MEDICAL CARE
Volume 41, Number 1, pp 84–99

Risk Adjustment Using Automated Ambulatory Pharmacy Data

The RxRisk Model

PAUL A. FISHMAN, PHD,* MICHAEL J. GOODMAN, PHD,† MARK C. HORNBROOK, PHD,‡
RICHARD T. MEENAN, PHD,‡ DONALD J. BACHMAN, MS,‡ AND
MAUREEN C. O'KEEFFE ROSETTI, MS‡

OBJECTIVES. Develop and estimate the RxRisk model, a risk assessment instrument that uses automated ambulatory pharmacy data to identify chronic conditions and predict future health care cost. The RxRisk model's performance in predicting cost is compared with a demographic-only model, the Ambulatory Clinical Groups (ACG), and Hierarchical Coexisting Conditions (HCC) ICD-9-CM diagnosis-based risk assessment instruments. Each model's power to forecast health care resource use is assessed.

DATA SOURCES. Health services utilization and cost data for approximately 1.5 million individuals enrolled in five mixed-model Health Maintenance Organizations (HMOs) from different regions in the United States.

STUDY DESIGN. Retrospective cohort study using automated managed care data.

SUBJECTS. All persons enrolled during 1995 and 1996 in Group Health Cooperative of Puget Sound, HealthPartners of Minnesota and the Colorado, Ohio and Northeast Regions of Kaiser-Permanente.

MEASURES. RxRisk, an algorithm that classifies prescription drug fills into chronic disease classes for adults and children.

RESULTS. HCCs produce the most accurate forecasts of total costs than either RxRisk or ACGs but RxRisk performs similarly to ACGs. Using the R^2 criteria HCCs explain 15.4% of the prospective variance in cost, whereas RxRisk explains 8.7% and ACGs explain 10.2%. However, for key segments of the cost distribution the differences in forecasting power among HCCs, RxRisk, and ACGs are less obvious, with all three models generating similar predictions for the middle 60% of the cost distribution.

CONCLUSIONS. HCCs produce more accurate forecasts of total cost, but the pharmacy-based RxRisk is an alternative risk assessment instrument to several diagnostic based models and depending on the nature of the application may be a more appropriate option for medical risk analysis.

Key words: Risk assessment; risk adjustment; pharmacy. (Med Care 2003;41:84–99)

*Center for Health Studies, Group Health Cooperative, Seattle, Washington.

†Health Partners Research Foundation, Minneapolis, Minnesota.

‡Center for Health Research, Kaiser-Permanente, Northwest Region, Portland, Oregon.

Supported by Merck Inc., the Ambulatory Pediatric Association, Cooperative Agreement No. 17-C-90864/9-01, "Reforming Medicare Risk Payment through Competitive Market Forces" from the Health Care Financing Administration (HCFA), by Grant No. 33039 from the Robert Wood Johnson Foundation (RWJF) and grant number R01HS11314 from the Agency for Healthcare Research and Quality. Additional support was provided by Kaiser Permanente.

The authors assume full responsibility for the accuracy and completeness of the information contained in this report.

Address correspondence and reprint request to: Paul A. Fishman, PhD, Center for Health Studies, Group Health Cooperative, 1730 Minor Avenue, Seattle, WA 98101. E-mail: fishman.p@ghc.org

Received December 3, 2001; initial decision January 22, 2002; accepted August 14, 2002.

A risk assessment is a forecast of a population's future health services cost or utilization based on its health status in a prior 'risk' period. Among other applications, risk assessments are used to adjust capitated health plan payments, as a case mix adjuster for clinical and health services research and as a tool for profiling health plans and providers.

Ideally, risk assessments would be based on a complete understanding of a population's health status and demographic profile. The prohibitive cost of administering surveys or clinically measuring health status for large populations has focused attention on risk assessment measures that use data available from computerized information systems. For example, the Hierarchical Coexisting Conditions (HCCs),[1] Ambulatory Clinical Groups (ACG),[2,3] Chronic and Disability Payment System (CDPS),[4,5] instruments are all based on International Classification of Disease, 9th Revision, Clinical Modification (ICD-9-CM),[6,7] codes available on automated information systems or from claims.

ICD-9-CM diagnosis-based risk assessment models are widely accepted as case mix adjusters and have been adopted by various payers as a means to adjust capitated payments in a variety of settings. Although not currently widely used, pharmacy-dispensing data available from automated sources are another source of information for risk assessment models. Several pharmacy-based risk models have been developed, most notably those by Clark et al,[8] Roblin,[9] Lamers,[10] Fishman and Shay,[11] and Gilmer et al[12] and these instruments may have several advantages over instruments based on ICD-9-CM diagnoses.

For populations that have a drug benefit as part of their insurance package pharmacy data are often more reliable, timely, and complete than diagnostic data. Some managed care organizations (MCOs), particularly group and staff-model health maintenance organizations (HMOs), do not routinely collect outpatient diagnostic data using standard ICD-9-CM codes, or do so in inconsistent and unreliable ways.[13] Most MCOs, however, maintain automated pharmacy data or can obtain this data from third parties, such as pharmacy benefit managers (PBMs).

A drug dispense requires that a provider write a prescription so risk data are linked to a specific clinical course of action. A dispense also records a specific drug, dose, and means of administration, which supports a more specific interpretation about the physician's clinical intent. Therefore, risk assessment models using routinely collected pharmacy data may be more appropriate than approaches relying on diagnostic data when chronic conditions are associated with clear pharmacotherapies and diagnostic data are either unavailable or incomplete.

There are several challenges to relying on ICD-9-CM diagnostic codes for assessing medical risk. Diagnostic data from automated databases may be imprecise, particularly if the chosen code is not used to assess payment or a clinical course of action, as may be the case on preprinted encounter forms commonly used in managed care environments that contain diagnoses codes from which providers choose. Automated databases may not capture fourth and fifth digits of ICD codes that are necessary to distinguish among diagnoses or indicate whether a diagnosis is intended as a "rule out" as opposed to a confirming diagnosis.[14]

A statistical concern with using diagnostic data for risk adjustment is that they link forecasts of resource use with current utilization: all diagnosis-based models require a visit to a health care provider and for a specific diagnosis to be made in order for relative risk characteristics to be captured. A pharmacy-based risk instrument may capture the health risk for persons with a stable, well-managed chronic disease even without a visit to a provider because many health plans do not require a visit with a provider to fill or refill a prescription. Therefore, a risk adjustment instrument based on pharmacy data are not necessarily linked to any specific type of utilization. Health plans may require annual reauthorizations for prescription drugs at a physician visit but there is no certainty that the ICD-9-CM code relevant for risk assessment purposes will be coded at that visit. A diagnosis based risk model will miss some well managed but expensive patients because providers do not feel compelled to record a recognized, underlying chronic medical problem. However, a pharmacy-based risk model will identify individuals with chronic conditions whose medical needs are managed with medications.[15–17]

Materials and Methods

Because of the increasing interest in pharmacy-based risk assessment, we revised and expanded the Chronic Disease Score (CDS),[8] a risk assessment instrument based on automated outpatient pharmacy data developed at Group Health Coop-

erative of Puget Sound (GHC). The revised and expanded model we report removes several barriers that prevented application of the CDS in a multisite risk adjustment context and these changes make the model universally relevant as a risk-adjustment tool. To distinguish the models we call the new instrument RxRisk. We briefly review the weaknesses in the CDS that motivated development of the RxRisk.

