Principal Inpatient Diagnostic Cost Group Model for Medicare Risk Adjustment

Gregory C. Pope, M.S., Randall P. Ellis, Ph.D., Arlene S. Ash, Ph.D., Chuan-Fen Liu, Ph.D., John Z. Ayanian, M.D., M.P.P., David W. Bates, M.D., M.Sc., Helen Burstin, M.D., M.P.H., Lisa I. Iezzoni, M.D., M.Sc., and Melvin J. Ingber, Ph.D.

The Balanced Budget Act (BBA) of 1997 required HCFA to implement health-status-based risk adjustment for Medicare capitation payments for managed care plans by January 1, 2000. In support of this mandate, HCFA has been collecting inpatient encounter data from health plans since 1997. These data include diagnoses and other information that can be used to identify chronic medical problems that contribute to higher costs, so that health plans can be paid more when they care for sicker patients. In this article, the authors describe the risk-adjustment model HCFA is implementing in the year 2000, known as the Principal Inpatient Diagnostic Cost Group (PIPDCG) model.

INTRODUCTION

The goal of implementing health-status-based risk adjustment for Medicare capitation payments is to fairly compensate health plans for the expected costs associated with the disease burden of their enrollees. In support of this BBA mandate, HCFA has been collecting inpatient encounter data from health plans with discharges occurring since July 1997. These data include diagnoses and other information that can be used for risk adjustment. Risk adjustment will initially be based only on inpatient diagnoses.

The current PIPDCG model is the culmination of more than a decade of research supported by HCFA (Ash et al., 1989; Ellis and Ash, 1995; Ellis et al., 1996). Previous publications describe analyses of many methodological issues and alternative models. Here, we describe the specific model developed for year 2000 implementation and assess its performance. More details on development of the PIPDCG payment model are available in Pope et al. (1999). The physician co-authors have discussed clinical classification and other issues elsewhere (Iezzoni et al., 1998).

In this article, we first describe and briefly review the role of risk adjustment in Medicare payments to managed care plans and how the PIPDCG model determines a beneficiary's relative risk factor. Second, we comment on the strengths and limitations of using inpatient encounter data to adjust capitation payments for health status. This section puts the PIPDCG model in broader context and presents some concerns that helped shape model development. Third, we describe model develop-

Gregory C. Pope is with Health Economics Research, Inc. Randall P. Ellis is with Boston University. Arlene S. Ash is with Boston University School of Medicine. Chuan-Fen Liu is with Veterans Administration Puget Sound Health Care System. John Z. Ayanian and David W. Bates are with Harvard Medical School and Brigham and Women's Hospital. John Z. Ayanian is a paid consultant to Health Economics Research, Inc. and DxCG, Inc. Helen Burstin is with the Agency for Healthcare Research and Quality. Lisa I. Iezzoni is with Harvard Medical School, Beth Israel Hospital, the Charles A. Dana Research Institute, and the Harvard Thorndike Laboratory. Melvin J. Ingber is with the Health Care Financing Administration (HCFA). This research was supported by HCFA Contract Number 500-95-048. The views expressed in this article are those of the authors and do not necessarily reflect the views of Health Economics Research, Inc., Boston University, Boston University School of Medicine, Veterans Administration Puget Sound Health Care System, Harvard Medical School, Brigham and Women's Hospital, DxCG, Inc., the Agency for Healthcare Research and Quality, Beth Israel Hospital, the Charles A. Dana Research Institute, the Harvard Thorndike Laboratory, or HCFA.

ment: our data, sample, and variable definitions, the PIPDCG diagnostic classification system, and how diagnoses are sorted into Diagnostic Cost Groups (DCGs). The analysis and calibration of demographic factors for the PIPDCG model is reported next. Excluding diagnoses from short hospital stays is then considered. Finally, we examine the predictive accuracy and stability of the model and draw some conclusions.

MEDICARE RISK ADJUSTMENT

Medicare pays health maintenance organizations (HMOs) a monthly capitated amount for the medical care of each enrolled Medicare beneficiary. In the year 2000, 10 percent of payment for most beneficiaries is based on the PIPDCG risk-adjustment model, with the percentage scheduled to rise in following years. (The other 90 percent of payment is based on Medicare's historical adjusted average per capita cost [AAPCC] payment methodology as modified by the BBA. Enrollees who have been entitled to Medicare for less than 18 months will be paid for using a demographic model for new entitlees.) This capitated payment is the product of a county rate, determined by the beneficiary's residence, and a PIPDCG risk factor for that beneficiary. That is:

Payment = (Beneficiary relative risk factor) * (county rate)

For example, if a beneficiary living in a county with a monthly rate of $500 has a relative risk factor of 1.10, Medicare will pay a managed care plan 1.10*$500 = $550 per month for that beneficiary's medical care. The relative risk factor reflects the expected relative costliness of providing medical services to beneficiaries in different health states. By paying more for sicker beneficiaries, managed care plans are encouraged to enroll and work to satisfy the needs of such people. In this article, we explain how the PIPDCG model calculates an individual's risk factor. The risk-adjustment model is also used in calculating the county rate, as explained by Ingber (2000).

PIPDCG RELATIVE RISK FACTORS

The central feature of the PIPDCG model is calculating each beneficiary's relative risk factor. A beneficiary whose Medicare expenditures are predicted to equal the national average has a relative risk factor of 1.00. Risk factors greater than 1.00 indicate above average expected costliness; factors below 1.00 indicate lower-than-average expected costs. Tables 1 and 2 can be used to construct an individual's relative risk factor, starting with a base year (year 1) of demographic and medical information:

- *Step 1*. Compute a demographic factor (Table 1) by adding up to three individual factors: (1) age and sex; (2) originally disabled status (for a person who is now over age 65 but was previously entitled to Medicare because of disability); (3) Medicaid status (for a person who was entitled to Medicaid at any time during the base year).
- *Step 2*. Select the PIPDCG factor (Table 2) by: (1) assigning each hospital stay of at least 2 days to a PIPDCG category based on the principal medical problem that led to the admission; then (2) identifying the relative risk factor associated with the highest numbered of these PIPDCG categories. Note that beneficiaries with no hospital stays of at least 2 days are assigned to PIPDCG 4, along with those whose only hospitalizations fall into the lowest numbered PIPDCG, that is, 4; both groups receive PIPDCG 4's factor of zero.

Table 1

Demographic Factors Used by HCFA[1], by Sex and Age Group

Sex and Age Group	Age/Sex Factor	Additive Factors	
		Originally Disabled	Medicaid
Male			
0-34 Years	0.367	—	0.125
35-44 Years	0.380	—	0.283
45-54 Years	0.487	—	0.370
55-59 Years	0.615	—	0.397
60-64 Years	0.760	—	0.418
65-69 Years	0.541	0.415	0.440
70-74 Years	0.705	0.398	0.457
75-79 Years	0.907	0.334	0.461
80-84 Years	1.077	0.287	0.445
85-89 Years	1.258	0.237	0.404
90-94 Years	1.376	0.189	0.331
95 or Over	1.357	0.141	0.242
Female			
0-34 Years	0.362	—	0.192
35-44 Years	0.403	—	0.312
45-54 Years	0.526	—	0.367
55-59 Years	0.643	—	0.397
60-64 Years	0.891	—	0.412
65-69 Years	0.453	0.605	0.433
70-74 Years	0.588	0.576	0.440
75-79 Years	0.747	0.519	0.454
80-84 Years	0.918	0.415	0.423
85-89 Years	1.096	0.313	0.327
90-94 Years	1.162	0.232	0.231
95 or Over	1.128	0.152	0.168

[1] Refer to Table 2 for PIPDCG add-on factors. Working-aged multiplicative factor = 0.21.

NOTES: HCFA is Health Care Financing Administration. PIPDCG is Principal Inpatient Diagnostic Cost Group. Factors shown are for people with at least 1 year of eligibility. HCFA requires 12 months of data. Medicare beneficiaries under age 65 are eligible because of disability. The Medicare population mean = 1.

SOURCE: Health Care Financing Administration: *Proposed Method of Incorporating Health Status Risk Adjusters into Medicare+Choice Payments*. Report to Congress. Baltimore, MD. March 1, 1999.

- *Step 3:* Add the demographic and PIPDCG factors to achieve a relative risk score. If Medicare is not this person's primary payer, multiply this score by 0.21 to represent the expected part of total health care costs for which HCFA is responsible.

As an example, a male 69 years of age, not enrolled in Medicaid, never eligible for Medicare by disability, and not covered by another insurer, receives a demographic factor of 0.541. If he was not hospitalized (in year 1), nothing (the factor associated with PIPDCG 4) is added to this, and his Medicare expenditures this year are expected to be 54.1 percent of average. However, beneficiaries hospitalized for serious illnesses are assigned to higher

Table 2

Add-On Factors for PIPDCGs

PIPDCG	Factor
4	0.000
5	0.375
6	0.458
7	0.697
8	0.822
9	0.915
10	1.170
11	1.271
12	1.662
14	2.000
16	2.438
18	2.656
20	3.392
23	3.823
26	4.375
29	5.189

NOTE: PIPDCG is Principal Inpatient Diagnostic Cost Group.

SOURCE: Health Care Financing Administration: *Proposed Method of Incorporating Health Status Risk Adjusters into Medicare+Choice Payments*. Report to Congress. Baltimore, MD. March 1, 1999.

numbered PIPDCGs. The PIPDCG numbers (4 through 29) reflect approximate mean Medicare expenditures in the year following hospitalization, in thousands of 1996 dollars. For example, if our sample beneficiary were hospitalized for leukemia, he would be assigned to PIPDCG 29, because people with leukemia in 1995 had average 1996 costs of $30,456. Classification in this PIPDCG increases his expected costs for year 2 by adding 5.189 to the demographic factor, for a total of 5.730. That is, such a person is expected to incur costs in year 2 that are nearly 6 times average.

