Case Selection for a Medicaid Chronic Care Management Program

Sharada Weir, D.Phil., Gideon Aweh, M.S., and Robin E. Clark, Ph.D.

Medicaid agencies are beginning to turn to care management to reduce costs and improve health care quality. One challenge is selecting members at risk of costly, preventable service utilization. Using claims data from the State of Vermont, we compare the ability of three pre-existing health risk predictive models to predict the top 10 percent of members with chronic conditions: Chronic Illness and Disability Payment System (CDPS), Diagnostic Cost Groups (DCG), and Adjusted Clinical Groups Predictive Model™ (ACG-PM™). We find that the ACG-PM™ model performs best. However, for predicting the very highest-cost members (e.g., the 99th percentile), the DCG model is preferred.

INTRODUCTION

As health care costs rise for all payers, Medicaid programs are faced with the dual challenges of providing appropriate health care for some of society's most vulnerable members and containing costs. Direct controls, such as limiting eligibility for programs or access to needed services, are unlikely to produce net budgetary relief, since public funds ultimately may be tapped for costly urgent care for those who cannot get treatment by other means. Targeting high-cost patients for enhanced care management is another approach to defending limited funds against increasing health care costs. The care management approach seeks to engage consumers in self-management of chronic conditions, while supporting physician adherence to evidence-based care guidelines in an effort to reduce the demand for future high cost, preventable utilization. Cost savings are not guaranteed, but the potential to simultaneously reduce costs and improve health outcomes is appealing.

As part of Act 191, Vermont's health care reform legislation, OVHA, which manages their publicly-funded health insurance programs, established a pair of integrated programs to offer care management services to all chronically ill members in the State who meet qualifying criteria (Maxwell, 2007; Office of Vermont Health Access, 2007a). In general, members must have been diagnosed with one or more of 11 chronic conditions[1]: (1) asthma, (2) diabetes, (3) chronic obstructive pulmonary disease, (4) low back pain, (5) congestive heart failure, (6) ischemic heart disease, (7) rheumatoid arthritis, (8) hypertension, (9) disorders of lipid metabolism, (10) depression, and (11) chronic renal failure.[2] Members with Medicare or other major third party insurance are ineligible, as are those in pharmacy-only benefit programs and programs that are paid from non-Medicaid funding sources.

The authors are with the Center for Health Policy and Research, University of Massachusetts Medical School. The research in this article was supported by the Office of Vermont Health Access (OVHA) under Contract Number 10733. The statements expressed in this article are those of the authors and do not necessarily reflect the views or policies of the Center for Health Policy and Research, University of Massachusetts Medical School, OVHA, or the Centers for Medicare & Medicaid Services (CMS).

[1] Members without one of the selected chronic conditions may gain access to the program through provider referral.

[2] Exploration of Vermont claims data determined that those with one or more of the selected chronic conditions were typically substantially more costly than others, overall and within specific diagnoses.

The Chronic Care Management Program (CCMP) serves the majority of the State's Medicaid beneficiaries with one or more chronic conditions. Member-level interventions range from mailings of condition-specific self-management literature and telephone access to health coaches to the development of an individual care plan and face-to-face patient support, depending on the needs of the member (Office of Vermont Health Access, 2008). The Care Coordination Program (CCP) is an intensive case management program staffed by teams of nurses and social workers and funded to serve the top 5 to 10 percent of eligible members with chronic conditions.

One challenge for the CCP is predicting which members are at greatest risk for costly and preventable service utilization. In the initial phase of program development, cases were selected based on prior cost and health care utilization (i.e., emergency department and acute inpatient hospitalization). Systems were put in place to facilitate referrals from hospitals and emergency departments, and cases could also be referred from outside sources (e.g., primary care physicians) (Office of Vermont Health Access, 2007b). Adoption of these recruitment techniques was expedient, helping the program get off the ground. However, as the CCP develops and matures, considerations of efficiency and equity become increasingly important. Although prior cost and utilization can help to predict future resource use, by themselves they are somewhat crude indicators in that they do not distinguish between transitory and chronic needs. Prediction of future cost can be greatly enhanced by taking into account prior diagnoses.

Health risk predictive modeling is a methodological approach that uses clinical diagnostic information to predict future cost. While there will always be room for referral systems in a program such as the CCP, basing initial case selection on health risk predictive modeling output enables the program to efficiently evaluate all eligible members. It may help to identify at-risk patients who do not have effective physician advocates and provide timely assessment of risk before patients require emergency department or inpatient care.

Predictive models can be customized for a particular insurer or Medicaid Program, implicitly incorporating the effects of benefits, practice patterns, and characteristics of the population served. Customization is costly and beyond the capacity of Medicaid budgets in many States; models developed in other States can be adopted at much lower cost. However, selecting a model can prove challenging, as the best model for the intended use generally is not obvious.

The purpose of this study was to help the State of Vermont choose a predictive-modeling software tool that would select members with one or more chronic illnesses who are most likely to need and benefit from participation in an intensive care management program. As a small State without the population necessary to create a custom model, Vermont needed to find the best model from among the readily available alternatives. Criteria for the selected tool included the ability to predict future costs with reasonable accuracy, cost effectiveness, and a clinically-meaningful condition classification system.

We used 2 years of historical claims data to evaluate the potential of the diagnosis-based health risk predictive modeling approach to select high cost cases for the CCP. Three models were chosen for testing based on their methodological attributes and history of use with Medicaid populations: The CDPS Version 2.5 (University of California, 2008); DCG RiskSmart™ Version 2.2 (DxCG, 2008); and ACG-PM™ Version 7.0 (Johns Hopkins University, 2008). We examined and compared the

potential of each pre-existing model to predict high cost Medicaid members in Vermont. Given scarce public resources, appropriate case selection is crucial.

