Using prior utilization to determine payments for Medicare enrollees in health maintenance organizations

by James Beebe, James Lubitz, and Paul Eggers

The Tax Equity And Fiscal Responsibility Act of 1982 is expected to make it more attractive for health maintenance organizations (HMO's) to participate in the Medicare program on an at-risk basis. Currently, payments to at-risk HMO's are based on a formula known as the adjusted average per capita cost (AAPCC). This article describes the current formula and discusses a modification, based on prior use of Medicare services, that endeavors to more accurately predict risk. Using statistical simulations, formulas incorporating prior use performed better for some types of biased groups than a formula similar to the one currently employed. Major concerns involve the ability to "game the system." The prior-use model is now being tested in an HMO demonstration. This article also outlines the limitations of a prior-use model and areas for future research.

Introduction

Health maintenance organizations (HMO's) have generally been considered to be more efficient systems for providing health care than the traditional fee-for-service system. Many studies have found that health care costs and hospital use in HMO's, particularly in prepaid group practices (PGP), are generally lower than in the fee-for-service sector (Manning et al., 1984; Wolinsky, 1980; Luft, 1978; and Roemer and Shonick, 1973). The studies often point to lower hospital use. An example is the recently published study from the RAND Health Insurance Experiment (Manning et al., 1984). The authors found that an HMO achieved significant savings from lower hospitalization rates when patients were randomly assigned to an HMO or to a fee-for-service health care insurance package.

It has also been suggested that one reason for the lower costs may be biased selection of healthier persons into HMO's (Luft, 1978 and 1982). Over the past few years, a number of studies have appeared that show biased selection of lower-than-average users of health care into HMO's, both by Medicare enrollees (Eggers, 1980; and Eggers and Prihoda, 1982), and by persons under 65 years of age (Jackson-Beeck and Kleinman, 1983; Roghman, Sorensen, and Wells, 1980; Arthur D. Little, 1983). These studies examined the health care use of HMO enrollees before they joined an HMO, and found that they had used less services on the average than persons not joining. This suggests that the HMO enrollees may have been in better-than-average health and would have been lower users even if they were not in HMO's.

The results of these studies should be qualified in three ways. First, all but one study reported on cases in which an HMO option was offered for the first time. The findings might be different if the studies were repeated after an HMO option had been available for some years. Second, the finding of lower pre-enrollment use only applied to prepaid group practices, not to the individual practice association

(IPA) type of HMO's in the studies. Persons may be attracted to IPA's because they can maintain their present ties to their physicians, and such persons may be in poorer than average health (Luft, 1981). Third, in all the studies, only one fee-for-service alternative was offered. A study of the Federal Employees Health Benefits Program, where a range of fee-for-service and HMO plans are available, did not find selection of low users into HMO's (Schuttinga, Fallik, and Steinwald, 1984). Instead, lower users were attracted into the low-premium, high-cost-sharing, fee-for-service plans. The authors emphasize that, "...the extent of adverse or beneficial selection into HMO's depends on the price and the comprehensiveness of benefits of each available fee-for-service option."

The problem of potential overpayments to HMO's has been made more critical by the HMO provisions of the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA). These provisions make it more attractive for HMO's to participate in Medicare on an at-risk basis and are expected to increase the number of at-risk HMO's in Medicare. TEFRA authorizes prospective reimbursement to HMO's under risk-sharing contracts at a rate equal to 95 percent of the adjusted average per capita cost (AAPCC). TEFRA defines the AAPCC to be the estimated average per capita amount that would be payable if Medicare services for HMO members were furnished in the local fee-for-service market.

TEFRA requires that the AAPCC formula take into account those factors that are likely to be associated with differences in health care use, such as age and sex, and thereby adjust for differences in expected health service use between HMO enrollees and other Medicare beneficiaries living in the HMO's service area. The current AAPCC formula uses age, sex, welfare status, and institutional status as adjustment factors.[1]

As previously noted, studies of Medicare enrollees joining HMO demonstrations raise the issue of biased selection. After adjusting for age, sex, welfare status, and institutional status, Eggers (1980) and

Eggers and Prihoda (1982) found in three out of four HMO's examined that Medicare enrollees still had lower use in their preenrollment years. In the fourth, which was structured much like an IPA, Medicare enrollees had about average use. These findings imply that a reimbursement formula that does not take health status into account may overpay HMO's when they experience a favorable selection of low users or, conversely, may underpay when HMO's experience an unfavorable selection of high users. As a result of the Eggers and Prihoda studies, the Congressional Budget Office (1982) estimated that under the present AAPCC, increased Medicare enrollment in HMO's would increase Medicare costs in the short run.[2]

Our purpose was to test, in a simulation, whether or not a regression model that uses available prior utilization and demographic predictor variables could produce more accurate predictions of future Medicare reimbursement than the current AAPCC formula. Although many variables, such as self-reported health status or functional status, are associated with health care use, we restricted our examination to variables currently available from administrative records. Thus, variables that might theoretically be more attractive but that would require added data gathering were not investigated.

The utility of prior use as a predictor of future use is suggested by studies showing that a person's past use of health care is correlated positively with subsequent use (Roos and Shapiro, 1981; Densen, Shapiro, and Einhorn, 1959; Eggers, 1980; Eggers and Prihoda, 1982; Anderson and Knickman, 1984b; and McCall and Wai, 1983). This is certainly plausible. Many persons have chronic conditions, such as cancer, stroke, or mental disorders, that require repeated treatments, often over a long time. In addition, some persons may have more of a tendency to seek medical care than others.

Incorporating a prior-use variable into reimbursement formulas for HMO's in Medicare has, in fact, been suggested by a number of authors (Trapnell, McKusick, and Genuardi, 1982; Anderson and Knickman, 1984a and b; and Thomas et al., 1983). These authors believe that a prior-use adjustment would substantially remove the possibility of financial losses to Medicare or the HMO in case of biased selection, thus increasing HMO participation. Adjustment for prior use could also remove the current disincentive for HMO's to enroll sicker than average people, since premiums would better reflect expected health care use.

