
Improving Health-Based Payment for Medicaid Beneficiaries: CDPS

Richard Kronick, Ph.D., Todd Gilmer, Ph.D., Tony Dreyfus, M.C.P., and Lora Lee, M.S.

This article describes the Chronic Illness and Disability Payment System (CDPS), a diagnostic classification system that Medicaid programs can use to make health-based capitated payments for TANF and disabled Medicaid beneficiaries. The authors describe the diversity of diagnoses and different burdens of illness among disabled and AFDC Medicaid beneficiaries. Claims from seven States are analyzed, and payment weights are provided that States can use when adjusting HMO payments. The authors also compare the taxonomy and statistical performance of CDPS to other leading diagnostic classification systems and find that the new model performs better in a number of respects.

INTRODUCTION

In previous work, we argued that health-based payment for Medicaid beneficiaries with disabilities is both important and feasible (Kronick, Zhou, and Dreyfus, 1995; Kronick et al., 1996). Among people with disabilities, health expenditures are strongly related to recent diagnoses, and health plans are well aware that attracting too many people with costly problems can lead to large financial losses. If a State Medicaid program does not pay more to health plans whose members have above-

average levels of need, it will penalize plans attractive to people with greater needs and jeopardize quality of care. The greater predictability of expenditures among people with disabilities compared with a general population both increases the importance of health-based payment and makes it easier to do well. The strong relationship between diagnoses and future expenditures allows Medicaid programs to use diagnoses to make good predictions of health care needs.

We can now report that Medicaid programs have been leaders in the implementation of health-based payment (Table 1). Using diagnoses from both ambulatory and inpatient encounters, Maryland implemented risk adjustment in May 1997, Colorado in July 1997, Oregon in June 1998, and Delaware in January 2000. Using inpatient data only, Utah implemented a limited version of health-based payment in June 1998. Utah is planning on expanding to full diagnostic risk adjustment if the encounter data supplied by health maintenance organizations (HMOs) are of sufficient quality. Minnesota implemented health-based payment in January of 2000, adjusting 5 percent of the capitation based on diagnostic case mix, with the remaining 95 percent based on traditional demographic rate cells. Michigan is using diagnostic adjustment as part of its competitive procurement process; Michigan divides a plan's bid by that plan's case mix in order to compare bids against each other on an equitable basis. New Jersey has announced its intentions to implement

Richard Kronick, Todd Gilmer, and Lora Lee are with the Department of Family and Preventive Medicine, University of California, San Diego. Tony Dreyfus is an independent consultant. This research was supported by grants from the Robert Wood Johnson Foundation and from the Health Care Financing Administration (HCFA). The views expressed in this article are those of the authors and do not necessarily reflect the views of the University of California or HCFA.

Table 1
Medicaid Health-Based Payment Activity

State	Population Covered	Date Implemented	Classification System	Data Source
Implemented				
Maryland	SSI + TANF	5/97	ACGs	Prior FFS Claims
Colorado	SSI + TANF	7/97	DPS	HMO Encounter Data
Oregon	SSI	6/98	DPS	HMO Encounter Data
Utah	SSI	6/98	Marker Diagnosis	Inpatient Only Encounters
Minnesota ¹	TANF	1/00	ACGs	HMO Encounter Data
Delaware	SSI + TANF	² 1/00	CDPS	HMO Encounter Data
Michigan	SSI	6/00	CDPS	HMO Encounter Data
Planned				
New Jersey	SSI	2000	DPS	Prior FFS
Delaware	SSI	2000	CDPS	HMO Encounter Data
Washington	TANF	2001	CDPS	HMO Encounter Data
Utah ³	SSI	2001	CDPS	HMO Encounter Data

¹ Affects 5 percent of total capitation.

² TANF on 7/00.

³ Dependent upon quality of encounter data.

NOTES: SSI is Supplemental Security Income. TANF is Temporary Assistance to Needy Families. ACGs is Adjusted Clinical Groups. FFS is fee-for-service. DPS is Disability Payment System. HMO is health maintenance organization. CDPS is Chronic Illness and Disability Payment System.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

risk-adjusted payments in 2000 and has done substantial work in preparation for implementation. Other State Medicaid programs, including Massachusetts, New York, and Pennsylvania, are seriously evaluating health-based payment options.

Medicaid programs are much more active than private employers in implementing health-based payment, and some are ahead of the Medicare program. HCFA began phasing in risk-adjusted payments to Medicare in January 2000, using only inpatient diagnoses, while most Medicaid programs implementing health-based payment are using or planning to use diagnoses from both ambulatory and inpatient encounters.

Concerns about health care for beneficiaries with disability account for much of the impetus for Medicaid health-based payment, although some States have extended health-based payment to beneficiaries of Temporary Assistance to Needy Families (TANF) as well. As health-based payment in Medicaid is implemented more widely, we expect increased interest in its use for TANF beneficiaries. Yet relatively little

information is available about the burden of disease among TANF beneficiaries, nor has much analysis been presented of the ability of diagnostic classification systems to do a good job of fairly compensating HMOs for this population (Weiner et al., 1998). In this article, we:

- Describe a diagnostic classification system, CDPS, that we have developed as a tool for State Medicaid programs to use when paying HMOs for Supplemental Security Income (SSI) and TANF beneficiaries.
- Describe the burden of disease among SSI and TANF beneficiaries and the extent to which expenditure effects of diagnoses are similar or different in the two groups.
- Compare the structure and statistical performance of CDPS with those of other leading diagnostic classification systems.
- Describe the extent to which diagnoses persist from year to year in fee-for-service claims data and discuss the implications for health-based payment systems.

DEVELOPMENT OF CDPS

The revision of the Disability Payment System (DPS) was intended to make the system more complete and more effective in its adjustment of payments for the TANF population. We used a much larger data base than previously, with claims records for nearly 4 million Medicaid beneficiaries from seven States. Effects of diagnoses on future expenditures were analyzed for all the 15,000 diagnosis codes in the *International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM)* (Public Health Service and Health Care Financing Administration, 1980), and physician specialists were consulted extensively to help determine the appropriateness and organization of diagnoses included in the new system.

The product of the revision, the CDPS, includes 20 major categories of diagnoses, which correspond to body systems or type of diagnosis. (For prospective estimation of payment weights, we exclude the categories for infants, leaving the model with 19 major categories.) Most of the major categories are further divided into several subcategories according to the degree of the increased expenditures associated with the diagnoses. For example, diagnoses of the nervous system are divided into three subcategories for high-cost, medium-cost, and low-cost conditions. (Refer to Table 2 for a list of the subcategories and sample diagnoses.)

CDPS includes many more diagnoses and works better than the original DPS in predicting expenditures for Medicaid beneficiaries of all types.¹ Medicaid programs can

¹The new name, Chronic Illness and Disability Payment System, instead of Disability Payment System, is intended to correct the mistaken impression that the payment system consists primarily of disabilities or could be used only for a disabled population. The vast majority of the diagnoses in the old and new models are not disabilities but diagnoses of disease. Some of these diagnoses are very serious and could be disabling, e.g., muscular dystrophy, but many others, e.g., migraines or uncomplicated adult-onset diabetes, are unlikely to be disabling conditions.

use the new CDPS with greater confidence that the system fully exploits diagnoses that predict significantly elevated future expenditures and that are sufficiently well defined for payment purposes. Software to implement CDPS is available at no charge at <http://www.medicine.ucsd.edu/fpm/cdps/>.

Data

The initial selection of diagnoses for CDPS was based on analysis of expenditure data for approximately 600,000 disabled Medicaid beneficiaries and 3.3 million Aid to Families with Dependent Children (AFDC) and AFDC-related Medicaid beneficiaries (Table 3).² The data base was substantially larger than the data base for the original DPS, which used only 120,000 SSI Medicaid beneficiaries in two States for identification of diagnoses and 400,000 beneficiaries in five States for final testing and determination of categories. Our data contained information on services and procedures, Medicaid payments, and diagnoses, including usually one diagnosis code for ambulatory claims and up to five diagnosis codes in most States for inpatient claims.

To focus on more complete diagnostic records, we included in the analysis only those beneficiaries with a full initial year of Medicaid eligibility and at least 1 month of subsequent-year data. We excluded beneficiaries with Medicare coverage and those enrolled in health plans, for whom Medicaid has partial or no claims information. We also excluded from the analysis beneficiaries in institutions and those in

²The disabled beneficiaries include people receiving SSI as well as others who are eligible for Medicaid because of disability. The AFDC program has been transformed to TANF, but all of our data predate the conversion. Welfare reform has resulted in a sharp decline in the Medicaid caseload, and there is speculation that the remaining caseload is sicker (Ellwood and Lewis, 1999). However, we think that the relationships between diagnoses and expenditures that we find among AFDC beneficiaries in the early and mid-1990s will be applicable to current Medicaid beneficiaries as well.

Table 2
Chronic Illness and Disability Payment System Categories with Sample Diagnoses

Diagnostic Category	Sample Diagnoses
Cardiovascular	
Very High	Heart transplant status or complications
Medium	Congestive heart failure, cardiomyopathy, tricuspid and pulmonary valve disease
Low	Endocardial disease, myocardial infarction, angina, coronary atherosclerosis, dysrhythmias
Extra Low	Hypertension
Psychiatric	
High	Schizophrenia
Medium	Bipolar affective disorder
Low	Other depression, panic disorder, phobic disorder
Skeletal and Connective	
Medium	Chronic osteomyelitis, aseptic necrosis of bone
Low	Rheumatoid arthritis, osteomyelitis, systemic lupus, traumatic amputation of foot or leg
Very Low	Osteoporosis, musculoskeletal anomalies, thoracic and lumbar disc degeneration
Extra Low	Osteoarthritis, skull fractures, other disc and vertebral disorders
Nervous System	
High	Quadriplegia, amyotrophic lateral sclerosis and other motor neuron disease
Medium	Paraplegia, muscular dystrophy, multiple sclerosis
Low	Epilepsy, Parkinson's disease, cerebral palsy, migraine, cerebral degeneration
Pulmonary	
Very High	Cystic fibrosis, lung transplant, tracheostomy status, respirator dependence
High	Respiratory arrest or failure, primary pulmonary hypertension, selected bacterial pneumonias
Medium	Other bacterial pneumonias, chronic obstructive asthma, adult respiratory distress syndrome
Low	Viral pneumonias, chronic bronchitis, asthma, COPD, emphysema
Gastrointestinal	
High	Peritonitis, hepatic coma, liver transplant
Medium	Regional enteritis and ulcerative colitis, chronic liver disease and cirrhosis, enterostomy
Low	Ulcer, hernia, GI hemorrhage, intestinal infectious disease, intestinal obstruction
Diabetes	
Type 1 High	Type 1 diabetes with renal manifestations or coma
Type 1 Medium	Type 1 diabetes without complications or with neurological or ophthalmic complications
Type 2 Medium	Type 2 or unspecified diabetes with complications, proliferative diabetic retinopathy
Type 2 Low	Type 2 or unspecified diabetes without complications
Skin	
High	Decubitus ulcer
Low	Other chronic ulcer of skin
Very Low	Cellulitis, burn, lupus erythematosus
Renal	
Very High	Chronic renal failure, kidney transplant status or complications
Medium	Acute renal failure, chronic nephritis, urinary incontinence, cystostomy or urinostomy
Low	Kidney infection, kidney stones, hematuria, urethral stricture, bladder disorders
Substance Abuse	
Low	Opioid, barbiturate, cocaine, amphetamine abuse or dependence, drug psychoses
Very Low	Alcohol abuse, dependence, or psychosis
Cancer	
High	Lung cancer, ovarian cancer, secondary malignant neoplasms, leukemia, multiple myeloma
Medium	Mouth, breast or brain cancer, malignant melanoma, radiation or chemotherapy
Low	Colon, cervical, or prostate cancer, carcinomas in situ
Developmental Disability	
Medium	Severe or profound mental retardation
Low	Mild or moderate mental retardation, Down's syndrome
Genital	
Extra Low	Uterine and pelvic inflammatory disease, endometriosis, hyperplasia of prostate

See footnotes at end of table.

Table 2—Continued
Chronic Illness and Disability Payment System Categories with Sample Diagnoses

Diagnostic Category	Sample Diagnoses
Metabolic	
High	Panhypopituitarism, pituitary dwarfism, non-HIV immunity deficiencies
Medium	Kwashiorkor, marasmus, and other malnutrition, parathyroid, and adrenal gland disorders
Very Low	Other pituitary disorders, gout
Pregnancy	
Incomplete	Normal pregnancy, complications of pregnancy
Complete	Normal delivery, multiple delivery, delivery with complications
Eye	
Low	Retinal detachment, choroidal disorders, vitreous hemorrhage
Very Low	Cataract, glaucoma, congenital eye anomaly, corneal ulcer
Cerebrovascular	
Low	Intracerebral hemorrhage, precerebral occlusion, hemiplegia, cerebrovascular accident
Infectious Disease	
AIDS, High	AIDS, pneumocystis pneumonia, cryptococcosis, Kaposi's sarcoma
Infectious, High	Staphylococcal or pseudomonas septicemia, cytomegaloviral disease
HIV, Medium	Asymptomatic HIV infection
Infectious, Medium	Other septicemia, pulmonary or disseminated candida, toxoplasmosis, typhus
Infectious, Low	Poliomyelitis, oral candida, herpes zoster, parasitic intestinal infections
Hematological	
Extra High	Congenital factor VIII and factor IX coagulation defects (hemophilia)
Very High	Hemoglobin-S sickle-cell disease
Medium	Other hereditary hemolytic anemias, aplastic anemia, splenomegaly, agranulocytosis
Low	Other white blood cell disorders, purpura, other coagulation defects

NOTES: COPD is chronic obstructive pulmonary disease. GI is gastrointestinal. HIV is human immunodeficiency virus. AIDS is acquired immunodeficiency syndrome. CDPS is Chronic Illness and Disability Payment System. CDPS also includes categories for infants and a more detailed categorization for pregnancy. A complete description of CDPS diagnostic categories by ICD-9-CM codes is available at <http://www.medicine.ucsd.edu/fpm/cdps/>. ICD-9-CM is *International Classification of Diseases, 9th Revision, Clinical Modification* (Public Health Service and Health Care Financing Administration, 1980).

