MEDICAL CARE
Volume 39, Number 11, pp 1188–1202
©2001 Lippincott Williams & Wilkins, Inc.

The Medicaid R_x Model

Pharmacy-Based Risk Adjustment for Public Programs

TODD GILMER, PHD,* RICHARD KRONICK, PHD,* PAUL FISHMAN, PHD,† AND THEODORE G. GANIATS, MD,*

BACKGROUND. Risk adjustment models typically use diagnoses from claims or encounter records to assess illness severity. However, concerns about the availability and reliability of diagnostic data raise the potential for alternative methods of risk adjustment. Here, we explore the use of pharmacy data as an alternative or complement to diagnostic data in risk adjustment.

OBJECTIVES. To develop and test a pharmacy-based risk adjustment model for SSI and TANF Medicaid populations.

RESEARCH DESIGN. Pharmacological review combined with empirical evaluation. We developed the Medicaid R_x model, a system that classifies a subset of the National Drug Codes into categories that can be used for risk-assessment and risk-adjusted payment.

SUBJECTS. Subjects consisted of 362,370 persons with disability and 1.5 million AFDC and TANF beneficiaries in California, Colorado, Georgia, and Tennessee during 1990–1999.

MEASURES. We compare pharmacy and diagnostic classification for three chronic diseases. We also compare R^2 statistics and use simulated health plans to evaluate the performance of alternative models.

RESULTS. Pharmacy and diagnostic classification vary in their ability to identify specific chronic disease. Using simulated plans, diagnostic models are better at predicting expenditures than are pharmacy-based models for disabled Medicaid beneficiaries, although the models perform similarly for TANF Medicaid beneficiaries. Models that combine diagnostic and pharmacy data have superior overall performance.

CONCLUSIONS. The performance of risk adjustment models using a combination of pharmacy and diagnostic data are superior to that of models using either data source alone, particularly among TANF beneficiaries. Concerns regarding variations in prescribing patterns and the incentives that may follow from linking payment to pharmacy use warrant further research.

Key words: Risk adjustment; pharmacy; Medicaid. (Med Care 2001;39:1188–1202)

Substantial effort has been made to develop and implement health-based payment systems that use diagnoses reported during inpatient stays and outpatient visits to adjust payments to health plans based on the health status of their enrollees.[1–4] These payment systems are designed to encourage health plans to develop systems of care that are responsive to those most in need, to reward those plans and providers that care for sicker than average enrollees, and to attenuate the strong financial in-

*From the Department of Family and Preventive Medicine, University of California, San Diego.

†From the Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, Washington.

Financial support was received from CalOptima and the Center for Health Care Strategies, Inc.

Address correspondence and reprint requests to: Todd Gilmer, PhD, Department of Family and Preventive Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0622. Email: tgilmer@uscd.edu

Received July 10, 2000; initial review October 16, 2000; accepted June 26, 2001.

centives for health plans to attract healthy patients if payment is not adjusted by health status.

Diagnostic information has several advantages over demographics or prior expenditure when used to adjust payments: diagnoses are strongly related to current and future expenditures; payment based on diagnoses is not much influenced by utilization levels (that is, payment based on diagnoses does not simply reward greater utilization of services); and diagnoses are routinely collected and reported in fee-for-service data and in many HMO data systems.

There are two main concerns regarding the use of diagnostic information in health-based payment. First, some payors have not been collecting encounter data from health plans, and thus do not have any diagnostic information on health plan enrollees. Second, even when these data are collected, the quality and completeness of diagnostic data are uncertain. Health plans making capitated payments to provider groups and plans with salaried physicians may not even have accurate information on the number of encounters. Although a study of health plans in Sacramento, California shows encounter data to be sufficient for risk adjustment, a remaining general uncertainty regarding the quality of encounter data and the potential for models based on combined sources of data suggest that we consider other methods of risk adjustment.[5,6]

We have been exploring the use of pharmacy data as an alternative or complement to diagnostic data. Similar to diagnostic data, pharmacy data are predictive of future expenditure and a pharmacy based payment system would not simply reward greater health care utilization. Unlike diagnostic data, which involve some degree of uncertainty and discretion in the assignment of ICD-9 diagnosis codes, the National Drug Codes (NDCs) used in pharmacy data are linked to a specific product and a clinical course of action. Also, in some settings pharmacy data may be available but diagnostic information may not be. For example, CalOptima, an Orange County, California Health Insuring Organization has been making capitated payments to health networks since 1996, but pays directly for prescription drugs. As a result, CalOptima has complete dispensing data, but incomplete diagnostic information from contracting organizations. There are other state Medicaid programs in a similar situation. In this article, we examine the potential for health-based payment within public programs using pharmacy-dispensing data.

Several pharmacy-based risk assessment models have been developed, including those by Clark et al,[7] Roblin,[8] Lamers,[9] and Fishman and Shay.[10] In our view the most useful starting point is the Chronic Disease Score (CDS) model.[7] The Medicaid R_x model we present here is a refinement of CDS.

CDS is a pharmacy-based risk assessment model developed at the Center for Health Studies, Group Health Cooperative of Puget Sound (GHC). A team of physicians, pharmacists, and health services researchers identified drugs in the GHC formulary that were clearly linked to the treatment or management of chronic conditions among adults. A similar method was later used to develop a Pediatric CDS model.[10] A person filling a prescription for a drug included in the algorithm is considered to have the chronic condition associated with that drug. Their chronic disease score is based on a regression model that reflects the likely future health services use associated with their age, sex, and CDS profile.

CDS was developed and estimated exclusively within the GHC staff model delivery system, based on data from a commercial population. However, much of the activity related to actual adjustment of payments made to health plans has occurred in the public sector among state Medicaid programs where it has been argued that health-based payment is both important and feasible.[1,11–13] This paper presents a model based on the CDS methodology, but that has been revised to reflect the patterns of disease in Medicaid populations.

