Risk Adjustment for Dually Eligible Beneficiaries Using Long-Term Care

Nelda McCall, M.A., and Jodi Korb, M.A.

This study explores use of the principal inpatient diagnostic cost groups (PIPDCG) and hierarchical coexisting conditions (HCC) risk-adjustment methodologies for a population of dually eligible beneficiaries receiving chronic long-term care (LTC). Measures of individual predictive accuracy for this population compared with the total Medicare population were similar for the PIPDCG models but somewhat smaller for the HCC models. Incorporating measures of functional status increased the R^2 values by only a small amount for Medicare expenditures but by a somewhat larger amount for total expenditures. Addition of other variables, especially placement, further improved the predictive power.

INTRODUCTION

The issue of examining risk adjustment for beneficiaries jointly eligible for the Medicare and Medicaid programs is important from multiple perspectives. The Federal Government is beginning a major reform in its Medicare managed care program to promote new managed care enrollment options and to develop a more equitable reimbursement system. Some States are also contemplating changes in how they might better serve the dually eligible population by integrating acute and LTC.

In each of these efforts, there is a need to understand the underlying health status of the populations being served so that payments can be set to ensure that the sickest groups receive adequate payment. Because of dually eligible beneficiaries' high level of disability and cost, there are substantial risks for both paying too little and promoting beneficiary underservice or paying too much and wasting limited government resources. In the States' efforts, there is the additional need to understand the full range of medical and social services their beneficiaries are receiving under public programs and to determine how best to measure these needs in setting capitation rates. Both these efforts suggest that understanding and quantifying the need for services among enrolled beneficiaries will be an important component of the future delivery and payment system.

The Medicare reform legislated through the 1997 Balanced Budget Act creates the Medicare+Choice program, expanding the range of managed care choices available to Medicare beneficiaries. It also requires HCFA to develop risk-adjusted Medicare capitation payments by January 2000. Without an adequate risk-adjustment system for making capitation payments to managed care plans, there is serious concern as to whether the Medicare+Choice program will endanger access and quality of care to high-risk beneficiaries (Medicare Payment Advisory Commission, 1998; Swartz, 1998).

Dually eligible beneficiaries are among the most vulnerable of these high-risk populations. They are more often females and minorities and consume disproportionate shares of both programs' resources (Feder, 1997). In 1995 dually eligible beneficiaries made up 15 percent of Medicare enrollees

The authors are with Laguna Research Associates. This article was funded under the Health Care Financing Administration (HCFA) Order Number HCFA-98-0228. The opinions expressed in this article are those of the authors and do not necessarily represent the views of Laguna Research Associates or HCFA.

(Physician Payment Review Commission, 1997) and consumed 30 percent of Medicare expenditures (Health Care Financing Administration, 1997). In the Medicaid program, they made up 15 percent of all enrollees (Physician Payment Review Commission, 1997) and consumed 35 percent of all expenditures (Health Care Financing Administration, 1997).

Dually eligible beneficiaries receiving chronic LTC services are an even more vulnerable group. Being deemed at risk of institutionalization requires either a functional-status evaluation with a defined level of disability (such as limitations in two or three activities of daily living [ADLs]) or a physician's certification. Twenty-eight percent of all dually eligible beneficiaries are estimated to have limitations in more than two ADLs, compared with 6 percent of the Medicare population excluding the dually eligible beneficiaries (Feder, 1997), and 19 percent of the dually eligible beneficiaries report having poor health status compared with 8 percent of other Medicare beneficiaries (Physician Payment Review Commission, 1997).

A parallel current at the State level is pushing toward integration of Medicaid and Medicare services for dually eligible beneficiaries. Faced with two disparate systems for delivery and payment of services with often perverse incentives promoting high-cost institutional care over consumer-desired home care services, States are beginning to consider whether they should take a more aggressive management role for their dually eligible Medicaid beneficiaries. Medicaid programs in Arizona, Colorado, Connecticut, Florida, Maine, Massachusetts, Minnesota, New Hampshire, New York, Oregon, Rhode Island, Texas, Vermont, and Wisconsin have seriously considered, planned, or implemented programs that attempt to integrate care, and in some cases funding, for Medicare and Medicaid services (Booth, Fralich, and Saucier, 1997; Kaye and Fralich, 1998; Meiners, 1998). Successful implementation of these programs requires, among other things, a way to set capitation payments to providers so that the payments reflect the needs of those served.

This study is an exploratory look at the use of the diagnostic cost group (DCG) risk-adjustment methodology planned for use by HCFA in implementing risk adjustment in 2000 (Greenwald et al., 1998). Both Medicare and total (Medicare plus Medicaid) expenditures for dually eligible beneficiaries receiving chronic LTC services are examined. Also examined is how the addition of functional-status measures and other variables (placement, urban status, ethnicity, cash-assistance status, State) contribute to the explanatory power of the risk-adjustment models.

DATA

This study focuses on elderly and physically disabled dually eligible beneficiaries in the Arizona Long-Term Care System (ALTCS) and New Mexico's Medicaid LTC fee-for-service (FFS) program. All of these individuals were assessed to be at risk of institutionalization by their respective Medicaid programs and were receiving home care or nursing home care. Data for Arizona and New Mexico beneficiaries were available in a data base previously constructed for the evaluation of the ALTCS. This data base contains Medicare and Medicaid utilization, enrollment, and assessment information for the 21-month period from January 1, 1991, through September 30, 1992. During this time period, the Medicaid LTC programs in Arizona and New Mexico covered a similar array of services, including both nursing home care and home and community-based services.

Utilization data examined for Arizona beneficiaries are encounter data submitted by the LTC program contractors, FFS claims paid by ALTCS, and Medicare data from HCFA's national claims history data base. New Mexico FFS program data include both Medicare data from the national claims history data base for New Mexico and New Mexico Medicaid claims. Medicaid claims are those processed by the New Mexico fiscal intermediary and home and community-based services claims processed by the Coordinated Community In-Home Care program.[1]

Assessment data for ALTCS beneficiaries are from a computerized system maintained by the State. This system contains detailed assessment information from an 18-page State-applied pre-admission screening instrument measuring functional (including ADLs and instrumental activities of daily living), psychosocial, and medical needs. These data are available for all ALTCS enrollees. The assessment used in this study is the one performed closest to the midpoint of the individual's enrollment during the study period.[2] New Mexico assessment information is more limited in the scope of what was collected and the people for whom it was collected. The assessment instrument is a one-page hard-copy form completed by the nursing home staff or a licensed physician and sent to the New Mexico professional review organization, contracted by the State to determine medical eligibility. The form includes information on ADL limitations (using different level definitions than Arizona) and medical conditions. The study team coded these data from hard-copy forms stored in the warehouses of the New Mexico professional review organization for a subset of the study population, specifically clients identified to be new admissions to the New Mexico LTC program during 1990 or 1991. As in Arizona, the assessment closest to the date of the midpoint of the individual's enrollment period is used.

The study population was restricted to elderly and physically disabled dually eligible beneficiaries who had a full year of dual eligibility in calendar year 1991 and at least 1 month of eligibility in 1992. This resulted in a total of 4,915 ALTCS beneficiaries and 3,019 New Mexico beneficiaries. Two-thirds (or 5,271) of the beneficiaries had assessment data available. Of these, more than 90 percent (4,824) were Arizona beneficiaries.