Tables 1 and 2 show that being older, male, enrolled in Medicaid, originally entitled to Medicare by disability, and hospitalized last year for more serious illnesses (i.e., assigned to a higher numbered PIPDCG) all increase a beneficiary's relative risk factor. "Working-aged" status, however, reduces the risk factor by almost four-fifths, because then Medicare is only responsible for paying for part of the beneficiary's health care.

Next, we examine the advantages and limitations of using principal hospital diagnoses to measure beneficiaries' health status. Then, we describe how we develop the PIPDCG model.

RISK-ADJUSTING USING INPATIENT ENCOUNTER DATA

The BBA mandated improved risk-adjustment formulas within 3 years. HCFA's year 2000 health-based risk-adjustment model uses inpatient hospital admissions records because only hospital data were feasible to collect, calibrate, and process within this time frame. Collecting inpatient records is seen as an interim step to collecting full encounter data from all (or most) care settings. Virtually all policy analysts support HCFA's intention to implement full-encounter risk adjustment as soon as feasible, and several all-encounter models have been developed and calibrated for Medicare (e.g., Ellis et al., 1996; Weiner et al., 1996). Yet, because risk adjustment in Medicare is beginning with inpatient data, it is useful to reflect on the strengths and weaknesses of an inpatient-only method.

The primary advantages of inpatient-based risk adjustment over all-encounter risk adjustment are practical. Inpatient diagnoses are obtained more easily and cheaply, and the data-collection burden on health plans and providers is substantially lower than with ambulatory encounter data. Inpatient diagnoses—especially principal diagnoses—are likely to be more accurate and are easier to audit and verify, and their quality is more nearly uniform across different systems. Because inpatient admission is also a proxy for severity of illness, it seems reasonable to begin the transition to risk-adjusted payments by focusing on the most severely ill and expensive enrollees, who are most likely to be hospitalized.

An inpatient admission—especially one of at least 2 days' duration—represents a significant expenditure by a health plan. Hospitalizing a patient who does not really need it, for the purpose of recording a diagnosis that would increase payments next year, is less likely because hospitalization is so disruptive and expensive. With an all-encounter model, a patient could easily be scheduled for extra ambulatory visits, during which additional, payment-increasing diagnoses could be recorded. Another benefit of inpatient-based risk adjustment is that capitated health plans, which receive no marginal payment for providing health care, may have an incentive to underprovide medical care, especially expensive ser-

vices such as hospitalizations (Newhouse, 1996). An inpatient-based risk-adjustment system partly mitigates this disincentive of capitation.

The primary disadvantage of inpatient-based risk adjustment is the distorted incentive for health plans to choose among sites of care. Plans obtain higher risk-adjusted payments only by admitting their enrollees to the hospital. Thus, plans can be penalized when they successfully avoid an unnecessary admission by providing appropriate ambulatory care. The incentive to admit is contrary to the usual tenets of managed health care, because managed care plans have achieved most of their cost savings by reducing inpatient hospital use (Miller and Luft, 1997). Mitigating inappropriate incentives for hospital admission is the motive for some aspects of the PIPDCG model development, as we discuss later.

Using only the principal diagnosis from inpatient stays to infer health status, as the PIPDCG model does, has related advantages and disadvantages. The principal diagnosis is likely to be of good quality because hospitals and attending physicians are accustomed to audits of this diagnosis in the diagnosis-related groups (DRGs) Medicare payment system, and it may be the only diagnosis available from some plans and hospitals. It is also a proxy for severity, because it is the problem that was the reason for the admission. A risk-adjustment model based only on the principal diagnosis is not sensitive to the completeness of coding of secondary diagnoses. On the other hand, using only the principal diagnosis makes the model particularly sensitive to resequencing of diagnoses. It is sometimes unclear which diagnosis should be principal, and plans have an incentive to reorder the diagnoses to maximize reimbursement. (Medicare DRG payments already create incentives for hospitals to consider the order of principal inpatient diagnoses for their payment incentives. The PIPDCG system may modify those incentives in some circumstances.) Also, because hospital admissions are often precipitated by acute health crises, the principal diagnosis may be more likely to represent an acute diagnosis than the underlying chronic illness. This consideration is somewhat at odds with the rationale for prospective risk adjustment, which seeks to predict year 2 costs from chronic illness. Finally, modeling choices (such as restricting attention to the single most costly principal diagnosis) made that were to mitigate the perverse incentives associated with hospital-based illness detection reduce predictive accuracy.

Even so, the PIPDCG model is far more powerful than the demographic factors in the AAPCC system used previously by HCFA to pay Medicare health plans. By providing greater fairness and accuracy in capitated payments, the current change in payment formulas is "a step in the right direction, albeit a modest step" (Iezzoni et al., 1998).

DEVELOPMENT OF THE PIPDCG MODEL

Data, Sample, Expenditures

The PIPDCG model has been developed and calibrated with data from the traditional Medicare fee-for-service (FFS) program, reflecting the legislative mandate that Medicare capitation be based on 95 percent of what a beneficiary is expected to cost in the FFS program, as it was under the previous AAPCC methodology. Risk-adjustment factors developed from FFS data are in this sense consistent with the historical basis of Medicare's capitation payment methodology. Yet even without such a mandate, FFS data represent the only comprehensive and representative

source of information on costs for treatment of Medicare beneficiaries. The implicit assumption is that relative costs of patients with specified levels of disease burden are similar in the FFS and managed care sectors.

The PIPDCG model was developed on a 5-percent sample of Medicare's FFS enrollees in 1995 and 1996. Diagnoses from hospitalizations in the base year of 1995 were used to predict 1996 Medicare expenditures. Beneficiaries who died in 1995 were excluded from the sample, but 1996 decedents were included. To ensure a complete base-year diagnostic profile and complete 1996 expenditures, beneficiaries in the sample had to be enrolled in Medicare FFS throughout 1995 and 1996. Beneficiaries eligible at any time during the sample period for the end stage renal disease program and beneficiaries for whom Medicare was not the primary payer were excluded.

Expenditures were aggregated from hospital inpatient, hospital outpatient, professional (physician/supplier), home health, and durable medical equipment claims. Hospice expenditures were excluded because managed care plans are not responsible for hospice care. (All months of hospice eligibility and expenditures during those months were excluded from the analysis.) Deductibles and copayments for Medicare-covered services that are the responsibility of beneficiaries were excluded from expenditures. Thus, expenditures are Medicare payments to providers. Indirect medical education expenditures were also excluded, because the BBA specifies that medical education payments are to be phased out of capitation payments to managed care plans and paid directly to teaching hospitals.[1]

To correctly estimate monthly payments for all beneficiaries, including those who died, we weighted observations by Medicare-eligible months in the prediction year (1996). First, we annualized 1996 payments, dividing actual total 1996 payments by the fraction of the year (rounded up to the nearest whole month) that each beneficiary was alive and eligible for Part A and Part B. Each observation was then weighted by the same fraction in all analyses. Annualizing and weighting observations results in unbiased estimates of the average and total payments for a group in which individuals are eligible for different fractions of the year.

Base-year diagnoses were obtained from 1995 hospitalizations in facilities eligible for Medicare's prospective payment system (PPS) and non-PPS facilities and units including psychiatric, rehabilitation, long-term, children's and other specialty hospitals. Diagnoses from skilled nursing facilities (SNFs) and skilled nursing units were not used. Demographic information, described in more detail later, was obtained from Medicare enrollment files.

Table 3 describes our 1995-1996 model development sample, which contains 1,387,105 beneficiaries. In 1996, no payment was made for 9.9 percent of our sample, and 4.8 percent died in that year.

Diagnostic Classification

The goal of our classification of diagnoses was to differentiate beneficiaries expected to have different levels of Medicare expenditures in year 2. Beneficiaries who are hospitalized for treatment of serious illnesses—for example, lung cancer—in year 1 are expected to have higher expenditures in year 2 than

[1] Medical education payments are comprised of indirect and direct payments. It was not feasible to exclude direct payments. Indirect payments, which we excluded, represent about two-thirds of total medical education payments.

Table 3

Statistics for the 1995-1996 Medicare 5-Percent Sample[1], by Beneficiary Characteristic

Characteristic	Number of Beneficiaries	Percent of Total Sample	Mean 1996 Payments	Ratio to the Mean
Overall Sample	1,387,105	100.0	$5,186	1.00
Disabled (Age ≤ 64)	154,784	11.2	4,636	0.89
Younger Disabled (Age ≤ 44)	55,579	4.0	3,846	0.74
Older Disabled (Age 45 - 64)	99,205	7.2	5,082	0.98
Aged (Age ≥ 65)	1,232,321	88.8	5,256	1.01
Originally Disabled	87,154	6.3	7,966	1.54
Younger Elderly (Age 65 - 84)	1,073,853	77.4	4,917	0.95
Older Elderly (Age 85 or Over)	158,468	11.4	7,685	1.48
Medicaid	204,267	14.7	7,290	1.41
Disabled (Age ≤ 64)	66,370	4.8	5,556	1.07
Elderly (Age ≥ 65)	137,897	9.9	8,161	1.57
Non-Medicaid	1,182,838	85.3	4,828	0.93
Disabled (Age ≤ 64)	88,414	6.4	3,944	0.76
Elderly (Age ≥ 65)	1,094,424	78.9	4,901	0.95
Female	812,354	58.6	5,098	0.98
Male	574,760	41.4	5,310	1.02

[1] Excludes working aged in 1995 and 1996.