A prior-use adjustment might also be applicable if a voucher system proposal is adopted. Such a system would restructure Medicare by giving vouchers to Medicare enrollees to purchase private health plans (or to remain in Medicare). Although a voucher strategy may promote competition among health plans, it can have the same undesirable effect found in the Federal Employees Health Benefits Program of encouraging enrollment of healthy persons into low-premium, high-cost-sharing plans, leaving sicker persons in high-premium, comprehensive-benefit plans and raising overall program costs (Luft, 1982; and Anderson and Knickman, 1984a). Adjusting the value of vouchers to reflect health status, perhaps through a prior-use adjustment, could reduce the undue financial rewards to plans that attract healthy persons and reduce the undue losses to other plans that attract sicker ones.

For our study, we first developed an AAPCC model incorporating prior-use variables in addition to most of the other factors used in the present AAPCC formula. Then, we tested this prior-use AAPCC model by comparing its predictive accuracy for groups of enrollees with various kinds of statistically simulated biased selection against the accuracy of an AAPCC formula similar to the current one. In this article, we also discuss the limitations of a prior-use AAPCC formula as well as areas for future research.
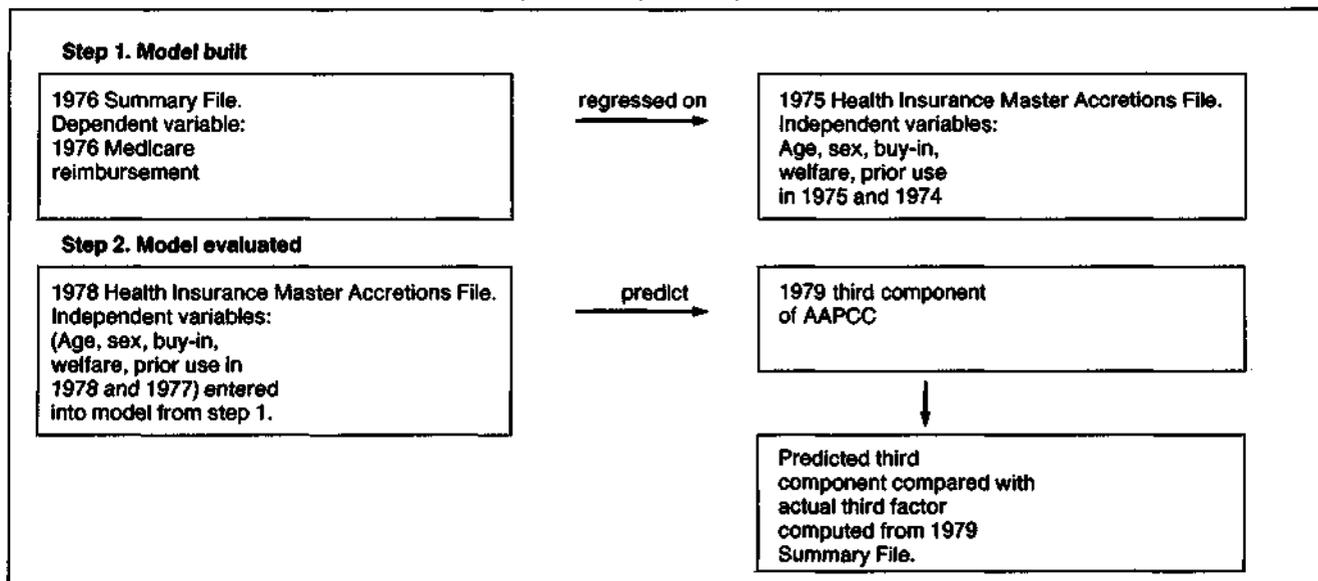
Methods

Study design

Before detailing the data sources and methods of our study, it is helpful to review the overall study design. The purpose of the study was to test how the inclusion of prior-use variables might improve the accuracy of the AAPCC formula. The actual constraints of data availability that would be faced if an AAPCC formula with prior-use variables were really to be employed to pay HMO's were duplicated as far as possible. Thus, we used data files that could actually be employed to implement an AAPCC formula with a prior-use adjustment, rather than data files that might, in theory, yield a better model.

In the first step of our study, models were built to predict an enrollee's 1976 total Medicare reimbursement (Figure 1). The models included ones using variables similar to those in the present AAPCC formula, as well as ones incorporating variables on prior hospital and physician use in 1975 and 1974. Then, the models were evaluated, using demographic data and 1978 and 1977 prior-use data for a variety of statistically simulated groups of enrollees, to predict the 1979 AAPCC third component for each group. The groups were statistically biased on a number of demographic and utilization variables. As explained in the following section, the third component indicated the ratio of the predicted Medicare reimbursement for a group of HMO enrollees to the reimbursement for the other

[1]Welfare status is measured by whether the Medicare enrollee is also covered by Medicaid, and institutional status reflects whether the enrollee is living in an institution (e.g., a nursing home). For a complete explanation of the current AAPCC formula, see Kunkel and Powell, 1981.
[2]The estimates assume that a biased selection effect would eventually dissipate over a period of 4 to 5 years. Two studies (Trapnell, McKusick, and Genuardi, 1982, and Welch, 1984) show that regression toward the mean is to be expected given biased selection on prior use. Welch contends that the actual degree of bias would be about one-half that observed by Eggers and Prihoda.

Figure 1
Study design: Building and testing models incorporating prior-use variables in an adjusted average per capita cost (AAPCC) formula.

Step 1. Model built

| 1976 Summary File.
Dependent variable:
1976 Medicare
reimbursement | regressed on → | 1975 Health Insurance Master Accretions File.
Independent variables:
Age, sex, buy-in,
welfare, prior use
in 1975 and 1974 |

Step 2. Model evaluated

| 1978 Health Insurance Master Accretions File.
Independent variables:
(Age, sex, buy-in,
welfare, prior use in
1978 and 1977) entered
into model from step 1. | predict → | 1979 third component
of AAPCC |

| | | Predicted third
component compared with
actual third factor
computed from 1979
Summary File. |

Medicare beneficiaries in the HMO's service area. Each model was then evaluated by comparing the predicted third component with the actual third component calculated from 1979 reimbursement data.