SOURCE: Kronick, R., et al., San Diego, California, 2000.

Table 3
**Number of Observations in Regression Analysis, by State and Beneficiary Group:
 Selected States**

State	Totals	Adults with Disability	Children with Disability	AFDC Adults	AFDC Children
Total	6,280,443	960,760	130,324	1,548,488	3,640,871
California	3,415,068	402,987	39,427	905,474	2,067,180
Colorado	338,925	50,454	15,701	85,221	187,549
Georgia	667,424	88,538	16,848	148,709	413,329
Michigan	1,007,649	118,996	16,758	288,462	583,433
Missouri	72,270	67,886	4,384	—	—
Ohio	137,451	118,700	18,751	—	—
Tennessee	641,656	113,199	18,455	120,622	389,380
Unduplicated Count of Beneficiaries	3,936,626	549,595	80,646	1,001,775	2,304,610

NOTES: AFDC is Aid to Families with Dependent Children. Data from Michigan, Ohio, and Tennessee are for 1991-1993; California and Georgia are for 1990-1992; Missouri, 1991-1994; Colorado, 1992-1996. Observations are included in the regression analyses for beneficiaries with 12 months of eligibility in the base year and at least 1 month in the rate year. Beneficiaries were excluded if they had Medicare coverage, were institutionalized, enrolled in a health maintenance organization, or in a home and community-based waiver program. The unduplicated count of beneficiaries is lower than the total number of observations because beneficiaries who were continuously eligible for more than 24 months account for 2 or more observations. The table lists the total number of observations; model development was performed on a 75-percent sample.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

home and community-based waiver programs, who are not often enrolled in managed care. We calculated expenditures for services typically included in an acute care HMO benefit package, excluding dental services and long-term care services. (Approximately 50 percent of expenditures for beneficiaries with disability are made for long-term care, but these services are not currently included under capitated contract.) We set aside one-quarter of the data for use as a validation sample, which left us with 3 million individuals for use in the development of the model.

In the regression analysis, we combined multiple years of data to increase the stability of the estimates. As a result, the number of observations was larger than the unduplicated count of beneficiaries. To obtain results that could be useful in a variety of Medicaid programs, we modified the data from the seven States in two ways. First, to minimize the effects of interstate and interyear variation in the level of expenditures, we normalized expenditures in each State-year to 1.0 by dividing the expenditures per month by the mean expenditures per month for the State-year. These normalized expenditures were then used as the dependent variable for regression against prior-year diagnoses. Second, we weighted the observations used in the regression, so that the set of observations from each State received equal weight. In this way, we avoided results dominated by the expenditure patterns in California, which had a large share of the total observations. We also reduced the weight for beneficiaries who were eligible for only part of the year in which expenditures were observed. We analyzed the residuals from an unweighted regression, found that the standard deviation of the residuals was approximately four times larger for persons eligible for 1 month than persons eligible for 12 months and weighted each

observation by $1 - 0.067 \times (12 - \text{number of eligible months})$. This method decreases the weight progressively more for beneficiaries with less eligibility in the second year and reduces the influence of the shorter period observations in rough proportion to their increasing variability.

Method of Analysis

The selection and grouping of diagnoses for CDPS depended upon analysis of our expenditure data and on the advice of 15 clinician consultants. The basic method of analysis was to use the presence of diagnoses recorded in the first year of individuals' claims as regression variables to predict expenditures in individuals' subsequent year of claims. We empirically identified diagnoses that are significantly associated with increased future health care costs. These diagnoses, largely chronic conditions, can serve the aim of health-based payment to provide additional resources to plans that enroll people with greater ongoing needs.

An important challenge to any effort to construct a diagnosis-based payment system is the defining of diagnoses in terms of ICD-9-CM codes. These 15,000 codes are organized under nearly 1,000 three-digit general codes, nearly all with four-digit or five-digit subcodes required for reference to a more specific diagnosis. For example, the ICD-9-CM code 428 refers to heart failure, but more specific four-digit codes are 428.0 for congestive heart failure and 428.1 for left heart failure. Some codes are vague, such as 429.2 for unspecified cardiovascular disease, while others are more specific, such as 250.43 for uncontrolled Type 1 diabetes with renal manifestations.

Creating the diagnostic classification system required decisions about what level of detail should be used in defining each diagnosis in the system. For example,

consider hemoglobin-S disease, a type of sickle-cell anemia. Should hemoglobin-S disease with crisis (282.62) be considered its own condition, separate from hemoglobin-S disease without mention of crisis (282.61)? Or should these two conditions be considered a single entity? Or should hemoglobin-S disease be combined with other sickle-cell anemias or with all the hereditary hemolytic anemias?

Defining a diagnosis more narrowly appears to give greater accuracy in predicting expenditures, but too narrow a definition could make it difficult for clinicians to agree whether an individual's condition justifies a given diagnosis, and could lead to unstable expenditure estimates. We were eager to distinguish among diagnoses that were associated with markedly different levels of elevated future cost, but we did not want to separate codes into different groups if our consulting clinicians believed that it might be difficult to distinguish between them. In general, we regarded groups of diagnoses coded with the same first three digits as likely candidates for the codes to define a single diagnosis, and we broke up three-digit groups only when both the data suggested and clinician judgment concurred that these were distinct conditions that could have different effects on future cost.

Even in cases where we found significantly different effects on expenditures for two separately coded diagnoses, we kept them together in a single variable if our clinical consultation suggested that the distinction between the diagnoses could not be easily made. We found, for example, that the diagnosis of paranoid schizophrenia was associated with more elevated future cost than catatonic schizophrenia, but our clinical consultants were not convinced that diagnoses of subtypes of schizophrenia were made consistently, so we kept all diagnoses of schizophrenia together as a single variable.

Excluding Ill-Defined Diagnoses

Much of our consultation with clinicians was intended to screen out diagnoses that are clinically not well defined. We made special efforts to exclude ill-defined diagnoses from CDPS in order to make the system more reliable and reduce the chances that health plans, clinicians, and Medicaid programs would find themselves questioning diagnoses. Given that health-based payment will naturally cause plans to make greater efforts to report diagnoses, a focus on well-defined diagnoses seems advisable to prevent difficult disagreements between payers and plans. Ill-defined diagnoses would make it difficult for States to audit plans and distinguish between accurate and inaccurate reporting. Our efforts to exclude ill-defined diagnoses parallel the work of other researchers in health-based payment. Notably, Ash and colleagues exclude many diagnoses from their Hierarchical Condition Category (HCC) model as too vague to be used for adjustment of payments (Ash et al., 1998).

We considered a diagnosis to be well defined if it had a clear, shared meaning among clinicians. The diagnosis should be distinctive enough that an auditing clinician could judge from a good medical record whether the diagnosis was made on an adequate clinical basis. Laboratory results or diagnostic imaging help make many diagnoses well defined, but many diagnoses rely significantly on physician observation. As a result, diagnosis-based payment resembles traditional billing in its reliance on honest physician observation, and States will need to audit diagnoses to keep diagnostic reporting honest.

Although many diagnoses that would help predict future cost were judged as not well defined and were excluded from the model, the majority of diagnoses that are predictive of elevated future costs were

judged as well defined. Many of the excluded diagnoses came from sections in the ICD-9-CM specifically set up for ill-defined descriptions of disease or for recording symptoms that have not yet been tied to a specific disease.

Some of the conditions judged to be ill-defined were diagnoses that lack clear definition for clinicians, such as chest pain or dyspnea. We also excluded many common symptoms that we judged to be too easily elicited in patient histories: Symptoms such as headache, backache, or joint pain might be recalled by any adult from some time in the months previous to a physician visit.

In a few instances we bent our standards to include diagnoses in CDPS that a clinician thought were not well defined. For example, because of our concern that some individuals with dementia or traumatic brain injury might not otherwise be recognized by CDPS, we included the diagnoses of Alzheimer's disease and of non-psychotic mental disorders due to organic brain damage despite clinician concerns that these conditions are not well defined. We also included asthma and chronic bronchitis, but not acute bronchitis, even though the distinction between chronic bronchitis and asthma on the one hand and acute bronchitis on the other may not always be well observed. We did so because asthma can be quite expensive, and for many people, asthma is likely their only CDPS diagnosis. We anticipate further research and controversy about the inclusion of more and less well-defined conditions in payment systems, because many of the conditions we exclude for being ill-defined are useful predictors of future cost.

The inclusion of ill-defined diagnoses may increase predictive accuracy but will likely reduce accuracy in implementation. In general, as more diagnoses are included in a payment system, a greater volume of diagnoses needs to be reported and audit-

ed, and a higher proportion of variation in level of need observed among plans would result from differences in plans' abilities to make and report diagnoses rather than from actual differences in their enrollees. It seems likely that the inclusion of ill-defined diagnoses would particularly make the payment system more vulnerable to aggressive plan efforts to increase reporting. The modest improvement in accuracy on a given data set that is gained through ill-defined diagnoses seems far less important than having a system that is more easily administered and probably more accurate in practice.

We believe that the exclusion from CDPS of ill-defined conditions is a virtue, and we encourage others to consider this issue more carefully. Our work is far from sufficient to settle the question of which diagnoses are well defined and which are not. For each diagnosis that our data indicated was predictive of elevated costs, our approach was to ask specialists directly how well defined they thought the diagnosis was. We did not ask about the many diagnoses that failed to show any association with elevated future cost. A much more intensive approach might involve asking clinicians to make diagnoses from sample medical records and regarding diagnoses as ill-defined where the clinicians' test diagnoses show weak agreement. It might be important to include both specialist and primary physicians because some diagnoses might be well defined for the specialist but not the generalist.

Excluding Low-Cost Diagnoses

A related question is whether groups of diagnoses with high frequency and very low-cost implications should be included in a payment system. For example, bladder and urethral infections were diagnosed for 95,000, or 10 percent, of our sample of adults

with disability and more than 185,000, or 12 percent, of AFDC adults. Estimated additional monthly costs in the next year associated with this diagnosis were only \$12 for adults with disability and \$11 for AFDC adults. A respiratory tract condition such as sinusitis, pharyngitis, acute bronchitis, or cough was diagnosed for 1.5 million, or 42 percent, of the AFDC children in the sample and for 33 percent of AFDC adults. Estimated additional monthly costs in the next year were \$8 for the AFDC children and \$13 for the AFDC adults.

These additional amounts are very small in comparison with the additional amounts in the range of \$200-800 for the more costly diagnostic groups. The cost effects of less than \$30 are also small relative to the average monthly expenditures for people with disability in our sample of \$416. For AFDC beneficiaries, however, these small cost effects are more significant because average AFDC adult expenditures are \$158 per month, and average expenditures for AFDC children are \$57 per month.

In theory, if the gathering and reporting of diagnoses were perfect, it would be advantageous to include high-frequency, low-cost diagnoses because their presence would increase accuracy and fairness by bringing more money to the plans that serve people with greater needs. In practice, however, plans' ability and eagerness to make and report diagnoses might vary. As we argued previously, the more high-frequency, low-cost diagnoses are included, the more apparent variation in need is likely to result from differences in reporting, not actual differences in need.

An important consideration in deciding whether to include diagnoses for payment purposes is whether there is reason to expect uneven distribution of enrollees with a certain diagnosis across plans. For many of the extremely low-cost diagnoses, such as bladder infections, minor upper

respiratory conditions, or ear infections, there is little reason to expect that people with these conditions will be distributed unevenly among plans. On the other hand, some other low-cost conditions, such as hypertension, migraines, or asthma, have somewhat higher cost effects and might be distributed unevenly among plans, and plans that have stronger specialist networks or are located in poorer neighborhoods might attract a disproportionate share. In addition, encouraging plans to diagnose these conditions has a value in itself because attention to them can be highly beneficial for individual health. The inclusion of hypertension is particularly important for this reason and also because 13 percent of disabled adults are coded with hypertension but not more serious cardiovascular disease.

As a result of all these considerations, we decided to eliminate many of the lower cost conditions from our recommended payment model. We recommend the use of 56 diagnostic subcategories for payment purposes and an additional 15 subcategories of high-frequency diagnoses with very small cost effects only for profiling purposes.

Counting Diagnoses with the CDPS Subcategories

The organization of diagnostic categories and the rules for counting diagnoses are somewhat different in CDPS than they were in the original DPS. The most obvious change is an increase in the number of diagnostic subcategories, from 43 in DPS to 56 in CDPS, which results partly from the more comprehensive and larger set of diagnoses included in CDPS. Some of the new subcategories result from increasing distinctions among diagnoses that were in DPS, while other new subcategories are in new major areas. New major

areas include infectious disease, pregnancy, and infants. Two new subcategories resulted from creating separate subcategories for Type 1 and Type 2 diabetes. And several other major categories gained a subcategory to reflect finer distinctions among cost levels.

A less obvious but equally significant change is in the rules used for counting diagnoses within major categories. In the original DPS, 10 of the 18 major categories were designated as “hierarchical” categories in which only the single most severe diagnosis within the major category was counted, while 8 were designated as “fully counted” categories in which multiple diagnoses could be counted. Our use of fully counted categories had been intended to capture the additional needs that arise from distinct diseases, but in revisiting this issue, we placed a higher value on limiting incentives for proliferative coding and on consistency across major categories. We also found relatively little predictive benefit in counting multiple diagnoses within major categories. As a result, every one of the major categories in CDPS is counted hierarchically. This change in the counting rules simplifies the model, strengthens its resistance to additional coding, and produces only small decreases in the accuracy of simulated payments.

Single counting within major categories is intended to avoid encouraging a proliferation of different diagnoses reported for a single disease process just to increase payment. For example, if someone is diagnosed with a significant cardiovascular disease such as congestive heart failure, an additional diagnosis of hypertension is probably not of much additional significance for cost. In other cases, additional coding would clearly have no implication for cost, for example, if someone with Type 1 diabetes with complications were subsequently coded with uncomplicated Type 1 diabetes.

Thus, if an individual has a medium-cost infectious disease and a low-cost infectious disease, he or she would be counted in CDPS as simply having medium-cost infectious disease. An individual with two different medium-cost psychiatric illnesses would be counted simply as having medium-cost psychiatric illness. As a result of this approach, the expenditures associated with people with multiple diagnoses in a single major category are loaded onto the single-highest category.