METHODOLOGY

In this study we explore how well the existing DCG risk-adjustment models predict expenditures for dually eligible beneficiaries at risk of institutionalization. Both 1991 expenditures (concurrent) and 1992 expenditures (prospective) are predicted as a function of beneficiary characteristics in 1991, the base year. Expenditure measures include Medicare expenditures as well as total Medicare and Medicaid expenditures. Medicare expenditures are measured by summing Medicare paid amounts. Total expenditures are equal to Medicare-allowed amounts plus Medicaid paid amounts for non-crossover services.[3] Dollar amounts used for the Arizona encounter data are the amounts that the program would have paid if the service had been delivered by a FFS provider and not through one of its capitated contractors.

[1] Combining Medicare and Medicaid utilization data was complicated by the two States' different systems for handling joint Medicare/Medicaid (crossover) claims. In addition, the crossover indicator in Arizona was incorrectly coded in the encounter data during the study period. In order to correct ALTCS encounter data for crossover activity, all services received by a given beneficiary with the same service type and service date as a service in the Medicare data files were excluded from the ALTCS data. For consistency, this exclusion was also performed for the New Mexico data.

[2] Because the study population had to be enrolled for the first 12 months of the 21-month study period, almost all of the assessment data used were for the base year (1991).

[3] Non-crossover services are services that are not covered by Medicare, such as pharmacy.

DCG Models

The study uses DCG models designed for the Medicare population as implemented in DxCG Software, Release 3. The DCG system is one of a number of risk-adjustment systems that rely on demographic and diagnostic information available on administrative data to predict resource use. The DCG methodology has evolved over the years incorporating increasingly more information (Ash et al., 1989; Ellis and Ash, 1995-96; Pope et al., 1998b). Most notably the first models used only inpatient diagnoses, whereas currently inpatient and ambulatory diagnoses can be used. Although the original model was developed for Medicare, the system has since added models for Medicaid and commercial populations.

Other risk-adjustment systems with similar data requirements include the ambulatory care groups methodology originally developed for Medicaid primarily as an ambulatory case-mix adjuster for non-aged populations and subsequently expanded (Starfield et al., 1991; Weiner et al., 1991; Weiner et al., 1996); and the disability payment system developed for the disabled Medicaid population (Kronick et al., 1996). In addition, other systems using survey-based data have been developed and evaluated (Gruenberg, Kaganova, and Hornbrook, 1996; Pope et al., 1998a), but these have disadvantages in terms of their data requirements. This study focuses on the DCG methodology because of its probable use by HCFA for implementing risk adjustment in the year 2000.

Two variants of the DCG models are examined, the PIPDCG model and the HCC model. As a comparison to the PIPDCG and HCC models, the software also includes models based only on age and sex. All models rely on demographic and diagnostic information available from administrative data to predict costs in either the current or subsequent period. The models were developed using claims data for 1991 and 1992 on a 5-percent national Medicare sample of approximately 1.4 million beneficiaries (the "benchmark-Medicare" sample).

The PIPDCG model classifies individuals based on their age, sex, and principal inpatient diagnoses. Individuals are assigned to a single category representing the most significant medical problem recorded (the inpatient diagnosis associated with the highest future costs). Only a single diagnosis per inpatient stay is used. An advantage of the PIPDCG model is that it requires a fairly limited set of data (i.e., age, sex, and one *International Classification of Diseases, 9th Revision, Clinical Modification* [ICD-9-CM] diagnostic code per inpatient stay). However, with only a single inpatient diagnosis, it cannot capture the full picture of a person's health status because it does not include medical problems treated exclusively in the ambulatory setting and information about multiple problems. This may be especially problematic for the elderly and disabled who tend to have multiple chronic problems.

The HCC model incorporates all inpatient and outpatient diagnoses (except those reported on laboratory and other ancillary tests, encounters with non-clinicians, and durable medical equipment claims). Individuals are assigned to condition categories that are groups of diagnostic codes based on clinical condition and expected resource use. Within a disease, a person is assigned only to the condition category associated with the most serious manifestation of the disease process, however, assignments are made for each distinct disease. By assigning individuals to multiple categories, the HCC methodology recognizes the cumulative effects of multiple problems.

Study Design

The study is composed of three parts. In the first part, we examine the predictive accuracy of DxCG software's predicted Medicare expenditures,[4] comparing them against actual expenditures and calculating various measures of predictive accuracy. Predictions are made for the PIPDCG model and the HCC model applied both concurrently and prospectively, as well as the prospective age-sex model.

In the second part, we examine demographic-calibrated[5] results obtained by using the relative risk scores produced by the DxCG software (i.e., the individual's predicted resource use relative to the average predicted resource use of the population) and age-sex groups in ordinary least squares regressions run on the Arizona and New Mexico sample data.[6] This is done for the PIPDCG and HCC models applied concurrently and prospectively and for both Medicare and total (Medicare plus Medicaid) expenditures. For the baseline age-sex model, expenditures are regressed against dummy variables for each of the age-sex groups. Regressions are weighted by the number of months of eligibility during the year.

In the third part, independent variables for functional status and other characteristics of interest such as placement in home or nursing home care, ethnicity, cash-assistance status, urban status, and State are incorporated into the demographic calibration models to see if they improve the models' predictive power. These are measured in the base year, 1991. Predictions are calculated only for the subset of the sample for which functional-status data were available. Functional status is specified by three variables: the number of ADLs for which the beneficiary is independent, the number of ADLs for which the beneficiary is totally dependent, and a dummy variable indicating mental disorientation. ADLs examined are bathing, eating, mobility, transferring, and toileting. The placement variable is specified as home care (reference group), nursing home, or mixed (i.e., both nursing home and home care placement during the base year). Cash assistance is a dummy variable indicating that the beneficiary received cash assistance for more than 50 percent of the year. Ethnicity is entered as a dummy variable indicating the beneficiary is of white, non-Hispanic origin. Urban is a dummy variable indicating whether or not the beneficiary resides in an urban area, and State is a dummy variable indicating residence in Arizona.

Measures of predictive accuracy at the individual level are the R^2, mean absolute prediction error, the standard deviation of the absolute prediction error, and the percent of absolute prediction errors within specified dollar ranges. In assessing a model's explanatory power at the individual level, it is important to bear in mind that models do not have to explain 100 percent of the variation. Many high-cost acute medical conditions are by nature unpredictable in a given year. Researchers have estimated that only between 15 and 25 percent of the variation in total costs can be explained by even the best prospective models (Dunn et al., 1996; Newhouse, 1994; Ellis et al., 1996a).

Also examined are predictive ratios for various subgroups. The predictive ratio is equal to the sum of predicted expenditures for a group divided by the sum of actual expenditures for the same group. The closer

[4] Expenditures are annualized for each beneficiary by multiplying actual expenditures by 12 and dividing by the number of months with any observed enrollment.

[5] The demographic calibration model is one where prediction = $a + b\,(\hat{y}) + c$ (demographic factors), and a, b, and c are determined by the new data. $\hat{Y}$ is the relative risk score predicted by the software.

[6] A better technique would have been to use a split-sample design, where one-half of the sample is used for parameter estimation and the other half to make predictions. However, this was not feasible given the sample size. Because predictions are based on sample-derived parameter estimates, their predictive power may be overstated due to overfitting.

