2] These new diagnoses feed into the official Grouper program to obtain the DRG associated with these patient characteristics.

The SuperPRO selects randomly the cases it sees from all the cases reviewed by the PROs. The only cases we use in this analysis, however, are those that are part of the PRO's 3 percent random sample. The other cases reviewed by the PROs have such idiosyncratic

[1]In the early part of our data period, the PROs for a few states (including Mississippi and Nevada) were selecting a 5 percent random sample of Medicare discharges for review rather than just a 3 percent random sample as occurred in other states and in the later period for all states.

[2]The SuperPRO also performs other review activities—for example, quality-of-care review—not relevant to this project.

sampling histories that using their data to infer population rates is impossible. Consequently, we use only random sample cases.

The SuperPRO's review activities are organized around cycles, which are timed to match each PRO's contract year. We will use data from cycles 3 through 6. Cycles 3 and 4 together span 13 months, typically including the second year of each PRO's operation.[3] Cycles 5 and 6 together almost always covered a consecutive 12 months of each PRO's operation, with each covering 6 months of the year.[4]

Because the PROs began operation at various times during the period from July 1984 to December 1984, the cases SuperPRO reviewed in any one cycle come from different calendar months in each PRO. In most states, our data span a 25-month period beginning between November 1985 and March 1986. Nine states are exceptions. In four states—Hawaii, Idaho, Ohio, and Pennsylvania—SuperPRO did no review for one cycle, and in New York, the cycle 3 medical records were unavailable for this study. In four other states—Louisiana, Missouri, New Mexico, and South Dakota—the data cover a time period longer than 25 months. The cycle's dates refer to the case's PRO review date; thus, many discharges occurred earlier than the cycle's beginning date. In sum, we have data on hospital discharges spanning all of fiscal years 1986 and 1987, with some cases from fiscal years 1985 and 1988.[5]

## SUPERPRO DATA

We had SuperPRO manually examine the PRO worksheet for each case to determine whether it was part of the random sample of cases. Table 2.1 shows the size of the random sample in each cycle and year. The cases listed in this table comprise the basic sample for the study, but not all information is available for all cases.

To examine the DRGs assigned by different Groupers, we need the complete diagnostic and procedural codes assigned by the coder. We have such information for the original SuperPRO coding of the entire sample and for the SuperPRO recoding of a random subset of cycle 3 and 4 cases; we will call this random subset the reabstraction sample.

---

[3]Cycle 3 covered 4 months of PRO operation and cycle 4 covered 5 months of PRO operation, typically with a 4-month hiatus between cycles. The exceptions are Pennsylvania, with 2 months elapsed between cycles; South Dakota, with 5 months elapsed between cycles; and Louisiana, with 7 months elapsed between cycles.

[4]The only exceptions are Missouri, with a 3-month elapsed period between cycles, and South Carolina, which covered only 11 months (with a single month being sampled in both cycles; the same case was never sampled twice).

[5]The data file also contained 30 records from fiscal year 1984, but this report did not use these.

Table 2.1

SUPERPRO RANDOM SAMPLE, BY CYCLE AND FISCAL YEAR OF DISCHARGE

| | Fiscal Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1985 | | 1986 | | 1987 | | 1988 | |
| Cycle | Cases | Percentage | Cases | Percentage | Cases | Percentage | Cases | Percentage |
| 3 | 954 | 94.8 | 1205 | 32.0 | 0 | 0.0 | 0 | 0.0 |
| 4 | 42 | 4.2 | 2114 | 56.1 | 445 | 10.6 | 0 | 0.0 |
| 5 | 10 | 1.0 | 406 | 10.8 | 1780 | 42.5 | 0 | 0.0 |
| 6 | 0 | 0.0 | 42 | 1.1 | 1961 | 46.8 | 275 | 100.0 |
| All | 1006 | 100.0 | 3767 | 100.0 | 4186 | 100.0 | 275 | 100.0 |

We have the codes assigned by the PRO and by the hospital for all cycle 5 and 6 cases and for the cycle 3 and 4 cases that are in the reabstraction sample. For the remaining cycle 3 and 4 cases, we have only the DRG assigned by the hospital and PRO as recorded in SuperPRO's database; we do not have the diagnostic and procedural codes necessary to use other Groupers to assign DRGs. As a result, we can use these cases in the constant Grouper analysis of hospital and PRO coding only when the discharge date corresponds to the Grouper being used in the analysis.

The second and third lines of Table 2.2 show the number of cases available for our analysis of the difference between the original SuperPRO coding and the combined hospital and PRO coding using the 1986

Table 2.2

SIZES OF VARIOUS SAMPLES, BY FISCAL YEAR

| Fiscal Year Sample | Fiscal Year | | | | |
|---|---|---|---|---|---|
| | 1985 | 1986 | 1987 | 1988 | Total |
| SuperPRO random sample | 1006 | 3767 | 4186 | 275 | 9234 |
| 1986 Grouper-hospital CMI | 388 | 3767 | 4120 | 275 | 8550 |
| 1987 Grouper-hospital CMI | 388 | 3264 | 4186 | 275 | 8113 |
| Reabstraction sample | 377 | 2815 | 385 | 0 | 3577 |
| Real change sample | 0 | 3263 | 4126 | 0 | 7389 |

and 1987 Groupers. We used the 1986 Grouper for all the 1986 cases; we could not use it for 66 (= 4186 – 4120) of the 1987 cases, however, because of missing diagnostic and procedural codes. For the same reason, when we used the 1987 Grouper, we had to omit 503 (= 3767 – 3264) fiscal year 1986 cases. The fourth line of the table gives the reabstraction sample's size. The calculation of the true CMI change between 1986 and 1987 uses both the reabstraction sample and the total sample of cycle 5 and 6 cases (7389 cases with discharge dates in 1986 and 1987).[6]

## SELECTING THE REABSTRACTION SAMPLE

The reabstraction sample is a subsample of the cycle 3 and 4 cases. The recoding of this sample began in October 1988. Two-thirds of the cases had been completed by mid-December, with the last abstraction completed by March 1, 1989. Because the SuperPRO work on the cycle 5 cases began in approximately January 1988, most of our cases were abstracted within the same calendar year, and all within 14 months of each other. Thus, we believe that the coding standards used for the reabstraction sample were similar to those used for cycle 5 and 6.[7]

The selection process for the reabstraction sample used state as a stratification variable because of the way SuperPRO selected its original sample. The selection of all cases seen by the SuperPRO is stratified by state, and roughly 13 percent of these cases (or an average of 43 cases per state per cycle) are part of the 3 percent random sample. However, the sample is constructed so that the SuperPRO obtains the same precision in each state; this precision results in a similar sample size from state to state. For our purposes, we prefer a sample more closely proportional to the number of Medicare cases in each state. Consequently, we oversampled the SuperPRO sample from the largest states.

Table 2.3 shows the number of Medicare PPS discharges, by state (estimated from a 20 percent sample of records), in fiscal years 1986 and 1987 and the number of cases in the total SuperPRO random

---

[6]The numbers in Table 2.2 are the effective sample sizes for each analysis. A small number of additional cases in the intended samples turned out to have missing data. One 1985 case and four 1986 cases were intended for the reabstraction sample, but were not reabstracted by SuperPRO (although we did obtain hospital and PRO codes for these cases). In addition, we were unable to get hospital diagnostic codes for two 1986 cases in the reabstraction sample.

[7]The list of codes used in the reabstractions was that of October 1, 1987, while SuperPRO's regular review used the codes approved as of the discharge date.