Data sources

Our study was possible because of the existence of the Health Insurance Master Accretions (HIMA) File. The HIMA File was the source of all independent variables in our study. The purpose of the File is to record every Medicare enrollee's use of benefits so that accurate determinations may be made about Part A benefit periods,[3] hospital coinsurance and lifetime reserve days, and exhaustion of hospital benefits. The File tracks use of skilled nursing facility days, home health agency visits, and whether or not the Part B deductible has been met. It also contains demographic data for all enrollees.

The information on the File is available for all enrollees and is updated daily. Thus, it would be possible to obtain prior-use data from the File for inclusion in the AAPCC calculation for a group of enrollees about to join an HMO. The currency and completeness of the HIMA File make it the only Medicare File that could be used at present in a practical situation as a source of prior-use information. However, the HIMA File does not contain utilization data for persons who had not been previously enrolled in Medicare or for Medicare beneficiaries currently enrolled in HMO's. A

different approach would be needed for these two groups.

There are some additional limitations to the HIMA File. Although it is updated every day, there are time lags between a hospital discharge and when the hospital bill reaches the Health Care Financing Administration (HCFA) for posting to the enrollee's HIMA record. This time lag occurs because of transit and processing time at the hospital, the fiscal intermediary, and HCFA's central office. The average time between discharge and posting in the HIMA File is 61 days. Based on the distribution of these posting time lags, we estimated that 80 percent of a previous year's hospital use is available at any given time.

The HIMA File that was used to build models for this study had a September 1975 cut-off date; thus, hospital-use variables for the 12 months from October 1974 to September 1975 reflect only 80 percent of hospital use in that period. For this study, we used a 0.1-percent sample of the HIMA File.

The dependent variable for our study, Medicare reimbursement, was not available from the HIMA File. This information was obtained from the Person Summary File, one of several files in the Medicare Statistical System. We chose the Person Summary File for two reasons. First, it was the most current source of reimbursement by individual beneficiary. When we began our study, 1979 was the latest available year. Second, it was a 5-percent probability sample of Medicare enrollees, of which the 0.1-percent HIMA File sample was a subset. Thus, data from each beneficiary record on the HIMA File sample were linked to data for the same beneficiary from the Person Summary File. The resulting 0.1-percent sample of merged records formed the data base for our study. The merged 1975 HIMA and 1976 Person Summary records, containing data on 1976 Medicare reimbursement, were the data base

---

[3]The Medicare program defines a benefit period as the time between the first day an enrollee was an inpatient of a hospital or skilled nursing facility and 60 days after the last day the enrollee was an inpatient of such facilities. Medicare pays for 90 hospital days in a benefit period.

used to develop the models. The merged 1978 HIMA and 1979 Person Summary records, containing data on 1979 Medicare reimbursement, comprised the data base to test the models.

The current AAPCC formula

The AAPCC formula currently in use consists of the product of three major components:[4]
1. The U. S. per capita Medicare cost as projected to the current year.
2. An adjustment based on the historical relationship between Medicare per capita reimbursements in the local area which the HMO serves and national Medicare costs.
3. An adjustment for the differences between persons who choose to enroll in an HMO and the population at large from which HMO enrollees are drawn.

At present, this third component adjusts for four factors known to be associated with health care use: age, sex, welfare status, and institutional status. It takes the form of a ratio whose numerator reflects the average characteristics of the HMO enrollees and whose denominator reflects the average characteristics of the population from which the HMO enrollees were drawn. Thus, a third component with a value of 1.00 indicates that the HMO has drawn a group of enrollees equivalent to the fee-for-service population in its service area with respect to the four AAPCC factors. A third component greater than 1.00 indicates the HMO enrolled a group expected to have higher-than-average levels of health care use. A third component less than 1.00 indicates a group of enrollees expected to have lower levels of health care use than the general population.

Building the models

Our goal was to incorporate prior use as an additional factor in calculating the third component. To accomplish this, regression models were developed to predict an enrollee's total Medicare reimbursement in a subsequent year using certain predictor (independent) variables in previous years. The predictor variables used in the models were age, sex, Part B buy-in status, and a variety of prior-utilization variables. Institutional status, a factor in the current AAPCC formula, was excluded because it was available only from special surveys. There is no reason to believe that the absence of an institutional factor had an important impact on the findings, because only about 5 percent of the Medicare population are institutionalized at any time. The variables used in developing the models are shown in Table 1. As previously noted, all predictor variables were derived from the HIMA File, and the dependent

---

[4]See the Technical Note for a detailed description of the AAPCC.

variable, Medicare reimbursement per person, came from the Person Summary File.

Although most variable definitions in Table 1 are self explanatory, a few may require further elaboration. Part B buy-in (variable B.3) indicates whether the Medicare enrollee is also a Medicaid eligible for whom a State has purchased Medicare Part B coverage. (Louisiana, Oregon, and Wyoming do not have buy-in programs.) This variable was used as a proxy for welfare status in our models. On the basis of the current AAPCC formula and a study by McMillan et al. (1983), we would expect Medicare reimbursements to be higher for such persons. Variables B.6 and B.7 indicate whether the enrollee met the yearly Part B (supplementary medical insurance) deductible in 1975 and 1974, respectively. In both years, the deductible was $60.00.

We built models using the linked HIMA and Summary Files. A record for each enrollee in the sample was created by linking a person's record from the September 1975 HIMA File, which provided the predictor variables, with their Medicare reimbursement in 1976 as shown on the 1976 Summary record. Persons who were not alive on January 1, 1976, were eliminated from the study group used to build the model because, of course, all HMO enrollees would be alive when they joined. Also eliminated were persons not eligible for both Part A (hospital insurance) and Part B (supplementary medical insurance) of Medicare. This was done because all Medicare HMO enrollees at present are required to have both Parts A and B. We eliminated disabled enrollees under 65 years of age from the study group to simplify the analyses. Applying these selection criteria resulted in a sample size of 20,773 enrollees.