We experimented with various counting approaches, stimulated in part by the example of Ash and colleagues (1998), who developed counting methods that mix counting and hierarchy within diagnostic categories. We tried an intermediate approach between full counting and single counting: Major categories were subdivided into different diagnostic areas and additional counts were allowed for diagnoses in the different areas. For example, we divided cardiovascular diagnoses into areas such as valvular, myocardial, dysrhythmic, and peripheral conditions. We found, however, that such subdivisions added substantially to the complexity of the model but yielded very little improvement in its performance.

Meanwhile, CDPS follows DPS in counting multiple diagnoses when they are from different major categories. Thus, if an individual had two medium-cost renal diagnoses and low-cost developmental disability, he or she would be counted as having medium-cost renal disease and low-cost developmental disability. By considering not only a person’s single most serious diagnosis but also diagnoses from other major categories, accuracy is substantially improved. Using diagnoses from multiple major categories improves accuracy because average expenditures are much higher for people with diagnoses from greater numbers of major categories.

In addition, the diagnosis of additional conditions from different major categories offers a greater potential benefit for individuals than the additional coding of highly related diagnoses. For a male already diagnosed with schizophrenia who receives an additional diagnosis of paranoid state, the additional diagnosis may not be associated with any new treatment he would otherwise not have received. But the diagnosis of a gastrointestinal or cardiovascular condition could bring him valuable additional care. By counting diagnoses from different major categories, CDPS encourages identification of additional chronic conditions that may deserve attention. Such attention to the whole individual is important for everyone but can be particularly important for people with serious mental illness or developmental disability, whose medical attention has often focused on their cognitive problems and neglected physical health (Druss et al., 2000).

Estimating Payment Weights for Subgroups

The population of Medicaid beneficiaries can be treated as a single group or divided into subgroups when estimating payment weights. If a single group for all Medicaid beneficiaries is used, then variables for category of assistance and age and interactions of these variables with diagnostic groups can be introduced to better fit the model. Alternatively, the model can be estimated separately on subgroups such as the AFDC adults or the SSI children. If the effects of diagnoses on expenditures are similar across groups or if there are relatively few people in a subgroup, then a single model with supplementary variables will work well. On the other hand, if the effects of diagnoses are dissimilar across subgroups and the subgroups are large, then separate models are preferable.

We estimated separate weights for persons with disability and AFDC beneficiaries because: (1) there are very large differences in the expenditure effects of CDPS categories between adults with disability and those enrolled in AFDC; (2) there are large differences in the average expenditures for the two groups; (3) we have large samples of both AFDC adults and those with disability; and (4) State Medicaid programs typically have separate base rates for SSI beneficiaries and TANF beneficiaries.

Among persons with disabilities, we estimated a combined regression for adults and children because we had found that parameter estimates from regressions estimated separately on adults and children with disabilities were relatively similar. So, too, were average expenditures per month and the average number of CDPS diagnoses. We also used the combined regression because most States do not have a separate base rate for non-Medicare adults and children with disabilities. Although the effects of CDPS diagnoses on expenditures are similar among adults and children with disabilities, they differ in some areas. We estimated a regression with interactions of age (coded as a dichotomous variable for under or over 18) and each of the CDPS categories, and retained in the final model 11 interactions with CDPS categories that appeared useful. (We selected interactions that had substantial numbers of children, *t*-statistics with absolute values above 2.0 [with the exception of diabetes], and reasonably stable estimates across States.) For 9 of these 11 categories, the effects of diagnoses are larger among children with disabilities than among adults with disabilities.

In contrast to our approach of using a combined model for adults and children with disabilities, we estimated separate regressions for AFDC adults and AFDC children for several reasons. First, the

expenditure effects of CDPS diagnoses are substantially different for AFDC children and adults in a large number of CDPS sub-categories.³ Second, many States have separate base rates for AFDC children and adults, making separate payment weights useful. Third, we have a very large sample of AFDC children and adults, supporting the estimation of reasonably stable payment weights for the two groups.

DISEASE BURDEN AND EXPENDITURE EFFECTS

CDPS provides a detailed diagnostic description of Medicaid beneficiaries in terms of the frequency of different kinds of diagnoses and their expenditure effects. Our analysis shows that only a small proportion of the SSI disabled have such salient conditions as paraplegia, acquired immunodeficiency syndrome (AIDS), or cystic fibrosis, which policymakers or the public may commonly associate with disability and Medicaid. Instead, Medicaid beneficiaries with disability experience a wide variety of serious and less serious conditions to which policymakers and health plans might give more attention.

Most of the more serious diagnoses are found much less frequently among AFDC adults than among adults with disability. Among children receiving Medicaid, children with disability have much greater relative frequency of serious diagnoses, but the absolute numbers of AFDC children enrolled in Medicaid are so much larger than the number of SSI children that the number of AFDC children with a given serious condition may be greater than the number of disabled children with that con-

³ We estimated a regression on AFDC adults and children combined, with interaction terms for age and CDPS categories. Using the decision rule that we would keep an interaction term in the model if the CDPS category had substantial numbers of children, had a significant *t*-statistic, and was similar in regressions on individual States, we would have kept interaction terms for slightly more than one-half of the 56 CDPS categories.

dition. (When we refer to “children with disability,” we mean those who are Medicaid beneficiaries because they receive SSI as well as other children who are eligible for Medicaid because of disability. As we show later, there are many AFDC children who also have chronic illnesses and disabilities.) The diagnostic description details the diverse challenges of providing health care to Medicaid beneficiaries with disability and to AFDC beneficiaries.

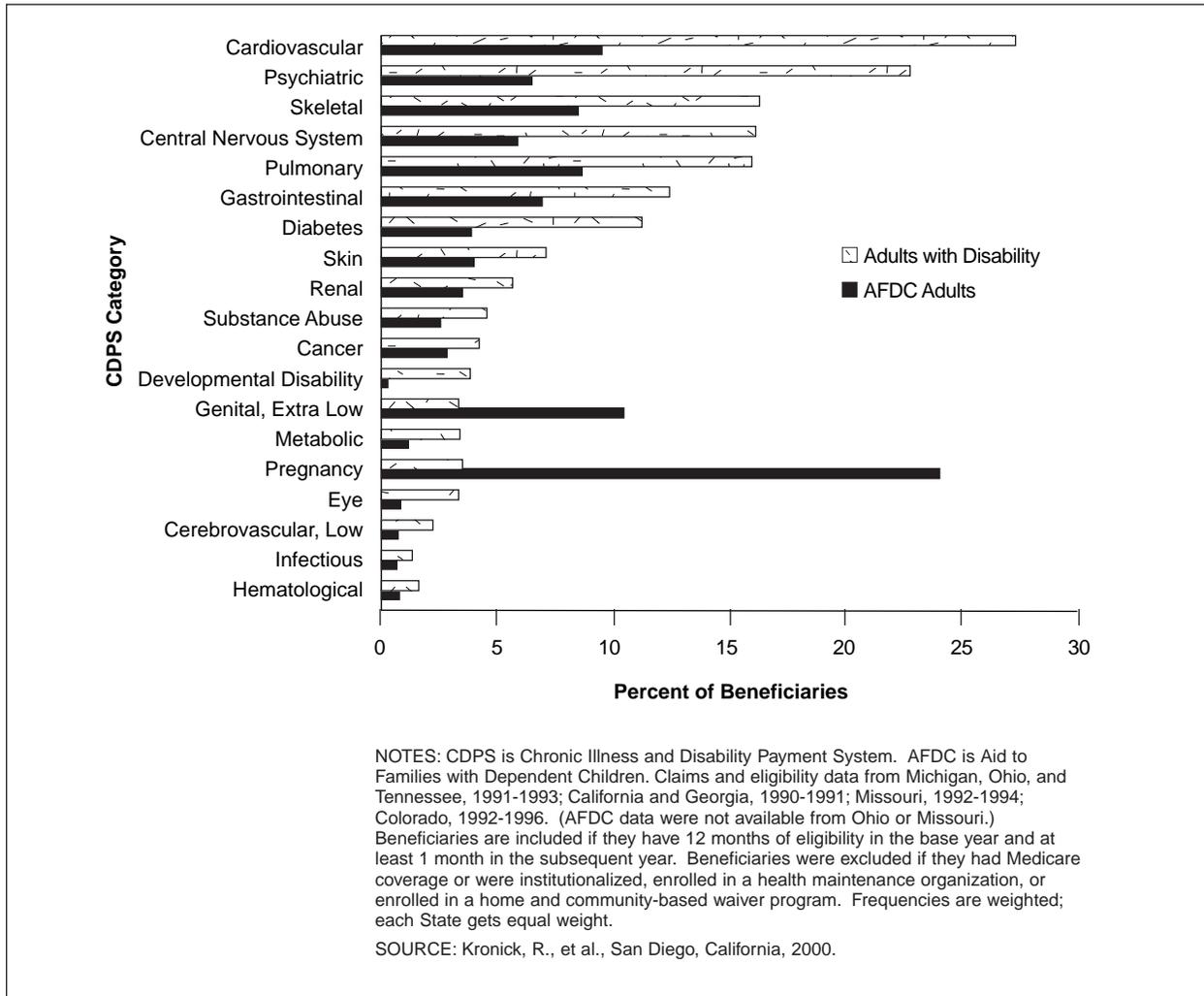
Disease Burden for Adults

Among adults with disabilities, approximately one-quarter of beneficiaries have a psychiatric or a cardiovascular diagnosis that is included in CDPS, and approximately 10-15 percent of beneficiaries have a diagnosis in the skeletal, central nervous system, pulmonary, gastrointestinal, or diabetes CDPS categories (Figure 1 and Table 4).⁴ Smaller numbers have a diagnosis in other CDPS categories such as renal, substance abuse, cancer, and metabolic. There are a few CDPS categories, such as infectious diseases (including AIDS) and hematology that are recorded for only 2 percent of adult beneficiaries with disabilities.

With the exceptions of pregnancy and diagnoses in the genital category, the prevalence of major CDPS categories for AFDC beneficiaries is uniformly lower than for persons with disabilities. For many major CDPS groups, the frequency among AFDC beneficiaries is approximately one-third to one-half of the rate among persons with disabilities. Pregnancy, how-

⁴ The frequencies in Table 4 are weighted frequencies, giving equal weight to the observations for each State. The weight for each observation is proportional to (total number of observations in a beneficiary group across all States)/(total number of observations in the State in the beneficiary group). The frequencies for most categories vary relatively little across States, so that the weighted and unweighted frequencies are, for the most part, similar. However, for a few CDPS categories, California is notably different, and the weighted frequencies give less weight to California than the unweighted.

Figure 1
Frequency of Major CDPS Categories for Adults with Disability and AFDC Adults



ever, is far more common for AFDC adults: 24 percent of AFDC adults who were continuously eligible for at least 12 months either complete a pregnancy or were pregnant during the year; in contrast, pregnancy is uncommon among persons with disability. (Completed pregnancy includes miscarriages, abortions, normal delivery, and delivery with complications; incomplete pregnancy includes normal and complicated pregnancies.)

There are a number of high-cost and very-high-cost CDPS categories that are rare among adults with disabilities but are close to non-existent among AFDC adults

(Figure 2). For example, very-high-cost cardiovascular problems (primarily heart transplants) are coded for 0.23 percent of persons with disabilities but only for 0.02 percent of AFDC adult beneficiaries; high-cost central nervous system problems (primarily quadriplegia) are diagnosed in 0.34 percent of persons with disabilities and less than 0.01 percent of AFDC beneficiaries. Even more striking, high-cost psychiatric problems (primarily schizophrenia) are relatively common among persons with disabilities (11.7 percent of beneficiaries) but are diagnosed in only 0.3 percent of AFDC adults.

Table 4
Frequency of CDPS Categories, by Beneficiary Group

Category	Disabled Adults (n) = 960,760	AFDC Adults (n) = 1,548,488	Disabled Children (n) = 130,324	AFDC Children (n) = 3,640,871
	Percent			
Cardiovascular	27.06	9.38	7.98	1.19
Very High	0.23	0.02	0.66	0.01
Medium	3.52	0.46	0.04	0.00
Low	11.13	3.80	6.53	1.00
Extra Low	12.18	5.10	0.75	0.18
Psychiatric	22.67	6.83	11.64	3.32
High	11.65	0.34	0.36	0.04
Medium	1.63	0.27	0.40	0.06
Low	9.39	6.22	10.88	3.22
Skeletal	16.81	8.23	12.01	3.08
Medium	0.26	0.03	0.12	0.01
Low	4.17	1.43	3.66	0.60
Very Low	4.29	3.09	7.25	1.90
Extra Low	8.09	3.68	0.98	0.57
Nervous System	16.65	5.87	31.02	2.78
High	0.34	0.01	0.68	0.00
Medium	1.86	0.27	5.37	0.10
Low	14.45	5.59	24.97	2.67
Pulmonary	16.10	8.66	15.36	9.91
Very High	0.21	(1)	1.56	(1)
High	0.94	0.21	0.61	0.22
Medium	0.87	0.27	1.17	0.24
Low	14.08	8.18	12.02	9.45
Gastrointestinal	12.59	6.96	7.48	3.98
High	0.29	0.09	0.44	0.02
Medium	2.11	0.68	1.31	0.15
Low	10.19	6.19	5.73	3.81
Diabetes	11.25	4.23	0.94	0.45
Type 1 High	0.11	0.01	(1)	(1)
Type 1 Medium	2.61	0.45	(1)	(1)
Type 2 Medium	0.63	0.10	(1)	(1)
Type 2 Low	7.90	3.67	0.94	0.45
Skin	7.88	4.37	4.93	3.49
High	0.48	0.02	0.28	0.01
Low	0.97	0.21	0.17	0.04
Very Low	6.43	4.14	4.48	3.44
Renal	5.67	3.33	4.96	1.32
Very High	0.63	0.05	0.25	0.02
Medium	1.70	0.42	0.24	0.06
Low	3.34	2.86	4.47	1.24
Substance Abuse	4.92	2.27	0.25	0.18
Low	1.75	1.25	0.11	0.07
Very Low	3.17	1.02	0.14	0.11
Cancer	4.55	2.79	3.55	0.33
High	1.15	0.26	1.33	0.06
Medium	2.20	0.78	1.47	0.17
Low	1.20	1.75	0.75	0.10
Developmental Disability	3.90	0.09	9.12	0.10
Medium	0.76	(1)	1.86	0.01
Low	3.14	0.09	7.26	0.09

See footnotes at end of table.

