Longitudinal Patterns of Spending Enhance the Ability to Predict Costly Patients

A Novel Approach to Identify Patients for Cost Containment

Julie C. Lauffenburger, PharmD, PhD,† Jessica M. Franklin, PhD,*† Alexis A. Krumme, MS,*† William H. Shrank, MD, MSHS,‡ Troyen A. Brennan, MD, JD, MPH,‡ Olga S. Matlin, PhD,‡ Claire M. Spettell, PhD,§ Gregory Brill, MS,* and Niteesh K. Choudhry, MD, PhD*†*

Background: With rising health spending, predicting costs is essential to identify patients for interventions. Many of the existing approaches have moderate predictive ability, which may result, in part, from not considering potentially meaningful changes in spending over time. Group-based trajectory modeling could be used to classify patients into dynamic long-term spending patterns.

Objectives: To classify patients by their spending patterns over a 1-year period and to assess the ability of models to predict patients in the highest spending trajectory and the top 5% of annual spending using prior-year predictors.

Subjects: We identified all fully insured adult members enrolled in a large US nationwide insurer and used medical and prescription data from 2009 to 2011.

Research Design: Group-based trajectory modeling was used to classify patients by their spending patterns over a 1-year period. We assessed the predictive ability of models that categorized patients in the top fifth percentile of annual spending and in the highest spending trajectory, using logistic regression and split-sample validation. Models were estimated using investigator-specified variables and a proprietary risk-adjustment method.

Results: Among 998,651 patients, in the best-performing model, prediction was strong for patients in the highest trajectory group (*C*-statistic: 0.86; R^2: 0.47). The *C*-statistic of being in the top fifth percentile of spending in the best-performing model was 0.82 (R^2: 0.26). Approaches using nonproprietary investigator-specified methods performed almost as well as other risk-adjustment methods (*C*-statistic: 0.81 vs. 0.82).

Conclusions: Trajectory modeling may be a useful way to predict costly patients that could be implementable by payers to improve cost-containment efforts.

Key Words: prediction, risk adjustment, longitudinal data, costs, trajectory

(*Med Care* 2016;00: 000–000)

Total national spending on health care in the United States has exceeded 17% of the gross domestic product and is projected to increase further in coming years.[1] Efforts to contain and manage costs rely, in part, on the ability to accurately predict patients' spending. There have been many previous attempts at prediction using a variety of patient features and data sources. However, these approaches have achieved modest predictive ability, rely on potentially cost-prohibitive proprietary platforms, and/or use data that may be challenging for payers and researchers to quickly obtain, such as medical claims.[2–5]

Current approaches have generally attempted to predict a single composite value of total health care spending and have often defined patients as "high-cost" if their expenditures are in the top fifth percentile of their beneficiary pool.[6,7] However, the threshold for determining what constitutes high health spending is somewhat arbitrary, and many patients may experience substantial increases or decreases in spending over the course of a year, which may not be reflected in single composite measures.[4,5,8–11] Understanding the dynamics of spending may be important for prediction and ultimately cost-containment.

In contrast, group-based trajectory modeling, an application of finite mixture modeling, is a promising alternative strategy that estimates changes over time in an outcome that is measured repeatedly and identifies individuals with similar longitudinal patterns.[12–14] This method has been applied in the context of long-term medication adherence, levels of disability, and clinical biomarkers and has been shown to be even better than single composite values.[12,13,15–17] Trajectories have also been successfully

From the *Division of Pharmacoepidemiology and Pharmacoeconomics; †The Center for Healthcare Delivery Sciences, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; ‡CVS Health, Woonsocket, RI; and §Data Science, Aetna, Hartford, CT.

Supported by an unrestricted grant from CVS Health to Brigham and Women's Hospital.

W.H.S., T.A.B., and O.S.M. were employees of CVS Health at the time of this manuscript preparation. C.M.S. is an employee of Aetna. The remaining authors declare no conflict of interest.

Reprints: Niteesh K. Choudhry, MD, PhD, Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston MA 02120. E-mail: nkchoudhry@bwh.harvard.edu.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website, www.lww-medicalcare.com.

ISSN: 0025-7079/16/000-000

predicted as outcomes in previous modeling approaches.[18] Group-based trajectory models have not been applied to health care spending, even though costs are as dynamic as these other outcomes.

METHODS

In this study, our objective was to apply group-based trajectory modeling to describe and classify overall spending patterns over a 1-year period as well as identify patients with particularly high utilization within these spending patterns for potential cost-containment. We then assessed the ability to predict high resource utilization using predictors from the prior year. Lastly, we compared the ability of models to predict these spending outcomes using variables derived from health care databases with a commonly used proprietary risk-adjustment method.

Setting and Study Design

This study used medical and prescription data for the years 2009–2011 for all fully insured adult members enrolled in Aetna, a large US nationwide insurer. These data contained complete paid claims data for all procedures, physician encounters, hospitalizations, and filled prescriptions (including dose dispensed and amounts paid by the insurer and the patient). These data were linked to eligibility data that included age, sex, and zip code of residence. Aggregate data on socioeconomic status, race, and educational attainment were obtained by linking zip code of residence with data from the 2010 US Census. The Institutional Review Board of Brigham and Women's Hospital approved the study.

Our cohort was constructed using 2 different "entry" years, 2010 and 2011, to enhance generalizability over different years (Supplemental Fig. 1, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). To be included, patients (18 y or above) were required to maintain continuous eligibility for the entire calendar year (January to December) and for ≥ 1 calendar year before their entry year. This prior year is hereafter referred to as the baseline year (eg, 2009 or 2010, depending on the entry year). In addition, they had to have ≥ 1 pharmacy and medical claim in both the entry year and baseline year to ensure active insurance use. Patients were excluded if they had missing sex and age information. If patients were eligible for both entry years, only their first year was examined.

Costs

We measured total monthly health care spending for each eligible patient by summing all of the allowed inpatient costs, outpatient medical and physician office visit costs, and outpatient drug costs. Monthly costs were generated by summing the costs per month, dividing by the number of days in that month and multiplying by 30. Costs were then logarithmically transformed to normalize their distribution, as frequently done in previous research.[11,19] All costs were inflated using the medical care component of the consumer price index to 2011 dollars when necessary.

