

How recalibration method, pricing, and coding affect DRG weights

by Grace M. Carter and Jeannette A. Rogowski

We compared diagnosis-related group (DRG) weights calculated using the hospital-specific relative-value (HSRV) methodology with those calculated using the standard methodology for each year from 1985 through 1989 and analyzed differences between the two methods in detail for 1989. We provide evidence suggesting that classification error and subsidies of higher weighted cases by lower weighted cases caused compression in the weights used for payment as late as the fifth year of the prospective payment system. However, later weights calculated by the standard method are not compressed

because a statistical correlation between high markups and high case-mix indexes offsets the cross-subsidization. HSRV weights from the same files are compressed because this methodology is more sensitive to cross-subsidies. However, both sets of weights produce equally good estimates of hospital-level costs net of those expenses that are paid by outlier payments. The greater compression of the HSRV weights is counterbalanced by the fact that more high-weight cases qualify as outliers.

Introduction

In this study, we addressed the issue of how to best control for the variation across hospitals in resource use and pricing policy during calculation of DRG relative weights. DRG weights are used for prospective payment of Medicare hospital operating costs and, starting with fiscal year (FY) 1992, for prospective payment of capital costs.

The DRG payment that each hospital receives for a case is proportional to the DRG weights. Ideally, the weights should reflect the relative cost of providing care in each DRG in an efficient hospital. If so, when a hospital's cost structure differs from its payments, it has an incentive to either become more efficient or to reduce its frequency of care for the DRGs for which its costs exceed payment. Therefore if payments are set at efficient costs, hospitals have the maximum incentives for efficiency. However, because the efficient cost structure is unknown, the DRG weights have been based on estimates of the average cost structure.

Considerations of beneficiary access and provider equity also imply weights should be proportional to costs. If the prospective payment system (PPS) paid so that average payment equaled average cost (or less than cost) and weights were not proportional to costs, then hospitals might refuse to provide care in those DRGs that were undervalued, thus seriously limiting access to needed care. Finally, inaccurate weights would cause inequity among providers, based on each hospital's DRG distribution. The equity issue is important because the case-mix index (CMI) (or average DRG weight) is the largest PPS payment adjustment and has

a substantial effect on the distribution of PPS payments.

It is not easy to measure the cost of care for cases in each DRG, however. Information about per diem costs and departmental costs and charges from the hospital's cost report can be used to transform charges into an estimate of the cost of each case. These cost estimates can then be standardized and averaged in the same way as in the usual method to produce cost-based weights. Although these cost-based weights account for consistent variations across hospitals and across departments within hospitals in percentage markup, they do not account for within-department variation in markups or for any variations across hospitals in cost. In addition, the cost reports take longer to acquire than charge data and so cost-based weights would be based on older data than the standard-method weights. Cost-based weights were used in the first 2 years of PPS.

Since 1986, the DRG weights have been calculated using charges standardized by the factors used to pay operating costs associated with teaching, disproportionate share, and input prices. The payment factors capture only a small part of the variation across hospitals in charges for any specific DRG. In addition, hospitals mark up each service differentially using inconsistent allocations of fixed costs and unknown pricing rules, thus possibly introducing further errors in the standard method.

Another method that would account for more of the cross-hospital variation in charges than the standard method is the HSRV methodology. This method would also account for cross-hospital variation in costs. The HSRV methodology differs from the current method in that hospital charges are not standardized using fixed-payment adjustments, but instead are standardized at the hospital level using hospital-specific charges and the hospital's CMI. If, within each hospital, charges were set in proportion to costs, then the HSRV method would be superior to the standard method in the extent to which it produces weights that reflect relative costs. However, because pricing rules vary and are unknown, there is no guarantee that the HSRV method will yield relative weights closer to relative costs.

The views expressed are the authors' and do not necessarily reflect those of the Health Care Financing Administration.

The research for this article was performed within the RAND/University of California, Los Angeles/Harvard Center for Health Care Financing Policy Research and funded by the Health Care Financing Administration under Cooperative Agreement Number 99-C-98489/9-08.

Reprint requests: Grace M. Carter, RAND, P.O. Box 2138, Santa Monica, California 90407-2138.

The most important criterion for choosing among methods should be the extent to which the method produces weights that reflect the relative costs of care across DRGs. Because only charges can be directly observed, however, it is not possible to measure the true relative costs of care across DRGs. For example, it is widely believed that relative weights used in the first PPS years were compressed (i.e., the resources needed by DRGs with high relative weights were underestimated and those with low relative weights were overestimated). Because we cannot measure relative costs precisely, we must fall back on secondary criteria.

One of the most important of these secondary criteria is the correspondence, at the hospital level, between average Medicare inpatient cost per case and the Medicare CMI. If the DRG weights are compressed, then the CMI for hospitals with a high CMI will be biased downward, the CMI for hospitals with a low CMI will be biased upward, and cost will not be in strict proportion to the CMI. Such CMI compression may be caused by factors in addition to DRG compression. CMI compression could also be caused by a correlation between the CMI and the likelihood that a hospital would receive cases that require greater than average resources for that DRG, or by a correlation between the CMI and a tendency to provide more resources to similar cases.

Another important criterion for comparing DRG weights is the extent to which the DRG weights improve the hospital-level correlation between payments and costs. This is not the same as the correlation between average cost per case and the CMI for two reasons: (1) outlier payments increase payment to specific DRGs over and above their share of DRG weight; and (2) the correlation between average cost per case and the CMI is usually measured via a regression of cost on CMI while controlling for teaching, disproportionate share, and input prices, and the value of these coefficients may differ from the amounts used for payment. For example, Congress currently mandates that teaching hospitals be paid for indirect medical education costs at a rate that exceeds estimates of these costs.

An additional criterion for comparing DRG weights is insensitivity to upcoding by a group of hospitals. Upcoding increases the national CMI and therefore increases payments to hospitals; upcoding by a subset of hospitals would introduce inequities in the payment rates. Yet another criterion is stability over time, which should allow hospitals additional information for planning, specialization, and/or investment decisions. The HSRV weights are not affected by changes in payment factors and thus perhaps may be more stable.

Previous research

The HSRV method was developed by Vertrees and Pettengill and is described in Lave et al. (1981). This method of calibration differs significantly from the method currently in use because it relies on each hospital's own charges to adjust for the hospital's relative costliness rather than relying on predetermined

hospital characteristics. We will describe the method in more detail in the Methods section.

Rogowski and Byrne (1990) compared HSRV and standard charge-based weights on FY 1984 data. This study showed that the two sets of weights were quite similar at a DRG level: For 89.7 percent of DRGs and 95.2 percent of cases, the two types of weights differed by no more than 5 percent. The congruence of the methodology was substantially less at the hospital level. In 1984, a shift from standard charge-based weights to HSRV weights would have changed the CMI more than 2 percent for approximately one-half of all hospitals. These hospitals accounted for 30 percent of cases.