SOURCE: Health Economics Research, Inc., analysis of 1995 and 1996 Medicare claims data, Waltham, MA, 1999.

beneficiaries who are not hospitalized or who are hospitalized for less serious illnesses. To begin, we classified all of the more than 15,000 *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) (Public Health Service and Health Care Financing Administration, 1980) diagnostic codes into 172 Principal Inpatient Diagnostic Groups (PIPDxGs). The primary criteria used in forming PIPDxGs were clinical coherence and adequate sample size (so that future mean expenditures could be estimated reliably, refer to Ash et al., 1989; Ellis and Ash, 1995; and Ellis et al., 1996, for further discussion). For example, PIPDxG 1 is central nervous system infections. In 1995, 222 Medicare beneficiaries in our sample were hospitalized with a principal diagnosis falling into this category, and they had mean 1996 Medicare expenditures of $11,291, more than double the sample average of $5,186.

The next step was to select diagnoses for inclusion in the payment model. Including all diagnoses in the risk-adjustment model creates incentives for even relatively healthy people to be admitted for minor diagnoses to obtain higher payment. We excluded diagnoses that may be minor, transitory, or non-specific. Examples of excluded diagnoses are sprains (minor), appendicitis (transitory, i.e., having definitive treatment), and fever (non-specific). By excluding such diagnoses, risk adjustment focuses on the burden of high-cost, chronic illness. Altogether, 75 of the 172 PIPDxGs were excluded from the payment model. (Note that the costs associated with these excluded diagnoses are not dropped. The model captures these costs in other factors, such as age and sex.)

Special attention was paid to two diagnoses, chemotherapy and human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS). The PIPDCG model is diagnosis-based and generally excludes treatments or procedures such as chemotherapy. However, we were reluctant to exclude chemotherapy because it identifies a group of very ill beneficiaries with high expected future costs, and in the

DRG-based payment system, hospitals are required by HCFA to code chemotherapy as the principal diagnosis when it is provided for hospitalized cancer patients. Our solution was to recognize chemotherapy admissions when coded with a "V-code" for chemotherapy as the principal diagnosis and then to classify such admissions by the most serious type of cancer present among the secondary inpatient diagnoses. HIV/AIDS is another special case, associated with very high future expenditures. If admission is for an HIV-related condition, HIV/AIDS may not be the principal diagnosis. We chose to allow classification into the HIV/AIDS PIPDxG based on either a principal or secondary diagnosis of HIV/AIDS. These are the only two situations in which secondary diagnoses are used in the Medicare PIPDCG model.

The result of our preliminary decisions was a list of diagnoses that were eligible to trigger increased payments. The next problem was how to differentiate beneficiaries by their expected future costliness based on this list of diagnoses. If each hospitalized beneficiary were hospitalized only once, ranking diagnoses by future costliness would be straightforward. However, approximately one-half of beneficiaries who are hospitalized during a year are hospitalized more than once. The PIPDCG model addresses the problem of multiple diagnoses per beneficiary by introducing a hierarchy; it identifies only the single diagnosis most predictive of higher future expenditures and ignores other diagnoses (admissions). To rank and group diagnoses by their expected future costliness, we used the DCG sorting algorithm (Ash et al., 1989; Ellis and Ash, 1995).

The DCG sorting algorithm starts by computing the mean 1996 expenditures of beneficiaries with each 1995 diagnosis. Note that the costs for a female hospitalized for two distinct medical problems will contribute to two PIPDxG means. All diagnoses are then ranked in descending order of their 1996 expenditures. The highest ranked diagnoses (PIPDxGs) are grouped into the highest ranked aggregated cost-based grouping or PIPDCG. The highest ranked PIPDCG is PIPDCG 29, which includes the PIPDxGs of HIV/AIDS and blood, lymphatic cancers/neoplasms. The PIPDCG number, 29, refers to the approximate mean 1996 Medicare expenditures in thousands of dollars for beneficiaries assigned to the PIPDCG. Each beneficiary admitted to the hospital in 1995 with a principal or secondary diagnosis of HIV/AIDS or a principal diagnosis of blood, lymphatic cancers, or neoplasms, is uniquely assigned to PIPDCG 29.

After the highest ranked PIPDCG is formed, beneficiaries assigned to it are removed from the sample, mean expenditures by diagnosis are recomputed, and the diagnoses are re-ranked. The next-highest-cost PIPDxGs are included in the second-ranked PIPDCG, number 26. This PIPDCG includes the diagnoses metastatic cancer and brain/nervous system cancers, with mean 1996 expenditures of approximately $26,000 each. Beneficiaries assigned to PIPDCG 26 are then removed from the sample, mean expenditures by diagnosis are recalculated, and the process is repeated. In this way, all diagnoses (PIPDxGs) are assigned to PIPDCGs.

The result of the DCG sorting algorithm is the 16 PIPDCGs (with number labels ranging from 29 on down to 4) presented in Table 4, which shows the diagnoses assigned to each of the cost-based groupings. (The physician co-authors reviewed the results of the sorting algorithm for clinical plausibility and incentives, and reassigned a small number of PIPDxGs. Refer to Pope et al., 1999 for details.) The relative risk factor of beneficiaries assigned to the base PIPDCG 4—which includes bene-

Table 4

Diagnoses Included in Each Principal Inpatient Diagnostic Cost Group (PIPDCG)

Group	Diagnosis
PIPDCG 29	HIV/AIDS[1]
	Blood, Lymphatic Cancers/Neoplasms[2]
PIPDCG 26	Metastatic Cancer[2]
	Brain/Nervous System Cancer[2]
PIPDCG 23	Liver/Pancreas/Esophagus Cancer[2]
	End Stage Liver Disorders
	Cardio-Respiratory Failure and Shock
	Decubitus and Chronic Skin Ulcers
PIPDCG 20	Diabetes with Chronic Complications
	Coma and Encephalopathy
	Aspiration Pneumonia
	Renal Failure/Nephritis
PIPDCG 18	Cancer of Placenta/Ovary/Uterine Adnexa[2]
	Blood/Immune Disorders
	Paralytic and Other Neurologic Disorders
	Polyneuropathy
	Gram-Negative/Staphylococcus Pneumonia
PIPDCG 16	Mouth/Pharynx/Larynx/Other Respiratory Cancer[2]
	Lung Cancer[2]
	Cirrhosis, Other Liver Disorders
	Congestive Heart Failure
	Atherosclerosis of Major Vessel
	Chronic Obstructive Pulmonary Disease
PIPDCG 14	Septicemia (Blood Poisoning)/Shock
	Adrenal Gland, Metabolic Disorders
	Delirium/Hallucinations
	Paranoia and Other Psychoses
	Anxiety Disorders
	Personality Disorders
	Degenerative Neurologic Disorders
	Spinal Cord Injury
PIPDCG 12	Tuberculosis
	Stomach, Small Bowel, Other Digestive Cancer[2]
	Rectal Cancer[2]
	Cancer of Bladder, Kidney, Urinary Organs
	Benign Brain/Nervous System Neoplasm
	Diabetes with Acute Complications/Hypoglycemic Coma
	Inflammatory Bowel Disease
	Rheumatoid Arthritis and Connective Tissue Disease
	Bone/Joint Infections/Necrosis
	Dementia
	Drug/Alcohol Psychoses
	Major Depression/Manic and Depressive Disorders
	Epilepsy and Other Seizure Disorders
	Cerebral Hemorrhage
	Stroke
	Peripheral Vascular Disease
	Pulmonary Fibrosis and Bronchiectasis
	Pleural Effusion/Pneumothorax/Empyema
PIPDCG 11	Gastrointestinal Obstruction/Perforation
	Gastrointestinal Hemorrhage
	Paroxysmal Ventricular Tachycardia
	Bacterial Pneumonia
	Cellulitis and Bullous Skin Disorders

See footnotes at end of table.

Table 4—Continued

Diagnoses Included in Each Principal Inpatient Diagnostic Cost Group (PIPDCG)

Group	Diagnosis
PIPDCG 10	Colon Cancer[2]
	Schizophrenic Disorders
	Post-Myocardial Infarction
	Unstable Angina
	Thromboembolic Vascular Disease
	Kidney Infection
	Vertebral Fracture Without Spinal Cord Injury
PIPDCG 9	Other Cancers[2]
	Pancreatitis/Other Pancreatic Disorders
	Acute Myocardial Infarction
	Transient Cerebral Ischemia
	Fractures of Skull/Face
	Pelvic Fracture
	Hip Fracture
	Internal Injuries/Traumatic Amputations/Third-Degree Burns
PIPDCG 8	Cancer of Uterus/Cervix/Female Genital Organs[2]
	Peptic Ulcer
	Valvular and Rheumatic Heart Disease
	Hypertension, Complicated
	Coronary Atherosclerosis
	Angina Pectoris
	Atrial Arrhythmia
	Precerebral Arterial Occlusion
	Aortic and Other Arterial Aneurysm
	Asthma
	Brain Injury
	Artificial Opening of Gastrointestinal Tract Status
PIPDCG 7	Central Nervous System Infections
	Abdominal Hernia, Complicated
	Alcohol/Drug Dependence
PIPDCG 6	Cancer of Prostate/Testis/Male Genital Organs[2]
PIPDCG 5	Breast Cancer[2]
	Ongoing Pregnancy with Complications
	Ongoing Pregnancy with No or Minor Complications
PIPDCG 4	No or Excluded Inpatient Admissions
	Ectopic Pregnancy
	Miscarriage/Terminated Pregnancy
	Completed Pregnancy with Major Complications
	Completed Pregnancy with Complications
	Completed Pregnancy Without Complications (Normal Delivery)

[1] Includes principal and secondary inpatient diagnoses of HIV/AIDS.

[2] Includes principal diagnoses and secondary diagnoses when the principal diagnosis is chemotherapy.

NOTES: HIV is human immunodeficiency virus. AIDS is acquired immunodeficiency syndrome.