MODELS

Well-established tools are available to estimate current illness burden from administrative claims data and predict future overall resource use. Health risk predictive models use demographics, diagnoses, prior cost and service utilization, or some combination of these data to predict cost and stratify the population into low, medium, and high-risk cases relative to the average health risk of the population. Statistical techniques distinguish between diagnoses that are likely to be costly in the future versus those that are transitory in nature.

Model properties vary. Models developed using data mining techniques, such as neural networks or artificial intelligence, were not considered for this study. Such models typically perform well, but may be "overfit" to the development population, performing less well when applied to other datasets. They are not methodologically transparent and often fail to produce clinically-meaningful findings (Hartnell and MacKinnon, 2003; Martin, Rogal, and Arnold, 2004). Models based solely on pharmacy data, such as MedicaidRx (Gilmer et al., 2001; University of California, 2007), are also excluded from consideration here. Although these models have the advantage of utilizing timely pharmacy claims data, which are not subject to the long run out period common with other types of claims (e.g., inpatient facility claims), their clinical classifications are less precisely defined, by nature, and they perform less well overall than models that include diagnoses from claims (Zhao et al., 2005). We have also excluded models that rely on data on prior health care service utilization and procedures (e.g., surgery), owing to data limitations, though such information would be expected to improve model performance.

We chose to focus on readily-available models developed using clinically-meaningful and transparent methods for Medicaid populations and utilizing *International Classification of Disease, 9th Edition, Clinical Modification* (ICD-9-CM) diagnosis data from claims. All three models chosen for comparison were developed for Medicaid populations and are being used in one or more States currently; none were developed with Vermont data. We compare model performance in selecting high cost cases in Vermont.

ACG-PM™

The ACG® Case-Mix System was developed at Johns Hopkins University using health care claims data from a combination of commercial managed care and Medicaid populations (Smith and Weiner, 1994; Weiner et al., 1996; Weiner et al., 1998). The ACG-PM™ uses demographics, diagnoses, and pharmacy expenditures to predict future cost. First, patients are assigned to ACG® actuarial cells, which are mutually exclusive patient clusters defined by age, sex, and combinations of diagnoses. Resource utilization scores are assigned to each cluster based on the mean cost of patients in that cell in the development dataset. In the second stage, risk scores are created using linear regression of future cost as a function of ACG® resource utilization score plus indicators of frailty, hospital dominant conditions, specific chronic conditions, and optionally, prior pharmacy spending (Weiner, 2005). This two-stage method, which allows the model to give higher weight to certain conditions and includes some prior expenditure data,

improves prediction compared with using ACG® resource utilization scores alone to forecast future resource use. Although ACG® cells contain broad groupings of disease, the software produces for each member a complete list of conditions, grouped into clinically-meaningful categories. A version of the ACG® software is available without a fee to State Medicaid agencies. For information on ordering: see http://www.acg.jhsph.edu/html/OrderACGProducts_Medicaid.htm#cons.

CDPS

The CDPS model was developed at the University of California, San Diego, and is made available at http://cdps.ucsd.edu/ (Kronick and Dreyfus, 1996; Kronick et al., 2000; Chronic Illness and Disability Payment System, 2005). The CDPS uses only diagnoses and demographic data to predict cost. The model was developed using linear regression under the assumption that the effects of different categories of diagnoses on costs are additive. That is, an individual's risk score generally increases with each additional separate condition identified. The most recent version of the model was developed with Medicaid claims from a group of mainly Midwestern States (California, Colorado, Georgia, Michigan, Missouri, Ohio, and Tennessee) covering the period from the early to middle 1990s. The CDPS disease classification system focuses on chronic conditions; a subset of ICD-9-CM diagnosis codes is grouped into condition categories. A potential advantage of CDPS is that its classification system was developed specifically for Medicaid populations, so particular emphasis was placed on classifying mental health and substance use disorders. Separate risk scores are produced for Temporary Assistance for Needy Families (TANF) and Supplemental Security Income (SSI) beneficiaries. The model also distinguishes between adults and children in creating risk scores.

DCG

The DCG Hierarchical Condition Category (HCC) approach was developed by researchers from Boston University who were part of the team that created the CMS-HCC model to adjust Medicare capitation payments (Pope et al., 2004). DxCG, Inc. offers an annual license arrangement for it's RiskSmart™ software which assesses illness burden and predicts future cost using the DCG methodology (http://www.dxcg.com/). Like CDPS, DCG models are additive across condition categories with cost predictions developed using linear regression. The DCG model includes several disease interaction terms and child-specific weights for certain conditions. DxCG's Medicaid model was developed using Massachusetts data and was most recently updated in 2007. Separate risk scores are available for Medicaid fee-for-service and managed care applications. Nearly all ICD-9-CM diagnoses are grouped into detailed sets of clinically meaningful disease and condition categories. Although the DCG classification system is sometimes criticized for its handling of behavioral health conditions, one recent study predicting mental health and substance abuse cost in a Veterans Affairs' population found slightly better performance for DCG models compared with CDPS (Sloan et al., 2006).

In recent research by Winkelman and Mehmud (2007) the DCG and ACG-PM™ commercial prospective models were found to perform similarly in terms of R^2 (ACG-PM™: 18.7 percent; ACG-PM™ without prior pharmacy cost: 16.2 percent; DCG: 17.4 percent) and mean absolute prediction error (MAPE) as a percentage of

total actual costs (ACG-PM™: 85.6 percent; ACG-PM™ without prior pharmacy cost: 90.4 percent; DCG: 88.0 percent). Both models outperform CDPS' Medicaid model with TANF cost weights (R^2 = 12.4 percent; MAPE = 95.8 percent). However, it is impossible to know whether these results, which used data from a commercially-insured study population and compared overall model performance, are applicable to Vermont's Medicaid population and the task of selecting high cost cases.