## Table 2.3

POPULATION AND SAMPLE SIZES, BY STATE

| State | Fiscal Year 1986 | | | | Fiscal Year 1987 | | | |
|---|---|---|---|---|---|---|---|---|
| | Discharges (1000) | SuperPRO Sample | Reabstraction Sample | Real Change Sample | Discharges (1000) | SuperPRO Sample | Reabstraction Sample | Real Change Sample |
| Ala. | 202 | 136 | 130 | 136 | 193 | 86 | 1 | 86 |
| Alaska | 5 | 62 | 14 | 14 | 5 | 25 | 4 | 25 |
| Ariz. | 121 | 42 | 39 | 41 | 119 | 100 | 0 | 100 |
| Ark. | 137 | 62 | 55 | 62 | 133 | 66 | 4 | 66 |
| Calif. | 816 | 79 | 71 | 78 | 797 | 92 | 20 | 92 |
| Colo. | 84 | 74 | 65 | 74 | 75 | 63 | 0 | 63 |
| Conn. | 107 | 57 | 57 | 57 | 99 | 120 | 23 | 120 |
| Del. | 25 | 51 | 14 | 51 | 24 | 68 | 0 | 68 |
| D.C. | 20 | 115 | 10 | 13 | 21 | 163 | 9 | 120 |
| Fla. | 595 | 53 | 50 | 53 | 555 | 51 | 0 | 51 |
| Ga. | 245 | 84 | 72 | 83 | 242 | 76 | 2 | 76 |
| Hawaii | 23 | 83 | 19 | 22 | 22 | 156 | 0 | 156 |
| Idaho | 34 | 45 | 32 | 38 | 34 | 49 | 4 | 48 |
| Ill. | 479 | 132 | 121 | 132 | 451 | 123 | 37 | 123 |
| Ind. | 222 | 136 | 129 | 136 | 220 | 117 | 4 | 117 |
| Iowa | 133 | 176 | 128 | 136 | 132 | 96 | 0 | 96 |
| Kans. | 122 | 56 | 45 | 56 | 118 | 81 | 0 | 81 |
| Ky. | 192 | 118 | 105 | 118 | 186 | 59 | 0 | 59 |
| La. | 190 | 94 | 74 | 94 | 181 | 98 | 36 | 98 |
| Maine | 52 | 25 | 23 | 25 | 50 | 95 | 23 | 95 |
| Mass. | 245 | 23 | 0 | 23 | 235 | 79 | 0 | 79 |
| Mich. | 329 | 120 | 99 | 120 | 329 | 137 | 2 | 137 |
| Minn. | 146 | 122 | 99 | 122 | 125 | 71 | 1 | 71 |
| Miss. | 147 | 153 | 147 | 153 | 146 | 157 | 2 | 157 |
| Mo. | 247 | 74 | 60 | 74 | 238 | 81 | 0 | 81 |

Table 2.3—continued

| State | Fiscal Year 1986 | | | | Fiscal Year 1987 | | | |
|---|---|---|---|---|---|---|---|---|
| | Discharges (1000) | SuperPRO Sample | Reabstraction Sample | Real Change Sample | Discharges (1000) | SuperPRO Sample | Reabstraction Sample | Real Change Sample |
| Mont. | 37 | 98 | 33 | 50 | 37 | 64 | 0 | 64 |
| Nebr. | 72 | 88 | 64 | 76 | 65 | 115 | 11 | 114 |
| Nev. | 26 | 42 | 20 | 41 | 26 | 29 | 0 | 29 |
| N.H. | 38 | 75 | 42 | 49 | 37 | 59 | 0 | 59 |
| N.Mex. | 44 | 17 | 16 | 16 | 44 | 90 | 25 | 88 |
| N.Y. | 493 | 7 | 0 | 7 | 654 | 88 | 0 | 88 |
| N.C. | 178 | 104 | 95 | 104 | 206 | 76 | 5 | 76 |
| N.Dak. | 35 | 54 | 36 | 37 | 33 | 75 | 0 | 74 |
| Ohio | 461 | 23 | 17 | 23 | 459 | 100 | 25 | 100 |
| Okla. | 147 | 59 | 56 | 58 | 142 | 121 | 35 | 121 |
| Oreg. | 79 | 69 | 57 | 69 | 100 | 56 | 1 | 56 |
| Pa. | 623 | 68 | 68 | 68 | 584 | 56 | 15 | 56 |
| R.I. | 45 | 47 | 45 | 46 | 43 | 104 | 4 | 104 |
| S.C. | 127 | 65 | 46 | 65 | 110 | 56 | 0 | 56 |
| S.Dak. | 40 | 4 | 3 | 4 | 38 | 40 | 12 | 40 |
| Tenn. | 249 | 108 | 100 | 107 | 239 | 85 | 0 | 85 |
| Tex. | 509 | 41 | 40 | 41 | 504 | 91 | 16 | 91 |
| Utah | 37 | 97 | 32 | 48 | 36 | 64 | 0 | 64 |
| Vt. | 21 | 65 | 13 | 13 | 20 | 61 | 7 | 49 |
| Va. | 182 | 114 | 108 | 114 | 203 | 150 | 43 | 150 |
| Wash. | 143 | 102 | 98 | 102 | 143 | 73 | 14 | 73 |
| W.Va. | 105 | 125 | 100 | 114 | 101 | 83 | 0 | 83 |
| Wis. | 199 | 73 | 53 | 73 | 192 | 93 | 0 | 93 |
| Wyo. | 17 | 50 | 15 | 27 | 18 | 48 | 0 | 48 |
| Total | 8825 | 3767 | 2815 | 3263 | 8764 | 4186 | 385 | 4126 |

sample. The sampling ratios varied from a low of 0.008 percent in New York in 1986 to a high of 0.678 percent in the District of Columbia in 1987. The reabstraction sample, also shown in the table, was selected entirely from cycles 3 and 4. The selection process undersampled the 1985 cases because the study's primary purpose was to measure the CMI change between 1986 and 1987.[8] We selected all available cases from the 12 states with the smallest ratio of available cases to population, and all available 1986 and 1987 cases from 24 additional states with lower-than-average sampling rates. In the remaining states, cases were sampled randomly with a sampling fraction that would produce equal sample-to-population ratios. The sample-to-population ratio in these states was higher than in any of the states where we selected all of the available sample, except that we ensured a sample of at least 20 cases from each state where it was available.[9]

The remaining columns in Table 2.3 show the sample available for our analysis of true change in the CMI. The sample consists of the recoded cases from cycles 3 and 4 and the total sample from cycles 5 and 6.

## ESTIMATION ISSUES

The SuperPRO data have three characteristics that raise issues concerning appropriate techniques for estimating population parameters: the unequal proportions of the population available in each state, the omission of outliers, and the censored 1987 sample.

### Stratified Sample

As we just noted, the sample contains widely varying proportions of the population in each state. Hence, if the parameter being estimated varies from state to state, the sample's mean will be a biased estimate of the population mean. The usual solution to this problem is to calculate a weighted sample mean, with the weights on each sample data point being proportional to the ratio of the population size to the sample size for the strata (state, in our case) containing the sample. However, when the sampling proportions are as varied as in our case, the weighted sample mean is imprecise because of the large weights given to just a few data points.

---

[8]We used 1985 data to determine how coding practices changed over time.

[9]The actual selection process occurred in two stages because of changes in our sample design. In the beginning of the project, we had hoped to obtain additional cases (that is, cases not part of the SuperPRO sample) from the larger states. This option turned out to be too costly. The effect of the process, however, was equivalent to a stratified random sample chosen in a single stage.

Our method of dealing with the stratification issue is first to examine whether we can establish the existence of state differences in the particular statistic. We do this using an analysis of variance (ANOVA) in which the explanatory variables include dummy variables for each state. If we reject the hypothesis that the collection of state effects are all zero, we will use the weighted mean. If, on the other hand, the state effects are statistically indistinguishable from zero, they might still exist and be small. In this case, we will compare the weighted and unweighted means. If they turn out to be similar, then we will feel confident using the unweighted mean, which will be more precise than the weighted mean.

## Outliers

Our sample omits all outlier cases. These are a tiny percent of the population (3.26 percent of 1986 cases; 4.16 percent of 1987 cases). They are, however, far from a random sample of cases because they typically occur in high-weighted DRGs. We use two different strategies for dealing with outliers depending on the estimation problem.

For analyses of the coding effect (comparing hospital and original SuperPRO coding; comparing original SuperPRO coding with SuperPRO recoding), we will first show that the difference in the weight assigned to a case by the two codings is related to the DRG weight. Then we will use the average DRG weight assigned to outlier cases to estimate what the coding effect would be if outliers had been included in the sample and if outliers were coded just like nonoutlier cases with the same DRG weight. We will demonstrate that the effect is small because the number of outlier cases is small.[10]

The absence of outliers is more important for reconciling the difference between the SuperPRO estimate of CMI total change and the estimate of total change in the CMI derived from the Patbill data. Because there were more outlier cases in 1987 than in 1986, the total CMI change was measurably higher (0.2 percent) than was the change in the CMI for nonoutlier cases. We will develop direct estimates of the percentage by which the CMI for nonoutlier cases increased because of coding and true change. Because the number of outlier

[10]Susan Marquis has pointed out that outlier payments may contain an effect of more upcoding by hospitals in higher-weighted cases, although the effect's direction is ambiguous. Suppose upcoding of an outlier case exists. For cost outliers, the inlier portion of the payment will be larger (the estimate of the CMI change accounts for this), but the cost threshold may be higher, thereby reducing the outlier portion (this is not accounted for by case mix change, because it is not part of the CMI). For day outliers, the inlier payment will again be larger; the outlier payment may be larger or smaller depending on how the day threshold changes and on how the geometric mean length of stay changes.

cases is so small, the CMI percentage increase resulting from coding for nonoutlier cases is also a good estimate of the CMI percentage increase resulting from coding for all cases. Thus, we will divide the increase in the total CMI minus the increase in the nonoutlier CMI into coding and true change in direct proportion to the estimated rate of increase resulting from each cause in the nonoutlier population.

