

---

# A Clinically Detailed Risk Information System for Cost

Grace M. Carter, Ph.D., Robert M. Bell, Ph.D., Robert W. Dubois, M.D., Ph.D., George A. Goldberg, M.D., Emmett B. Keeler, Ph.D., John S. McAlearney, Edward P. Post, M.D., and J. David Rumpel

---

*The authors discuss a system that describes the resources needed to treat different subgroups of the population under age 65, based on burden of disease. It is based on 173 conditions, each with up to 3 severity levels, and contains models that combine prospective diagnoses with retrospectively determined elements. We used data from four different payers and standardized the cost of most services. Analyses showed that the models are replicable, are reasonably accurate, explain costs across payers, and reduce rewards for biased selection. A prospective model with additional payments for birth episodes and for serious problems in newborns would be an effective risk adjuster for Medicaid programs.*

## INTRODUCTION

We present a Clinically Detailed Risk Information System for Cost (CD-RISC), which describes the resources needed to treat different subgroups of the population under age 65 based on burden of disease. It could be used to adjust payments, to aid in negotiations between insurers and payers or providers, or to perform policy analyses.

Newhouse (1986) listed four criteria for judging such systems: strength of prediction of utilization; ease of collection; ease of audit and difficulty of gaming; and size of

incentives for inefficient care. We would add acceptability to all those with a stake in the outcome of the risk adjustment, including payers, plans, providers, and patients.

The strength of prediction, or the ability to predict variations in the expected cost of care due to observable patient characteristics, is widely seen as important to deter selection of low-cost patients by plans (Newhouse et al., 1989). In addition, this ability should provide plans with incentives to provide good care and especially good care for ill patients who tend to be more expensive. If capitated firms must bear 100 percent of the costs of the care they provide, they have strong incentives to cut costs (Ellis and McGuire, 1986). The same information barriers that enabled fee-for-service providers to order and profit from excess care may prevent patients or their purchasing agents from realizing that the patients are being undertreated. Because physicians value providing additional care, additional payment for expensive patients should provide incentives for providers to increase the quantity of care that they deliver to such patients (Ellis and McGuire, 1993).

Risk-adjusted payment can increase efficiency by deterring socially wasted effort to select patients (Newhouse, Buntin, and Chapman, 1997) and by replacing competition for healthy patients with price competition. Currently, most employers who offer a choice of plans subsidize higher cost plans, despite economists who argue that this removes incentives for efficiency in the production of health care (Hunt et al., 1997). Although it is not clear why

---

Grace M. Carter, Robert M. Bell, Emmett B. Keeler, John S. McAlearney, and J. David Rumpel are with RAND. Robert W. Dubois and George A. Goldberg are with Value Health Sciences. Edward P. Post is with the University of Pittsburgh. This work was supported by the Health Care Financing Administration (HCFA) under Contract Number 500-92-0023. The views expressed in this article are those of the authors and do not necessarily represent the views of RAND, Value Health Sciences, the University of California, Los Angeles, or HCFA.

employers subsidize, it may be that, in the absence of risk adjustment, equal payments might be perceived as unfair to plans with sicker patients and even to the patients themselves. At least theoretically, plans that provide more generous care are likely to be more attractive to sicker patients who will get more benefit from the more generous plan (Keeler, Carter, and Newhouse, 1998). Adequate risk adjustment would allow employers to pay for the extent to which illness differs across those choosing different plans but not pay extra for more expensive practice styles or higher prices (Robinson et al., 1991).

In designing a risk-adjustment system, we would like to account for the cost of efficiently providing health care to each person. However, it is an impossible task to determine efficient care for each possible disease—and indeed, efficient care varies over time as technology changes. Instead, in our design of CD-RISC, we estimate the average cost of care for population groups in a large population, including indemnity payers, Medicaid, and health maintenance organizations (HMOs). This is similar to the use of average charges in creating diagnosis-related group (DRG) weights as a surrogate for the cost of efficient care. In order not to confound the resource cost of care with variation in prices, we eliminate price variation among plans by standardizing costs.

Previous risk adjustment has been based on demographics, diagnoses derived from claims, past utilization and treatments, and survey data. Demographics and diagnoses are the most readily available data and we incorporate them into CD-RISC. We use only diagnoses from inpatient records and physician bills for visits and surgery. We decided not to use survey data because of the cost and because of the relatively greater susceptibility to gaming. Using retrospective data on current-year utilization

has advantages and disadvantages. Including it gives incentives to provide unnecessary services (Ellis et al., 1996). However, these variables also increase incentives to provide needed care and accuracy.

CD-RISC provides a selection of models that use different amounts of information about care delivered during the payment year and therefore provide different points on the tradeoff curve. By choosing a specific model, each user controls how much priority is put on possible risk-adjustment goals. If a prospective model is used for HMO payment, all the financial risk for care is on the health plan. Using retrospective data transfers part of the risk back to the payer. We combine both prospective and retrospective data in some models. Retrospective data can allow us to capture the large expected expenditures associated with births. Choosing a model with more retrospective data includes more diseases, both acute episodes and chronic diseases that are diagnosed for the first time during the year, and therefore makes the model more accurate. This greater accuracy should reduce incentives to select healthier patients into a plan. Because physicians value providing good care, episodes of illness and outlier payments can encourage an increase in the provision of care to vulnerable populations. On the other hand, retrospective models are likely more influenced by demand or “taste” for care (e.g., whether care is sought for minor problems) and by provider practice styles (e.g., whether telephone consultations where diagnoses are not recorded are encouraged). Further, more use of retrospective information may reduce incentives to prevent disease.

CD-RISC describes each patient’s clinical characteristics using conditions and severity levels with enough detail that physicians should believe that they reflect

the patient's resource needs. Our system also limits the ability to increase payment by certain kinds of upcoding.

## **DATA**

We received complete claims (or transactions) and eligibility data for 2 years from four different payers. The claims data include hospital bills, physician bills, other supplier bills, and prescription drug bills. We use all bills to determine the cost of health care for each enrollee. We use diagnoses recorded on inpatient and selected physician bills to determine clinical characteristics. The eligibility data identify each enrollee and list dates of enrollment and disenrollment.

The payers consist of the Michigan Medicaid agency, two managed care organizations (which we designate as "national HMO" and "western HMO" to preserve their anonymity), and an indemnity plan. The sample consisted of persons who were either continuously enrolled for a 2-year period or were born or died during the period and were continuously enrolled while alive. We used data on all such persons for the private payers and a 40-percent sample of Michigan Medicaid participants. The data covered approximately 360,000 persons, of whom 48 percent were insured by Medicaid.

## **METHODS**

### **Statistical Models**

We used weighted least-squares regression to fit a variety of risk-adjustment payment models to the costs for each sample member. The dependent variable in each regression was the patient's annualized, standardized cost during the second year of our data. The independent variables described the patient's age (in categories),

sex, and clinical conditions. Thus, the coefficient on a clinical condition provides an estimate of the marginal cost of that condition. We weighted those who were born or died by the fraction of the year that the patient was observed. This results in unbiased estimates of monthly payment rate and also compensates for the higher variance in the estimate of cost for these patients.

A split-sample technique was used for validation. We evaluated each model with respect to its ability to explain costs within and across population subgroups, to protect plans and/or providers from financial risk, and to reduce the rewards from selecting patients.

### **Model Portfolio Overview**

Each of the CD-RISC models varies in the amount of information that it uses about care delivered during the payment year. The models include a prospective model, which is based solely on diagnoses recorded before the start of the payment year, and a retrospective model, based on diagnoses recorded during the year for which costs are estimated. The diagnoses are grouped into conditions, with an attached severity level. The condition-severity groups are then organized into body systems, with only the most expensive group in each body system affecting prediction. However, other conditions in both the same and other body systems may affect severity level. When a higher severity level is due to the presence of other, lower cost, diseases in the same body system, the extra cost of the higher severity level is the sum of the effects of the lower cost disease and the interaction of the two conditions.

Other models combine the prospective model's diagnoses with information about high-cost episodes of illness that occur during the payment year or with outlier pay-

ments for high-cost cases. The episode-of-illness payment would be determined prospectively from the regression and would be added to the per member per month payment amount from the same regression.

We include three models with prospective diagnoses, supplemented with information about three different episode-of-illness components that pay for: (1) only birth episodes and the baby's condition, (2) a selected subset of episodes of illness judged by authors Dubois, Goldberg, and Post to be non-discretionary treatment and non-preventable illnesses, and (3) all expensive episodes of illness. A demographic model is described for comparison. Including the demographic model, we report results concerning a total of six models based on regression. For one of these models, the prospective model with birth episodes, we also analyzed the effect of adding outlier payments.

### Annualized Standardized Costs

In order to obtain relative resource costs for each person in our sample, we standardized the costs of each service so that they are the same for each occurrence of the same service. We used the resource-based relative value scale from the Medicare fee schedule to standardize most physician claims. For each hospital, we calculated a standardization factor that was proportional to the average allowed payment per discharge divided by the hospital's case-mix index. The case-mix index is the average of the weights for the DRGs assigned to its sample cases. The standardized cost for each case is calculated as the allowed payment for the case divided by the standardization factor, with the proportionality factor set so that the total estimated hospital payment in the private sec-

tor sample equaled the actual total private sector payment. This method adjusts for systematic variation in charges across hospitals while allowing the standardized cost of each case to vary with the charges for the individual case. (The ratio of the standardized costs for any two cases within the same hospital is the same as the ratio of allowed charges for the same two cases. Further details on the standardization algorithm can be found in Carter et al., 1997.)

We did not standardize claims for drugs or non-physician services such as durable medical equipment or facility charges for ambulatory surgery. The assumption of a national market for drugs and durable medical equipment is not unreasonable. The small size of the remaining non-standardized expenditures should limit the importance of our inability to standardize.

Costs for patients who were born or died during the year were annualized by dividing standardized costs by the fraction of the year that the patient was observed.

### Clinical Group Definitions

CD-RISC is based on an initial set of 173 conditions, each with up to 3 severity levels: usual or low, medium, and high. Seven of the conditions are further split by age. The conditions and severity levels are from the Practice Review System developed by Value Health Sciences to profile physician practices and were based initially on the subjective judgment of physician panels.