SOURCE: Health Economics Research, Inc., Waltham, MA, 1999.

ficiaries with no prior-year hospitalizations or excluded low-future-cost admissions only—is based solely on demographic factors. A beneficiary's relative risk factor increases from its demographic baseline if and only if he or she is assigned to one of the PIPDCGs numbered 5-29.

Table 5 shows the number of admissions used in different stages of the PIPDCG modeling. Final assignment to a PIPDCG is based on diagnostic information relating to only 40 percent of base-year admissions. Diagnoses on 37 percent of admissions are omitted because their principal diagnoses

Table 5

Statistics for Admissions Used in Different Stages of PIPDCG Modeling

Admissions in 1995	Number of Admissions	Percentage	Percentage of Admissions in PIPDCG Sorting Algorithm
Total Admissions[1]	415,231	100.0	—
Excluded Admissions Due to Exclusions of PIPDxG	153,276	36.9	—
Admissions of People with an Included Admission	55,210	13.3	—
Admissions of People Without an Included Admission	98,066	23.6	—
Remaining Admissions Participate in the PIPDCG Sorting Algorithm	261,955	63.1	100.0
Exclusions Due to Multiple Admissions per Person[2]	85,389	20.6	32.6
Excluded Short-Stay Admissions[3]	8,889	2.1	3.4
Admissions Used to Define PIPDCG	167,677	40.4	64.0

[1] There were 258,363 persons with at least one admission in 1995.

[2] For a person with multiple admissions in 1995, only the admission with the highest future cost was used to determine the PIPDCG. However, one person may have multiple admissions for the same principal inpatient diagnosis.

[3] Zero- or 1-day stays.

NOTES: PIPDCG is Principal Inpatient Diagnostic Cost Group. PIPDxG is Principal Inpatient Diagnostic Group.

SOURCE: Health Economics Research, Inc., analysis of 1995 and 1996 Medicare claims data, Waltham, MA, 1999.

put them in PIPDxGs that are excluded (although one-third of these are associated with people who have other hospitalizations that do increase their expected costs). Another 21 percent of hospitalizations do not affect payment because their PIPDCG assignment was no higher than a PIPDCG associated with another admission for the same person. Finally, another 2 percent of admissions are ignored because they are associated with stays of less than 2 days (discussed later).

Table 6 shows frequencies and mean expenditures of the PIPDCGs in our 1995-1996 FFS data. In the end, 12 percent of beneficiaries—and 70 percent of beneficiaries hospitalized (with length of stay greater than 1 day) in year 1—were assigned to PIPDCGs that raise payments. Because the PIPDCG model is fully hierarchical and assigns beneficiaries to a single diagnostic category, readmission for the same diagnosis, or rehospitalization for the same or lower-future-cost diagnoses, does not affect PIPDCG assignment. Although incentives for hospitalization are inherent in any inpatient-based risk-adjustment model, the PIPDCG model does not reward multiple hospitalizations.

Demographic Factors

Although the main focus of the PIPDCG model is on using diagnostic information, demographic variables remain important predictors of subsequent-year spending. Demographic information explains a significant amount of variation of spending that is unrelated to observed hospital diagnoses, and hence was included in the model. As explained previously, a beneficiary's relative risk score is determined by adding a demographic factor and a hospital diagnosis factor. The incremental effect of diagnostic category and demographic factors on future expenditures was estimated using linear regression.

Medicare's original capitation payment methodology, the AAPCC, employed the following demographic factors: age, sex, Medicaid enrollment, residence in an institution, and working-aged status (where Medicare is the secondary payer to a private group health insurance plan). We examined all these factors, plus one additional factor, "originally disabled" status.

Table 6

Descriptive Statistics on PIPDCGs, by Group

Group	Frequency	Percentage of Sample	Percentage of Those Hospitalized1 in 1995	Mean 1996 Expenditures	Standard Error of the Mean
Entire Sample	1,387,105	100.00	—	5,186	12
Those with at Least One Admission[1]	241,495	17.41	100.00	11,472	47
PIPDCG 4[2]	1,217,773	87.79	—	4,162	11
PIPDCG 5	1,677	0.12	0.69	5,897	292
PIPDCG 6	1,640	0.12	0.68	6,489	338
PIPDCG 7	1,771	0.13	0.73	7,406	315
PIPDCG 8	25,977	1.87	10.76	8,628	116
PIPDCG 9	21,077	1.52	8.73	9,540	139
PIPDCG 10	14,226	1.03	5.89	10,366	158
PIPDCG 11	21,012	1.51	8.70	11,427	161
PIPDCG 12	26,592	1.92	11.01	13,124	142
PIPDCG 14	7,016	0.51	2.91	15,102	317
PIPDCG 16	29,378	2.12	12.17	17,348	178
PIPDCG 18	5,611	0.40	2.32	18,381	415
PIPDCG 20	5,731	0.41	2.37	22,385	473
PIPDCG 23	4,339	0.31	1.80	24,294	641
PIPDCG 26	1,869	0.13	0.77	26,461	975
PIPDCG 29	1,416	0.10	0.59	30,456	1,597
Sum[3] of PIPDCGs 5-29	169,332	12.21	70.12	—	—

[1] With length of stay greater than 1 day.

[2] Contains beneficiaries with no 1995 hospital admissions, excluded admissions only, short-stay admissions only, and certain other low-cost admissions only.

[3] Contains all beneficiaries whose hospital admission (diagnosis) results in a higher capitation payment the following year.

NOTE: PIPDCG is Principal Inpatient Diagnostic Cost Group.

SOURCE: Health Economics Research, Inc., analysis of 1995 and 1996 Medicare claims data, Waltham, MA, 1999.

Age and Sex

The AAPCC used 10 age categories, each split into male and female, for example, female, age 65-69, male, age 35-44. We adopted these cells, with one exception. To reflect the rapidly growing numbers of very old Medicare beneficiaries, we split the AAPCC's single "age 85 or over" category into three: age 85-89, age 90-94, and age 95 or over. Each age category is again stratified by sex (male/female) to produce a total of 24 age/sex categories, as shown in Table 1.

Medicaid Status

Medicaid status is a factor in the AAPCC methodology. Mean Medicare expenditures for Medicare-Medicaid dually eligible beneficiaries are 29 percent higher than predicted by age, sex, and principal hospital diagnosis. Setting accurate relative Medicare risk scores for Medicare-Medicaid enrollees requires an explicit adjustment for the higher expenditures of dually eligible persons. Such an adjustment prices this vulnerable subgroup of Medicare beneficiaries accurately, encouraging health plans to enroll persons who are also eligible for Medicaid in Medicare. Medicaid status is routinely available in HCFA administrative files, and it is relatively immune to manipulation by health plans. (Plans could attempt to enroll in Medicaid as many of their Medicare enrollees as possible to obtain higher Medicare capitation payments. This is not necessarily undesirable.)

Medicaid status does have disadvantages as a risk adjuster. Medicaid eligibility rules vary widely across States, resulting in divergent proportions of the poor enrolled in Medicaid programs. Thus,

Medicaid status is an imperfect indicator of poverty status, which has historically been linked to higher health care costs. The non-Medicaid poor may also incur higher Medicare expenditures, but this group cannot be identified using HCFA's administrative files. Also, it is not entirely clear why Medicaid status predicts higher future health care expenditures. It could be a proxy for aspects of health or functional status that are not captured by other available measures. (One subgroup of Medicaid enrollees, the medically needy, are eligible for Medicaid by virtue of high health care expenses. For this subgroup, Medicaid enrollment is clearly a proxy for poor health.) Or it could be related to socioeconomic characteristics, such as reduced literacy or inadequate social support, of poorer beneficiaries that result in higher health care expenses.

We believe the advantages of including Medicaid status outweigh the disadvantages, so we included it in the PIPDCG model. We conducted several analyses to determine the form of the Medicaid adjustment in the PIPDCG model. The primary issues were:

- Should the adjustment be prospective or concurrent?
- Should the proportion of the year in Medicaid status be taken into account?
- Should the adjustment differ for subgroups of Medicaid beneficiaries?
- Should the adjustment differ by age and sex?

In the old AAPCC, the Medicaid adjustment was based on Medicaid enrollment in the month of Medicare payment. That is, the Medicaid adjustment was concurrent with payment. We adopted an alternative approach: making the Medicaid adjustment prospective. A prospective adjustment is more consistent with the prospective risk-adjustment framework of the PIPDCG model, where adjustment is made for factors observable to health plans in year 1 when they enroll beneficiaries. A prospective approach also has administrative advantages, allowing Medicaid status to be observed and payment rates determined in advance of payment. A prospective approach is consistent with the perspective that Medicaid status is a socioeconomic indicator. (In fact, because Medicaid status adds to costs even in the payment year, Medicaid status does not appear to be measuring substitutions of Medicaid care for Medicare care.) We found empirically that accounting for the number of Medicaid-eligible months in year 1 had little effect on the magnitude of the Medicaid adjustment. Hence, we adopted the simple approach of assigning Medicaid status to any beneficiary with at least 1 month of Medicaid enrollment in the base year.

We investigated whether to differentiate the Medicaid adjustment by four broad categories of reason for Medicaid eligibility. Unfortunately, however, State reporting of such reasons to HCFA is problematic. Inaccurate assignment of dually eligible beneficiaries to eligibility categories is one reason we do not favor differentiating the Medicaid adjustment by eligibility status. In addition, when we estimated separate adjustments using available data, their magnitudes did not vary substantially.[2] This small gain in payment accuracy from making eligibility distinctions is a second reason for adopting a single adjustment for Medicaid status. Should the reporting of eligibility categories improve, a more refined adjustment for Medicaid status could be reconsidered.

