HSR Health Services Research

DOI: 10.1111/1475-6773.12222
RESEARCH ARTICLE

Using Self-Reported Health Measures to Predict High-Need Cases among Medicaid-Eligible Adults

Laura R. Wherry, Marguerite E. Burns, and Lindsey Jeanne Leininger

Objective. To assess the ability of different self-reported health (SRH) measures to prospectively identify individuals with high future health care needs among adults eligible for Medicaid.

Data Sources. The 1997–2008 rounds of the National Health Interview Survey linked to the 1998–2009 rounds of the Medical Expenditure Panel Survey ($n = 6{,}725$).

Study Design. Multivariate logistic regression models are fitted for the following outcomes: having an inpatient visit; membership in the top decile of emergency room utilization; and membership in the top cost decile. We examine the incremental predictive ability of six different SRH domains (health conditions, mental health, access to care, health behaviors, health-related quality of life [HRQOL], and prior utilization) over a baseline model with sociodemographic characteristics. Models are evaluated using the *c*-statistic, integrated discrimination improvement, sensitivity, specificity, and predictive values.

Principal Findings. Self-reports of prior utilization provide the greatest predictive improvement, followed by information on health conditions and HRQOL. Models including these three domains meet the standard threshold of acceptability (*c*-statistics range from 0.703 to 0.751).

Conclusions. SRH measures provide a promising way to prospectively profile Medicaid-eligible adults by likely health care needs.

Key Words. Medicaid, prediction models, self-rated health measurement, risk assessment

Medicaid agencies are increasingly adopting care coordination and management strategies for their high-need beneficiaries, often with the twin goals of enhancing quality of care and reducing costs (Verdier, Byrd, and Stone 2009; Kim et al. 2013). This trend is likely to accelerate as states implement the Affordable Care Act's (ACA) new options to coordinate the care of chronically ill Medicaid beneficiaries (Kaiser Family Foundation 2011) and, in at

least 28 states, expand Medicaid coverage to a larger population (Advisory Board 2013). To target high-need individuals, care management initiatives typically apply a case-finding approach that uses predictive models to stratify patients along likely future health care utilization (Knutson, Bella, and Llanos 2009; Verdier, Byrd, and Stone 2009). There is a mismatch, however, between the data requirements of these models, which rely on historical medical claims, and the data available to Medicaid programs for both new and "churning" beneficiaries, of whom states expect increasing numbers under the ACA (Knutson, Bella, and Llanos 2009). This article proposes and tests a practical alternative to a claims-based algorithm for identifying high-need Medicaid beneficiaries, one that relies upon well-validated self-reported health (SRH) measures that states or managed care organizations (MCOs) may potentially collect during the initial application and enrollment process.

Background

Under Medicaid, many states have developed care coordination and management programs for "high-cost, high user" beneficiaries (Centers for Medicare and Medicaid Services 2010). Comprehensive reviews of these types of programs consistently find that program success critically depends on identifying and targeting individuals with complex health care needs (Bodenheimer and Berry-Millett 2009; Brown 2009). Indeed, the Center for Medicare and Medicaid Services (2012) encourages states to use risk stratification to identify high-risk Medicaid patients and to help prioritize enrollment in enhanced coordinated care. Over the past several decades, sophisticated algorithms have been developed employing recent claims history for such case-finding purposes (Weiner et al. 1991; Kronick et al. 2000; Ash et al. 2001; Weir, Aweh, and Clark 2008).

However, the reliance on the availability of recent claims history is a vulnerability in the context of serving Medicaid populations. Recent claims data are unavailable for new beneficiaries, which poses a serious limitation as states implement the optional Medicaid expansion for low-income adults under the ACA with an estimated 9 million new enrollees in 2014 (Congressional Bud-

---

Address correspondence to Lindsey Jeanne Leininger, Ph.D., University of Illinois-Chicago School of Public Health, 1603 W. Taylor Street, Chicago, IL 60612; e-mail: lindseyl@uic.edu. Laura R. Wherry, Ph.D., is with the Division of General Internal Medicine and Health Services Research, David Geffen School of Medicine at the University of California, Los Angeles, CA. Marguerite E. Burns, Ph.D., is with the Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI.

get Office 2013). Moreover, the enrollment patterns of this expansion population will likely be characterized by high levels of churn, presenting an ongoing challenge for Medicaid agencies aiming to prospectively profile beneficiaries. Sommers and Rosenbaum (2011) estimate that approximately one-half of all income-eligible adults will experience a shift in eligibility between Medicaid and the exchanges within a 1-year period. Finally, even for continuing Medicaid members, the availability of claims data for analytic purposes may lag up to a year or more after care use depending on the claims submission process (Hilltop Institute 2003; Stanek and Takach 2010).