Each condition is a grouping of diagnostic codes from the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) (Public Health Service and Health Care Financing Administration, 1980). Examples include breast cancer, diabetes, urinary tract infection, and hypertension. We used diagnoses from inpatient bills and from physician bills for visits or

surgery to assign clinical groups. Diagnoses associated with services of pathology, radiology, immunization injections, anesthesia, and assistants at surgery were not allowed to assign conditions because we believe these are often incorrect.

Each ICD-9-CM code is assigned to, at most, one condition. Severity levels are assigned based on both the ICD-9-CM codes that define the condition and other codes that represent complications or comorbidities that increase the resources required to care for a condition. Thus, one ICD-9-CM code can affect the severity level of more than one condition. After all conditions have been assigned, the set of all ICD-9-CM codes for the patient are searched a second time to determine severity.

For example, diabetes is assigned based on the recording of an ICD-9-CM code beginning with 250. These codes also affect severity, as the code 250.00 for uncomplicated adult onset diabetes is assigned to low severity, but code 250.01 for uncomplicated juvenile diabetes is assigned to medium severity. Codes for diabetes with complications are assigned to either high or medium, depending upon the kind of complication or manifestation. For example, an ophthalmic manifestation (250.50) is assigned to medium severity, but either ketoacidosis (250.1x) or hyperosmolar coma (250.2x) is assigned high severity. In addition to the 250 codes, certain non-diabetes codes also affect assignment of severity. For example, patients with the codes 362.02 (proliferative diabetic retinopathy) are assigned to high severity. Those with specific coronary problems such as acute myocardial infarction (codes 410.xx) and skin problems that either complicate management of diabetes or indicate an advanced stage of the disease are assigned to at least medium severity. The patient is assigned only to the highest severity for which he or she qualifies, so a

person whose only relevant codes were 250.50 and 362.02 would be assigned to high severity, while a person with 250.50 and a code of the form 410.xx or 250.00 and 410.xx would be assigned to medium severity. Further information is available in Carter et al. (1997). The ICD-9-CM rules are available upon request from the authors.

Although 3 severity levels are defined for each disease condition, we combined severity levels with fewer than 30 cases in our analysis half-file. We also split 10 condition-severity combinations based on age. In each model, we tested 369 combinations of condition, severity, and age and then eliminated variables with insignificant coefficients. In the prospective model, we expected that statistically significant conditions would be chronic, recurrent, or of long duration. On the other hand, in the retrospective model, many acute conditions were expected to be also statistically significant.

### **Clinical Hierarchy**

The combinations of condition, severity, and age are organized into hierarchies within 16 body systems so that the model prediction for each patient depends upon, at most, one combination per body system. This reduces incentives for multiple coding of the same or related conditions. The body systems are based on the ICD-9-CM coding structure. Hierarchical systems were also used in the original Diagnostic Cost Group risk-adjustment model.

Two separate rankings were derived, one for the prospective model and one for the retrospective model. The ranking of condition-severity combinations within each hierarchy is in order of their costs as determined from the regression. The rankings were derived using an iterative procedure. When coefficients on higher severity group(s) were smaller than coeffi-

cients on lower severity group(s), the levels were combined. The prospective model ranking was used for all mixed models.

### Episodes of Illness

Episodes of illness are retrospective elements added to the prospective models. They consist of either selected high-cost conditions or specific clinical events (e.g., birth or surviving a heart attack) or treatments (e.g., bone marrow transplant.) Of the 27 episodes of illness, 11 are determined from the patients' retrospective ICD-9-CM codes, by assigning condition-severity combinations, just like prospective condition-severity combinations. The other 16 episodes are determined from records of treatment, either procedures or hospitalizations. Angioplasty and mastectomy were determined from either physician bills or hospitalizations. The 14 other clinical events or treatments classified as episodes of illness were determined from hospitalization records because all of them would require hospitalization under current standards of treatment.

These 27 episodes of illness are the only episodes of illness that we have modeled. They were chosen based on the literature and knowledge of expensive hospitalization episodes. Of course, the fully retrospective model includes many more retrospectively determined condition-severity combinations.

### Outlier Payments

Two models include outlier payments for a small number of patients who are exceptionally expensive relative to their predicted payment amount. The rationale is that these outlier cases are so unusual that their illness cannot be categorized by a payment system built with a reasonable number of payment categories. Outlier

payments can be viewed as reinsurance against the costs of receiving a person with extremely high medical costs.

We examined two outlier policies in which outlier payments constituted 2 and 4 percent of total payments, respectively. Each policy used a "fixed-loss" threshold, so that payments were made for part of the costs above the sum of the fixed loss and the pre-outlier payment (Keeler, Carter, and Newhouse 1988). The pre-outlier payment is from the prospective model, with birth episodes and baby's condition, with a multiplicative budget-neutrality adjustment.

### Split-Sample Validation

Preliminary versions of our regression models were fit with a randomly chosen half-sample and used to determine the hierarchical form of the final equations and to decide on which condition-severity combinations were to be included in each model. After deciding on the final form of the models, we evaluated their fit to the other half-sample (the validation sample) in two ways: (1) we used the equation (i.e., coefficients as well as variables) from the analysis sample to predict costs for the members of the validation sample and examined its accuracy, (2) we refitted the model on the validation sample and examined both its  $R^2$  and the consistency of the predictions between the model fit on the two samples.

We use the Efron  $R^2$ , which accounts for bias as well as variance in the error, as our measure of the accuracy of prediction (Efron, 1978). The Efron  $R^2$  for a subsample is merely:

$1 - \frac{\sum((\text{cost}_i - \text{prediction}_i)^2)}{\sum((\text{cost}_i - \bar{c})^2)}$ , where

$\bar{c}$  is the average cost for the subsample and the summations are over all members  $i$  of the subsample. If the subsample is the entire fitting sample, the Efron  $R^2$  will yield the traditional  $R^2$ .

## Simulations

Other validation activities were conducted using simulated payments based on the predictions from each of our models in order to assess their ability to explain costs within and across population subgroups, to protect plans and/or physician groups from financial risk, and to reduce rewards from selecting patients. We chose the individual as the basis of analysis rather than the beneficiary year (which we use in the regressions), with payments and costs adjusted to the fraction of the year that the person was present and no weighting required in the simulations. The individual provides the obvious metric for analyzing selection, financial risk, and outlier payments. The evaluation of the accuracy of payments for members of a subgroup is based on the Efron  $R^2$ , which accounts for both the bias in the mean prediction for the subgroup and the variance of the prediction error. The Efron  $R^2$  for the total population differs from the regression  $R^2$  because its unit of analysis is a person rather than a person-year.

The simulations assume no change in the amount of care delivered to each beneficiary in response to a change in payment.

## SPLIT-SAMPLE VALIDATION RESULTS

We summarize the results of the validation for three models. The independent variables in the prospective model are age categories, sex, age-sex interactions, and combinations of condition and severity based on diagnoses recorded in the year preceding the year in which costs were incurred. The second model, the newborn model, starts with the same variables as the prospective model and adds variables describing whether a female gave birth during the cost year and whether the

new baby suffered from either congenital anomalies or severe diseases of the newborn. The retrospective model replaces the condition-severity combinations in the prospective model with ones based on diagnoses recorded in the year in which costs were incurred.

The first three rows of Table 1 give the number of condition-severity combinations used to predict costs in each of these three models. As expected, more condition-severity combinations predict costs retrospectively and thus are represented in the model.

## Explanatory Power

Table 2 summarizes the ability of our three models to explain cost. The first two columns provide the  $R^2$  when each model was fit on each of the two different samples. The  $R^2$  on the validation half-sample is close to the fit on the analysis sample in all cases and even better than the analysis sample for the prospective model.

The last two columns of Table 2 show the ability of the equation from the analysis sample model to predict the costs of the validation sample. The bias is just the mean value of the annual cost of the validation sample minus the mean prediction and is quite small. For each model, the  $R^2$  is a little smaller than the  $R^2$  for the model from the same sample, indicating imprecision in the measurement of the costs of some diseases between the two samples. There is no sign of a grossly overfitted model.

## Consistency of Prediction

For each member of the validation sample and each model, we compared the prediction from the analysis-sample model with the prediction from the model estimated with the validation sample. The cor-

**Table 1**  
**Number of Each Type of Variable in Each Model**

Model	Condition-Severity Variables		Event or Treatment Variables
	Prospective	Retrospective	
<b>Split-Sample</b>			
Prospective	175	0	0
Add Birth and Baby's Condition	177	0	1
Retrospective	0	319	0
<b>Full-Sample</b>			
Demographics Only	0	0	0
Prospective	133	0	0
Add Birth and Baby's Condition	134	0	2
Add Non-Discretionary Episodes	135	2	7
Add All Potential Episodes	135	9	17
Retrospective	0	256	0

NOTE: All models also included 21 dummy variables to capture 22 age-sex cells.

SOURCE: Carter et al., Santa Monica, California, 1999.

**Table 2**  
**Predictive Power of Preliminary Model from Half-Samples**

Model	Same-Sample $R^2$		Validation-Sample Prediction from Analysis Model	
	Analysis Sample	Validation Sample	Bias	Efron
	Prospective	0.080	0.083	\$15
New-Baby	0.178	0.176	13	0.162
Retrospective	0.370	0.369	1	0.352

SOURCE: Carter et al., Santa Monica, California, 1999.

relations between the predictions are high—from 0.94 for the prospective model to 0.98 for the retrospective model. Nevertheless, there were noticeable differences between the predictions from the models fit on the two samples for some individuals, particularly for the retrospective model. For 9 percent of the validation sample, the absolute value of the difference between retrospective model predictions from the analysis and validation sample exceeded the prediction amount by 50 percent or more. For the prospective model, only 1.7 percent of the sample had such large relative differences between the predictions.