Longitudinal comparisons between weighting methodologies have focused on differences over time between cost and charge weights. Cost-based weights account for differences across hospitals and across departments in the same hospital in average markup of charges over costs. They do not account for differences among hospitals in costs or differences in the markup of individual services within the same department. Carter and Farley (1992) showed that the differences between cost and charge weights increased only slightly from 1985 to 1987. However, the study indicated that the degree of divergence was sensitive to details of the methods used to calculate each set of weights, and especially to the rules used to eliminate statistical outliers. The growth in the national CMI was somewhat higher from 1985 to 1987 because charge weights were used rather than cost weights. Thus, overall expenditures for Medicare rose slightly faster from the use of charge weights than they would have if cost weights had been used.

A substantial amount of past research has focused on the issue of whether or not the weights are compressed. Three reasons for the hypothesized compression have been given by several authors (Pettengill and Vertrees, 1982; Lave, 1985):

- Most hospitals assign just one per diem for routine cost and one per diem for a special care unit, yet per diem nursing costs almost surely vary by DRG.
- Many believe that charges are set so that low-cost services subsidize higher cost services. To the extent that this is true, then the weights of DRGs that use those overpriced low-cost services (which tend to be low-cost DRGs) will be overestimated and the weights of high-weight DRGs that use the underpriced higher cost services will be underestimated.
- Errors in classification of cases into DRGs will tend to make the weights more similar than they should be.

Pettengill and Vertrees (1982) used simulation to show the effect of varying amounts of classification error on weight compression. Although classification error decreased substantially with the introduction of PPS, Carter, Newhouse, and Relles (1991) showed that improvements in coding occurred in response to the substantial refinements in the grouper in 1988.¹ Thus,

¹The grouper is the computer program that assigns a case to a DRG based on diagnoses and procedures and less frequently on age and discharge destination.

classification errors may still be a source of compression in the weights.

Cotterill, Bobula, and Connerton (1986) found that 1981 hospital cost per case showed little CMI compression. Thorpe, Cretin, and Keeler (1988) used a similar model to estimate CMI compression from 1984 data. They found that charge-based CMIs appear significantly compressed when one controlled only for factors affecting payment. However, the magnitude of the estimate of compression declined substantially when bed size was included in the model. Cotterill et al. had included bed size in their model.

Rogowski and Byrne (1990) showed that in 1984 cost weights were even more compressed than charge weights. Cost-based weights calculated with the HSRV methodology were the most compressed among the tested methods used to calculate weights. Further, charge-based weights calculated with the HSRV methodology were more compressed than were charge weights computed with the standard methodology.

Research questions

We examine whether HSRV and standard weights for the same DRG diverged during the period 1985-89. Similarly, we examine the degree of congruence between the CMIs of individual hospitals produced by each method. We also examine the characteristics of hospitals most affected by differences in the weights and explicate how these characteristics cause the two methods to produce different weights for certain DRGs. In particular, we develop evidence that suggests that cross-subsidies and other hospital pricing policies are an important source of the differences between the two sets of weights.

We examine whether CMI compression remained during the fifth year of PPS. We hypothesize that DRG compression and, therefore, CMI compression declined over time as a result of reduction of the classification errors, and we demonstrate that the improvement in coding that occurred in FY 1988 led to a measurable decrease in compression. We also show how hospital pricing policies affect compression.

We determine which set of weights is more compressed and which would produce the highest correlation between payment (under current rules) and cost. Finally, we examine which set of weights would have produced a larger increase in the CMI from 1986 to 1988 when coding improvements were a substantial component of the increase.

Data

The primary data set used for this project is a 20-percent sample of inpatient hospital bills for Medicare discharges occurring in fiscal years 1985 through 1989. The records used for this analysis exclude cases at exempt hospitals and units in PPS States and also exclude cases in Maryland and New Jersey where hospitals were not paid under PPS. Although New York and Massachusetts did not join PPS until FY 1986, their bills are included in the sample in all 5

fiscal years. Puerto Rican hospitals are included only since they joined PPS.

In order to calculate charge-based weights, we needed data on the factors used for standardization of charges: the PPS wage index, cost-of-living adjustment, teaching payment factor, and disproportionate share payment factor. For each year, we used an estimate of the factors actually in use in that year. For 1985 and 1986, we used data from RAND's extract of an early version of the Provider-Specific file. For 1987 through 1989, we used a file provided by HCFA that contained the payment adjusters for these years.

In addition to our longitudinal analyses, we present cross-sectional analyses of cases discharged during FY 1989 in order to explicate how the methods result in different weights. Given the constant definition of DRGs during a single Federal fiscal year, this comparison will demonstrate clearly the difference between the two methods. We also analyze cost and other characteristics of cases discharged during PPS5 because cost data are gathered only for a hospital's fiscal year.

In the latter cross-sectional analysis, we draw the total cost per case at each hospital during PPS5 from the "capital regression public use file," which is available from HCFA. All needed data were available on 4,890 hospitals during PPS5.

Methods

The HSRV methodology was applied to charge data from each year's bills to create weights. In the HSRV method, charges are standardized at the hospital level using hospital-specific charges. The total charge for each case is divided by the average charge for the hospital in which the case occurred. The resulting ratio is then multiplied by the hospital CMI to produce a hospital-specific relative charge. The process of calculating the weights is iterative. Initial values are chosen for the CMI of each hospital (for example, the CMI based on the standard charge-based weights). DRG weights are then set in proportion to the average value of the hospital-specific relative charges. Using the new DRG weights, a new CMI can be calculated for each hospital and therefore new hospital-specific relative charges. The process is continued until there is convergence between the weights produced at adjacent steps, for instance when the maximum difference is less than 1 percent.

Rogowski and Byrne (1990) showed that the algorithm is not sensitive to starting values of the CMI. We verified this analysis by comparing HSRV weights that used a starting value of 1 for each hospital's CMI and those that used a starting value from the actual paid weight. The two results were indistinguishable to 4 significant digits after five iterations. We report those that used a starting value of 1, but the alternative results are equivalent.

In calculating HSRV weights, we excluded cases in each DRG that fell outside a 3 standard deviation range of the first iteration of hospital-specific relative

Table 1
Standard deviation of HSRV weights and standard weights: Fiscal years 1985-89

Fiscal year	Case-weighted standard deviation			DRG-weighted standard deviation		
	Number of cases	HSRV	Standard	Number of DRGs	HSRV	Standard
1985	1,796,084	0.63667	0.66239	421	0.7891	0.8216
1986	1,804,519	0.64125	0.66569	424	0.7806	0.8139
1987	1,835,664	0.64826	0.67696	423	0.9318	0.9725
1988	1,845,844	0.73914	0.77619	426	1.0582	1.0868
1989	1,777,818	0.76122	0.80044	426	1.0702	1.1073

NOTES: HSRV is hospital-specific relative-value. DRG is diagnosis-related group.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

charges. We investigated whether an exclusion rule based on later HSRV iterations would produce measurably different weights, but found that it had no effect. For example, we analyzed the 1988 cases that were trimmed under each method. There are a total of 1,859,598 PPS cases in our analysis file. Comparing the first and last iterations of the HSRV, almost exactly the same cases are trimmed in both instances. There were 13,679 cases trimmed in the first iteration and 13,817 trimmed in the last iteration. Of these, 13,435 were trimmed in both iterations; thus 98.2 percent of the first iteration of trimmed cases remained at the end; and fewer than 400 additional cases were removed.