With regard to the issue of differentiation by age and sex, we found reasonably large differences in the Medicaid factor by

[2] The lack of larger differences in predicted expenditures by Medicaid eligibility category may be attributable to the inaccurate assignment of beneficiaries to these categories. Random error in assignment can obscure real differences.

age and sex even after addressing differences in illness burden with the PIPDCG factor. To calibrate the Medicaid adjustment, we interacted Medicaid with each of the 24 age and sex cells. Actuarial smoothing was used to specify the Medicaid adjustment for some extreme-age cells with few beneficiaries.

Originally Disabled Status

"Originally disabled" refers to beneficiaries currently entitled to Medicare by age (i.e., 65 years of age or over) but originally entitled to Medicare when under age 65 by disability. This demographic factor was not used in Medicare's AAPCC payment methodology. However, originally disabled beneficiaries are 43 percent more expensive than predicted by age, sex, and principal hospital diagnoses. The advantages of making an explicit adjustment for originally disabled status are similar to the advantages of adjusting for Medicaid. Medicare managed care should be paid a fair price so that members of this vulnerable group will be as attractive to health plans as other Medicare enrollees. Originally disabled status is routinely available from HCFA administrative files and relatively immune to manipulation by health plans.

An additional advantage of an adjustment for originally disabled status is that it establishes an appropriate age profile of relative risk factors for the Medicare disabled population. If an adjustment is not made, the relative risk factor for a disabled Medicare beneficiary falls when that beneficiary turns 65 (because the average disabled beneficiary of age 64 has higher health care needs than the average new enrollee of age 65 to Medicare). With adjustment, relative risk factors rise appropriately as disabled beneficiaries age. This effect can be seen in Table 1, where, for example, the demographic factor for a disabled male age 64 is 0.760, while for an originally disabled male age 65 it is 0.956 = 0.541 + 0.415.

Just as Medicaid status is an imperfect proxy for poverty, originally disabled is an imperfect proxy for disability in the elderly Medicare population. It does not capture beneficiaries who become disabled after age 64. Also, qualifying for Medicare disability benefits requires a total inability to work for at least 2 years, and eligibility rules for disability benefits and their interpretation vary over time. Finally, not all individuals who qualify for benefits actually apply for them. In short, although originally disabled status identifies an important group of higher cost Medicare beneficiaries, it is not a comprehensive population-based measure of disability.

We included originally disabled status in the PIPDCG model. Its effects are differentiated by age/sex cell through interaction terms in a regression model and subsequent actuarial smoothing of cells with small sample sizes. We investigated interaction effects between originally disabled and Medicaid but found that they were not significant. Hence, the originally disabled and Medicaid adjustments are additive. The incremental payments associated with being originally disabled decrease with increasing age.

Working-Aged Status

Though historically termed "working-aged," this factor is more accurately a measure of Medicare's financial liability rather than the employment status of the beneficiary. When a Medicare beneficiary is enrolled in a private group health insurance plan, by law the private plan is the primary payer for the beneficiary's health care. Medicare will only pay for Medicare-covered services that are not covered by the private plan or for cost-sharing

imposed by the private plan that exceeds Medicare cost-sharing. Beneficiaries with private group health insurance are said to be in working-aged status, because typically they obtain private insurance through employment or the employment of a spouse. (Working-aged status refers only to beneficiaries entitled to Medicare by age [i.e., those who are 65 years of age or over]. By definition, the Medicare disabled population under age 65 is rarely employed.) Working-aged was an adjustment factor under HCFA's AAPCC capitation methodology. Unlike Medicaid status, it is inherently a year 2 (payment year) adjuster, because Medicare expenditures are lower during the months when Medicare is a secondary payer. Medicare expenditures for working-aged beneficiaries may also be lower because employment may be a marker for better health status. (As previously noted, not all working-aged beneficiaries are employed, because some beneficiaries may obtain private insurance coverage through their spouses.)

We developed a multiplicative adjustment for working-aged status. The multiplier is applied to the relative risk factor derived for a beneficiary as if he or she were not in working-aged status. The multiplier scales the risk factor downward for the lower expenditures of beneficiaries with private insurance. We used the data to determine empirically a working-aged factor of 0.21. Our calculation creates annual payments that are correct, on average, assuming that beneficiaries receive the full amount suggested by their PIPDCG factor in months that Medicare is the primary payer, and 21 percent of that amount in months where they have working-aged status.

The relative risk factor for a male beneficiary age 68 who is not enrolled in Medicaid, was not previously entitled by disability, and was not hospitalized in the base year is 0.541. If this male is employed with private group health insurance, his relative risk factor is reduced to 21 percent of 0.541, or 0.114. That is, this beneficiary is expected to cost Medicare only 11 percent as much as the average beneficiary.

Our estimate of a working-aged factor is based on FFS experience in the traditional Medicare program, the only nationally representative data available. Describing the costs of working-aged Medicare beneficiaries enrolled in Medicare managed care plans is a subject for future research. If these beneficiaries are also enrolled through their employer in private managed care plans with comprehensive benefits, their Medicare-covered costs may be even lower than we have estimated.

Institutional Status

The final demographic factor we considered was institutional status, another factor used in calculating Medicare's AAPCC. If a beneficiary is a resident of a qualifying institution on the last day of a month and for the 29 prior consecutive days, the beneficiary qualifies for the higher institutional payment rate in the following month. Institutional status is thus primarily a concurrent (year 2), rather than prospective (year 1) adjustment. Institutional status is not routinely available in Medicare administrative files, but is reported by health plans for managed care enrollees to HCFA.

Because institutional status is not available in HCFA's claims or enrollment files, we used the Medicare Current Beneficiary Survey (MCBS) to analyze expenditures of institutionalized beneficiaries. The MCBS is a population-based in-person panel survey of about 12,000 Medicare beneficiaries per year. Medicare claims data for MCBS respondents are merged to the survey files. Detailed information on institutional status is collected by the survey. We ana-

Table 7

Actual Compared with Predicted Expenditures, by Institutional Status

Group Status	Observations	Average Annual Months Institutionalized	Mean Annualized Expenditures		Predictive Ratio[2]
			Actual	Predicted[1]	
Entire Sample	32,228	—	5,186	5,186	1.00
Neither Medicaid nor Institutionalized	25,369	—	4,547	4,723	1.04
All Institutionalized	2,715	—	8,570	8,534	1.00
Nursing Home Only	2,044	10.28	6,371	8,766	1.38
All Other Situations	671	7.52	18,412	7,496	0.41
SNF Only	77	1.32	46,159	7,209	0.16
SNF/Nursing Home	261	6.60	33,364	10,278	0.31
Mixed	27	8.60	34,315	8,833	0.26
ICF/MR	195	11.45	3,040	4,407	1.45
Mental Health	82	10.85	6,312	6,268	0.99
Hospital	21	9.88	3,463	7,356	2.12
Rehabilitation	8	10.83	8,313	8,613	1.04
Medicaid	5,471	—	7,131	7,612	1.07
Medicaid and Institutionalized	1,327	—	7,617	9,164	1.20
Not Medicaid and Institutionalized	1,388	—	9,605	7,849	0.82

[1] By PIPDCG model.

[2] Predicted divided by actual expenditures.

NOTES: PIPDCG is Principal Inpatient Diagnostic Cost Group model. SNF is skilled nursing facility. ICF/MR is intermediate care facility for the mentally retarded.

SOURCE: Health Economics Research, Inc., analysis of 1991-1994 Medicare Current Beneficiary Survey data, Waltham, MA, 1999.

lyzed expenditures in 3 years: 1992, 1993, and 1994. Expenditures in each year were normalized to the same mean, $5,186, as in our 5-percent sample. There were 32,228 observations in our MCBS sample. Observations are person-years, many of which represent the same person in multiple years. Expenditure data for beneficiaries who exited from the survey were obtained from Medicare claims files. Sample selection criteria and variable definitions for our MCBS analysis were essentially equivalent to those for our analysis of the Medicare 5-percent sample. Results are weighted using the product of the survey sampling weight and (for persons who died) the fraction of the year eligible for Medicare. For the institutionalized, expenditures are also weighted by the fraction of the year institutionalized.

Table 7 compares mean actual and predicted expenditures in the MCBS sample by institutional and Medicaid status. Expenditures are predicted using the PIPDCG model, including age, sex, principal inpatient diagnoses, and Medicaid and originally disabled statuses. For the institutionalized population as a whole, the PIPDCG model (excluding institutional status) predicts mean Medicare expenditures accurately. It is not necessary to include an explicit, separate adjustment for institutional status to ensure that the PIPDCG model reflects the average expenditures of this vulnerable group accurately.

Expenditures are not predicted accurately for important subgroups of the institutionalized, however. Expenditures are overpredicted by 38 percent for residents of nursing homes, most of whom are long-term residents (average months institutionalized per nursing home resident per year is 10.3). These long-term residents may have significant health care costs, but many of these costs are the responsibility of Medicaid or private insurance rather than Medicare, or are paid out of pocket. Conversely, expenditures are substantially underpredicted for beneficiaries who reside in a SNF. Beneficiaries with a SNF

stay are presumably post-acute care patients, because Medicare requires a hospitalization prior to a covered SNF stay. (Average months institutionalized per year for SNF only is 1.3, and for SNF/nursing home is 6.6.) The long-term versus post-acute dichotomy is also reflected in the overprediction of expenditures for institutionalized Medicaid beneficiaries versus the underprediction of expenditures for beneficiaries not enrolled in Medicaid. As beneficiaries reside longer in nursing homes, they are more likely to become impoverished and qualify for Medicaid.