The CDS was developed exclusively for an adult population. Attempts to apply the model to children proved inadequate because children have a different set of chronic conditions than adults and may receive the same medical diagnosis but follow different pharmacological treatments for that condition. We developed a Pediatric Chronic Disease Score (PCDS),[11] which identifies chronic condition categories and drug assignments that reflect the unique challenges of risk model development using pharmacy data among pediatric populations but the adult and pediatric models had not been unified into a single instrument.

The CDS was developed using the GHC formulary as a basis for drug classification. Researchers and health plan decision makers had applied the model to other data sets but this required developing individual crosswalks. The CDS had not been intended as a capitated payment adjuster and included several categories that are inappropriate in a model used for finance purposes. Specifically, the "Pain" and "Pain and Inflammation" categories among adults included drugs that are prescribed less systematically than is appropriate for a finance model. The CDS was developed and estimated exclusively within the GHC staff model delivery system so risk weights may reflect practice pattern and drug use bias present in GHC, limiting the applicability of the model in a wider setting.

The CDS and the PCDS were validated relative to the ACG instrument. The drug-based approaches produced total cost forecasts that were similar or superior to the ACGs. The CDS explained 10% of the prospective variance in cost compared with 8% for ACGs,[8] although these results were not based on a split sample validation design. The PCDS explained 6% of the prospective variance in cost compared with 2% for ACGs for an independent validation sample.[11]

Development of RxRisk

RxRisk is an all-ages and market segment pharmacy-based risk assessment model that can be easily replicated in multiple health care settings. RxRisk produces estimates of future health care cost based on an individual's age, sex, and chronic condition profile measured by pharmacy dispenses linked to chronic conditions or clinically homogenous groups. An individual can be classified into any RxRisk category through a single dispense of a drug identified by the algorithm as linked to that conditions or clinical group. We considered and rejected requiring multiple dispenses before establishing a link between an individual and a condition or group because we cannot be certain using only automated data that the number of dispenses observed for an individual reflects a clinical choice or is the consequence of an insurance arrangement.

Development of the RxRisk algorithm used the GHC formulary and the classification algorithm used by the CDS and the Pediatric CDS as a departure point. First, we established a relationship between prescription drugs and chronic conditions and clinical groupings using the specific drugs found in GHC's formulary. To create the RxRisk we built tables that reported therapeutic classes and representative drugs used to identify each condition category. We then consulted physicians from our research centers to review the classification system to assess the face validity of the manner in which conditions are identified, the clinical validity of the way classes were mapped to each category and whether therapeutic uses have changed since the algorithms were last updated. Next, we received input from physicians at several clinics in Group Health Cooperative to get reactions about the risk categories and the manner in which drugs are assigned to these categories. These sessions included physicians practicing in both the closed and network portions of the Group Health delivery system. We then mapped drugs from the GHC formulary by ingredient and means of administration to drugs included on databases developed by Multum and FirstData and drugs on the formularies of the other health plans included in the study. FirstData and Multum are commercial software products, but GHC subscribes to FirstData for pharmacy administration purposes and Multum provides its databases through its Web site at no cost to the user.

TABLE 1. Description of R_xRisk Classes

RxRisk Class	Representative Drug Class(es)
Acne, pediatric	Anti-acne peroxides, anti-acne tretinoin, retinoids, topical macrolides
Allergic rhinitis, pediatric	Anti-inflammatory glucocorticoids
Amino acid disorders, pediatric	Amino acids
Anxiety and tension, adult	Salicylate combinations, barbiturates, benzodiazepines, meprobamate, miscellaneous hypnotics, paraldehyde
Anxiety and tension, pediatric	Anticholinergics, benzodiazepines
Asthma, adult	Anti-inflammatory glucocorticoids, isoproterenol, bronchodilators, cromolyn, xanthines
Asthma, pediatric	Anti-inflammatory glucocorticoids, bronchodilators, cromolyn, xanthines
Attention deficit disorder, pediatric	Anorexics/analeptics
Bipolar disorder, adult and pediatric	Lithium
Cardiac disease, adult	Class I a antiarrhythmic, Class I c antiarrhythmics. Class III antiarrhythmic, procainamide, disopyramide, quinidine, vasodilator nitrates, diuretic loops
Cardiac disease, pediatric	Beta adrenergic blockers, Class I a antiarrhythmic, Class I c antiarrhythmics, Class III antiarrhythmic, digitalis glycosides, dipyridamole, procainamide, vasodilator nitrates, calcium channel blockers, diuretic loop
Central line supplies, pediatric	Fibrinolysin antagonists, heparin
Congenital adrenal hypoplasia, pediatric	Anti-inflammatory glucocorticoids
Coronary/peripheral vascular disease, adult	Antiplatelet, oral anticoagulants, trental
Cystic fibrosis, adult	Anti-inflammatory glucocorticoids, enzymes
Cystic fibrosis, pediatric	Aminoglycosides, quinolones, antibiotic urinary tract anti-infective agents, mucolytics
Depression, adult	Monoamine oxidase inhibitors, phenothiazine combinations, tricyclic anti-depressants, SSRIs
Depression, pediatric	Monoamine oxidase inhibitors, tricyclic antidepressants,
Diabetes, adult	biguanides, insulins, sulfonylureas
Diabetes, pediatric	Insulins
Eczema, pediatric	Anti-inflammatory glucocorticoids, antipsoriasis combinations, topical steroids
Epilepsy, adult	Anti-convulsants
Epilepsy, pediatric	Anticonvulsant barbiturate + cogenerators, hydantoins
ESRD, adult	Marrow stimulants, human erythropoeitin
Gastric acid disorder, adult	Histamine H_2 blockers, prostaglandins, proton pump inhibitor
Gastric acid disorder, pediatric	Histamine H_2 blockers, proton pump inhibitor
Gout, adult	colchicine, uric acid inhibitors
Growth hormone deficiency, pediatric	Human growth hormone
Heart disease/hypertension, adult	Beta adrenergic blockers, dopamine, calcium channel blockers
Hemophilia, pediatric	Hemostatics
HIV, adult and pediatric	Miscellaneous anti-protozoal, antivirals, pentamidine
Hyperlipidemia, adult and pediatric	Antilipemic clofibrate, antilipidemic exchange resins, HMG coagulant reductase inhibitors
Hypertension, adult	Ace inhibitors, antihypertensive vasodilators, clonidine, ganglionic blockers, guanethidine, methyldopa, rauwolfia alkaloids, alpha/beta blockers, diuretic combinations, diuretic k+ depleting agents, diuretic k+ sparing agents
Immunodeficiency, pediatric	Immune serums
Iron overload, pediatric	Heavy metal antagonists
Irritable bowel syndrome, adult and pediatric	Sulfonamides