METHODS

We used 2 years of prior Vermont Medicaid claims data to test the ability of each model to correctly identify future high-cost members. Diagnoses recorded during State fiscal year (SFY) 2005 (July 2004–June 2005) were used to predict cost in SFY 2006 (July 2005-June 2006).

Data for this study were supplied by the OVHA, which granted the authors access to six distinct data sets linkable using unique member and claim identifiers: (1) professional claims (e.g., emergency room, office visits, and other outpatient services); (2) institutional claims (e.g., hospital and nursing home utilization); (3) pharmacy claims; (4) eligibility data (e.g., demographic information, dates of eligibility, eligibility category, and type of coverage); (5) third party insurance coverage data; and (6) provider data.

After excluding members with fewer than 30 days of Medicaid eligibility and those with Medicare or other major insurance or with pharmacy-only benefits, we were left with data on 126,616 members in SFY 2005 and 121,059 members in SFY 2006, or approximately 80 percent of Vermont's Medicaid population. Keeping only those with at least 11 months of Medicaid eligibility in both years leaves 59,384 members. Of these, 16,708 had one or more of the qualifying chronic conditions.

Winkelman and Mehmud (2007) compared predictive modeling tools using both the offered models with out-of-the-box cost weights, which reflect the relationship between illness burden and cost from the model development population, and customized models with cost weights that were recalibrated to the study's commercially-insured population dataset. The customized models were modified to include prior cost data; all outperformed the offered models. Prospective R^2 results for each of the three customized models were nearly indistinguishable (ACG-PM™: 22.1 percent; CDPS: 21.2 percent; and DCG: 22.9 percent). Unfortunately, we were not able to recalibrate the three models to fit Vermont Medicaid data owing to the small population of members with chronic conditions. Model calibration is usually performed with populations in the hundreds of thousands to 1 million or more members.

Running each software product produced a set of prospective risk scores for the study population using the model's out-of-the-box cost weights. Since all three models were developed on beneficiaries from other States, risk scores were renormalized to the mean of the study population for ease of interpretation. The population norm is 1.0, by definition. A score below 1.0 indicates that the individual is expected to be less costly than average. A score in excess of 1.0 indicates a higher-than-average health care cost risk (e.g., a score of 2.0 means that the person is predicted to be twice as costly as the population average).

We stratified the population by risk score, assigning cases to the top 10 and top 1 percent risk groups for each model based on their relative risk scores. We then examined the top percentiles of actual year 2 cost and evaluated each model in terms

of its ability to correctly predict future high-cost members. We computed the sensitivity of the model to correctly pick those who were high cost, the specificity of the model to correctly exclude those who were not high cost, and the positive predictive value of the case selection tool.

We computed the sensitivity of each model, or the ability to correctly identify high cost members, defined as the number of true positives (cases selected based on year 1 data that turn out to be high cost in year 2) as a percentage of true positives plus false negatives (members not selected in year 1 that turn out to be high cost in year 2; also known as Type II error). By definition, sensitivity declines with increasing thresholds of risk, from a lower extreme where all members are selected and sensitivity is 100 percent to the upper extreme where no members are selected and sensitivity is 0.

We were also interested in the specificity of the model, or the ability to correctly exclude low-cost members, defined as the number of true negatives (members not selected in year 1 that are not high cost in year 2) as a percentage of true negatives plus false positives (cases selected in year 1 that are not high cost in year 2, also known as Type I error). Specificity increases as fewer cases are selected.

Sensitivity and specificity are traded off as progressively smaller groups of cases are selected. To evaluate the overall model performance on both tests, we produced receiver operating characteristic (ROC) curves (Figure 1) for each model. The ROC curve plots sensitivity against 1-Specificity. The best model will simultaneously exhibit the highest sensitivity and specificity combinations, bowing out from the 45° null line. Model performance was evaluated by calculating the area under the ROC curve for each model and comparing confidence intervals.

We also were interested in the positive predictive value (PPV), or the proportion of cases selected in year 1 (true + false positives) that are high cost in year 2 (true positives). This ratio is of practical value to care management program planners since it indicates the expected proportion of actual high-cost members among all selected cases for a given model. It can be used to determine the approximate number of appropriate cases that will be selected for care management for a given number of cases included in the intervention group or to determine the size of the intervention group needed to ensure that a certain number of appropriate cases will be identified.

Adapting the PPV test, we classified all cases picked by each model into four outcome categories, (1) low-cost, (2) moderate-cost, (3) high-cost and (4) very high-cost, depending on actual cost in year 2. This provided more information on the ability of the model to select the most appropriate cases for care management. We defined a low-cost pick as a case selected in year 1 that does not meet the 50th percentile of actual cost in year 2 ($4,061), since at that threshold, intervention costs may exceed potential savings.[3] A moderate-cost pick was one that ranged from the 50th to < 90th percentiles. A high-cost pick was one that ranged from the 90th ($24,597) to < 99th percentiles and a very high-cost pick met or exceeded the 99th percentile ($96,332).[4]