Table 1

Age and Sex Composition of the At-Risk and Benchmark-Medicare Samples

Sex and Age	Sample At-Risk	Sample Benchmark-Medicare
	Percent	
Both Sexes	100	100
Under 65 Years	9	9
65-69 Years	6	23
70-74 Years	10	25
75-84 Years	35	32
85 Years or Over	41	11
Female	76	59
Under 65 Years	4	4
65-69 Years	4	13
70-74 Years	6	15
75-84 Years	27	20
85 Years or Over	35	8
Male	24	41
Under 65 Years	5	6
65-69 Years	2	10
70-74 Years	3	11
75-84 Years	9	12
85 Years or Over	6	3

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991-September 30, 1992; statistics calculated from DxCG software output.

the ratio is to 1.00, the better the implied predictive accuracy. Subgroup predictive ratios are important to examine because subgroups that are underpredicted are those for which capitated plans have financial incentives to selectively avoid enrolling. Overpredicted subgroups are ones for which the financial incentives are to selectively enroll members of the subgroup. Subgroups examined are levels of expenditures in the base year (divided into quartiles), placement, age, sex, number of ADLs requiring total assistance, mental disorientation, and number of inpatient hospitalizations.

DESCRIPTION OF THE STUDY POPULATION

Beneficiary Characteristics

Table 1 shows the age and sex compositions of the sample of Arizona and New Mexico elderly and physically dually eligible beneficiaries assessed to be at risk of institutionalization and receiving home care or nursing home care (the at-risk sample) and the benchmark-Medicare sample. Not surprisingly, the at-risk sample skews toward the older age ranges and is predominately female (76 percent). Seventy-six percent of the at-risk sample were at least 75 years of age, with 41 percent 85 years or over. Beneficiaries in the at-risk sample were notably older than those in the benchmark-Medicare sample. The mean age of the at-risk beneficiaries was 80 years, compared with 72 years for beneficiaries in the benchmark-Medicare sample. The two populations also differed in their sex distribution, with the at-risk population having 17 percent more females. Of special note, 35 percent of the at-risk sample were females 85 years of age or over, compared with only 8 percent of the benchmark-Medicare sample.

Table 2 shows the percent distributions of selected functional and demographic characteristics for the at-risk sample in the base year. The distributions reveal a population that, on average, had considerable functional limitation. Functional-status data were available for approximately two-thirds of the sample. Of these beneficiaries, 59 percent had total assistance needs in at least one ADL, with 32 percent requiring total assistance in three or more ADLs. Only 12 percent were independent in at least three of the five ADLs examined. Eighteen percent had a positive indicator of mental disorientation.

Eighty-three percent of the beneficiaries resided in a nursing home for all of 1991, compared with 12 percent who received home care and 5 percent who changed placements during the year. The majority (74 percent) of the beneficiaries were white, and only 11 percent received cash assistance. Urban residents accounted for 57 percent, and Arizona beneficiaries comprised 62 percent of the sample.

Table 2

Distribution of Selected Functional and Demographic Characteristics of the At-Risk Sample: 1991

Characteristic	Statistic
Number of Persons	7,934
	Percent
Number of ADLs – Independent	
0	37.4
1-2	18.4
3-5	11.6
Unknown	32.7
Number of ADLs – Total Assistance	
0	27.4
1-2	18.6
3-5	21.4
Unknown	32.7
Mental Disorientation	
Yes	12.2
No	54.3
Unknown	33.4
Placement	
Nursing Home	82.7
Home Care	12.4
Mixed	4.9
Ethnicity	
White	73.7
Black	2.8
Hispanic	18.5
Native American	3.3
Other/Unknown	1.6
Cash Assistance	
Yes	10.5
No	89.5
Urban/Rural Status	
Urban	57.3
Rural	42.7
State	
Arizona	61.9
New Mexico	38.1

NOTES: ADLs is activities of daily living. Percents may not add to 100.0 because of rounding.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–December 31, 1991.

Rates per 10,000 for the 30 most common condition categories in the at-risk sample are shown in Table 3. These conditions are the basis of the HCC categories. The rates reflect the total counts of individuals who had a condition, whether or not they also had a more serious manifestation of that condition. For example, a person with both 050: heart attack and 057: hypertension is counted in both categories.

More than 3 in 10 at-risk beneficiaries were categorized as having other dermatological disorders (such as cellulitis or abscess) or other musculoskeletal and connective tissue disorders (such as gout or osteoarthritis). One of the more notable differences between the at-risk and benchmark-Medicare samples was for dementia (2,880 per 10,000), which was more than six times as common among at-risk beneficiaries. The at-risk sample also had a significantly higher incidence of psychosis and high-cost mental disorders (2.7 times), and depression and moderate-cost mental disorders (3.3 times). Some of the other conditions for which rates were higher for the at-risk sample included neurological conditions such as Parkinson's disease (3.3 times) and high- and low-cost cerebrovascular disease (2.3 times). Included among conditions for which the at-risk sample had lower rates compared with the benchmark-Medicare sample were conditions related to cancer; hypertension; and screening, observation, and special exams.

Expenditures

The at-risk sample's mean actual expenditures in 1991 were $3,827 for Medicare and $25,064 for Medicare plus Medicaid (Table 4). More than 95 percent of the expenditures fell within two standard deviations of the mean, with the highest amounts reported equal to $158,105 and $182,042, respectively. Annualized expenditures for 1992 were somewhat higher: $4,870 for Medicare and $26,900 for Medicare plus Medicaid. The largest amounts were $400,842 for Medicare and $448,039 for Medicare plus Medicaid, although more than 99 percent of the cases fell within two standard deviations of the mean.

Table 3

Rates per 10,000 Individuals for the 30 Most Common Condition Categories in the At-Risk and Benchmark-Medicare Samples

		Sample	
Code	Condition Category	At-Risk	Benchmark-Medicare
092	Other Dermatological	3,372	2,347
026	Other Musculoskeletal/Connective Tissue	3,022	3,879
030	Dementia	2,880	468
080	Other Urinary System	2,446	1,628
048	Congestive Heart Failure	1,961	1,238
097	Other Injuries and Poisonings	1,950	1,953
059	Low-Cost Cerebrovascular Disease	1,945	821
100	Minor Symptoms, Signs, Findings	1,844	3,354
099	Major Symptoms	1,828	3,312
073	Low-Cost Eye	1,731	2,521
004	Other Infectious Disease	1,703	1,468
060	High-Cost Vascular Disease	1,700	882
075	Low-Cost Ear, Nose, and Throat	1,685	1,976
052	Chronic Ischemic Heart Disease	1,647	1,710
117	Screening/Observation/Special Exams	1,557	4,263
043	Moderate-Cost Neurological	1,447	440
032	Psychosis/Higher Cost Mental	1,419	518
015	Diabetes With No or Unspecified Complications	1,362	1,093
064	Chronic Obstructive Pulmonary Disease	1,142	1,201
023	Low-Cost Gastrointestinal	1,133	1,680
029	Iron Deficiency and Other Anemias	1,123	962
057	Hypertension (High Blood Pressure)	1,088	2,071
062	Atherosclerosis	964	444
067	Low-Cost Pneumonia	962	551
033	Depression/Moderate-Cost Mental	943	288
018	Other Endocrine, Metabolic, Nutritional	920	2170
017	Moderate-Cost Endocrine/Metabolic/Fluid-Electrolyte	892	825
091	Chronic Ulcer of Skin	870	310
094	Hip Fracture/Dislocation	838	201
022	Moderate-Cost Gastrointestinal	662	683

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–September 30, 1992; statistics calculated from DxCG software output.