We also applied the standard methodology to charge data from each year's bills to create standard-method weights based on our sample of cases. We standardized each sample case using our best estimate of the payment adjustments in use in that year. The standardized charge for the case is given by:

$$\text{charges} * (\text{frac} / \text{wageindex} + (1 - \text{frac}) / \text{cola}) / (1 + \text{dsh} + \text{ime}),$$

where:

- frac* = fraction of Federal payment that is labor related,
- wageindex* = wage index for the hospital,
- cola* = cost-of-living adjustment for the hospital (this is 1 everywhere except Alaska and Hawaii),
- dsh* = payment rate for disproportionate share at the hospital, and
- ime* = payment rate for the indirect cost of medical education at the hospital.

In calculating standard weights, we excluded cases that were outside 3 standard deviations in the distribution of the log of standardized charges in the DRG.

For both methods, we used the DRG definitions that were used for payment of each year's cases. In addition to the calculation of annual weights, we also calculated a second set of weights from our FY 1986 Medicare provider analysis and review (MEDPAR) file where each case was classified by the FY 1988 DRG definitions rather than by the DRG definitions used for FY 1986 payments.² We used this second set of weights to

measure the CMI increase from FY 1986 to FY 1988 and to verify our finding that weights estimated on the 1988 file were much less compressed than weights estimated on the 1986 file.

We also simulated payments during 1988 and 1989. The simulations assumed a fully implemented capital PPS with a rate based on deflating the FY 1992 capital payment rate (*Federal Register*, 1991). The operating payment rates correspond to the national standardized amounts published in each year's *Federal Register*. We used the FY 1992 outlier payment methodology. To calculate payments, we multiplied our relative weights (which have a mean of 1 in each year) by the case-weighted average of the weights that were actually used for payment each year.

Longitudinal comparison

Table 1 shows the standard deviation of the DRG weights produced by each method in each year from FY 1985 through FY 1989. Refinement of the DRGs has increased the spread of weights across DRGs since 1986. The large change in DRG definition that occurred in FY 1988 is apparent in the contemporaneous increase in the standard deviation of the case-weighted weights.

In each year, the standard deviation of the HSRV weights is smaller than the standard deviation of the standard weights. However, the magnitude of the difference is small compared with the temporal increase in the standard deviation. For example, from FY 1986 to FY 1989, the standard deviations of the HSRV and standard weights grew 18.7 and 20.2 percent, respectively. The standard deviation of the standard weight exceeded that of the HSRV weight in FY 1989 by only 5.2 percent (derived from table).

In the first column of Table 2, we report the case-weighted average of the absolute value of the difference between HSRV and standard weights. Because payments are roughly proportional to the DRG weight, this measure is roughly proportional to the fraction of dollars that would be redistributed across cases if one moved from one system of weights to the other. Thus, the use of HSRV weights would have changed payment by 2.8 percent in FY 1989 for the average case. The last four columns provide the percent of DRGs and cases with HSRV weights that were within at least 5 percent and within at least 10 percent of the standard weight. These numbers are roughly consistent with the 1984 findings from Rogowski and Byrne (1990) who found that 95.2 percent of cases had HSRV charge-based

²Before applying the 1988 grouper, we recoded the codes from the *International Classification of Diseases, 9th Revision, Clinical Modification* on the FY 1986 files to account for changes in these codes.

Table 2
Mean value of absolute difference between HSRV and standard weight,
and percent of cases and DRGs with absolute difference of less than or equal to
5 and 10 percent: Fiscal years 1985-89

Fiscal year	Mean absolute difference	≤ 5 percent		≤ 10 percent	
		DRGs	Cases	DRGs	Cases
1985	0.0238	90.97	93.01	98.10	99.66
1986	0.0238	91.04	95.97	98.11	99.94
1987	0.0253	86.52	95.55	98.11	99.57
1988	0.0264	80.99	90.61	96.71	99.39
1989	0.0280	76.06	86.93	95.31	99.05

NOTES: HSRV is hospital-specific relative-value. DRGs are diagnosis-related groups.
 SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

Table 3
Percent difference between HSRV and standard weights,
by magnitude of weight and year: Fiscal years 1985-89

Fiscal year	HSRV minus standard weight as percent of standard weight					
	Percentile of standard weight			Percentile of HSRV weight		
	<25 percent	25-75 percent	>75 percent	<25 percent	25-75 percent	>75 percent
1985	1.93	1.12	-1.33	1.49	1.23	-1.30
1986	1.79	0.81	-1.02	1.39	0.89	-1.02
1987	2.07	1.49	-1.57	2.04	1.24	-1.19
1988	2.87	1.56	-1.77	2.39	1.68	-1.76
1989	2.96	1.74	-1.95	2.85	1.76	-1.94

NOTE: HSRV is hospital-specific relative-value.
 SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

weights that were within 5 percent of standard weights and that 99.5 percent of cases were with 10 percent.³

The difference between HSRV weights and standard weights has widened over time. However, even in FY 1989 the difference between the HSRV and current weights was substantially smaller than the difference between charge and cost weights found in all earlier studies of PPS cases. For example, Carter and Farley (1992) found that only 66.6 percent of FY 1987 cases had cost-based weights that were within 5 percent of standard weights, and that 99.5 percent of cases were within 10 percent, and that the mean absolute difference between cost and charge weights was 4.5 percent (rather than the 2.8 percent between HSRV and standard weights).

The difference between an HSRV weight and the corresponding standard weight is strongly related to the magnitude of the weights. The first three columns of Table 3 divide cases according to their percentile ranking based on the standard weight. The first column shows that the FY 1985 HSRV weights averaged 1.93 percent higher than the standard weight for the 25 percent of cases with the lowest standard weight. In all years, the HSRV weights tend to be larger than standard weights for low-weight DRGs and smaller for high-weight DRGs. The magnitude of the effect increases somewhat over time. The last three columns

of Table 3 show that the results are similar if one uses the HSRV weight to define low- and high-weight DRGs.

Table 4 shows that the differences between HSRV and standard weights in each DRG translate into only modest differences between the CMI's that each hospital would experience under HSRV and standard weights. The typical hospital would have experienced a 1.25-percent difference in its CMI (and therefore in its payment) in FY 1989 by using weights calculated by the HSRV method compared with weights calculated by the standard method. The case-weighted absolute differences are only slightly smaller. The Pearson correlation coefficient between the CMI's calculated under the standard and HSRV methods is also very high, 0.999. The widening over time of the difference between the two CMI's is particularly evident in the increasing proportion of hospitals (and cases) where the two CMI's differ by more than 2 percent.

Why method affects case-mix indexes

Table 5 shows the distribution of changes across individual hospitals. The first column shows that HSRV weights increase the CMI for a large majority of hospitals. Only about 11.9 percent of hospitals experience more than one-half of 1 percent decline in CMI. The hospitals that would lose under HSRV tend to be larger than the average hospital. These 11.9 percent of hospitals care for almost 30 percent of cases and have almost 200 more beds than the average hospital (derived from table). As is to be expected from our DRG-level findings, the amount of the loss under HSRV is strongly correlated with the level of a

³The slightly greater congruence of the weights from the two methods in Rogowski and Byrne (1990) may be attributed, at least in part, to the fact that they used exactly the same cases for both algorithms. As explained in the methodology section, we removed statistical outliers separately for each method.