Given the extent of the "no history" problem among Medicaid populations, it is crucial to assess whether a nonclaims-based alternative exists that can support the prospective profiling of members by their likely health care use. The collection of SRH measures from new beneficiaries at the time of enrollment may be a promising alternative. A compelling literature employing samples from Medicare (Perrin et al. 2011), the VA (DeSalvo et al. 2009), and nationally representative adult populations (Fleishman et al. 2006; Fleishman and Cohen 2010) demonstrates that SRH measures can effectively predict future medical care utilization and medical expenditures, typically approaching but not quite reaching the performance of claims-based algorithms. To our knowledge, there is no analogous research examining the performance of SRH measures among the national adult Medicaid population.

For SRH measures to be a viable case-finding alternative for the Medicaid population, two conditions must be met: (1) it must be feasible for Medicaid programs to collect and analyze the input data; and (2) the resulting model must be sufficiently predictive of high health care use. Fortunately, there are precedents that suggest the feasibility of states implementing self-reported health risk assessments (HRAs) for Medicaid beneficiaries. Agencies in states as diverse as Delaware, Maine, Maryland, Oklahoma, and Washington have collected SRH measures as part of the application or enrollment process for subsets of their adult members (Rawlings-Sekunda, Curtis, and Kaye 2001; Delaware Division of Medicaid and Medical Assistance 2010; Washington State Health Insurance Pool 2012; Maryland Department of Health and Mental Hygiene 2013). Often Medicaid HRAs are administered by MCOs, who have considerable experience collecting and using these data for case-finding purposes (America's Health Insurance Plans's 2005).

Furthermore, a recent evaluation of an initiative in Wisconsin Medicaid to collect SRH measures demonstrated their success in prospectively classifying high-need beneficiaries among a nonelderly adult population. The Wisconsin Medicaid program embedded a brief SRH screener in its application

for childless adult coverage with the express purpose of using these data to help identify beneficiaries likely to need, and use, costly health care (Wisconsin Department of Health Services 2008). An evaluation of the screener's performance in predicting high utilization in the first year of coverage found predictive ability comparable to that of claims-based algorithms (Leininger et al. 2014).

The findings from the Wisconsin Medicaid experience, in combination with the existing evidence of the predictive power of SRH measures in other populations, substantiate the promise of using SRH-based predictive algorithms to identify high-need beneficiaries within the ACA Medicaid expansion population. While recent systematic reviews have suggested ideal SRH measures for diagnostic purposes (Goetzel et al. 2011; Oremus, Hammill, and Raina 2011), no standard exists regarding the optimal SRH measures for predictive purposes. In sum, there is little evidence to guide states regarding *which* SRH measures or combination of measures will effectively predict high health care use among Medicaid-eligible adults.

Contribution

This article constitutes a first step toward establishing a predictive standard for case-finding applications for adult Medicaid members who lack recent claims history. To our knowledge, it is the first article to assess which SRH measures are meaningfully predictive of high health care utilization among a nationally representative sample of Medicaid-eligible adults. The current lack of a predictive standard likely arises in part from the proprietary nature of HRAs administered by MCOs and other primary vendors (Rothstein and Harrell 2009). The overarching goal for this article is to provide a predictive standard that is informed by scientific theory, subject to rigorous empirical assessment, and made freely available in the public domain.

To do so, we identify and test a series of promising constructs identified from a thorough review of the scientific literature. Contrasting with the typical approach of examining one construct in isolation (e.g., Fleishman and Cohen 2010) or all constructs as a whole (e.g., Perrin et al. 2011), we compare the incremental predictive ability of a variety of candidate SRH constructs. The results quantify the expected gain in predictive ability that the addition of each construct (or set of measures) confers. This approach helps concretize the trade-offs between predictive capacity and respondent burden that Medicaid programs will inevitably face in the design of an SRH tool.