Dollars have not been adjusted to account for the differential costs of service provision in different geographic regions. Ellis et al. (1996b) found that incorporation of a geographic adjustment based on input prices from Medicare's prospective payment system's area hospital wage index and the Medicare fee schedule geographic adjustment factor for physician payment had little effect on the explanatory power of the HCC model, increasing the R^2 only from 8.62 percent to 8.88 percent.

Table 4 shows the distribution of 1991 expenditures by service type for the at-risk sample. Combined Medicare plus Medicaid expenditures were more than 6.5 times as high as those reported only for Medicare. Nursing facility dollars accounted for 87 percent of this difference. Another 8 percent was attributable to home and community-based services and drugs, neither of which is covered by Medicare. Although inpatient services accounted for 44 percent of Medicare expenditures, these services represented only 8 percent of total expenditures.

RESULTS

Uncalibrated Models

Individual Predictive Accuracy

Table 5 summarizes the individual predictive accuracy for Medicare using the prospective age-sex model, the concurrent and prospective PIPDCG models, and the

Table 4

Distribution of Medicare and Medicare-Plus-Medicaid Expenditures, by Service Type for the At-Risk Sample: 1991

	Medicare		Medicare Plus Medicaid	
Service Type	Dollars	Percent	Dollars	Percent
Mean Expenditures	$3,827	100	$25,064	100
Standard Deviation	7,923	—	11,164	—
Institutional				
Inpatient	1,666	44	1,949	8
Outpatient	692	18	953	4
Nursing Facility	267	7	18,640	74
Home Health Care	91	2	93	0
Non-Institutional				
Evaluation and Management	357	9	549	2
Procedures	201	5	305	1
Tests	75	2	81	0
Imaging	52	1	76	0
Therapies	1	0	36	0
Home and Community-Based Services	0	0	1,140	5
Other	425	11	762	3
Drugs	0	0	480	2

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–December 31, 1991.

concurrent and prospective HCC models. The age-sex model is included for comparison purposes. These statistics were calculated from output generated from the DxCG software.

As can be seen in the table, for Medicare expenditures the PIPDCG concurrent model explains 43.3 percent of the individual variation, and the PIPDCG prospective model explains 2.8 percent of the individual variation. Applied concurrently, the HCC model explains a smaller amount of the individual variation than the PIPDCG model, 31.0 percent versus 43.3 percent. This is explained by the fact that the PIPDCG model explicitly takes into account that a person has been hospitalized (compared with the HCC model, which does not), combined with the large percentage (44 percent) of Medicare dollars spent on hospital care for this population. Applied prospectively, the HCC model explains a larger amount of the individual variation than the PIPDCG model, 3.8 percent compared with 2.8 percent.

The mean absolute prediction error, which minimizes the effect of a relatively small number of individuals with large prediction errors, is smallest for the PIPDCG concurrent model and largest for the age-sex prospective model. It is almost identical for the PIPDCG and HCC prospective models. The standard deviation of the absolute prediction error, which describes how dispersed the prediction errors are around the mean, is very large for all of the prospective models. For the concurrent models, it is smaller for PIPDCG than HCC.

With respect to the distribution of the errors, the concurrent models have proportionately more predictions that are within $1,000 of actual: 65.9 percent for the PIPDCG concurrent model and 37.1 percent for the HCC concurrent model. The prospective age-sex model has the smallest percentage of errors within $1,000, only 5.5 percent. The percentage for the PIPDCG prospective model is 6.6 percent and slightly better at 6.9 percent for the HCC prospective model. At the other end of the distribution,

Table 5

Summary of Individual Predictive Accuracy for Uncalibrated Predictions of Medicare Expenditures

Measure	Age-Sex Prospective	PIPDCG Model Concurrent	PIPDCG Model Prospective	HCC Model Concurrent	HCC Model Prospective
Mean Actual Expenditures	$4,870	$3,827	$4,870	$3,827	$4,870
Mean Predicted Expenditures	4,518	3,752	4,699	4,084	5,432
R^2	-2.1	43.3	2.8	31.0	3.8
Mean Absolute Error	6,196	2,512	5,891	3,360	5,893
Standard Deviation of Absolute Error	27,886	5,410	27,492	5,658	27,300
Percent Absolute Error Within $1,000	5.5	65.9	6.6	37.1	6.9
Percent Absolute Error Within $5,000	63.7	84.4	73.3	79.9	67.0
Percent Absolute Error Within $10,000	90.2	94.3	87.7	93.6	89.0

NOTES: PIPDCG is principal inpatient diagnostic cost group. HCC is hierarchical coexisting conditions.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–September 30, 1992; statistics calculated from DxCG software output.

the prospective PIPDCG model has the largest proportion of absolute errors greater than $10,000, 12.3 percent, compared with 11.0 percent for the prospective HCC model and 9.8 percent for the age-sex model. The concurrent PIPDCG model has 5.7 percent of its absolute errors greater than $10,000, and the HCC concurrent model has 6.4 percent.

Thus for Medicare expenditures for this study population, the PIPDCG concurrent model is explaining more individual variation than the HCC concurrent model. The HCC prospective model is explaining more individual variation than the PIPDCG prospective model. The individual predictive accuracies (R^2 values) estimated by both prospective models for Medicare expenditures are less than one-half those that have been observed for the general Medicare population (DxCG, Inc., 1998) as shown in Table 6.

Group Predictive Accuracy

Table 7 shows group predictive ratios for subgroups for Medicare expenditures. A predictive ratio closer to 1.00 indicates better prediction; a ratio greater than 1.00 indicates overprediction, and less than 1.00 indicates underprediction. Subgroups examined are defined by placement (nursing home or home care), sex, age, number of ADLs needing total assistance, mental disorientation, level of base-year expenditures (divided into quartiles), and hospitalization experience in the base year.

In general all of the models tend to overpredict Medicare expenditures for nursing home placements, females, those 80 years of age or over, those with mental disorientation, those having lower expenditures and, in the HCC model, those with greater numbers of ADL limitations and those with no hospitalizations. In the PIPDCG model, those having no hospitalizations are underpredicted. This pattern would be expected because the PIPDCG model uses inpatient diagnoses only. The largest overprediction

Table 6

Individual Predictive Accuracies Estimated, by Concurrent and Prospective Models

Sample	Concurrent PIPDCG	Concurrent HCC	Prospective PIPDCG	Prospective HCC
At-Risk	43.3	31.0	2.8	3.8
Benchmark-Medicare	43.4	43.2	5.7	8.5

NOTES: PIPDCG is principal inpatient diagnostic cost groups. HCC is hierarchical coexisting conditions.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991-September 30, 1992; statistics calculated from DxCG software output.