It is clear that the Medicare institutionalized population contains two distinct subpopulations, a very expensive post-acute care population receiving SNF care and a much-less-expensive long-term care population in nursing homes. An adjustment for the long-term nursing home population would be negative because Medicare expenditures are overpredicted for this group. It seems undesirable to discourage health plans from enrolling this vulnerable group through a negative adjustment. Also, as a practical matter, HCFA would rely on plans to self-report the data used to reduce payments—a problematic scenario. Although use of post-acute care SNF services in the payment year is an indicator of illness severity, it is inconsistent with a prospective, diagnosis-based capitation model to increase payments for current-year service use. Because the SNF stay is precipitated by a hospitalization, the prospective PIPDCG will identify some of these high-cost beneficiaries. For these reasons, we do not favor adjustments for institutionalized subpopulations. No adjustment for institutional status is included in the PIPDCG risk-adjustment model.

Excluding Short Hospital Stays

Because of concerns that health plans may overadmit patients in order to increase payments, we explored the sensitivity of the model to including and excluding diagnoses from short hospital stays. The concern is that payments based only on hospitalizations may give health plans an incentive to hospitalize enrollees to increase future payments. In its utilization review of hospital admissions, the health plan weighs the expected marginal revenue and marginal costs of the hospitalization along with the benefits of treatment to the patient. Even if plans are largely delivering appropriate care, raising the marginal cost of getting a hospital diagnosis that "counts" for risk adjustment reduces the incentive to hospitalize. Requiring hospitalizations of at least 2 days raises the marginal cost of getting a hospital diagnosis counted and discourages this undesirable strategy.

Excluding short-stay diagnoses has advantages and disadvantages. To the extent that length of stay is a proxy for severity of illness, short-stay patients are less severely ill and less expensive in the future than longer stay patients. Excluding short-stay diagnoses does not degrade very much the accuracy of the risk-adjustment model in predicting future expenses. Short-stay hospitalizations may be highly substitutable for certain diagnostic and therapeutic procedures, drug administration, observation, screening, and rule-out diagnoses that are often appropriately performed in the outpatient setting. Attaching a substantially greater future payment to inpatient, as opposed to outpatient, testing and procedures gives health plans a strong incentive to perform these activities in the hospital. Excluding short-stay diagnoses limits this incentive. Finally, excluding diagnoses arising from short stays eliminates penalties to health plans that avoid unnecessary short-stay admissions; models with such exclusions may thus provide a more level playing field among plans.

Excluding short-stay diagnoses also has disadvantages. Although less expensive in the future than longer stay patients, short-stay patients are more expensive than the non-hospitalized patients. Excluding short stays reduces the risk-adjustment model's predictive accuracy. Although excluding short-stay diagnoses lessens incentives to admit, paying more only when diagnoses come from longer stays gives health plans an incentive to increase length of stay. Manipulating length of stay is probably easier than increasing the admission rate. Excluding short-stay diagnoses penalizes plans that are effective in reducing inappropriately long hospital stays. Short stays are often clinically appropriate and indicated. Clinical review of the appropriateness of hospital admissions—such as by the peer review organizations under Medicare's PPS for hospitals—is a more refined approach to excluding inappropriate admissions than a blanket exclusion of diagnoses from short stays.

We developed empirical evidence on the frequency and future costs associated with short-stay admissions in our 1995-1996 FFS 5-percent sample Medicare data. We defined short-stay admissions as zero-day stays (i.e., the same admission and discharge dates), or 1-day overnight stays (i.e., discharge date 1 day later than admission date).[3] In assigning beneficiaries to PIPDCG 5 or above (thus raising a beneficiary's relative risk factor), after removing stays with excluded diagnoses and multiple admissions per beneficiary, only 2.1 percent of total 1995 admissions were excluded because they were short stays (Table 5). As for costliness, we found that the future expenses of beneficiaries whose PIPDCG is assigned by a 1-day stay is 17 percent lower than predicted by the PIPDCG model. So the future expenditures associated with short stays are significantly lower than future expenditures associated with longer stays. However, the future expenditures of beneficiaries with short-stay hospitalizations are greater than future expenditures of the never-hospitalized. Excluding diagnoses from short stays reduced the predictive power (R^2) of the PIPDCG model from 5.72 to 5.69 percent.

HCFA made a policy decision to exclude diagnoses from short stays in the PIPDCG payment model. For the initial implementation of risk adjustment, this approach focuses on the more severely ill longer stay patients, while eliminating the sensitivity of payments to increased short-stay admissions of less ill patients. The short-stay patients are included in the lowest paying DCG, along with those who are not hospitalized and those with excluded diagnoses. Ignoring short stays raises the coefficients on the demographic variables slightly.

REGRESSION MODEL

The final PIPDCG model includes the following risk factors to explain future Medicare expenditures: 24 age/sex cells, Medicaid status interacted with age/sex cells, originally disabled status interacted with age/sex cells, working-aged status, and the 16 PIPDCG diagnostic categories assigned from prior-year principal hospital diagnoses. The incremental effects of these beneficiary characteristics on Medicare expenditures were estimated in a linear multiple regression model, with annualized 1996 Medicare expenditures as the dependent variable. All risk factors except for working-aged were used as predictors. Beneficiaries with working-aged months in 1995 or 1996 were excluded from the regression sample, and a multiplicative adjustment for working-aged status was determined as already described.

[3] In Medicare's PPS for hospitals, length of stay is determined by adding 1 to discharge date minus admission date. Thus, what we call "zero-day stays" would be referred to as "1-day stays" in PPS terminology, and our "1-day stays" in PPS are "2-day stays."

The regression was estimated on the 1995-1996 5-percent random sample of Medicare beneficiaries, representing 1,387,105 individuals, and was weighted for months of beneficiary Medicare eligibility in 1996. After actuarial smoothing of a few demographic coefficients estimated from small sample sizes in the extreme age ranges, the relative risk factors shown in Table 1 were derived by dividing the regression coefficients by mean expenditures. The population mean of $5,100 was used rather than the sample mean of $5,186 for this calculation. (HCFA provided the mean expenditures for the full FFS Medicare population. The sample used for model development did not include new enrollees because a full year of data was not available for them.) For example, the estimated coefficient of PIPDCG 6, which contains the diagnosis prostate cancer, was $2,333. This coefficient means that, holding constant a beneficiary's age, sex, Medicaid status, and originally disabled status, a 1995 hospitalization for prostate cancer was associated with an average $2,333 higher Medicare expenditures in 1996 (assuming the hospitalization for prostate cancer was the beneficiary's only 1995 hospitalization or that any other hospitalizations were assigned to lower ranked PIPDCGs). Dividing this coefficient by the mean (2,333/5,100) yields the relative risk factor of 0.458 (or about 46 percent of the average expenditure) for PIPDCG 6 shown in Table 1. This hospitalization risk factor is added to the beneficiary's demographic risk factor to determine his or her total relative risk factor.

EVALUATION OF MODEL

After developing the PIPDCG payment model, we evaluated its predictive accuracy and stability. We used two samples to judge predictive accuracy and stability: our 5-percent 1995-1996 random sample of Medicare beneficiaries (n = 1,387,105), and our 1991-1994 sample of beneficiaries from the MCBS (n = 32,021 person years). The 5-percent sample used to judge predictive accuracy is the same as our model development sample. We did not reserve a portion of our sample for validation because similar PIPDCG models have been extensively validated in previous work (Ellis et al., 1996). Because hospitalizations for many diagnoses, even in the Medicare population, are relatively rare events, we felt that it was more important to maximize the sample available for model development. In addition, our previous work has shown that with the large sample sizes we use, predictive accuracy calculated for model development and validation samples is similar, although estimation sample statistics may slightly overstate the performance of the model in new data (Ellis et al., 1996).

The MCBS sample, on the other hand, is an independent validation sample. But statistics computed from it are subject to much greater random variation because its sample size is so much smaller. We prefer predictive ratios from the much larger 5-percent sample and present them where possible. We focus on predictive accuracy for groups not defined as risk factors in the PIPDCG model. It is not surprising, especially for the 5-percent estimation sample, that the model performs well for groups defined by risk factors included in the model, and we do not present such results here (refer to Pope et al., 1999, for these statistics).

Predictive Accuracy

We present two statistics to judge model predictive accuracy: the R^2 statistic and predictive ratios. R^2 is the percentage of total variance in Medicare expenditures that is predicted by the risk-adjustment model. It measures the predictive accura-

cy of the model for individual beneficiaries. A predictive ratio is the ratio of mean predicted expenditures to mean actual expenditures for a subgroup of beneficiaries. The predictive ratio measures the accuracy of model prediction for groups of beneficiaries, with a ratio closer to 1.00 indicating better prediction.

R^2 Statistic

The estimation R^2 of the PIPDCG model was 6.2 percent. The R^2 of a demographic model including age, sex, Medicaid status, and originally disabled status was 1.5 percent. Thus, the PIPDCG model represents a fourfold improvement in predictive accuracy over a demographic model in our data. Models including diagnoses from all care settings and accounting for the presence of multiple clinical conditions per person have achieved R^2 values as high as approximately 9 percent (Ellis et al., 1996). The PIPDCG model, relying only on the single most predictive principal inpatient diagnosis, achieves about two-thirds of the performance of models using data from all settings and multiple conditions.

The theoretical maximum R^2 for a prospective risk-adjustment model (i.e., a model predicting expenditures based on prior-year individual characteristics) has been estimated to be 20-25 percent (Newhouse et al., 1989). Hence, the PIPDCG model appears to predict about one-quarter to one-third of the potentially explainable variance in expenditures. Although this may seem low, no prospective risk-adjustment model that predicts anywhere close to the theoretical maximum has been devised without using variables that directly measure prior utilization. Also, increasing next year's payments because more services were used this year rewards plans because they spend more money—a feature that is inconsistent with health-risk-based payments. A Medicare risk-adjustment model only needs to predict expenditures approximately as well as health plans can, so that plans find that they receive payments that equal their expected costs for enrollees.