METHODS

Data and Sample

Our sample is drawn from the 1997–2008 rounds of the National Health Interview Survey (NHIS) linked with the 1998–2009 rounds of the Medical Expenditure Panel Survey (MEPS). The MEPS follows a subsample of NHIS participants longitudinally; accordingly, we use the cross-sectional data collected in the NHIS to create our baseline predictors and the first year of MEPS participation as the follow-up period over which utilization and expenditures are calculated (see Figure A1 for a graphical depiction). In the analysis that follows, we employ analytic weights that account for pooling across years in addition to differential sample selection probabilities and nonresponse. When weighted, data from the MEPS provide nationally representative estimates of health care utilization, expenditures, and insurance coverage for the civilian, noninstitutionalized population.

The analytic sample is restricted to individuals who served as sample adults in the NHIS and have valid data for both surveys. We define the analytic sample to resemble the adult population gaining Medicaid coverage under the ACA expansions: U.S. citizens of the ages 19–64 with family incomes below 138 percent of the federal poverty line. We also exclude from the sample any women who are currently pregnant, as well as any individuals reporting receipt of Medicare or Supplemental Security Income benefits, as the Medicaid eligibility pathways for these groups are historically distinct. All exclusions are made based on characteristics reported in the NHIS survey. Finally, we exclude individuals with missing data for any predictor or outcome measures, which totaled 615 observations (or 8 percent of remaining observations). We are left with a final analytic sample of 6,725 low-income adults.

Measures

The utilization outcomes of interest are emergency room (ER) visits and hospitalizations, as these types of health care have long served as the focus of care management efforts and subsets of both (especially ambulatory-sensitive ER visits and hospital readmissions) are emerging as potential health care performance indicators (Ash and Ellis 2012). As discussed above, Medicaid care management initiatives often seek to target the highest cost users (Weir, Aweh, and Clark 2008); accordingly, we examine the incurring of high costs as an additional outcome of interest.

We operationalize the ER and hospitalization outcome measures by creating binary variables that indicate if the respondent had at least one inpatient visit or if s/he had at least two ER visits during the first year of the MEPS, both of which approximate membership in the top utilization decile. Specifically, during the first year of MEPS participation, 9.47 percent of the sample experienced one or more inpatient visits, and 7.18 percent experienced two or more ER visits. Analogously, we construct an indicator of membership in the top decile of total annual health care expenditures, as defined within each panel in the study sample. This corresponds to roughly $7,833 (in 2011 dollars) or more in health care expenditures during the year.

We consider the predictive ability of seven different sets of predictors (see Table 1). Our categorization groups domains identified as potentially important risk factors for elevated health care utilization. For each domain, we select measures of risk based on their potential predictive performance as assessed through careful literature review, as well as their feasibility of collection within standard Medicaid application and enrollment processes. Each of our primary specifications has fewer than 25 predictors to ensure adequate power to detect differences in predictive capacity of our models as recommended by Harrell, Lee, and Mark (1996).[1]

The first set of predictors is a standard set of *baseline characteristics* that are currently collected by many Medicaid systems. This set includes age (entered into all models with an additional squared term); sex; family structure (married, number of adults in family, number of children in family); geographic region; and family income (0–50 percent FPL, 51–100 percent FPL, 101–138 percent FPL). We impute family income for individuals missing this information using imputation files provided by the National Center for Health Statistics (NCHS). We also include indicators of public or private health insurance coverage during the first year of the MEPS in the baseline set of characteristics.

The next set of predictors indicates the presence of *health conditions* that are established predictors of future health costs and utilization (e.g., Fleishman and Cohen 2010). We include self-reported conditions identified by Chaudhry, Jin, and Meltzer (2005) that approximate the enumerated conditions of the Charlson Comorbidity Index (CCI) (Charlson et al. 2008). Dummy variables indicate respondent report of eight different conditions. Two CCI-related conditions (arthritis and HIV) are not included due to their inconsistent availability in the data; we later examine sensitivity to their exclusion as well as alternative formulations of the enumerated health conditions