Materials and Methods

The Medicaid R_x model was developed with the goal of improving the performance of a pharmacy-based classification system for Medicaid populations, including beneficiaries eligible due to medical disability or receipt of Social Security Income (SSI) and those eligible under Temporary Assistance to Needy Families (TANF). As a beginning, we adopted the basic structure of the combined CDS and Pediatric CDS.

The development of the Medicaid R_x model involved both a pharmacological review and an empirical evaluation. The pharmacological review involved reading clinical monographs for all drugs in the CDS system, as well as for drugs not in CDS, but in therapeutic classes related to CDS categories. For the empirical evaluation, we used data on 362,370 persons with disabilities and 1.5

million AFDC and TANF beneficiaries across four states and 9 years to examine the existing system and to identify additional drugs that were predictive of future cost.

We estimate the Medicaid R_x model and evaluate its performance relative to a diagnostic-based classification system. For the diagnostic comparison, we use the Chronic Illness and Disability Payment System (CDPS), a risk adjustment model that groups diagnoses according to chronic and disabling disease. CDPS assigns ICD-9 diagnoses from claims or encounter data to one of 56 subcategories in 18 major categories relating to bodily systems (pulmonary) or prevalent disease (cancer). These categories are hierarchical, so that a person may have multiple categories across major groups (diabetes and cardiovascular disease) but only a single subcategory per major group. CDPS was developed using Medicaid claims data and is currently the risk adjustor most widely applied by state Medicaid programs.

Pharmacological Review

For the pharmacological review, we used Clinical Pharmacology (CP),[14] a reference of clinical monographs, and the American Hospital Formulary Service (AHFS) Drug Information,[15] a clinical reference that also assigns drugs to therapeutic classes. We first reviewed CP monographs for each drug included in CDS to confirm that the drug is still used for the conditions indicated by CDS and to determine if there are other conditions for which the drug is commonly used. Next we used AHFS to identify all other drugs in therapeutic classes that included at least one CDS drug and then reviewed the CP monographs to determine the use of each of these additional drugs. In conducting this review, we consulted with pharmacist and physician specialists to determine the organization and appropriateness of drugs included in or excluded from the new system.

As a result of this review, we added some drugs and excluded some others. For example, we added the newer atypical dibenzodiazepines to the antipsychotics, added four new protease inhibitors to human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS), created new categories for PCP Pneumonia and Hepatitis, expanded anti-Parkinsonian agents to include those prescribed for tremor, and created a category for Alzheimers. We excluded from HIV/AIDS a macrolide antibiotic recently approved for the treatment of certain ulcers and community acquired pneumonia in children. Following Lamers,[9] we dropped from Cystic Fibrosis enzymes used as adjunct therapies in the absorption of other agents and combined Renal Disease with the End Stage classification.

We also combined categories where we found drugs that were often used together to treat the same underlying disease. For example, lithium is often prescribed along with antipsychotics to treat some manifestations of schizophrenia (this is supported in our data where we find that 73% of disabled adults filling a prescription for lithium also receive an antipsychotic). Similarly, certain antidepressants are effective for anxiety, and cardiovascular related diseases are often treated with a polypharmacy of β-blockers, diuretics, and vasodilators. To provide a more clinically valid description of population-based illness and to reduce the effects of prescribing patterns on reimbursement, Medicaid R_x has a single category for cardiovascular disease that includes the CDS categories for vascular disease, heart disease, heart failure, and hypertension. Medicaid R_x also has a single category for psychotic and bipolar illness that is hierarchical above a single category for depression and anxiety.

Empirical Review

The empirical review used Medicaid claims data to suggest additional categories that are predictive of future costs in Medicaid populations. We used data on SSI and TANF Medicaid beneficiaries in California, 1995 to 1999; Colorado, 1992 to 1996; Georgia, 1990 to 1992; and Tennessee, 1991 to 1993 (Table 1). We excluded from this analysis Medicaid beneficiaries on whom we were likely to have incomplete information on expenditure due to either third party coverage (including beneficiaries dually eligible for Medicaid and Medicare) as well as beneficiaries enrolled in HMO plans that did not provide claims or encounter data (the decline in the number of California TANF beneficiaries over time is due to mandated HMO enrollment in several counties). We also excluded those beneficiaries residing in nursing homes or enrolled in home and community based waiver programs because these beneficiaries are rarely enrolled in health plans accepting capitated payments.

We estimate a prospective regression, in which drug information in an initial (base) year is used to

TABLE 1. Medicaid Beneficiaries by State and Year

State	Year	Disabled	TANF Adults	TANF Children
California	1995	14,362	21,110	44,741
	1996	15,069	14,184	30,516
	1997	14,798	7,023	16,679
	1998	14,946	4,688	12,191
Colorado	1992	15,221	25,810	55,566
	1993	17,338	24,610	54,793
	1994	17,631	20,727	46,391
	1995	15,965	14,074	30,799
Georgia	1990	49,962	66,830	181,977
	1991	55,424	81,879	231,352
Tennessee	1991	65,484	58,893	185,136
	1992	66,170	61,729	204,244
Total	1990–1998	362,370	401,557	1,094,385

predict expenditure in a subsequent (rate) year. To ensure complete drug information, we included in the analysis only beneficiaries with a full year of Medicaid eligibility in the base year, and at least 1 month of eligibility in the following rate year. Parameters are estimated by linear regression. Although nonlinear risk-adjustment models can provide better individual-level prediction than linear models, the largest gains to nonlinear modeling occur when sample sizes are small or when there are a few extreme cost cases.[16] We follow the convention in the literature an estimate regression models using weighted ordinary least squares.