Still, the R^2 values of prospective risk-adjustment models are quite low in absolute terms (i.e., relative to 100 percent, or perfect prediction) and modest relative to the supposed theoretically attainable maximums. The low R^2 values remind us that there is considerable variability in medical expenses due to the random onset of acute illness that must remain the province of insurance risk-pooling. The PIPDCG and other prospective risk-adjustment models cannot be expected to, and do not, predict expenditures accurately for individual beneficiaries. There is also substantial room for improvement in risk-adjustment models, much of which probably will have to await the availability of more clinically detailed and precise data on beneficiaries, such as their functional status or severity of illness within specific diagnostic groups. In the meantime, health plans (or beneficiaries) possessing more accurate information can engage in profitable biased selection against capitation payments incorporating even health-based risk adjustment. Imperfect clinical risk adjustment, nevertheless, is better than no clinical risk adjustment, as even imperfect risk adjustment will limit selection opportunities and may be "good enough" to deter the most flagrant and injurious forms of risk selection.

The R^2 is a measure of the proportion of individual variability that is explained by a risk-adjustment model. It can be an overly pessimistic measure, however, because health plans are generally only able to increase their enrollment within categories of similar individuals (for example, by advertising the excellence of their care to cardiac patients) and cannot easily specifically recruit, say, only the healthier cardiac

cases. To better understand the ability of the PIPDCG model to correctly predict payments for groups of enrollees rather than each individual, we examined its predictive ratios for selected groups of enrollees. As in Ash et al. (1989), the predictive ratios are defined as the ratio of predicted to actual spending for patients in specific clinical groups.

Predictive Ratios

Table 8 shows predictive ratios from two models: (1) a demographic model that uses age, sex, Medicaid status, and originally disabled status to predict expenditures; and (2) the PIPDCG model, which includes, in addition to these demographic factors, prior-year principal inpatient diagnoses grouped into the 16 PIPDCG categories. We compare the PIPDCG model with a demographic model to show the effect of adding a health-status measure (prior-year hospital diagnoses) to demographic risk adjusters. Ratios are shown for groups defined by prior-year Medicare expenditure percentiles, number of prior-year hospital admissions, chronic conditions diagnosed during inpatient or ambulatory encounters, beneficiary self-rated general health status, and beneficiary self-reported difficulty in activities of daily living (ADLs), a measure of functional status. (ADLs include eating, bathing, dressing, using the toilet, walking, and getting in and out of chairs.) The latter two groups can be defined only for the smaller MCBS survey sample. The other predictive ratios were calculated for the much larger 5-percent 1995-1996 sample. The chronic condition groups were defined by a prior-year diagnosis in any care setting, ambulatory (physician office, hospital outpatient) as well as inpatient.

The PIPDCG model predicts expenditures more accurately than the demographic model for all groups defined by prior-year expenditure percentiles, except for the fourth quintile of prior-year expenditures.[4] The effect of adding hospital diagnoses is most dramatic for the most expensive, and presumably sickest, beneficiaries. For the beneficiaries with the top 1 percent of prior-year expenditures, the demographic model projects expenditures that, on average, are only 19 percent of actual average expenditures. The PIPDCG model improves the average prediction for the most expensive beneficiaries to 47 percent. The improved predictive accuracy of the PIPDCG model is significant, yet it still overpredicts average expenditures by a factor of 2 for beneficiaries in the lowest quintile of prior-year expenditures and underpredicts expenditures for beneficiaries with the top 1 percent of prior-year expenditures by approximately one-half.

The PIPDCG model predicts average expenditures fairly accurately for beneficiaries with no, one, and two prior-year hospitalizations, only underpredicting by a substantial amount for beneficiaries with three or more hospitalizations in the prior year. (Recall that the model uses the principal diagnosis from the single hospitalization most predictive of future expenses.) The PIPDCG model improves substantially upon the demographic model for all prior-year hospitalization categories. Although underpredicting for all chronic disease groups, the PIPDCG model does better for each diagnosis than the demographic model. The PIPDCG model does best relative to demographics when the diagnosis is most likely to be the reason for a hospital admission, such as lung or pancreatic cancer and intracerebral hemorrhage. Its predictive accuracy exceeds that of the demographic model the least for diagnoses with more outpatient-oriented treatment, such as arthritis and hypertension.

[4] Because less than 20 percent of Medicare beneficiaries are hospitalized in a year, the fourth expenditure quintile presumably includes mostly beneficiaries with large outpatient expenditures. These beneficiaries are not distinguished by the PIPDCG model, which includes only hospital inpatient diagnoses.

Table 8

Predictive Ratios[1] for Demographic and PIPDCG Risk-Adjustment Models, by Model and Beneficiary Group

Beneficiary Group	Demographic Model[2]	PIPDCG Model
Prior-Year Expenditures[3]		
First Quintile (Lowest)	2.57	2.09
Second Quintile	1.88	1.54
Middle Quintile	1.35	1.10
Fourth Quintile	0.96	0.84
Fifth Quintile (Highest)	0.47	0.75
Top 5 Percent	0.29	0.61
Top 1 Percent	0.19	0.47
Prior-Year Hospital Admissions[3]		
No Admissions	1.31	1.07
One Admission	0.66	1.02
Two Admissions	0.50	0.91
Three or More Admissions	0.31	0.69
Chronic Conditions[3,4]		
Any Chronic Condition Below	0.84	0.89
Depression	0.59	0.77
Alcohol/Drug Dependence	0.44	0.78
Hypertensive Heart/Renal Disease	0.65	0.81
Benign/Unspecified Hypertension	0.83	0.90
Diabetes with Complications	0.47	0.63
Diabetes Without Complications	0.63	0.73
Heart Failure/Cardiomyopathy	0.51	0.74
Acute Myocardial Infarction	0.47	0.78
Other Heart Disease	0.66	0.80
Chronic Obstructive Pulmonary Disease	0.63	0.79
Colorectal Cancer	0.59	0.78
Breast Cancer	0.75	0.81
Lung/Pancreatic Cancer	0.35	0.61
Other Stroke	0.53	0.74
Intracerebral Hemorrhage	0.42	0.73
Hip Fracture	0.59	0.83
Arthritis	0.79	0.84
Self-Rated General Health Status[5]		
Poor	0.54	0.67
Fair	0.81	0.86
Good	1.03	1.01
Very Good	1.36	1.27
Excellent	1.74	1.57
Functional-Status[5,6] **Difficulty in:**		
5-6 ADLs[7]	0.61	0.74
3-4 ADLs[7]	0.69	0.76
1-2 ADLs[7]	0.83	0.85
None	1.33	1.26

[1] Mean predicted expenditures for a group divided by mean actual expenditures.

[2] Includes age/sex, Medicaid, originally disabled.

[3] Calculated from 5-percent 1995-1996 Medicare sample.

[4] Defined as beneficiaries with a 1995 diagnosis on a Medicare hospital inpatient, outpatient, physician, or other professional health claim.

[5] Calculated from 1991-1994 Medicare Current Beneficiary Survey sample.

[6] Measured as difficulty with activities of daily living (ADLs).

[7] Activities of daily living include eating, bathing, dressing, using the toilet, walking, and getting in and out of chairs.

NOTE: PIPDCG is Principal Inpatient Diagnostic Cost Group.

SOURCE: Health Economics Research, Inc., analysis of 1995-1996 Medicare data, and 1991-1994 Medicare Current Beneficiary Survey, Waltham, MA, 1999.

The PIPDCG model improves demographic predictions for all self-rated general health-status and functional-status groups. The proportional improvement is the greatest for beneficiaries reporting the worst (poor) health or the most functional

limitations (difficulty in performing five or six ADLs). Nevertheless, as with chronic conditions, the model still underpredicts for beneficiaries reporting the worst health or the most functional limitations.

Overall, the impression from the predictive ratios is that the PIPDCG model improves predictive accuracy significantly compared with a demographic model for almost all groups but that it falls well short of predicting future expenditures accurately in an absolute sense for many groups. If predictive accuracy were the only criterion for a risk-adjustment model, we would improve the predictions of the PIPDCG model by incorporating additional information into the model, such as that used to define some of the evaluation groups. However, in the design of the PIPDCG model, predictive accuracy was sometimes sacrificed to improve the behavioral incentives of the model. For example, accounting for multiple prior-year hospitalizations or higher prior-year expenditures would reward health plans that rehospitalized enrollees or were inefficient in their expenditures. Moreover, some types of information that can be used to improve predictive accuracy, such as base-year expenditures, ambulatory diagnoses, and survey health-status measures, are not currently available for most Medicare managed care enrollees. (HCFA intends to collect encounter data from all care settings, which will provide ambulatory diagnoses and imputed expenditures, as soon as practical. And HCFA does collect survey health-status measures for a sample of enrollees in each Medicare managed care plan through its Health Outcomes Survey.)

Stability

In addition to predictive accuracy, a desirable property of a risk-adjustment model is stability. The model's predictions should be stable, replicable, and predictable from year to year for a population that is not changing over time. A county's Medicare enrollees or the enrollees of a long-established health plan might constitute a stable population. The amount of instability or random variation in a population's average relative risk score depends upon the size of the population. The incidence of acute illness is largely random, and in a small population, the number and type of hospitalizations, and hence PIPDCG assignments, will vary from year to year.

We analyzed the amount of random variation in mean risk scores as a function of populations of different sizes, such as the enrollees of small and large health plans. Using the normal distribution approximation, the 95-percent confidence interval for the mean relative risk score is given by

$$RRS \pm 1.96*CV/\sqrt{n},$$

where RRS = the calculated mean relative risk score for a population, CV = the coefficient of variation of risk-adjustment model predictions, and n = the number of beneficiaries in the population. Table 9 tabulates 95-percent confidence-interval factors for demographic and PIPDCG models for populations of different sizes. For example, if the mean demographic risk score for a health plan with 1,000 enrollees is 1.05, the 95-percent confidence interval is 1.050 $\pm$ 2.2 percent (that is, from 1.028 to 1.072).