Table 7

Predictive Ratios for Subgroups for Uncalibrated Predictions of Medicare Expenditures

Characteristic	Number	Mean 1991 Expenditures	Mean 1992 Expenditures	Age-Sex Prospective	PIPDCG Model Concurrent	PIPDCG Model Prospective	HCC Model Concurrent	HCC Model Prospective
Placement								
Nursing Home	6,558	$3,216	$4,093	1.21	1.06	1.15	1.15	1.18
Home Care	984	5,500	6,739	0.65	0.90	0.76	0.68	0.68
Mixed	392	9,860	13,170	0.35	0.78	0.50	0.61	0.46
Sex								
Male	1,940	4,352	6,214	0.81	1.00	0.85	1.02	0.86
Female	5,994	3,658	4,437	1.09	1.00	1.07	0.99	1.06
Age								
Under 70 Years	1,042	6,708	7,688	0.45	0.69	0.56	0.80	0.65
70-80 Years	1,999	4,396	5,562	0.76	0.97	0.84	0.97	0.88
80 Years or Over	4,893	2,982	3,966	1.37	1.17	1.28	1.12	1.22
ADLs –Total Assistance								
0	2,174	4,593	6,051	0.80	0.84	0.79	0.85	0.84
1-2	1,474	3,698	4,994	0.99	0.95	0.95	1.06	1.00
3-5	1,695	3,986	4,890	0.98	0.88	0.97	1.08	1.04
Unknown	2,591	3,155	3,784	1.30	1.33	1.34	1.08	1.19
Mental Disorientation								
Yes	971	3,177	4,182	1.19	0.97	1.12	1.07	1.12
No	4,310	4,310	5,658	0.85	0.87	0.85	0.96	0.90
Unknown	2,653	3,281	3,837	1.28	1.29	1.32	1.06	1.18
Quartiles of Use								
Lowest	1,984	182	2,421	2.04	4.56	1.49	9.32	1.45
26-50	1,983	551	3,008	1.65	1.63	1.21	4.44	1.41
51-75	1,984	1,835	4,922	0.98	1.24	0.86	2.03	1.01
Highest	1,983	12,743	9,302	0.51	0.89	0.87	0.58	0.73
Hospitalizations								
0	6,091	1,344	3,882	1.26	0.66	0.92	2.01	1.11
1	1,290	8,905	6,715	0.73	1.39	1.29	0.71	0.91
2 or More	553	19,337	11,808	0.39	0.84	0.90	0.54	0.70

NOTES: PIPDCG is principal inpatient diagnostic cost group. HCC is hierarchical coexisting conditions. ADLs is activities of daily living.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–September 30, 1992; statistics calculated from DxCG software output.

is for the lowest expenditure quartile and the largest underprediction is for the highest expenditure quartile.

Demographic Calibration Models

Demographic calibration of the DCG model involved regressing actual expenditures against the relative risk scores generated by the DxCG software and age-sex variables and examining the individual and group predictive accuracy. This is done both for Medicare and total (Medicare plus Medicaid) expenditures. Results for individual predictive accuracy of the demographic calibration models are shown in Table 8 and group predictive accuracy in Tables 9 and 10.

Medicare Expenditures

Both the concurrent and prospective PIPDCG and HCC models improve their individual predictive power over that observed for the benchmark-Medicare sample estimates (Table 5 compared with Table 8). All of the R^2 values increase, especially for the prospective models. The R^2 for the concurrent PIPDCG model increases from 43.3 percent to 47.0 per-

Table 8

Summary of Individual Predictive Accuracy for Demographic Calibrated Predictions of Medicare and Medicare-Plus-Medicaid Expenditures

	Age-Sex[1]		PIPDCG Model		HCC Model	
Measure	Concurrent	Prospective	Concurrent	Prospective	Concurrent	Prospective
Medicare						
R^2	4.1	2.0	47.0	5.6	32.3	5.6
Mean Absolute Error	4,375	6,011	2,543	5,803	3,395	5,825
Standard Deviation of Absolute Error	6,408	27,431	5,180	27,092	5,566	27,060
Percent Absolute Error Within $1,000	8.5	5.6	54.8	6.2	36.3	8.2
Percent Absolute Error Within $5,000	81.6	71.7	85.8	68.7	79.5	66.9
Percent Absolute Error Within $10,000	92.5	90.1	95.1	88.9	93.6	89.2
Medicare Plus Medicaid						
R^2	1.4	1.2	25.6	2.9	25.0	5.7
Mean Absolute Error	6,734	8,743	5,949	8,764	6,275	8,722
Standard Deviation of Absolute Error	8,803	34,562	7,570	34,074	7,358	33,396
Percent Absolute Error Within $1,000	15.2	10.9	16.3	9.9	14.1	10.8
Percent Absolute Error Within $5,000	59.3	49.2	62.3	48.7	57.7	47.1
Percent Absolute Error Within $10,000	78.1	68.3	81.1	68.4	80.9	69.6

[1] This model is generated running age-sex groups as the independent variables.

NOTES: PIPDCG is principal inpatient diagnostic cost group. HCC is hierarchical coexisting conditions.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991-September 30, 1992; statistics calculated from DxCG software output.

cent, and for the prospective PIPDCG model, from 2.8 percent to 5.6 percent. The HCC concurrent model's R^2 increases from 31.0 percent to 32.3 percent, and the HCC prospective model increases from 3.8 percent to 5.6 percent. The other measures of individual predictive accuracy also improved slightly for the demographic calibrated model over the benchmark-Medicare uncalibrated model. This suggests that the relationship between expenditures and demographic characteristics is different for this population of persons eligible for Medicaid LTC services in Arizona and New Mexico than for the national Medicare population.

The pattern of overprediction and underprediction for subgroups in the calibrated models is similar to that observed for the benchmark-Medicare model (Table 7 compared with Table 9). Overprediction is generally observed for those in nursing homes, those experiencing mental disorientation, and those with lower expenditures in the base year. The amount of variation by subgroups between predicted and actual expenditures for the calibrated versus the benchmark-Medicare model seems to, for some subgroups, become narrower and wider in other cases. In the PIPDCG models, demographic calibration results in more underprediction for the highest expenditures group and for the most ADL-impaired group.

Total Expenditures

These measures for total expenditures are shown in Table 8. The demographic calibrated models explain 25 percent of the variation in both the concurrent PIPDCG and HCC model. These models explain 2.9 percent of the variation in the prospective PIPDCG model and 5.7 percent of the variation in the prospective HCC model.

Mean absolute prediction errors and standard deviations of the absolute prediction error for total expenditures are higher than those for Medicare expenditures. For total expenditures, the PIPDCG concurrent model has the lowest mean absolute prediction error while the age-sex prospec-

Table 9

Predictive Ratios for Subgroups for Uncalibrated Predictions of Medicare Expenditures