Table 4—Continued
Frequency of CDPS Categories, by Beneficiary Group

Category	Disabled Adults (n) = 960,760	AFDC Adults (n) = 1,548,488	Disabled Children (n) = 130,324	AFDC Children (n) = 3,640,871
			Percent	
Genital, Extra Low	3.59	10.39	0.93	0.66
Metabolic	3.37	1.11	8.24	1.09
High	0.79	0.21	1.32	0.11
Medium	0.77	0.35	0.98	0.14
Very Low	1.81	0.55	5.94	0.84
Pregnancy	3.53	24.12	0.75	0.87
Incomplete	2.21	17.19	0.41	0.48
Complete	1.32	6.93	0.34	0.39
Eye	3.20	0.53	1.43	0.21
Low	0.46	0.13	(1)	(1)
Very Low	2.74	0.40	1.43	0.21
Cerebrovascular, Low	2.39	0.43	1.89	0.15
Infectious	1.18	0.41	1.11	1.64
AIDS, High	0.40	0.09	0.17	0.03
Infectious, High	0.11	0.03	0.14	0.02
HIV, Medium	0.12	0.06	0.03	0.01
Infectious, Medium	0.55	0.23	0.77	0.24
Infectious, Low	0.89	0.66	0.86	1.34
Hematological	1.74	0.65	2.93	0.43
Extra High	0.06	0.01	0.29	0.01
Very High	0.29	0.02	1.32	0.07
Medium	0.53	0.26	0.90	0.20
Low	0.86	0.36	0.42	0.15
With No CDPS Diagnosis	28.6	53.3	35.1	72.4

¹ Subcategories were combined with the subcategory or subcategories below for the purposes of the regression, because the numbers of beneficiaries in the category were too small to allow a reliable estimate of the expenditure effect. For example, the pulmonary very-high-cost subcategory was combined into the pulmonary high-cost category for AFDC adults and AFDC children. For both disabled children and AFDC children, all subcategories of diabetes were collapsed into a single category.

NOTES: CDPS is Chronic Illness and Disability Payment System. AFDC is Aid to Families with Dependent Children. HMO is health maintenance organization. Individuals can be counted in more than one diagnostic category. Data from Michigan, Ohio, and Tennessee are for 1991-1993; California and Georgia are for 1990-1992; Missouri, 1991-1994; Colorado, 1992-1996. (AFDC data were not available from Ohio or Missouri.) Beneficiaries are included if they have 12 months of eligibility in the base year and at least 1 month in the subsequent year. Beneficiaries were excluded if they had Medicare coverage, were institutionalized, enrolled in an HMO, or in a home and community-based waiver program. Frequencies are weighted; each State gets equal weight. AIDS is acquired immunodeficiency syndrome. HIV is human immunodeficiency virus.

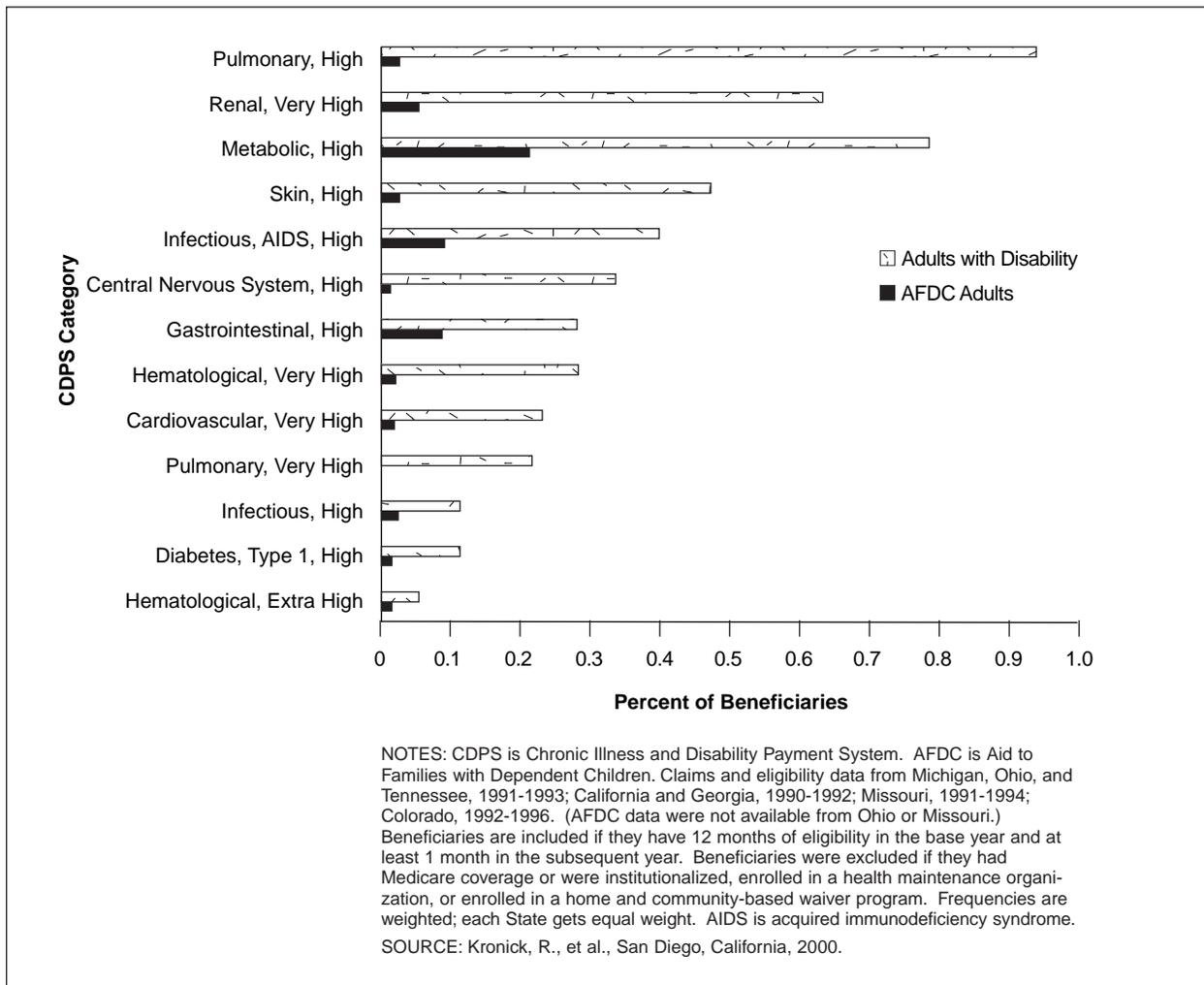
SOURCE: Kronick, R., et al., San Diego, California, 2000.

To contrast the diagnostic profile of adults with disabilities and AFDC adults, we selected the 32 highest cost subcategories. Among adults with disabilities, 33 percent have a diagnosis in at least one of these relatively high-cost categories, a rate that is almost five times the 7 percent rate among AFDC adults (Table 5). Considering all CDPS categories (excluding pregnancy), adults with disabilities average approximately twice as many CDPS diagnoses as AFDC adults. The higher proportion of adults with disabili-

ties who have CDPS diagnoses is as expected: Adults with disabilities are Medicaid beneficiaries because of a significant physical or mental condition, often a chronic physical or psychiatric illness.

One result that at first seems surprising is the significant proportion of people with disability who appear to have no diagnosis in a CDPS category: 29 percent for adults with disability and 35 percent for children with disability (Table 5). A few of these individuals may be receiving SSI because of conditions that were judged not suffi-

Figure 2
Frequency of Selected High-Cost CDPS Categories for Adults with Disability and AFDC Adults



ciently well defined to be included in CDPS. Many of these individuals, however, probably had or could have had a CDPS condition diagnosed at some point, but the diagnosis was not recorded in claims during the year in which we counted diagnoses. (Analysis of disabled adults with continuous eligibility of more than 24 months shows 20 percent with no CDPS diagnosis; over a 36-month period, 15 percent with no CDPS diagnosis.) In particular, we suspect that a significant number of beneficiaries with mental retardation are not coded as such. In the section “Adjusting to Change in Diagnostic

Reporting,” Medicaid fee-for-service (FFS) data show significant underreporting of diagnoses, with many individuals coded with a significant chronic illness not coded with it again in the following year.

Disease Burden for Children

Children are, on average, much healthier than adults and are much less likely to have most of the conditions that are included in CDPS (Table 4 and Figure 3). For example, cardiovascular problems, diabetes, skeletal and connective diagnoses, and psychiatric problems are all relatively

Table 5
Number of CDPS Categories, by Recipient Group and Type of CDPS Diagnosis

Type of Recipient	All CDPS Categories, Excluding Pregnancy			Relatively High-Cost Categories Only		
	Average Number of Categories per Recipient	Percent with No Categories	Average Number of Categories for Recipients with at Least One Category	Average Number of Categories per Recipient	Percent with No Categories	Average Number of Categories for Recipients with at Least One Category
Adults with Disability	1.66	29	2.34	0.44	67	1.31
AFDC Adults	0.77	54	1.68	0.08	93	1.12
Children with Disability	1.27	35	1.96	0.31	75	1.28
AFDC Children	0.34	74	1.30	0.02	98	1.07

NOTES: CDPS is Chronic Illness and Disability Payment System. AFDC is Aid to Families with Dependent Children. Analysis of Medicaid claims and eligibility data from seven State Medicaid programs for persons with disability; five States for AFDC recipients. For disabled adults, $n = 960,760$; for AFDC adults, $n = 1,548,480$; for children with disability, $n = 130,324$; for AFDC children, $n = 3,640,871$. Relatively high-cost categories exclude 22 low-cost CDPS categories.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

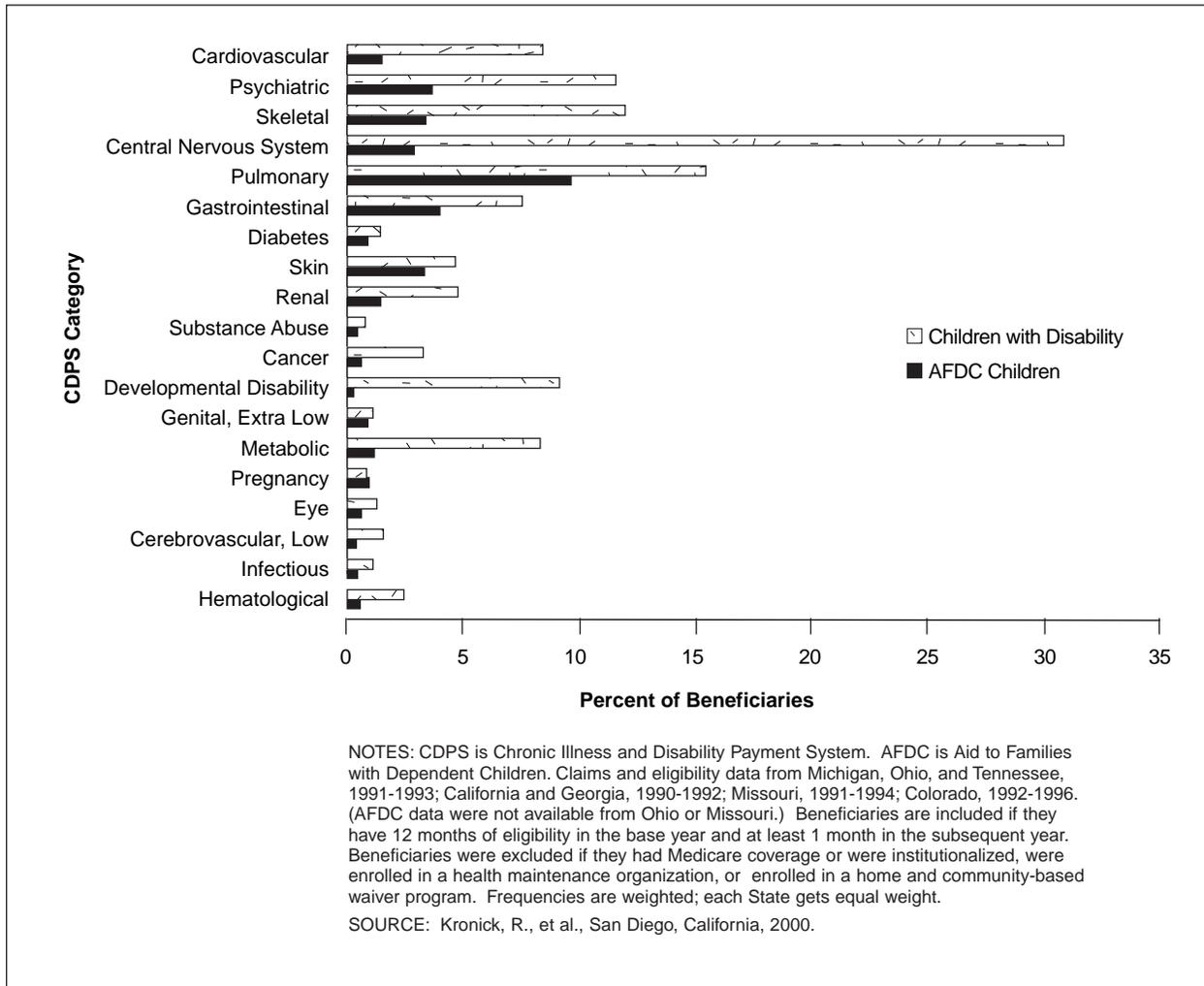
common among adults with disabilities but are much less common among children. There are some areas, such as pulmonary conditions, in which the prevalence of CDPS diagnoses is similar among children and adults, and a few areas, such as developmental disabilities and metabolic conditions, in which diagnoses are more frequent among children than adults. Low-cost central nervous system diagnoses are made for 25 percent of children with disabilities—far higher than the rate of such diagnoses among adults—primarily reflecting high rates of cerebral palsy, spina bifida, and epilepsy.