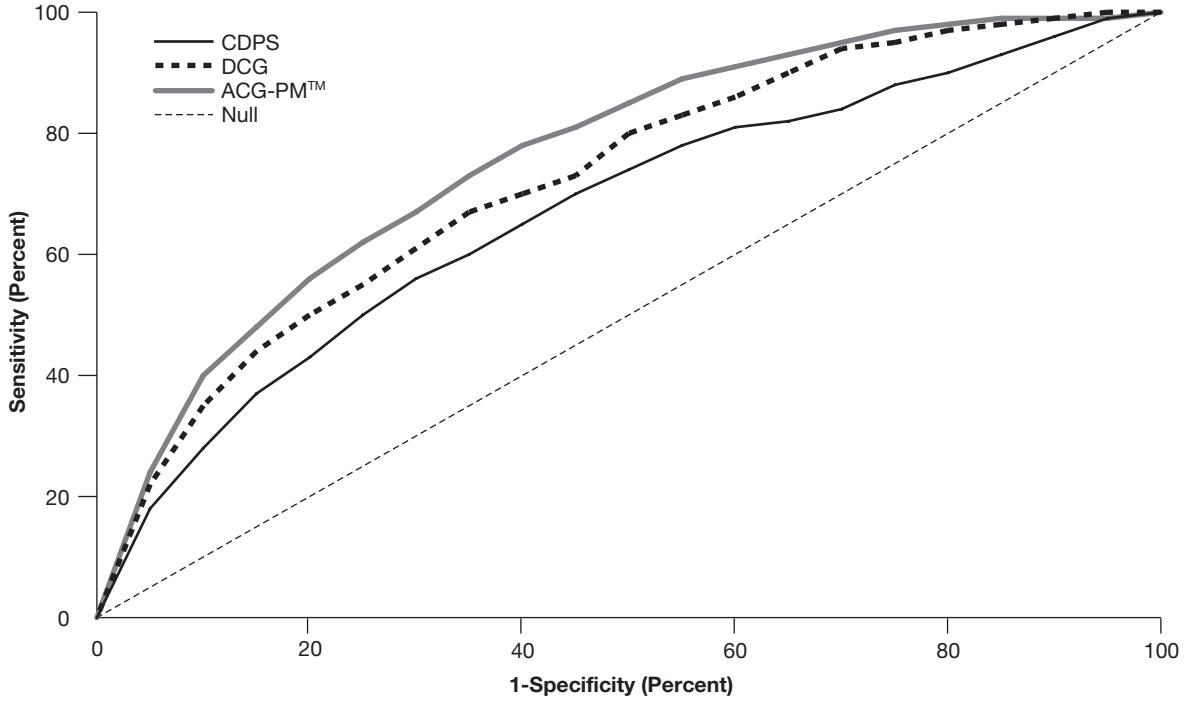
We evaluated the distribution of cases picked in the top 10 percent risk group, since that corresponds roughly to the number of members that Vermont budgeted to serve in its most intensive intervention, the CCP, and in the top 1 percent risk group,

[3] Actual program costs were not available at the time this article was prepared. Reported costs of multiple disease management programs vary widely (Goetzel et al., 2005).

[4] Categories could be adjusted to reflect different break-even points.

Figure 1

Evaluation of Model Performance Using the Receiver Operating Characteristics (ROC) Curve: 90th Percentile Cost Threshold



NOTES: CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™. The population was stratified by risk score and members assigned to the top 10 percent risk group for each model. After computing the sensitivity of each model, or the ability to correctly identify high-cost members, we also examined the specificity of the model, or the ability to correctly exclude low-cost members.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

since care management programs in States serving larger populations typically are forced to be more selective owing to budget constraints. Next, we examined the percentage of all year 2 high-cost members found in the top 10 and top 1 percent risk groups selected using year 1 data.

As a final practical test, we examined the performance of each model in terms of selecting cases who were at risk of future hospitalization. Avoiding preventable hospitalizations is often a key goal of care management, and members at risk of hospitalization may be particularly suitable candidates for intervention. We examined the percentage of cases selected by each model who had one or more inpatient admissions in year 2 and the percentage of all hospitalizations in the data that were accounted for by cases in the high risk groups.

RESULTS

Table 1 shows sensitivity, specificity, and positive predictive value of each of the models evaluated at the 90th percentile of year 2 actual cost. We find that the three models have similar performance on the sensitivity test at the lowest cutoff points, but model performance diverges as fewer cases are selected, with the ACG-PM™ model performing best. Performance on the specificity test is indistinguishable for the three models. The corresponding ROC curve suggests that the ACG-PM™

Table 1

Sensitivity, Specificity, and Positive Predictive Value Evaluated at the 90th Percentile of Cost, by Selected Risk Percentile and Model

Model	Risk Percentile						
	5th	10th	25th	50th	75th	90th	95th
Sensitivity							
CDPS	0.99	0.96	0.88	0.74	0.50	0.28	0.18
DCG	1.00	0.99	0.95	0.80	0.55	0.35	0.22
ACG-PM™	0.99	0.99	0.97	0.85	0.62	0.40	0.24
Specificity							
CDPS	0.05	0.10	0.26	0.53	0.78	0.92	0.96
DCG	0.05	0.11	0.27	0.53	0.78	0.93	0.97
ACG-PM™	0.00	0.11	0.27	0.54	0.79	0.93	0.97
Positive Predictive Value							
CDPS	0.03	0.07	0.08	0.09	0.12	0.21	0.35
DCG	0.01	0.02	0.03	0.14	0.09	0.26	0.44
ACG-PM™	0.01	0.01	0.04	0.08	0.13	0.32	0.47

NOTES: CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

Table 2

Area Under the Receiver Operating Characteristics (ROC) Curve, Evaluated at the 90th Percentile of Cost

Model	Area	95 Percent Confidence Intervals	
		Lower Bound	Upper Bound
CDPS	0.69	0.67	0.70
DCG	0.75	0.74	0.76
ACG-PM™	0.79	0.78	0.80