### Censored Sample

As we have discussed, cases are reviewed at SuperPRO according to the date at which the PRO reviewed them, rather than according to the discharge date. As Table 2.1 showed, some cases with fiscal year 1986 discharge dates underwent review in cycles 5 and 6. Similarly, we expect that some cases with fiscal year 1987 discharge dates will undergo review in cycles 7 and 8, which were not available for this project. We will test whether such late-arriving cases are statistically different from the earlier-arriving cases and show that they are not different. Thus, we expect that the missing data have not introduced bias into our estimates.

## PATBILL DATA

We used a 20 percent sample of Medicare cases chosen from the Patbill file based on the last digit of the beneficiary's identification number.[11] We used this file to estimate the CMI level at which Medicare paid for its cases (which we will call the paid CMI) over the period from July 1, 1984 (when almost all hospitals were on the PPS), and the end of calendar year 1987. We then used the information about total CMI change found in the Patbill file to adjust our estimates of the components of case mix change from the SuperPRO sample, as we discuss below.

We also used the Patbill file for three other purposes: to develop sampling weights to make national estimates from sample statistics; to estimate the CMI change for outlier cases (because these cases are omitted from the random sample cases the SuperPRO reviews); and to estimate the effect on the CMI of procedures covered by Medicare for the first time in 1987.

---

[11]The Patbill file records information about hospitalizations that the HCFA receives from intermediaries. The Medicare Provider Analysis and Review (MEDPAR) file is the research file the HCFA constructs from the Patbill file.

## ESTIMATING THE COMPONENTS OF CASE MIX CHANGE

We compute the components of case mix change from the SuperPRO sample by determining case-level estimates of these components, then computing weighted means over the entire sample. These components add up to total change—an unbiased estimate of the total CMI change in nonoutlier cases found in the Medicare data. Because we have just a sample of cases, the sample mean differs from the more precise population mean estimate from the Patbill data. We exploit this more precise population mean estimate to adjust simultaneously our estimates of the components of change according to the regression model postulated below.

Consider the following regression problem, where the CMI is measured using any specific Grouper. Let

$\mu 1$ = change in the paid CMI from 1986 to 1987,
$\mu 2$ = true change from 1986 to 1987 in the CMI,
$\mu 3$ = coding change from 1986 to 1987 resulting from changes in implicit or explicit standards (that is, coding standards change common to SuperPRO and hospitals),
$\mu 4$ = coding change from 1986 to 1987 that results from hospital practices only,

y1 = estimate of $\mu 1$ derived from the 20 percent Patbill sample,
y2 = estimate of $\mu 2$ derived from the SuperPRO sample,
y3 = estimate of $\mu 3$ derived from the SuperPRO sample, and
y4 = estimate of $\mu 4$ derived from the SuperPRO sample.

The $\mu$'s correspond to the equations defined in the introduction. Define $\epsilon 1$ as the error term in the equation $y1 = \mu 1 + \epsilon 1$. Because by definition $\mu 1 = \mu 2 + \mu 3 + \mu 4$, we have

$$y1 = \mu 2 + \mu 3 + \mu 4 + \epsilon 1.$$

Also,

$$y2 = \mu 2 + \epsilon 2,$$

$$y3 = \mu 3 + \epsilon 3,$$

$$y4 = \mu 4 + \epsilon 4,$$

where $E(\epsilon i) = 0$, and $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$ are each independent of $\epsilon 1$ and have a covariance matrix we can estimate from the SuperPRO data.

If the $\epsilon$i's were all independent, this would be a standard weighted regression problem, with the weight on each of the four data points yi inversely proportional to the estimated variance of $\epsilon$i. But the $\epsilon$i's are not independent; thus, we have a nondiagonal weight matrix, and so must use generalized least squares (GLS).

Formally, let

$$y = \begin{bmatrix} y1 \\ y2 \\ y3 \\ y4 \end{bmatrix}, X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mu = \begin{bmatrix} \mu 2 \\ \mu 3 \\ \mu 4 \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon 1 \\ \epsilon 2 \\ \epsilon 3 \\ \epsilon 4 \end{bmatrix},$$

and define $\Sigma$ to be the covariance matrix of $\epsilon$. In matrix terms, our formulation is

$$y = X\mu + \epsilon,$$

where $\epsilon$ has mean zero and an estimated covariance matrix $\Sigma$. According to the well-known principles of GLS (see Wonnacott and Wonnacott, 1979), the minimum variance, linear, unbiased estimate of $\mu$ is

$$\hat{\mu} = (X'\Sigma^{-1}X)^{-1} \quad (X'\Sigma^{-1}y);$$

this has covariance matrix

$$\text{var}(\hat{\mu}) = (X'\Sigma^{-1}X)^{-1}.$$

In our section on empirical results, we report both the initial and the GLS-adjusted estimates of $\mu$.

## GROUPER

The Grouper is a computer program that assigns a DRG to each case based on the case's principal diagnosis (the diagnosis found, after study, to be the principal reason for the hospitalization), secondary diagnoses, procedures, age, sex, and discharge destination. As we already mentioned, each DRG is associated with a weight; the CMI is just the average of the DRG weights for a set of cases. Consequently, the Grouper is instrumental in calculating any CMI.

Different versions of the Grouper have been issued to correspond to DRG system refinements and revisions that have occurred over time. The refinements typically underwent implementation at the beginning of a fiscal year, and different Groupers were used for each of fiscal

years 1986, 1987, and 1988.[12] This procedure raises the issue of how the Grouper's revisions affect the CMI.

The DRG weight calculation is designed so that the CMI will be the same under each Grouper. In particular, weights for each successive Grouper are calibrated so that the CMI for an annual set of Medicare data is the same for the new Grouper as for the existing Grouper. Because of delays in creating data files, however, the calibration data set usually contains only cases that were paid under an earlier Grouper. For example, weights for both the fiscal year 1986 Grouper (version 3) and the fiscal year 1987 Grouper (version 4) were established using cases from fiscal year 1984.[13] Because the case distribution changes over time, different Groupers usually give different CMIs when they are used on an annual data set other than that used for calibration.

To test the sensitivity of our results to Grouper choice, we used three Grouper versions: those for fiscal years 1986, 1987, and 1988. The first two Groupers are those that were actually in use during most of our data period. The fiscal year 1988 Grouper represents the largest change in Grouper methodology since the PPS beginning and should be more similar to future Groupers than either of the other two. We calculate CMIs from our Patbill file with the three Groupers. To examine whether our estimates of the CMI's components depend on Grouper choice, we will also examine the DRGs each Grouper assigned to each of the SuperPRO sample cases.

Using different Groupers on the same cases is not a completely straightforward process because of changes over time in the ICD-9-CM (International Classification of Diseases, 9th Revision, Clinical Modification) coding system used to input diagnostic and procedural information about the case. For most changes in ICD-9-CM codes, developing a mapping that allows us to be reasonably sure which DRG the case would have received under a different Grouper is possible. For example, at the beginning of fiscal year 1987, the procedure code 36.0 (removal of coronary artery obstruction) split into several codes. Using the 1986 Grouper on cases after this date necessitates replacing each occurrence of the new subcodes (for example, 36.01) with the previous code 36.0. Also, since most subcodes are assigned to the same DRG, we can use the 1987 and 1988 Groupers on a 1986 case with code 36.0 by arbitrarily selecting one subcode to replace it. The appendix lists the translations we used for the ICD-9-CM codes.

---

[12]The 1985 Grouper was used during the first part of fiscal year 1986 because of a delay in implementing the new regulations. For convenience, we used the 1986 Grouper to measure CMI throughout the year.

[13]Department of Health and Human Services (DHHS; 1985) and DHHS (1986).

Most ICD-9-CM changes can be resolved to allow different Groupers to evaluate the cases. A few cases exist, however, for which ascertaining which DRG would have been assigned by a Grouper from another fiscal year is logically impossible. For example, at the beginning of fiscal year 1988, the primary diagnosis 996.5 (mechanical complication of other specified prosthetic device, implant, and graft) split into five subcodes based on what kind of device, implant, or graft had the complication. And, unlike most earlier coding changes, the subcodes were assigned to different DRGs. Thus, determining from the Medicare stay record of a 1986 or 1987 case with primary diagnosis 996.5 which code the case would have received in fiscal year 1988 and which DRG the 1988 Grouper would have given the case is impossible. The HCFA, in normalizing the weights for the 1988 Grouper, decided what translation to use for these codes; we believe we have followed the HCFA-developed rules. We also examined the frequency of these ambiguous cases in the SuperPRO database and found that they constituted fewer than 3 percent of the cases. Thus, any errors the ambiguous cases might have introduced will be small.

The final problem in evaluating different Groupers is that the HCFA has added coverage for some new procedures. For example, it first added coverage in 1987 for cardiac defibrillators and for cochlear implants. The 1986 Grouper does not recognize these procedure codes (see list 3 in the appendix). We applied the 1987 DRG weights for the 1986 Grouper to the 142 cases in our Patbill sample with these codes. The 142 cases raised the CMI by less than one-tenth of 1 percent, and this change is true change. None of these codes appeared in the SuperPRO sample.