### Implications for Final Models

The variability in predictions, especially in the retrospective model, caused us to use more severe pruning rules for insignif-

icant variables in our final models than in the split-sample models and to enforce monotonicity in the coefficients on different severity levels within the same condition. As shown in the “Full-Sample” section of Table 1, the prospective model on the pooled sample uses only 133 combinations of condition-severity-age, instead of the split sample’s 175, and the retrospective model on the pooled sample uses 256 combinations of condition-severity-age, instead of the split sample’s 319. Despite the reduced randomness in the model because of its fewer parameters, we found only a modest improvement in the consistency of prediction between models fit on the analysis sample and the same model fit on the validation sample. Substantial random variability remains in the predictions due to the inherent variability in the cost of expensive, relatively rare, diseases.



## FINAL MODEL RESULTS

In addition to the three models in the validation analyses, our final payment models include a demographic model, two more episode models, and two payment systems that add outlier payments to estimated costs from the new-baby model. The first new-episode model adds only episodes that the physician authors of this article (Dubois, Goldberg, and Post) judged to be non-discretionary. The second new-episode model uses all 35 episodes of illness. For comparison, we use an age-and-sex model and a flat-capitation model that pays the average cost for each sample member.

We analyzed each of these models in order to show the value of the additional information being incorporated in terms of:

- Its ability to explain costs within and across our four different payers.
- The extent to which it subjects plans and other risk-bearing entities such as physicians and physician groups to random risk.
- The extent to which it limits rewards from various kinds of selection behavior on the part of plans.

To the extent possible, we compared our findings with those from other published risk-adjustment models.

### Explanatory Power

Table 3 provides the  $R^2$  from the fitting of each regression model on the complete data set. The age-and-sex  $R^2$  is very small, but this is consistent with the published literature. Weiner et al. (1991) report an  $R^2$  of total charges on age group and sex of 0.04 for one group-model HMO, and Smith and Weiner (1994) report a range of 0.03 to 0.06 for demographics explaining total

Table 3

#### $R^2$ Values from Final Models on Full Sample

Model	$R^2$
Age-Sex	0.031
Prospective	0.078
Add Birth and Baby's Condition	0.175
Add Non-Discretionary Episodes	0.194
Add All Episodes	0.288
Retrospective	0.370

SOURCE: Carter et al., Santa Monica, California, 1999.

charges in a variety of settings. Kronick et al. (1996) report an  $R^2$  range from 0.004 to 0.015 across the States in their sample of Medicaid disabled persons. The disabled have a much higher variance of costs than our general population.

The  $R^2$  values in Table 3 increase steadily as different kinds of information are added. The literature provides only a general comparison because the  $R^2$  depends upon the population being studied (Hadorn et al., 1993) and on specific study design parameters. For example, our decision to limit our analysis to those insured throughout the year plus those who were born or who died may have increased our  $R^2$  somewhat. (Ash et al., 1997, report an  $R^2$  of 0.094 for a large indemnity insurer for a prospective model and 0.396 for a retrospective model.) Smith and Weiner (1994) report that Ambulatory Care Groups, operating in a prospective model with demographics, predict 0.10 to 0.15 percent of the variance in total charges, which is better than our purely prospective model  $R^2$  of 0.078 but worse than the 0.175 of the prospective model that includes the new baby's condition and payment for all birth episodes.

On the other hand, our retrospective model performs notably better, with an  $R^2$  of 0.366, than the 0.12 to 0.20 reported in the same paper for Ambulatory Care Groups. Predictions based on the Disability Payment System do slightly bet-

**Table 4**  
**Net Sex and Age Effects on Annual Dollar Expenditures After Controlling for Clinical Variables, by Model**

Sex and Age Group	Demographic Only	Purely Prospective	New Baby	Non-Discretionary Episodes	All Episodes	Purely Retrospective
<b>Males</b>						
To Be Born	\$7,176	\$7,176	\$3,583	\$3,773	\$3,514	\$2,606
Up to 1 Year	1,945	1,430	1,361	1,236	1,030	271
1 Year	1,028	542	509	520	515	47
2 Years	790	462	444	442	425	82
3 Years	677	412	406	398	397	67
4 Years	836	560	553	522	512	183
5-12 Years	596	354	358	334	339	101
13-18 Years	846	541	546	480	469	205
19-34 Years	687	396	398	370	324	113
35-44 Years	1,284	840	841	802	665	363
45-54 Years	1,807	1,222	1,221	1,180	908	516
55-64 Years	2,843	1,805	1,807	1,775	1,143	648
<b>Females</b>						
To Be Born	6,230	6,230	2,967	3,164	2,867	1,956
Up to 1 Year	1,544	1,125	1,080	959	834	174
1 Year	872	478	486	469	418	114
2 Years	696	391	388	382	368	51
3 Years	664	409	415	393	367	105
4 Years	538	325	329	328	323	131
5-12 Years	473	261	267	248	248	66
13-18 Years	1,197	772	577	545	541	208
19-34 Years	2,017	1,149	782	768	686	129
35-44 Years	1,889	1,075	1,026	978	898	242
45-54 Years	2,241	1,337	1,339	1,282	1,063	434
55-64 Years	3,087	1,874	1,879	1,803	1,329	625

SOURCE: Carter et al., Santa Monica, California, 1999.

ter in the disabled population than our predictions for the general population, with an  $R^2$  range of 0.16 to 0.22 for a prospective model and from 0.38 to 0.46 for a retrospective model (Kronick et al., 1996). Predictions from the Hierarchical Coexisting Conditions model in the Medicare population are more similar to our models: Ellis et al. (1996) report an  $R^2$  of 0.086 in a purely prospective model and 0.4074 in a retrospective model based solely on conditions.<sup>1</sup>

<sup>1</sup> Weiner et al. (1996) also discuss a risk-adjustment model for the Medicare population. However, they do not appear to provide an  $R^2$  that is comparable to those discussed in this article. However, they provide payment-to-cost ratios in randomly selected groups of 5,000. A comparison of these to our analogous results to be presented later suggests that their  $R^2$  is lower than ours.

## Model Coefficients

### Demographics

The age and sex coefficients in the models are shown in Table 4 (no intercept appears in the models). Expenditures decline in the first few years of life and then remain about constant through childhood. Expenses for females begin to rise in the adolescent years and are substantially higher in the years of highest fertility. Average expenses for males do not begin to rise until about the mid-thirties.

Annualized costs for the care of a newborn male for the first 6 months of life are estimated as \$7,176 in the sex-age model. Female newborns' annualized expenses

average \$946 less. The few babies that are born with the serious problems controlled for in the new-baby model account for about one-half of the total expenses associated with newborns in this population. A male baby without these problems will have annualized costs of only \$3,583.

In all demographic groups, the size of the coefficients generally declines as one moves across the table. The difference between the demographic-only coefficient and the coefficient in any other model is the amount of money that is explained by the disease variables in that model in that age group. This difference is largest for the oldest group in the prospective model and is substantial for children only in the retrospective model (except for the “to be born” group previously discussed). The retrospective model shows much smaller effects of age and sex than the other models, but the remaining effects are large enough to be important.

#### Prospective and New-Baby Models

The new-baby model is identical to the purely prospective model, except that it includes a variable for a birth episode and two variables for the condition of the newborn in the current year. The coefficients for each of the condition-severity variables in both models are similar, so we present details only for the new-baby model. Coefficients for 3 of the 16 body-system hierarchies are shown in Table 5, with all variables being shown in the Technical Note. The most expensive diseases are lung cancer, metastatic cancer, congestive heart failure in a child, human immunodeficiency virus (HIV) infection, high- and medium-severity mental retardation, renal failure, high- and medium-severity congenital anomalies in the newborn, and high-severity diseases of the newborn; all these increase next-year costs by at least \$15,000.

There is a large difference in the cost of congenital anomalies between newborns and those born even a year earlier. Although these conditions persist, many treatments, such as surgical corrections, are concentrated in the period soon after birth.

About two-thirds of the conditions in the prospective models are chronic conditions. The non-chronic conditions are usually either protracted or recurrent. When multiple severity levels appear in these models, the effect of severity is usually quite important. Examples include breast cancer (medium or high severity=\$11,521; usual severity=\$5,436) and diabetes (high=\$6,358; medium=3,952; low=\$1,547). Also, in many conditions, only the high- and medium-severity levels predict increased cost.

#### Retrospective Model

Many more condition-severity combinations explain costs retrospectively (256) than prospectively (133). The magnitudes of the coefficients, also shown in Table 5 and the Technical Note, seem plausible. Some of the most expensive diseases are similar to those in the prospective model: metastatic cancer, congestive heart failure in a child, high- and medium-severity mental retardation and renal failure. Other expensive conditions are acute conditions or chronic conditions with intense acute episodes: ischemic heart disease, cerebrovascular disease, high-severity conduction or rhythm problems, high-severity gastrointestinal tract disorders, and septicemia. The majority of conditions that appear in both models have larger effects in the retrospective model.

Hypertension is one of the exceptions to the general rule that chronic conditions have larger coefficients in the retrospective model than in the prospective model. Hypertension helps predict future costs because those with this condition are at risk for more expensive vascular diseases.

**Table 5**  
**Illustrative Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
<b>Blood</b>					
Non-Deficiency Anemia	High	10,662	17.35	8,875	19.46
Other Blood Disorder	High	5,601	7.08	16,721	30.36
Other Blood Disorder	Low or Medium	1,796	6.48	7,335	31.23
Deficiency Anemia	Medium or High	1,281	3.15	4,198	12.14
Non-Deficiency Anemia	Low or Medium	NA	NA	4,242	6.85
Deficiency Anemia	Low	NA	NA	1,171	6.55
<b>Neoplasm</b>					
Lung Cancer	Any	22,867	21.53	14,120	18.9
Metastatic Cancer	Any	18,522	27.38	34,898	85.7
Breast Cancer	Medium or High	11,521	6.52	13,318	11.55
Hematological or Lymphatic Cancer	High	11,489	19.89	14,515	32.96
Other Cancer	High	7,985	12.21	10,921	22.69
Breast Cancer	Low	5,436	11.69	5,450	14.17
Unspecified Neoplasm	Medium or High	4,676	8.4	3,188	5.78
Colorectal Cancer	Any	3,257	2.94	12,888	16.15
Other Cancer	Low or Medium	2,627	5.87	6,072	17.85
Benign Neoplasm	Medium or High	NA	NA	3,886	17.79
Cancer in Situ	Any	NA	NA	2,721	5.65
Hematological or Lymphatic Cancer	Low or Medium	NA	NA	1,793	2.94
Benign Neoplasm	Low	NA	NA	1,353	6.36
Unspecified Neoplasm	Low	NA	NA	810	2.6
<b>Newborn<sup>1</sup></b>					
Congenital Anomaly	High	4,561	19.81	4,472	18.99
Congenital Anomaly	Medium	2,211	8.36	2,586	9.69
<b>Newborn<sup>2</sup></b>					
Congenital Anomaly	High	36,878	93.04	14,394	33.7
Congenital Anomaly	Medium	21,918	47.82	4,207	9.2
<b>Episode</b>					
Birth	Any	5,526	66.43	3,485	34.68

<sup>1</sup> Infant born in previous year.