Table 1: Domains of Predictors

Domain	*Measures*	*Adapted from*
(1) Baseline characteristics	Age	Leininger et al. (2014)
	Sex	
	Marital status	
	Number of adults in family	
	Number of children in family	
	Geographic region	
	Family income	
(2) Health conditions	Asthma, emphysema, or chronic bronchitis	Chaudhry, Jin, and Meltzer (2005)
	Cancer	
	Diabetes	
	Ulcer	
	Heart trouble	
	Weak or failing kidneys	
	Liver condition	
	Stroke	
(3) Mental health	Serious mental illness	Kessler et al. (2003)
(4) Access to health care	Prior health insurance coverage	Andersen (1995)
	Uninsured during last year	
	Usual source for sick care	
(5) Health-related behaviors	Tobacco use	Zarkin et al. (2004)
	Excessive alcohol use	
	Obesity and overweight	
(6) Health-related quality of life	Health and Activity Limitation Index	Erickson, Wilson, and Shannon (1995)
(7) Prior health care utilization	Emergency room care	Naessens et al. (2005)
	Overnight hospital stays	
	10+ health care visits	

that rely upon weighted scoring across conditions (Elixhauser et al. 1998; Chaudhry, Jin, and Meltzer 2005; Austin et al. 2013).

Next, we rely on the Kessler-6 psychological distress scale to construct a measure of *mental health.* Although the Kessler-6 scale is not intended to identify specific disorders, previous work has demonstrated its utility as a screening tool for both depression and anxiety disorders (Gill et al. 2007), which have been shown to be highly predictive of medical care utilization (Manning, Newhouse, and Ware 1982). This is consistent with a large body of research demonstrating the association between comorbid mental illness with relatively greater frequency and intensity or cost of health care use (Egede, Zheng, and Simpson 2002; Himelhoch et al. 2004; Li et al. 2008). Consistent with prior research (e.g., Strine et al. 2005; Kling, Liebman, and Katz 2007), we

use an indicator of serious mental illness that is derived from the composite score on individual responses on the Kessler-6 instrument (Kessler et al. 2003).

The next set of predictors provides information on *access to health care*. A large body of work indicates that health insurance coverage and access to health care affect medical care utilization (e.g., Andersen 1995; Babitsch, Gohl, and von Lengerke 2012; Finkelstein et al. 2012). Included measures indicate whether the respondent had health insurance and the type of coverage (public, private, and uninsured); whether he or she was uninsured at some point during the last 12 months; and whether he or she has a usual source for sick care. We include information on uninsurance spells as prior work shows that intrayear gaps in insurance coverage may have important implications for the utilization of medical care (Banerjee, Ziegenfuss, and Shah 2010). Each of these variables measures the respondent's status at baseline (i.e., the NHIS survey interview).

The next block includes self-reported measures of *health-related behaviors* associated with increased health care utilization. This category of risk factors includes indicators of obesity and overweight (Sturm 2002; Finkelstein, Fiebelkorn, and Wang 2003; Thorpe et al. 2004; Finkelstein et al. 2009), smoking (Miller, Ernst, and Collin 1999; Sturm 2002; Bertakis and Azari 2006), and alcohol abstention (Rice et al. 2000; Polen et al. 2001; Zarkin et al. 2004). We subsequently consider a predictor block reflecting *health-related quality of life* (HRQOL), a construct including activity limitations and perceived health status that is associated with high future health care expenditures (Fleishman et al. 2006). The Health and Activities Limitation Index (HALex) is a single score index developed by NCHS to estimate a person's overall HRQOL, with values ranging from 1.00 for individuals in excellent health with no activity limitations to 0.10 for those in poor health with more severe limitations (Erickson, Wilson, and Shannon 1995). Activity limitation is defined using questions about a person's ability to perform activities of daily living, instrumental activities of daily living, major activities in terms of work or housework, and any other reported limitations.

Our final domain includes information on *prior health care utilization*, where "prior" refers to health care use reported in the NHIS during the baseline year. Measures of prior year utilization collected on the Wisconsin HNA proved strongly predictive of high health care use (Leininger et al. 2014). This is consistent with prior research demonstrating the persistence of high medical spending (Cohen, Ezzati-Rice, and Yu 2006) and the reliance of claims-based algorithms on the prior year's health care

encounters (e.g., Kronick et al. 2000). We include three measures of prior year utilization based on questions in the NHIS that indicate whether the respondent had at least one visit to the ER, stayed in the hospital for at least one night, and received health care from professionals at least 10 times during the past 12 months. We consider the latter measure to indicate high outpatient utilization, which is similar to the approach taken by Naessens et al. (2005). Additional details on how predictor measures are operationalized are available in the Appendix (Table A1).