AFDC children have a very low incidence of chronic illness and disability relative either to AFDC adults or to children with disability. Among AFDC children, 74 percent have no CDPS diagnoses, and 98 percent have no relatively high-cost CDPS diagnoses (Table 5). Children with disabilities have approximately 75 percent as many diagnoses as adults with disabilities, but AFDC children have only 44 percent as many diagnoses as AFDC adults. The relatively healthy diagnostic picture of AFDC children is consistent with the relative average monthly expenditures of the

groups: \$57 per eligible month for AFDC children, \$158 per month for AFDC adults, \$363 for children with disabilities, and \$425 for adults with disabilities.

Although children with disabilities are much more likely than AFDC children to have serious diagnoses, there are so many more AFDC children than there are disabled children that in absolute numbers there are more AFDC children with a variety of severe problems than there are children with disabilities. For example, in our sample, just under 0.3 percent of children with disabilities have extra-high-cost hematological problems (hemophilia with deficiencies in clotting factors VIII or IX), compared with 0.01 percent of AFDC children. However, because there are 30 times as many AFDC children as there are children with disabilities, there are slightly more AFDC children with hemophilia than there are children with disabilities. Similarly, high-cost metabolic problems are diagnosed among approximately 1.3 percent of children with disabilities and 0.11 percent of AFDC children but affect more than twice the number of AFDC children as the number of children with disabilities.

Figure 3
Frequency of Major CDPS Categories, Children with Disability and AFDC Children



Expenditure Effects of Diagnostic Categories

Among beneficiaries with disabilities, diagnoses made in a given year have strong associations with expenditures in the subsequent year (refer to Table 6, which shows parameter estimates from prospective regressions, in which diagnoses from year 1 are used to predict expenditures in year 2).⁵

⁵The regressions reported in Table 6 are based on the full sample—combining the development and validation sample. For presentation purposes, we converted the original regression results, which were ratios of expenditure effects to average monthly expenditures, into annual dollar amounts. To do so, we multiplied the ratios by the weighted average annualized expenditures: \$4,980 for persons with disability, \$1,884 for AFDC adults, and \$684 for AFDC children.

A variety of relatively rare, extra-high-cost and very-high-cost conditions increase subsequent year expenditures by \$10,000 per year or more, for example, very-high-cost cardiovascular conditions (primarily transplants), very-high-cost pulmonary conditions (primarily cystic fibrosis), renal failure, AIDS, and certain extra-high and very high-cost hematological conditions. Diagnoses in each of these very-high-cost categories are found in fewer than 1 percent of beneficiaries with disabilities.

A variety of slightly more frequent but still relatively rare CDPS categories have expenditure effects of approximately \$4,000 to \$9,000 per year. By far the most

Table 6
Subsequent-Year Annual Expenditure Effects of CDPS Categories, by Beneficiary Group

Category	Adults and Children with Disability	AFDC Adults	AFDC Children
Cardiovascular			
Very High	\$14,939	\$7,343	\$9,459
Medium	4,444	2,345	2,947
Low	1,799	943	890
Extra Low	708	701	489
Psychiatric			
High	4,841	22,477	6,037
Medium	3,770	22,477	3,322
Low	1,671	1,076	1,550
Skeletal			
Medium	5,313	3,822	1,365
Low	1,886	1,027	587
Very Low	1,233	2809	369
Extra Low	545	2809	225
Nervous System			
High	9,726	2,699	10,518
Medium	3,314	1,737	3,343
Low	1,582	954	654
Pulmonary			
Very High	13,586	(1)	(1)
High	7,548	1,991	2,422
Medium	5,163	2,268	1,385
Low	1,852	891	496
Gastrointestinal			
High	8,677	22,021	3,231
Medium	3,353	22,021	1,451
Low	1,506	798	304
Diabetes			
Type 1 High	9,911	10,312	(1)
Type 1 Medium	3,787	2,863	(1)
Type 2 Medium	3,111	2,514	(1)
Type 2 Low	1,452	664	729
Skin			
High	7,049	2,523	1,698
Low	2,594	1,122	787
Very Low	867	407	175
Renal			
Very High	14,741	8,387	2,270
Medium	2,536	1,465	646
Low	1,183	650	472
Substance Abuse			
Low	2,253	1,506	2,393
Very Low	1,115	821	967
Cancer			
High	5,114	3,080	4,661
Medium	1,727	1,153	1,199
Low	431	204	766
Developmental Disability			
Medium	5,314	(1)	5,328
Low	1,642	412	2,118

See footnotes at end of table.

Table 6—Continued
Subsequent-Year Annual Expenditure Effects of CDPS Categories, by Beneficiary Group

Category	Adults and Children with Disability	AFDC Adults	AFDC Children
Genital			
Extra Low	\$175	\$464	\$559
Metabolic			
High	4,946	1,670	3,550
Medium	3,156	1,079	1,019
Very Low	1,089	883	582
Pregnancy			
Incomplete	560	492	951
Complete	1,114	1,903	2,231
Eye			
Low	2,199	1,174	(1)
Very Low	1,018	707	686
Cerebrovascular			
Low	1,109	1,066	688
Infectious			
AIDS, High	211,477	23,125	21,282
Infectious, High	211,477	23,125	21,282
HIV, Medium	4,200	1,159	834
Infectious, Medium	4,200	1,159	834
Infectious, Low	1,369	285	145
Hematological			
Extra High	62,576	7,821	12,137
Very High	13,874	6,634	3,350
Medium	3,972	1,047	854
Low	1,967	982	500
Baseline	1,382	944	429

¹ Subcategories were combined with the subcategory or subcategories below for the purposes of the regression because the numbers of beneficiaries in the category were too small to allow a reliable estimate of the expenditure effect.

² The coefficient of this subcategory was constrained to be equal to the coefficient of the subcategory above it because the unconstrained coefficients were not different enough to justify the estimation of separate coefficients.

NOTES: CDPS is Chronic Illness and Disability Payment System. AFDC is Aid to Families with Dependent Children. AIDS is acquired immunodeficiency syndrome. HIV is human immunodeficiency virus. Refer to the Technical Note for an explanation of regression details. Claims and eligibility data from Michigan, Ohio, and Tennessee, 1991-1993; California and Georgia, 1990-1991; Missouri, 1991-1994; Colorado, 1992-1996. (AFDC data were not available from Ohio or Missouri.) Beneficiaries are included if they have 12 months of eligibility in the base year and at least 1 month in the subsequent year. Beneficiaries were excluded if they had Medicare coverage, were institutionalized, enrolled in a health maintenance organization, or enrolled in a home and community-based waiver program. Regressions are weighted; each State gets equal weight.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

common of these high-cost conditions are high-cost psychiatric conditions (primarily schizophrenia), which are diagnosed in nearly 12 percent of beneficiaries with disabilities, and medium-cost cardiovascular conditions (primarily congestive heart failure), which are diagnosed in 4 percent of beneficiaries. Many of the more commonly occurring CDPS categories have expenditure effects of \$1,000-\$2,000 per year. These expenditure effects associated with

additional diagnoses are on top of a baseline amount of \$1,382 for beneficiaries with disability. For context, the average expenditure for beneficiaries with disability in our sample is \$4,980 per year. (The regression includes dummy variables for age and gender. Coefficients for these variables are listed in the Technical Note. The baseline amount of \$1,382 per year is the average value of the intercept plus the age-gender effects.)

The expenditure effects of CDPS diagnoses are much smaller among AFDC adults than among adults with disabilities. For most CDPS categories, the diagnoses that are expensive among beneficiaries with disability also tend to be expensive for AFDC beneficiaries but with much lower levels of added expense. For example, medium-cost cardiovascular problems add \$4,444 for beneficiaries with disability but \$2,345 for AFDC adults. We suspect that the effects of CDPS diagnoses are much larger for persons with disabilities because any given chronic illness or disability is more severe or more advanced, on average, for persons with disabilities than for AFDC beneficiaries. In addition, the average poorer health or functional status of adults with disabilities may increase their need for medical services during an illness. Comparing expenditure effects for AFDC adults and children, the picture is mixed, with only a small majority of categories showing higher expenditures for adults. For example, the skeletal, renal, and infectious disease categories show higher expenditures for adults, while the psychiatric and cancer categories show higher expenditure effects for children.

Maternity and Mental Health Services

An important element of health care for the AFDC population is for pregnancy, delivery, and neonates. There is little advantage, however, in using prospective health-based payment to cover maternity and neonatal care. Most of the costs associated with pregnancy, delivery, and neonates are incurred within the year. Thus, when variables for pregnancy, delivery, or premature birth are used in prospective regressions, they are associated with very small additional expenditures in the following year, far less than the costs associated with even a routine delivery. Use of prospective

weights for maternity and neonatal care would cause States to significantly underpay plans with a disproportionately high share of beneficiaries who give birth. Some States already use supplemental payments on top of capitation to cover the cost of each delivery, and we recommend that this approach be continued under health-based payment. Alternatively, weights for maternity and neonatal care that we have calculated on concurrent regressions could be substituted in the prospective model. Similarly, the high costs of neonatal care might be paid for on a supplementary basis.

Many States do not make mental health services the responsibility of HMOs and carve expenditures for such services out of the capitation. The main effect of re-estimating the regression to reflect this policy is to sharply reduce the coefficients for the psychiatric and substance abuse categories with little change in coefficients in other categories.

Predicted Expenditures Using Multiple Diagnoses

The predicted total expenditure for an individual is the sum of the baseline payment and the additional expenditure amounts for all of the diagnostic subcategories in which an individual is counted. (The total would also include the expenditures associated with an age-sex variable and any interaction between age and diagnostic categories.) Because many individuals with serious diagnoses are counted in multiple categories, the average predicted expenditures for individuals in most CDPS categories are substantially higher than the parameter estimates shown in Table 6. For example, the average predicted payment for beneficiaries in the medium-cost cardiovascular CDPS category is \$13,044, although the additional expenditure associated with this category is \$4,444; similarly,

the average predicted expenditure for beneficiaries with medium-cost Type 1 diabetes is \$12,084, while the regression coefficient is only \$3,787.

We have presented a model that assumes that the effects of diagnoses in different major CDPS categories are additive. We assume, for example, that the effects of having cardiovascular and central nervous system problems simultaneously are equal to the sum of the effects of having each of these problems individually. But this assumption of additivity may be incorrect: In some cases, having two or more problems might be more expensive than the sum of the independent effects, while in other cases, having two or more problems might be less costly than the sum of the independent effects.

As a rough test of the additivity assumption, we examined the actual and predicted expenditures for beneficiaries with disability categorized by the number of CDPS subcategories in which they have a diagnosis (Figure 4). On average, the additivity assumption appears reasonable: as the number of CDPS categories increases, both the actual and predicted expenditures increase. For the few beneficiaries with five to nine CDPS categories, actual expenditures are slightly higher than predicted, suggesting that the effects of having many different diagnoses are slightly greater than the sum of the individual effects. Preliminary work with a variety of interaction terms of CDPS subcategories yielded only slight improvements in overall model performance.

Goodness of Fit

Diagnostic information does a better job of predicting expenditures among persons with disabilities than among AFDC beneficiaries. As we have shown previously (Kronick et al., 1996), expenditures are

concentrated among a small number of high-cost beneficiaries among persons with disability, just as they are for AFDC beneficiaries: Among both AFDC beneficiaries and persons with disability, the most expensive 20 percent of beneficiaries account for 80 percent of expenditures. The difference between the two populations is the extent to which diagnostic information helps identify those who will be expensive. Diagnostic information does a better job for persons with disabilities because both the prevalence and the severity of diagnoses are much greater than among AFDC beneficiaries.

The statistical summary of this discussion is that the R^2 statistics are substantially higher for regressions on persons with disability than for AFDC beneficiaries. Estimated on our validation sample (the 25 percent of the data that was reserved from use while we were developing CDPS), the R^2 for the disabled is 0.18, compared with 0.08 for AFDC adults and 0.04 for AFDC children (Table 7). The difference in R^2 between AFDC adults and children results primarily from the lower prevalence of chronic illness among children and secondarily from smaller effects of diagnoses on expenditures for a variety of CDPS categories.

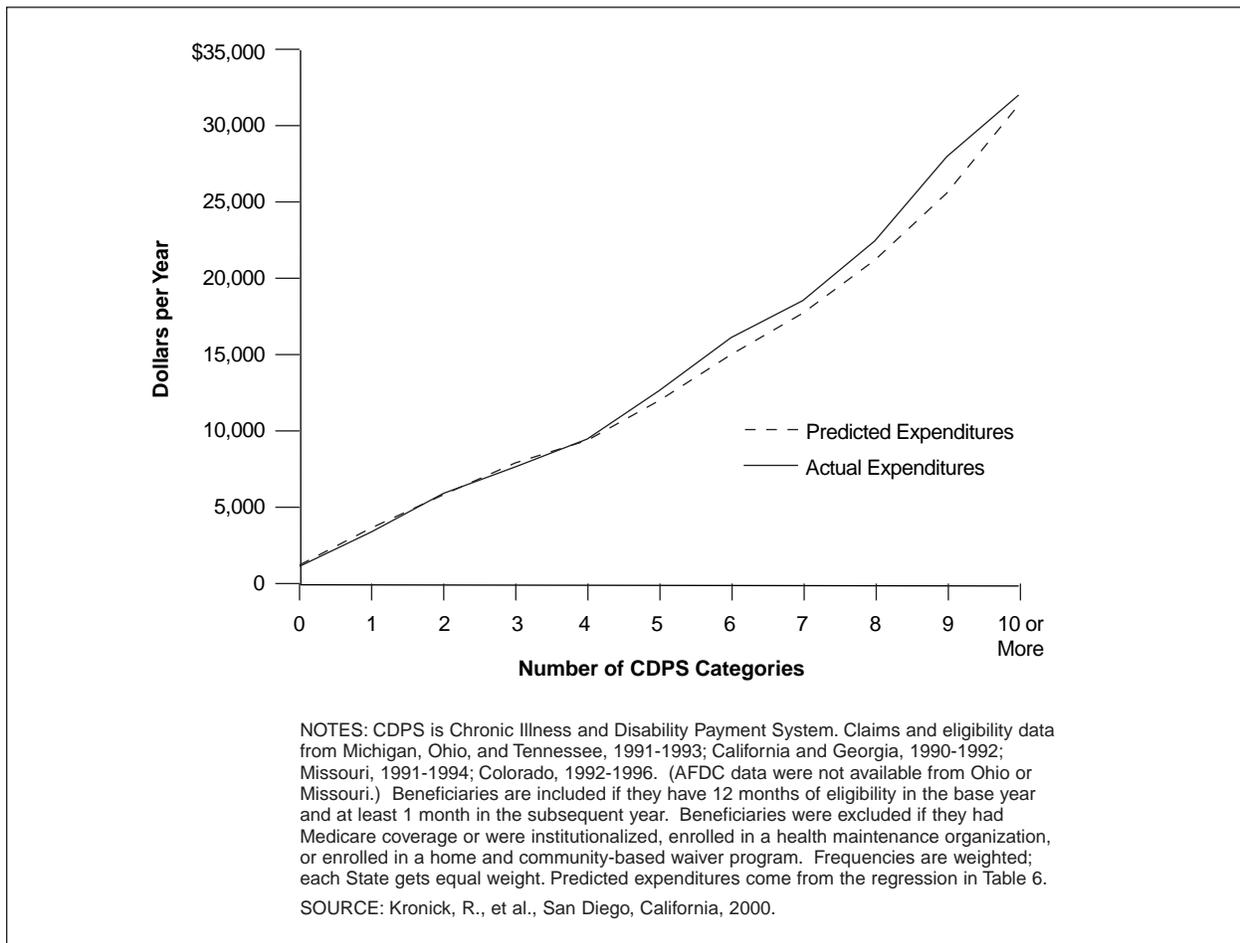
The R^2 for the validation sample is similar to the R^2 for the entire sample, suggesting that we did not significantly overfit the data.