(*Continues*)

TABLE 1. (Continued)

RxRisk Class	Representative Drug Class(es)
Lead poisoning, pediatric	Heavy metal antagonists
Liver disease, adult and pediatric	Ammonia detoxicants
Malabsorbtion, pediatric	Dietary supplements
Malignancies, adult	Leucovorin, monoclonal, miscellaneous antinauseants, antineoplastic alkylating, antineoplastic antibiotics, antineoplastic mao inhibitors, antineoplastic progesterones, antineoplastic pyrimidines, antineoplastics misc, bladder protectant, methotrexate, purine antimetabolites, colony stimulating factors
Malignancies, pediatric	Leucovorin, miscellaneous antinauseants, anti-neoplastics: alkylating, antibiotics, MAO inhibitors, pyrimidines, methotrexate, purine anti-metabolites, colony stimulating factors
Ostomy, pediatric	Ostomy supplies
Pain and inflammation, pediatric	Antiinflammatory non steroidals
Pain, pediatric	Acetaminophen combinations, narcotic agonist/antagonists, narcotic analgesics, propoxyphene, salicylate combinations
Parkinsons disease, adult	Dopamine, MAO b inhibitors
Pituitary hormone, pediatric	Pituitary hormones
Psychotic illness, adult and pediatric	Miscellaneous antipsychotics, butyrophenones, phenothiazines, thiothixenes
Renal disease, adult	Potassium removing resins
Renal disease, pediatric	Marrow stimulants, Potassium removing resins, dietary supplements, human erythropoeitin
Rheumatoid arthritis, adult and pediatric	Antiinflammatory glucocorticoids, gold salts-injectable, gold salts-oral
Sickle cell anemia, pediatric	Penicillin derivatives, vitamin combinations
Steroid dependent diseases, pediatric	Anti-inflammatory glucocorticoids, thyroid replacement
Thyroid disorder, adult	Thyroid replacement
Tracheostomy, pediatric	Mechanical devices
Transplant, adult	Immunosuppressive agents
Transplant, pediatric	Monoclonalpurine anti-metabolites, immunosuppressive agents
Tuberculosis, adult	Anti-tuberculosis antibiotics, isoniazide

Table 1 provides a summary of the classification process and readers may obtain the entire classification algorithm from the lead author. For ease of presentation, the table identifies drugs using the GHC formulary therapeutic class designation, although the algorithm does not necessarily use all the drugs within a specific class to identify a chronic condition. Although most therapeutic classes from the GHC formulary contribute to only one RxRisk category, some therapeutic classes contribute drugs to multiple RxRisk categories, so the same class may appear in several conditions. However, any particular drug can only map to a single RxRisk category even if other drugs from the same therapeutic class map to other RxRisk categories. Some categories are unique to either adult or pediatric populations, whereas some conditions have both adult and pediatric components. In this latter category, some conditions rely on the same set of drugs to identify adults and children, whereas some have pediatric and adult specific drug identifiers. These categories are identified as either "Adult" or "Pediatric" specific or "Adult and Pediatric".

The "Pain" and "Pain and Inflammation" categories were excluded from the adult portion of the instrument because drugs in these categories are prescribed with too much discretion than would be appropriate in a payment adjustment model. These categories were retained in the pediatric instrument because of the greater likelihood that pain management drugs prescribed to children reflected a physician's, and often a parent's, concern about addressing a child's chronic pain.

Analysis Plan and Data Sources

Equation 1 summarizes the model for developing risk weights for the RxRisk:

$$\text{Medical Risk}_t = f(\text{Age, Sex, Benefit Status, RxRisk Categories})_{t-1} \quad (1)$$

We assume that medical risk during time period t is a function of each individual's age and sex, the source and extent of their health insurance (commercial, Medicare, or Medicaid) and the set of chronic conditions they are being treated for as measured by the RxRisk algorithm during a previous time period. We estimate Equation 1 using health services data for approximately 1.5 million individuals enrolled in several large HMOs: GHC, HealthPartners of Minnesota and the Northeast Ohio and Rocky Mountain regions of Kaiser Permanente. The Kaiser Permanente plans are primarily closed-group model HMOs, but several regions also have network components. HealthPartners provides services through both staff and network models with approximately two thirds of its enrollees receiving care through the network component. GHC is a mixed model HMO with 85% of its enrollees receiving care in a closed panel group model and the rest through a network. Subjects were all individuals enrolled for 1 month during 1995 and at least 1 day in 1996 through a medical plan that included a pharmacy benefit as part of their benefit package. The pharmacy benefit requirement is a proxy for a comprehensiveness health benefit package, which increases the chance that we are capturing the full range of health services subjects use during the year. We also excluded Medicare eligible enrollees receiving care through a cost contract.

We estimate Equation 1 using total costs as a proxy for medical risk. Therefore, total health care costs are the dependent variable in an single-equation weighted least squares (WLS) regression given by Equation 2:

$$\text{Risk}_{i,t} = X_{i,t-1}\beta_i + U_i \ldots\ldots \quad (2)$$

where: $\text{Risk}_{i,t}$ are health care costs for the ith person in year t, $= X_{i,t-1}$ are the demographic and RxRisk characteristics for the ith person in year $t-1$, β_i are the weights associated with each of the demographic and RxRisk characteristics and U_i is a disturbance term.

Several researchers have proposed methods other than single-equation linear regression to estimate health services utilization and cost because the distributional properties of these data may require transformation.[18–21] Our previous research has demonstrated that ordinary (OLS) or weighted (WLS) least squares perform better in developing forecasting models than other functional forms.[22] Regression using OLS or WLS also allows easy calculation of a risk profile in a meaningful metric (ie, dollars) for each subject through summing coefficients for each RxRisk variable. Regression weights were a function of individual eligibility in the risk period. Individual observations are weighted by the length of enrollment in the risk period. Forecast year costs are annualized for persons enrolled for less than 12 months.

We apply the RxRisk algorithm to prescription drug fills from 1995 to predict medical risk (annual per capita expense) in 1996. Costs are truncated at $400,000 because of data-quality concerns with outlier cases, but this decision affected less than one tenth of 1% of subjects. Costs for individuals enrolled less than 12 months in the risk years were then eligibility adjusted based on months enrolled during the risk year. We annualized costs for persons enrolled for less than 12 months during the forecast year.

Risk assessment models do not reflect predicted cost associated with babies born in the forecast year because they do not have any pharmacy (or diagnostic) experience in the risk year. As a result, forecast models often ignore a set of enrollees that are potentially among the more expensive in the health plan. We attempted to correct this oversight by incorporating newborns' costs with those of their mother. We made this choice for two reasons. First, there may be a link between a newborn's costs and the risk experience of his or her mother and the child's cost may be reflected in the mother's risk profile. Second, health plans often report the costs of providing care for mothers and newborns, at least until the child is provided their own enrollment number and it is difficult to distinguish the costs of the mother and child in some cases.