Analyses

We begin by examining the incremental performance of each domain of predictor variables. We use logistic regression to estimate a series of models for the outcome variables that add each domain to a baseline model. In keeping with standard practice, model building and testing are performed using a split-sample approach. We randomly divide the full sample into two and use the first subsample to fit each model; the resulting parameter estimates are then applied to the second subsample to compute measures of performance.

We assess the incremental predictive ability of each augmented model by comparing the *c*-statistic and discrimination slope with those of the baseline model. While the *c*-statistic (or area under the receiving operating curve) is the more traditional measure of model discrimination, measurements of integrated discrimination improvement (IDI) have been proposed as an advantageous alternative when quantifying improvements in model performance. Equivalent to the improvement in discrimination slopes between a new and existing model, the IDI is a more sensitive measure than the c-statistic and offers a direct interpretation—an increase in the difference in mean predicted probabilities of events and nonevents (Pencina et al. 2012). We compute confidence intervals for the *c*-statistic and discrimination slope using a 500 replicate bootstrap procedure. We also bootstrap the difference in these statistics between each new model and the baseline model to determine the statistical significance of the incremental gain in predictive performance.

After assessing the incremental predictive performance of each domain, we examine different combinations of the most promising domain candidates. For the best-performing models, we further evaluate their performance by estimating the sensitivity, specificity, and predictive values when using different risk thresholds, or cutoffs of predicted probability, to define a likely future high-need case.

RESULTS

Descriptive Analysis

Table 2 presents descriptive statistics for the full analytic sample, as well as those classified as "high need" in the follow-up period under each of our outcome measures. Overall, the demographic characteristics of the full sample and the high users are somewhat similar. Health and health behaviors, however, are very different for sample members with high medical utilization when compared to the full sample. Larger proportions of high users report the presence of health conditions and are assessed to have serious mental illness. The HALex scores are also lower on average, indicating poor HRQOL. In terms of behavioral risk factors, larger proportions of high-need individuals are obese and have a history of smoking, although the share that abstains from alcohol is smaller than is reported in the full sample.

Previous interactions with the health care system are also distinct for those with high health care needs. The proportion with a usual source for sick care is high, although reports of health insurance coverage vary by outcome. The majority with high ER use reports being uninsured during the past 12 months, while a smaller proportion of top spenders reports being uninsured. Prior health care utilization is consistently elevated for all high users with a higher proportion reporting an ER visit, overnight hospitalization, or 10+ health care visits in the last 12 months when compared to the full sample.

Multivariate Analysis

We first examine the incremental predictive performance of models that include each domain of predictors separately (Models 1–7). The *c*-statistics for the baseline model (Model 1) range from 0.60 to 0.69 for the three outcome variables (Table 3). Subsequent models reveal higher *c*-statistics, with the highest values occurring under the model incorporating past year utilization measures (Model 7). The additions of the HRQOL (Model 6) and health conditions (Model 2) domains provide the next greatest incremental increases over Model 1, although the rank ordering varies by outcome. The changes in the *c*-statistic for models featuring all other domains are smaller and not statistically significant.

Comparing the discrimination slopes across models reveals a similar story. Table 4 confirms that Model 7 yields the largest incremental gain over Model 1. In general, Models 2 and 6 have the next largest incremental

improvements. While the health behaviors (Model 5) and the mental health domains (Model 3) also provide statistically significant gains, they are typically of a more modest magnitude.

Models with combinations of the three best-performing domains (Models 8–11) are examined next. Across performance indicators and outcomes, models including prior year's utilization (Models 9, 10, and 11) perform the best. For the outcome of high ER use, the model with all three of the top domains (Model 11) has the highest *c*-statistic. When predicting high costs, Models 10 and 11 perform equally well. The ordering of *c*-statistics differs slightly for the hospitalization outcome with Model 10 providing the largest incremental improvement over Model 1. However, when comparing discrimination slopes across models and outcomes, Model 11 consistently yields the largest incremental gain.

To further evaluate the best-performing models (Models 9, 10, and 11), we calculate the sensitivity, specificity, and positive and negative predictive values under each model by risk threshold (Table 5). Across all outcome measures, we see performance improvement under the augmented models at each risk threshold when compared to the baseline model. For the most part, Models 9, 10, and 11 perform comparably well. Model 11 is the most successful in identifying high ER and high-cost members at each risk threshold. However, the preferred model for predicting hospitalizations appears to depend on the risk threshold in use.