<sup>2</sup> Infant born in prediction year.

NOTES: Remaining coefficients are found in Table A in the Technical Note. NA is not applicable.

SOURCE: Carter et al., Santa Monica, California, 1999.

But if none of those develop during the year, then hypertension causes only a modest increase in costs (medium or high severity=\$1,687, usual severity=\$576). The effects of diabetes and HIV infection also are smaller in the retrospective model than in the prospective model. This is likely because the complications of these diseases show up in other body systems, and their costs are added there.<sup>2</sup>

<sup>2</sup> Another reason for lower costs of some diseases in the retrospective model than in the prospective model is the different hierarchies used in the two models. In the prospective model, lung cancer takes precedence over metastatic cancer. In the retrospective model, metastatic cancer takes precedence because of the high costs associated with treatment episodes. This moves the most expensive lung cancer cases into this group, lowering the cost of the remaining lung cancer cases.

## Episode-of-Illness Models

As shown in Table 6, several of the episodes of illness are associated with extremely large expenditures: bone marrow transplant (\$148,500), kidney transplant (\$94,600), and tracheotomy-ventilator costs (\$90,600) are the most expensive. Six of the 10 non-discretionary episodes and 17 of the entire 27 episodes cost more than \$15,000.

Most of the condition-severity variable coefficients in the episode models are similar to those of the new-baby model (Carter et al., 1997); thus, we do not report the details in this article. However, a few drop

**Table 6**  
**Episode-of-Illness Coefficients for Dollar Expenditures, by Model**

Episode of Illness	Severity	Only Non-Discretionary Episodes		All Episodes	
		Coefficient	t-Statistic	Coefficient	t-Statistic
Birth	Any	5,399	65.35	5,373	69.18
Extensive Burns	NA	19,790	14.51	20,044	15.64
Craniotomy	NA	35,339	49.62	32,818	49.00
Bone Fracture	High	11,805	35.59	10,560	33.83
Internal Trauma or Injury	Medium or High	10,571	19.07	9,209	17.66
Neonatal Disorder, Episode	High	23,803	138.68	22,918	141.88
Kidney Transplant	NA	94,624	34.20	96,189	36.99
Major Multiple Trauma	NA	25,836	23.16	23,755	22.65
Mastectomy	NA	12,750	22.74	13,014	24.70
Bone Marrow Transplant	NA	148,535	69.56	145,833	72.62
Inguinal Abdominal Hernia	Medium or High	NA	NA	7,425	15.43
Gall Bladder Disease	Medium or High	NA	NA	11,757	21.05
Gall Bladder Disease	Low	NA	NA	6,039	33.32
Gastrointestinal Tract Hemorrhage	Any	NA	NA	8,837	22.15
Septicemia	High	NA	NA	31,455	48.96
Septicemia	Medium	NA	NA	11,229	25.30
Septicemia	Low	NA	NA	10,918	42.08
Acute Myocardial Infarction	NA	NA	NA	19,512	30.62
Back or Neck Operation	NA	NA	NA	16,437	54.48
Major Joint Replacement	NA	NA	NA	27,552	41.50
Major Kidney Surgeries	NA	NA	NA	15,401	17.82
Large Bowel and Related Surgeries	NA	NA	NA	25,851	50.59
Major Chest Surgery for Respiratory Disease	NA	NA	NA	48,592	53.08
Open Heart Surgery	NA	NA	NA	46,860	89.88
Angioplasty	NA	NA	NA	27,439	47.51
Stroke	NA	NA	NA	25,082	24.93
Ventilator Use and/or Tracheostomy	NA	NA	NA	90,617	115.98

NOTE: NA is not applicable.

SOURCE: Carter et al., Santa Monica, California, 1999.

substantially: metastatic cancer goes from \$18,522 to \$11,400; congestive heart failure in children from \$15,740 to \$4,852 and in adults from \$9,041 to \$5,876; high-severity ischemic heart disease from \$4,620 to a negative value; high-severity osteoarthritis from \$7,777 to \$1,929. Expenditures for these chronic diseases are concentrated in episodes of intense care.

### **Payer Analysis of Pooled Predictions**

The regression models just presented estimate the standardized cost of conditions averaged across the combined patient population for all of our payers. We examined the extent to which these models explain the costs within and among each of our payers using the simulation model. Then we examined the value of several strategies for improving payer-specific predictions.

Table 7 shows the bias in the average payment amount for each payer and model. For all except the models with outlier payments, the payment is merely the prediction from the regression model adjusted for the fraction of the year that the person was enrolled. The first row just shows the difference between average standardized costs for the group (which would be paid under a flat capitation) and the overall average cost. As is well known, the per capita utilization of the Medicaid population exceeds that of private payers. Among private payers, the indemnity population is the most expensive. The next row shows that the age and sex of the populations explain little of the difference in costs among payers. However, the mostly chronic conditions in the prospective model are a major reason for the large differences between Medicaid and private

**Table 7**  
**Bias in Prediction of Mean Cost for Each Payer, by Model**

Payment Model	Mean Cost Minus Mean Payment				
	Medicaid	WHMO	Indemnity	NHMO	All Private
Flat Capitation	\$275	-\$313	-\$91	-\$390	-\$251
Age-Sex	279	-217	-210	-456	-255
Prospective	93	-122	45	-238	-85
Prospective+Births, New Baby's Condition	63	-117	98	-192	-57
Add Non-Discretionary Episode	60	-123	105	-179	-55
Add All Episodes	47	-113	120	-165	-43
Add 2-Percent Outlier	70	-128	87	-182	-64
Add 4-Percent Outlier	76	-126	72	-176	-68
Retrospective	-75	-14	284	-30	84

NOTES: The data covered 171,591 Medicaid beneficiaries, 93,210 members of the WHMO, 63,637 insured by the indemnity plan, and 31,254 members of the NHMO. The standardized cost for each of the four payers (in order) was \$1,827; \$1,136; \$1,351; and \$1,060 for an overall average annual standardized expenditure of \$1,488. HMO is health maintenance organization. WHMO is western HMO. NHMO is national HMO.

SOURCE: Carter et al., Santa Monica, California, 1999.

payers. The bias of the prediction for Medicaid declines from underestimating costs by \$279 using age and sex to an underestimate of only \$93. The addition of the new baby's condition and births continues to lower the bias.

Interestingly enough, in the retrospective model, the bias actually turns to a relatively small overestimate of costs. The reason for this change in the bias of costs is not clear without additional data. It could be due to Medicaid patients actually getting less care than private patients with the same disease or to Medicaid patients receiving less care per visit and thus needing more visits for the same services (which we would count as more services.) In any case, despite the possibility of some measurement error in our valuing services, it is clear that the great majority of the extra costs of care for Medicaid patients is linked to their increased burden of illness.

The other interesting contrast is among the private payers. Comparing the prospective model with the age-sex model, one finds that controlling for chronic diseases widens the difference in estimated payments between the indemnity plan and the HMOs, with the cost of the indemnity plan being underestimated and the cost of the HMO being overestimated. The addition of controls for episodes of illness or for

all the conditions in the retrospective model continues to magnify the difference between the indemnity plan and HMOs while simultaneously narrowing the difference among the HMOs. This is consistent with the expected result that HMOs have more episodes of illness but lower costs per episode compared with indemnity plans (Keeler and Rolph, 1988).

Outlier payments, which represent only 2 or 4 percent of payments, have little effect on our ability to pay in proportion to the costs of the different payer groups.

Table 8 shows the ability of our pooled model to predict costs within each payer's population. It addresses the extent to which a payer or health plan can examine the relative costliness of patients using a pooled model. The demographic model explains little in any of the private payers and only 3 percent of the variance in the Medicaid population. The prospective model provides some predictive power in all groups. The addition of the new baby's condition increases  $R^2$  in proportion to the proportion of the group who are new babies: It greatly increases the ability to predict Medicaid beneficiaries' costs and also helps substantially for the national HMO and western HMO. However, the addition of this condition hardly improves prediction at all for the indemnity population, for whom less than 1 percent of mem-

**Table 8**  
**Ability to Predict Costs Within Each Payer, by Model**

Payment Model	Efron $R^2$				
	Total	Medicaid	WHMO	Indemnity	NHMO
Age-Sex	0.02	0.03	0.01	0.01	0.00
Prospective	0.09	0.08	0.07	0.12	0.04
Prospective+Births, New Baby's Condition	0.16	0.21	0.11	0.13	0.09
Add Non-Discretionary Episode	0.19	0.20	0.19	0.17	0.14
Add All Episodes	0.32	0.26	0.37	0.37	0.38
Add 2-Percent Outlier	0.44	0.43	0.55	0.36	0.30
Add 4-Percent Outlier	0.55	0.54	0.65	0.50	0.43
Retrospective	0.42	0.39	0.41	0.50	0.53

NOTES: HMO is health maintenance organization. WHMO is western HMO. NHMO is national HMO.  
SOURCE: Carter et al., Santa Monica, California, 1999.

bers are newborns. The  $R^2$  of this model is quite adequate in all the payer groups, ranging from 0.09 through 0.13 for the private payers, and is 0.21 for our Medicaid sample with its high fraction of newborns.

The addition of episodes of illness or the information in the retrospective model has a much larger effect on the private payers than on the Medicaid population. This is related to the much higher percentage of adults in these populations. The episode and retrospective models explain cost much more in the adult population than in children and explain cost most for the oldest adults.