Baseline Predictors

We specified 36 clinically relevant characteristics using enrollment files and medical and pharmacy claims in the 12 months before cohort entry. All investigator-specified predictors are shown in Table 1. Sociodemographic characteristics included age, sex, region, health plan, and community-level variables based on member's ZIP code of residence, such as median household income, race/ethnicity, and educational attainment. Clinical characteristics were measured using International Classification of Diseases Ninth edition codes in medical files and included human immunodeficiency virus infection and acquired immune deficiency syndrome, cancer, end-stage renal disease, coronary artery disease, chronic obstructive pulmonary disease/asthma, hypertension, hyperlipidemia, congestive heart failure, stroke, major depression, diabetes, liver disease, chronic kidney disease, atrial fibrillation, Alzheimer/dementia, osteoporosis, alcohol use, obesity, stress, and tobacco use (Supplemental Table 1, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). Each member's numbers of unique prescriptions (by generic name), outpatient physician office visits, emergency room visits not leading to hospitalizations, days spent in the hospital, and log-transformed spending in the baseline year, and generosity of medical and prescription benefits (copayments and deductibles/total net payments) were also measured.[20]

The use of and adherence to 25 common, chronic medication classes (eg, beta-blockers) were measured in the baseline year (Supplemental Table 2, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). For each class with ≥ 1 prescription filled in the first 6 months of the baseline year, we created a "supply diary" beginning with the first fill. This diary linked all of the observed fills based on dispensing date and days' supply; switching was allowed within each class of molecularly related drugs. From this diary, we calculated the proportion of days covered (PDC) as a weighted average across any of the 25 classes that the patient filled, to yield one average PDC.[21,22] For the applicable prediction models, use of any chronic disease medication was included as a binary indicator variable; average PDC was included as an interaction variable so the logistic regression would provide estimates on all patients and not just those with ≥ 1 chronic medication.

Risk Adjustment Instruments

We also classified patients using several common, validated risk adjustment methods. The Johns Hopkins Adjusted Clinical Group (ACG) case-mix system uses a proprietary algorithm based on medical and pharmacy claims to predict costs as well as the likelihood of being a high-cost patient. This risk adjustment method is thought to perform very well compared with other proprietary and nonproprietary instruments.[23] As commonly applied in the literature, the individual Aggregated Diagnosis Groups and Rx-defined Morbidity Groups generated from this case-mix software were entered as separate binary indicators.[3,7,23] In addition, we estimated the combined comorbidity score, a validated diagnosis-based tool, as a proxy for comorbidity and each patient's Rx-risk

TABLE 1. Baseline Characteristics of High-Cost Patients by Spending Definition

		Percentile-based Cost Definition		Trajectory-based Cost Definition	
Characteristics	**Full Cohort (N = 998,651)**	**Top 5th Percentile (N = 49,934)**	**Lower 95th Percentile (N = 948,717)**	**High-cost Trajectory (N = 314,271)**	**Other Trajectories (N = 684,380)**
Demographic					
Female (%)	59.6	56.1	59.8	61.1	58.9
Age [mean (SD)] (y)	45.8 (13.7)	51.9 (12.8)	45.5 (13.7)	50.9 (13.0)	43.5 (13.5)
West region (%)	17.8	19.0	17.7	18.0	17.7
South region (%)	41.9	39.3	42.1	42.0	41.9
Midwest region (%)	8.9	9.0	8.9	8.7	9.0
Northeast region (%)	31.4	32.7	31.3	31.4	31.4
Benefits' generosity [mean (SD)]					
Prescription	0.5 (0.3)	0.4 (0.3)	0.5 (0.3)	0.4 (0.2)	0.5 (0.3)
Medical	0.3 (0.2)	0.2 (0.1)	0.3 (0.2)	0.2 (0.1)	0.3 (0.2)
Health plan type (%)					
PPO	22.1	20.4	22.2	20.1	23.0
POS	19.3	19.1	19.3	19.2	19.4
HMO	32.9	34.6	33.8	33.2	32.7
Managed choice	24.6	24.8	24.6	26.2	23.9
Other	1.1	1.2	1.1	1.3	1.0
% Black [mean (SD)]	12.1 (18.3)	12.5 (19.2)	12.1 (18.3)	11.9 (18.2)	12.2 (18.4)
Income median [mean (SD)]	70,992 (25,968)	70,157 (26,322)	71,036 (25,948)	71,308 (26,239)	70,847 (25,841)
% HS graduate [mean (SD)]	88.6 (8.3)	88.3 (8.3)	88.5 (8.3)	88.8 (8.0)	88.5 (8.4)
Clinical					
Alcohol abuse (%)	4.2	8.3	4.0	5.4	3.7
Atrial fibrillation (%)	0.1	0.6	0.1	0.3	0.1
Coronary artery disease (%)	0.7	3.2	0.5	1.6	0.2
Cancer (%)	5.5	16.7	4.9	9.3	3.7
Congestive heart failure (%)	0.1	0.8	0.1	0.3	0.1
Chronic kidney disease (%)	1.4	6.2	1.1	3.1	0.6
COPD/asthma (%)	9.0	15.4	8.7	13.2	7.1
Dementia (%)	0.3	1.0	0.2	0.6	0.1
Depression (%)	8.2	14.5	7.9	14.4	5.3
Diabetes (%)	7.2	15.8	6.8	14.6	3.8
End-stage renal disease (%)	0.1	1.8	0.1	0.4	0.1
HIV/AIDS (%)	0.3	3.0	0.2	0.8	0.1
Hyperlipidemia (%)	32.2	43.6	31.6	46.7	25.5
Hypertension (%)	26.8	42.6	26.0	42.3	19.7
Liver disease (%)	0.1	0.8	0.1	0.3	0.1
Obesity (%)	5.3	9.6	5.1	7.5	4.3
Osteoporosis (%)	6.4	11.6	6.1	10.1	4.7
Stress (%)	3.6	5.8	3.5	5.6	2.7
Stroke (%)	0.2	0.7	0.1	0.4	0.1
Tobacco use (%)	4.6	9.4	4.4	6.0	4.0
ER visits, [mean (SD)]	0.3 (0.8)	0.6 (1.7)	0.2 (0.7)	0.3 (1.0)	0.2 (0.6)
Office visits [mean (SD)]	10.0 (11.7)	22.0 (25.3)	9.3 (10.2)	15.9 (16.3)	7.2 (7.4)
Hospital days [mean (SD)]	0.6 (16.5)	3.2 (33.1)	0.5 (15.1)	1.2 (24.6)	0.3 (10.9)
Costs					
Year 1 costs [mean (SD)]	5968 (16,566)	27,702 (51,120)	4824 (10,941)	11,390 (25,524)	3478 (9031)
Medication use and adherence					
Prescriptions [mean (SD)]	5.5 (4.3)	10.0 (7.0)	5.2 (4.0)	8.4 (5.3)	4.1 (3.0)
Medication use (%)	45.1	63.4	44.1	70.2	33.6
Adherence [mean (SD)]	0.7 (0.3)	0.7 (0.3)	0.7 (0.3)	0.8 (0.2)	0.7 (0.3)