We simulated the actual limitations of data availability that would be faced if we really had to use HIMA data to set HMO reimbursement. The main limitation is the 15- to 24-month lag in obtaining the Summary File, the source of the dependent variable, Medicare reimbursement, used in developing the model. If a prior-use model were actually employed to set reimbursement, there would be at least a 2-year difference between the data used to develop the model and the year for which HMO reimbursement was being set. Thus, there is a 24-month period between the years used to build the model (1975 and 1976) and the years used to test it (1978 and 1979). In addition, we built in a 3-month lag between the HIMA File, the source of data for the independent variables, and the year for which use is predicted. This was done to duplicate the time that would be necessary in practice for enrollees to choose an HMO and for their data on the HIMA File to be passed through the predictive model to set a reimbursement amount. Thus, we use data from the HIMA File with a September 1978 cut-off to predict calendar year 1979 reimbursement.

Table 1
Variables used in regression models

Variable	Definition	Mean	Standard deviation
A. **Dependent variable**[1]			
Reimbursement per enrollee in 1976	Total Medicare reimbursement per enrollee in 1976	$649	$1,891
B. **Independent variables**[2]			
Demographic variables:			
1. Age in 1975	Age as of 1975 minus 65	8.9 years	6.7 years
2. Sex	Dummy variable, 0 if male	0.6 (60 percent female)	0.49
3. Part B buy-in status in 1975	Dummy variable, 1 if State bought Part B Medicare coverage for enrollee, 0 if not	.102 (10.2 percent buy-in)	.30
Prior-use variables:			
4. Hospital use in last year	Dummy variable, 1 if had hospital admission from October 1974 to September 1975, 0 if not	0.15 (14.9 percent of enrollees used hospital)	.36
5. Hospital days in last 2 years	Number of hospital days used from October 1973 to September 1975	4.26 days	11.16 days
6. Part B deductible met, 1975	Dummy variable, 1 if met Part B deductible in 1975, 0 if not	0.32 (32 percent of enrollees met Part B deductible in 1975)[3]	0.47
7. Part B deductible met, 1974	Dummy variable, 1 if met Part B deductible in 1974, 0 if not	0.43 (43 percent of enrollees met Part B deductible in 1974)	0.49

[1]Source—1976 Person Summary File.
[2]Source—1975 Health Insurance Master Accretions File.
[3]The percent meeting the deductible is low because only three-quarters of 1975 data were used.

The models selected for detailed analysis are shown in Table 2. The first model, "Demographic," tests the predictive power of three of the four current AAPCC variables (age, sex, and buy-in) with no prior-use information. The second model, "Hospital Use," tests the simplest prior-use model by using as a prior-use variable whether or not the enrollee was hospitalized in the last year. This variable could probably be accurately gathered from a survey of enrollees with no HIMA record, i.e., those just joining Medicare. The last model, "Hospital Days, Part B," uses the number of hospital days used in the past 2 years and whether the Part B deductible was met in the past 2 years. Ordinary least squares methods were used to estimate the coefficients.[5]

Table 2 also shows the percent of variance explained (R^2) for each of the models. None of the R^2 values exceeds 5 percent, indicating that the models could be expected to give very poor predictions for the annual reimbursement of individuals. Such low R^2 values are typical of models predicting future health care use of individuals. However, as we shall

see, these R^2 values do not reflect the expected accuracy of predictions of group averages. Because the random errors of predictions for individuals tend to cancel out for large groups, predictions for groups can be much more accurate.

Testing the models

The file used to test the models was constructed like the one used for model development. For each enrollee, it linked their record in the 1978 0.1-percent sample HIMA File with their Medicare reimbursement from the 1979 Summary File. The sample size was 22,513.

We evaluated the model both at the national level and then for smaller groups of enrollees, such as States, to simulate how the model might actually be used. We also simulated biased selection, by testing the models with groups of enrollees with higher- and lower-than-average proportions of high or low users, and higher- and lower-than-average proportions of older enrollees and buy-in enrollees.

The biased groups were selected on the basis of 1978 data. Figure 2 shows the groups used to test the models. The States and groups of States were chosen to provide a variety of sizes and geographic areas. The biased groups were selected to test what we thought might be likely types of biased selection. The biasing procedure was quite simple. For example, to get a group with a low percentage meeting the

---

[5]Models using a number of other variables were tested but eliminated because the additional variables contributed an insignificant amount to the explanatory power of the model finally selected. Some variables tested but not included in the final models were measures of skilled nursing facility use, the length of time since the last bill, and number of hospital days in the last year (as opposed to the last 2 years). Models using logarithmic transformations and interaction terms were also tested but found to be no improvement, for prediction purposes, over the linear models.

deductible, we selected at random 50 percent of the persons meeting the deductible and combined them with 100 percent of persons not meeting the deductible. The group with a high percentage meeting the deductible was formed by combining all persons meeting the deductible with a randomly selected 50 percent of persons not meeting the deductible. Other groups were formed in a similar manner.

Table 2
Regression coefficients of models relating 1976 Medicare reimbursement per enrollee to 1975 and 1974 characteristics

Variable	Type of model		
	Demographic	Hospital Use	Hospital Days, Part B
Age	16.7	14.7	9.66
	(1.97)	(1.96)	(1.96)
Sex	-123	-116	-137
	(2.26)	(26.4)	(26.3)
Buy-in	251	229	153
	(26.8)	(43.6)	(43.3)
Hospital use in last year	—	676	—
		(36.6)	
Hospital days in last 2 years	—	—	20.7
			(1.24)
Part B deductible, 1974	—	—	191
			(29.1)
Part B deductible, 1975	—	—	341
			(30.4)
Constant	547	463	349
	Percent		
R^2 (corrected)	0.6	2.2	4.3

NOTES: The coefficients of all variables of all models are statistically significant at the p<.05 level. The sample size equals 20,773 Medicare enrollees 65 years of age or over. Standard errors are shown in parentheses.

Although these procedures resulted in groups with biased average reimbursement, the groups are not necessarily representative of those who actually join an HMO. HMO enrollees may differ from other persons in ways that affect their use of health care but are not readily measurable.