Sensitivity Analysis

Our study sample is comprised of individuals both with and without insurance coverage when we examine their utilization in the MEPS. Given our interest in the behavior of this population when insured, and the possibility that prediction models may perform differently for those with and without health insurance, we rerun the analysis when restricting the sample to individuals with insurance (see Table A2). Our findings are similar to those in the analysis above with the best-performing models including the health conditions, HRQOL, and prior utilization domains.

As mentioned earlier, we examine several alternative formulations of the health conditions domain. Specifications employing a simple count of the total number of health conditions perform comparably to those employing eight separate dummy variables reflecting the enumerated conditions. We also test the performance of a weighted CCI comorbidity score following the methodology of Chaudhry, Jin, and Meltzer (2005). The CCI model performs

Table 2: Descriptive Statistics for Analytic Sample

	Full Sample	*Sample Defined by MEPS Outcome Variable*		
	N = 6,725	*Two + ER Visits N = 483*	*Any Hospitalization N = 637*	*Top Expenditure Decile N = 642*
NHIS predictor variables (%, unless otherwise stated)				
(1) Baseline characteristics (administrative data)				
Age in years, mean (SD)	35.97 (14.65)	34.96 (13.97)	37.51 (16.06)	41.01 (15.60)
Gender, male	38.34	28.13	25.35	28.21
Married	23.17	17.11	24.67	21.54
Number of adults in family, mean (SD)	1.55 (0.88)	1.52 (0.91)	1.58 (0.97)	1.53 (0.92)
Number of children in family, mean (SD)	1.00 (1.49)	1.14 (1.52)	1.09 (1.50)	0.80 (1.26)
Income (% FPL)				
0–50%	29.58	33.46	30.24	28.05
51–100%	36.09	38.55	35.41	34.97
101–138%	33.74	26.93	34.28	35.79
(2) Health conditions/comorbidities				
Asthma, emphysema or chronic bronchitis	17.44	33.13	21.28	24.36
Cancer	4.43	5.48	6.83	8.48
Diabetes	4.43	9.25	10.49	13.60
Ulcer	8.86	17.79	13.00	14.33
Heart trouble	7.69	12.14	14.36	16.41
Weak or failing kidneys	1.19	2.77	3.26	3.58

Continued

Table 2. *Continued*

	Full Sample	*Sample Defined by MEPS Outcome Variable*		
	N = 6,725	*Two + ER Visits N = 483*	*Any Hospitalization N = 637*	*Top Expenditure Decile N = 642*
Liver condition	1.83	3.00	4.30	4.64
Stroke	1.31	3.12	4.29	4.57
(3) Mental health				
Serious mental illness	6.51	12.18	11.33	14.81
(4) Access to health care				
Health insurance coverage preceding enrollment				
Private only	39.29	22.42	30.55	41.84
Public	20.88	34.43	30.68	30.06
Uninsured	39.83	43.15	38.76	28.10
Has been uninsured at any time during past 12 months	48.97	56.86	51.30	39.86
Has usual source for sick care	74.81	78.02	81.50	88.15
(5) Health-related behaviors				
Smoking status				
Current smoker	39.97	58.12	43.65	40.34
Former smoker	12.18	9.54	14.09	15.22
Never smoked	47.85	32.34	42.25	44.44
Alcohol consumption				
Abstainers	38.19	40.43	49.04	47.40
Light and medium drinkers	37.69	35.58	33.27	35.30
Heavy drinkers	24.12	23.98	17.69	17.30
Body weight				
Underweight	2.22	2.27	1.73	1.75

Continued

Table 2. *Continued*

	Full Sample	*Sample Defined by MEPS Outcome Variable*		
	N = 6,725	*Two + ER Visits N = 483*	*Any Hospitalization N = 637*	*Top Expenditure Decile N = 642*
Normal	40.58	31.62	33.58	32.63
Overweight	29.50	25.88	29.63	30.31
Obese	27.70	40.23	35.05	35.30
(6) Health-related quality of life				
HALex, mean (SD)	0.82 (0.23)	0.71 (0.30)	0.72 (0.30)	0.67 (0.31)
(7) Prior health care utilization				
Had visit to hospital ER in past yr	27.50	52.90	44.84	44.23
Stayed in hospital overnight in past yr	10.17	19.32	23.26	24.88
Had 10+ health care visits in past yr	13.42	25.29	28.27	39.77
MEPS health care utilization outcome variables (calculated over first year of MEPS participation)				
Any ER visits	21.39	100.00	55.64	46.93
Two or more ER visits	6.51	100.00	22.84	18.25
Any inpatient discharges	8.61	30.22	100.00	51.51
Expenditures, mean (SD)	3,666 (11,938)	9,965 (21,148)	19,840 (33,327)	23,253 (28,614)

Note. HALex = Health and Activity Limitation Index. Medical expenditures put in 2011 $ using the GDP price index per AHRQ recommendation. Region dummies not shown.