(*Continued*)

TABLE 1. Baseline Characteristics of High-Cost Patients by Spending Definition (*continued*)

Characteristics	Full Cohort (N = 998,651)	Percentile-based Cost Definition		Trajectory-based Cost Definition	
		Top 5th Percentile (N = 49,934)	Lower 95th Percentile (N = 948,717)	High-cost Trajectory (N = 314,271)	Other Trajectories (N = 684,380)
Risk adjustment measures					
Comorbidity score [mean (SD)]	0.1 (0.9)	8.7 (1.9)	0.1 (0.8)	0.3 (1.3)	0.1 (0.7)
Rx-risk [mean (SD)]	2.4 (2.0)	4.3 (2.7)	2.2 (1.9)	3.8 (2.3)	1.8 (1.5)
No. ADGs [mean (SD)]	5.3 (3.3)	8.0 (4.1)	5.1 (3.2)	6.9 (3.6)	4.5 (2.8)
No. Rx-MGs [mean (SD)]	4.4 (3.0)	7.5 (4.4)	4.2 (2.8)	6.3 (3.5)	3.4 (2.2)

ADG indicates Adjusted Diagnosis Group; COPD, Chronic Obstructive Pulmonary Disease; ER, emergency room; HIV/AIDS, human immunodeficiency virus infection and acquired immune deficiency syndrome; HS, high school; Rx-MG, RX-defined morbidity groups.

score, a commonly used risk adjustment tool using pharmacy claims.[9,24]

Identification of High-Cost Patients

Costly patients were defined 2 different ways. We first identified a "high-cost" group of patients using trajectory modeling based on consistently high longitudinal spending.[12,13,15] Second, we classified patients as "high-cost" if they were within the top fifth percentile of the most costly patients based on annual spending, which is similar to other approaches in the literature.

In the trajectory approach, we used the SAS procedure Proc Traj, a free SAS add-on that fits a semiparametric mixture model to longitudinal data, as previously described.[13,15,25] In brief, group-based trajectory models were developed as an application of finite mixture modeling and identify clusters of individuals who follow similar patterns over time.[13] This modeling approach analyzes longitudinal data by fitting a semiparametric (discrete) mixture model and assigns groupings in data based on their probability distributions.[25] We modeled longitudinal cost trajectories in the entry year (2010 or 2011) using calendar month as the time variable and a censored normal model. The models were estimated using a "forward" classifying approach using from 2 to 7 groups, each time investigating model fit using the Bayesian information criterion (BIC), whereby lower BIC indicates better model fit.[13] The number of groups investigated was capped at 7, owing to the ability to interpret separate trajectory groups, lowest BIC obtained, $\geq 5\%$ of the sample in each group, and the abilities of models to converge.[14] All these factors are considerations in choosing the best trajectory model.[25–27] The group-based trajectory modeling approach estimates each individual's probability of membership in each group and assigns individuals to the trajectory group based on their highest probability of membership.

Statistical Analysis

To assess the ability to predict costly patients, we compared the performance of models that categorized patients within the top fifth percentile of annual spending versus the best-fitting trajectory model that identified a "high-cost" group against all other trajectory groups. For each of these 2 cost outcomes, we fit 5 different prediction models using logistic regression. As shown in Table 2, model 1 utilized patient demographic and clinical characteristics as predictors, model 2 used characteristics that were obtainable using only pharmacy claims and enrollment data, and model 3 utilized all investigator-specified baseline characteristics. In addition, model 4 included the predictors from the Johns Hopkins ACG software alone, and model 5 incorporated the ACG software predictors with other baseline predictors, a

TABLE 2. Ability of Models to Predict High-Cost Spending Groups

Model	Predictors	*C*-statistic (95% CI)		Pseudo R^2 (95% CI)		Net Reclassification Index (95% CI)	
		Top 5th Percentile	High-cost Trajectory	Top 5th Percentile	High-cost Trajectory	Top 5th Percentile	High-cost Trajectory
1	Demographic+clinical	0.77 (0.76, 0.78)	0.80 (0.80–0.80)	0.17 (0.16–0.17)	0.32 (0.32–0.32)	Referent	Referent
2	Pharmacy claims*	0.77 (0.77, 0.77)	0.86 (0.86–0.86)	0.17 (0.16–0.18)	0.46 (0.46–0.47)	0.14 (0.13, 0.15)	0.62 (0.61, 0.63)
3	All IS†	0.81 (0.81, 0.82)	0.85 (0.85–0.85)	0.21 (0.21–0.22)	0.44 (0.43–0.44)	0.60 (0.59, 0.62)	0.67 (0.66, 0.68)
4	ADGs+Rx-MGs	0.80 (0.80, 0.81)	0.83 (0.83, 0.83)	0.22 (0.21, 0.22)	0.39 (0.39, 0.39)	0.45 (0.43, 0.46)	0.52 (0.52, 0.53)
5	ADGs+Rx-MGs+all IS†	0.82 (0.81, 0.83)	0.86 (0.86, 0.86)	0.26 (0.23, 0.30)	0.47 (0.46–0.47)	0.63 (0.62, 0.64)	0.75 (0.75, 0.76)

*Pharmacy claims: demographic+medication use/adherence+year 1 pharmacy costs.
†Investigator-specified (IS): demographic+clinical+medication use/adherence+year 1 (baseline) total costs.
ADG indicates adjusted diagnosis group; CI, confidence interval; Rx-MG, RX-defined Morbidity Groups.