State-Specific Versus Multistate Weights

A question of concern to a number of States is whether they should estimate payment weights on their own data or use payment weights already estimated on data from a number of States. To use the multistate weights in Table 4, a State would need to adjust them to reflect average expenditures in the State in order to ensure budget

Figure 4

Actual and Predicted Expenditures for Beneficiaries with Disability, by Number of CDPS Categories



neutrality. The decision about State-specific or multistate weights depends in part upon whether patterns of care in a given State are thought to be notably different from national patterns and in part on whether there are enough beneficiaries in the State to estimate payment weights reliably.

To explore this question, we use the 75-percent development sample to estimate separate sets of State payment weights using data from individual States, and sets of multistate weights that exclude each State in turn. We then used both the multistate weights and the State-specific weights to predict expenditures for beneficiaries from the validation sample grouped into simulated health plans (refer to the section “Comparison with Other Payment Systems” for a description of

the plans). In four of the seven States, the multistate weights led to better predictions than the State-specific weights; in one State, the predictions were very similar, and in two States, the State-specific weights were better. Only in California did the State-specific weights show a clear superiority, probably because of the very large number of beneficiaries and perhaps in part because of the small proportion of diagnoses for which the detailed fifth digit of the ICD-9-CM code was retained in California data. It is likely that all but the largest States would do better using payment weights estimated from multistate data.

A Medicaid program need not view payment weights as set in stone. Large changes in treatment costs due to changes

Table 7
R² Statistics for CDPS, by Beneficiary Group

Group	Persons with Disability	AFDC Adults	AFDC Children
Development Sample	0.191	0.087	0.032
Validation Sample	0.183	0.083	0.041

NOTES: CDPS is Chronic Illness and Disability Payment System. AFDC is Aid to Families with Dependent Children. The model development sample used 75 percent of the data. The validation sample used 25 percent of the data.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

in drugs and other technology might justify ad hoc modification of individual category weights.

COMPARISON WITH OTHER PAYMENT SYSTEMS

States and other payers implementing health-based payment could consider a variety of diagnostic classification systems that have been created for various payment and profiling purposes, but only three approaches are now publicly available that were designed for adjustment of capitated payments to health plans.⁶ The two approaches other than CDPS are the Diagnostic Cost Groups (DCGs) and the Adjusted Clinical Groups (ACGs). (In previous versions of the ACG models, ACG stood for Ambulatory Care Groups.) This section describes how CDPS differs from these other systems in method and the next section presents differences in predictive performance based on regressions we ran using the various models on our Medicaid data.

The new CDPS and recent versions of the DCG and ACG models are similar in important respects, and all three models can be used to adjust payments to health

⁶ A fourth, the Clinical Risk Groups, will soon be available from 3M Health Information Systems. A fifth, the Clinically Detailed Risk Information System for Cost (CD-RISC), was developed by Grace Carter et al., (2000).

plans with diagnostic data far more effectively than traditional risk adjustment through demographic data alone. The additive DCG model, known as the Hierarchical Condition Category (HCC) model, and CDPS are similar. The ACG model is substantially different from CDPS and the HCCs both in how it classifies diagnoses and in its inclusion of many ill-defined and high-frequency, low-cost diagnoses. As we show later, for persons with disability, CDPS performs better than the HCCs, which in turn perform better than the ACGs. For AFDC beneficiaries, the performance of the ACGs and CDPS is fairly similar, with HCCs performing somewhat less well.

We believe that CDPS is preferable to the HCCs both because CDPS performs better for people with disabilities and for TANF beneficiaries and because it organizes and distinguishes among diagnoses in a number of cases more appropriately. In addition, Medicaid payment weights estimated from a variety of State Medicaid programs are now available for the CDPS categories and not yet for the HCCs.⁷ We believe that both CDPS and the HCC model are preferable to the ACGs because the ACGs include many ill-defined and high-frequency, low-cost diagnoses that we think should not be used in a payment system and because its performance for people with disabilities is less good. In this section, we examine in detail some of the differences between HCCs and CDPS in order to stimulate further discussion about the variety of decisions involved in constructing health-based payment models. We then look more briefly at the ACG model.

⁷ The Medicaid payment weights for HCCs provided in Ash et al. (1998) were estimated on data from one State. These payment weights are estimates of the expenditure effects of diagnoses on a combined population of AFDC and SSI beneficiaries. We argued previously that it is important to have separate payment weights for the two population groups.

Comparison of CDPS with the HCCs

The HCCs are one of several DCG models. The HCC model organizes ICD-9-CM codes into 543 diagnosis definitions (“DxGROUPS”). Many diagnoses are defined simply by a three-digit ICD-9-CM code and all its subcodes, other diagnoses are defined by multiple three-digit codes, and still others by individual four-digit codes or combinations of three-, four-, and five-digit codes. Despite many differences in detail, the HCC diagnosis definitions and the CDPS single-diagnosis variables, such as heart failure, schizophrenia, or hemoglobin-S disease, are parallel concepts.

The HCCs then cluster their DxGROUPS into 118 “condition categories,” which correspond to the CDPS diagnostic subcategories. Despite differences in nomenclature and content, the HCC condition categories and CDPS diagnostic categories are fairly similar. For example, the HCC has a category for quadriplegia, which is very similar to the CDPS category central nervous system high-cost. Both the CDPS and the HCC categories function as dummy or zero-one variables. If an individual’s record contains a diagnosis code in one of the defined diagnoses in the category, the model initially sets the category to one for that individual; otherwise, the category is set to zero.

Both models count multiple diagnoses across categories that are different from each other, for example, a cardiovascular diagnosis and a psychiatric diagnosis. Thus, both models share the assumption that the cost effects of multiple different types of diagnoses should be added together in order to produce an accurate prediction of total expenditures. (By contrast, the Principal Inpatient DCG model predicts expenditures using only the single most serious diagnosis made in the inpatient setting.⁸)

Both models also limit the counting they do within body systems or types of disease. Like CDPS, which counts only the most serious subcategory in a major category, the HCC model uses single counting in many areas. But in a few diagnostic areas, for example metabolic disorders, HCCs impose no counting rules, so that a different diagnosis in each of several condition categories can be counted. For several areas, the HCC model uses special counting rules that compromise between single counting and unrestricted counting. For the heart disorders, up to five different condition categories can each be counted separately. With separate HCC categories for vascular disease, a maximum of seven diagnoses could be counted, where CDPS would count only the single most serious cardiovascular subcategory. We considered multiple counting of this kind for CDPS in some areas but rejected it because it rewards increased coding while giving very little benefit in improved predictive power.

We think that one advantage of CDPS is its more precise identification and categorization of a number of medium- and high-cost diagnoses. Distinguishing a more expensive diagnosis from other related but less expensive diagnoses can help make better cost predictions and get resources to plans that attract people with greater needs. For example, the HCCs combine all of the diagnoses of hereditary hemolytic anemias, including hemoglobin-S disease, which we found to have far higher cost

⁸ The Principal Inpatient DCG model has been selected by HCFA for its current implementation of health-based payment for Medicare’s managed care reimbursement. Although health-based payment with only inpatient data sharply reduces the amount of data required for implementation, almost all analysts, including the creators of the DCG models, support a prompt movement by Medicare to adjustment by diagnoses from both inpatient and ambulatory settings. Using both sources of diagnostic data brings much greater accuracy and, more importantly, avoids creating inappropriate incentives to hospitalize and inappropriate penalties for plans that have successfully reduced hospital use (Iezzoni and Ayanian, 1998). We strongly concur (Dreyfus and Kronick, 1999).

implications than the other hereditary hemolytic anemias. In CDPS, hemoglobin-S disease is put in our very-high-cost hematological category, which is associated for 4,546 disabled beneficiaries with \$13,874 per year in additional cost. In the HCC model, the hereditary hemolytic anemias are grouped with other blood and immune disorders, which together are associated in our sample with only \$6,504 of additional cost. (The estimates cited in this section for HCCs come from a regression we have estimated using HCCs. Further information on the regression is provided later.)

Among pulmonary conditions, we found cystic fibrosis associated with unusually high costs for the disabled and placed it in the subcategory of very-high-cost pulmonary diagnoses, which predicts \$13,586 per year. For adults, the HCCs have grouped cystic fibrosis with other fibroses and other chronic lung disorders, a category that has an annual coefficient of only \$2,844 when we ran the HCC model on our Medicaid disabled sample. For children, the HCC model does recognize the high costs of cystic fibrosis, putting it with other diagnoses in a separate category for very high-cost pediatric disorders, but it is associated with only \$4,440 per year. Other examples of conditions we categorized separately from neighboring codes, where the HCCs do not, include decubitus ulcer, bacterial endocarditis, primary pulmonary hypertension, acute cor pulmonale, and certain coagulation disorders (congenital factor VIII and factor IX disorders). Our larger data set with so many disabled Medicaid beneficiaries may have allowed us to pick up notable differences in cost effects that were missed in the development of the HCCs.

In addition, we see one area in which the CDPS organization of diagnoses seems much more appropriate than that of the HCCs. The HCC model seems to have

erred in its counting for psychiatric and substance abuse diagnoses, which are grouped together in the area of mental disorders: If the single category for drug and alcohol psychosis or dependence is counted, then none of the psychiatric disorders are counted. Both our data and the HCC regression indicate that schizophrenia is much more costly than alcohol dependence among Medicaid beneficiaries. But, as a result of the HCC hierarchy, an individual with both alcohol dependence and schizophrenia would have less money paid for him or her than an individual with only schizophrenia. (Also, a plan that enrolls someone with schizophrenia would receive the same money as a plan enrolling someone with bipolar, affective, and other depressive disorders, which our analysis suggests would not be appropriate.)

CDPS may also have a modest advantage over the HCC model in retaining some helpful low-cost diagnoses. For the HCC payment model, more than 30 HCC categories were excluded from the payment model because of vagueness or because of suspected or observed lack of significant cost effects (Ash et al., 1998). But the elimination of vague or no-cost diagnoses was made using entire categories, and some useful diagnoses were lost from the HCCs. For CDPS, we evaluated diagnoses one at a time, based on the magnitude of their cost effects and the clarity of their definition. For example, the HCC Medicaid payment model excludes central nervous system infection and other infectious disease, losing diagnoses of polio, bacterial meningitis, almost all tuberculosis, herpes zoster, herpes simplex, and oral thrush—diagnoses that we included in various low-cost categories. Other conditions lost to the HCCs that we thought worthy of use include heart conduction disorders, cellulitis, adrenal gland disorders, and kidney infections and stones.

Comparison of CDPS with ACGs

The ACG models use a different approach from CDPS and the HCCs to predict expenditures from diagnostic data. Like CDPS and the HCCs, the ACG models assign diagnosis codes to various categories, which can be used to predict expenditures for individuals. But the ACG models differ sharply from the other two in the method of categorizing diagnoses. We see two problems in the ACG approach: poor separation of high-cost diagnoses from other conditions and inappropriate use of all ICD-9-CM codes, including many ill-defined and extremely low-cost diagnoses.

Nonetheless, ACG-based models have seen considerable use and have helped bring acceptance of case-mix analysis to health plans and payers. ACG models have been used by the State of Maryland and the Minnesota Buyers' Health Care Action Group for payment purposes and extensively by health plans for evaluating case mix. Maryland's use of an ACG model made it the first State to set diagnostically adjusted rates for large numbers of Medicaid beneficiaries.

The ACG approach assigns ICD-9-CM codes to 32 diagnostic categories called Ambulatory Diagnostic Groups (ADGs) based primarily on expectations about the condition's effect on individual health and resource needs (Johns Hopkins University, 1998). Likelihood of persistence, disability, reduced life expectancy, and need for diagnostic, specialist, therapeutic, and hospital care are considered to assign ICD-9-CM codes to the ADGs. As a result, many of the ADGs include conditions that appear unrelated but are thought to have similar effects on future resource use. For example, cerebral thrombosis and acute pancreatitis are both in the group for progressive conditions that are likely to recur.

Other groups are used for discrete conditions that are likely to recur, unstable chronic medical conditions, and time-limited minor psychosocial conditions. A few of the ADGs are based on more specific types of diagnoses, for example, separate groups for asthma, dermatologic conditions, malignancy, and pregnancy. Some groups are defined with combinations of resource expectation and type of condition, such as stable orthopedic conditions, unstable eye conditions, or unstable recurrent or persistent psychosocial conditions.