Outliers occur in all populations and thus, the outlier policies substantially improve  $R^2$  for all payers. Outlier payments account for random variation in costs, rather than systematic variation in costs.

### Adding Payer Effects

The CD-RISC models fit to data that have been pooled across payers provide reasonable explanatory power within each of the payer's data sets. Thus, we believe these or other estimates derived by averaging across payer groups are best for setting payment rates when beneficiaries may choose among plans. For other purposes, one might want the best estimate for an individ-

ual plan. Would the models do better if we allowed each payer's data to affect the prediction differently? We explored this question using the new-baby model and the retrospective model. Several of the episodes do not have a large enough sample to obtain separate estimates for all payers.

Many condition-severity combinations also do not have sufficient sample size to allow precise estimates of payer-specific effects. Thus, fitting completely separate models for each payer removes biases in the pooled model caused by ignoring practice-pattern differences across payers—but at a cost of introducing substantial overfitting. Consequently, we examined the effect of using a much smaller number of parameters to capture payer effects.

We report in Table 9 the effect of different methods of adding payer effects:

- Method 1: A regression for each payer of annualized standardized cost as a linear function of the prediction from the pooled model (which we call the payer-specific linear correction).
- Method 2: A pooled regression allowing for the interaction of payer and all 13 condition-severity combinations with at least 30 persons within each payer and with a coefficient of \$1,900 or more (the expensive, high-volume interaction).
- Method 3: Separate regressions for each payer (the full regression). Although the

**Table 9**  
**Comparison of  $R^2$  Values for Payer-Specific Model with Models from Pooled Data**

Payment Model	Efron $R^2$				
	Total	Medicaid	WHMO	Indemnity	NHMO
<b>Prospective with Birth, New Baby's Conditions</b>					
Pooled Data	0.162	0.206	0.109	0.131	0.090
Payer-Specific Linear Correction	0.166	0.210	0.109	0.140	0.100
Expensive, High-Volume Interaction	0.163	0.207	0.110	0.134	0.099
Full Regression	0.178	0.214	0.129	0.158	0.129
<b>Retrospective</b>					
Pooled Data	0.420	0.386	0.408	0.505	0.533
Payer-Specific Linear Correction	0.166	0.210	0.109	0.140	0.100
Expensive, High-Volume Interaction	0.427	0.390	0.414	0.521	0.540
Full Regression	0.459	0.403	0.462	0.583	0.626

NOTES: HMO is health maintenance organization. WHMO is western HMO. NHMO is national HMO.

SOURCE: Carter et al., Santa Monica, California, 1999.

small sample size for some conditions introduces substantial variability into the predictions of this model, we include this model for completeness.

As expected from Table 7, payer-specific effects exist. In the models using Method 1 to capture payer effects, the intercept terms are significantly different from zero in seven of the eight cases, and the slope coefficients are significantly different from 1.0 in all cases and numerically important.

Table 9 compares the Efron  $R^2$  for the pooled data with the similarly calculated  $R^2$  for the models with payer-specific effects. The first method of payer-specific effects causes only a modest increase in  $R^2$ . The  $R^2$  values for the second method that allows payer main effects and an interaction of payer with high-volume, expensive diseases are roughly comparable to the simple linear correction. The payer-specific full regressions fit more closely than the linear model because they can exploit random variation in the data. Indeed, the regressions are likely overfitting the data. Still, no matter how we specify the payer effect, most of the explainable variation in per person costs is attributable to diseases, not to practice patterns associated with particular payers.

The first method appears to capture most of the differences across payers that we can measure with any precision. In general, the same diseases tend to be more expensive for all four payers. Only 4 of the 39 terms interacting payer with expensive high-volume conditions are significant at the  $p=0.05$  level, a result explainable by chance ( $p=0.13$  using the Poisson approximation to the binomial). Thus, we cannot measure the cost of individual diseases precisely enough to have any confidence in the payer-specific estimates.

### Random Risk

Another problem with any prospectively set price for the health care of groups of individuals is that a plan might receive patients who are costlier than average just by chance. We randomly assigned each of our patients into groups of size 1,000, 5,000, and 50,000, until we had defined 500 groups of each size, without regard for the identity of the patients' payer. Then we simulated the payment each group would receive under each payment system.

Risks are substantial for groups of 1,000 or 5,000 patients. For example, 5 percent of the smallest groups would have received



**Table 10**

**Bias in Prediction of Mean Second-Year Cost for Persons Grouped by Quintiles of First-Year and Second-Year Expenses, by Model**

Payment Model	Mean Cost Minus Mean Payment			
	First Year		Second Year	
	Quintile 1	Quintile 5	Quintile 1	Quintile 5
Flat Capitation	-\$946	\$2,024	-\$1,437	\$4,422
Age-Sex	-913	1,646	-1,242	3,780
Prospective	-421	439	-846	3,103
New-Baby	-379	455	-763	2,592
Add Non-Discretionary Episode	-362	430	-731	2,509
Add All Episodes	-318	382	-637	2,124
Add 2-Percent Outlier	-367	441	-748	2,512
Add 4-Percent Outlier	-358	423	-733	2,439
Retrospective	-119	243	-222	885
Actual Cost in Year 2	348	3,318	0	5,858
Actual Cost in Year 1	0	5,402	270	3,732
Sample Size	71,889	65,616	71,939	71,938

NOTES: First-year sample excludes babies born during second year. The first quintile of first-year expenses contains all those with zero first-year expenses.

SOURCE: Carter et al., Santa Monica, California, 1999.

payments covering less than 81.4 percent of the cost of care for their patients under a flat per member per month system. Even for groups of 5,000 patients, 5 percent of groups receive payments that cover less than 91.3 percent of costs. Models that better match costs modestly decrease this risk, and the retrospective payment model and the outlier payment policies reduce risk most.

Outlier payments or the retrospective model do modestly reduce the financial risk associated with care for small groups of patients. For example, the proportion of groups of 5,000 patients that are reimbursed at less than 91 percent of cost declines from 5 percent to 1 percent (Carter et al., 1997). Nevertheless, substantial risk remains, with 5 percent of cases of 1,000 patients receiving less than 87 percent of cost and 5 percent of cases of 5,000 patients receiving less than 94 percent of costs. Thus, individual physician practices or groups of two or three doctors should probably not bear full risk. When partial risk is borne, accounting for outlier cases can lower the chance that a doctor will receive reduced income because of the chance draw of expensive patients and improve fairness.

The groups of 50,000 experience much reduced risk under all payment models. Although the outlier payments and retrospective models again provide the most reduction in risk, the amount is relatively small.

**Selection Incentive**

We examined the extent to which the payment models reduce the rewards that plans receive by avoiding persons needing expensive care or selecting those who are exceptionally inexpensive. One way in which plans can affect their beneficiary population is by observing the actual expenses of their members and by subtly encouraging disenrollment of the expensive persons or by aggressively working to retain the less expensive. Table 10 shows how much the plan would gain or lose, under each payment plan, by enrolling or retaining members in the lowest and highest quintiles of first-year expenses.

Each person in the lowest quintile of first-year expenses would have second-year expenses \$946 less than a flat per member per month system would pay for them, while those in the highest quintile of

first-year expenses would cause the plan to lose \$2,024. The age-and-sex adjustment reduces the rewards from selection by only a small amount. The prospective model, however, causes a large reduction in the rewards from selection. Profits from the lowest cost cases are more than cut in half. The loss from the highest quintile members is cut by 78 percent, compared with the flat-payment system, and by 73 percent, compared with an age-and-sex-adjusted payment system. The addition of births and conditions associated with babies born in the second year has little effect in groups based on first-year expenses because newborns are not included in the sample analyzed there.

The episode models with payments for non-discretionary episodes and outliers reduce the rewards to selection by little compared with the new-baby model. This implies that the previous year's costs are not predictive of these expenses, and therefore these payments do not help against selection based on the previous year's expenses.

Interestingly, the retrospective model does further reduce rewards from selection based on the previous year's expenses. Compared with the new-baby model, the retrospective model reduces the reward for selection of the least expensive members during the first year by 69 percent and the loss from the most expensive quintile by almost one-half.

Last year's expenses are not the only way plans could identify which individuals they wish to enroll and whom they wish to avoid or disenroll. Grouping patients by actual second-year expenses provides an upper bound on the reward from these selection behaviors. It also provides an indication of the extent to which the payment system will be fair to plans and providers who, for whatever reason, obtain more than their share of costly beneficia-

ries. Table 10 also shows the losses associated with patients grouped by the first and last quintiles of cost during the prediction year. Adjustment for age and sex, chronic conditions, and births and new babies' conditions each produce a modest decrease in the heterogeneity of losses with a substantial cumulative effect. The new-baby model reduces profits on the lowest cost quintile by 47 percent and losses on the highest quintile of costs by 41 percent compared with flat capitation.

The episode models or outlier payments add little by this measure. These payments affect too few people to have a measurable effect at this level of aggregation. The retrospective model has a substantial effect on decreasing the heterogeneity of losses. Profit on the lowest quintile is reduced by 85 percent and losses on the highest quintile are reduced by 80 percent, compared with flat capitation.

## **DISCUSSION**

### **Model Quality**

The validation exercise showed that the general method is sound. The hierarchies, which were determined based solely on the analysis sample, are robust and predict well in an independent sample.

Although it is difficult to make exact comparisons with other models in the literature because of differences in methods and data sets, the CD-RISC models appear to have an explanatory power at least equal to those reported in the literature. Further, and perhaps more importantly, the models are effective at reducing rewards from selection.

We believe that the models work well because of the combination of the definition of severity level with the use of the body-system hierarchy. The severity-of-illness level in our model appears to be a powerful

predictor of future and current costs. It depends upon both the details of the disease, as recorded in the precise diagnoses assigned, and on the presence of comorbidities that either signal an advanced stage of the disease or complicate the treatment of the disease as judged by physicians. Thus, high-severity cases are often the interaction of a major disease with a somewhat less expensive one in the same body system, whose total costs for management are less than would be estimated by the sum of two independent costs. When costs are in different body systems, they are more likely to be additive, and only rarely does a diagnosis from a different body system affect severity level. Further analysis that used the same data to compare CD-RISC models with models with a different structure are needed to verify this conjecture.