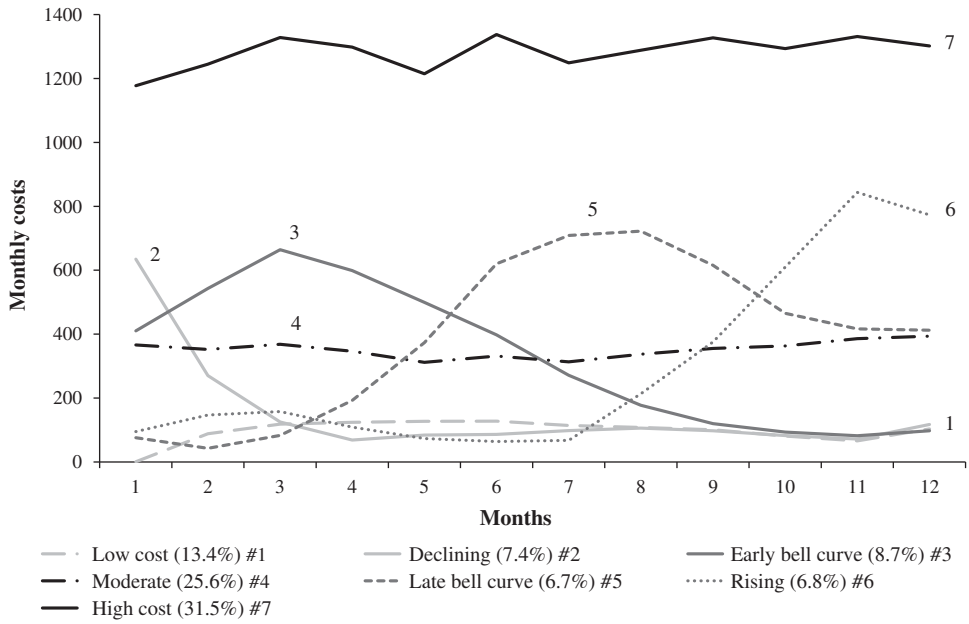


FIGURE 1. Seven spending phenotypes using the trajectory model approach. The mean spending levels using 7-group trajectory modeling in the full sample are plotted. The percentages refer to the number of patients who belong to each trajectory group out of the full cohort.

frequently recommended approach.[3,23] Secondary models with other predictor combinations and risk-adjustment methods were also conducted.

To avoid "over optimism" bias, we used internal split-sample validation by randomly dividing the full cohort into two halves as an initial derivation sample and a validation sample.[28] We evaluated each model through discrimination, performance, and reclassification measures.[29] Discrimination, the ability of the model to distinguish between patients who do and do not experience the outcome, was measured by the *C*-statistic, which ranges from 0.5 (non-informative model) to 1.0 (perfect prediction).[30,31] Pseudo R^2 were used to assess model performance by examining the variation explained by the model, ranging from 0 (no variation explained) to 1.0 (all explained).[32] The continuous net reclassification improvement index (NRI) was used to assess changes in the predicted outcome classification of expanded models versus model 1 as the base model.[22,24] The NRI measures the net proportion of case and control patients that are correctly reclassified by other models and is thought to enhance interpretations from *C*-statistics due to its increased sensitivity to changes in model predictions.[33] The *C*-statistics, R^2, and the NRIs were calculated with bootstrapped 95% confidence intervals using 1000 replications with replacement sampling.[34,35]

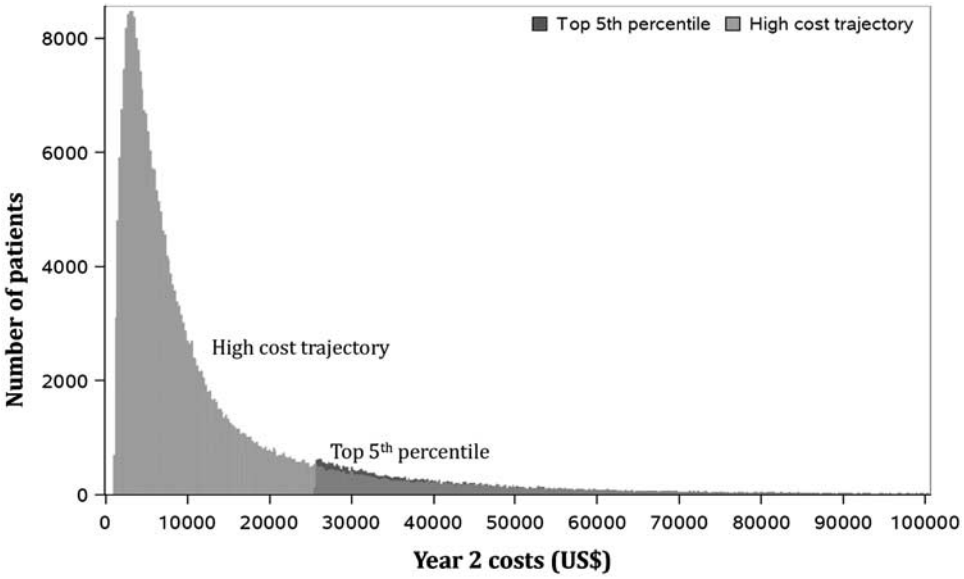
We also performed several sensitivity analyses. First, analyses among clinically relevant patient subgroups were conducted, including males and females separately, patients aged below and above 65 years, and among patients with diabetes, hypertension, or depression. The best non-proprietary predictive model (defined by the highest *C*-statistic) and the model using only variables obtainable from pharmacy claims data were used for these subanalyses. We also estimated models that predicted spending in the top 10th percentile and models that included an indicator of use and a PDC interaction term for the medication classes individually rather than an average PDC. For clinical context, we explored the relationship between potentially intervenable patient characteristics (eg, alcohol use) and high-cost trajectory membership and prediction in model 3. Finally, we measured the spending in the entry year associated with modifiable risk factors[36] and estimated additional trajectory models (Supplement, Supplemental Digital Content 1, http://links.lww.com/MLR/B288).

All analyses were performed using SAS 9.4 (Cary, NC).