The dependent variable for this analysis is annualized expenditure for services typically included in an acute care HMO benefit package, including inpatient and outpatient costs, drug costs, and lab, x-ray, and ancillary services, but excluding dental services and long-term care. We calculate individual expenditure by summing the amount paid by Medicaid for these services by person, per year. We then divide this amount by the number of months eligible in the year and multiply by 12. We normalize expenditure, dividing by the state and year specific means to account for inflation across years and cross sectional differences in the both price of health services and the generosity of Medicaid coverage across states. We also weight observations to account for differing lengths of program eligibility in the rate year and differences in practice patterns and sample sizes across states. The weight is a product of the estimated variation in expenditure that is due to length of enrollment and the sample size of each state. Persons with longer periods of eligibility have larger influence on the parameter estimates and each state has an equal influence on the final results.

To determine additional candidates for the Medicaid R_x model, we used the Multum Lexicon,[17] to combine the NDC codes not included in the Medicaid R_x model into approximately 1,000 categories based on drug names. We added these categories to the set of Medicaid R_x model covariates, and regressed the entire set on subsequent year expenditure. We then reviewed 72 categories that had at least 100 persons and a statistically significant parameter estimate among disabled adults. We found a number of drugs indicated primarily for chronic conditions that were associated with substantial expenditure in the following year, and added several new categories: Alcoholism, Burns, Folate and Iron Deficiency Anemias, Gallstones, Glaucoma, Herpes, Multiple Sclerosis/Paralysis, Nausea, Neurogenic Bladder, and Osteoporosis/Paget's bone disease. We performed a similar exercise using the AHFS therapeutic class designations and added categories for Infections, Irrigating Solution, Replacement Solution, and for Eye, Ear, Nose, and Throat (EENT) Disorders in TANF Children. Table 2 provides a list of Medicaid R_x model categories.

In contrast to CDS, the Medicaid R_x model includes a few categories that are predictive of future expenditure, but that are not necessarily related to a specific chronic disease. For example,

TABLE 2. Medicaid R_x Model Categories

Medicaid R_x	Summary Drug Descriptions	Restrictions
Alcoholism	Disulfiram	
Alzheimers	Tacrine	Adults only
Anticoagulants	Heparins	
Asthma/COPD	Inhaled glucocorticoids, bronchodilators	
Attention deficit disorder	Methylphenidate, CNS stimulants	TANF Children only
Burns	Silver Sulfadiazine	
Cardiovascular	Ace inhibitors, beta blockers, nitrates, digitalis, vasodilators	
Cystic fibrosis	Pancrelipase	
Depression/anxiety	Antidepressants, antianxiety	
Diabetes	Insulin, sulfonylureas	
Eyes, ears, nose, throat (EENT)	Anti-infectives for EENT related conditions	TANF Children only
ESRD/renal	Erythropoietin, Calcitriol	
Folate deficiency	Folic acid	
Gallstones	Ursodiol	Adults only
Gastric acid disorder	Cimetidine	
Glaucoma	Carbonic anhydrase inhibitors	Adults only
Gout	Colchicine, Allopurinol	Adults only
Growth hormone	Growth hormones	Children only
Hemophilia/von Willebrands	Factor IX concentrates	
Hepatitis	Interferon beta	
Herpes	Acyclovir	
HIV/AIDS	Antiretrovirals	
Hyperlipidemia	Antihyperlipidemics	Adults only
Infections, high	Aminogycosides	
Infections, medium	Vancomycin, Fluoroquinolones	
Infections, low	Cephalosporins, Erythromycins	
Inflammatory/autoimmune	Glucocorticosteroids	
Insomnia	Sedatives, Hypnotics	
Iron deficiency	Iron	
Irrigating solution	Sodium chloride	
Liver disease	Lactulose	
Malignancies	Antinoeplastics	
Multiple sclerosis/paralysis	Baclofen	
Nausea	Antiemetics	
Neurogenic bladder	Oxybutin	
Osteoporosis/pagets	Etidronate/calcium regulators	Adults only
Pain	Narcotics	
Parkinsons/tremor	Benztropine, Trihexyphenidyl	Adults only
PCP pneumonia	Pentamidine, Atovaquone	Disabled only
Psychotic illness/bipolar	Antipsychotics, lithium	
Replacement solution	Potassium chloride	
Seizure disorders	Anticonvulsants	
Thyroid disorder	Thyroid hormones	
Transplant	Immunosuppressive agents	
Tuberculosis	Rifampin	
Childrens cardiovascular	Interaction: Cardiovascular, 18 and younger	
Childrens HIV/AIDS	Interaction: HIV/AIDS, 18 and younger	
Childrens MS/paralysis	Interaction: MS/Paralysis, 18 and younger	

irrigating solutions are used to clean burn wounds or decubitus ulcers, but are not indicative of a specific chronic disease. Further, some of the categories included in the Medicaid R_x model not only can be used for more than one underlying disease, but also may be used primarily for acute conditions. For example, the infection categories consist mainly of antibiotics used to treat acute, rather than chronic, infections. Although the infections may be acute, the data indicate that persons using them in 1 year have higher costs in the subsequent year. Most diagnostic based risk adjustment models use diagnoses for acute infections as well. For example, the model proposed to risk adjust payments to HMOs enrolling Medicare beneficiaries would pay an additional $1,300 per year for each person diagnosed with viral pneumonia.[2]

Estimating the Medicaid R_x model

Following the approach we have taken in previous work, we estimate regression coefficients separately for three groups based on age and category of assistance: persons with disability, TANF adults, and TANF children. We base this decision on three criteria. First, these three groups have different base rates of utilization and expenditure. In our sample, adults and children with disability have an average annual expenditure of $4,400; TANF adults $1,800; and TANF children $660. Second, sample size considerations suggest combining the disabled adults and children, where only 20% of the population is children and the entire group is relatively small; but not TANF, where 65% are children and each group is relatively large. Third, when the groups are combined by aid category with interactions of Medicaid R_x categories with age, we find few significant interactions in the disabled (we retain three statistically significant interactions with similarly sized effects in each state—Cardiovascular, HIV/AIDS, and Multiple Sclerosis/Paralysis). More than half the categories in a combined TANF regression would require interactions.