Validating forecasting results on an independent sample is an important step in risk model selection because a model's predictive power might be overstated if it is evaluated using the data on which it is estimated.[23] For example, an empirical model may fit the data on which it was

TABLE 2. Description of Subjects Included in Estimation and Validation Samples

	Estimation Sample			Validation Sample		
	Children	Adults	All Subjects	Children	Adults	All Subjects
N	415,458	1,015,164	1,430,622	30,596	75,649	106,245
Percent Male	51.1%	45.8%	47.4%	51.2%	45.9	47.5
Mean (SD) Age as of December 31, 1995	8.7 (5.1)	42.5 (14.8)	32.7 (19.6)	8.78 (5.2)	42.4 (14.8)	32.6 (19.6)
Mean (SD) Total Costs—1996	676.71 (3268)	1,824 (6245)	1,491.47 (5572)	699.35 (3318)	1,838.11 (6213)	1510.17 (5561)
Mean (SD) Pharmacy Costs—1996	57.06 (374)	241.03 (755)	187.61 (672)	61.84 (581)	238.48 (941)	187.61 (857)

estimated well, but subsequently perform poorly with other data, including that from the same population. We assessed model over fitting by withholding a 7% random validation sample 106,245 and estimating the model on the remaining 93% of subjects (1.43 million). The forecasting power of this model is based on its predictive performance on the 106,245-person validation sample using the following six measures:

1. **R^2.**[23] R^2 measures the percent of individual variance explained by the model and indicates how much of the variance in the dependent variable is explained by the model.
2. **Mean Prediction/Mean Prediction Error.** The mean prediction error (MPE) assesses how well on average the model produces a forecast of this mean.[23] The MPE is calculated as: (sum(predicted—actual costs)/N), where N is the number of individuals in the validation sample. The strength of a prediction is gauged by how close the MPE is to zero. The MPE summarizes how well on average a model predicts mean total costs for a defined population.
3. **Mean Absolute Prediction Error.** The mean absolute prediction error (MAPE)[23] is the absolute value of the difference between actual and predicted costs or utilization: $(\Sigma_i$ |predicted—actual costs/)N $|_i$. Using the absolute value means that predictions that are greater or less than actual costs cannot cancel each other out, as can happen with the MPE.
4. **Prediction Ratio.**[24] The Prediction Ratio is the mean quotient of the predicted to the actual cost values for each subject and complements the mean and absolute prediction error in assessing the accuracy of a prediction. A risk instrument's forecasting power is evaluated by the degree its prediction ratio deviates from 1. Because some subjects have no costs in the forecast year, we add a value of 1 to both predicted and actual costs for all subjects to avoid undefined ratios.
5. **Fit of Actual to Predicted Costs (Mincer-Zarnowitz Test).**[24] A perfect forecasting model would reveal exact correspondence between actual and predicted costs. However, the actual degree of this correspondence can be determined with an OLS regression using predicted costs against actual costs for the validation sample. A regression line for an unbiased forecast would produce an intercept that is not significantly different from 0 and a slope coefficient that is not significantly different from 1. A model's relative performance can be gauged by examining the deviation of the intercept and slope estimates for cost distribution ranges and overall. The regression equation used to conduct the Mincer-Zarnowitz test was used to derive R^2 for the validation sample, which is the proportion of variance in observed annual per capita expense explained by the forecasting model.
6. **Cost Quintile Analysis.**[24] Risk instruments may generate more accurate predictions for different ranges of the cost distribution. We assess the relative strength of the RxRisk model to predict costs for relatively high, medium, and low cost subjects by examining the predictive performance of the model by cost quintile. Subjects are grouped into five equally populated segments based on their actual costs and we calculate Predicted Expense, Mean

TABLE 3. RxRisk Model

Variable Label	Parameter Estimate	Standard Error	P value
Demographic Factors			
Intercept (Females 4 ≤ Age ≤ 9)	409.75	22.06	0.0001
Medicare: Age < 65	3999.86	113.43	0.0001
Medicare: Age = 65 and older	303.96	50.29	0.0001
Medicaid: Age < 65	417.12	30.03	0.0001
Medicaid: Age = 65 and older	3325.88	133.67	0.0001
Females, Age = 0	1091.00	63.88	0.0001
Females, 1 ≤ Age ≤ 3	171.07	39.04	0.0001
Females, 10 ≤ Age ≤ 17	197.94	28.92	0.0001
Females, 18 ≤ Age ≤ 24	847.83	32.82	0.0001
Females, 25 ≤ Age ≤ 34	1432.35	27.78	0.0001
Females, 35 ≤ Age ≤ 44	658.10	26.796	0.0001
Females, 45 ≤ Age ≤ 54	653.44	28.20	0.0001
Females, 55 ≤ Age ≤ 64	895.94	32.33	0.0001
Females, 65 ≤ Age ≤ 74	1118.73	56.69	0.0001
Females, Age ≥ 75	1827.76	66.97	0.0001
Males, Age = 0	1385.94	62.09	0.0001
Males, 1 ≤ Age ≤ 3	293.45	38.56	0.0001
Males, 4 ≤ Age ≤ 9	37.04	30.49	0.2245
Males, 10 ≤ Age ≤ 17	87.89	28.74	0.0022
Males, 18 ≤ Age ≤ 24	70.69	34.59	0.041
Males, 25 ≤ Age ≤ 34	76.99	29.25	0.0085
Males, 35 ≤ Age ≤ 44	201.31	27.40	0.0001
Males, 45 ≤ Age ≤ 54	520.07	28.39	0.0001
Males, 55 ≤ Age ≤ 64	1205.25	32.92	0.0001
Males, 65 ≤ Age ≤ 74	1810.56	56.04	0.0001
Males, Age ≥ 75	2842.19	72.67	0.0001
Adult RxRisk Factors			
Anxiety and tension	922.37	23.81	0.0001
Asthma	774.85	24.55	0.0001
Bipolar disorder	857.81	103.74	0.0001
Cardiac disease	2574.47	35.96	0.0001
Coronary/peripheral vascular disease	2662.23	56.51	0.0001
Cystic fibrosis	5100.00	209.49	0.0001
Depression	1188.59	21.98	0.0001
Diabetes	2229.10	36.52	0.0001
Epilepsy	1774.11	61.82	0.0001
ESRD	30941.00	428.78	0.0001
Gastric acid disorder	1413.26	27.08	0.0001
Glaucoma	805.77	60.56	0.0001
Gout	1147.71	65.90	0.0001
Heart disease/hypertension	711.49	24.47	0.0001
HIV	12149.00	267.55	0.0001
Hyperlipidemia	611.47	38.36	0.0001
Hypertension	622.14	22.19	0.0001

(*Continues*)

TABLE 3. (Continued)