# III. RESULTS

In this section, we will partition the 1986–1987 CMI increase into changes in hospital coding practices that were not replicated at SuperPRO, changes in coding practices at both hospitals and SuperPRO, and true change. We will perform the analysis using the fiscal year 1986 Grouper to measure the CMIs. The principal results using the 1987 Grouper are similiar to those of the 1986 Grouper and are summarized after we finish describing the 1986 results. Section IV considers the 1988 Grouper.

## COMPARISON OF SUPERPRO AND HOSPITAL CODING

### Descriptive Analysis

The first section of Table 3.1 shows the percentage of cases for which SuperPRO, during its normal PRO review, assigned the same DRG as the hospital assigned at discharge. In recent years, the two codings agree in approximately 85 percent of the cases. Since 1986, somewhat surprisingly, the SuperPRO has been almost as likely to raise the case's DRG weight as to lower it. On the other hand, in 1985 SuperPRO was much more likely to assign the case a lower weight than did the hospital. The 1985 data are consistent with the Hsia et al. (1988) study, showing that measurable upcoding by hospitals existed at that time.

Table 3.1 also compares the SuperPRO coding with that of the PRO. The findings are generally similar in this case to those in the previous SuperPRO-hospital comparison, with much more upcoding indicated in 1985 than in later years. Only slightly greater agreement exists between PRO and SuperPRO than between the hospital and PRO. The bottom section of Table 3.1 shows that the PRO only rarely disagrees with the hospital-assigned DRG.

Table 3.2 shows the CMI from the three codings, based on the 1986 Grouper and giving each case equal weight (that is, not correcting for the differing sampling proportions by state). In light of Table 3.1's findings, it is not surprising that in both 1986 and 1987 the three CMIs are all within one-tenth of 1 percent of each other. SuperPRO's CMI is actually the highest of the three. Although the CMI assigned by the hospital in 1988 is somewhat higher than the CMI assigned by SuperPRO, the 1988 sample is very small and the difference is not statistically significant.

## Table 3.1

COMPARISON OF WEIGHTS OF DRGs ASSIGNED BY DIFFERENT CODERS, CURRENT YEAR'S CODING
(In percent)

A. Hospital versus SuperPRO

| | Hospital | | SuperPRO | |
|---|---|---|---|---|
| Fiscal Year | Higher | Same | Higher | Total |
| 1985 | 15.4 | 76.9 | 7.7 | 100 |
| 1986 | 8.5 | 83.5 | 8.0 | 100 |
| 1987 | 8.2 | 84.3 | 7.6 | 100 |
| 1988 | 8.0 | 85.8 | 6.2 | 100 |

B. PRO versus SuperPRO

| | PRO | | SuperPRO | |
|---|---|---|---|---|
| Fiscal Year | Higher | Same | Higher | Total |
| 1985 | 11.1 | 82.4 | 6.5 | 100 |
| 1986 | 6.8 | 86.6 | 6.6 | 100 |
| 1987 | 6.9 | 87.0 | 6.1 | 100 |
| 1988 | 6.2 | 87.6 | 6.2 | 100 |

C. Hospital versus PRO

| | Hospital | | PRO | |
|---|---|---|---|---|
| Fiscal Year | Higher | Same | Higher | Total |
| 1985 | 5.3 | 91.5 | 3.3 | 100 |
| 1986 | 3.1 | 94.0 | 2.9 | 100 |
| 1987 | 2.8 | 94.6 | 2.6 | 100 |
| 1988 | 1.8 | 97.8 | 0.4 | 100 |

NOTE: The table is based on entire SuperPRO sample.

Table 3.2

UNWEIGHTED CMIs USING 1986 GROUPER, CURRENT YEAR'S CODING

| Fiscal Year | SuperPRO | Hospital | PRO | Percentage Increase over SuperPRO: Hospital | Percentage Increase over SuperPRO: PRO |
|---|---|---|---|---|---|
| 1985 | 1.1177 | 1.1598 | 1.1620 | 3.63[a] | 3.62[a] |
| 1986 | 1.1838 | 1.1830 | 1.1828 | −0.07 | −0.07 |
| 1987 | 1.1956 | 1.1953 | 1.1945 | −0.02 | −0.02 |
| 1988 | 1.2525 | 1.2706 | 1.2534 | 1.44 | 0.07 |

[a]Difference is statistically significant ($p < .05$).

The data presented so far strongly suggest that if SuperPRO had coded all the Medicare cases, it would have produced virtually the same CMI as the hospitals and PROs produced in 1986 and 1987. We next analyze the extent to which the sampling problems we discussed in Section II could be influencing this finding.

## Effect of Sample Characteristics

Table 3.3 analyzes the variance in the difference between the case's SuperPRO-assigned DRG weight and its hospital-assigned weight. We

Table 3.3

ANALYSIS OF COVARIANCE IN THE LOGARITHM OF THE RATIO OF THE HOSPITAL-ASSIGNED DRG TO THE SUPERPRO-ASSIGNED DRG, 1986 GROUPER

| Variable | Degrees of Freedom | Univariate: F-statistic | Univariate: Significance Level | Multivariate: F-statistic | Multivariate: Significance Level |
|---|---|---|---|---|---|
| State | 48 | 1.01 | 0.46 | 1.04 | 0.41 |
| Discharge month | 11 | 1.76 | 0.06 | 1.49 | 0.13 |
| Fiscal year | 1 | 0.01 | 0.93 | 0.47 | 0.50 |
| Processing delay | 3 | 0.37 | 0.78 | 0.24 | 0.87 |
| Surgical case | 1 | 0.79 | 0.37 | 21.43 | 0.0001 |
| Hospital weight | 1 | 212.98 | 0.0001 | 274.75 | 0.0001 |
| Surgical-weight interaction | 1 | 122.12 | 0.0001 | 122.12 | 0.0001 |

NOTE: The table is based on 7885 cases from fiscal years 1986 and 1987.

used a logarithmic transform because we judged it more reasonable to believe effects are multiplicative, although the implications from estimating a linear model are similar (results not shown). The first F-statistic reports a univariate ANOVA with each explanatory variable alone in the equation. The second F-statistic reports the marginal contribution of that explanatory variable to an equation containing all the other explanatory variables.

The table shows that there is no measurable effect of state, discharge month, or processing delay on the difference between the hospital's coding and the SuperPRO's coding.[1] Thus, we expect that the stratification by state in our sample has not distorted our findings and that the CMI changes reported in Table 3.2 are representative of what would have happened if SuperPRO had coded all the nonoutlier cases. Indeed, we show below that accounting for the varying sampling fractions by state causes little change in the estimates.

The strong effects that do show up in the table relate to the case's DRG weight and to whether it is medical or surgical. As the hospital-assigned DRG weight increases, the percentage amount by which SuperPRO will lower the weight increases (see Table 3.4). This effect is much more pronounced for medical cases than for surgical cases; for the latter, the difference in the two weights increases by only 2.7 percent of the hospital weight when the hospital weight increases by one CMI point, rather than by the 11.4 percent (exp(0.1077) − 1) increase found in medical DRGs. The positive coefficient on the surgical dummy compensates for the difference between medical and surgical cases in the relationship between the DRG weight and upcoding. In

Table 3.4

SELECTED COEFFICIENTS IN ANALYSIS OF COVARIANCE, 1986 GROUPER

| Variable | Coefficient | t-statistic |
|---|---|---|
| Hospital weight | 0.1077 | 16.58 |
| Surgical case | 0.0462 | 4.63 |
| Surgical-weight interaction | −0.0811 | 11.05 |

NOTE: See Table 3.3 for complete model specification.

[1]Processing delay is entered as a four-level category depending on whether the discharge occurred after the PRO cycle's start, between 1 and 6 months before the cycle's start, 7 to 12 months before the cycle's start, or more than a year before the cycle's start. In other analysis (not shown here), we examined more detailed categorizations, but also found no significant effects.

fact, little overall coding difference between medical and surgical cases exists.[2]

The strong relationship between the DRG weight and the difference between hospital and SuperPRO weights raises the possibility that the missing outlier cases, which have higher weights than the average case, could have affected our results. To determine whether any effect is quantitatively important, we estimated the average difference between the hospital's CMI and the SuperPRO's CMI on outlier cases, assuming that outlier cases would be coded by both parties just like nonoutlier cases with equivalent DRG weights.