Table 4
Mean value of absolute difference between HSRV and standard CMI, percent of providers,
and percent of cases with absolute difference of less than or equal to 2 and 4 percent:
Fiscal years 1985-89

Fiscal year	Mean absolute difference in CMI		Absolute different ≤ 2 percent		Absolute difference ≤ 4 percent	
	Hospital weighted	Case weighted	Providers	Cases	Providers	Cases
1985	0.0085	0.0072	98.7	99.5	99.9	100.0
1986	0.0084	0.0074	98.9	99.4	99.9	100.0
1987	0.0098	0.0087	97.6	98.7	99.8	99.9
1988	0.0112	0.0093	92.4	98.2	99.4	99.9
1989	0.0125	0.0101	86.7	96.1	99.0	99.9

NOTES: HSRV is hospital-specific relative-value. CMI is case-mix index.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

hospital's CMI, no matter which method is used to measure the CMI.

The column headed "standardization factor" in Table 5 gives the average value of the multiplier used to transform charges to standardized charges in the standard method. A smaller value means that charges are more heavily discounted because the hospital is paid based on a higher than average multiplier. Perhaps surprisingly, the standardization factor is negatively correlated with the hospital's loss under HSRV. The hospitals that would lose under HSRV already have their own charges heavily discounted by the standard method. Thus, the HSRV method discounts their charges even further in calculating weights. The last three columns of the table show three characteristics that have a large influence on the standardization factor. The losing hospitals are 96 percent urban and are much more likely than winner hospitals to be teaching and disproportionate-share hospitals. Indeed, roughly three-quarters of the losing hospitals receive either a teaching or disproportionate-share supplement.

Is there something specific about the case mix of the small number of losing hospitals that leads them to lose case weight and therefore payment under HSRV? The answer turns out to be clearly yes, at least for the vast majority of losing hospitals. Table 6 decomposes the change in case weight for all cases in the hospitals that would lose under HSRV, by major diagnostic category (MDC) and whether the DRG is surgical or not. This decomposition is compared with a similar one for other hospitals. For each hospital group, the change in case weight contributed by each DRG is the HSRV weight minus the standard weight multiplied by the number of cases. In Table 6, the change in the case weight is summarized over DRGs in each category and expressed as a percent of the total change in case weight for the group—roughly 8,500 weighted cases in each group. A positive number means that the average HSRV weight is larger than the standard weight in the category. In general, medical cases receive higher weights under HSRVs, and many surgical DRGs received lower weights under HSRVs. As one can see from the table, the hospitals that lose under HSRV lose more case weight on MDC 5 (circulatory system surgery cases) than they lose overall. They make some of it back on their medical cases. The hospitals that win under HSRV

weights gain on their medical cases, particularly in MDCs 4 (respiratory system), 5 (circulatory system), and 6 (digestive system).

Part of the difference between losing hospitals and others is due to the losing hospital's higher involvement in circulatory system surgery. The losing hospitals care for only 30 percent of all of Medicare's cases but they performed 63.5 percent of the circulatory system surgery. However, much of the difference is also a result of differences in the kind of circulatory system surgery performed in the two kinds of hospitals. The losing hospitals are heavily involved with the most expensive cardiac surgery. These hospitals performed over 90 percent of all Medicare surgery in each of DRGs 103 (heart transplant), 104, 105 (cardiac valve procedures with and without catheterization), 107 (coronary bypass without catheterization), 108 (other cardiothoracic or vascular procedures with pump), and 88.4 percent of cases in DRG 106 (coronary bypass with catheterization). If one ranks the DRGs on their contribution to the change in case weight within the losing hospitals, all of DRGs 104 through 108 are counted within the top 10 (DRG 103 does not count because of its small volume). The remaining DRGs within the top 10 and the fraction of cases in the losing hospitals are: 112 (75 percent), 110 (46 percent), 468 (35 percent), 124 (59 percent), and 474 (34 percent). DRGs 110, 468, and 474 caused more case-weight loss to the other hospitals than to the losing hospitals. Further, 83.1 percent of the hospitals that lost case weight under HSRV engaged in cardiac surgery in these very expensive DRGs, compared with only 5.0 percent of other hospitals. These very expensive cardiac surgery cases account for 79 percent of the total loss in case mix that all the losing hospitals experienced under HSRV.

The loss in case mix under HSRV occurs because this difference in case mix is combined with charges for cases that are much higher than average in more typical DRGs in the losing hospitals. Standardized charges (i.e. charges adjusted by payment factors for input prices, teaching, and disproportionate share as described in the methodology section) per unit of HSRV case-mix weight average \$8,385 for the hospitals that would lose case weight under HSRV versus only \$7,598 for other hospitals. The difference in standardized

Table 5
Characteristics of hospitals, by the difference between the fiscal year 1989 HSRV and standard CMIs

HSRV CMI compared with standard CMI	Number of hospitals	Cases in thousands	Case-weighted average						Type of hospital		Disproportionate share
			Hospitals	Cases	Beds	HSRV CMI	Standard CMI	Standardization factor	Urban	Teaching	
			Percent		Percent		Percent		Percent		
All hospitals	5,580	1,791	100.0	100.0	307	1.000	1.000	0.978	55.1	21.3	27.9
At least 3 percent less	54	33	1.0	1.8	562	1.387	1.428	0.865	96.2	71.7	52.8
2 percent less	153	122	2.7	6.8	487	1.253	1.278	0.879	96.1	68.0	49.0
1 percent less	460	373	8.2	20.8	494	1.116	1.127	0.906	96.3	60.0	46.3
Equal	919	476	16.5	26.6	322	1.004	1.004	0.927	92.9	43.8	45.4
1 percent more	1,764	565	31.6	31.5	206	0.920	0.912	1.024	60.3	17.2	28.4
2 percent more	1,541	186	27.6	10.4	85	0.820	0.805	1.163	22.9	3.3	15.2
At least 3 percent more	689	37	12.3	2.0	57	0.713	0.691	1.153	22.8	3.4	12.7

NOTES: HSRV is hospital-specific relative-value. CMI is case-mix index.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

Table 6

HSRV weight minus standard weight, by major diagnostic category (MDC), medical or surgical category, and whether hospitals would lose case weight under HSRV as a percent of total change in case weight in a hospital group: Fiscal year 1989