Variable Label	Parameter Estimate	Standard Error	P value
Irritable bowel syndrome	1159.11	107.39	0.0001
Liver failure	9365.55	267.84	0.0001
Malignancies	4088.35	64.22	0.0001
Parkinsons disease	2658.37	353.02	0.0001
Psychotic illness	2420.26	46.47	0.0001
Renal disease	25372.00	533.97	0.0001
Rheumatoid arthritis	1395.73	39.37	0.0001
Thyroid disorder	359.49	31.15	0.0001
Transplant	10364.00	258.16	0.0001
Tuberculosis	2610.97	201.77	0.0001
Pediatric RxRisk Factors			
Acne	240.37	55.76	0.0001
Allergic rhinitis	319.55	56.29	0.0001
Amino acid disorders	17413.00	1631.12	0.0001
Anxiety and tension	1769.33	175.71	0.0001
Asthma	600.66	58.14	0.0001
Attention deficit disorder	471.77	56.96	0.0001
Bipolar disorder	1151.04	337.28	0.0006
Cardiac disease	2914.48	209.35	0.0001
Central line supplies	11808.00	1035.07	0.0001
Congenital adrenal hypoplasia	2246.44	770.21	0.0035
Cystic fibrosis	5054.63	265.34	0.0001
Depression	1123.95	90.02	0.0001
Diabetes	2196.17	222.46	0.0001
Eczema	167.79	47.39	0.0004
Epilepsy	2411.72	160.73	0.0001
Gastric acid disorder	2339.19	143.67	0.0001
Growth hormone deficiency	11979.00	588.42	0.0001
HIV	12149.00	267.55	0.0001
Irritable bowel syndrome	1159.11	107.39	0.0001
Immunodeficiency	3669.82	1397.31	0.0086
Liver disease	2441.04	520.18	0.0001
Malignancies	14172.00	415.74	0.0001
Pain	1012.57	148.63	0.0001
Pain and inflammation	363.02	68.67	0.0001
Pituitary hormone	740.13	284.23	0.0092
Psychotic illness	744.89	371.19	0.0448
Renal disease	28369.00	941.16	0.0001
Rheumatoid arthritis	627.31	443.53	0.1573
Sickle cell anemia	618.89	194.92	0.0015
Steroid dependent diseases	1052.09	127.41	0.0001
Thyroid disorder	1385.78	264.60	0.0001
Transplant	9110.11	834.54	0.0001

TABLE 4. Model Performance—Entire Validation Sample

	Estimation Sample	Validation Sample
N	1,430,622	106,245
Mean Total Cost	$ 1,491	$ 1,510

Risk Measure	Percent of Subjects Assigned a Risk Factor in Addition to Age and Sex	Estimation R^2	Mean Prediction	Mean Prediction Error	Mean Absolute Prediction Error	Prediction Ratio	Mincer-Zarnowitz Constant (St. Error)	Mincer-Zarnowitz Coefficient (St. Error)	Validation R^2
Age/Sex	0%	3.84%	1438.81	−71.19	1646.97	152.2	216.69 (26.66)*	0.89 (0.0144)§	3.5%
RxRisk	28%	9.39%	1438.15	−71.85	1530.36	109.3	155.78 (21.14)*	0.94 (0.009)§	8.74%
ACGs	85%	9.24%	1450.69	−59.31	1504.38	102.3	57.18 (20.91)*	1.001 (0.009)§	10.15%
HCCs	75%	13.57%	1446.84	−63.16	1449.85	90.3	1.57 (19.07)*	1.04 (0.0074)§	15.42%

*$P < 0.01$, §$P < 0.05$

Prediction Error, Absolute Prediction Error and the Prediction Ratio for each quintile.

To assess the relative forecasting power of the RxRisk model, we compare its performance with an age and sex only model (henceforth referred to as the Demographic instrument), and to the Hierarchical Co-existing Conditions (HCCs)[1] and Ambulatory Clinical Groups.[2,3] (ACGs)

TABLE 5. Model Performance—Cost Quintile Analysis

Quintile	Model	Mean Prediction Mean (SD)	Mean Prediction Error Mean (SD)	Absolute Prediction Error Mean (SD)	Prediction Ratio Mean (SD)
1st	Actual	7.21 (15.66)			
	Age/Sex	963.09 (764.14)	955.88 (763.85)	955.88 (763.84)	746.86 (774.09)
	RxRisk	701.69 (586.04)	694.48 (585.23)	694.49 (585.22)	534.50 (569.38)
	ACG	676.25 (623.83)	669.04 (622.02)	669.04 (622.01)	498.85 (552.62)
	HCC	598.50 (706.78)	591.30 (705.95)	591.33 (705.57)	443.31 (622.51)
2nd	Actual	152.17 (57.33)			
	Age/Sex	1,073.77 (790.05)	921.61 (789.51)	923.18 (787.67)	8.17 (7.34)
	RxRisk	856.80 (704.29)	704.63 (702.59)	707.11 (700.14)	6.46 (6.57)
	ACG	906.16 (819.53)	753.99 (818.03)	756.35 (815.85)	6.83 (7.09)
	HCC	795.85 (855.91)	643.68 (854.24)	651.58 (848.23)	5.97 (7.18)
3rd	Actual	403.02 (93.50)			
	Age/Sex	1,311.22 (931.68)	908.20 (927.69)	922.54 (913.43)	3.38 (2.51)
	RxRisk	1,114.45 (917.74)	711.43 (912.70)	735.90 (893.10)	2.86 (2.41)
	ACG	1,133.47 (1,023.25)	730.46 (1,019.39)	748.80 (1,005.99)	2.91 (2.64)
	HCC	1,075.81 (1,054.25)	672.79 (1,049.23)	738.47 (1,004.08)	2.75 (2.76)
4th	Actual	914.32 (233.74)			
	Age/Sex	1,654.02 (1,134.05)	739.70 (1,134.63)	924.46 (989.91)	1.89 (1.36)
	RxRisk	1,613.79 (1,341.41)	699.48 (1,325.46)	937.83 (1,169.00)	1.82 (1.51)
	ACG	1,590.78 (1,492.66)	676.47 (1,481.86)	908.10 (1,325.35)	1.80 (1.71)
	HCC	1,572.03 (1,460.52)	657.72 (1,447.68)	928.73 (1,290.66)	1.78 (1.67)
5th	Actual	6,074.21 (11,316.44)			
	Age/Sex	2,192.88 (1,552.38)	−3,881.33 (11,031.30)	4,508.76 (11,065.93)	0.72 (.68)
	RxRisk	2,904.93 (2,932.91)	−3,169.27 (11,214.02)	4,576.48 (10,716.99)	0.90 (.94)
	ACG	2,947.70 (2,839.10)	−3,126.50 (11,085.47)	4,439.60 (10,627.89)	0.88 (.94)
	HCC	3,192.94 (3,620.77)	−2,881.27 (10,793.02)	4,339.14 (10,293.80)	0.91 (1.00)