Reimbursement for 1979 was predicted using the models developed in Step 1 for each biased group and for the whole group (i.e., State or the United States) from which the biased group was selected. These values were used to form a ratio similar to the third component of the AAPCC (as described earlier and in the Technical Note). The numerator of the ratio was the predicted 1979 reimbursement for a biased group and the denominator was the predicted

1979 reimbursement for the group from which the biased group was selected:

$$\frac{\text{Predicted 1979 reimbursement for biased group}}{\text{Predicted 1979 reimbursement for total group}}$$

These third components derived from the models were then compared with the ratio of actual 1979 reimbursement for each biased group to the actual reimbursement for the whole group. If the two ratios were close, the regression model was considered to be a good predictor. Finally, third components were calculated using three (age, sex, and buy-in) of the four underwriting factors presently used to reimburse HMO's (referred to as the "Underwriting" model). Since our sample contained no information on institutionalization, the fourth underwriting factor, the factors were modified by collapsing the institutional factor into the age, sex, and welfare factors.[6] Thus, the Underwriting model is similar to the Demographic model in that they use the same factors. However, they differ in form in that the Underwriting AAPCC model uses the factors in 60 discrete cells, whereas the Demographic model uses a regression equation.

Findings

Each model was evaluated by comparing the third component it predicted for a variety of biased groups with the third component for the same groups based on actual reimbursement data. The closer the predicted third component was to the actual third component, the better the model performed. Table 3 shows the third components of the AAPCC for biased groups selected from the full national sample. Column 1 contains the number of persons out of the full 1978-1979 sample of 22,513 persons in each biased group. Column 2 contains the actual third component (reimbursement for biased group divided by reimbursement for full sample) for each biased group. Columns 3, 5, 7, and 9 show the third components calculated from the Underwriting model and the three regression models, respectively. Columns 4, 6, 8, and 10 contain the ratio of the third components calculated from the models to the actual third component of Column 2. The closer these ratios are to 1, the better the model performs. A ratio less than 1 indicates that the model's prediction is too low. A ratio greater than 1 indicates that the model's prediction is too high.

The first six rows of Table 3 show results for groups biased on prior use. The actual third component ranges from .87 for the low prior-reimbursement group to 1.15 for the group with a high percentage meeting the Part B deductible. The Underwriting and Demographic models predict third components near 1. The ratio of these predictions to actual third components ranges from 1.14 or 14 percent too high, to .88, or 12 percent too low. The

---

[6]Both the current underwriting factors and the modified underwriting factors are available upon request. The authors have also derived underwriting factors from one of the regression equations.

Table 3

AAPCC[1] third component based on selected models and ratio to actual third components, by type of biased group: United States summary, 1979

Type of biased group	Number of persons in group (1)	1979 actual third component (2)	Underwriting		Demographic		Hospital Use		Hospital Days, Part B—2 years	
			Third component (3)	Ratio to actual (4)	Third component (5)	Ratio to actual (6)	Third component (7)	Ratio to actual (8)	Third component (9)	Ratio to actual (10)
Biased on prior use										
Part B										
Low	17,889	.88	.99	1.13	.99	1.13	.95	1.08	.87	1.00
High	15,859	1.15	1.01	.88	1.01	.88	1.06	.92	1.14	.99
Hospital use										
Low	20,747	.92	1.00	1.09	1.00	1.09	.92	1.00	.94	1.02
High	13,029	1.12	1.01	.90	1.01	.90	1.12	1.00	1.10	.98
Prior reimbursement										
Low	17,041	.87	.99	1.14	.99	1.14	.96	1.10	.91	1.05
High	16,772	1.14	1.01	.89	1.01	.89	1.04	.91	1.09	.96
Biased on demographics										
Age										
Low	17,881	.93	.93	1.00	.95	1.02	.95	1.02	.96	1.03
High	15,865	1.07	1.08	1.01	1.05	.98	1.05	.98	1.05	.98
Buy-in										
Low	21,481	.98	.97	.99	.98	1.00	.98	1.00	.98	1.00
High	12,285	1.04	1.06	1.02	1.04	1.00	1.04	1.00	1.03	.99

[1] Adjusted average per capita cost.

Figure 2

Groups used to test how the prior-use models would perform

1. Geographic areas

   a. New York
   b. California
   c. Pennsylvania
   d. Texas and Oklahoma
   e. Illinois and Indiana
   f. Michigan and Ohio
   g. Georgia
   h. Florida

2. Biased groups (selected on the basis of 1978 data)

   a. Part B, Percent meeting Part B deductible in a year (41 percent average)
      Low percentage (26 percent) meeting deductible
      High percentage (58 percent) meeting deductible

   b. Hospital use, Percent using hospital in a year (16 percent average)
      Low percentage (9 percent) using hospital
      High percentage (27 percent) using hospital

   c. Prior reimbursement, Medicare reimbursement (49 percent had some reimbursement in a year)
      Low percentage (32 percent) with reimbursement
      High percentage (66 percent) with reimbursement

   d. Age, Percent 75 years of age or over (41 percent average)
      Low percentage (26 percent) 75 years of age or over
      High percentage (58 percent) 75 years of age or over

   e. Buy-in (9 percent average)
      Low percentage (5 percent) with buy-in
      High percentage (17 percent) with buy-in

smallest ratio for these two models is 1.09 (a 9-percent error) for the low-hospital-use group.

The Hospital Use model predicts third components identical to the actual components for the two groups biased on hospital use. However, it does less well for the groups biased on Part B deductible and prior reimbursement. For these groups, the Hospital Use model has prediction errors ranging from 8 percent to 10 percent.

The Hospital Days, Part B model does better, in general, than any of the other three models. The prediction errors are 0 and 1 percent for the low- and high-Part B groups and 2 percent for both hospital-use groups. The largest prediction errors for this model are 5 percent and 4 percent for the low- and high-prior-reimbursement groups.

The last four rows of Table 3 show results for groups biased on the AAPCC factors of age and buy-in status. For these groups, we find that all four models predict about equally well and all predict quite close to the actual component. Many of the predictions are exactly equal to the actual component to two significant digits. The largest error is 3 percent for the Hospital Days, Part B model in predicting the low-age group.

We also calculated third component predictions for each of eight States or pairs of States for each biased group discussed in Figure 2. (The results for the Underwriting model are not displayed since our analysis found that it predicted nearly exactly the same as the Demographic model.) As a measure of

predictive accuracy, we used the average absolute relative error. For example, if the actual third component was .86 and the model predicted .94, the prediction error is the absolute value of 1-.94/.86 = .09. These errors were then averaged across the eight State groups. Table 4 shows these average errors.