Some Medicaid R_x categories are not applied to children. For example, children generally do not have glaucoma, gallstones, gout, osteoporosis, hyperlipidemia, or Parkinsons; in the rare instances in which a drug in these categories is prescribed for a child, we ignore the prescription in estimating regression coefficients. Similarly, we exclude growth hormones when assigned to adults. Other drugs are prescribed so rarely among TANF adults (eg, those for hemophilia, Parkinsons, PCP pneumonia), or among TANF children (eg, those for alcoholism, hemophilia, hepatitis, PCP pneumonia, tuberculosis) that we are unable to estimate a reliable regression coefficient. ADHD is dropped for the disabled and TANF adults because it had a small negative coefficient and we did not want to discourage treatment. Drugs for Alzheimers are included in the Medicaid R_x model as we expect that they are associated with substantial expenditure, although they are not found in our younger-than-65 Medicaid data and consequently we are unable to estimate a regression coefficient for Alzheimers.

The Restricted Medicaid R_x model

Some categories of drugs in the full Medicaid R_x model are especially susceptible to gaming. Thus, although we consider them to be clinically valid (they tell us something about illness) and predictive of expenditure, their inclusion in a risk adjustment model would create incentives for increasing prescriptions and potentially for overuse. We are primarily concerned with two types of medications: those that are prescribed for persons with severe illness but that are also commonly used either for less severe manifestations or even prevention; and medications that are highly susceptible to practice pattern variation.

A good example of this first type is cimetidine, prescribed for gastric acid disorder (GAD), but commonly purchased over the counter to treat or prevent heartburn. If physicians were to be reimbursed substantially for prescribing cimetidine, it is reasonable to assume that many will encourage their heartburn susceptible patients to patronage the pharmacy rather than purchasing their medication over the counter at their local drugstore. Similar concerns surround the use of pharmacy data to reimburse for folate and iron deficiency anemias (treated with folic acid and iron) and eye, ear, nose, and throat disorders (EENT) in children.

A second type of medication is one that lacks a clear consensus for use, and for which prescriptions are likely to reveal variations in practice patterns rather than illness severity. These are medications for insomnia (sedatives and hypnotics), pain (narcotics), and some antibiotics (erythromycins, cephalosporins). Although including

these drugs in a model is useful for profiling, using them for health-based payment may result in an inequitable redistribution of resources, or, at worst, over medication of patients. Thus, for purposes of health-based payment, we suggest a restricted version of the model that excludes GAD, folate and iron deficiency anemias, EENT, insomnia, pain, and low-cost infections.

Evaluating the Medicaid R_x Model: Comparison to Diagnostic-Based Classification

Identifying disease alternatively through pharmacy and diagnostic data will show different results. We examine the relative merits of each approach using three analyses. The first analysis compares the average health care costs of disabled Medicaid beneficiaries identified by drugs to those identified by diagnoses for three chronic conditions that are highly prevalent both among Medicaid beneficiaries and other populations: diabetes, cardiovascular disease, and mental illness. The second analysis compares individual-level predictions for four drug and diagnostic models using R^2 statistics: Medicaid R_x, restricted Medicaid R_x, CDPS, and CDPS combined with the restricted Medicaid R_x. The third analysis evaluates the predictive ability of these at the group level using simulated health plans.

For the health plan simulation, we randomly divide our data into a 75% estimation sample and a 25% validation sample. Using the 75% sample, we estimate regression coefficients for the four models. We then sample without replacement from the validation data, assigning persons to a set of hypothetical health plans based on their prior year expenditure. The probability that a person is assigned to any given plan varies with the decile of their prior year expenditure. For example, a person in the highest decile of prior year expenditure is more likely to be assigned to the highest cost plan than a person in any of the lower deciles of prior year expenditure. The probabilities of selection were adjusted to create a set of 5 health plans: two with adverse selection, one with average selection, and two with favorable selection. We compute actual expenditures for the enrollees in each hypothetical plan, and normalize these expenditures to 1.0.

We apply the coefficients estimated using to 75% sample to the drug-based and diagnostic-based categories in the validation sample, normalize to 1.0 across persons, and summarize actual and predicted expenditure for each hypothetical plan and risk adjustment model. We also calculate a summary measure of risk adjustment performance that is the mean absolute difference between predicted and actual expenditures across the simulated plans. We believe that this approach more realistically approximates the selection process of persons to health plans that simply dividing persons by decile or quartile of prior year expenditure.

Results

The prevalence of Medicaid R_x categories by eligibility status is shown in Table 3. As expected, Medicaid R_x categories are generally more common among the disabled than among TANF adults or children. For most chronic diseases, we find high rates among the disabled, somewhat lower rates among TANF adults, and very low rates among children: 25% of the disabled receive medications in cardiovascular, compared to 10% of TANF adults and 0.5% of TANF children; 19% of the disabled receive medications for depression/anxiety, compared to 13% of TANF adults and 1% of TANF children; and 8% of the disabled receive medication for diabetes, compared to 2% of TANF adults and 0.3% of TANF children. Asthma/COPD is a notable exception, where we identify 13% of the disabled, 8% of TANF adults, and 12% of TANF children.

Several rare, yet severe conditions appear predominantly among the disabled: the prevalence of cystic fibrosis is 0.67%; ESRD/renal disease is 0.23%; hemophilia is 0.04%, multiple sclerosis/paralysis is 0.66%; and transplantations are 0.37%. Similarly, the disabled are most likely to be prescribed medications for the most severe infections: HIV/AIDS, PCP pneumonia, tuberculosis, and high-cost infections. Severe mental illness (SMI) is a common cause of disability, and we find that 13% of disabled Medicaid beneficiaries receive medications for schizophrenia/bipolar illness, compared to 1% of TANF adults and 0.3% of TANF children (some of these drugs are also used to treat severe emotional disturbance in children).