TABLE 6. Validation Results—Sub Populations

	Men: N = 50,520			
	Mean Prediction Error	Absolute Prediction Error	Prediction Ratio	Validation R^2
Risk Measure				
Demographic	−55.64	1467.55	175.56	3.7%
RxRisk	−56.93	1357.13	122.77	9.3%
ACGs	−12.80	1390.28	131.42	9.7%
HCCs	−51.31	1316.68	107.31	15.0%
	Women: N = 55,717			
	Mean Prediction Error	Absolute Prediction Error	Prediction Ratio	Validation R^2
Risk Measure				
Demographic	−85.29	1809.68	131.04	3.1%
RxRisk	−85.39	1687.46	97.10	8.0%
ACGs	−101.49	1607.85	75.81	10.3%
HCCs	−73.91	1570.61	76.11	15.5%
	Children (aged 17 and younger): N = 30,596			
	Mean Prediction Error	Absolute Prediction Error	Prediction Ratio	Validation R^2
Risk Measure				
Demographic	−36.79	707.85	85.67	0.6%
RxRisk	−35.37	675.06	71.86	4.6%
ACGs	−152.27	777.17	90.16	7.3%
HCCs	−29.58	662.36	50.98	11.1%
	Adults (aged 18 and older): N = 75,649			
	Mean Prediction Error	Absolute Prediction Error	Prediction Ratio	Validation R^2
Risk Measure				
Demographic	−85.09	2026.79	179.13	2.9%
RxRisk	−86.61	1876.28	124.45	8.3%
ACGs	−144.88	1798.49	107.15	9.7%
HCCs	−76.74	1768.35	107.11	15.1%
	Medicare: N = 5726			
	Mean Prediction Error	Absolute Prediction Error	Prediction Ratio	Validation R^2
Risk Measure				
Demographic	−119.84	5282.58	116.83	0.2%
RxRisk	−110.00	4872.69	67.52	7.7%
ACGs	−91.99	4911.14	68.03	7.4%
HCCs	−88.03	4683.98	65.15	15.9%
	Medicaid: N = 3579			
	Mean Prediction Error	Absolute Prediction Error	Prediction Ratio	Validation R^2
Risk Measure				
Demographic	−60.13	1620.83	160.36	11.4%
RxRisk	−83.89	1556.90	140.43	14.8%
ACGs	−58.91	1520.49	125.72	16.5%
HCCs	−76.03	1448.76	109.14	26.3%

HCCs assigns ICD-9-CM codes to "DxGroups" that are clinically related and similar with respect to levels of resource use. DxGroups are aggregated into 136 Coexisting Conditions (CCs) that include DxGroups belonging to a major body system or disease type, grouped by cost and clinical relation. To avoid double counting within related CCs, the HCC algorithm then imposes hierarchies on the CCs based on disease severity, choosing only the highest ranked CC among sets of related conditions. Adjusted Clinical Groups (ACGs) are 51 mutually exclusive health status categories defined by morbidity, age, and gender. The ACG algorithm assigns all ICD-9-CM codes to one of 32 diagnosis groups, known as Ambulatory Diagnosis Groups (ADGs). Diseases are placed in an ADG based on five clinical dimensions: duration (acute, recurrent or chronic), severity (minor/stable vs. major/unstable), diagnostic certainty (symptoms vs. diseases), etiology (infectious, injury or other), and likely use of specialty care (medical, surgical, obstetric, hematology, etc.). Both HCCs and ACGs are currently used to risk adjust health plan payments from a variety of public and private purchasers. The Demographic model resembles the actuarial approach commonly used to adjust for differences among pediatric and adult populations and we include this approach as a reference point to allow for easy comparison with other risk assessment research. The empirical weights for the Demographic, ACG, and HCC models are based on the same WLS regression described in Equation 2 using the same population on which the RxRisk was estimated.

Results

Descriptive data on the multi-HMO estimation sample are provided in Table 2. Regression results used to estimate RxRisk weights are provided in Table 3. Validation tests performed for the 106,245 person validation sample are provided in Table 4. Validation results by Cost Quintile are provided in Table 5 and separate validation results for men and women, children and adults, and Medicare and Medicaid eligible subjects are provided in Table 6.

HCCs produce the most accurate prediction of total future health care costs for the entire validation sample but this relative advantage over ACGs and RxRisk differs across the validation criteria. HCCs explain 13.5% of the prospective variance in cost whereas ACGs and RxRisk have an R^2 of 9.4% and 9.2% respectively for the estimation sample. The Demographic instrument explains relatively little of the prospective variance in cost with an R^2 of 3.9%. R^2 values for the validation sample show a similar pattern: HCCs explain 15.4% of the prospective individual variation in cost among subjects in the validation sample, compared with 10.1% for ACGs, 8.7% for RxRisk, and 3.5% for the Demographic model.

These R^2 values must be placed in context, because although they might seem low they are consistent with results obtained in other risk based forecasting exercises. Risk based forecasts produce relatively low R^2 because a great deal of health care costs occur randomly and cannot be predicted, on an individual basis, with great confidence. Newhouse et al[29] argues that the maximum explainable R^2 for risk assessment models may be approximately 30%. If this is the case, the various models we test explain about one third to one half of the explainable prospective variance in cost.

The four instruments perform similarly in predicting mean cost for the validation sample with $6.37 in annual predicted cost separating HCCs from Demographics alone. The range for the Mean Prediction Errors is also small: $8.71 in annual costs separates HCCs and Demographics alone. Based on the Mean Absolute Prediction Error criteria, the HCCs produce a 3.7% more accurate prediction than ACGs, 5.4% more accurate than RxRisk, and 12% more accurate than Demographics.

Each instrument fails the Mincer-Zarnowitz test. Intercepts for OLS regressions of predicted on actual costs for all four models are significantly different from 0 and slope coefficients are different from 1. This result is primarily caused by the large number of observations in the validation sample, but the values of the intercept and slope coefficients still reveal important information about the performance of each model. The demographic and RxRisk models have larger intercepts than ACGs or HCCs and slope coefficients less than one so these models will tend to over-predict low cost and under-predict high cost enrollees. This result is driven by the smaller set of subjects whose risk is established through means other than age and sex. HCCs and ACGs both have relatively smaller intercepts and slope coefficients greater than one and are therefore more likely to under-predict low cost subjects.

HCCs are also the most accurate risk instrument when we evaluate validation results for the quintile

analysis, but all four instruments do a poor job predicting costs for the least and most expensive subjects. HCCs outperform ACGs and RxRisk at each quintile but the differences among these instruments are small for the middle 60% of the cost distribution. The greatest difference among HCCs, ACGs, and RxRisk subjects in the middle three cost quintiles occurs in the 2nd cost quintile where HCCs are 14% more accurate than ACGs and 8% more accurate than RxRisk for the Mean Prediction Error criterion.

HCCs have the highest R^2 and most accurate Prediction Ratio in validation tests conducted separately on men (n = 50,520) and women (n = 55,717). ACGs and RxRisk perform similarly for men and women on the R^2 and Prediction Ratio criteria for men but ACGs produce a more accurate Prediction Ratio for women. ACGs produce the most accurate mean predictions for men (MPE = −$12.80), but this advantage does not carry over to women where ACGs produce the least accurate prediction (MPE = −$101.49). Demographics alone generate Mean Prediction Errors for men and women similar to the other approaches but age and sex alone perform significantly worse on the other validation criteria.