For the six groups biased on prior use, it is apparent that the prior use models do better on the average than the Demographic model. For every biased group, the Hospital Use and the Hospital Days, Part B models had lower average errors than the Demographic model. For the Demographic model, the average error ranged from a low of 8 percent for the low-hospital-use group to a high of 16 percent for the low-prior-reimbursement group. The range for the Hospital Use model is 3 percent to 11 percent, and the range for the Hospital Days, Part B model is from 3 percent to 6 percent. By this measure, the latter model is from two to three times better than the Demographic model for every group.

Looking at individual States and groups of States (not shown) we found considerable variation in how well the models predict for the six groups biased on prior use, although the prior-use models predict better in most cases than the Demographic model. However, even the best model, Hospital Days, Part B, had prediction errors as high as 13 percent in some cases.

For groups biased on prior-use, the results of comparing the prior-use models with the

Demographic model for the 48 (eight State groups times six biased groups) State-group cells are summarized in Table 5. The Hospital Use model predicts better than the Demographic model 44 times out of the 48. The Demographic model does better three times and there is one tie. The Hospital Days, Part B model predicts better than the Demographic 42 times out of 48. The Demographic model does better three times and there are three ties. Finally, the Hospital Days, Part B model does better than the Hospital Use model 28 times out of 48. The Hospital Use model does better 17 times out of 48 and there are three ties.

For the four groups biased on age and buy-in status, we find much the same results for the States as was found nationally—none of the models stand out as a better predictor. The average error for all models is about 5 percent. Table 6 summarizes the predictions for the 32 State-group cells and further reinforces the impression of similarity in the models.

An observation of some interest is that predicted components have less variation from State to State than the actual components. The maximum range of actual third components among States for the biased group (high buy-in) with the greatest variation was .21 (not shown in the table). The maximum range of predicted third components was .05 for the Demographic model, .05 for the Hospital Use model, and .08 for the Hospital Days, Part B model.

One possible reason for the considerable range in the actual components might be random variation that can occur in the selection of relatively small subsets. If this is the primary cause of the variation, then stability in the predictions may be an asset so long as the predictions approximate the average of

the actual components that would be found from repeated sampling of small subsets. However, a second reason for the difference in range between the actual and predicted could be the interaction between local use patterns and the variables in the equations or the exclusion from the models of variables which influence the actual components. If this is the case, the difference reflects a deficiency in the models.

Discussion

The results of this simulation study provide a number of insights into the consistency of use over time and the efficacy of incorporating prior use into an operational AAPCC-like mechanism. The first important finding is that, in statistical simulations, prior use improves the ability to predict reimbursement in a subsequent time period. In developing a model of reimbursement, the addition of one or more measures of prior use to a model containing only demographic variables more than doubled the explanatory power (as measured by R^2) of the regression equations. Of the variables tested, prior use is the single best predictor of subsequent Medicare reimbursement.

Table 5
Number of times each model predicts better or ties with every other model for groups biased on prior use

Type of model	Number
Demographic versus Hospital Use	
Demographic better	3
Hospital Use better	44
Ties	1
Demographic versus Hospital Days, Part B	
Demographic better	3
Hospital Days, Part B better	42
Ties	3
Hospital Use versus Hospital Days, Part B	
Hospital Use better	17
Hospital Days, Part B better	28
Ties	3

Table 4
Average absolute error of the AAPCC[1] third component predictions from prior-utilization models for 8 States or pairs of States,[2] by type of biased group

Type of biased group	Type of model		
	Demographic	Hospital Use	Hospital Days, Part B
Biased on prior use			
Low Part B use	.09	.04	.04
High Part B use	.11	.06	.03
Low hospital use	.08	.03	.04
High hospital use	.10	.03	.04
Low prior reimbursement	.16	.11	.06
High prior reimbursement	.12	.09	.04
Biased on demographics			
Low age	.05	.04	.05
High age	.05	.05	.05
Low buy-in	.02	.02	.02
High buy-in	.06	.06	.06

[1] Adjusted average per capita cost.
[2] The 8 States or pairs of States are New York, California, Pennsylvania, Texas and Oklahoma, Illinois and Indiana, Michigan and Ohio, Georgia, and Florida.

Table 6
Number of times each model predicts better or ties with every other model for groups biased on demographics

Type of model	Number
Demographic versus Hospital Use	
Demographic better	6
Hospital Use better	9
Ties	17
Demographic versus Hospital Days, Part B	
Demographic better	7
Hospital Days, Part B better	11
Ties	14
Hospital Use versus Hospital Days, Part B	
Hospital Use better	8
Hospital Days, Part B better	6
Ties	18

In addition, it seems that the simple measure of presence or absence of hospitalization in a previous period is nearly as effective a measure of prior use as more detailed measures, including number of hospital days and use of Part B services. This could be advantageous in developing an AAPCC based on prior use, because some individuals, such as those enrolling as soon as they become 65 years of age, could have no utilization data in the HIMA File. Clearly defined and easily collected data are a necessity if such a mechanism is to be operationally possible. Data on whether or not a person was hospitalized in the past year could probably be collected through a simple questionnaire.

An AAPCC based solely on administrative data maintained on 100 percent of the Medicare beneficiaries (i.e., the HIMA File) shows promise of being efficient to implement and operate. A note of caution should be made here, however. The HIMA File has never been used for this purpose before and it is not known what idiosyncratic variations may be found when looking at small areas. For example, if a county or group of counties had bill submissions that were faster, or slower, than the national average, predictions for that group of counties could be biased. Before this mechanism could be implemented at the national level, considerable testing would have to be done.

The prior-use models developed in this study are at least as good and, for some groups, better in a predictive sense than a model similar to the current AAPCC formula. The simulated groups for which the prior-use models were found in this study to be better than the current AAPCC are groups which are under- or over-represented by persons with high prior use. This could be a particularly salient finding in that studies of Medicare enrollment in HMO demonstration projects paid by the AAPCC mechanism have shown significant differences in prior use between enrollees and the average Medicare population (Eggers, 1980; Eggers and Prihoda, 1982). An AAPCC formula based partly on prior use might help eliminate this source of bias in HMO enrollment. Even so, none of the prior-use models entirely eliminate the effects of biased selection. Losses to Medicare from biased selection of low users into HMO's may be reduced, but not eliminated, with prior-use models.