MDC	Number of cases		All cases		Medical		Surgical ¹	
	Hospitals losing under HSRV	Other hospitals	Hospitals losing under HSRV	Other hospitals	Hospitals losing under HSRV	Other hospitals	Hospitals losing under HSRV	Other hospitals
Total	528,203	1,262,898	-100.0	100.0	31.9	143.0	-131.9	-43.0
1	40,935	100,966	-4.5	-0.9	1.3	4.6	-5.8	-5.5
2	5,803	9,030	2.0	3.8	0.1	0.2	2.0	3.6
3	5,774	14,880	0.9	3.2	0.7	3.0	0.1	0.2
4	54,904	186,274	-1.0	24.2	8.2	38.8	-9.2	-14.6
5	160,813	317,316	-105.4	13.7	6.1	43.2	-111.4	-29.5
6	53,762	153,311	7.2	23.8	6.6	20.6	0.6	3.2
7	15,464	41,293	3.1	10.3	1.4	4.2	1.7	6.1
8	51,002	115,598	6.1	18.1	2.5	7.0	3.7	11.1
9	13,628	36,240	1.0	3.4	1.4	4.5	-0.4	-1.0
10	17,566	53,067	1.4	5.9	2.0	6.7	-0.6	-0.8
11	25,543	62,187	-0.3	3.3	1.9	5.9	-2.2	-2.6
12	14,543	36,777	1.5	3.9	0.4	1.1	1.1	2.7
13	7,615	15,285	1.7	3.6	0.2	0.3	1.5	3.3
14	355	613	0.1	0.2	0.1	0.1	0.0	0.0
16	5,311	13,218	0.5	1.5	0.5	1.5	0.0	-0.1
17	17,921	22,899	-4.1	-4.7	-3.3	-3.7	-0.8	-1.0
18	9,767	24,969	-2.0	-3.3	0.2	0.2	-2.3	-3.5
19	5,833	15,219	1.0	2.7	1.2	3.1	-0.2	-0.4
20	1,746	6,373	0.1	0.7	0.0	0.0	0.1	0.7
21	10,527	18,243	-2.5	-1.7	0.0	0.8	-2.4	-2.5
22	327	697	0.6	1.0	0.0	0.1	0.6	0.9
23	2,076	4,443	0.0	0.3	0.2	0.7	-0.2	-0.4
99	6,988	14,000	-7.4	-13.0	0.0	0.0	-7.4	-13.0

¹Includes DRGs that contain both medical and surgical cases.

NOTES: HSRV is hospital-specific relative-value. Definitions of the MDCs may be found in 3M Health Information Systems (1992).

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

charges for the two groups of hospitals is reduced by about 20 percent using standard weights, but still the hospitals that lose under HSRV charge more per case. Because the HSRV hospitals charge more than average for typical cases, their high charges for the very expensive cardiac surgery cases are downweighted in calculating the HSRV weights. As these hospitals account for almost 90 percent of the cases in DRGs 103 through 108, the HSRV weights in these DRGs are substantially lower than their standard weights.

Interestingly, the hospitals that lose under HSRV do not charge more than would be expected for their expensive cardiac surgery cases. They actually charge less than other hospitals for cases in DRGs 103 through 108. One cannot make too much of these lower charges because of their near monopoly on these cases and because they typically have much higher volume over which to amortize the high fixed costs associated with this type of surgery. However, it is interesting that their charges for expensive cardiac procedures where there are competing hospitals are quite in line with the competition. For example, in DRGs 109 through 112, standardized charges per unit of HSRV case weight average \$8,408 in the hospitals that lose under HSRV and \$8,349 in other hospitals. The hospitals that lose under HSRV perform 63 percent of the surgery in these DRGs.

The finding that the hospitals that lose under HSRV charge more than expected for their typical cases but not for their expensive cardiac surgery cases is consistent with these hospitals subsidizing very expensive services with excess revenue from less expensive services. This kind of cross-subsidization has been hypothesized to cause compression in the charge weights. Although the findings are consistent with cross-subsidization, they are also consistent with the hospitals that lose under HSRV providing more intense services than other hospitals to patients in typical DRGs. This second explanation is plausible because the hospitals that lose under HSRV are mostly teaching hospitals engaged in very high-technology cardiac surgery and thus may engage in a more resource-intensive style of medicine or may receive sicker patients.

In order to examine these two competing hypotheses, we decided to look for cross-subsidizing behavior among all hospitals. If other hospitals cross-subsidize, then it increases the likelihood that cross-subsidization is responsible for the charges at losing hospitals that are higher than expected for typical cases but not for expensive cardiac surgery cases. This test provides only weak evidence, not certain proof, that the losing hospitals engage in similar behavior. We divided DRGs into two groups at the case-median weight and then we

Table 7

**Standardized charge per unit of DRG weight, by DRG weight group and hospital characteristic:
Fiscal year 1989**

Hospital characteristic	Standardized charge divided by HSRV CMI			Standardized charge divided by standard CMI		
	Low DRG weight	High DRG weight	All DRGs	Low DRG weight	High DRG weight	All DRGs
All hospitals	\$7,698	\$6,803	\$7,249	\$7,886	\$6,628	\$7,255
CMI percentile						
0-25	6,954	5,950	6,522	7,158	5,897	6,615
25-75	7,979	6,904	7,436	8,171	6,748	7,453
75-100	8,029	7,275	7,602	8,187	6,968	7,497
HSRV CMI minus standard CMI, percent						
≤ -3	8,739	7,578	8,043	8,876	7,107	7,816
-2	8,006	7,399	7,652	8,149	7,045	7,505
-1	8,103	7,365	7,706	8,271	7,081	7,630
0	8,277	7,221	7,736	8,468	7,049	7,741
1	7,563	6,362	6,993	7,768	6,272	7,056
2	6,307	5,297	5,869	6,505	5,292	5,979
≥ 3	5,535	4,218	5,063	5,752	4,235	5,208

NOTES: DRG is diagnosis-related group. HSRV is hospital-specific relative-value. CMI is case-mix index.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

calculated the standardized charge per unit of DRG weight for each hospital and each of the two DRG groups. If the DRG weights provide an unbiased estimate of relative resource use in this group and there were no cross-subsidies, then one would expect the standardized charge per unit of DRG weight to be about the same for the two DRG groups within each hospital. Instead, we found that 76 percent of all hospitals had standardized charges per unit of standard DRG weight that were higher for the low-weight DRG group than for the high-weight DRG group. On average, standardized charges per unit of DRG weight in the low-weight DRG group exceeded that in the high-weight group by 19 percent. Using the HSRV weights, charges are higher than expected in the group of low-weight DRGs in 71 percent of all hospitals. The average difference between the two groups in standardized charges per unit HSRV weight was 13 percent. If the DRG weights are unbiased estimates of resource use, then this would demonstrate that cross-subsidization occurs. It has been said that the subsidies may not be deliberate, but might instead be caused by a lack of sophistication of management information systems. The inability to accurately cost out procedures may tend to average costs out across high- and low-weight DRGs. Management using this cost information to set prices may then level out charges accordingly.

Compression of the DRG weights cannot cause this apparent cross-subsidization. If the weights are actually compressed, then the true relative resource use of low-weight DRGs is actually lower and the ratio of standardized charges to "true" DRG weight for low-weight DRGs is even higher, and even more cross-subsidization is occurring than one would estimate from the assumption of unbiased weights.

Table 7 shows that the apparent cross-subsidization was not limited to any one segment of the CMI

distribution and does not determine whether hospitals would win or lose under HSRV. Table 7 does, however, illuminate the causes of the variation among "other" hospitals in the change in CMI under the HSRV method. Among other hospitals, the average standardized charge per unit of DRG weight (either method) decreases with an increase in the size of the gain under HSRV. The low charges for higher weight DRGs of the hospitals that won the most contributed to causing their typical cases (which are predominantly very low-weight medical cases) to have higher weights under the HSRV method than under the standard method.