Characteristic	Number	Mean 1991 Expenditures	Mean 1992 Expenditures	Age-Sex		PIPDCG Model		HCC Model	
				Concurrent	Prospective	Concurrent	Prospective	Concurrent	Prospective
Placement									
Nursing Home	6,558	$3,216	$4,093	1.13	1.15	1.04	1.12	1.12	1.15
Home Care	984	5,500	6,739	0.89	0.86	1.01	0.90	0.78	0.81
Mixed	392	9,860	13,170	0.44	0.41	0.75	0.52	0.63	0.48
Sex									
Male	1,940	4,352	6,214	1.00	1.00	1.00	1.00	1.00	1.00
Female	5,994	3,658	4,437	1.00	1.00	1.00	1.00	1.00	1.00
Age									
Under 70 Years	1,042	6,708	7,688	1.00	1.00	1.00	1.00	1.00	1.00
70-80 Years	1,999	4,396	5,562	0.99	0.98	1.00	0.99	1.01	0.99
80 Years or Over	4,893	2,982	3,966	1.01	1.01	1.00	1.01	0.99	1.01
ADLs –Total Assistance									
0	2,174	4,593	6,051	0.83	0.80	0.84	0.80	0.85	0.83
1-2	1,474	3,698	4,994	1.01	0.96	0.94	0.93	1.04	0.98
3-5	1,695	3,986	4,890	0.99	1.01	0.92	1.00	1.09	1.05
Unknown	2,591	3,155	3,784	1.21	1.28	1.30	1.32	1.08	1.20
Mental Disorientation									
Yes	971	3,177	4,182	1.14	1.12	0.96	1.08	1.04	1.08
No	4,310	4,310	5,658	0.90	0.87	0.88	0.86	0.97	0.90
Unknown	2,653	3,281	3,837	1.17	1.27	1.26	1.31	1.06	1.20
Quartiles of Use									
Lowest	1,984	182	2,421	20.2	1.96	6.89	1.55	9.23	1.50
26-50	1,983	551	3,008	6.46	1.53	2.21	1.21	4.27	1.37
51-75	1,984	1,835	4,922	2.14	1.01	1.41	0.91	2.05	1.03
Highest	1,983	12,743	9,302	0.32	0.56	0.80	0.83	0.59	0.73
Hospitalizations									
0	6,091	1344	3,882	2.81	1.24	1.01	0.99	2.01	1.12
1	1,290	8,905	6,715	0.43	0.74	1.22	1.16	0.70	0.88
2 or More	553	19,337	11,808	0.23	0.46	0.75	0.83	0.55	0.69

NOTES: PIPDCG is principal inpatient diagnostic cost group. HCC is hierarchical coexisting conditions. ADLs is activities of daily living.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–September 30, 1992; statistics calculated from DxCG software output.

Table 10

Predictive Ratios for Subgroups for Demographic Calibrated Predictions for Medicare-Plus-Medicaid Expenditures

Characteristic	Number	Mean 1991 Expenditures	Mean 1992 Expenditures	Age-Sex		PIPDCG Model		HCC Model	
				Concurrent	Prospective	Concurrent	Prospective	Concurrent	Prospective
Placement									
Nursing Home	6,558	$26,508	$28,134	0.94	0.95	0.93	0.95	0.94	0.95
Home Care	984	15,831	17,319	1.62	1.58	1.66	1.59	1.57	1.55
Mixed	392	24,081	30,496	1.06	0.90	1.19	0.94	1.16	0.94
Sex									
Male	1,940	25,714	28,550	1.00	1.00	1.00	1.00	1.00	1.00
Female	5,994	24,854	26,367	1.00	1.00	1.00	1.00	1.00	1.00
Age									
Under 70 Years	1,042	26,954	28,755	1.00	1.00	1.00	1.00	1.00	1.00
70-80 Years	1,999	25,657	27,524	1.00	1.00	1.00	1.00	1.00	1.00
80 Years or Over	4,893	24,419	26,233	1.00	1.00	1.00	1.00	1.00	1.00
ADLs –Total Assistance									
0	2,174	23,436	25,542	1.07	1.05	1.07	1.05	1.07	1.06
1-2	1,474	25,662	28,195	0.99	0.95	0.96	0.95	0.98	0.96
3-5	1,695	28,193	30,329	0.89	0.89	0.88	0.89	0.91	0.90
Unknown	2,591	24,043	25,115	1.04	1.07	1.05	1.08	1.02	1.05
Mental Disorientation									
Yes	971	26,693	27,767	0.93	0.97	0.91	0.96	0.92	0.98
No	4,310	25,192	27,673	1.00	0.97	0.99	0.97	1.01	0.96
Unknown	2,653	24,260	25,208	1.03	1.07	1.05	1.07	1.01	1.05
Quartiles of Use									
Lowest	1,984	14,658	19,176	1.72	1.40	1.60	1.38	1.58	1.34
26-50	1,983	22,251	26,187	1.12	1.02	1.03	1.00	1.07	1.00
51-75	1,984	25,461	27,981	0.98	0.96	0.93	0.94	0.96	0.95
Highest	1,983	37,892	34,534	0.67	0.79	0.79	0.83	0.76	0.84
Hospitalizations									
0	6,091	22,478	26,104	1.11	1.03	1.00	1.00	1.05	1.00
1	1,290	30,697	28,348	0.82	0.95	1.06	1.04	0.92	1.00
2 or More	553	40,407	32,564	0.63	0.84	0.90	0.96	0.83	0.96

NOTES: PIPDCG is principal inpatient diagnostic cost group. HCC is hierarchical coexisting conditions. ADLs is activities of daily living.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–September 30, 1992; statistics calculated from DxCG software output.

Table 11

Effect of Inclusion of Functional Status and Other Variables on the R^2 of the Demographic Calibrated Predictions

	Age-Sex[1]		PIPDCG Model		HCC Model	
Variable	Concurrent	Prospective	Concurrent	Prospective	Concurrent	Prospective
Medicare						
Basic Model	4.59	2.85	48.55	6.38	31.72	6.13
DCG + Functional Status[2]	—	—	48.59	6.62	32.03	6.36
DCG + Functional Status[2] + Other[3]	—	—	49.26	8.50	33.48	8.43
Medicare Plus Medicaid						
Basic Model	1.43	1.50	27.89	3.18	27.25	6.01
DCG + Functional Status[2]	—	—	31.50	4.76	29.81	7.53
DCG + Functional Status[2] + Other[3]	—	—	46.79	17.71	40.61	18.45

[1] This model was generated running age-sex groups as the independent variables.

[2] Functional-status variables were: number of independent ADLs, number of totally dependent ADLs, and an indicator of mental disorientation.

[3] Other variables were placement, ethnicity, cash-assistance status, urban status, and State.

NOTES: This table includes only beneficiaries having coded functional status (5,271 of the 7,934 total population). PIPDCG is principal inpatient diagnostic cost group. HCC is hierarchical coexisting condition. DCG is diagnostic cost group. ADLs is activities of daily living.

SOURCE: Data from Arizona's Prepaid Medicaid Management Information System and New Mexico Medicaid data, January 1, 1991–September 30, 1992; statistics calculated from DxCG software output.

tive model has the highest. The HCC models have smaller standard deviations of the absolute prediction error than the PIPDCG models. The proportion of absolute prediction errors within $1,000 of actual expenditures ranges from 10.8 percent for the HCC prospective model to 16.3 percent for the PIPDCG concurrent model. A little more than 80 percent of the absolute prediction errors are within $10,000 for the concurrent models versus slightly less than 70 percent for the prospective models.

These measures for subgroups for total expenditures are shown in Table 10. Nursing home beneficiaries, those at higher levels of ADL disability, with mental disorientation, in the highest quartile of use, and having two or more hospitalizations are consistently underpredicted in both the concurrent and prospective PIPDCG and HCC models. The group with the largest underprediction is the group with three or more ADLs, where the HCC concurrent model underpredicts their use by 24 percent. The biggest overprediction is for home care beneficiaries across all models. The models all overpredict by 55-66 percent.

Functional Status and Other Variables

Approximately two-thirds of the study population had information on functional status available. For these individuals, concurrent and prospective age-sex, PIPDCG and HCC demographic calibrated models were estimated. Next, these models were estimated including functional-status variables (number of independent ADLs, mental disorientation, number of totally dependent ADLs). Finally, the models were estimated including functional status and other variables (i.e., placement, ethnicity, cash assistance, urban status, and State). R^2 values for these models are shown in Table 11.[7] For Medicare expenditures, the R^2 values are slightly larger for all the models except the HCC concurrent model than those reported in Table 8 for the full population. For total expenditures, R^2 values are larger in all the models for those having functional-status information, compared with the R^2 values reported for the full population.