NOTES: The ROC curve plots sensitivity against 1-Specificity. CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

Table 3

Distribution of Picks[1] Selected in the Top 10 and Top 1 Percent Risk Groups, by Cost Threshold[2] and by Model

Model	Low-Cost (<50th Percentile)	Moderate-Cost (≥50th-<90th Percentile)	High-Cost (≥90th-<99th Percentile)	Very High-Cost (≥99th Percentile)
Top 10 Percent Risk Group				
CDPS	22.2	49.8	23.4	4.6
DCG	17.2	47.6	30.1	5.1
ACG-PM™	10.1	50.3	35.0	4.7
Top 1 Percent Risk Group				
CDPS	18.5	39.3	31.5	10.7
DCG	3.6	23.8	54.2	18.5
ACG-PM™	3.0	32.1	51.8	13.1

[1] Picks are members whose health risk score places them in the top risk group for each model.

[2] Cost thresholds are as follows: 50th percentile, $4,061; 90th $24,597; and 99th $96,332.

NOTES: CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

model is best, followed by the DCG model (Figure 1). This is confirmed in Table 2, which shows that the 95 percent confidence interval (CI) for the ACG-PM™ model excludes the 95 percent CI of the DCG model, which, in turn, excludes the 95 percent CI of the CDPS model. That is, the area under the ACG-PM™ model's ROC curve is significantly larger than the area under the DCG model's ROC curve, which in turn is significantly larger area than the area under the CDPS model's ROC curve. Areas under the ROC curves for CDPS and DCG both improve (CDPS: 0.73; DCG: 0.78) and the differences between all models become insignificant when evaluated at the top 1 percent of total year 2 cost.

The findings for PPV are less definitive. At the lower risk cutoff points, where many cases are selected and the rate of false positives is high for all models, CDPS performs best. At the highest cutoff points, where fewer cases are selected, the ACG-PM™ model has the highest PPV. Since, in practice, care management programs must choose relatively small percentages of members to serve, the lower range of risk percentiles is not as relevant and the ACG-PM™ model can be said to perform best overall. Differences between the models are less pronounced when evaluated at the top 1 percent of cost (not shown), though the DCG model appears to perform slightly better than the other models at the highest risk cutoff.

Taking the top 10 percent risk group (Table 3), we find that the ACG-PM™ model has the smallest percentage of low-cost picks (10.1 percent) and CDPS the highest (22.2 percent). The ACG-PM™ model also has the highest percentage of moderate and high-cost picks, but the DCG model produces the largest percentage of very high-cost picks. With the top 1 percent risk group, we find that the ACG-PM™ model again has the fewest low-cost picks and most moderate-cost picks and the DCG model has the highest number of high and very high-cost picks.

Examining the percentage of all high-cost members that are found in the top 10 percent risk group (Figure 2), we find that the ACG-PM™ model selects the highest percentage of all year 2 high-cost members when evaluated at the 50^{th} - 90^{th} percentile cost thresholds. The ACG-PM™ and DCG models pick similar percentages at the 95^{th} percentile, and the DCG model picks the highest percentage at the 99^{th} percentile. When the top 1 percent risk group is considered (Figure 3), the ACG-PM™ and DCG models perform similarly up to the 90^{th} percentile cost threshold, but the DCG model picks higher percentages of high-cost members at the 95^{th} and 99^{th} percentile cost thresholds.

Finally, we compared the ability of the models to identify members at risk of hospitalization. Not surprisingly, we found that the number of hospitalizations and total cost in year 2 are strongly correlated in our data (Pearson correlation coefficient = 0.4288; $p<0.0001$). Nevertheless, comparing performance in terms of predicting hospitalizations helps to further differentiate the models.

Table 4 shows the percentage of cases selected in each models' high-risk groups (top 10 and top 1 percent) with hospitalizations in year 2, the mean number of hospitalizations per person, and the percentage of all hospitalizations that were accounted for by cases in the top risk groups for each model. For both the top 10 and top 1 percent risk groups, we find that the ACG-PM™ model performs best, with the top 10 percent risk group accounting for 36 percent, and the top 1 percent risk group accounting for 7 percent, of all hospitalizations in the data. The DCG model outperforms CDPS for the top 10 percent risk group, with 27 versus 23 percent of

Table 4

Number and Percentage of Inpatient Admissions Identified in the Top 10 and Top 1 Percent Risk Groups, by Model