These diagnostic categories can be used in two ways for predicting expenditures. One approach is to use the ADGs much as one would the diagnostic categories of the CDPS or HCC models, with each ADG operating as a dummy or zero-one variable set to one if a diagnosis in the ADG is found in an individual's record. By regressing individual expenditure data against the counts of the ADGs for each individual, one can produce estimates of the additional future costs associated with the diagnoses in each of the ADGs.

Another approach is the standard recommended ACG model, which uses an algorithm to convert the dummy variables of the ADGs into more than 80 mutually exclusive categories (from 82 to 93 categories, depending upon some choices by the model user). These mutually exclusive categories, or ACGs, are defined through combinations of different classes of the ADGs and age variables. Individuals of the same age with diagnoses in multiple ADGs are classified into different ACGs depending upon how many of their ADGs are considered major. For example, for adults, the ADG for progressive, likely-to-recur conditions is considered major, while the ADG for allergies is not.

An advantage of the mutually exclusive ACGs is that they can be used in a traditional ratesetting fashion, with average

costs for past members of each group trended forward to predict the future costs of people assigned to the group. The ADG regression model works better in predicting expenditures than the standard ACG model, but like CDPS and the HCCs, it requires estimating a regression and setting up a system that can calculate an average casemix for each health plan.

Both the ADG and ACG models have the disadvantage of putting conditions with very dissimilar cost implications together into single groups. For example, the unstable chronic medical conditions mix together in one category cystic fibrosis, multiple sclerosis, unspecified cardiac dysrhythmia, unspecified heart disease, ulcer of lower limbs, and degeneration of intervertebral disc. Our analysis shows that these conditions have vastly different implications for future cost.

A more important problem is that the ACG models use all the ICD-9-CM diagnostic codes, many of which are ill-defined or so common that they could be elicited from almost any patient. A straightforward application of the ACG model would provide substantially more resources for individuals who have a medical visit for even the most minor problems. For example, for a disabled adult with a sore throat or headache, the ACG model run on our sample would add \$348 in annual payment and for a TANF adult, \$228 in annual payment. For another example, the diagnosis of a backache would add \$636 for a disabled adult and \$276 for a TANF adult. Similarly, a swollen ankle would add \$1,128 for the disabled adult and \$456 for the TANF adult. The combination of a backache and a swollen ankle would add \$1,356 for a disabled adult and \$636 for the TANF adult.

These additional amounts seem far in excess of what plans should receive for such relatively minor problems. Meanwhile, the increased payment for a

disabled adult or child with a diagnosis of sickle-cell anemia or cystic fibrosis would be only \$3,828. The inclusion of the high-frequency, low-cost conditions helps the model in simulated predictive accuracy but rewards plans for increased coding of minor problems and does relatively little to direct resources to plans that enroll a needier-than-average membership. The problem of modest additional payments for high-cost diagnoses can be attenuated if the ACGs are implemented with separate payments for specified high-cost diagnoses, as was done in Maryland and as is recommended by the developers of ACGs (Weiner, 1998).

Predictive accuracy is improved by using high-frequency, low-severity conditions, because past users of even small amounts of service are more likely than non-users to seek services in the future. Although Medicaid programs might well want to encourage contact between physicians and beneficiaries, attempting to accomplish this through health-based payment seems unwise. Using so many very-low-cost diagnoses in setting rates also requires a huge volume of diagnoses and exacerbates problems of auditing and of favoring plans that report better than others.

Comparison of Statistical Performance

We present comparisons of statistical performance for six models: our CDPS model, an ADG model, an ACG model, an HCC model, our original DPS model, and a CDPS “catch-all” model, in which we include variables for a set of extremely low-cost and ill-defined diagnoses excluded from our recommended payment model.

For each model we estimated three regressions: one for beneficiaries with disabilities (adults and children combined), one for AFDC adults, and one for AFDC

children. We estimated these regressions on the validation sample—the 25 percent of the data set aside before CDPS model development. For the HCC model, we use the subset of 85 HCCs that Ash et al. (1998) recommend for the Medicaid prospective payment model.⁹ In the ADG model, we included a number of ADGs that have statistically significant, negative parameter estimates and that would likely be excluded if an ADG payment model were implemented (it would be awkward to reduce plan payments because of additional diagnoses). Keeping these ADGs in the model, however, improves statistical performance, and we were eager not to disadvantage any of the models.¹⁰

For persons with disability, the CDPS regression has a significantly higher R^2 than HCCs, 0.183 compared with 0.143, and almost twice the predictive power of ACGs, 0.183 compared with 0.098 (Table 8). For persons with disabilities, the revised CDPS and the original DPS have similar explanatory power, and the catch-all CDPS model does not increase the R^2 statistic by much compared with the recommended CDPS model. Among AFDC adults, the R^2 statistics from the various models are more similar than for persons with disability, but the CDPS R^2 is slightly higher than the HCC statistic, and higher than the ACG and ADG R^2 statistics as well. The revisions in CDPS to make it more sensitive for an AFDC population are evident in the improvement in R^2 from the original DPS model. Predictive performance can be

⁹ In each model we include the same set of age-sex dummy variables. When estimating HCCs for persons with disability, we include the set of child interactions suggested by Ash et al., (1998). When estimating for AFDC adults and children separately, no interactions of age and HCC categories are needed. The HCC model estimated with our data set produces a better result in predicting expenditures than would the published HCC model parameter estimates.

¹⁰ In the ADG regression for persons with disability, we include interactions of each ADG with age. When estimating the original DPS, we do not include interactions with age. When estimating the CDPS catch-all model, we include the same set of age interactions as is included for CDPS.

Table 8
 R^2 from Alternative Models, by Beneficiary Group

Model	Persons with Disability	AFDC Adults	AFDC Children
CDPS	0.183	0.083	0.041
HCC	0.143	0.080	0.031
ACG	0.098	0.069	0.031
ADG	0.111	0.077	0.042
DPS	0.175	0.065	0.037
CDPS, Catch-All	0.188	0.093	0.047

NOTES: AFDC is Aid to Families with Dependent Children. CDPS is Chronic Illness and Disability Payment System. HCC is Hierarchical Condition Category. ACG is Adjusted Clinical Group. ADG is Ambulatory Diagnostic Group. DPS is Disability Payment System. All regressions estimated on the 25-percent validation sample.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

improved measurably by inclusion of a variety of ill-defined and high-frequency, low-severity diagnoses, as indicated in the catch-all model. Overall R^2 statistics are much lower among AFDC children than for AFDC adults, and the models show similar explanatory power.

A more important issue than R^2 is whether the models get the right amount of money to plans with biased enrollment. Does it matter which system is used, or do the leading classification systems give much the same results when used for payment? To explore this question, we simulated a universe of health plans by assuming that beneficiaries with expensive diagnoses will sort themselves disproportionately into different plans. To avoid favoring any one system, we identified high-cost diagnoses using three-digit ICD-9-CM code groups as regression variables rather than the diagnoses in any one system's high-cost categories. We selected as high-cost diagnoses the 100 highest cost ICD-9-CM codes with more than 100 disabled beneficiaries in the validation sample. We also identified the 100 most common ICD-9-CM codes with very moderate cost effects (between \$30 and \$100 per month).

We then constructed five hypothetical health plans by assigning disproportionate shares of beneficiaries with the high-cost

diagnoses and with the moderate cost diagnoses to the five plans. (The use of the very moderate-cost diagnoses was intended to favor systems such as the ACGs and the CDPS catch-all model, which should do a good job of predicting costs for plans that attract disproportionate shares of people with light diagnoses.) For beneficiaries with disability, the five simulated plans include one plan with actual expenditures that are almost 1.6 times the average expenditures, one plan with expenditures 1.3 times average, one plan very close to the average, one plan slightly below average, and one plan with very healthy enrollees (Table 9). Each of these simulated plans has about 50,000 members. (For AFDC adults, each of the 5 plans has approximately 80,000 members; for AFDC children, 180,000.)

For beneficiaries with disability, the CDPS model predicts expenditures that are very close to the actual expenditures in each simulated plan: It slightly underpredicts actual expenditures for the plan with sick enrollees and slightly overpredicts expenditures for the health plan with healthy enrollees. But the differences between actual and predicted expenditures are relatively small. The HCC model is close to the CDPS model in performance but does less well for the two extreme plans. The ACG model would not be nearly so satisfactory: It would underpay the plan with sick enrollees by 13 percent and overpay the plan with healthy enrollees by 19 percent. On average, the absolute value of the payment error for CDPS is 3 percent, for the HCCs 5 percent, for ACGs 10 percent. The ADGs would perform somewhat better than the ACGs but not so well as CDPS or the HCCs.

For AFDC adults, there is much less difference among models in performance. CDPS would do a good job of paying plans appropriately, with payments within 3 per-

cent of actual expenditures for three of the plans, and a 5-percent overpayment for one plan. HCCs perform similarly to CDPS but slightly worse for two of the five plans. ADGs perform very well for this set of plans; ACGs are slightly closer than CDPS to actual for plan 2, but slightly further away on plan 4 and the high-cost plan 1. The CDPS catch-all model has predictive expenditures very close to actual, indicating that it is possible to predict very accurately, if one is unconcerned about the use of ill-defined and low-severity diagnoses.

Performance comparisons for AFDC children are similar to AFDC adults. CDPS does reasonably well, though ADGs, like the CDPS catch-all model, get consistently closer to actual. HCCs appear to have a problem for AFDC children, with significant underpayments for the plan with sick enrollees and significant overpayments for the plan with healthy enrollees.

It appears from this work that there is some difference among the models in the extent to which they will pay appropriately, although the differences are not overwhelming. For beneficiaries with disability, CDPS appears to perform significantly better than the ACG-ADG models and somewhat better than the HCC model. For AFDC children and adults, CDPS appears to perform somewhat better than HCCs and has similar performance to ACGs. For both AFDC children and adults, both the ADGs and the CDPS catch-all model perform better than the other models, but the inclusion of all diagnoses in these models makes them less appropriate for payment purposes.

The performance comparisons are not conclusive, but they support the propositions that CDPS:

- Can be used to make equitable payments for both SSI and TANF beneficiaries.
- Performs better than the similar HCC model.
- Is significantly better than the ACG

Table 9
Predicted Expenditures for Simulated Health Plans, by Model and Beneficiary Group

Group and Plan	Actual Expenditures per Month	CDPS	HCCs	ACGs	ADGs	DPS	CDPS Catch-All
Beneficiaries with Disability							
1	1.58	1.52	1.49	1.38	1.43	1.51	1.52
2	1.31	1.28	1.27	1.21	1.24	1.27	1.28
3	0.96	0.97	0.97	0.97	0.97	0.97	0.96
4	0.82	0.83	0.83	0.91	0.89	0.83	0.84
5	0.59	0.63	0.66	0.70	0.66	0.65	0.62
Average Percentage Error	—	0.03	0.05	0.10	0.07	0.04	0.03
AFDC Adults							
1	1.25	1.21	1.21	1.18	1.21	1.18	1.22
2	1.17	1.12	1.11	1.14	1.15	1.07	1.14
3	0.99	1.00	1.00	1.00	1.00	1.00	1.00
4	0.86	0.89	0.89	0.90	0.88	0.91	0.88
5	0.81	0.85	0.86	0.85	0.83	0.89	0.83
Average Percentage Error	—	0.03	0.04	0.04	0.02	0.06	0.02
AFDC Children							
1	1.36	1.29	1.20	1.22	1.29	1.27	1.31
2	1.26	1.22	1.15	1.20	1.25	1.20	1.24
3	1.01	1.00	1.00	1.00	1.00	1.00	1.00
4	0.89	0.90	0.93	0.93	0.90	0.91	0.90
5	0.82	0.87	0.91	0.88	0.85	0.87	0.85
Average Percentage Error	—	0.03	0.07	0.06	0.02	0.04	0.02

NOTES: CDPS is Chronic Illness and Disability Payment System. HCC is Hierarchical Condition Category. ACG is Adjusted Clinical Group. ADG is Ambulatory Diagnostic Group. DPS is Disability Payment System. AFDC is Aid to Families with Dependent Children.

SOURCE: Kronick, R., et al., San Diego, California, 2000.

model for beneficiaries with disability and is similar to the ACGs for AFDC beneficiaries.

Given the much greater vulnerability of the ACG model to proliferative coding and the availability of CDPS payment weights estimated on multistate data, we see a clear overall advantage to Medicaid programs in the implementation of CDPS. In addition, CDPS is public-use software, and most leading actuaries that work with Medicaid programs have experience in the use of the closely related DPS model and software.