Prospective regression coefficients for the Medicaid R_x model are shown in Table 4. These coefficients are the expected annual marginal cost associated with a particular category. Hemophilia/

TABLE 3. Prevalence of Medicaid Beneficiaries in Medicaid R_x Categories

Medicaid R_x	Disabled	TANF Adults	TANF Children
Alcoholism	0.42	0.23	0.00
Alzheimers	0.00	0.00	0.00
Anti-coagulants	1.52	0.26	0.02
Asthma/COPD	13.01	8.02	12.07
Attention deficit disorder	—	—	1.32
Burns	0.69	0.38	0.42
Cardiovascular	24.78	9.73	0.47
Cystic Fibrosis	0.67	0.28	0.03
Depression/anxiety	19.24	12.86	1.01
Diabetes	7.59	2.08	0.33
Eyes, ears, nose, throat (EENT)	—	—	14.55
ESRD/renal	0.23	0.02	0.00
Folate deficiency	0.96	0.42	0.06
Gallstones	0.08	0.03	—
Gastric acid disorder	15.71	8.80	0.58
Glaucoma	1.71	0.28	—
Gout	1.15	0.33	—
Growth hormone	0.06	—	0.01
Hemophilia/von Willebrands	0.04	0.00	0.00
Hepatitis	0.05	0.04	0.00
Herpes	0.91	1.12	0.24
HIV/AIDS	0.42	0.05	0.01
Hyperlipidemia	3.74	1.05	—
Infections, high	0.14	0.03	0.01
Infections, medium	6.83	6.17	0.29
Infections, low	45.33	53.92	54.86
Inflammatory/autoimmune	9.20	6.56	7.06
Insomnia	4.25	1.59	0.21
Iron Deficiency	2.27	5.36	2.05
Irrigating Solution	0.05	0.00	0.00
Liver Disease	0.22	0.01	0.02
Malignancies	0.89	0.19	0.03
Multiple sclerosis/paralysis	0.66	0.09	0.01
Nausea	11.54	9.94	8.28
Neurogenic bladder	0.93	0.19	0.09
Osteoporosis/pagets	0.31	0.02	—
Pain	27.27	35.14	6.37
Parkinsons/tremor	6.77	—	—
PCP pneumonia	0.03	0.00	0.00
Psychotic illness/bipolar	13.14	0.92	0.27
Replacement solution	5.68	1.76	1.19
Seizure disorders	11.16	1.11	0.56
Thyroid disorder	3.56	1.67	0.08
Transplant	0.37	0.03	0.01
Tuberculosis	0.16	0.06	—
Childrens cardiovascular	0.86	—	—
Childrens HIV/AIDS	0.03	—	—
Childrens MS/paralysis	0.12	—	—
Percentage with no MRX category	24.0	20.1	33.2

Note (—): These MRx categories are excluded from these aid categories.

TABLE 4. Prospective Regression Coefficients: Annual Marginal Cost by Medicaid R_x Category

Medicaid R_x	Disabled	TANF Adults	TANF Children
Alcoholism	2,574	576	—
Alzheimers	—	—	—
Anticoagulants	4,511	2,478	10,556
Asthma/COPD	1,554	614	342
Attention deficit disorder	—	—	674
Burns	2,427	577	90
Cardiovascular	1,127	641	1,018
Cystic fibrosis	2,794	440	2,527
Depression/anxiety	1,028	884	1,233
Diabetes	2,184	1,656	1,361
Eyes, ears, nose, throat (EENT)	—	—	224
ESRD/renal	13,355	4,387	5,684
Folate deficiency	3,524	667	1,307
Gallstones	4,411	1,804	—
Gastric acid disorder	1,499	891	1,070
Glaucoma	1,070	989	—
Gout	468	595	—
Growth hormone	20,709	—	21,938
Hemophilia/von Willebrands	88,169	—	—
Hepatitis	9,605	3,291	—
Herpes	1,345	360	222
HIV/AIDS	10,490	7,320	12,807
Hyperlipidemia	534	1,130	—
Infections, high	15,838	3,726	7,293
Infections, medium	2,462	798	840
Infections, low	340	276	177
Inflammatory/autoimmune	1,063	376	266
Insomnia	1,180	798	931
Iron deficiency	1,769	576	285
Irrigating solution	8,756	3,881	23,783
Liver disease	4,481	4,395	2,656
Malignancies	3,709	3,029	10,117
Multiple sclerosis/paralysis	4,358	2,750	1,679
Nausea	872	530	134
Neurogenic bladder	3,756	1,091	597
Osteoporosis/pagets	9,746	5,455	—
Pain	592	273	162
Parkinsons/tremor	1,523	—	—
PCP pneumonia	6,960	—	—
Psychotic illness/bipolar	2,831	1,838	1,993
Replacement solution	1,957	443	549
Seizure disorders	2,356	1,358	1,428
Thyroid disorder	519	541	614
Transplant	5,904	3,437	6,025
Tuberculosis	6,942	1,326	—

(*continues*)

TABLE 4. (Continued)

Medicaid R_x	Disabled	TANF Adults	TANF Children
Childrens cardiovascular	2,144		
Childrens HIV/AIDS	8,497		
Childrens MS/paralysis	11,515		
Age under 1	3,654		105
Age 1 to 5	2,116		−24
Male 5 to 14	212		—
Female 5 to 14	191		−29
Male 15 to 24	−166	−157	119
Female 15 to 24	34	766	737
Male 25 to 44	—	—	
Female 25 to 44	−307	280	
Male 45 to 64	−83	478	
Female 45 to 64	−456	288	
Intercept	1,196	386	286
R^2	15.25	10.87	5.85

Note (—): These MRx categories are excluded from these aid categories.