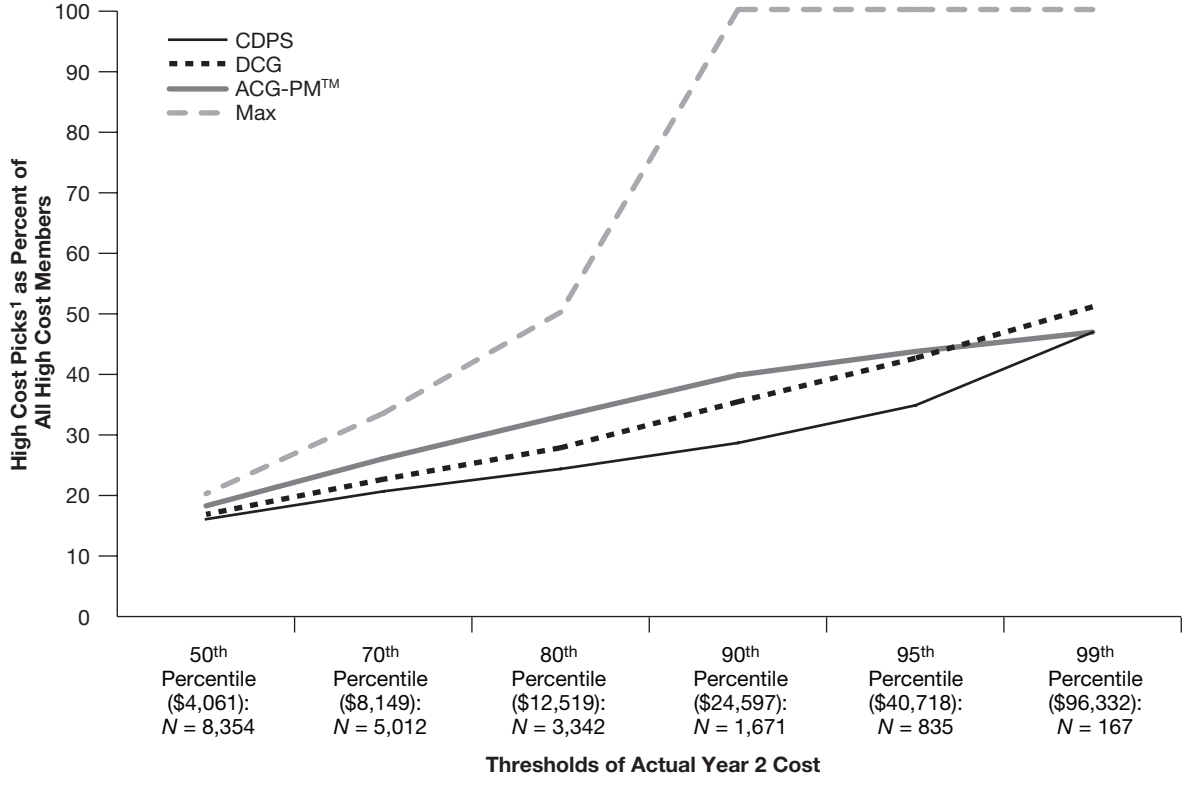
Model	*N*	Total Acute Inpatient Admissions in Year 2	Percent of High Risk Group with an Admission in Year 2	Mean Inpatient Admissions, Per Member with At Least 1 Admission	Percent of Inpatient Admissions Found in High Risk Group
Top 10 Percent Risk Group					
CDPS	1,697	700	20	2.1	23
DCG	1,671	829	24	2.0	27
ACG-PM™	1,671	1,115	32	2.1	36
Top 1 Percent Risk Group					
CDPS	168	122	28	2.6	4
DCG	168	116	31	2.2	4
ACG-PM™	168	223	51	2.6	7

NOTES: CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

Figure 2

Percentage of All High-Cost Members Found in Top 10 Percent Risk Group, by Actual Year 2 Cost Threshold



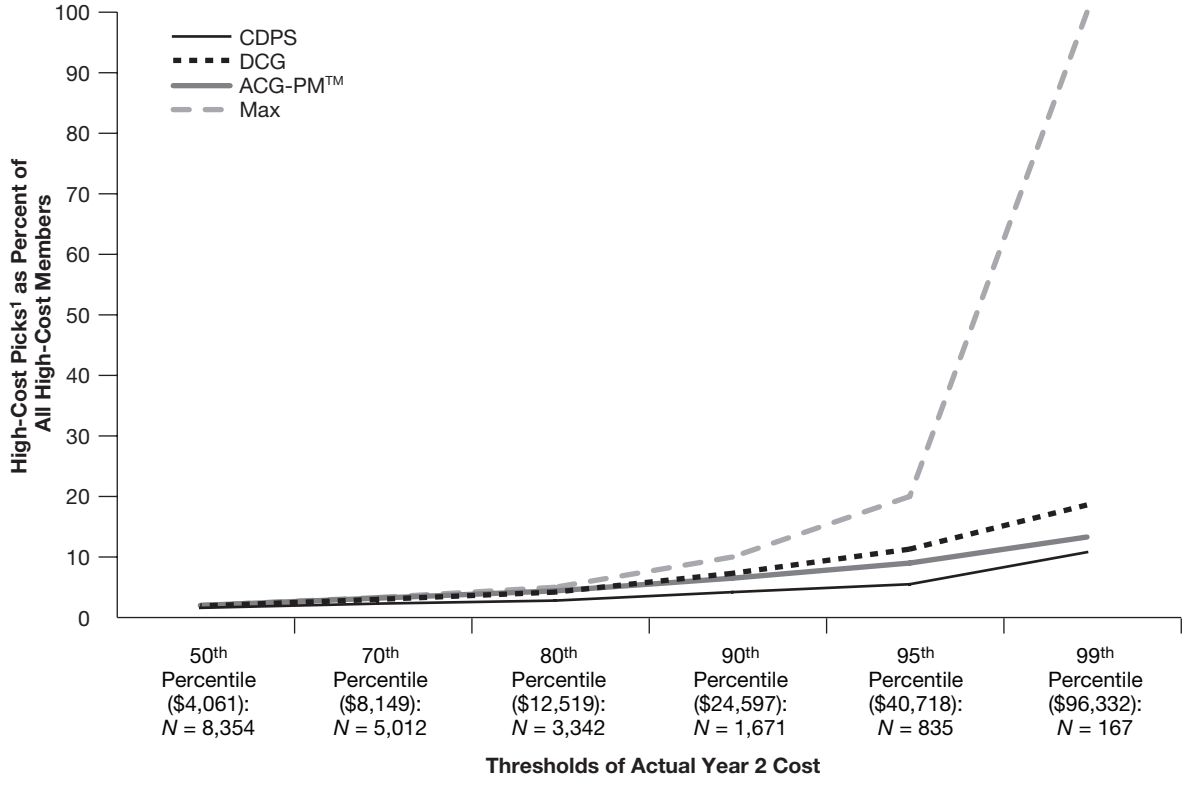
[1] High-cost picks are members in the top 10 percent risk group who were actually high cost in year 2 as evaluated at each cost threshold.