[7] Detailed tables presenting regression results and predictive ratios for these analyses are available upon request from the authors.

Medicare Expenditures

For Medicare expenditures, addition of the functional-status variables slightly improves the predictive power of all the models. The R^2 increases by less than 1 percent in the two concurrent models (from 48.55 percent to 48.59 percent in the PIPDCG model and from 31.72 percent to 32.03 percent in the HCC model); the R^2 increases by a slightly larger amount, 4 percent, in the two prospective models (from 6.38 percent to 6.62 percent in the PIPDCG model and from 6.13 percent to 6.36 percent in the HCC prospective model).

In both the concurrent and prospective PIPDCG models, none of the functional-status variables are significant at the 5-percent level, although being mentally disoriented is significant and has a negative effect on expenditures at the 10-percent level of significance in the prospective PIPDCG model. In the concurrent HCC model, both ADL variables are significant. The number of independent ADLs has a positive effect, and the number of totally dependent ADLs has a negative effect on Medicare expenditures. This suggests that among this impaired population, those less disabled are larger consumers of the primarily acute care Medicare services, controlling for diagnosis, while those more impaired are relatively smaller users of such services. Because the more disabled are more likely to be institutionalized, this may suggest substitution of nursing home care for inpatient hospitalizations and other acute care services among this population. The prospective HCC model includes a significant positive effect on Medicare expenditures for the number of independent ADLs at the 10-percent level but no other significant effects.

The addition of other variables further increased the predictive power of the model, although it should be remembered that some of these variables are not ones that would be appropriate to include in a risk-adjusted capitation methodology. The most significant effects are found for the placement variables. In the HCC models, being in a nursing home compared with home care is associated with lower Medicare expenditures. Being in mixed placement (both home care and nursing home care over the period of study) relative to being in home care is associated with more Medicare expenditures. Urban status and State are significant in some of the models. Cash-assistance status and ethnicity are never significant. Thus, neither cash-assistance status nor ethnicity significantly impacts Medicare expenditures in a model controlling for diagnoses, functional status, and placement.

Total Expenditures

The R^2 increases more substantially in the total-expenditures model than in the Medicare model, increasing 13 percent for the PIPDCG concurrent model, 50 percent for the PIPDCG prospective model, 9 percent for the HCC concurrent model, and 25 percent for the HCC prospective model.

The number of totally dependent ADLs is significant in all four models, resulting in higher expenditures as the number of totally dependent ADLs increases. Mental disorientation results in more expenditures in the concurrent models and fewer in the prospective models. Two of these four results are only significant at the 10-percent level. The number of ADLs in which an individual is independent has significant negative effects in the concurrent models but no significant effects in the two prospective models.

The R^2 values increase substantially when other factors are added to the models, especially the prospective models. The two concurrent models' R^2 values increase

49 percent for PIPDCG and 36 percent for HCC. The PIPDCG prospective model's R^2 increases 272 percent and the HCC prospective model's R^2 increases 145 percent. As for Medicare expenditures, placement has an important effect on total expenditures. For nursing home residents, total expenditures are much larger, as they are for those with mixed placement. State and urban status often, but not always, affect the results, and being white compared with minority has significant negative effects in both prospective models. Cash-assistance status never has a significant effect.

DISCUSSION

This study reports findings from an exploratory study of the effect of using the DCG risk-adjustment methodology for a population assessed to be at risk of institutionalization living in the community and in nursing homes. It studies a population of dually eligible elderly and physically disabled beneficiaries for whom Medicare and Medicaid data had been integrated as part of a previously funded HCFA evaluation. The sample is restricted to those receiving chronic LTC services. The sample size is small and results should be interpreted cautiously.

Two kinds of expenditures were examined: Medicare expenditures and total (Medicare plus Medicaid) expenditures. With respect to Medicare expenditures, the uncalibrated PIPDCG and HCC models generally do not do as well in predicting Medicare expenditures for this population at risk of institutionalization as for the general Medicare population. In the concurrent models, R^2 values were 43.3 percent in the PIPDCG model and 31.0 percent in the HCC model. For the prospective analyses, the HCC model explained a larger percentage of the variation than the PIPDCG model: 3.8 percent versus 2.8 percent. These numbers compare with R^2 values of 43.4 percent for the PIPDCG and 43.2 percent for the HCC concurrent models, and 5.7 percent for the PIPDCG and 8.5 percent for the HCC prospective models found for the benchmark-Medicare population by the DxCG software developers.

Demographic calibration improves the predictive power of the models—slightly for the concurrent models (the PIPDCG model increased from 43.3 percent to 47.0 percent; the HCC model increased from 31.0 percent to 32.3 percent) but more substantially for the prospective models. The R^2 values for the PIPDCG model increased from 2.8 percent to 5.6 percent and for the HCC model from 3.8 percent to 5.6 percent. This suggests that the demographic calibrated PIPDCG model for this sample of individuals explains about the same amount of the variation as that reported for the full Medicare population. The HCC model estimated explains about two-thirds of the individual variation that is explained in the full Medicare population. Despite these encouraging findings, it should be remembered that considerable variation for this frail population remains unexplained, suggesting the importance of replicating these analyses on larger samples of the at-risk dually eligible population.

With respect to total expenditures, using demographic calibrated models, R^2 values for PIPDCG and HCC concurrent models are both 25 percent. The prospective PIPDCG model's R^2 is 2.9 percent, and the prospective HCC model's is 5.7 percent. Note that this 5.7 percent is slightly larger than the 5.6 percent for the prospective HCC model of Medicare expenditures. Although these are relatively encouraging R^2 values, given that the relative risk scores were developed from Medicare expenditures, more thought needs to be given to modeling expenditures that include nursing home costs. Nursing home costs (which accounted for nearly 75 percent of this population's total expenditures in 1991) are not so much related to the individual's disease history but to the event of being institutionalized. Once placed in a nursing home, the variation in nursing home costs is only over a narrow range relative to the large increase due solely to being in a nursing home. This suggests the need to study utilization data more closely by type of service and to develop strategies for capitation that take account of different predictors affecting different types of services.

When measures of functional status were added to the calibrated models, the R^2 values increased by only a small amount for Medicare expenditures (in all cases, less than 4 percent) and a somewhat larger amount for total expenditures with percentage increases of 9-50 percent. These findings suggest that, for Medicare expenditures, once variables capturing patient diagnoses are in the models, only a small amount of added predictive power is achieved by additional functional-status variables. For total expenditures, more predictive power is gained because the Medicaid costs, which are primarily for home and nursing home care, are more affected by functional status. Because this study used only three functional-status measures, results point to the importance of more research with a broader range of functional-status measures to investigate more closely functional status and the interrelationships between functional-status and diagnoses measures in predicting expenditures.

Addition of other available variables (placement, ethnicity, cash-assistance status, urban status, State) further improved the model's predictive power. Placement, as was expected, had a big and consistent effect on expenditures across all the models. Being in a nursing home, compared with home care, was associated with smaller Medicare costs. For total expenditures, being in a nursing home had significant positive effects. Consideration of the appropriate role of placement in the design of a risk-adjusted capitation system for populations such as these is of substantial importance.