Table 3.5 shows the results of this analysis. The columns headed "Nonoutlier CMI" are from the SuperPRO data and derive directly from Table 3.2. The "Hospital/PRO" column is a weighted average of the hospital and PRO coding, with a weight on the PRO CMI of one-third because approximately one-third of all cases are reviewed by the PRO, and the PRO-assigned DRG governs. Thus, this is an estimate of the CMI we would find in Medicare records for cases found at the SuperPRO. The hospital/PRO CMI for outlier cases is calculated directly from the Medicare records as the paid CMI for outlier cases. We used the regression Tables 3.3 and 3.4 describe to estimate the logarithm of the ratio of the paid CMI to the CMI SuperPRO would have found for these cases.[3] We then applied this ratio to the paid CMI to estimate a "SuperPRO" CMI for outlier cases. As may be seen in the table, if outlier cases were coded as nonoutlier cases with the same weight, hospitals would be upcoding them a substantial amount (0.014 CMI points in 1986 and 0.005 CMI points in 1987). The last columns of Table 3.5 are a weighted average of the data for nonoutlier cases and the outlier estimate. Although outlier cases are different, they are such a small proportion of the cases that our estimate for all cases is not appreciably different from the data for nonoutlier cases.

The last row of Table 3.5 shows the increase from 1986 to 1987 in each of the CMIs and in the CMI difference. The increase in the CMI difference in the SuperPRO data—0.00040—is our estimate of the effect of hospital-specific coding change on the CMI increase from 1986

[2]The hospital CMI minus the SuperPRO CMI is 0.0016 for medical cases and −0.0007 for surgical cases. Surgical cases, however, have a higher weight; this higher weight combined with the lower slope yields a higher intercept for surgical cases.

[3]The only data beyond those in the table needed to derive the outlier SuperPRO CMI estimates are the fraction of surgical cases and the average CMI value for surgical cases. In the SuperPRO data, the fraction of surgical cases was 26.7 percent in 1986 and 26.1 percent in 1987; the average CMI for surgical cases in 1986 and 1987 was 1.8922 and 1.9063, respectively. Among outlier cases, 50 percent were surgical in 1986 and 48 percent were surgical in 1987; the average CMI for surgical outlier cases in 1986 and 1987 was 2.5074 and 2.5220, respectively.

## Table 3.5

### EFFECT OF EXCLUSION OF OUTLIER CASES ON DIFFERENCE BETWEEN HOSPITAL/PRO CMI AND SUPERPRO CMI, 1986 GROUPER

| | Nonoutlier CMI | | | Outlier CMI | | | | Total CMI | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Fiscal Year | Hospital/ PRO | SuperPRO | Difference | Hospital/ PRO | SuperPRO | Difference | Percentage Outlier | Hospital/ PRO | SuperPRO | Difference |
| 1986 | 1.18293 | 1.18384 | −0.00091 | 1.82629 | 1.81207 | 0.01422 | 3.26 | 1.20390 | 1.20432 | −0.00042 |
| 1987 | 1.19504 | 1.19555 | −0.00051 | 1.76354 | 1.75819 | 0.00535 | 4.16 | 1.21869 | 1.21896 | −0.00027 |
| Increase | 0.01211 | 0.01171 | 0.00040 | −0.06275 | −0.05387 | −0.00888 | 0.90 | 0.01479 | 0.01464 | 0.00015 |

NOTES: Nonoutlier CMI is from SuperPRO data, with the hospital/PRO average being the sum of two-thirds the hospital CMI and one-third the PRO CMI from Table 3.2. The percentage of outlier cases and the hospital/PRO average CMI for outlier cases are estimated from the Patbill file. The SuperPRO CMI for outlier cases is estimated from the regression of Tables 3.3 and 3.4 assuming that the difference between hospital/PRO and SuperPRO coding for outlier cases is the same as that for nonoutlier cases at the same DRG weight. The total columns average the nonoutlier and outlier columns, with the weight on the outlier cases equal to the percentage of cases that are outliers.

to 1987. This estimate is the y4 component of the CMI change or, in the notation of the introduction, [CMI87(H,C) – CMI87(S,C)] – [CMI86(H,C) – CMI86(S,C)]. After adjusting for outlier cases, the difference in the total CMI increase as determined by the hospital/PRO and SuperPRO remains very small (0.00015). In sum, although substantial uncertainty exists about exactly how much upcoding occurs in outlier cases, the small number of such cases makes the possibility that their absence has substantially distorted our y4 estimate very unlikely.

For completeness, Table 3.6 compares the means and standard deviations of the case mix change estimate using weighted and unweighted estimates. The 1987 weighted estimates are almost identical to the unweighted estimates. The 1986 weighted estimates show SuperPRO raising the hospital-assigned CMI, but given the size of the standard errors, the data are consistent with the hospital and SuperPRO using similar coding practices. Note also that the standard errors of the weighted differences are roughly twice those of the unweighted differences. Thus, when weighting is necessary, the SuperPRO sample is approximately as useful as a proportional sample approximately one-fourth its size. In Sec. V, we will discuss how future SuperPRO samples could be made more powerful.

Table 3.6

COMPARISON OF WEIGHTED AND UNWEIGHTED ESTIMATES OF THE DIFFERENCE IN CMIs ASSIGNED BY DIFFERENT CODERS, 1986 GROUPER

| | Unweighted | | Weighted | |
|---|---|---|---|---|
| Fiscal Year | Average | Standard Error | Average | Standard Error |
| | Hospital Minus SuperPRO | | | |
| 1985 | 0.0421 | 0.0179 | 0.0396 | 0.0424 |
| 1986 | –0.0008 | 0.0042 | –0.0073 | 0.0079 |
| 1987 | –0.0002 | 0.0036 | 0.0000 | 0.0058 |
| | PRO Minus SuperPRO | | | |
| 1985 | 0.0443 | 0.0177 | 0.0542 | 0.0368 |
| 1986 | –0.0011 | 0.0040 | –0.0042 | 0.0069 |
| 1987 | –0.0011 | 0.0033 | 0.0000 | 0.0058 |

## CHANGES IN SUPERPRO CODING OVER TIME

In the previous subsection, we found that SuperPRO's coding of fiscal year 1986 and 1987 cases was similar to that of hospitals (our notation's y4 component). Here we investigate whether SuperPRO's coding has itself changed over time (our notation's y3 component). We use the 2815 discharges from fiscal year 1986 that SuperPRO first coded during cycles 3 and 4 (in calendar year 1987) and then recoded as part of this study, the reabstraction sample.

SuperPRO's coding behavior definitely changed between the two earlier and two later cycles. In the more recent coding, SuperPRO coded many more secondary diagnoses than it coded in its first coding (see Fig. 3.1). In the first coding, an average of 3.39 diagnoses were coded per case; in the second coding, the average rose to 3.87. This difference in the two codings is statistically significant ($t = 26.2$; $p < 0.0001$).

This difference translates into a modest increase in the CMI produced by the two codings. In this set of cases, the 1986 Grouper (unweighted by state) CMI produced by the original coding was 1.16484 and that produced by the recoding was 1.17122—an increase of 0.55 percent. Although the difference is only marginally significant ($t = 1.77$; $p < 0.10$ using a two-tailed test), the evidence that some change in coding behavior occurred is overwhelming.

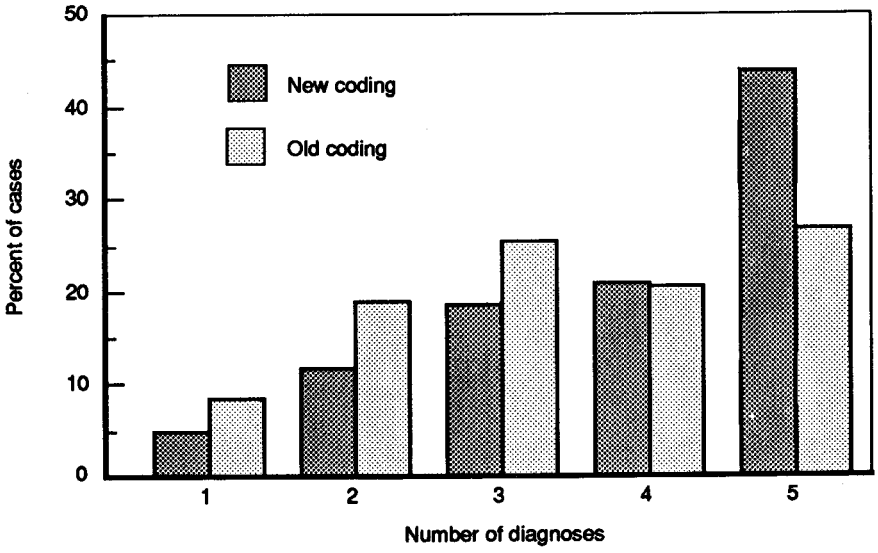



Fig. 3.1—Comparison of two SuperPRO codings of the same cases

We used an unweighted CMI estimate in the previous paragraph because we presume there is no reason to believe that SuperPRO's own coding behavior should be influenced by the state in which the discharge occurred. Just to be sure, we did an ANOVA of the logarithm of the ratio of the DRG weights assigned in the two recodings similar to the ANOVA Table 3.3 reports. Not surprisingly, we found no significant effect of state ($F = 1.18$; $p = 0.20$); consequently, we will not report a weighted estimate for the effect of changes in SuperPRO's coding practices.