RESULTS

Cohort Characteristics

In total, 998,651 patients were included in the study (Supplemental Fig. 1, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). Their mean age was 45.8 (SD: 13.7), and 59.6% were female. Almost 50% of patients in the top quartile of spending were not previously costly (Supplemental Fig. 2, Supplemental Digital Content 1, http://links.lww.com/MLR/B288), indicating dynamic movement across years. The 7-group trajectory model had the lowest BIC and is presented in Figure 1; a sample patient case for each group is shown in Supplemental Table 3 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288). Baseline characteristics differed between the "high-cost" spending groups and the other less costly groups (Table 1). Baseline characteristics for the other trajectory groups are



Spending definition	No. of patients	Mean (SD)	Median	Minimum	Maximum	Q1	Q3
Top 5th percentile (US$)	49,934	62,893 (71,358)	42,047	24,574	2,491,003	31,617	65,335
High cost trajectory (US$)	314,271	15,392 (33,791)	7,058	953	2,491,003	3,871	14,563

FIGURE 2. Spending in year 2 by high-cost spending definition.

shown in the Appendix Table. Trajectory modeling approaches using fewer groups are shown in Supplemental Figure 3 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288).

In the 7-group model, the trajectory with the highest average spending included 31.5% of patients and accounted for 68.1% of total spending. In comparison, the top fifth percentile of patients, who all had annual costs exceeding $25,000, accounted for 44.2% of the cohort's total spending. Compared with the high-cost trajectory, the patients in the top fifth spending percentile had more variation and right skew in their costs (Fig. 2).

Cost Prediction

Table 2 shows the results of the 5 main prediction models in the validation sample (N = 499,780). The ability to accurately predict costly patients was good across all models, including those using only demographic and clinical characteristics. Models in which high-cost was defined using the high-cost trajectory had consistently higher *C*-statistics and R^2 than using the top fifth percentile. For example, the best-performing high-cost trajectory prediction model had a *C*-statistic of 0.86 while the best-performing top fifth percentile prediction model had a *C*-statistic of 0.82. Similarly, up to 47% of the variation in costly patients could be explained using trajectory modeling while only 26% of the variation could be explained using the top fifth percentile approach.

Overall, the best-performing models for both spending outcome definitions (model 5) included the Johns Hopkins ACG software predictors along with age, sex, and baseline costs (fifth percentile *C*-statistic: 0.82, R^2: 0.26; high-cost trajectory *C*-statistic: 0.86, R^2: 0.47). For both spending outcome definitions, models including non-proprietary model predictors (eg, model 3) still had very high predictive ability. These models outperformed models with other predictor combinations and risk adjustment instruments (Supplemental Table 4, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). The multivariate results from the best-performing models (model 3) are given in Supplemental Table 5 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288).

Similarly, for both spending outcome definitions, the model that incorporated only predictors from pharmacy claims and enrollment information (model 2) was still strongly predictive of costly patients (top fifth percentile *C*-statistic: 0.77; high-cost trajectory *C*-statistic: 0.86). In fact, using the trajectory approach, the use of pharmacy claims alone outperformed the model that included all investigator-specified characteristics, even when examining the bootstrapped 95% confidence intervals. The estimates from other models in the validation sample are shown in Supplemental Table 4 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288). The predictive accuracies in the initial derivation sample (N = 498,871) were similar to the validation sample (Supplemental Table 6, Supplemental Digital Content 1, http://links.lww.com/MLR/B288).

Sensitivity and Exploratory Analyses

The *C*-statistics of the prediction models among patient subgroups are shown in Figure 3. The trajectory modeling approach demonstrated good model discrimination within

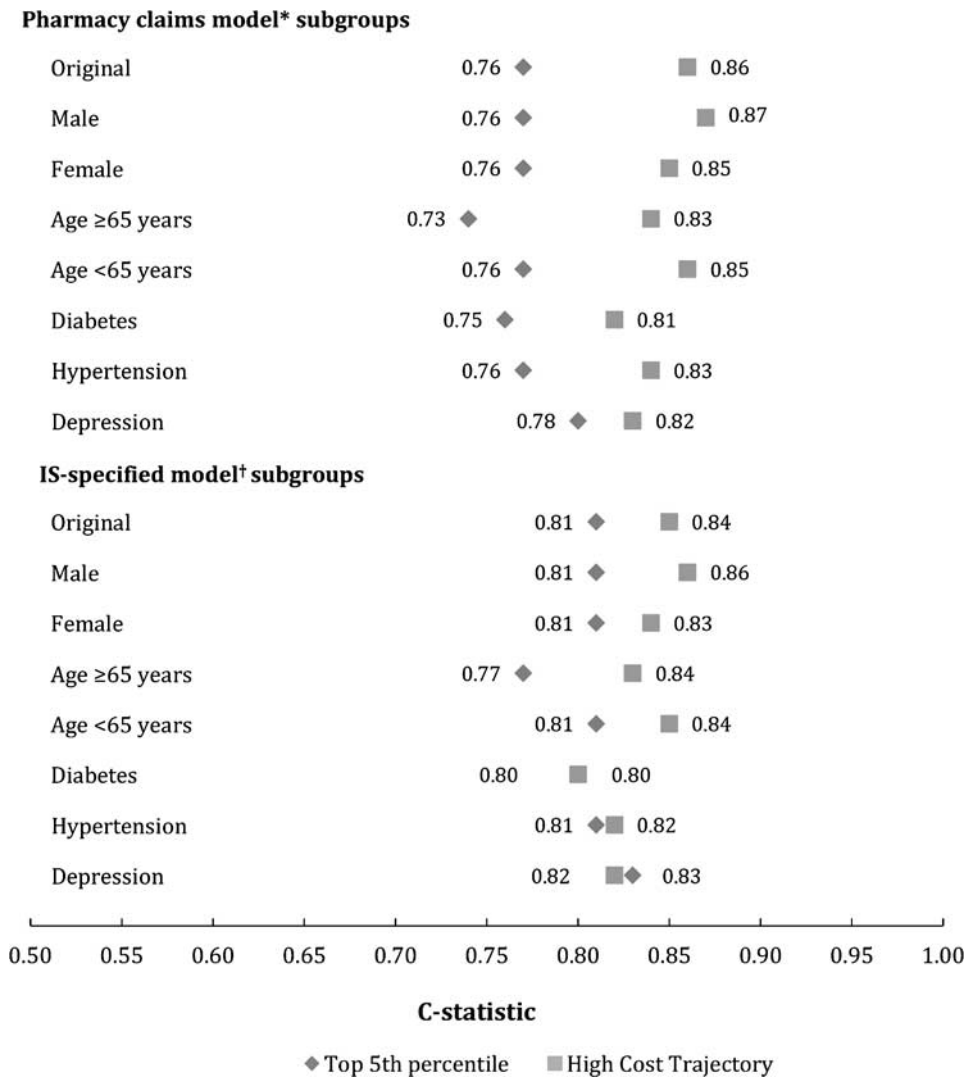


FIGURE 3. Ability of models to predict high-cost spending within patient subgroups. *Demographic + medication use/adherence + year 1 pharmacy costs. †investigator-specified (IS): demographic + clinical + medication use/adherence + year 1 total costs.