Most of the variation in per person costs is attributable to disease incidence, not to particular payers. Thus, standardizing costs and pooling across payers to develop a risk-adjustment model allows one to estimate the average resource costs associated with diseases within a population without distortion from the differential prices paid by the payers and different administrative costs.

### **Model Uses**

The CD-RISC contains a variety of individual risk-adjustment models that differ in the extent to which they achieve different goals. We believe that there is not one “best” model but rather that several of our models could be optimum in different situations, depending upon the goals and values of the stakeholders in the outcome of the risk adjustment.

Improving the accuracy of individual predictions does not necessarily imply reducing incentives for some selection

behaviors or reducing financial risk. All of the CD-RISC models have the advantage of describing the patients’ clinical characteristics using conditions and severity levels that provide enough detail that physicians should believe that they accurately describe resource needs. All the models should have more face validity than models where patients are grouped based purely on the average costs of their disease.

We recommend that those responsible for developing payment systems for Medicaid programs or any other large population of those under age 65 give consideration to risk adjusting with the prospective model that includes birth episodes and payments for serious problems in newborns. This model has strong cost-control incentives because plans bear 100 percent of the marginal costs of the care they deliver. It should be perceived to be fair by plans and providers because it is based solely on patients’ clinical needs. We believe the inclusion of information about birth episodes and the condition of the baby will not induce poor care but will protect plans and patients against adverse selection. It has satisfactory explanatory power, explains differences across payer groups quite well, and substantially reduces the reward from selection.

Whether payments for non-discretionary episodes of care should be added to this payment system depends upon the circumstances. Episode payments should be considered when one desires incentives to avoid the underprovision of valuable care such as transplants or if one believes that selection against the relevant population is likely. If payers are concerned about episode payments inducing care of low value, they have the option of using a blended model that pays less than the expected cost of an episode. Episode payments should also be considered in dealing

with payments to small organizations such as physician groups. Assignment of patients may not be random; indeed some practices may have reputations for dealing with particular types of problems. In these circumstances, the use of episode payments could improve fairness.

Although the retrospective model is based on ICD-9-CM diagnoses and not treatment, it is probably more affected by practice styles and patients' taste for health care relative to other goods than the prospective model. It has weaker cost-control incentives than the prospective models. However, the retrospective model also has some advantages that may offset these drawbacks for particular uses. It has a much higher  $R^2$  than any of our other models and the lowest rewards from selection. Thus, it might be useful for payment or allocation of bonuses among or within physician practices that agreed on style of care and/or are concerned about adequately paying the physicians who care for the sickest patients. It is also useful for analysis because of its greater accuracy and because the drawbacks related to incentives do not apply to analyses of past behavior.

Outlier payments have almost no effect on bias in payment for large groups of patients. Such payments affect only risk reduction and incentives to provide care to expensive patients. Outlier cases are so unusual it is impossible to predict their costs from our clinical or demographic data and, we suspect, it may be impossible to predict these costs sufficiently in advance from any data. Because risk declines, in percentage terms, with the size of the population being paid for, outlier payments are really important for risk reduction only for small groups. Even for large groups, one might wish to add outlier payments to ensure that expensive patients receive adequate care.

A limitation of all these models, but especially the retrospective model, is the need for a large data set to accurately calibrate the model and obtain good coefficients. Larger data sets will soon be available from data warehouses, and it would be interesting to determine how large a data set is needed to obtain relatively stable coefficients. For most purposes, the predictions from the models can be viewed as only relative and can be recalibrated by fitting costs for a new payer as a linear function of older CD-RISC predictions derived from a very large data set.

## TECHNICAL NOTE

Table A presents all of the coefficients and  $t$ -statistics for the clinical conditions in the prospective model with new-baby episodes that depend upon the condition of the baby at birth and in our retrospective model. Each coefficient gives the estimated marginal expenditure for a person with that condition and severity level. Within each body system, the condition-severity level variables are sorted in decreasing order of their coefficients in the new-baby model. Many more conditions were included in the retrospective model than in the prospective model.

## REFERENCES

- Ash, A., Ellis, R.P., Yu, W., et al.: *Risk Adjustment for the Non-Elderly*, Boston Medical Center. Submitted to the Health Care Financing Administration under Contract Number 18-C-90462/1-02. Boston University Medical School. Boston. Sept. 4, 1997.
- Carter, G.M., Bell, R.B., Dubois, R.W., et al.: *A Clinically Detailed Risk Information System for Cost*. DRU-1731-1-HCFA. RAND. Santa Monica, CA. 1997.
- Efron, B.: Regression and ANOVA with Zero-1 Data: Measures of Residual Variation. *Journal of the American Statistical Association* 73:113-121, 1978.
- Ellis, R.P., and McGuire, T.G.: Provider Behavior Under Prospective Payment. *Journal of Health Economics* 5(2):129-151, June 1986.

- Ellis, R.P., and McGuire, T.G.: Supply-Side and Demand-Side Cost Sharing in Health Care. *Journal of Economic Perspectives* 7(4):135-151, Fall 1993.
- Ellis, R.P., Pope, G.C., Iezzoni, L.I., et al.: Diagnosis-Based Risk Adjustment for Medicare Capitation Payments. *Health Care Financing Review* 17(3):101-128, Spring 1996.
- Hadorn, D.C., Keeler, E.B., Rogers, W.H., and Brook, R.H.: *Assessing the Performance of Mortality Prediction Models*. MR-181-HCFA. RAND. Santa Monica, CA. 1993.
- Hunt, K.A., Singer, S.J., Gabel, J., et al.: Paying More Twice: When Employers Subsidize the Difference in Prices Among the Insurance Plans They Offer their employees? *Health Affairs* 16(6):150-156, Nov/Dec 1997.
- Keeler, E., Carter, G., and Newhouse, J.: A Model of the Impact of Reimbursement Schemes on Health Plan Choice. *Journal of Health Economics* 17:297-320, 1988.
- Keeler, E. B., Carter, G.M., and Trude, S.: Insurance Aspects of DRG Outlier Payments. *Journal of Health Economics* 7(3):193-214, Sept. 1988.
- Keeler E.B., and Rolph, J.: The Demand for Episodes of Treatment in the Health Insurance Experiment. *Journal of Health Economics* 7(4):337-367, Dec. 1988.
- Kronick, R., Dreyfus, T., Lee, L., and Zhou, Z.: Diagnostic Risk Adjustment for Medicaid, the Disability Payment System. *Health Care Financing Review* 17(3):7-34, Spring 1996.
- Newhouse, J.P.: Rate Adjusters for Medicare Capitation. *Health Care Financing Review*. 1986 Annual Supplement, HCFA Pub. No. 03225. Health Care Financing Administration. December 1986.
- Newhouse, J.P., Buntin, M.B., and Chapman, J.D.: Risk Adjustment and Medicare. *Health Affairs* 16(3): 26-43, 1997
- Newhouse, J.P., Manning, W.G., Keeler, E.B., and Sloss, E.M.: Adjusting Capitation Rates Using Objective Health Measures and Prior Utilization. *Health Care Financing Review* 10(3):41-54, Spring 1989.
- Public Health Service and Health Care Financing Administration: International Classification of Diseases, 9th Revision, Clinical Modification. U.S. Department of Health and Human Services. Washington, DC. U.S. Government Printing Office, September 1980.
- Robinson, J.C., Luft, H.S., Gardner, L.B., and Morrison, E.M.: A Method for Risk-Adjusting Employer Contributions to Competing Health Insurance Plans. *Inquiry* 28(2):107-116, Summer 1991.
- Smith, N.S., and Weiner, J.P.: Applying Population-Based Case Mix Adjustment in Managed Care: The Johns Hopkins Ambulatory Care Group System. *Managed Care Quarterly* 2(3):21-34, 1994.
- Weiner, J.P., Starfield, B.H., Steinwachs, D.M., and Mumford, L.M.: Development and Application of a Population-Oriented Measure of Ambulatory Care Case-Mix. *Medical Care* 29(5):452-472, May 1991.
- Weiner, J.P., Dobson, A., Maxwell, S.L., et al.: Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses. *Health Care Financing Review* 17(3):77-100, Spring 1996.

---

Reprint Requests: Grace M. Carter, RAND, 1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138. E-mail: Grace\_Carter@rand.org