NOTES: CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

Figure 3

Percentage of All High-Cost Members Found in Top 10 Percent Risk Group, by Actual Year 2 Cost Threshold



[1] High-cost picks are members in the top 1 percent risk group who were actually high cost in year 2 as evaluated at each cost threshold.

NOTES: CDPS is Chronic Illness and Disability Payment System. DCG is Diagnostic Cost Groups. ACG-PM™ is Adjusted Clinical Groups, Predictive Model™.

SOURCE: Office of Vermont Health Access: Medicaid claims data from State fiscal years 2005-2006.

all hospitalizations, but the CDPS and DCG models perform equally well for the top 1 percent risk group, each containing 4 percent of all hospitalizations.

DISCUSSION

We find that the ACG-PM™ outperforms the other models in terms of sensitivity, specificity, and positive predictive value at the 90th percentile cost threshold, but not at the 99th percentile. We performed three additional practical tests. First, we classified cases selected in the top 10 percent risk group based on year 1 data into low, moderate, high, and very high-cost picks, depending on actual cost observed in year 2. Second, we determined the percentage of high-cost members found in the top 10 percent risk group for each model. The results showed that the ACG-PM™ model generally outperforms the other models through the 90th percentile of cost. However, the DCG model produces the highest percentage of high-cost and very high-cost picks and finds the largest percentages of high-cost members when evaluated at the 99th cost percentile. Finally, we compared the number of hospitalizations in the highest risk groups and found that

the ACG-PM™ model outperformed the other two models for both the top 10 and top 1 percent risk groups. Putting these results together, we conclude that the ACG-PM™ model is generally preferred. However, if the care management program were required to concentrate its efforts on the top 1 percent of members in terms of total expected future cost, the DCG model would be preferred.

Our findings tend to support prior research. On the one hand, regression-based models are more flexible and predict better than cell-based models (Rosen et al., 2001). The performance of the DCG model indicates the strength of the DCG classification system and regression methodology to predict the highest-cost members. On the other hand, while the diagnosis-based approach has been found to do a better job of finding future high-cost members and those with manageable chronic illness than simply looking at prior cost, methods that combine information both on diagnoses and prior cost tend to do best (Ash et al., 2001; Zhao et al., 2002; Zhao et al., 2003). This is supported by our findings that the ACG-PM™ model, which includes pharmacy expenditure data, tends to perform best overall.

Neither of the three models we tested were developed to predict hospitalizations. The superior performance of the ACG-PM™ model on this outcome may be explained by the inclusion of variables indicating the number of hospital dominant conditions, particular chronic conditions, and a frailty flag, since these indicators of future cost are also likely indicators of risk of future hospitalization.

Our mandate for choosing a particular health risk predictive modeling technology for Vermont included a combination of methodological, performance, and cost criteria. The CDPS model is available free to all and a version of the ACG-PM™ model is made available without a fee to State Medicaid agencies. However, to the extent that the DCG model does a better job of selecting high-cost cases, it will be a more fiscally prudent choice over time. Another consideration was the level of detail provided by the condition grouper. CDPS groups only a partial set of diagnoses. By contrast, the ACG® and, particularly, the DCG software packages provide richly detailed clinical classification systems.

It was determined that if either CDPS or ACG-PM™ outperformed DCG, the choice would be clear. If DCG was found to be better or more flexible, the choice of model would depend on the magnitude of the differences in performance, expected savings from care management, and the cost of the software license. Ultimately, the DCG model may be a cost-saving technology if found to be much more accurate in picking cases for care management. Based on our analyses and consideration of the circumstance that the CCP provides coverage for the top 10 percent of members, the State chose the ACG-PM™ model.

One limitation of this study is that we are only able to test currently available technology. The field of health risk predictive modeling is one that has expanded in recent years and we may expect new developments in the field, including both improvements to existing models and the entrance of new models, to continue to occur in the coming years.

Another study limitation is the possibility that our specific findings would not hold for other State Medicaid programs. Although health risk predictive models are generally robust, if a health system is sufficiently different from that of the development population, different results may be found. Nonetheless, our approach to the selection of a predictive modeling tool for case selection may be of general interest to other States facing the problem of selecting

cases for care management and choosing among the available alternatives. In particular, our finding that differences between the models are small when predicting the very highest cost members may be helpful to States facing budgetary constraints.

Our findings indicate that although most of the cases selected by each model exceeded median cost in year 2, none of the models is able to predict with great accuracy the highest cost members. At best, the health risk predictive models we tested are able to identify approximately one-half of the very highest cost members (i.e., those costing >$96,000 per year). This performance is in line with what is typically seen with prospective models in terms of predicting cost. A barrier to predictive accuracy in general is the volatile nature of serious injuries and illnesses that result in high cost health care encounters or episodes. A potential barrier to accuracy in this instance in particular is the lack of a customized model fit on a sample of Vermont's Medicaid population with chronic conditions. Whether this level of performance is adequate for the purposes of a care management program is a larger issue. Another question is whether the data produced by these models is sufficiently timely to meet program needs. Some evidence suggests that models using real-time hospital admissions data could produce substantial savings for Medicaid programs (Billings and Mijanovich, 2007).

The currency of hospital data may improve with increased adoption of electronic medical records; presently, many States experience a lag of several months before such data are reasonably complete. Future research will consider whether adoption of predictive modeling technology was an improvement over previous case selection methods which focused on prior cost and service utilization. Health risk predictive models offer opportunities for efficiency and equity, but lack the immediacy and individual detail—particularly information on social factors—that care managers may require.