Compression

An important criterion on which to compare recalibration methods is the extent to which cost per case and CMI correspond at the hospital level. If high-weight DRGs are systematically underweighted and low-weight DRGs systematically overweighted, then compression will occur for CMIs calculated at the hospital level as well as for the DRGs. CMI compression can also be caused by problems in the classification system or by a correlation between the CMI and a tendency to provide more resources per case. Although we can't observe DRG compression, it is possible to test for CMI compression using a regression of each hospital's average cost per case on its CMI. In the absence of CMI compression, the coefficient on the CMI would be 1.

The first four columns of Table 8 present regressions of total Medicare cost per case on the CMI and other payment factors. (The last four columns of Table 8 will be discussed in the next section.) Because the hospital cost data we are using is from PPS5, it covers different calendar periods. Thus, to control for the effects of

Table 8
Regression of Ln (cost per case) and Ln (cost per case net of outlier payments) on
various case-mix indexes (CMIs) and payment factors: PPS5

Variable	Total Medicare cost per case				Medicare cost per case net of outlier payments			
	HSRV method	Standard method	Paid weights	Capital file CMI	HSRV method	Standard method	Paid weights	Capital file CMI
Ln (case-mix index)	1.083 (0.024)	1.020 (0.022)	1.060 (0.023)	1.116 (0.023)	1.035 (0.024)	0.974 (0.023)	1.018 (0.024)	1.078 (0.024)
Ln (wage index)	0.575 (0.024)	0.581 (0.023)	0.576 (0.024)	0.579 (0.023)	0.490 (0.024)	0.494 (0.024)	0.490 (0.024)	0.492 (0.024)
Resident-to-day ratio	0.290 (0.019)	0.273 (0.019)	0.283 (0.019)	0.243 (0.019)	0.254 (0.019)	0.247 (0.019)	0.256 (0.019)	0.215 (0.019)
Percent of low-income beneficiaries	0.167 (0.022)	0.165 (0.022)	0.178 (0.022)	0.165 (0.021)	0.122 (0.022)	0.119 (0.022)	0.131 (0.022)	0.120 (0.022)
Large urban	0.185 (0.011)	0.180 (0.011)	0.187 (0.011)	0.174 (0.011)	0.180 (0.011)	0.176 (0.011)	0.181 (0.011)	0.169 (0.011)
Other urban	0.117 (0.009)	0.112 (0.009)	0.116 (0.009)	0.103 (0.009)	0.110 (0.009)	0.106 (0.009)	0.109 (0.009)	0.096 (0.009)
Time dummy variable in years	0.093 (0.008)	0.093 (0.008)	0.090 (0.008)	0.106 (0.008)	0.130 (0.009)	0.131 (0.009)	0.128 (0.009)	0.144 (0.009)
Intercept	7.995 (0.009)	8.018 (0.008)	8.000 (0.009)	8.014 (0.008)	7.948 (0.009)	7.969 (0.009)	7.951 (0.009)	7.964 (0.009)
N	4,890	4,890	4,890	4,890	4,890	4,890	4,890	4,890
R-square	0.71	0.71	0.71	0.72	0.66	0.67	0.66	0.68

NOTES: Ln is the natural logarithm function. All regressions are case-weighted. Standard errors are shown in parentheses. HSRV is hospital-specific relative-value.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

differing time periods over which costs are measured, a variable giving the fraction of the year from the start of Federal FY 1988 until the start of the hospital's PPS5 year is included in the regressions. (Mean values for the variables in the regressions may be found in Carter and Rogowski, 1993.)

Under the HSRV method, the CMIs are more compressed than under the standard method. The coefficient on the HSRV CMI, 1.083, is significantly different from 1. However, the coefficient based on the standard weights, 1.020, is not significantly different from 1.

In FY 1988 there were substantial improvements in coding as a response to refinements in the grouper. The effects of this change on compression can be seen in Table 8. Our methodology calculates weights for the grouper in effect in each year. For actual payment purposes, HCFA calculates weights for each grouper based on the cases in an earlier year's file. The weights used for payment in FY 1988 were calculated based on HCFA's FY 1986 file, before the coding improvements occurred. These paid weights are compressed when applied to cases classified after the coding improvements occurred (with a coefficient of 1.060), whereas the weights calculated using the same methodology on the file with the improved coding are not compressed (with a coefficient of 1.020, not significantly different from 1). In order to test whether our sampling methodology affected this conclusion, we also created DRG weights based on the 1988 grouper for our sample of FY 1986 data. Again, we found more compression with the 1986 file weights than with the 1988 file weights.

The fourth column of Table 8 shows that the CMI on the capital file, which is based on all cases at the

hospital, has an even larger coefficient than the CMI calculated from the same weights for the sample cases in our file. We attribute this to the fact that random-measurement error tends to bias the coefficient toward zero when only a sample is used. Similarly, because of measurement error in the CMI when a sample of cases is used, the HSRV CMIs, if calculated on a full sample, are probably more compressed than the regression in Table 8 indicates. However, because the standard CMIs are less compressed than the HSRV CMIs when calculated on the identical sample of cases, the same result would probably hold true in a 100-percent sample. Thus, it is likely that the HSRV CMIs, if calculated on a full sample, would be more compressed than the CMIs currently in use.

We saw earlier that hospitals tend to charge substantially more than expected for low-weight DRGs and less than expected for high-weight DRGs. A priori, one would expect that this would lead to DRG compression in both standard and HSRV weights. Because high CMI hospitals specialize in high-weight cases, this would in turn lead to compression in both the standard and HSRV CMIs. In order to isolate the cause of compression in the HSRV weights and not in the standard weights, we need to consider the relationship between a hospital's CMI and its cost-to-charge ratio. These two variables are negatively correlated (correlation coefficient = $-.30$). As shown in Table 9, hospitals in the lowest 25 percentile of the CMI distribution have an average (case-weighted) ratio of costs to charges (RCC) of 0.728 but those in the highest CMI quartile have an RCC of only 0.623. This relationship means that, in the calculation of the standard weight, hospitals with high-weight DRGs have such high charges that they counterbalance their shift of

Table 9

Standardized charge and cost per unit DRG weight for HSRV and standard weight, by HSRV CMI minus standard CMI: PPS5

Hospital characteristic	Number of cases	Standardized charge		Standardized cost		RCC
		Standard CMI	HSRV CMI	Standard CMI	HSRV CMI	
All hospitals	1,689,900	\$7,203	\$7,198	\$4,656	\$4,650	0.661
HSRV CMI minus standard CMI, percent						
-3	20,250	7,758	7,991	4,234	4,360	0.575
-2	112,738	7,504	7,653	4,797	4,891	0.642
-1	363,103	7,614	7,687	4,646	4,690	0.623
0	483,250	7,551	7,546	4,693	4,689	0.625
1	524,767	6,981	6,919	4,673	4,631	0.683
2	165,999	5,955	5,846	4,524	4,441	0.752
3	19,793	5,201	5,058	4,248	4,131	0.845
CMI percentile						
0-25	389,244	6,546	6,455	4,602	4,537	0.728
25-75	907,847	7,343	7,329	4,686	4,676	0.647
76-100	392,809	7,528	7,631	4,640	4,703	0.623

NOTES: Charges and costs are case-weighted. DRG is diagnosis-related group. CMI is case-mix index. HSRV is hospital-specific relative-value. RCC is ratio of costs to charges. The standardization factors used in this table are derived from the first two regressions in Table 8.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

charges away from high-weight DRGs. Similarly the low-CMI hospitals that specialize in low-weight DRGs have charges that are lower than expected, but the weight for these DRGs is increased by charges that are higher than expected in other hospitals because of cross-subsidies. This resulted in a rough balance of CMI-adjusted standardized costs across CMI groups despite the disparity of charges (Table 9).