Were the differences in SuperPRO's coding over time systematic? Just as in the case of the hospital-SuperPRO comparison, a statistically significant relationship exists between the increase in coding found in the second SuperPRO coding and the DRG weight assigned to medical cases in the first coding. The effect is much smaller, however, than we found for the difference between the hospital and SuperPRO codings.[4] Although outlier cases (if the sample had included them) might have received a slightly greater amount of upcoding in the second coding, the effect on the measured CMI would be insignificant because of the small percentage of outlier cases. Our estimate of the effect of coding change common to both SuperPRO and the hospital is therefore 0.00638 = 1.17122 − 1.16484. This result is y3 or, in the introduction's notation, CMI86(S,R) − CMI86(S,C).

## TRUE CHANGE

The final component of case mix change is the true change—that is, the CMI change that results from changes in DRGs because of a different mix of illness or a different mix of procedures (y2 in our notation). We can estimate this change using SuperPRO's coding of cases from cycles 5 and 6 and the contemporaneous recoding of cases from cycles 3 and 4. Because SuperPRO used essentially similar coding practices for all these codings, the difference between the 1986 and 1987 cases is an estimate of the CMI's true change.

As is well known, different parts of the country have measurably different CMIs; this fact shows up in the SuperPRO sample. In an analysis of variance in the CMI, state is statistically significant ($F = 1.45$, 48, 7342 degrees of freedom, $p < 0.03$); consequently, we will weight our estimates to produce an unbiased estimate of true change. Table 3.7 shows the results . The difference in the two contemporaneous recodings' case mix

[4]The DRG weight coefficient for medical cases for the two SuperPRO codings was 0.034; we found no relationship between DRG weight and the difference in the two SuperPRO codings for surgical cases.

Table 3.7

ESTIMATE OF TRUE CHANGE IN CMI OF NONOUTLIER CASES

| Fiscal Year | Number of Cases | CMI | Standard Error |
|---|---|---|---|
| 1986 | 3263 | 1.1901 | 0.024 |
| 1987 | 4126 | 1.1960 | 0.016 |
| Difference | | 0.0059 | 0.029 |

NOTE: Weighted based on sampling proportions in each state.

is 0.0059 points on the CMI scale, or an increase of 0.50 percent. Thus, y2 is 0.0059 or, in the introduction's notation, CMI87(S,R) – CMI86(S,R). The estimate's standard error is relatively large because of the loss in precision resulting from weighting. Indeed, the difference is not statistically different from zero.

We will deal with the omission of outliers after we reconcile our estimates of the increase's components with the CMI measured from the Patbill data.

## RECONCILIATION WITH PATBILL DATA

We next modify our estimates of the CMI increase's components from the SuperPRO data by taking into account the CMI change measured from the 20 percent sample of Medicare discharges, the Patbill data. We restrict this analysis to nonoutlier cases because only for such cases do we have reasonable quantitative methods for estimating the estimates' standard error. After completing this analysis, we will adjust for the missing outlier cases.

The estimates y2, y3, and y4 that we have just developed are unbiased estimates of the CMI increase's components from 1986 to 1987. Therefore, the sum of these three terms is an unbiased estimate of the CMI increase for nonoutlier cases. Because the SuperPRO data are just a random sample of Medicare cases, the sample estimate of the CMI increase may differ from the population increase. The Patbill file provides us with a much more precise estimate of the CMI increase. Using the GLS methodology Sec. II discussed, we can use the Patbill estimate of the total change to improve our estimates of the components of CMI change.

Table 3.8 displays both the two alternative estimates of total CMI change, and the CMI change's components. The set of estimates on

Table 3.8

PRELIMINARY AND IMPROVED ESTIMATES OF THE COMPONENTS OF CASE MIX CHANGE, 1986 GROUPER

| | Original Estimate | | GLS Estimate | |
|---|---|---|---|---|
| Component | CMI Change | Standard Error | CMI Change | Standard Error |
| Total change | .0246[a] | .0008 | 0.0246 | .0008 |
| True change | .0059[b] | .0310 | 0.0176 | .0062 |
| Coding: | | | | |
| Both SuperPRO and hospital | .0064[b] | .0036 | 0.0064 | .0036 |
| Hospital only | .0004[b] | .0052 | 0.0005 | .0052 |

[a]From Patbill.
[b]From SuperPRO sample.

the table's left is obtained directly from Patbill and our SuperPRO sample[5]; the set on the right is based on GLS.

The GLS estimate uses four separate pieces of information to make more precise the estimates of the three-parameter vector $\mu$. In the original set of estimates, the estimate of true change (y2) has by far the largest standard error: The standard error of the true change estimate in the SuperPRO samples is much larger than are standard errors in the changes resulting from coding practices. (This difference reflects the weighting of the true change estimate, whereas the coding changes common to the hospital and SuperPRO were unweighted.) The GLS procedure recognizes this higher variability of the true change component and chooses as an estimate of $\mu 2$ a linear combination of y that reflects this variability. This linear combination is approximately

$$y1 - y3 - y4.$$

This estimate of $\mu 2$ has substantially less variability than does the original one. Its value of .0176 is about three-fourths of the total change, and it has a t-value of almost 3.

Thus, the GLS estimate assigns a higher fraction of the total change to true change because, in effect, it assigns virtually all the increases in the total CMI as measured by the Patbill sample that was not replicated in the SuperPRO sample to true change. If we were to estimate

[5]The estimate of the hospital-specific coding effect used a weighted average of hospital and PRO coding to estimate the paid CMI, giving two-thirds of the weight to the hospital coding (see Table 3.5).

the fraction of true change from only the SuperPRO sample, we would estimate just under half, but then the estimate of total change would be low by almost a factor of two.

## SENSITIVITY ANALYSIS

Table 3.9 shows data similar to those in Table 3.8 but uses the 1987 Grouper instead of the 1986 Grouper. The raw estimate of true change from the SuperPRO data is somewhat larger than the 1986 Grouper estimate, but the final estimate of true change is 0.0164, which is very similar to the 0.0176 finding from the 1986 Grouper.

We also tested how sensitive our 1986 Grouper findings were to the weighting procedures we used. When we reestimated the GLS model using a case-weighted estimate of true change (one that did not account for differential sampling fractions by state), the GLS estimate of true change actually increased slightly from the .0176 estimate. When we returned to a weighted estimate of true change (that is, to the one that accounted for differential sampling fractions across states) and replaced the unweighted estimate of the hospital-specific coding effect with a weighted estimate, the true change estimate declined but remained responsible for half the change.

Table 3.9

PRELIMINARY AND IMPROVED ESTIMATES OF THE COMPONENTS OF CASE MIX CHANGE, 1987 GROUPER

| | Original Estimate | | GLS Estimate | |
|---|---|---|---|---|
| Component | CMI Change | Standard Error | CMI Change | Standard Error |
| Total change | 0.0246[a] | .0008 | 0.0246 | .0008 |
| True change | 0.0109[b] | .0311 | 0.0164 | .0061 |
| Coding change: | | | | |
| Both SuperPRO and hospital | .0043[b] | .0035 | 0.0043 | .0035 |
| Hospital only | .0039[b] | .0052 | 0.0039 | .0052 |

[a]From Patbill.
[b]From SuperPRO sample.

The y1 in these analyses is the CMI change of nonoutlier cases. The CMI of all cases in the Patbill sample increased by 0.0021 more than the CMI of nonoutlier cases because of the greater fraction of outlier cases in fiscal year 1987 than in fiscal year 1986 (see Table 3.10). The coding changes with the largest impact are those that occurred at both the hospital and SuperPRO; for these changes, little relationship between the coding change and the DRG weight existed. Although a strong relationship between DRG weight and hospital-specific coding behavior was evident, the hospital-specific effect is sufficiently small that adjusting that component (y4) for the inclusion of outliers changes the overall CMI very little (recall Table 3.5). Thus, assuming that the CMI percentage increase resulting from coding for all cases is similar to the CMI percentage increase resulting from coding for nonoutlier cases is reasonable.

## EFFECT OF GROUPER

Table 3.11 shows the CMIs calculated on the 20 percent sample of Medicare cases using the three Groupers. The 1986 and 1987 Groupers gradually grew apart during 1985 and 1986 from their starting position in 1984 when they were normalized to the same CMI. In 1986, the CMI measured by the 1987 Grouper was 0.002 (or 0.17 percent) higher than the CMI measured by the 1986 Grouper. From 1986 to 1987, the rate of increase measured by the two Groupers was virtually indistinguishable (see Table 3.12).

The total change of 2.39 percent in the paid CMI is the sum of an increase of approximately 2.23 percent (2.22 percent measured with the

Table 3.10

CMI FOR NONOUTLIER CASES AND FOR TOTAL CASES—PATBILL DATA, 1986 GROUPER

| | Nonoutlier Cases | | All Cases | |
|---|---|---|---|---|
| Fiscal Year | Number of Cases | CMI | Number of Cases | CMI |
| 1986 | 1,765,135 | 1.18988 | 1,824,569 | 1.21061 |
| 1987 | 1,752,621 | 1.21449 | 1,828,669 | 1.23733 |
| Increase | | .02461 | | .02672 |

NOTE: Data for 1987 exclude 142 cases that could not be grouped with the 1986 Grouper because they received newly approved technology.