HCCs have the highest R^2 in validation tests conducted separately on adults (n = 75,649) and children (n = 30,596) but HCCs and ACGs have similar Prediction Ratios for both age groups. RxRisk generates similar Mean and Absolute Prediction Errors as do HCCs for both children and adults. Based on the Mean Prediction Error criterion, RxRisk is approximately $6 per year per person less accurate in predicting cost among children than HCCs and approximately $10 less accurate among adults. Based on the Absolute Prediction Error criterion, RxRisk is approximately $10 per year per person less accurate in predicting cost among children than HCCs and approximately $100 less accurate among adults.

Although our validation sample has a small number of Medicare (n = 5726) and Medicaid (n = 3579) eligible subjects, we report separate validation tests for these market segments. Demographics explain little of the variance in prospective cost for the Medicare population (0.2%) whereas ACGs and RxRisk produce similar R^2 values (R^2 = 7.4 and 7.7% respectively). HCCs do considerably better on this criterion with an R^2 of 15.9%. HCCs, ACGs, and RxRisk perform similarly on the Prediction Ratio criterion and the Absolute Prediction Error, but the HCCs and ACGs perform better using the Mean Prediction Error Criteria.

A different pattern holds for the Medicaid population. Although HCCs, ACGs, and RxRisk continue to generate more accurate predictions, Demographics produce more accurate cost prediction than was the case for the general population. This may be caused by the Medicaid population being dominated by children and younger women for whom chronic conditions may not as significant as the general population.

Discussion

We reported a risk-assessment instrument based on automated pharmacy and demographic data. We evaluated the performance of the RxRisk model against an age and sex only model, and the HCC and ACG diagnosis risk instruments to assess the ability of RxRisk to forecast future health care costs. No single test can determine the predictive power of a risk assessment model so we conducted several tests to gauge the performance of RxRisk relative to these other instruments. On the basis of these analyses we conclude that HCCs produce a more accurate prediction of future total health care costs relative to RxRisk and ACGs. This advantage, however, is not as clear when the instrument's performance is examined for different segments of the population. We also conclude that RxRisk is comparable to ACGs in predicting total future health care costs and, on the basis of statistical performance, we find little to distinguish the ability of these two instruments to forecast total health care costs.

There are several reasons why HCCs out-perform RxRisk. First, HCCs use more information to create risk categories than RxRisk. We report in Table 4 that 75% of subjects class into an HCC category but only 28% of subjects class into an RxRisk category. HCCs are, therefore, better able to discriminate risk along the risk spectrum.

Although HCCs outperform RxRisk throughout the risk spectrum, this relative advantage is most evident among people in the lowest and highest cost quintiles. Using the Mean Prediction Error criteria as an example, the greatest difference between HCCs and RxRisk is in the extreme cost quintiles. This is evidence of the HCC's greater ability to distinguish risk among persons with relatively low and high future medical care expense.

The relative strength of HCC's ability to predict cost is also demonstrated among persons with risk factors that cannot be identified with pharmacy data. Pharmacy- and diagnosis-based risk models each

have relative strengths and weakness with respect to their ability to identify medical need associated with specific conditions. Drug models will do a better job identifying risk for some medical conditions but diagnoses will be consistently better with others and there are some cases where the absence of a drug marker can be critical.[12] Pregnancy is a good example of this because it is often the most expensive medical event among commercially insured women of child-bearing age. However, we did not feel confident assigning a drug marker for pregnancy. As a result, RxRisk age and sex risk weights for women of child-bearing age are larger and have a greater variance than similar age and sex weights for HCCs. As we report in Table 2, the risk weight for women aged 25 to 34 in RxRisk is $1432 but in the course of estimating an HCC model for our sample we found that these women had a risk weight of $878. The variance inflation factor, which measures the relative contribution of a parameter estimate to the overall explanatory power of the, is 2.5 for the RxRisk weight is 2.5 but 1.6 for HCCs. This difference is the largest between the two models for any age, sex risk weight.

A statistical analysis is only the first step in determining whether a risk instrument is appropriate to use in specific situations and the analyses we present cannot determine whether the diagnosis-based HCCs or ACGs or the pharmacy-based RxRisk instruments are appropriate in specific circumstances. HCCs out-perform ACGs and RxRisk on the validation tests we conducted but other criteria should be a part of the decision about which risk instrument to use and one of the key factors is the confidence users have in the availability and quality of the data source on which a particular instrument relies.

Users may choose a drug-based instrument if pharmacy data are of greater quality and more appropriate to the task than diagnosis data. Pharmacy data are either more readily available and of higher quality in some managed care environments because capitation has effected data collection efforts. In capitated environments providers have less incentive to report diagnostic and procedural data and health plans are less likely to collect these data when it is not necessary for reimbursement. At a minimum, the quality and completeness of diagnostic data may be an issue and we have demonstrated that the RxRisk model is an alternative risk assessment tool that health plans and payers may find valuable.

Pharmacy-based risk assessment instruments may be less subject to gaming because drug dispenses are linked to a specific clinical course of action and require that a provider write a prescription that a patient fills. No risk assessment model can eliminate gaming, but the goal of every risk model is to increase the cost of gaming or decrease the cost of monitoring provider and health plan behavior to make it easier to detect any possible gaming that might occur. If the RxRisk algorithm is correctly specified, then an individual is linked through a pharmacy dispense to a clinical course of action. Some subjects may be incorrectly assumed to have a chronic condition because the RxRisk algorithm has assigned them to that category based on a drug dispense but these misclassifications are inevitable in any risk assessment instrument. The pharmacy-based approach avoids the likelihood of systematic manipulation of codes because gaming the system would require that patients fill prescriptions that are not medically indicated. Regardless of the likelihood that it would occur, the consequences of gaming a pharmacy based risk model are much greater than is the case with diagnosis based risk assessment. Gaming a diagnosis based risk model involves little or no impact on the patient or the plan. However, gaming a pharmacy-based risk model by dispensing drugs that are not necessarily medically indicated involves costs to both the health plan and consumer (if copayments or a deductible apply to the dispense) and potential health risks to the patient.

There are several limitations to applying the RxRisk model. First, a chronic condition must have a specific pharmacological treatment for pharmacy data to serve as an illness-identifying vehicle. This limitation is a particular concern among children because many costly pediatric conditions do not have specific or clearly defined drug regimens.[11,25–28]

A second problem is that health plans and pharmacy benefit mangers use different methods to classify drugs. Most information systems classify drugs by therapeutic class but there is no gold standard for this type of classification. Rather than create plan specific crosswalks for RxRisk we used NDC codes because every approved drug in the US is assigned an NDC. Relying on NDCs still presents some challenges in applying the RxRisk setting because there are still health plan specific idiosyncrasies. Some health plans buy drugs in large quantities and these drugs are often repackaged before distribution to consumers. Health plans may also provide drugs in compounds not otherwise commercially available. In both of these cases there are no standardized NDC codes and the health plan will likely use 'home-grown' codes to identify the fill.

Furthermore, patients on clinical protocols may receive experimental drugs, or may not have their drug use appear on typical MCO automated information systems. The RxRisk algorithm will not classify these dispenses correctly without modifications to incorporate these homegrown codes.