The formula given in Table 9 shows that the 95-percent confidence interval is proportional to the coefficient of variation (CV) of predictions. Because its predicted values have a wider range, the CV of PIPDCG model predictions is twice as large as the CV of demographic model predictions. (Note that this increased CV is a strength, not a weakness, of the PIPDCG model; if diagnostic information were a perfect predictor of future resource use, then

Table 9

95-Percent Confidence Intervals for Relative Risk Scores, by Model and Population Size

Item	Demographic Model[1]	PIPDCG Model
Coefficient of Variation*100	34.717	68.295
Population Size[2]	Percent[3]	
50	±9.6	±18.9
100	±6.8	±13.4
500	±3.0	±6.0
1,000	±2.2	±4.2
5,000	±0.96	±1.89
10,000	±0.68	±1.34
50,000	±0.30	±0.60
100,000	±0.22	±0.42
500,000	±0.10	±0.19
1,000,000	±0.07	±0.13

[1] Includes age/sex, Medicaid, and originally disabled.

[2] For example, number of health plan enrollees, or county residents.

[3] Percentage points plus or minus for a population's mean score.

NOTES: Calculated as 1.96*(CV/SQRT(Population Size)). For example, if the mean demographic risk score for a health plan with 1,000 enrollees is 1.050, the 95-percent confidence interval is (1.028, 1.072). PIPDCG is Principal Inpatient Diagnostic Cost Group.

SOURCE: Health Economics Research, Inc., analysis of 1995-1996 Medicare data, Waltham, MA, 1999.

the CV would be the same as that of total spending.) For any population size, the 95-percent confidence interval of the PIPDCG prediction is approximately twice as large as the confidence interval of the demographic prediction. That is, because it incorporates an additional predictive factor, hospital diagnoses, predictions from the PIPDCG model are subject to more random variation than demographic predictions. Nevertheless, in absolute terms, the degree of random variation is small for expected real-world health plan sizes. For a 5,000-person plan, the random variation in PIPDCG predictions is $\pm$ 1.9 percent, for a 10,000-person plan it is $\pm$ 1.3 percent, and for a 50,000-person plan, it is only $\pm$ 0.6 percent.

Nevertheless, for very small populations, say of under 1,000, random variation in mean PIPDCG risk scores is significant. Might it be better for very small populations to use a demographic risk adjuster than the PIPDCG adjuster? A common measure of predictive accuracy of models is the mean square error, or MSE. A well-known relationship from statistics (e.g., Mood, Graybill, and Boes, 1974) is that

$$MSE = variance + (bias)^2$$

Inaccuracy in predictions due to random small sample error is captured by the variance component of the MSE. Inaccuracy due to systematic misprediction by the model is captured by the $(bias)^2$ term. In our application, systematic misprediction would occur because a demographic model would not accurately capture the health status of a population, no matter how large the sample size. In large populations, the variance of the mean risk scores approaches zero, and the MSE equals the square of the bias. In large samples, clearly, an unbiased health-status-based risk-adjustment model predicts more accurately than a (biased) demographic model. But in very small samples, a demographic model, despite its bias, might have an advantage in MSE predictive accuracy because of the lower variance of its predictions.

We simulated the MSEs of the PIPDCG model and an age/sex demographic model for various sample sizes, degrees of bias, and mean risk scores. The simulations showed that, unless the bias in the demographic (age/sex) model is very small (e.g., 1 percent), the PIPDCG prediction is expected to be more accurate for any realistic plan size. For example, if the bias in the age/sex model is 5 percent, the mean PIPDCG risk scores have a smaller MSE than the mean age/sex risk score for populations (plans) with 250 or more members (enrollees).

To confirm our stability results empirically, we used the MCBS to simulate and compare national and regional mean risk scores for each of the 3 years available in

our sample: 1992, 1993, and 1994. The MCBS provides a reasonable representation of a plan because it is a longitudinal survey with substantial year-to-year overlap but replenishment (turnover) for death and attrition, with the goal of continuously representing the national Medicare population. The national MCBS sample size is about 12,000 per year, simulating enrollees in a small-to-moderate size health plan, while the regional sample sizes range from about 1,500 to 4,000, simulating enrollees in very small to small health plans. The results were consistent with predictions from the simulations. The PIPDCG model was slightly more unstable than demographic models, especially in the very small regional samples, but the degree of instability was well within acceptable bounds.

CONCLUSION

The PIPDCG model is a conservative risk-adjustment model. In simulations with Medicare FFS data, it adjusts payments for only 12 percent of enrollees based on their health status. To measure health status, it utilizes only principal inpatient diagnoses, which are the most widely available and highest quality diagnoses. The PIPDCG model focuses on beneficiaries hospitalized for serious illnesses with longer hospital stays. For the initial implementation of risk adjustment, the focus on the most severely ill and expensive beneficiaries and the most available and high-quality data seems appropriate. As more experience is gained with risk adjustment and its data and operational requirements, more refined and comprehensive risk-adjustment models can be implemented.

The PIPDCG model's exclusive use of inpatient diagnoses raises some issues related to incentives and fairness. Health plans have an incentive to admit enrollees to have diagnoses counted, and plans that avoid unnecessary admissions may be penalized. More fully understanding these issues will require analysis of managed care encounter data and observing plans' behavioral responses to implementation of the PIPDCG model.

Several factors mitigate concern over the incentive to admit. HCFA has announced the PIPDCG model as transitional to full-encounter risk adjustment, which HCFA intends to implement as soon as feasible. Will plans significantly alter their behavior in response to the incentives of a transitional system that may be in place for only a few years? Moreover, in the first year of implementation, 2000, clinically risk-adjusted payments will be only 10 percent of total payments. A plan is not guaranteed a higher future payment if it admits an enrollee. The enrollee may die or disenroll before the next payment period. The uncertain higher future payment coupled with the certain immediate cost of hospitalization also limits the incentive to unnecessarily hospitalize enrollees. The incentives of most non-Medicare insurers continue to work against hospitalization, perhaps too much so. Will plans and providers attempt to differentiate their behavior for Medicare versus non-Medicare enrollees? Finally, we have designed the PIPDCG model with concerns about incentives in mind. Specifically, none of the following lead to increased PIPDCG-based payments: short-stay hospitalizations; hospitalizations for diagnoses that may be minor, transitory, or non-specific; rehospitalizations; or multiple hospitalizations.

When the DRG PPS for Medicare hospitalizations was implemented in the early 1980s, there was considerable concern about inappropriate increases in admissions. Providers could benefit from the lump sum per admission payment. Yet the expected increases in admission rates

never materialized (Medicare Payment Advisory Commission, 1998). Health plan admission behavior will need to be carefully monitored by HCFA, just as hospital admissions under the PPS were carefully monitored by the HCFA-funded peer review organizations.

Virtually all policy analysts agree that health-based risk adjustment promotes the successful long-term operation of competitive capitated health insurance markets. Health-based risk adjustment has been long advocated and researched but is only starting to be implemented. Medicare's implementation of the PIPDCG model will provide valuable experience with real-world risk adjustment.

ACKNOWLEDGMENTS

The authors thank anonymous reviewers for helpful comments on this article.

REFERENCES

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

Ellis, R.P., and Ash, A.: Refinements to the Diagnostic Cost Group Model. *Inquiry* 32:1-12, Winter 1995.

Ellis, R.P., Pope, G.C. , Iezzoni, L.I., et al.: Diagnosis-Based Risk Adjustment for Medicare Capitation Payments. *Health Care Financing Review* 12(3):101-128. 1996.

Iezzoni, L.I., Ayanian, J.Z., Bates, D.W., Burstin, H.: Paying More Fairly for Medicare Capitated Care. *New England Journal of Medicine* 339(26):1933-1937, December 24, 1998.

Ingber, M.J.: Implementation of Risk Adjustment for Medicare. *Health Care Financing Review* 21(3):119-126, Spring 2000.

Luft, H.S.: Health Maintenance Organizations and the Rationing of Medical Care. *Milbank Memorial Fund Quarterly/ Health and Society* 60:268-306. 1987.

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

Miller, R.H., and Luft, H.S.: Does Managed Care Lead to Better or Worse Quality of Care? *Health Affairs* 16(5):7-25, 1997.

Mood, A.M., Graybill, F.A., and Boes, D.C.: *Introduction to the Theory of Statistics. Third Edition*. New York: McGraw-Hill Book Company. 1974.

Newhouse, J.P.: Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection. *Journal of Economic Literature* 34:1236-1263, 1996.

Newhouse, J.P., Manning, W.G., Keeler, E.B. and Sloss, E.M.: Adjusting Capitation Rates Using Objective Health Measurers and Prior Utilization. *Health Care Financing Review* 10(3):41-54, 1989.

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

Pope, G.C., Liu, C.F., Ellis, R.P., et al.: *Principal Inpatient Diagnostic Cost Group Models for Medicare Risk Adjustment.* Final Report to the Health Care Financing Administration under Contract No. 500-95-048. Waltham, MA: Health Economics Research, February, 1999.

Public Health Service and Health Care Financing Administration: *International Classification of Diseases, 9th Revision, Clinical Modification*. U.S. Department of Health and Human Services. Washington, DC. U.S. Government Printing Office, September 1980.

Weiner, J.P., Dobson, A., Maxwell, S., et al.: Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses. *Health Care Financing Review* 17:77-100, 1996.

Reprint Requests: Gregory C. Pope, Health Economics Research, Inc., 411 Waverley Oaks Road, Suite 330, Waltham, Massachusetts 02452. E-mail: GPope@her-cher.org