Table 3.11

CMI FROM MEDICARE DATA, BY GROUPER AND FISCAL YEAR

| Fiscal Year | Number of Cases | Grouper 1986 | Grouper 1987 | Grouper 1988 |
|---|---|---|---|---|
| 1985 | 1,709,511 | 1.17620 | 1.17767 | 1.16956 |
| 1986 | 1,824,569 | 1.21061 | 1.21262 | 1.21070 |
| 1987 | 1,829,811 | 1.23760 | 1.23955 | 1.24005 |

NOTE: The table is based on a 20 percent sample of PPS cases. The 1987 Grouper weights were used for all cases with procedures first allowed in 1987.

1987 Grouper) and a Grouper increase of 0.17 percent. The Grouper effect results from the difference in case mix measured by the two Groupers in 1986 at the start of the period. Thus, the 0.17 percent Grouper effect on the 1986–1987 CMI increase is purely administrative and not the result of true change. Whether the 1986 or 1987 Grouper more accurately measured the increase in CMI from 1984 to 1986 is unknown.

Table 3.12 shows one other point: The fiscal year 1988 Grouper measured a higher rate of CMI change than did either of the other Groupers. We will explore the interpretation of this finding in the next section, but first we will summarize this section's findings.

Table 3.12

PERCENTAGE INCREASE IN CMI, BY GROUPER AND FISCAL YEAR

| Fiscal Years | Grouper 1986 | Grouper 1987 | Grouper 1988 |
|---|---|---|---|
| 1985–1986 | 2.93 | 2.97 | 3.52 |
| 1986–1987 | 2.23 | 2.22 | 2.42 |

NOTE: The table is based on a 20 percent sample of PPS cases.

## SUMMARY OF THE COMPONENTS OF CASE MIX CHANGE

Table 3.13 shows our decomposition of the 0.0289 CMI change (2.4 percent) between 1986 and 1987. The first set of figures for nonoutlier cases comes from our estimates from the GLS model using the SuperPRO data (Table 3.8). The next set of figures for total groupable cases (1986 Grouper) comes from assuming coding change is similar among outlier cases, but also shows the Grouper effect resulting from changing to the 1987 Grouper. The column on the far right adds the increase of .0003 from the 142 Patbill file cases that were for procedures first covered by Medicare in 1987. We estimated these cases' CMI from the 1987 Grouper. None of these cases appeared at SuperPRO; thus, we did not include them in the GLS estimates.

## Table 3.13

COMPONENTS OF CMI CHANGE, 1986 GROUPER

| Component | Nonoutlier Cases | | Total Groupable Cases | | CMI Increase Resulting from Procedures Approved in 1987 | Total CMI Increase | |
|---|---|---|---|---|---|---|---|
| | CMI Increase | Percentage | CMI Increase | Percentage | CMI Increase | CMI Increase | Percentage |
| True change | 0.0176 | 71.6 | 0.0191 | 71.6 | 0.0003 | 0.0194 | 66.9 |
| Coding change at both SuperPRO and hospital | 0.0064 | 26.2 | 0.0070 | 26.2 | 0.0000 | 0.0070 | 24.1 |
| Coding change Hospital only | 0.0005 | 2.2 | 0.0006 | 2.2 | 0.0000 | 0.0006 | 2.1 |
| Subtotal | 0.0246 | 100.0 | 0.0267 | 100.0 | 0.0003 | 0.0270 | 93.1 |
| Grouper effect | | | 0.0020 | | | 0.0020 | 6.9 |
| Total | | | 0.00287 | | | 0.0289 | 100.0 |

NOTE: Numbers do not always add because of rounding.

# IV. FISCAL YEAR 1988 GROUPER

As we have just seen, the fiscal year 1988 Grouper measures a slightly greater increase in CMI from 1985 to 1987 than did earlier Groupers. Which is a more accurate reflection of the real increase in patients' resource needs? On the one hand, the 1988 Grouper was a substantial refinement of the DRG system—one that increased the correlation between DRG weight and each case's accounting cost. From this perspective, the 1988 Grouper would appear to be the better instrument for measuring CMI change. On the other hand, the 1988 Grouper eliminated age as equivalent to a complication or comorbidity and provided a greater increase in the DRG weight for coding additional comorbidities and complications. Thus, from this perspective, the 1988 Grouper would be more sensitive to changes in coding behavior and hence the worse instrument for measuring CMI change.

Our data include 3577 cases that were first coded by SuperPRO during cycles 5 and 6 and that were recently recoded using the fiscal year 1988 rules. By examining the CMI produced under each coding by the different Groupers, we can determine how sensitive the Groupers are to changes in coding behavior.

Table 4.1 shows the CMI produced at each coding by each Grouper. Table 4.2 expresses these CMIs in the percentage CMI increase at the new coding over the CMI from the original coding. The 1988 Grouper produced a coding change on the 1986 cases (the year for which we have the most data) that was roughly twice as large as the change produced by the 1986 Grouper (1.03 percent versus 0.55 percent). Thus,

Table 4.1

CMI PRODUCED BY EACH GROUPER UNDER OLD AND NEW SUPERPRO CODINGS OF THE SAME CASES

| Fiscal Year | Number of Cases | 1986 Grouper | | 1987 Grouper | | 1988 Grouper | |
|---|---|---|---|---|---|---|---|
| | | Old | New | Old | New | Old | New |
| 1985 | 377 | 1.1142 | 1.1475 | 1.1142 | 1.1497 | 1.1231 | 1.1542 |
| 1986 | 2815 | 1.1648 | 1.1712 | 1.1677 | 1.1720 | 1.1741 | 1.1862 |
| 1987 | 385 | 1.2041 | 1.2008 | 1.2067 | 1.2040 | 1.1994 | 1.1992 |
| Total | 3577 | 1.1637 | 1.1719 | 1.1663 | 1.1731 | 1.1715 | 1.1842 |

Table 4.2

COMPARISON OF THE GROUPER'S SENSITIVITY TO CODING CHANGE

| Fiscal Year of Discharge | Percentage CMI Increase with Recent Coding Compared to Original Coding | | | | | |
|---|---|---|---|---|---|---|
| | 1986 Grouper | | 1987 Grouper | | 1988 Grouper | |
| | Increase | t-statistic | Increase | t-statistic | Increase | t-statistic |
| 1985 | 2.99 | 2.30[a] | 3.19 | 2.50[a] | 2.77 | 2.03[a] |
| 1986 | 0.55 | 1.77 | 0.37 | 1.23 | 1.03 | 2.71[a] |
| 1987 | −0.27 | 0.61 | −0.22 | 0.52 | −0.02 | 0.04 |
| Total | 0.70 | 2.49[a] | 0.58 | 2.15[a] | 1.08 | 3.27[a] |

[a]The increase is statistically different than 0 ($p < .05$).

the 1988 Grouper appears more sensitive to changes in coding practices than earlier Groupers.

One major reason why the 1988 Grouper produced a greater CMI increase than the other Groupers did is that more comorbidities and complications were listed in the recodings, and the 1988 Grouper is more sensitive to these additional diagnoses. Another reason is that the recoding noted the presence of ventilator support in nine cases, while none had been noted in the earlier coding. Consequently, the 1988 Grouper placed the case in DRG 475 with a DRG weight of 3.1757. These nine cases accounted for approximately one-third of the 1988 Grouper CMI increase between the two codings. Ventilator support does not change DRG assignment using either of the other two Groupers.

As just noted, in fiscal year 1986, where most of our data concentrate, using the 1988 Grouper on cases coded by present standards produced an increase of 1.0 percent in the CMI compared to using the same Grouper on the same cases coded by 1986 standards. It produced a 1.8 percent CMI increase compared to using both the 1986 Grouper and the 1986 standards. The magnitude of this coding effect suggests that coding effects are likely to have substantially raised the CMI between 1987 and 1988. Thus, our finding that most of the CMI change between 1986 and 1987 was real may not be generalizable to later years.

# V. CONCLUSION

Our principal finding is that between 1986 and 1987, approximately two-thirds of the CMI change was true; the remainder was attributable to more complete coding and to the 1987 Grouper's having produced a higher CMI in 1986 than did the 1986 Grouper. Table 5.1 summarizes the components of case mix change from Table 3.13, but transforms the increase in each component into percentage terms rather than just CMI points.

As we mentioned in the introduction, our estimates assume that the medical record did not become more complete between 1986 and 1987. We believe this assumption to be approximately correct. To the extent that it is incorrect, less of the measured change is true. Unfortunately, we do not see any method to quantify the effect on the CMI of any possible improvement in medical recordkeeping that may have occurred.