Any pharmacy-based risk instrument relies on more frequently changing data than diagnoses-based approaches and will require more frequent updates. A drug model must deal with new uses of approved drugs, discontinued uses of older medications, and providers reacting to new clinical evidence that could change their prescribing patterns. Other pharmacy instruments avoid this problem by basing algorithms on a therapeutic class designation so new drugs in that class are automatically captured by the model, but we do not feel this is an adequate solution to the updating problem for two reasons. First, there is no single therapeutic class mechanism, and many formularies use idiosyncratic methods to organize approved medications. A second problem is that the RxRisk algorithm does not use all drugs in a therapeutic class to identify medications linked to chronic conditions. Changing pharmacy databases are not an intractable problem because we anticipate that RxRisk will use retrospective data and users can make needed changes to the drug classification process specific to their data before applying the model.

Perhaps the greatest challenge to pharmacy-based risk adjustment is the extent of drug benefit coverage among the target population. This is particularly relevant among seniors receiving medical care through the standard Medicare program that does not include an outpatient drug benefit. This may not be as great a problem in integrated health systems that provide full service health centers with pharmacies, because they collect drug dispensing data regardless of coverage if consumes fill their prescriptions in plan run pharmacies. This will not be the case for the fee-for-service Medicare program, which may be the most obvious place to institute a pharmacy-based risk adjuster because of the high prevalence of chronic conditions treated pharmacologically among seniors. Future research needs to determine whether RxRisk is biased when measuring risk in populations with less generous drug benefit coverage.

The key to risk assessment is data quality. Pharmacy-based risk models have the advantage of using data that is often more reliable and complete than the diagnostic data available in many managed-care organizations. As clinical information systems evolve, diagnostic data available from automated sources will improve and risk adjustment itself will create incentives for health plans to collect diagnostic data in a more timely and accurate way. However, pharmacy data will continue to be a reliable and timely source of risk data because of it provides a direct link to a clinical decision made by a health care provider.

Acknowledgments

The authors thank Dr. David McCulloch, MD, Dr. Paula Lozano, MD MPH, Cassandra Luce, Ella Thompson, and Joan Fanning for their assistance.

References

1. **Ash A, Powell F, Gruenberg L, et al.** Adjusting Medicare payments using prior hospitalization data. Health Care Financing Review 1989;10:17.

2. **Starfield B, Weiner J, Mumford L, et al.** Ambulatory Care Groups: A. categorization of diagnoses for research and management. Health Serv Res 1991;26:53–74.

3. **Weiner JP, Starfield BH, Steinwachs DM, et al.** Development and application of a population-oriented measure of ambulatory case-mix. Med Care 1991;29:452–472.

4. **Kronick R, Dreyfus T, Lee L, Zhou Z.** Diagnostic risk adjustment for Medicaid: the disability payment system. Health Care Finan Rev 1996;17:7–33.

5. **Kronick R, Gilmer T, Dreyfus T, Lee L.** Improving health-based payment for Medicaid beneficiaries: The Chronic and Disability Payment System. Health Care Finan Rev 2000;21:29–64.

6. **Hornbrook MC, Goodman MJ, Fishman PA, et al.** Global risk-assessment models: demographic and diagnostic approaches. Presented at the Annual Meeting of the Association for Health Serv Res. Atlanta, GA, June 1996.

7. **American Medical Association, ICD-9-CM.** International Classification of Diseases, 9th Revision, Volumes 1 and 2. Chicago, IL: American Medical Association; 2002.

8. **Clark DO, Von Korff M, Saunders K, et al.** A chronic disease score with empirically derived weights. Med Care 1995;33:783–795.

9. **Lamers LM.** Pharmacy costs groups: a risk-adjuster for capitation payments based on the use of prescribed drugs. Med Care 1999;37:824–830.

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

11. **Fishman P, Shay D.** Development and estimation of a pediatric chronic disease score from automated pharmacy data. Med Care 1999;37:872–880.

12. **Gilmer T, Kronick R, Fishman P, et al.** The Medicaid RX model: Pharmacy-based risk adjustment for public programs. Med Care 2001;39:1188–1202.

13. **Fishman PA, Wagner EH.** Managed care data and public health: The experience of Group Health Cooperative of Puget Sound Annual Review of Public Health 1998;19:477–491

14. **Goodman M, Bachman D, Fishman P, et al.** Variations in Diagnosis Coding Practices Across Multiple HMOs. Paper presented at the 16th Annual Association for Health Serv Res Annual Meetings Chicago, IL, June 1999:26.

15. **Johnson RE, Hornbrook MC, Nichols GA.** Replicating the Chronic Disease Score (CDS) from automated pharmacy data. J Clin Epidemiol 1994;47:1191–1199.

16. **AHFS 2002.** Drug Information. American Hospital Formulary Service, Bethesda, MD, 2001.

17. **Hornbrook MC, Goodman MJ.** Health Plan Case Mix: Definition, Measurement, And Use. In: Hornbrook MC, editor. Advances in Health Economics and Health Serv Res, Vol. 10 Greenwich, CT: JAI Press, 1989:41.

18. **Duan N, Manning W, Morris C, Newhouse J.** A comparison of alternative models for the demand for medical care. J Bus Econ Stat 1983;1:115.

19. **Duan N, Manning W, Morris C, Newhouse J.** Choosing between the sample selection model and the multi-part model. J Bus Econ Stat 1984;2:283.

20. **Duan N.** Smearing estimate: A non-parametric retransformation method. J Am Stat Assoc 1983;78:605.

21. **Blough DK, Madden CW, Hornbrook MC.** Modeling risk using generalized linear models. J Health Econ 1999;18:153–171.

22. **Fishman PA, Goodman MG, Hornbrook MC, et al.** Selecting Functional Forms to Assure Appropriate Payment for Chronic Disease. Paper presented at 15th Annual Association for Health Serv Res Meeting, Atlanta, GA, June 1996:9.

23. **Mincer J, Zarnowitz V.** The evaluation of economic forecasts. Economic Forecasts and Expectations New York: National Bureau of Economic Research, 1969.

24. **Pope GC, Ellis RP, Ash AS, et al.** Principal inpatient diagnostic cost group model for Medicare risk adjustment. Health Care Finan Rev 2000;21:93–118.

25. **Anderson GF, Bilenker JH.** Capitation payment rates and implications for the general pediatrician. Curr Opin Pediatr 1998;10:480–485.

26. **Boyle CA, Decoufle P, Yeargin-Allsop M.** Prevalence and impact of developmental disabilities in US children. Pediatrics 1994;93:399–403.

27. **Hobbs N, Perrin J, Ireys H.** Chronically Ill Children and their Families. San Francisco, CA: Jossey-Bass; 1985.

28. **Neff JM, Anderson G.** Protecting children with chronic illness in a competitive marketplace. JAMA 1995;274:1866–1869.

29. **Newhouse JP, Manning WG, Keeler EB, et al.** Adjusting capitation rates using objective health measures and prior utilization. Health Care Financ Rev 1989;10:41–54.