The RCC distribution across winning and losing hospitals changes the large difference in the distribution of standardized charges per unit of DRG weight into a much smaller difference in the distribution of standardized cost per unit of DRG weight. Table 9 shows that a hospital's gain in HSRV weight is strongly and positively related to its cost-to-charge ratio. There is only a small amount of variation in standardized cost per unit of DRG weight despite the large variation in charges. Thus there is only slightly more compression in the HSRV weights than in the standard weights.

Cost and payment

Another important criterion in comparing calibration methods is the extent to which the method improves the correlation between payments and costs at the hospital level. This criterion measures provider equity better than the correlation between CMI and costs. Whether it also measures provider incentives for efficiency and access better depends on the extent to which hospital administrators consider only DRG payments or also outlier payments as well when they make planning decisions. As explained in the Introduction, the correlation between payment and costs may be different than the correlation between the CMI and costs because of outlier payments and because payments for indirect medical education are in excess of costs. In addition, the payment factors that would be associated with the HSRV method would differ somewhat from those with the standard method.

The first four regressions in Table 8 examined the relationship between cost per case and each method's CMI while controlling for other payment variables, but not for outlier payments. Since outlier payments are concentrated in high-weight DRGs, it is possible that PPS payments to hospitals with large CMIs under the HSRV methodology might match hospital costs more accurately despite the apparent compression in HSRV weights. In order to find out, we ran regressions where the dependent variable is the log of the hospital's per case average of cost minus outlier payments. Because we enter the payment adjustment variables rather than constraining the coefficients to the actual payment factors, the equation shows how well each weight method could do at making payment proportional to costs rather than how well each method would do with current payment adjustments. As shown in the last four columns of Table 8, the effect of accounting for the part of costs that are paid through outlier payments is an apparent decompression of the HSRV CMIs. Once outlier payments are controlled for, the coefficient on the CMI for the HSRV weights drops to 1.035, insignificantly different from 1 and indeed almost as close to 1 as the coefficient on the standard CMI. (The distance is 0.035 versus -0.026.)

Using actual payment adjustments we reach a similar conclusion. The correlation between simulated payment and cost for PPS5 is quite similar for the two weight methods, with a Pearson correlation coefficient of 0.6860 for the HSRV weights compared with 0.6859 for the standard weights.

Table 10 shows additional detail about our simulated payment. It shows that the hospitals that perform very expensive cardiac surgery have higher margins than other hospitals under the standard weights and existing payment adjustments. Although their margins drop under HSRV, they remain greater than those of other hospitals.

Table 10

Cost per case and simulated payment per case, by weight method and whether hospital performs very expensive cardiac surgery

Type of hospital	Cost per case	Payment			
		HSRV weight		Standard weight	
		Per case	Margin	Per case	Margin
		Percent	Percent		
All hospitals	\$4,888	\$5,035	2.9	\$5,039	3.0
Hospitals performing very expensive cardiac surgery	5,944	6,189	4.0	6,248	4.9
Other hospitals	4,233	4,320	2.0	4,290	1.3

NOTES: Very expensive cardiac surgery is defined as diagnosis-related groups 103 to 108. HSRV is hospital-specific relative-value.

SOURCE: Carter, G.M., and Rogowski, J.A., RAND, Santa Monica, California, 1993.

Change in case-mix index

The final question we address is whether the DRG weight methodologies would provide different estimates of the CMI increase as a result of the improved coding that occurred in response to the 1987 and 1988 grouper refinements. If the improved coding were concentrated at specific hospitals with specific case-mix characteristics, then the HSRV methodology might cause a smaller increase in CMI.

We used our file of FY 1986 cases to determine relative weights for the FY 1988 grouper DRGs under both the HSRV and standard methodologies. Then we applied these relative weights to the cases found in our FY 1988 MEDPAR file. The resulting national CMI, therefore, measures the rate of increase in the CMI from FY 1986 and FY 1988.

The two methods measure almost identical rates of increase in the CMI. With the standard methodology, the increase was 5.975. With the HSRV methodology, the increase was 5.779. The difference in the rate of increase is only 3 percent.

Conclusions

We have provided evidence that two of the mechanisms that have been hypothesized to cause DRG compression were effective as late as the fifth year of PPS. The first mechanism is classification error. Coding errors on the FY 1986 file used to calculate case weights for FY 1988 payments led to measurable compression that disappeared when weights were calculated on later files that have elsewhere been shown to contain improved coding (Carter, Newhouse, and Relles, 1991). The second mechanism is the subsidization of higher weighted cases by lower weighted cases. The hospitals engaged in very expensive cardiac surgery were demonstrated to have charges that were higher than expected for typical cases, but not for their cardiac surgery cases. The vast majority of hospitals have charges that are lower than expected for cases in high-weight DRGs and charges that are higher than expected for cases in low-weight DRGs.

Although the weights used for payment in PPS5 exhibit CMI compression, weights calculated on the PPS5 files by the standard method do not show CMI

compression.⁴ The lack of compression is in contrast to analyses that showed that standard-charge weights calculated on FY 1984 data had compressed CMIs (Thorpe, Cretin, and Keeler, 1988). The correlation between high CMIs and high CMI-adjusted charges offset cross-subsidies within individual hospitals during the FY 1988 and FY 1989 timeframe. Because the cause of high charges was a high markup rather than high costs, the result was a lack of CMI compression of standard-method weights calibrated on files from this time period.

What does the lack of CMI compression say about the likelihood that DRG compression still exists? In addition to DRG compression, CMI compression could be caused by a correlation between the CMI and the likelihood that a hospital would receive cases that require greater than average resources for that DRG, or by a correlation between the CMI and a tendency to provide more resources to the same case. There is a general consensus in the literature that these correlations are non-negative. Thus, DRG compression should be less than that measured by CMI compression, and we expect that current weights are no longer compressed at the DRG level.

If there is a trend toward decreasing cost-to-charge ratios in hospitals with high CMIs, then sometime in the future the standard method will produce weights that overvalue higher weighted DRGs and undervalue lower weighted DRGs despite cross-subsidies within individual hospitals. The effect of additional DRG refinement on compression is somewhat more complicated. The *Federal Register* (1992) provides evidence that the CMI increased substantially from 1989 to 1991 as a result of improved coding in response to grouper changes, which introduced additional high-weight DRGs in 1991. Some of the cases belonging in the new high-weight DRGs were miscoded on earlier files. This probably reintroduced DRG compression into the weights paid for FY 1991 (by overvaluing DRGs from which the newly created high-weight DRGs received cases), but the improved coding probably returned the weights to their uncompressed state in subsequent years. If cases in the new high-weight DRGs were disproportional in hospitals with high CMIs, then

⁴PPS5 cases are from FY 1988 and FY 1989. Weights from these fiscal year files were used for payment in FY 1990 and 1991.

the FY 1991 DRG definitions and subsequent coding improvement would mean that decompression of the CMI has already occurred.