We recommend that there be continued monitoring and analysis of the type we have carried out. For example, the CMI change between 1987 and 1988 is considerably larger than that between 1986 and 1987, and our findings about the 1988 Grouper suggest that less of it may be true change. Ascertaining just how much of the change is true, however, will require recoding 1987 claims with 1988 coding standards.

Yale University researchers have recently proposed substantial changes in the Grouper—changes that will alter all rules for assigning cases to DRGs—and DRG weights (Health Systems Management Group, 1988). If these changes undergo implementation, repeating our study using that Grouper will be desirable.

Table 5.1

SUMMARY OF COMPONENTS OF PERCENTAGE CMI CHANGE, 1986–1987

| | Rate of Increase (Percentage) | Percentage of Increase Resulting from Cause |
|---|---|---|
| True change | 1.60 | 67 |
| Coding change | 0.63 | 26 |
| Grouper change | 0.17 | 7 |
| Total | 2.39 | 100 |

NOTE: Table 3.13 shows further details.

If modest changes could be made in data collection, studies such as ours and others involving monitoring of the PPS effects would be considerably easier and more precise. We recommend the following changes:

- Routine identification of the random sample of cases that comes to the SuperPRO should occur. We underwent considerable expense to identify such cases retrospectively, but distinguishing such cases when the PRO sends them to the SuperPRO would be an inconsequential task.
- For studies such as this one, having a sample closer to a proportionate sample would be better; indeed, for our purposes having an exactly proportionate sample—that is, a random sample proportional to the number of Medicare discharges in a state—would be good. Our weighted results were much less precise because a few observations became quite important. The larger states are underrepresented in the sample, in part because of a requirement that SuperPRO monitor each PRO. However, the performance of a PRO for a state with many beneficiaries (such as Florida) should be of more importance than the performance of a PRO in a small state. Hence, sampling rates that are more uniform across states may be in everyone's interests.
- The random sample of cases collected at SuperPRO should include outlier cases. Because of the small number of cases the cost would be modest. Alternatively, the sampling rate for nonoutlier cases could be reduced slightly by SuperPRO at HCFA's direction so that no new expense was involved.
- For other purposes, having the random sample include cases from exempt hospitals and units would be useful. Although not important for this study, the random sample is the only readily available sample of charts on patients in the Medicare program. For purposes of understanding care in exempt hospitals, having such cases in the random sample at the SuperPRO would be helpful .

# Appendix

## ICD-9-CM CODE TRANSLATIONS FOR USE WITH ALTERNATIVE GROUPERS

### LIST 1

For all cases discharged on or before September 30, 1986, to use the 1987 or 1988 Grouper,

| if the old procedure code is: | then feed the 87 or 88 Grouper the procedure code: |
|---|---|
| 360 | 3600 |
| 805 | 8050 |
| 442 | 433 |

### LIST 2

For all cases discharged on or after October 1, 1986, to use the 1984, 1985, or 1986 Grouper,

| if the procedure code is: | then feed the older Grouper the procedure code: |
|---|---|
| 3596 | 3599 |
| 3601, 3602, 3603, 3604, or 3609 | 360 |
| 5197, 5198 | 5199 |
| 8050, 8051, 8052, or 8059 | 805 |
| 4421, 4422, or 4429 | 442 |
| 4493, 4494 | 4499 |
| 5503, 5504 | 5502 |
| 5996 | 5995 |
| 5893 | 5899 |
| 6497 | 6495 |
| 8627 | 8622 |

## LIST 3

The following procedures were first paid for in 1987: 3794, 3795, 3796, 3797, and 3798 (defibrillator implant and replacement), and 2096, 2097, and 2098 (cochlear implants).

## LIST 4

For all cases discharged on or after October 1, 1987, to use an earlier Grouper,

| if the procedure code is: | then feed the older Grouper the procedure code: |
|---|---|
| 3324 | 3324 |
| 3325 | 3325 |
| 3327 | 3329 |
| 3605 | feed 1987 Grouper 3609<br>feed 1986 or earlier Grouper 360 |
| 3771 | 3774 |
| 3772 | 3774 |
| 3774 | 3776 |
| 3775 | 3789 |
| 3776 | 3781 |
| 3777 | 3783 |
| 3778 | 3771 |
| 3779 | 8609 |
| 3780 | 3773 |
| 3781 | 3773 |
| 3782 | 3773 |
| 3783 | 3773 |
| 3786 | 3785 |
| 3787 | 3785 |
| 3882 | 3829 |
| 4526 | 4526 |
| 4825 | 4825 |
| 4595 | 4593 |
| 5634 | 5633 |
| 5733 | 5733 |
| 6815 | 6814 |
| 6816 | 6813 |
| 8595, 8596 | 8599 |
| 8693 | 8699 |
| 8606 | 8609 |
| 8693 | 8699 |

**LIST 5**

For all cases discharged on or after October 1, 1987, to use an earlier Grouper,

| if the diagnosis code is: | then feed the older Grouper the diagnosis code: |
|---|---|
| 51881 | 7991 |
| 51882, 51889 | 5188 |
| 99651, 99652, 99653, 99654, or 99659 | 9965 |
| 99680, 99681, 99682, 99683, 99684, 99686, or 99689 | 9968 |

**List 6**

For cases discharged before October 1, 1987, to use the fiscal year 1988 Grouper,

| if the procedure code is: | then feed the newer Grouper the procedure code: |
|---|---|
| 3324 | 3324 |
| 3325 | 3325 |
| 3327 | 3328 |
| 3328 | 3329 |
| 3770 | 3770, 3780 |
| 3773 | 3770, 3780 |
| 3774 | 3770, 3780 |
| 3775 | 3770, 3780 |
| 3776 | 3770, 3780 |
| 3777 | 3770, 3780 |
| 3771 | 3778 |
| 3772 | 3778 |
| 3781 | 3789 |
| 3782 | 3789 |
| 3783 | 3789 |
| 3784 | 3789 |
| 3786 | 3778 |
| 5633 | 5634 |

| | |
|---|---|
| 6813 | 6816 |
| 4526 | 8899 |
| 4825 | 8899 |
| 5733 | 5733 |

NOTE: for fiscal year 1986 and earlier, also use list 1.

## LIST 7

For all cases discharged before October 1, 1987, to use the fiscal year 1988 Grouper,

| if the diagnosis code is: | then feed the newer Grouper the diagnosis code: |
|---|---|
| 7991 | 51881 |
| 5185 | 5185 |
| 5188 | 51889 |
| 9965 | 99659 |
| 9968 | 99689 |

# REFERENCES


Carter, Grace M., and Paul B. Ginsburg, *The Medicare Case Mix Index Increase: Medical Practice Changes, Aging, and DRG Creep,* The RAND Corporation, R-3292-HCFA, June 1985.

Department of Health and Human Services, *Federal Register,* Vol. 51, No. 170, September 3, 1986, p. 31488.

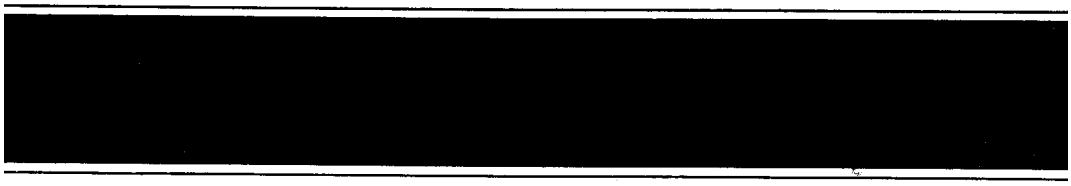
Department of Health and Human Services, *Federal Register,* Vol. 50, No. 111, June 10, 1985, p. 24373.

Health Care Financing Administration, "Inpatient Hospital Prospective Payment System and 1990 FY Rates," *Federal Register,* 1989, pp. 36452–36589.

Health Systems Management Group, *DRG Refinement with Diagnostic Specific Comorbidities and Complications: A Synthesis of Current Approaches to Patient Classification. Final Report,* Health Care Financing Administration Cooperative Agreement Nos. 15-C-98930/1-01 and 17-C-98930/1-02S1, Health Systems Management Group, School of Organization and Management, Yale University, New Haven, Conn., 1988.

Hsia, David C., W. Mark Krushat, Ann B. Fagan, Jane A. Tebbutt, and Richard P. Kusserow, "Accuracy of Diagnostic Coding for Medicare Patients Under the Prospective Payment System," *New England Journal of Medicine,* February 1988, Vol. 318, No. 6, pp. 352–355.

Simborg, D. W., "DRG Creep: A New Hospital-Acquired Disease," *New England Journal of Medicine,* Vol. 304, 1981, pp. 1602–1604.

Wonnacott, Ronald J., and Thomas H. Wonnacott, *Econometrics,* 2d ed., John Wiley & Sons, New York, 1979, pp. 427–428.

RAND/R-3826-HCFA