strata based upon age, sex, and presence of common chronic conditions. Specifically, the best-performing investigator-specified model (model 3) and the pharmacy claims-only model yielded very similar *C*-statistics as the full cohort for both spending outcome definitions. Models that predicted spending in the top 10th percentile performed similarly to the top fifth percentile (Supplemental Table 7, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). Models that included use and adherence predictors for the medication classes individually were very similar to the primary analyses (Supplemental Table 8, Supplemental Digital Content 1, http://links.lww.com/MLR/B288). Several potentially intervenable characteristics, such as depression or obesity, were associated with high-cost trajectory membership (Supplemental Table 9, Supplemental Digital Content 1, http://links.lww.com/MLR/B288); the ability to predict high-cost trajectory membership was still high with only potentially intervenable characteristics. We also provide trajectories of spending on potentially modifiable risk factors in Supplemental Figure 4 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288) and their prediction statistics in Supplemental Table 10 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288).

DISCUSSION

In this cohort of commercially insured patients, we found that the use of group-based trajectory modeling substantially improved the ability to predict who will be costly compared with a more conventional approach of defining high-cost users as those in the top fifth percentile of spending. This novel cost-modeling approach had high predictive accuracy, even when using easily implementable and non-proprietary baseline patient characteristics.

Prior efforts to predict health care spending have had modest accuracy and primarily assessed model performance using

R^2 values; these R^2 values range from, approximately, 0.10 to 0.30, depending on predictors and patient population.[6,9,37,38] This accuracies may have resulted, in part, because of how the outcome of interest, costs, has been categorized or calculated. Existing approaches have often focused on the use of a single measure of costs, such as total yearly costs or a threshold-based measure, like being in the top fifth spending percentile, which collapse an entire year's spending into a single static variable. As a result, patients who have consistently high spending throughout the year are grouped together with other, potentially different patients, who may have extremely high but short-lived periods of resource consumption.

In this study, we sought to evaluate a novel, cost-modeling approach that has been shown to better classify and improve the ability to predict other similarly dynamic outcomes.[15,27] This approach, group-based trajectory modeling, estimates changes over time in an outcome that is measured repeatedly and identifies individuals who have similar longitudinal patterns. One advantage of this approach is that it allows the data to define the outcome, rather than using an arbitrarily selected threshold of costly patients. This way of empirically defining costly patients may explain why trajectory models in our study performed better than more conventionally defining patients in the top fifth percentile. In addition, because this cost-modeling approach incorporates the timing of incurred costs, we may better identify when to intervene on costly patients.

The size of the high-cost trajectory group itself warrants further reflection. The trajectory modeling method classified 31% of the patients into the high-cost group, as opposed to, by definition, a much smaller proportion of patients using the top fifth spending percentile approach, which is why the average spending is lower. Although the top fifth percentile of patients accounts for an extremely disproportionate amount of spending, targeting more patients for cost-containment might be beneficial for several reasons. First, lowering the spending of a larger subgroup of costly patients will lower overall spending in the population, even if the average spending per patient only slightly decreases. Second, these patients were identified as being consistently high spenders, rather than episodically high spenders.

Of course, it may be infeasible to intervene upon a group as large as the high-cost trajectory, and not all costs may be preventable. Therefore, identifying additional subgroups may be necessary, and this trajectory approach may be a starting point. For example, a prediction rule, such as Supplemental Table 8 (Supplemental Digital Content 1, http://links.lww.com/MLR/B288), could be applied to the high-cost trajectory to identify patients with potentially modifiable characteristics who might be the ultimate target for cost-containment. Additional approaches that estimate trajectories of modifiable costs could also be useful.

We also compared the ability of models to predict costly patients using a variety of available variables derived from health care databases with commonly used risk-adjustment methods. One of proprietary approaches, the Johns Hopkins ACG system, has been shown to have favorable predictive accuracy in total costs and top fifth percentile of annual spending compared with other approaches but can still be cost-prohibitive for researchers and other health care organizations.[6,7] We found that commonly-used baseline characteristics also remarkably predicted costly patients. Using demographic, clinical, and prior year cost characteristics to predict either spending outcome was only slightly inferior to the ACG methodology based on all of the model performance measures. Thus, regardless of how spending is defined, costly patients may still be readily predicted even when proprietary candidate predictors may be infeasible or unavailable. This good model performance persisted even when predicting costly patients within subgroups of age, sex, and common chronic conditions, indicating good applicability across different types of patients.

Similarly, the models incorporating only predictors available in pharmacy claims data were also highly predictive of costly patients, performing only slightly worse than models using all available information, consistent with prior analyses from the Society of Actuaries showing that pharmacy claims data are powerful predictors of future costs.[39] These findings support that even pharmacy benefit managers, without access to all clinical claims data, may be able prospectively identify patients who are likely to be costly. These data are frequently more easily and quickly accessible to payers for policy decision-making, enabling these organizations to accurately target disease management programs and cost-containment.

These findings should be interpreted in light of several limitations. First, we restricted the analyses to patients with complete covariate information; while only a small proportion of the original cohort was lost (1.5%), these patients were also less likely to costly. We also only included patients if they used their insurance at least once to ensure active benefits. Although these patients would not have been classified as costly, the trajectories and predictive ability could have differed if all patients had been included. We also examined trajectories over the calendar year; patients with incomplete enrollment or other policy-level start and end dates may have different patterns. Second, due to differences in how the outcomes are categorized, the model performance of predicting a cost trajectory (a binary outcome) cannot be directly compared with predicting total costs (a continuous outcome), which is why the trajectory models were compared with the top fifth spending percentile rather than total costs; however, this threshold-based method is also frequently used. The candidate variables that were included in these prediction models may also not be exhaustive. Lastly, these results cannot be generalized to other payment systems, such as Medicaid or Medicare beneficiaries. Future work should consider trajectories with longer time horizons, stratifications by clinical condition, or further identifying and estimating potentially modifiable or discretionary costs.