HSRV weights are more compressed than standard-method weights, a finding consistent with earlier research on FY 1984 data (Rogowski and Byrne, 1990). We believe that this compression arises primarily because of within-hospital cross-subsidies. Because the HSRV method standardizes on charges within individual hospitals, it is not affected by the distribution of adjusted charges across hospitals, which offsets compression in the standard method. Thus, the weights produce a measure of relative-resource consumption across DRGs that is more compressed than standard weights.

Despite the compression of HSRV weights, our regressions show that HSRV weights and standard weights provide equally good measures of hospital-level costs net of those expenses that are paid by outlier payments. Outlier payments provide additional payment in high-weight DRGs and thus compensate for the compression of HSRV weights. The current payment system includes both outlier payments and payment in excess of costs for indirect medical education. The result is that the hospitals performing the very expensive cardiac surgery had substantially higher PPS margins than other hospitals. These cardiac surgery hospitals include almost all of the hospitals that have lower case weight under HSRV than under standard weights; almost all other hospitals have lower PPS margins and would gain under HSRV weights. Despite the fact that the cardiac surgery cases would lose case weight under HSRV weights, the cardiac surgery hospitals would still have higher PPS margins than other hospitals if HSRV weights were used and these other payment policies remained in force.

At the DRG level, the largest difference between the two methods is that the weights for very expensive cardiac surgery DRGs are lower under the HSRV method than under the standard method. Thus, the HSRV method would lower incentives for hospitals to start or expand such programs. The hospitals that engage in this expensive surgery are doing substantially better than other hospitals given existing payment adjustments and outlier rules, so considerations of provider equity would argue that the HSRV weights are superior in this respect.

We cannot definitively determine the desirability of changing incentives for expensive cardiac surgery because we cannot observe case-specific costs and because we do not completely understand the relative importance of net revenue and other factors in hospital decisionmaking. First, let us continue the arguments in the Introduction which assume that net revenue is the primary element in the hospital's objective function.⁵ Then, if, as appears unlikely to us, charges reflect each hospital's relative costs per DRG, then the HSRV incentives for efficiency and access would be superior. If one believes that our estimate of CMI compression is

a good estimate of DRG compression (because coding problems and DRG definition problems were largely solved) and hospital administrators ignored outlier payments, then the standard-weight incentives are slightly superior. If, instead, hospital administrators assess expected outlier payments for these cardiac surgery cases, then either method provides equally appropriate incentives.

In theory, the optimal policy depends on all well defined factors that affect hospital behavior, not just revenue. Hospital objective functions may depend on prestige derived from high-weight, high-technology DRGs. This is consistent with the subsidization pattern of charges that we observed in this study. The current DRG weighting system, because of the accident of the correlation between high-weight DRGs and low RCCs, compensates on average for the distortion caused by the subsidies. Because the net revenue expectations are the same for high- and low-weight Medicare patients, hospitals will prefer to treat high-weight Medicare patients rather than low-weight Medicare patients. Thus, the policy that avoids incentives to discriminate against low-weight patients would increase the DRG weight for low-weight cases toward the relative charges for the DRGs, which are presumably equilibrium prices that clear the market and that balance the hospital's double objectives of obtaining revenue and the enhanced prestige that comes from providing high technology services. This is another argument in favor of the HSRV weights.

Acknowledgments

We thank Daniel Byrne for his help in constructing the weights used in this article. We also thank all three anonymous reviewers for helpful comments on an earlier version of this article.

References

- Carter, G.M., and Farley, D.O.: A longitudinal comparison of charge-based weights with cost-based weights. *Health Care Financing Review* 13(3):53-63. HCFA Pub. No. 03329. Office of Research and Demonstrations, Health Care Financing Administration, Washington. U.S. Government Printing Office, Spring 1992.
- Carter, G.M., and Rogowski, J.A.: *The Hospital Specific Relative Value Method as an Alternative for DRG Recalibration*. Report No. MR-156-HCFA. RAND. Santa Monica, CA. 1993.
- Carter, G.M., Newhouse, J.P., and Relles, D.A.: *Has DRG Creek Crept Up? Decomposing the Case Mix Index Change between 1987 and 1988*. RAND Corporation Report R-4098-HCFA/ProPAC. Santa Monica, CA. 1991.
- Cotterill, P., Bobula, J., and Connerton, R.: Comparisons of alternative relative weights for diagnosis-related groups. *Health Care Financing Review* 7(3):37-51. HCFA Pub. No. 03222. Office of Research and Demonstrations. Health Care Financing Administration, Washington. U.S. Government Printing Office, Spring 1986.

⁵Another implicit assumption is that all appropriate care provided in an efficient hospital is of social value at least equal to its costs.

Federal Register: Prospective Payment System for Inpatient Capital Related Costs. Final Rule. Vol. 56, No. 169, 43358-43524. Office of the Federal Register, National Archives and Records Administration. Washington. U.S. Government Printing Office, Aug. 30, 1991.

Federal Register: Inpatient Hospital Prospective Payment System and 1993 FY Rules. Vol. 57, No. 108, 23618-23839. Office of the Federal Register, National Archives and Records Administration. Washington. U.S. Government Printing Office, June 4, 1992.

Lave, J., Pettengill, J., Schmid, L., and Vertrees, J.: Measurement Issues in the Development of a Hospital Case Mix Index for Medicare. *American Statistical Association Proceedings of the Social Statistics Section*, pp. 57-62, 1981.

Lave, J.R.: Is Compression Occurring in DRG Prices? *Inquiry* Vol. 22, pp. 142-147, Summer 1985.

3M Health Information Systems: *Diagnosis-Related Groups Definitions Manual, Version 10*. Pub. No. 92-054. Wallingford, CN. 1992.

Pettengill, J., and Vertrees, J.: Reliability and validity in hospital case-mix measurement. *Health Care Financing Review* 4(2):101-128. HCFA Pub. No. 03149. Office of Research and Demonstrations, Health Care Financing Administration. Washington. U.S. Government Printing Office, Winter 1982.

Rogowski, J.R., and Byrne, D.J.: Comparison of alternative weight recalibration methods for diagnosis-related groups. *Health Care Financing Review* 12(2):87-101. HCFA Pub. No. 03316. Office of Research and Demonstrations. Health Care Financing Administration. Washington. U.S. Government Printing Office, Winter 1990.

Thorpe, K.E., Cretin, S., and Keeler, E.B.: Are the diagnosis-related group case weights compressed? *Health Care Financing Review* 10(2):37-46. HCFA Pub. No. 03276. Office of Research and Demonstrations, Health Care Financing Administration. U.S. Government Printing Office, Winter 1988.