**Table A**  
**Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
<b>Blood</b>					
Non-Deficiency Anemia	High	10,662	17.35	8,875	19.46
Other Blood Disorder	High	5,601	7.08	16,721	30.36
Other Blood Disorder	Low or Medium	1,796	6.48	7,335	31.23
Deficiency Anemia	Medium or High	1,281	3.15	4,198	12.14
Non-Deficiency Anemia	Low or Medium	NA	NA	4,242	6.85
Deficiency Anemia	Low	NA	NA	1,171	6.55
<b>Neoplasm</b>					
Lung Cancer	Any	22,867	21.53	14,120	18.90
Metastatic Cancer	Any	18,522	27.38	34,898	85.70
Breast Cancer	Medium or High	11,521	6.52	13,318	11.55
Hematological-Lymphatic Cancer	High	11,489	19.89	14,515	32.96
Other Cancer	High	7,985	12.21	10,921	22.69
Breast Cancer	Low	5,436	11.69	5,450	14.17
Unspecified Neoplasm	Medium or High	4,676	8.40	3,188	5.78
Colorectal Cancer	Any	3,257	2.94	12,888	16.15
Other Cancer	Low or Medium	2,627	5.87	6,072	17.85
Benign Neoplasm	Medium or High	NA	NA	3,886	17.79
Cancer in Situ	Any	NA	NA	2,721	5.65
Hematological-Lymphatic Cancer	Low or Medium	NA	NA	1,793	2.94
Benign Neoplasm	Low	NA	NA	1,353	6.36
Unspecified Neoplasm	Low	NA	NA	810	2.60
<b>Circulatory</b>					
Congestive Heart Failure, Age Under 18	Any	15,749	12.59	42,644	44.32
Congestive Heart Failure, Age 18 or Over	Any	9,041	19.46	8,499	17.59
Cerebrovascular	High	6,669	10.31	21,270	38.31
Peripheral Vascular Disease	Medium or High	4,801	5.42	4,969	9.75
Ischemic Heart Disease	High	4,620	3.81	23,174	25.26
Ischemic Heart Disease	Medium	4,215	16.67	12,296	63.16
Hypertension	High	3,733	6.84	1,687	7.75
Thrombophlebitis/Deep Vein Thrombosis	Medium or High	2,801	6.41	6,343	17.81
Ischemic Heart Disease	Low	2,561	7.84	2,614	10.20
Arteriosclerosis	Any	2,369	1.75	NA	NA
Hypertension	Medium	2,236	7.53	1,687	7.75
Varicose Veins	Any	2,049	3.89	3,324	7.61
Hypertension	Low	994	7.43	576	5.40
Conduction/Rhythm	High	NA	NA	31,050	49.14
Cerebral Degeneration	Any	NA	NA	18,928	30.74
Other Cardiovascular Disease	High	NA	NA	17,831	43.79
Other Heart Disease	Medium or High	NA	NA	12,676	31.86
Conduction/Rhythm, Age 18 or Under	Medium	NA	NA	5,680	8.66
Cerebrovascular	Low or Medium	NA	NA	4,202	10.19
Conduction/Rhythm, Age Over 18	Medium	NA	NA	3,338	11.02
Thrombophlebitis/Deep Vein Thrombosis	Low	NA	NA	3,124	5.48
Other Cardiovascular Disease	Medium	NA	NA	3,051	5.99
Other Cardiovascular Disease	Low	NA	NA	2,061	2.58
Other Heart Disease	Low	NA	NA	1,742	8.02
Conduction/Rhythm, Age 18 or Under	Low	NA	NA	822	3.39
Conduction/Rhythm, Age Over 18	Low	NA	NA	576	5.40
<b>Digestive</b>					
Other Gastrointestinal Tract Disorder	High	11,016	13.93	35,103	50.94
Abdominal Pain	High	7,218	13.48	11,133	26.50
Lower Gastrointestinal Tract Problem	High	5,760	10.38	3,867	7.77
Liver Disorder	Any	4,882	8.76	NA	NA
Liver Disorder	Medium or High	NA	NA	13,251	24.71
Liver Disorder	Low	NA	NA	7,894	10.09
Non-Ulcer Peptic Disease	High	3,464	7.42	2,573	5.01
Peptic Ulcer	Medium or High	3,136	8.21	7,488	23.55
Gall Bladder, Biliary Disease	Medium or High	3,042	4.19	8,843	16.01

See footnotes at end of table.

**Table A—Continued**  
**Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
Gastrointestinal Tract Hemorrhage	Medium or High	2,098	4.20	5,705	16.49
Lower Gastrointestinal Tract Problem	Medium	2,074	5.00	3,164	8.83
Gastrointestinal Tract Hemorrhage	Low	1,658	2.48	1,440	2.63
Abdominal Pain	Medium	1,495	10.35	1,085	8.70
Non-Ulcer Peptic Disease	Low or Medium	1,005	8.21	950	8.30
Abdominal Pain	Low	848	8.40	563	6.88
Gall Bladder, Biliary Disease	Low	623	2.34	5,837	32.64
Other Gastrointestinal Tract Disease	Medium	NA	NA	9,380	37.77
Hepatitis	Medium or High	NA	NA	5,519	9.28
Inguinal Abdominal Hernia	Medium or High	NA	NA	4,239	8.65
Pancreatic Disease	Any	NA	NA	3,954	8.48
Other Gastrointestinal Tract Disease	Low	NA	NA	3,816	22.51
Inguinal Abdominal Hernia	Low	NA	NA	3,236	18.29
Rectal/Anal Conditions	High	NA	NA	2,997	6.54
Hepatitis	Low	NA	NA	2,448	5.31
Hemorrhoids	Any	NA	NA	1,440	4.06
Peptic Ulcer	Low	NA	NA	1,416	4.94
Rectal/Anal Conditions	Low or Medium	NA	NA	854	3.24
Lower Gastrointestinal Tract Problem	Low	NA	NA	538	2.86
Oral or Dental Disease	Any	NA	NA	408	3.16
<b>Infections</b>					
Human Immunodeficiency Virus	Any	15,099	22.48	10,192	23.65
Herpes Zoster	Any	1,568	2.51	NA	NA
Other Venereal Disease	Medium or High	790	5.17	639	4.65
Selected Infectious Disease	Medium or High	741	3.45	NA	NA
Selected Infectious Disease	High	NA	NA	4,651	15.23
Selected Infectious Disease	Medium	NA	NA	1,135	5.08
Selected Infectious Disease	Low	390	2.33	807	5.91
Septicemia	High	NA	NA	26,610	48.71
Septicemia	Low or Medium	NA	NA	9,734	39.20
Gonococcal Infection	Medium or High	NA	NA	2,742	2.72
Syphilis	High	NA	NA	2,740	2.32
Fever	Any	NA	NA	735	8.36
Flu/Virus	Any	NA	NA	336	3.50
<b>Injury</b>					
Internal Traumatic Injury	Medium or High	NA	NA	18,741	38.17
Internal Traumatic Injury	Low	NA	NA	4,560	9.62
Fracture	Any	284	3.05	NA	NA
Fracture	High	NA	NA	10,009	32.49
Fracture	Medium	NA	NA	2,107	17.34
Fracture	Low	NA	NA	882	9.87
Burn, Age 18 or Under	High	NA	NA	4,063	8.67
Burn, Age Over 18	High	NA	NA	2,512	3.78
Burn	Low or Medium	NA	NA	787	5.39
Vehicle Accident	Any	NA	NA	2,505	4.50
Other Accident	Any	NA	NA	3,074	16.49
Joint Dislocation	High	NA	NA	2,437	8.10
Joint Dislocation	Low or Medium	NA	NA	595	5.09
Sprain or Strain	Any	658	9.60	NA	NA
Sprain or Strain	High	NA	NA	3,702	11.39
Sprain or Strain	Medium	NA	NA	2,249	9.13
Sprain or Strain	Low	NA	NA	299	4.84
Wound or Injury	High	NA	NA	1,615	10.04
Wound or Injury	Low or Medium	NA	NA	351	7.56
Head Trauma	Medium or High	NA	NA	1,393	7.93
Head Trauma	Low	NA	NA	504	2.23
Adverse Effects of Medication	Any	NA	NA	607	2.28
Poisoning/Toxic Effect	Any	NA	NA	498	4.30
Superficial Injury or Contusion	High	1,038	2.56	NA	NA
Superficial Injury or Contusion	Low or Medium	380	5.17	NA	NA

See footnotes at end of table.

**Table A—Continued**  
**Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
<b>Mental</b>					
Mental Retardation	High	19,093	18.88	18,307	30.75
Mental Retardation	Medium	16,726	17.30	15,784	26.18
Mental Retardation	Low	10,111	13.56	7,022	12.89
Bipolar Disorder	Any	4,308	16.88	5,609	24.39
Schizophrenia	Medium or High	3,557	4.49	8,683	17.10
Psychosis/Major Depression	High	3,426	12.71	7,259	39.63
Psychosomatic Disorder	High	2,805	2.97	2,128	5.37
Alcohol Use Disorder	Medium or High	2,672	6.26	3,351	9.66
Psychosis/Major Depression	Medium	2,380	11.28	3,121	22.69
Psychosis/Major Depression	Low	2,053	7.26	3,121	22.69
Substance Use Disorder	Medium or High	1,884	5.58	NA	NA
Substance Use Disorder	High	NA	NA	3,958	6.59
Substance Use Disorder	Medium	NA	NA	2,196	8.69
Depression	Medium or High	1,751	3.81	NA	NA
Depression	High	NA	NA	9,899	14.87
Depression	Medium	NA	NA	1,775	4.11
Dementia/Delirium	Any	1,739	2.30	8,708	17.41
Non-Adult Psychiatric Disorder	Medium or High	1,680	9.21	NA	NA
Non-Adult Psychiatric Disorder	High	NA	NA	2,594	3.04
Non-Adult Psychiatric Disorder	Medium	NA	NA	955	6.86
Anxiety Disorder	Any	1,661	8.26	1,432	8.19
Other Psychiatric Disorder	Any	1,486	11.48	NA	NA
Other Psychiatric Disorder	Medium or High	NA	NA	2,415	16.00
Other Psychiatric Disorder	Low	NA	NA	1,218	10.08
Alcohol Use Disorder	Low	1,363	4.98	1,658	8.63
Depression	Low	1,073	6.71	1,731	14.73
Other Psychiatric Disorders	Any	NA	NA	11,808	23.01
Schizophrenia	Low	NA	NA	5,215	10.77
Sleep Disorder	Medium or High	NA	NA	3,730	8.92
Tobacco Use	Any	NA	NA	2,749	3.94
Suicide/Self-Inflicted Injury	Any	NA	NA	2,604	3.21
Personality Disorder	Any	NA	NA	1,739	5.32
Substance Use Disorder	Low	NA	NA	1,300	5.79
Non-Adult Psychiatric Disorder	Low	NA	NA	638	3.55
<b>ENMDD</b>					
Selected ENMDD	High	12,359	15.04	7,555	13.41
Immunological Disorder <sup>1</sup>	Any	9,919	12.71	15,549	27.20
Diabetes <sup>1</sup>	High	6,358	16.25	5,162	6.45
Malnutrition, Age 18 or Over	Any	6,256	7.40	19,115	27.20
Selected ENMDD	Medium	4,278	7.98	5,494	13.41
Diabetes	Medium	3,952	18.53	1,750	10.75
Obesity	Any	1,805	4.48	1,166	3.54
Diabetes	Low	1,547	7.99	634	4.11
Hypoglycemia	Any	1,439	2.59	NA	NA
Fluid/Electrolyte Abnormality	Any	1,343	8.42	3,464	28.49
Thyroid Disease	Medium or High	1,281	3.70	1,107	3.80
Vitamin/Mineral Disorder	Any	NA	NA	4,026	6.45
Lipid/Cholesterol Problem	Medium or High	NA	NA	2,147	7.26
Thyroid Disease	Low	NA	NA	487	2.13
<b>Musculoskeletal</b>					
Osteoarthritis	High	7,777	8.53	7,555	13.67
Rheumatoid Arthritis	Medium or High	7,320	10.49	2,692	3.31
Selected Musculoskeletal Disorders	Medium or High	6,022	10.33	12,709	31.01
Osteoarthritis	Medium	3,612	6.62	3,367	15.93
Osteoarthritis	Low	3,559	11.28	3,367	15.93
Rheumatoid Arthritis	Low	2,742	5.81	1,477	3.66
Other Arthritic or Collagen Vascular Disorders	Medium or High	2,367	5.68	3,077	8.62
Osteoporosis	Any	2,340	1.69	2,206	1.95
Low Back Pain	Medium or High	2,168	12.23	NA	NA
Low Back Pain	High	NA	NA	8,705	16.66

See footnotes at end of table.