ACKNOWLEDGMENTS

We appreciate the encouragement for this research provided by the OVHA and we particularly wish to thank Mary Day for her expertise and support in using the claims and eligibility data and for creating the ACG-PM™ output dataset for this study. We also would like to thank the three anonymous reviewers for their constructive comments.

REFERENCES

Ash, A.S., Zhao, Y., Ellis, R.P. et al.: Finding Future High-Cost Cases: Comparing Prior Cost Versus Diagnosis-Based Methods. *Health Services Research* 36(6):194-206, December 2001.

Billings, J. and Mijanovich, T.: Improving the Management of Care for High-Cost Medicaid Patients. *Health Affairs* 26(6):1643-1655, November/December 2007.

Chronic Illness and Disability Payment System: *Chronic Illness and Disability Payment System Software, Version 2.5*. 2005.

DxCG: *DxCG RiskSmart™ Stand Alone Software, Version 2.2*. 2007.

DxCG: *DxCG Methodology*. Internet address: http://www.dxcg.com/research-resources/index.asp (Accessed 2008.)

Gilmer, T., Kronick, R., Fishman, P., et al.: The Medicaid Rx Model: Pharmacy-Based Risk Adjustment for Public Programs. *Medical Care* 39(11):1188-1202, November 2001.

Goetzel, R.Z., Ozminkowski, R.J., Villagra, V.G., et al: Return on Investments in Disease Management: A Review. *Health Care Financing Review* 26(4):1-19, Summer 2005.

Hartnell, N. and MacKinnon, N.: Neural Networks: From Science Fiction to Pharmacy. *American Journal of Health-System Pharmacy* 60:1908-1909, September 15, 2003.

Johns Hopkins University, School of Public Health: *The Johns Hopkins University ACG Case-Mix System*. Internet address: http://www.acg.jhsph.edu/ html/AboutACGs.htm (Accessed 2008.)

Kronick, R. and Dreyfus, T.: Diagnostic Risk Adjustment for Medicaid: The Disability Payment System. *Health Care Financing Review* 17(3):7-33, Spring 1996.

Kronick, R., Gilmer, T., Dreyfus, T., et al.: Improving Health-Based Payment for Medicaid Beneficiaries: CDPS. *Health Care Financing Review* 21(3):29-64, Spring 2000.

Martin, K., Rogal, D., and Arnold, S.: Health-Based Risk Assessment: Risk-Adjusted Payments and Beyond. *Robert Wood Johnson, Changes in Health Care Financing and Organization (HCFO) Program*. January 2004.

Maxwell, J.: Comprehensive Health Care Reform in Vermont: A Conversation with Governor Jim Douglas. *Health Affairs* 26(6):W697-W702, November/December 2007.

Office of Vermont Health Access: *Clinical Initiatives*. 2007a. Internet address: http://ovha.vermont.gov/for-consumers/ clinical_initiatives_brochure_6-21-07.pdf (Accessed 2008.)

Office of Vermont Health Access: Care Coordination: Identification, Verification, and Stratification. Personal communication. 2007b.

Office of Vermont Health Access: Risk Stratification and Intervention Services. Personal communication. 2008.

Pope, G.C., Kautter, J., Ellis, R.P., et al.: Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model. *Health Care Financing Review* 25(4):119-141, Summer 2004.

Rosen, A.K., Loveland, S., Anderson, J.J., et al.: Evaluating Diagnosis-Based Case-Mix Measures: How Well Do They Apply to the VA Population? *Medical Care* 39(7):692-704, July 2001.

Sloan, K.L., Montez-Rath, M.E., Spiro, A., et al.: Development and Validation of a Psychiatric Case-Mix System. *Medical Care* 44(6):568-580, June 2006.

Smith, N. and Weiner, J.: Applying Population-Based Case Mix Adjustment in Managed Care: The Johns Hopkins ACG System. *Managed Care Quarterly* 2(3):21-34, Summer 1994.

University of California, San Diego: *Medicaid Rx*. Internet address: http://medicaidrx.ucsd.edu/ (Accessed 2008.)

University of California, San Diego: *Chronic Illness and Disability Payment System*. Internet address: http://cdps.ucsd.edu/index.html (Accessed 2008.)

Weiner, J.: *The Johns Hopkins ACG Case-Mix System Reference Manual, Version 7.0.* Health Services Research & Development Center, The Johns Hopkins University Bloomberg School of Public Health. Baltimore, Maryland. 2005.

Weiner, J., Dobson, A., Maxwell, S., et al.: Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses. *Health Care Financing Review* 17(3):77-99, Spring 1996.

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

Winkelman, R. and Mehmud, S.: *A Comparative Analysis of Claims-Based Tools for Health Risk Assessment.* A Research Study Sponsored by the Society of Actuaries. April 2007.

Zhao, Y., Ash, A., Ellis, R., et al.: Predicting Pharmacy Costs and Other Medical Costs Using Diagnoses and Drug Claims. *Medical Care* 43(1):34-43, January 2005.

Zhao, Y., Ash, A.S., Ellis, R.P., et al.: Disease Burden Profiles: An Emerging Tool for Managing Managed Care. *Health Care Management Science* 5(3):211-219, August 2002.

Zhao, Y., Ash, A.S., Haughton, J., et al.: Identifying Future High-Cost Cases Through Predictive Modeling. *Disease Management & Health Outcomes* 11(6):389-397, June 2003.

Reprint Requests: Sharada Weir, D.Phil., Center for Health Policy and Research, University of Massachusetts Medical School, 333 South Street, Shrewsbury, MA 01545. E-mail: sharada.weir@umassmed.edu