Von Willebrands is by far the most expensive at $88,000 per year. ESRD/renal disease is also a costly illness at $13,000 for the disabled and approximately $5,000 for the TANF populations. Growth hormone deficiency is expensive among both disabled and TANF children, approximately $21,000.

Some of the more common chronic conditions show similar expenditure effects across aid categories. The cost of diabetes ranges from $1,400 for TANF children to $2,200 for the disabled; depression/anxiety is least expensive for TANF adults at $900 and most expensive for TANF children at $1,200; cardiovascular disease is $600 for TANF adults, approximately $1,000 for disabled adults and TANF children, but more than $3,000 for disabled children (the expenditure effect for disabled children includes an interaction effect with age). Psychotic/bipolar illness costs Medicaid programs an additional $1,800 to $2,800 per beneficiary with SMI.

Serious infections are highly expensive among the disabled. High-cost infections are $16,000, hepatitis is $10,000, and PCP pneumonia and tuberculosis are each $7,000 per year. The marginal cost of HIV/AIDS is relatively high for all aid categories: $19,000 for disabled children, $10,000 for disabled adults, $7,000 for TANF adults, and $13,000 for TANF children. The average predicted costs for HIV/AIDS will be higher than the marginal predicted cost, because it will include age effects and marginal expenditures for comorbid conditions. The average predicted cost for disabled Medicaid beneficiaries with HIV/AIDS is $20,000.

Some of categories that we have added to the original CDS model are shown to predict substantial additional expenditure. Multiple sclerosis/paralysis predicts $16,000 for disabled children, $4,400 for disabled adults, $2,800 for TANF adults, and $1,700 for TANF children. Osteoporosis/Pagets bone disease predicts $9,700 for disabled adults and $5,500 for TANF adults. Disabled Medicaid beneficiaries receiving irrigating and replacement solutions cost an additional $8,800 and $2,000 respectively. Disabled Medicaid beneficiaries treated for burns cost an additional $2,400 per year.

Pharmacy Versus Diagnostic Identification

Next we compare pharmacy to diagnostic classification for three chronic diseases that are prevalent among Medicaid beneficiaries: diabetes, mental illness, and cardiovascular disease. For each disease, we identified disabled Medicaid beneficiaries that were either assigned to a Medicaid R_x category, because they filled a prescription for an associated drug, or to a CDPS category, because they received a related ICD-9 diagnosis on a fee for service claim. CDPS categories are hierarchical,

TABLE 5. Chronic Illness Case Studies Among Disabled Medicaid Beneficiaries: Average Annual Expenditure for Persons Identified by Either Diagnoses or Drugs

	Diabetic Coma	Type 1	Type 2 with complications	Type 2 without complications	No Diagnosis
Diabetes					
Diabetes	$26,184	$10,128	$9,288	$6,804	$6,900
	1.0%	21.3%	4.4%	41.1%	10.4%
No MRx	$32,700	$9,288	$12,300	$6,528	
	0.1%	1.7%	1.1%	18.8%	
	Schizophrenia	Bipolar illness	Depression/ anxiety	No Diagnosis	
Mental Illness					
Psychotic illness/bipolar	$7,596	$7,056	$6,648	$4,788	
	17.7%	2.5%	6.1%	9.1%	
Depression/anxiety	$7,140	$6,876	$6,624	$6,744	
	1.3%	0.5%	13.0%	40.2%	
No MRx	$4,200	$5,880	$4,656		
	2.0%	0.3%	7.3%		
	Transplant	CHF	AMI	Hypertension	No Diagnosis
Cardiovascular Disease					
Cardiovascular	$31,560	$11,232	$8,328	$5,460	$6,288
	0.5%	8.2%	17.9%	28.7%	28.2%
No MRx	$22,656	$9,408	$7,056	$4,884	
	0.6%	1.2%	9.1%	5.6%	

as are the Medicaid R_x categories for mental illness. Table 5 presents a matrix for each disease that shows for each diagnostic and pharmacy combination, the percentage identified and average costs. The average cost for all disabled Medicaid beneficiaries is $4,400.

Diabetes presents a scenario where the diagnostic method identifies disease and models severity better than the pharmacy method. Disabled adults experiencing diabetic coma, or with type 1 diabetes or type 2 diabetes with complications, cost more than those with type 2 diabetes without complications; expenditures in these groups are similar for those receiving and those not receiving a diabetic medication. Among all disabled adults identified with diabetes by either drugs or diagnoses, 10% do not have a diagnosis for diabetes in a given year and 22% have no record of filling a prescription for a diabetes drug. Of those without a prescription, 87% are diagnosed with type 2 diabetes without complications, a condition that may be managed by diet and exercise.

Mental illness illustrates the strengths of the pharmacy method in an area of health services where there exists severe under-reporting of diagnoses. Of the disabled diagnosed with a mental disease, 81% have a record of receiving a psychotropic medication. However, 55% of beneficiaries filling prescriptions do not receive a diagnosis related to mental illness. Most the undiagnosed (82%) receive medications used to treat depression and anxiety. However, 26% of beneficiaries prescribed anti-psychotics or lithium have no diagnosis of mental disease. Finally, although there is some discrimination of costs by diagnostic category, the largest costs differential is between those receiving versus those not receiving a psychotropic medication, where average costs are approximately $6,700 and $4,600, respectively.