In conclusion, we found that a dynamic approach using group-based trajectory models may be highly accurate and outperforms other conventional methods of identifying high-cost patients. This strategy is easily implementable and could optimize predictions that identify patients for potential cost-containment.

APPENDIX

TABLE 1. Baseline Characteristics of the Other 6 Trajectory Groups

Characteristics	Low Cost (N = 133,378)	Declining (N = 74,208)	Early Bell Curve (N = 86,726)	Moderate Cost (N = 255,922)	Late Bell Curve (N = 66,604)	Rising (N = 67,542)
Female (%)	48.9	54.0	62.1	63.8	61.6	59.3
Age [mean (SD)] (y)	39.7 (12.8)	40.8 (13.1)	42.8 (13.4)	46.5 (13.3)	43.7 (13.4)	43.2 (13.0)
West region (%)	17.9	17.7	17.8	17.7	17.3	17.7
South region (%)	42.5	41.4	41.3	42.0	42.0	41.4
Midwest region (%)	9.1	9.2	8.7	9.2	8.6	9.1
Northeast region (%)	30.5	31.7	32.2	31.2	32.1	31.8
Rx benefits' generosity [mean (SD)]	0.6 (0.3)	0.6 (0.3)	0.5 (0.3)	0.5 (0.3)	0.6 (0.3)	0.6 (0.3)
Medical benefits' generosity [mean (SD)]	0.3 (0.2)	0.3 (0.3)	0.3 (0.3)	0.3 (0.2)	0.3 (0.2)	0.3 (0.3)
Health plan type (%)						
PPO	24.5	23.5	23.0	22.1	22.6	23.6
POS	18.7	19.4	19.6	20.0	19.3	19.1
HMO	32.1	32.5	33.1	33.0	33.2	32.5
Managed choice	23.8	23.6	23.3	24.1	24.9	23.8
Other	0.9	1.0	1.0	1.1	1.0	1.1
% Black [mean (SD)]	12.4 (18.5)	12.2 (18.3)	12.4 (18.6)	12.0 (18.2)	12.3 (18.5)	12.2 (18.4)
Income median [mean (SD)]	70,060 (25,504)	70,787 (25,832)	70,964 (25,890)	70,898 (25,771)	71,629 (26,437)	71,349 (26,072)
% HS graduate, [mean (SD)]	88.1 (8.8)	88.4 (8.5)	88.5 (8.4)	88.7 (8.1)	88.7 (8.3)	88.6 (8.3)
Alcohol abuse (%)	3.6	3.8	3.8	3.7	3.5	3.6
Atrial fibrillation (%)	0.1	0.1	0.1	0.1	0.1	0.1
Coronary artery disease (%)	0.1	0.1	0.3	0.4	0.2	0.2
Cancer (%)	2.0	2.8	3.9	5.0	3.7	3.2
Congestive heart failure (%)	0.1	0.1	0.1	0.1	0.1	0.1
Chronic kidney disease (%)	0.2	0.3	0.6	0.9	0.5	0.4
COPD/asthma (%)	5.6	6.5	7.0	8.1	6.9	6.9
Dementia (%)	0.1	0.1	0.2	0.2	0.1	0.1
Depression (%)	2.6	4.0	6.2	7.4	4.6	4.0
Diabetes (%)	1.2	2.0	3.6	6.2	3.4	2.6
End-stage renal disease (%)	0.1	0.1	0.1	0.1	0.1	0.1
HIV/AIDS (%)	0.1	0.1	0.1	0.2	0.1	0.1
Hypertension (%)	9.3	12.4	18.4	29.2	18.0	15.1
Hyperlipidemia (%)	15.1	18.8	24.6	34.2	24.7	22.4
Liver disease (%)	0.1	0.1	0.1	0.1	0.1	0.1
Obesity (%)	3.0	3.8	4.7	5.1	4.1	4.0
Osteoporosis (%)	2.4	3.3	4.7	6.4	4.9	4.3
Stress (%)	1.7	2.4	3.3	3.2	2.4	2.4
Stroke (%)	0.1	0.1	0.1	0.1	0.1	0.1
Tobacco use (%)	3.8	4.0	4.1	4.1	3.8	3.9
ER visits [mean (SD)]	0.2 (0.6)	0.2 (0.7)	0.2 (0.7)	0.2 (0.6)	0.2 (0.6)	0.2 (0.6)
Office visits [mean (SD)]	4.9 (5.1)	6.3 (6.6)	8.1 (8.3)	8.7 (8.3)	7.0 (6.9)	6.5 (6.3)
Hospital days [mean (SD)]	0.2 (1.7)	0.4 (20.4)	0.4 (13.7)	0.4 (9.1)	0.3 (1.7)	0.3 (12.9)
Year 1 costs [mean (SD)]	2189 (5931)	2947 (7062)	3896 (9806)	4342 (11,023)	3220 (7195)	2961 (7351)
Prescriptions [mean (SD)]	3.0 (2.2)	3.5 (2.5)	4.3 (3.0)	5.0 (3.3)	3.9 (2.7)	3.7 (2.6)
Medication use (%)	13.4	19.6	31.9	51.8	30.2	24.9
Adherence [mean (SD)]	0.5 (0.3)	0.6 (0.3)	0.7 (0.3)	0.7 (0.2)	0.6 (0.3)	0.6 (0.3)
Comorbidity score [mean (SD)]	0.1 (1.2)	0.1 (0.6)	0.1 (0.7)	0.1 (0.8)	0.1 (0.6)	0.1 (0.6)
Rx-risk [mean (SD)]	1.2 (1.2)	1.4 (1.3)	1.8 (1.5)	2.3 (1.6)	1.7 (1.4)	1.5 (1.4)
No. ADGs [mean (SD)]	3.6 (2.4)	4.1 (2.6)	4.7 (2.8)	5.1 (3.0)	4.5 (2.8)	4.3 (2.7)
No. Rx-MGS [mean (SD)]	2.6 (1.7)	3.0 (1.9)	3.5 (2.2)	4.1 (2.4)	3.3 (2.1)	3.1 (2.0)

ADG indicates adjusted diagnosis group; ER, emergency room; HIV/AIDS, human immunodeficiency virus infection and acquired immune deficiency syndrome; HS, high school; COPD, chronic obstructive pulmonary disease; Rx-MG, RX-defined Morbidity Groups.