**Table A—Continued**  
**Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
Low Back Pain	Medium	NA	NA	3,587	24.61
Neck Problem	Medium or High	1,616	5.67	NA	NA
Neck Problem	High	NA	NA	7,345	9.84
Neck Problem	Medium	NA	NA	2,981	11.99
Selected Musculoskeletal Disorders	Low	1,343	7.00	1,680	10.67
Other Joint or Disc Disorder	Any	1,089	9.55	NA	NA
Other Joint or Disc Disorder	High	NA	NA	4,997	9.89
Other Joint or Disc Disorder	Medium	NA	NA	2,183	10.42
Other Joint or Disc Disorder	Low	NA	NA	1,129	10.33
Neck Problem	Low	1,010	3.99	1,072	4.94
Muscle Disorder	Any	952	5.57	859	6.04
Bursitis/Synovitis	Any	858	6.62	NA	NA
Bursitis/Synovitis	High	NA	NA	1,921	3.32
Bursitis/Synovitis	Low or Medium	NA	NA	664	6.03
Low Back Pain	Low	796	5.18	702	5.53
Bunion	Medium or High	NA	NA	3,047	6.23
Bunion	Low	NA	NA	2,348	3.31
Hammertoe	Any	NA	NA	1,827	3.35
Scoliosis	Any	NA	NA	1,754	4.36
Other Arthritic or Collagen Vascular Disorders	Low	NA	NA	1,553	1.80
<b>Neurologic</b>					
Selected Neurological Disorders	High	8,995	28.99	10,970	46.83
Selected Neurological Disorders	Medium	3,482	6.25	5,097	11.71
Cataract/Aphakia	Any	3,133	7.23	5,007	13.98
Seizure Disorder	Medium or High	1,792	4.16	NA	NA
Seizure Disorder	High	NA	NA	5,947	9.25
Seizure Disorder	Medium	NA	NA	4,729	13.05
Other Eye Problems	High	1,762	6.20	13,207	59.52
Headache	Medium or High	1,643	10.10	902	6.59
Visual Loss	Medium or High	1,574	2.80	NA	NA
Glaucoma	Any	1,469	4.27	NA	NA
Other Eye Problems	Medium	1,181	3.39	1,174	8.91
Seizure Disorder	Low	1,066	5.84	2,081	14.79
Peripheral Neuropathy	Medium or High	1,013	4.48	NA	NA
Peripheral Neuropathy	Any	NA	NA	1,388	9.00
Headache	Low	800	5.68	417	3.69
Other Eye Problems	Low	688	3.56	1,174	8.91
Ear Problem Except Hearing Loss	High	NA	NA	3,028	5.44
Ear Problem Except Hearing Loss	Medium	NA	NA	1,361	5.23
Ear Problem Except Hearing Loss	Low	NA	NA	299	2.90
Selected Neurological Disorders	Low	NA	NA	2,405	4.80
Hearing Loss	Any	NA	NA	1,477	6.56
<b>Other</b>					
Post-Therapy Complications	Medium or High	1,733	8.18	10,166	60.80
Laboratory/Pathology/X-ray Abnormality	Any	803	5.55	1,224	11.30
Nose Deformity	Any	NA	NA	6,417	5.31
Post-Therapy Complications	Low	NA	NA	3,417	16.16
<b>Genitourinary</b>					
Renal Failure	Any	26,024	35.35	36,132	19.90
Urinary Tract Infection	High	5,439	7.13	1,359	7.10
Other Kidney Disease	Any	3,749	8.89	NA	NA
Other Kidney Disease, Age 1 or Under	Any	NA	NA	8,615	11.66
Other Kidney Disease, Age Over 1	Medium or High	NA	NA	6,542	14.92
Other Kidney Disease, Age Over 1	Low	NA	NA	2,309	4.18
Other Male Genital Disorder	High	2,228	3.81	1,274	2.65
Benign Prostatic Hypertrophy	Any	2,094	3.66	1,642	3.46
Urinary Tract Infection	Medium	730	3.48	1,359	7.10
Other Female Genital Disorder	Any	668	6.4	NA	NA
Other Female Genital Disorder	Medium or High	NA	NA	1,958	17.83
Other Female Genital Disorder	Low	NA	NA	1,693	12.42

See footnotes at end of table.

**Table A—Continued**  
**Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
Urinary Tract Infection	Low	508	5.45	NA	NA
Prenatal Problem Eclampsia	High	NA	NA	5,528	10.34
Urinary Tract Stone	Any	NA	NA	3,976	15.15
Other Urinary Tract Disorder	Medium or High	NA	NA	3,851	21.20
Prenatal Problem Eclampsia	Medium	NA	NA	1,894	4.37
Prenatal Problem Eclampsia	Low	NA	NA	1,554	4.03
Menopausal Disorder	Medium or High	NA	NA	1,405	2.87
Non-Malignant Breast Disorder	Any	NA	NA	1,238	7.54
Menstrual Disorder	High	NA	NA	1,223	3.49
Family Planning, Infertility	Medium or High	NA	NA	1,057	9.65
Other Urinary Tract Disorder	Low	NA	NA	923	2.30
Other Male Genital Disorder	Low or Medium	NA	NA	718	3.72
Menstrual Disorder	Low or Medium	NA	NA	449	4.58
Vaginitis or Cervicitis or Vulvitis	Any	NA	NA	257	2.88
Family Planning, Infertility	Low	NA	NA	229	2.74
<b>Respiratory</b>					
Lower Respiratory Infection, Age 18 or Over	High	9,923	16.51	12,671	23.54
Selected Respiratory Diseases	High	4,746	10.97	32,306	113.38
COPD/Bronchitis/Emphysema	High	3,541	7.71	4,903	13.63
Asthma	Medium or High	1,909	12.18	NA	NA
Asthma	High	NA	NA	3,639	21.70
Asthma	Medium	NA	NA	1,997	10.62
COPD/Bronchitis/Emphysema	Medium	1,370	6.18	668	2.34
Selected Respiratory Diseases	Medium	1,252	7.09	2,493	10.62
Lower Respiratory Infection, Age 18 or Over	Medium	1,089	6.01	1,204	8.03
Asthma	Low	999	8.11	916	9.52
Sinusitis	Any	668	7.35	344	4.24
Acute Pharyngitis/Tonsillitis	Medium or High	425	5.37	NA	NA
Pleurisy/Pleural Effusion	Medium or High	NA	NA	18,706	31.57
Selected Respiratory Diseases	Low	NA	NA	1,829	7.81
Tonsils/Adenoids	Any	NA	NA	1,812	12.45
Lower Respiratory Infection, Age Under 18	Medium or High	NA	NA	983	11.90
Allergy, Hay Fever	Medium or High	NA	NA	847	2.17
Allergy, Hay Fever	Low	NA	NA	573	5.19
Otitis Media	Medium or High	NA	NA	436	5.99
COPD/Bronchitis/Emphysema	Low	NA	NA	424	3.88
<b>Skin</b>					
Chronic Skin Ulcer	Any	6,893	11.95	8,444	20.08
Selected Skin Disorders	Medium or High	938	3.52	174	1.80
Non-Fungal Skin Infection	Medium or High	585	4.28	NA	NA
Non-Fungal Skin Infection	High	NA	NA	2,475	8.62
Non-Fungal Skin Infection	Medium	NA	NA	539	4.25
Benign Skin Neoplasm	Medium or High	NA	NA	686	4.25
Acne	Any	NA	NA	486	3.38
Benign Skin Neoplasm	Low	NA	NA	335	1.90
<b>Episode</b>					
Birth	Any	5,526	66.43	3,485	34.68
<b>Pregnancy</b>					
Pregnancy	Any	229	3.63	NA	NA
Pregnancy	High	NA	NA	4,226	16.92
Pregnancy	Medium	NA	NA	2,427	23.80
Pregnancy	Low	NA	NA	1,417	15.37

See footnotes at end of table.

**Table A—Continued**  
**Clinical Variables in New-Baby and Retrospective Models**

Body System and Condition	Severity	New-Baby		Retrospective	
		Coefficient	t-Statistic	Coefficient	t-Statistic
<b>Newborn</b>					
Congenital Anomaly <sup>1</sup>	High	4,561	19.81	4,472	18.99
Congenital Anomaly <sup>1</sup>	Medium	2,211	8.36	2,586	9.69
Neonatal Disease Episode <sup>2</sup>	High	27,431	150.32	27,988	166.26
Congenital Anomaly <sup>2</sup>	High	36,878	93.04	14,394	33.70
Congenital Anomaly <sup>2</sup>	Medium	21,918	47.82	4,207	9.20

<sup>1</sup> Infant born in previous year.

<sup>2</sup> Infant born in prediction year.

NOTES: NA is not applicable. Based on sample size of 359,692 persons in each regression.  $R^2$  values are 0.175 and 0.370, respectively. Regressions fit by weighted least squares with weights proportional to length of time in the sample for persons who died or were born during the year. ENMDD is endocrine, nutritional, metabolic diseases and immunity disorders. COPD is chronic obstructive pulmonary disease.

SOURCE: Carter et al., Santa Monica, California, 1999.