If diabetes is a disease for which the diagnostic method provides superior identification and differentiation of cost, and mental illness is an area where the pharmacy-based method has better capture of both illness and severity, then cardiovascular disease illustrates the strengths and weaknesses of each system. Of disabled Medicaid beneficiaries identified with a cardiovascular disease by either method, 17% are not recorded as filling a prescription for a cardiovascular related medication and 28% do not receive a cardiovascular related diagnosis. A substantial number of persons identified and receiving treatment are missed by either method. Also, average costs are increasing by severity of diagnosis and by medi-

TABLE 6. R^2 Comparison

	MRx	MRx − R	CDPS	CDPS + MRx
Disabled	15.3	14.6	23.6	25.7
TANF adults	10.9	9.6	12.4	15.1
TANF children	5.9	5.4	6.6	8.7

Note: R^2 values are from regressions on the full sample.

cation use within severity of diagnosis. A "best" model for risk adjustment of cardiovascular disease may resemble this diagnostic-pharmacy matrix, where payment is based on a combination of diagnoses and pharmacy use.

Individual-level Prediction

A summary measure of the individual level predictive ability of a risk adjustment model estimated by linear regression is the R^2. In Table 6, we compare the R^2 statistic for four models: Medicaid R_x, restricted Medicaid R_x, CDPS, and CDPS combined with the restricted Medicaid R_x. We learn three things from this simple analysis. First, the loss in individual predictive ability when moving to the restricted to the unrestricted Medicaid R_x model is relatively minor, an average 8% decline across models. Second, for persons with disabilities, the much higher R^2 statistics for the CDPS model compared to the Medicaid R_x model indicates that diagnoses carry significantly more information about future health care needs than does pharmaceutical data for this population; among TANF beneficiaries, there is a smaller difference in R^2 between the pharmaceutical and diagnostic approaches. Third, among persons with disabilities, adding pharmaceutical information to a diagnostic model does not add much explanatory power; in contrast, for TANF beneficiaries, the combination of the two data types increases explanatory power substantially.

TABLE 7. Actual and Predicted Expenditure for Medicaid R_x and CDPS Using Simulated Health Plans

Aid category						
Disabled	Plan	Actual Expenditure	MRx	MRx − R	CDPS	CDPS + MRx
	1.00	1.51	1.26	1.25	1.31	1.35
	2.00	1.17	1.09	1.09	1.12	1.13
	3.00	1.00	1.00	1.00	1.00	1.00
	4.00	0.78	0.88	0.89	0.86	0.84
	5.00	0.67	0.83	0.84	0.79	0.77
	Mean absolute error		0.12	0.13	0.09	0.07
TANF adults	Plan	Actual Expenditure	MRx	MRx − R	CDPS	CDPS + MRx
	1.00	1.22	1.13	1.10	1.13	1.16
	2.00	1.09	1.05	1.04	1.05	1.07
	3.00	0.99	1.00	1.00	1.00	1.00
	4.00	0.91	0.95	0.96	0.95	0.93
	5.00	0.85	0.91	0.93	0.91	0.89
	Mean absolute error		0.05	0.06	0.05	0.03
TANF children	Plan	Actual Expenditure	MRx	MRx − R	CDPS	CDPS + MRx
	1.00	1.35	1.17	1.13	1.18	1.21
	2.00	1.15	1.07	1.06	1.07	1.09
	3.00	1.01	1.00	1.00	1.00	1.00
	4.00	0.83	0.92	0.93	0.91	0.90
	5.00	0.75	0.89	0.91	0.88	0.86
	Mean absolute error		0.10	0.12	0.09	0.08

Plan Simulation

When paying health systems, the important measure of the performance of health-based payment is whether it gets the right amount of money to diverse plans that have disproportionately healthy or disproportionately sick enrollment. An evaluation of pharmacy and diagnostic classification systems should answer the question of whether, in practice, it matters which system is used, or whether the methods produce similar results when used in a payment system. In Table 7, we compare Medicaid R_x, restricted Medicaid R_x, CDPS, and CDPS combined with the restricted Medicaid R_x by calculating predicted and actual expenditure for a set simulated health plans, and by summarizing model performance using the mean absolute error between predicted and actual expenditure across plans. We find that pharmacy based classification performs well relative to diagnostic based classification, and that a combined model proves to have the best overall performance. The major strength of the diagnostic model, as shown in the predictive ratios and summarized by the mean absolute error, is its ability to more distinctly differentiate higher from lower cost disabled Medicaid beneficiaries. The addition of pharmacy data improves this further. Pharmacy and diagnostic models perform similarly for TANF populations, and again a combined model shows the best predictive ability. Overall, the restricted Medicaid R_x performs similarly to the full Medicaid R_x. Given this small difference and our concerns regarding the appropriateness of including certain types of drugs in a payment model, we recommend a restricted model for use in health-based payment.

Discussion

Pharmacy information has a number of important advantages, as well as some significant disadvantages, relative to diagnostic information for use in health-based payment. A major advantage of pharmacy data are that is much more complete than diagnostic information. For Medicaid beneficiaries without third party coverage, we can be confident that almost every prescription filled by the beneficiary will be reflected in the data. In contrast, because physicians have not historically been paid based on diagnosis, we know that diagnostic information is often incomplete. Lack of persistence from year to year in the diagnosis of serious chronic conditions provides evidence of incomplete diagnostic recording—for example, of the Medicaid beneficiaries with a diagnosis of quadriplegia recorded in 1 year, more than 40% will not have the diagnosis of quadriplegia appear on any claim during the subsequent 12 months.

A second advantage of pharmacy data is that methods of physician payment should have little effect on the comparability of data across plans. A major concern about diagnostic information is that if some health plans pay physicians using a discounted fee-for-service system whereas other plans pay sub-capitated rates to provider groups, the plans using discounted fee-for-service may have a more complete record of encounters and better information on diagnoses than the plans using capitated payments. Similarly, health plans that pay physicians a fixed annual salary may have less complete information on diagnoses than plans paying fee-for-service. Further, the comparability of data across health plans and provider groups may be affected by the specificity (and number) of diagnoses that are listed on the 'super bill' used by the provider group, and on the procedures in place for coding 'other' diagnoses if the physician does take the time to write them. Differences across plans and provider groups in both methods of payment and of collecting diagnostic information has the potential for creating inequities when diagnoses are used in health-based payment.