REFERENCES

1. Centers for Disease Control. US health expenditures. Available at: www.cdc.gov/nchs/fastats/health-expenditures.htm. Accessed January 21, 2016.
2. Maciejewski ML, Liu CF, Fihn SD. Performance of comorbidity, risk adjustment, and functional status measures in expenditure prediction for patients with diabetes. *Diabetes Care*. 2009;32:75–80.
3. Kuo RN, Lai MS. Comparison of Rx-defined morbidity groups and diagnosis-based risk adjusters for predicting healthcare costs in Taiwan. *BMC Health Serv Res*. 2010;10:126.
4. Forrest CB, Lemke KW, Bodycombe DP, et al. Medication, diagnostic, and cost information as predictors of high-risk patients in need of care management. *Am J Manag Care*. 2009;15:41–48.
5. Yarger S, Rascati K, Lawson K, et al. Analysis of predictive value of four risk models in Medicaid recipients with chronic obstructive pulmonary disease in Texas. *Clin Ther*. 2008;30:1051–1057.
6. Kuo RN, Dong YH, Liu JP, et al. Predicting healthcare utilization using a pharmacy-based metric with the WHO's Anatomic Therapeutic Chemical algorithm. *Med Care*. 2011;49:1031–1039.
7. Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. *J Clin Epidemiol*. 2004;57:1040–1048.
8. Sales AE, Liu CF, Sloan KL, et al. Predicting costs of care using a pharmacy-based measure risk adjustment in a veteran population. *Med Care*. 2003;41:753–760.
9. Fishman PA, Goodman MJ, Hornbrook MC, et al. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. *Med Care*. 2003;41:84–99.
10. Powers CA, Meyer CM, Roebuck MC, et al. Predictive modeling of total healthcare costs using pharmacy claims data: a comparison of alternative econometric cost modeling techniques. *Med Care*. 2005;43:1065–1072.
11. Mihaylova B, Briggs A, O'Hagan A, et al. Review of statistical methods for analysing healthcare resources and costs. *Health Econ*. 2011;20:897–916.
12. Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychol Methods*. 2001;6:18–34.
13. Jones BL, Nagin DS. Advances in group-based trajectory modeling and a SAS procedure for estimating them. *Sociol Methods Res*. 2007;35:542–571.
14. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Ann Rev Clin Psychol*. 2010;6:109–138.
15. Franklin JM, Shrank WH, Pakes J, et al. Group-based trajectory models: a new approach to classifying and predicting long-term medication adherence. *Med Care*. 2013;51:789–796.
16. Gill TM, Gahbauer EA, Han L, et al. Trajectories of disability in the last year of life. *N Engl J Med*. 2010;362:1173–1180.
17. Mustillo S, Worthman C, Erkanli A, et al. Obesity and psychiatric disorder: developmental trajectories. *Pediatrics*. 2003;111:851–859.
18. Franklin JM, Shrank WH, Lii J, et al. Observing versus predicting: initial patterns of filling predict long-term adherence more accurately than high-dimensional modeling techniques. *Health Serv Res*. 2016;51:220–239.
19. Austin PC, Ghali WA, Tu JV. A comparison of several regression models for analysing cost of CABG surgery. *Stat Med*. 2003;22:2799–2815.
20. Artz MB, Hadsall RS, Schondelmeyer SW. Impact of generosity level of outpatient prescription drug coverage on prescription drug events and expenditure among older persons. *Am J Public Health*. 2002;92:1257–1263.
21. Benner JS, Glynn RJ, Mogun H, et al. Long-term persistence in use of statin therapy in elderly patients. *JAMA*. 2002;288:455–461.
22. Choudhry NK, Shrank WH, Levin RL, et al. Measuring concurrent adherence to multiple related medications. *Am J Manag Care*. 2009;15:457–464.
23. Haas LR, Takahashi PY, Shah ND, et al. Risk-stratification methods for identifying patients for care coordination. *Am J Manag Care*. 2013;19:725–732.
24. Johnson ML, El-Serag HB, Tran TT, et al. Adapting the Rx-risk-V for mortality prediction in outpatient populations. *Med Care*. 2006;44:793–797.
25. Jones BL, Nagin DS, Roeder KA. SAS procedure based on mixture models for estimating developmental trajectories. *Sociol Methods Res*. 2001;29:374–393.
26. Li Y, Zhou H, Cai B, et al. Group-based trajectory modeling to assess adherence to biologics among patients with psoriasis. *Clinicoecon Outcomes Res*. 2014;6:197–208.
27. Franklin JM, Krumme AA, Tong AY, et al. Association between trajectories of statin adherence and subsequent cardiovascular events. *Pharmacoepidemiol Drug Saf*. 2015;24:1105–1113.
28. Steyerberg EW, Harrell FE Jr., Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54:774–781.
29. Waljee AK, Higgins PD, Singal AG. A primer on predictive models. *Clin Transl Gastroenterol*. 2014;5:e44.
30. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128–138.
31. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115:928–935.
32. Czado C, Gneiting T, Held L. Predictive model assessment for count data. *Biometrics*. 2009;65:1254–1261.
33. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30:11–21.
34. Cook NR. Clinically relevant measures of fit? A note of caution. *Am J Epidemiol*. 2012;176:488–491.
35. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med*. 2004;23:2109–2123.
36. Goetzel RZ, Pei X, Tabrizi MJ, et al. Ten modifiable health risk factors are linked to more than one-fifth of employer-employee health care spending. *Health Aff*. 2012;31:2474–2484.
37. Liu CF, Sales AE, Sharp ND, et al. Case-mix adjusting performance measures in a veteran population: pharmacy- and diagnosis-based approaches. *Health Serv Res*. 2003;38:1319–1337.
38. Zhao Y, Ash AS, Ellis RP, et al. Predicting pharmacy costs and other medical costs using diagnoses and drug claims. *Med Care*. 2005;43:34–43.
39. Winkelman R. A Comparative analysis of claims-based tools for health risk assessment. Available at www.soa.org/Files/Research/Projects/risk-assessmentc.pdf. Accessed May 11, 2016.