It should be kept in mind that this analysis employed simulated biased groups. It is unknown what effect "real-world bias" would have on the efficacy of these models. It is likely that real-world bias will depend, in part, on the final form of the AAPCC formula, which will change the incentives for HMO's and influence their behavior. This means that any revised AAPCC formula should be tested before implementation and evaluated after implementation to identify possible unexpected effects on HMO performance.

The major potential problem with the inclusion of a prior-use measure is the possibility of new selection bias problems. It is conceivable that health plans

could selectively market to persons with hospitalizations for self-limited acute conditions (such as broken arms or legs or cataract removal) and avoid those with hospitalizations for conditions which are likely to require repeat admissions (e.g., cancer and heart disease). The possibility that a prior-use model would allow HMO's to bias enrollment in their favor to a much greater extent than under the present AAPCC has caused many actuaries to be cautious toward the concept of using prior use in an AAPCC formula.

This analysis has implications for the development of any kind of payment system based on actuarial categories. The Tax Equity and Fiscal Responsibility Act of 1982 states that the Department of Health and Human Services will develop an AAPCC which will insure "actuarial equivalence" with the general Medicare population. No AAPCC will ever be actuarially perfect. If it were true that persons were randomly drawn from different age and sex categories into health plans, then age and sex adjustment alone would be sufficient. Similarly, if one could be certain that persons will be randomly drawn from the universe of hospitalized patients, then the prior-use model will be sufficient. It is likely that experience will show this not to be the case. Nevertheless, it seems evident that a prior-use model of an AAPCC represents a potential advance over the current methodology. Efforts should be made to continue its development.

Future research

More precise measures of prior hospitalization than those used in the work reported here may lessen the chance of gaming an AAPCC formula with a prior-use model as well as improve its accuracy. The coding of Medicare hospital diagnosis on a 100-percent basis beginning in 1984 makes an adjustment based on hospital diagnosis feasible. Work is underway on models using hospital diagnosis to classify stays into those for self-limiting conditions and those for conditions, like cancer, indicative of chronic, recurring problems. Such a model should improve on the prior-use models reported here that simply employ the presence or absence of a prior hospital stay or the number of prior hospital days and do not distinguish between hospitalizations for chronic conditions and those for acute ones.

Any HMO prospective rate-setting system using prior-reimbursement variables would have to deal with five groups of Medicare eligible HMO enrollees:
1. Disability beneficiaries under 65 years of age.
2. HMO enrollees over 65 years of age who were in the Medicare program prior to enrollment long enough to have had the opportunity for any prior use to be recorded on the HIMA file.
3. Persons joining an HMO as soon as they turn 65 years of age and become entitled to Medicare.
4. Members of an HMO cost-reimbursement plan who switch their coverage to an at-risk HMO.

5. Persons who have been HMO members under an at-risk, AAPCC-type reimbursement system for more than 1 year.

The data base used for current research is not a large enough sample to permit the development of prior-use models for disabled enrollees under 65 years of age. However, we would expect that the same kinds of models as developed for the aged would be appropriate for the disabled. Thus, it would only be necessary to collect a larger sample to develop these models.

Persons in Group 2 are the only ones for which the research described in this article is directly applicable, although they constitute the largest of the five groups by far. Persons in Group 3 who simultaneously age into Medicare and join an HMO have no prior-use recorded on the HIMA File. Perhaps these people could be asked a simple question such as whether or not they were hospitalized in the past year. Their response could be used in the Hospital Use model discussed earlier. Alternatively, the current AAPCC could be used.[7]

Persons in Group 4 present a problem because no system exists to report Part B utilization data for cost-reimbursement HMO enrollees. For these people, a prior-use AAPCC employing only hospital use may have to be used. In any case, HMO's might argue that the HMO effect would have reduced utilization levels below that of comparable fee-for-service enrollees and would thus penalize HMO's for medically appropriate reductions in utilization levels.

Group 5 presents two problems for a prior-use AAPCC. First, paying HMO's on the basis of prior use of their own members may be counter to the incentives for utilization reduction inherent in the HMO concept. Second, even without adverse incentives, no system is established to report complete, person-level, utilization data from at-risk HMO's under TEFRA.

We are pursuing two lines of research in the hope of finding solutions to the problems of setting premiums for enrollees in some of these groups. First, we will develop a prior-use model that predicts reimbursement for a second year into the future. This, if successful, will provide a payment method for the second year of HMO enrollment. Second, we are conducting research that will shed light on the degree to which the average reimbursement of high- or low-cohorts of users regress toward the mean reimbursement over time. If such a phenomenon exists and is consistent, it may provide a method for reimbursing HMO enrollees beyond the second year. Data on regression to the mean could tell us how to set payments after an enrollee's second year in an HMO so the payments would approach the mean

reimbursement by a set amount each year until the mean was reached after a pre-determined number of years.

Along another line, an HMO demonstration is being conducted for which payment is based on prior-use models for enrollees from Group 2 for their first 2 years. It has not yet been determined how this group will be reimbursed in the third and subsequent years. Payment for enrollees from all other groups will be based on the current AAPCC.

Our study has shown that although a prior-utilization AAPCC could be superior to the current AAPCC in dealing with biased selection on the basis of prior use, questions still exist about how to implement it for some groups. The AAPCC should be viewed as an evolving system that will change on the basis of additional research, experience, and shifts in program goals, rather than as a formula that will be set for many years to come.

Acknowledgments

Technical note

In the reimbursement formula for HMO's under risk, the ratio of the underwriting index for the enrolled group to the underwriting index for the county population is used to adjust the county per capita cost to reflect the characteristics of the HMO enrollees.