Attaining higher levels of predictive accuracy for total (Medicare plus Medicaid) expenditures will be difficult, mainly because of the nature of nursing home costs (which include additional medical service costs and housing costs) and our lack of knowledge about the substitution effects of all kinds of services, and especially of how services like home care substitute for other types of care. Nonetheless, retaining a system that segments payment into acute (primarily Medicare) and LTC (primarily Medicaid) does not promote optimal care.

One approach to consider is removing the housing component of nursing home costs as part of the amount to be risk-adjusted. Once a patient has reached a certain threshold defined by a health, functional, and social assessment performed by an incentive-motivated case manager, this fixed amount would be added to the capitation payment and could be used either to pay for nursing home care or to finance home care or other services. But for this system to operate efficiently, it would have to be managed closely, and those receiving the capitation payments would need to have performance targets of the percentage of patients in less restrictive settings such as home care, with strong financial or other incentives tied to these target percentages. An approach like this is being used in Arizona and deserves more attention as payment reforms for capitating acute and LTC costs move forward (Wrightson, 1994; Wilkin et al., 1997; McCall, 1997).

The challenge of finding an appropriate payment system is great, but what is now necessary is to begin to formulate a starting point. Any system that is created will require evolution as markets change and better ways to use information are developed. As Congress has nudged progress forward by setting a challenging time frame on implementation of Medicare payment reform, so we must consider what a starting point for payment for integrated acute and LTC, including both Medicare and Medicaid, might be. The perfect should not be the enemy of the good as we begin to develop ways to pay for these populations so vulnerable and so costly. It is an inescapable priority necessary to ensure equity to the sickest members of our society.

ACKNOWLEDGMENTS

The authors would like to express their appreciation to Stanley Moore, who did the computer programming for this study. We are also grateful to the staff at the Health Care Financing Administration, especially our project officer, Jay Bae, James Lubitz, and Gerald Riley. Arlene Ash, Sean Aherne, and Randy Ellis of DxCG provided advice and consultation throughout the project. The authors would also like to acknowledge Pamela Turner and Morgan Tubby. Special thanks to Arlene Ash for her insightful comments and suggestions and for developing the terminology for describing the calibration technique employed.

REFERENCES

Ash, A., Porell, F., Gruenberg, L., et al.: Adjusting Medicare Capitation Payments Using Prior Hospitalization Data. *Health Care Financing Review* 10(4):17-29, Summer 1989.

Booth, M., Fralich, J., and Saucier, P.: *Integration of Acute and Long-Term Care for Dually Eligible Beneficiaries through Managed Care*. College Park, MD: University of Maryland Center on Aging, 1997.

Dunn, D., Rosenblatt, A., Taira, D., et al.: *A Comparative Analysis of Methods of Health Risk Assessment*. SOA Monograph M-HB96-1. Schaumburg, IL: Society of Actuaries, October 1996.

DxCG, Inc.: *Guide to the Diagnostic Cost Groups (DCGs) and DxCG Software: Release 3, SAS Version*. Waltham, MA: DxCG, Inc., April 1998.

Ellis, R., and Ash, A.: Refinements to the Diagnostic Cost Group (DCG) Model. *Inquiry* 32(4):418-429, Winter 1995/96.

Ellis, R., Pope, G., Iezzoni, L., et al.: Diagnosis-Based Risk Adjustment for Medicare Capitation Payments. *Health Care Financing Review* 17(3):101-128, Spring 1996a.

Ellis, R., Pope, G., Iezzoni, L., et al.: *Diagnostic Cost Group (DCG) and Hierarchical Coexisting Conditions (HCC) Models for Medicare Risk Adjustment Volume I: Final Report*. Waltham, MA: Health Economics Research, April 26, 1996b.

Feder, J.: *Medicare/Medicaid Dual Eligibles: Fiscal & Social Responsibility for Vulnerable Populations*. Washington, DC: The Kaiser Commission on the Future of Medicaid, May 1997.

Greenwald, L., Esposito, A., Ingber, M.J., and Levy, J.M.: Risk Adjustment for the Medicare Program: Lessons Learned from Research and Demonstrations. *Inquiry* 35(2):193-209, Summer 1998.

Gruenberg, L., Kaganova, E., and Hornbrook, M.: Improving the AAPCC with Health-Status Measures from the MCBS. *Health Care Financing Review* 17(3):59-75, Spring 1996.

Health Care Financing Administration: *A Profile of Dually Eligible Beneficiaries*. Washington, DC: Health Care Financing Administration, March 1997.

Kaye, N., and Fralich, J.: *Collection and Use of Data: State-Operated Managed Care Programs for Dual Eligibles*. College Park, MD: University of Maryland Center on Aging, 1998.

Kronick, R., Dreyfus, T., Lee, L., and Zhou, Z.: Diagnostic Risk Adjustment for Medicaid: The Disability Payment System. *Health Care Financing Review* 17(3):7-33, Spring 1996.

McCall, N.: Lessons from Arizona's Medicaid Managed Care Program. *Health Affairs* 16(4):194-199, July/August 1997.

Medicare Payment Advisory Commission: *Report to the Congress: Medicare Payment Policy, Volume1, Recommendations.* Washington, DC. March 1998.

Meiners, M.: *Medicare/Medicaid Integration Program Grantee Highlights.* Internet address: http://www.inform.umd.edu/aging. University of Maryland Center on Aging, October 8, 1998.

Newhouse, J.: Patients at Risk: Health Reform and Risk Adjustment. *Health Affairs* 13(1):132-146, Spring 1994.

Physician Payment Review Commission: *Annual Report to Congress.* Washington, DC. 1997.

Pope, G., Adamache, K.W., Walsh, E.G., and Khandker, R.W.: Evaluating Alternative Risk Adjusters for Medicare. *Health Care Financing Review* 20(2): 109-129, Winter, 1998.

Pope, G., Ellis, R., Liu, C.F., et al.: *Revised Diagnostic Cost Group (DCG)/Hierarchical Coexisting Conditions (HCC) Models for Medicare Risk Adjustment: Final Report.* Waltham, MA: Health Economics Research, Inc., February 6, 1998b.

Starfield, B., Weiner, J., Mumford, L., and Steinwachs, D.: Ambulatory Care Groups: A Categorization of Diagnoses for Research and Management. *Health Services Research* 26(1):59-68, April 1991.

Swartz, K.: Risk Selection and Medicare + Choice: Beware. *Inquiry* 35(2):101-103, Summer 1998.

Weiner, J., Starfield, B., Steinwachs, D., and Mumford, L.: Development and Application of a Population-Oriented Measure of Ambulatory Care Case-Mix. *Medical Care* 29(5):452-472, May 1991.

Weiner, J., Dobson, A., Maxwell, S.L., et al.: Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses. *Health Care Financing Review* 17(3):77-99, Spring 1996.

Wilkin, J., McCall, N., Wade, A., and Wrightson, C.W.: *Capitating Long-Term Care Medicaid Beneficiaries: More or Less Costly than Traditional Medicaid.* Discussion Paper 97-3. San Francisco: Laguna Research Associates, January 1997.

Wrightson, C.W.: *Method of Setting Capitation Payments.* Chapter 4 in Evaluation of Arizona's Health Care Cost Containment System Demonstration, Fourth Implementation and Operation Report. San Francisco: Laguna Research Associates, October 1994.

Reprint Requests: Laguna Research Associates, 1604 Union Street, San Francisco, CA 94123. E-mail: LagunaR@aol.com.