A third advantage of pharmacy data are that it is available on a more timely basis than diagnostic information: for many health plans and insurers, a 6-month lag is needed to have reasonably complete information from health care encounters, whereas pharmacy data are available soon after the prescription is filled.

Pharmacy data have disadvantages as well. A technical concern is that new drugs come on the market much more quickly than new ICD-9 codes are adopted; further, the uses of pharmaceuticals change more quickly than the uses of ICD-9 codes. Pharmacy based models will require more frequent updating than diagnosis based models. In addition, the use of 'home-grown' pharmacy codes in some state Medicaid programs may necessitate state-specific adaptations of pharmacy models.

These technical concerns can be largely resolved by devoting sufficient resources to model development and maintenance.

More worrisome are static and dynamic problems related to equity and incentives. The static concern is that pharmacy-based risk adjustment may reward those plans and providers that prescribe drugs liberally, and punish those that have adopted more conservative prescribing practices. The small area analysis literature has taught us that there are wide variations across providers in how similar patients are treated. Most of this literature analyzes variations in hospital and ambulatory care, and not in prescribing patterns. However, it seems likely that there are wide variations across physicians (and physician groups) in the propensity to prescribe pharmaceuticals for a given patient. To the extent that significant practice pattern variation exists, a pharmacy-based payment system may reward those plans and physicians with liberal prescribing practices rather than those that take care of people with greater than average health care needs.

Even more serious are concerns about incentives to increase the use of pharmaceuticals. For many drugs included in Medicaid R_x, the increased reimbursement from prescribing a drug is far greater than the cost of providing the drug itself. Especially in the many scenarios in which a physician may be uncertain about whether to prescribe a given pharmaceutical, knowing that deciding to prescribe the drug will result in a big increase in payments to the health plan may well change physician prescribing behavior. In some scenarios this may be desirable—for example, in the decision to prescribe β-blockers for patients who have had a heart attack. In other scenarios this may be less desirable—for example, the increased use of antibiotics. A countervailing incentive is the potential for costly and harmful overdosing. Physicians take on high financial and malpractice risks if they seek to increase dispensing of drugs for high cost conditions. Concerns about effects on prescribing patterns will be heightened if a large part of the market is using pharmacy-based risk adjustment, and will be attenuated if only Medicaid is adjusting payment based on pharmacy data: physicians do not change their prescribing patterns easily, and are not likely to do so if a relatively small share of their income is affected.

Given the superior performance of the combined model, we recommend future research into the use of risk adjustment models that use both diagnostic and pharmacy data. Not only is a combined model more predictive of future expenditure, but it also incorporates the advantages of pharmacy data, such as completeness, timeliness, and standardization across plan and payer types, with the more detailed disease description provided by diagnostic data. Such a hybrid model should be more equitable and less susceptible to gaming. Additional research is needed to determine the effect of practice pattern variation on payment rates using pharmacy and combined pharmacy and diagnostic risk adjustment models. The incentive concerns in a dynamic environment can only be fully addressed through a demonstration of a pharmacy-based payment system.

Acknowledgments

The authors acknowledge the assistance of Kristin Gericke, PharmD, Cheri Rice, and Richard Whittaker. Helpful comments and suggestions were received from reviewers. All remaining errors are the responsibility of the authors.

References

1. **Kronick R, Gilmer T, Dreyfus T, et al.** Improving payment for TANF and SSI Medicaid recipients: The chronic illness and disability payment system. Health Care Fin Rev 2000;21:29–64.

2. **Ash A, Ellis R, Pope G, et al.** Using diagnoses to describe populations and predict cost. Health Care Fin Rev 2000;21:7–28.

3. **Carter G, Bell R, Dubois R, et al.** A clinically detailed risk information system for cost. Health Care Fin Rev 2000;21:65–92.

4. **Weiner J, Dobson A, Maxwell S, et al.** Risk-adjusted medicare capitation rates using ambulatory and inpatient diagnoses. Health Care Fin Rev 1996;17:77–100.

5. **William M. Mercer, Inc.** Joint Purchasers' Risk Assessment Project: Phase I. Report. August, 2000.

6. **Wray, Hollingsworth J, Peterson N, et al.** Case-mix adjustment using administrative databases: a paradigm to guide future research. MCRR 1997;54:326–356.

7. **Clark DO, Von Korff N, Saunders K, et al.** A chronic disease score with empirically derived weights. Med Care 1995;33:783–95.

8. **Roblin DW.** Physician profiling using outpatient pharmacy data as a source for case mix measurement and risk adjustment. J Ambulatory Care Manage 1998;21:68–84.

9. **Lamers LM.** Risk-adjusted capitation payments: developing a diagnostic cost groups classification for the Dutch situation. Health Policy 1998;45:15- 32.

10. **Fishman PA, Shay DK.** Development and estimation of a pediatric chronic disease score using automated pharmacy data. Med Care 1999;37:874- 83. 27

11. **Weiner J, Tucker A, Collins A, et al.** The development of a risk-adjusted capitation payment system: the Maryland Medicaid Model. J Ambulatory Care Manage 1998;21:29–52.

12. **Kronick R, Dreyfus T, Lee L, et al.** Diagnostic risk adjustment for Medicaid: the disability payment system. Health Care Fin Rev 1996;17:7–33.

13. **Kronick R, Z. Zhou, T. Dreyfus.** Making risk adjustment work for everyone. Inquiry 1995;32:41–55.

14. **Clinical Pharmacology.** Gold Standard Multimedia, 2000.

15. **AHFSfirst.** First Databank, 2001.

16. **Gifford G, Manning WG, Finch M, et al.** Risk adjustment for persons with disabilities: evaluating functional form. Report, 1999.

17. **Multum Lexicon.** Multum Information Services, 2001.