ADJUSTING TO CHANGE IN DIAGNOSTIC REPORTING

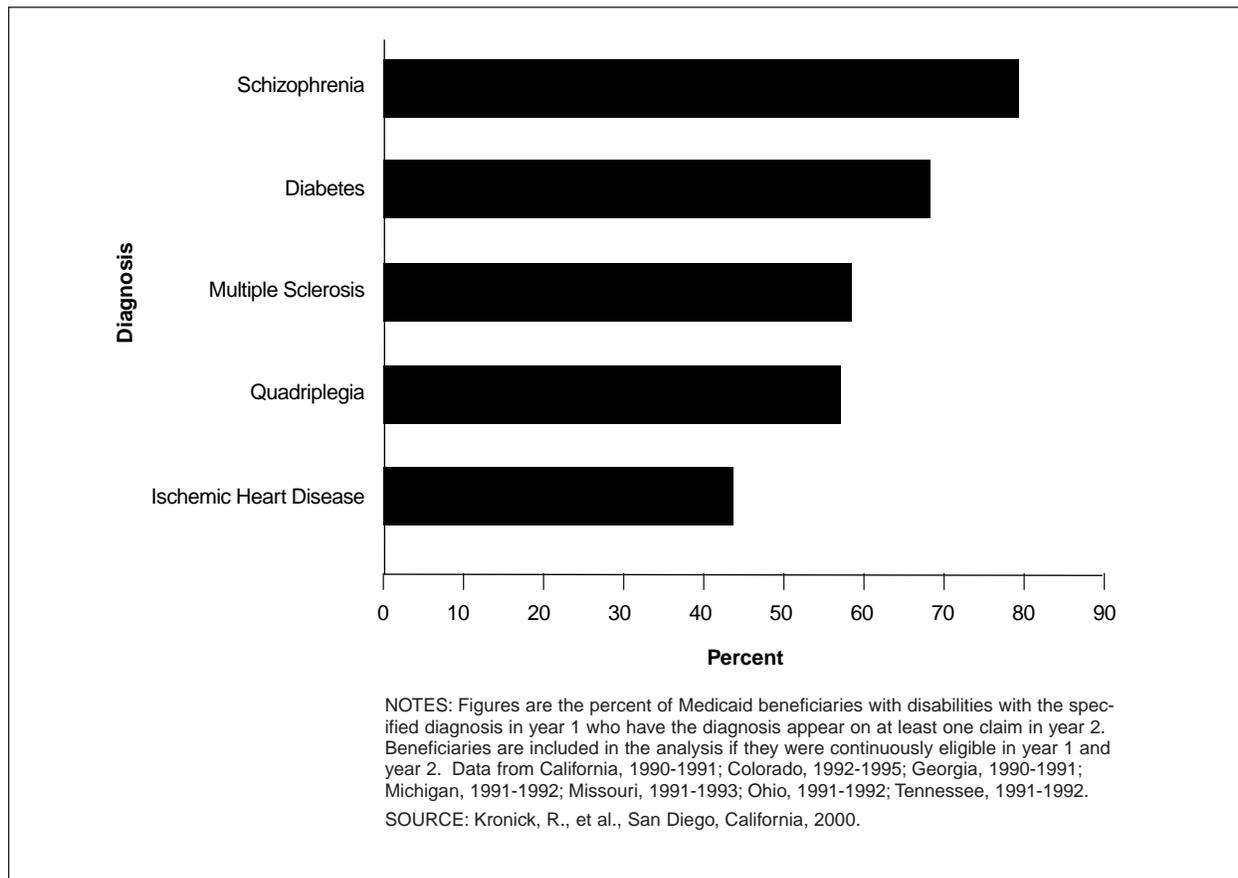
States face a number of important operational issues in implementing health-based payment: the timing of risk assessment and payment adjustment, the use of prospective or concurrent weights, evalu-

ating risk for newly eligible and dually eligible persons, and limiting profits and losses through risk-sharing or stoploss. We have addressed these issues in detail elsewhere (Kronick et al., 1996, Kronick and Dreyfus, 1997). In this section, we briefly address another implementation issue that also deserves attention.

One key challenge in implementing health-based payment is obtaining data that can be used to equitably measure health status across plans (Dreyfus and Kronick, 1999). We can aspire eventually to have data that accurately reflect enrollees' diagnoses, but we know that, in the short run, the data will be incomplete.

FFS data show significant underreporting of diagnoses. For many chronic diagnoses, between 20 and 60 percent of beneficiaries with these diagnoses appearing on a FFS claim in a given year do not have that diagnosis appear on a claim during the

Figure 5
Persistence of Diagnosis from Year 1 to Year 2



subsequent 12-month period (Figure 5). Examining beneficiaries with disabilities who were continuously eligible for Medicaid for a 24-month period, we find that 20 percent of beneficiaries with schizophrenia on at least one claim during the first 12 months do not have a diagnosis of schizophrenia during the second 12 months. And schizophrenia is the most persistent of the major diagnoses we have analyzed. The Medicare Payment Advisory Commission has found an even greater lack of persistence in Medicare fee-for-service data (Medicare Payment Advisory Commission, 1998).

As Medicaid, Medicare, and other payers begin to base payment on diagnoses, we expect that this lack of persistence will change. Plans being paid based on diag-

noses will certainly want to increase the number of diagnoses they report. One can imagine the enterprising health care consultant selling software that will maximize persistence: Before each scheduled visit, the software might search through last year's encounter records to identify the diagnoses that could lead to additional payment and then print a customized medical service record to prompt the physician to determine if any of these diagnoses are still present and could reasonably be construed as contributing to the need for care.

There are several potential responses to the problem of lack of persistence and error in measuring diagnoses. A payer might simply assume that, at least in the short run, plans will underreport more or less equally. In this case, payment will be

equitable, even though based on light reporting of diagnoses. Changes in coding practices will only come slowly in physician offices, especially if Medicaid is the only payer using ambulatory diagnoses. On the other hand, it is reasonable for payers and plans to be concerned that coding practices might be different in plans in which physicians are on salary or receiving subcapitation. In such environments, more active strategies by the payer may be needed.

One strategy is for payers to monitor coding practices at competing plans by comparing records for individuals over time. If the enrollees at one plan seem to be getting sicker significantly more quickly than enrollees at other plans, then it is likely that coding is changing. Similarly, if enrollees appear to get sicker at a faster-than-expected rate as they switch from fee-for-service into a particular health plan, then closer scrutiny of health-plan coding is needed. If some plans are clearly coding differently than others, then payment can be adjusted to offset these coding differences. Colorado uses such a process—a “data-reporting adjustment”—that was originally implemented to correct for underreporting but that works just as well to address increased intensity of diagnostic reporting (Tollen and Rothman, 1998). In addition, the auditing of diagnoses that is needed to keep health-based payment honest will also help keep reporting rates more consistent.

Finally, States could explore using longer periods for accumulating diagnostic information. Once a diagnosis of quadriplegia is made, for example, there is little need to have this diagnosis confirmed on an annual basis. We are beginning work on analyzing models with multiyear periods for accumulating diagnoses.

Although health-based payment can be started using diagnoses from FFS claims, its ongoing use depends upon health plans

reporting their members’ diagnoses. Some plans are concerned that health-based payment will require them to expend additional resources on gathering the data. Many plans can use their current systems to report diagnoses without too much additional effort, while others will need to make large improvements. Some plans may also be concerned that health-based payment will unfairly cause their revenues to decline.

Despite plans’ concerns, we remain convinced that the effort to implement health-based payment is well justified by the improved incentives to create equality for people with significant illness or disability. The experiences of the Colorado and Oregon Medicaid programs and the Washington State Health Care Authority, using HMO-supplied encounter data to assess health status, reinforce our opinion that gathering diagnostic data from inpatient and ambulatory encounters is feasible. The Medicare, Medicaid, and SCHIP Budget Reconciliation Act of 1999 directs the U.S. Department of Health and Human Services to study how to make diagnostic reporting work better. Although this and other elements of the Act may slow down the implementation of health-based payment in the Medicare program, further study on the issue of diagnostic reporting should, in the long run, be helpful.

CONCLUSION

We have shown that CDPS can be used with confidence by States to evaluate health status and pay HMOs equitably for both SSI and TANF Medicaid beneficiaries. We have argued previously that States contracting with HMOs for the care of persons with disabilities should implement health-based payment. The growing number of States that have begun health-based payment indicate that it can be done.

The question of whether health-based payment for TANF beneficiaries is worth the effort is harder to settle. The sizable effects of diagnoses on next-year expenditures for AFDC beneficiaries mean that health-based payment would be useful for discouraging risk selection and for motivating better services to TANF beneficiaries. But the potential of health-based payment to improve incentives for TANF beneficiaries is not so great as it is for SSI beneficiaries because a plan is less likely to attract a very biased selection of TANF beneficiaries through its service design and provider network. If a State is not using health-based payment for SSI beneficiaries, then the advantages of health-based payment for the TANF population may not be worth the administrative effort. But if a State is using health-based payment for SSI beneficiaries, then the additional work of implementation for TANF beneficiaries is likely worthwhile.

We provide a set of payment weights that States can use. We also provide a diagnostic profile of the SSI and AFDC beneficiaries that indicates the diversity of diagnoses and the greater level of illness and disability among the SSI population.

Our comparison of the taxonomy and statistical performance of the CDPS model to other leading classification systems indicates that CDPS has advantages in both structure and predictive accuracy for Medicaid beneficiaries. CDPS shows moderate advantage over the HCC model and significant advantage over the ACG model, which includes many ill-defined or extremely frequent diagnoses that improve statistical performance but reduce the reliability of health-based payment.

A key challenge in implementing health-based payment is getting data from health plans that can equitably be used in adjusting payments. In the transition to adjustment by diagnoses, payers and plans will have to make considerable efforts to ensure that reporting is consistent across plans. As health status measured by diagnosis is increasingly used to determine payment levels, more efforts will be needed to understand the process of diagnosis and to improve diagnostic reporting.

ACKNOWLEDGMENTS

We wish to acknowledge the contributions of our clinical consultants: Larry Berger, M.D., Stephen Bonfilio, Ph.D., Donald Duncan, M.D., Jonathan Epstein, M.D., A. Stuart Hanson, M.D., Martin Lee, M.D., Richard Lentz, M.D., Michael Macaulay, M.D., Hal Martin, M.D., Michael Shaw, M.D., Gregg Strathy, M.D., Janette Strathy, M.D., Frederick Taylor, M.D., Anton Willerscheidt, M.D., and especially Terry Wingert, M.D., and David Knutson, who helped arrange the clinical consultations. The clinical consultants, however, bear no responsibility for the content of the article. Finally, David Strom of the University of California at San Diego provided excellent support in the preparation of the article's figures and tables. We would also like to thank those who provided valuable comments on an earlier draft: Arlene Ash, John Bertko, Randall Ellis, Christopher Hogan, Sandi Hunt, Lisa Iezzoni, Tony Tucker, Jonathan Weiner, Pete Welch, and three anonymous reviewers.

REFERENCES

Ash, A.S., Ellis, R.P., Yu, W., et al.: *Risk Adjustment for the Non-Elderly*. Boston University, Department of Medicine. Report to the Health Care Financing Administration. Boston. June 1998.

Carter, G.M., Bell, R.M., Dubois, R.W., et al.: A Clinically Detailed Risk Information System for Cost. *Health Care Financing Review* 21(3):65-91, Spring 2000.

Dreyfus, T., and Kronick, R.: Paying Plans to Care for People with Chronic Illness. In Kronick, R., and de Beyer, J., eds: *Medicare HMOs: Making Them Work for the Chronically Ill*. Chicago. Health Administration Press, 1999.

Druss, B.G., Bradford, D.W., Rosenheck, R.A., et al.: Mental Disorders and Use of Cardiovascular Procedures After Myocardial Infarction. *Journal of the American Medical Association* 283(4):506-511, January 26, 2000.

Ellwood, M.R., and Lewis, K.: The Ins and Outs of Medicaid: Enrollment Patterns for California and Florida in 1995. Abstracts. Chicago: Association of Health Services Researchers, 1999.

Iezzoni, L., and Ayanian, J.: Paying More Fairly for Medicare Capitated Care. *New England Journal of Medicine* 139(26):1933-1938, December 24, 1998.

Johns Hopkins University School of Hygiene and Public Health, Health Services Research and Development Center.: *The Johns Hopkins University ACG Case Mix Adjustment System*. Documentation for PC-DOS and Unix. Baltimore, Maryland. January 1998.

Kronick, R., and Dreyfus, T.: *The Challenge of Risk Adjustment for People with Disabilities: Health-Based Payment for Medicaid Programs*. Princeton, NJ. Center for Health Care Strategies, 1997.

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, 1996.

Kronick, R., Zhou, Z., and Dreyfus, T.: Making Risk Adjustment Work for Everyone. *Inquiry* 32(1):41-55, 1995.

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.

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

Tollen, L., and Rothman, M.: Case Study: Colorado Medicaid HMO Risk Adjustment. *Inquiry* 35(7):154-170, Summer 1998.

Weiner, J.P., Tucker, A.M., Collina, A.M., et al.: The Development of a Risk-Adjusted Capitated Payment System: The Maryland Medicaid Model. *Journal of Ambulatory Care Management* 21(4):29-52, 1998.

Reprint Requests: Richard Kronick, Dept. of Family & Preventive Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0622. E-mail: rkronick@ucsd.edu

TECHNICAL NOTE

Regressions include an intercept term and age-sex dummy variables. Parameter estimates for the age-gender terms: In the regression for persons with disabilities: under 1 year of age, \$697; age 1-4, -\$630; age 5-15 male, -\$433; age 5-14 female, -\$387; age 15-24 male, -\$124; age 15-24 female, \$168; age 25-44 male, \$0 (reference category); age 25-44 female, \$348; age 45-64 male, \$395; age 45-64 female, \$730; intercept, \$1,126. For AFDC adults, age 18-24 male, -\$231; age 18-24 female, \$478; age 25-44 male \$0 (reference category); age 25-44 female, \$259; age 45-64 male \$496; age 45-64 female, \$595; intercept, \$624. For AFDC children, under 1 year of age, \$107; age 1-4, -\$8; age 5-15 male \$0 (reference category); age 5-14 female, -\$45; age 15-17 male, -\$125; age 15-17 female, \$537; intercept, \$400. The “baseline” amount shown in the bottom of Table 6 is the sum of the intercept plus the weighted average of the age-sex terms.

The regression for persons with disability includes interaction terms for age (coded as 1 for persons under age 19; 0 otherwise) and selected CDPS categories. The values for these interaction terms are: very high cardiovascular, -\$5,084; central nervous system, medium, \$1,201; very

high-cost pulmonary, \$5,239; high-cost gastrointestinal, \$4,679; medium-cost gastrointestinal, \$3,738; low-cost gastrointestinal, \$1,314; diabetes (all children are coded as Type 2, low), \$666; metabolic high, \$7,844; metabolic medium, \$947; infectious medium, \$5,457; hematological very high, -\$5,645.

For AFDC adults and children, very high and high-cost pulmonary conditions are combined. For AFDC children and children with disability, all diabetes categories are combined into a single category and all eye conditions are combined into a single category.

The weighted average annual expenditures for the groups were \$4,980 for persons with disability, \$1,884 for AFDC adults, and \$684 for AFDC children.

Using the significance statistics generated by the weighted least-squares regression, all coefficients shown in Table 6 are significant at $P < 0.0001$, except genital extra low-cost at $P < 0.01$ for the disabled. All the coefficients of the variables for interaction between age and diagnostic categories for the disabled described above were also significant at $P < 0.0001$, except diabetes low-cost at $P < 0.05$. However, because the expenditure data are not normally distributed, the true standard errors and P values are higher.