Equation 1 represents the AAPCC formula:

$$(1)\ AAPCC = USPCC \times \frac{APCC_{Co}}{APCC_{US}}$$

$$\times \frac{\sum_{i=1}^{30} U_i E_{iHMO} \Big/ \sum_{i=1}^{30} E_{iHMO}}{\sum_{i=1}^{30} U_i E_{iCo} \Big/ \sum_{i=1}^{30} E_{iCo}}$$

Where:
AAPCC = adjusted average per capita cost
USPCC = U.S. average per capita cost to the Medicare program
$\frac{APCC_{Co}}{APCC_{US}}$ = ratio of per capita reimbursement in the county to the United States (5-year average)

And:
U = a unique underwriting index which represents the ratio of risk for a particular subset of the Medicare population to the national average. Thirty population subsets are defined by the four variables: age (five groups), sex (two groups), and institutional and welfare status (three groups).
E = the number of Medicare beneficiaries in a unique underwriting index cell.

[7]Work is under way on an additional adjustor for the AAPCC. It uses an indicator of whether or not the enrollee was previously entitled to social security disability benefits. Aged enrollees previously entitled to disability benefits use more Medicare services than other enrollees. If the promise of initial work is born out, this adjustment could be incorporated into an AAPCC formula for Groups 2 and 3.

Thus, the last component represents the ratio of relative cost differences in the HMO to that of non-HMO enrollees because of demographic characteristics. It is calculated by dividing the relative risk of HMO enrollees by the relative risk of non-HMO enrollees in a given county. The product of all these components gives the AAPCC.

References

Anderson, G., and Knickman, J.: Adverse selection under a voucher system: Grouping Medicare recipients by level of expenditure. *Inquiry* 21(2):135-143, Summer 1984a.

Anderson, G., and Knickman, J.R.: Patterns of expenditures among high utilizers of medical care services: The experience of Medicare beneficiaries from 1974 to 1977. *Med Care* 22(2):143-149, Feb. 1984b.

Arthur D. Little, Inc.: *Evaluation of the Impact of Competitive Incentives on Employee's Choice of Health Care Coverage, Executive Summary.* Contract No. HHS-100-81-0067. Prepared for Office of the Secretary, Department of Health and Human Services. Cambridge, Mass. May 1983.

Congressional Budget Office: Cost estimate for H.R. 3399, May 26, 1982.

Densen, P., Shapiro, S., and Einhorn, M.: Concerning high and low utilizers of a service in a medical care plan, and the persistence of utilization levels over a three year period. *Milbank Memorial Fund Quarterly* 37:217-239, 1959.

Eggers, P.: Risk differential between Medicare beneficiaries enrolled and not enrolled in an HMO. *Health Care Financing Review.* Vol. 1, No. 3. HCFA Pub. No. 03027. Office of Research, Demonstrations, and Statistics, Health Care Financing Administration. Washington. U.S. Government Printing Office, Winter 1980.

Eggers, P., and Prihoda, R.: Pre-enrollment reimbursement patterns of Medicare beneficiaries enrolled in "at-risk" HMO's. *Health Care Financing Review.* Vol. 4, No. 1. HCFA Pub. No. 03146. Office of Research and Demonstrations, Health Care Financing Administration. Washington. U.S. Government Printing Office, Sept. 1982.

Jackson-Beeck, M., and Kleinman, J.H.: Evidence for self-selection among health maintenance organization enrollees. *JAMA* 250(20):2826-2829, Nov. 25, 1983.

Kunkel, S. A., and Powell, C. K.: The adjusted average per capita cost under risk contracts with providers of health care. *Transactions of the Society of Actuaries* 33:57-66, 1981.

Luft, H.S.: Health maintenance organizations and the rationing of medical care. *Milbank Memorial Fund Quarterly* 60(2):268-306, Spring 1982.

Luft, H.S.: How do health maintenance organizations achieve their "savings"? Rhetoric and evidence. *N Engl J Med* 298:1336-1343, June 15, 1978.

Luft, H.S.: The self-selection issue. *Health Maintenance Organizations: Dimensions of Performance.* New York John Wiley and Sons, 1981.

Manning, W.G., Leibowitz, A., Goldberg, G.A., et al.: A controlled trial of the effect of a prepaid group practice on use of services. *N Engl J Med* 310(23): 1505-1510, 1984.

McCall, N., and Wai, H.S.: An analysis of the use of Medicare services by the continuously enrolled aged. *Med Care* 21(6): 567-585, June 1983.

McMillan, A., Pine, P., Gornick, M., and Prihoda, R.: A study of the crossover population: Aged persons entitled to both Medicare and Medicaid. *Health Care Financing Review.* Vol. 4, No. 4. HCFA Pub No. 03152. Office of Research and Demonstrations, Health Care Financing Administration. Washington. U.S. Government Printing Office, Summer 1983.

Roghmann, K.J., Sorensen, A.A., and Wells, S.: Hospitalizations in three competing HMOs during their first two years: A cohort study of the Rochester experience. *Group Health Journal* 1(1):26-33, Winter 1980.

Roemer, M.I., and Shonick, W.: HMO performance: The recent evidence. *Milbank Memorial Fund Quarterly* 271-317, Summer 1973.

Roos, N. P. and Shapiro, E.: The Manitoba longitudinal study on aging: Preliminary findings on health care utilization by the elderly. *Med Care* 19(6):644-657, June 1981.

Schuttinga, J. A. Falik, M., and Steinwald, B.: Health plan selection in the Federal Employees Health Benefits Program. Office of Health Policy, Assistant Secretary for Planning and Evaluation, Department of Health and Human Services. Unpublished paper, Sept. 1984.

Thomas, J.W., Lichtenstein, R., Wyszewianski, L., et al.: Increasing Medicare enrollment in HMOs: The need for capitation rates adjusted for health status. *Inquiry* 20:227-239, Fall 1983.

Trapnell, G.R., McKusick, D.R., and Genuardi, J.S.: An evaluation of the adjusted average per capita cost (AAPCC) used in reimbursing risk-basis HMOs under Medicare. Unpublished report. Washington, D.C. Actuarial Research Corporation, April 1982.

Welch, W.P.: Regression Towards the Mean in Medical Care Cost: Implications for Biased Selection in HMOs. Unpublished paper. Graduate School of Public and International Affairs, University of Pittsburgh, 1984.

Wolinsky, F.D.: The performance of health maintenance organizations: An analytic review. *Milbank Memorial Fund Quarterly* 58(4):537-587, 1980.