Table 3: Discriminative Ability by Type of Predictor: *c*-statistic

	High ER	*Any Hospitalization*	*High Cost*
Model 1: Baseline	0.600 (0.555, 0.649)	0.663 (0.606, 0.678)	0.690 (0.647, 0.729)
Model 2: Baseline + conditions	0.654** (0.599, 0.685)	0.686* (0.629, 0.701)	0.705 (0.664, 0.743)
Model 3: Baseline + mental health	0.603 (0.564, 0.658)	0.664 (0.610, 0.683)	0.699 (0.634, 0.742)
Model 4: Baseline + access	0.610 (0.568, 0.658)	0.669 (0.595, 0.669)	0.693 (0.653, 0.734)
Model 5: Baseline + behaviors	0.629 (0.567, 0.660)	0.674 (0.609, 0.683)	0.693 (0.648, 0.736)
Model 6: Baseline + HRQOL	0.644** (0.600, 0.690)	0.697** (0.633, 0.704)	0.729** (0.687, 0.772)
Model 7: Baseline + prior utilization	0.676** (0.615, 0.705)	0.706** (0.652, 0.732)	0.746** (0.706, 0.780)
Model 8: Baseline + conditions + HRQOL	0.671** (0.610, 0.697)	0.698** (0.644, 0.711)	0.730** (0.682, 0.770)
Model 9: Baseline + conditions + prior utilization	0.697** (0.625, 0.718)	0.706** (0.657, 0.735)	0.746** (0.706, 0.783)
Model 10: Baseline + HRQOL + prior utilization	0.694** (0.627, 0.715)	0.716** (0.654, 0.733)	0.751** (0.714, 0.787)
Model 11: Baseline + conditions + HRQOL + prior utilization	0.703** (0.631, 0.714)	0.711** (0.653, 0.733)	0.751** (0.710, 0.793)

Note. $N = 3{,}363$ for validation sample. Confidence intervals of 95% in parentheses.
*Indicates that *c*-statistic differs from base model at $p < .05$; **Indicates that *c*-statistic differs from base model at $p < .01$.

Table 4: Discriminative Ability by Type of Predictor: Discrimination Slope

	High ER	*Any Hospitalization*	*High Cost*
Model 1: Baseline	0.016 (0.008, 0.025)	0.034 (0.025, 0.045)	0.059 (0.045, 0.072)
Model 2: Baseline + conditions	0.033** (0.021, 0.045) *106%*	0.050** (0.038, 0.063) *47%*	0.094** (0.073, 0.114) *59%*
Model 3: Baseline + mental health	0.020* (00.011, 0.028) *25%*	0.038* (0.027, 0.049) *12%*	0.072** (0.056, 0.089) *22%*
Model 4: Baseline + access	0.018 (0.010, 0.026) *13%*	0.038 (0.027, 0.049) *12%*	0.063 (0.048, 0.079) *7%*
Model 5: Baseline + behaviors	0.031** (0.019, 0.044) *94%*	0.038** (0.028, 0.047) *12%*	0.064** (0.049, 0.079) *8%*
Model 6: Baseline + HRQOL	0.030** (0.020, 0.040) *88%*	0.047** (0.035, 0.059) *38%*	0.107** (0.087, 0.129) *81%*
Model 7: Baseline + prior utilization	0.039** (0.028, 0.050) *144%*	0.057** (0.045, 0.071) *68%*	0.125** (0.102, 0.151) *112%*
Model 8: Baseline + conditions + HRQOL	0.038** (0.025, 0.051) *138%*	0.057** (0.043, 0.071) *68%*	0.118** (0.094, 0.143) *100%*
Model 9: Baseline + conditions + prior utilization	0.047** (0.035, 0.061) *194%*	0.067** (0.050, 0.083) *97%*	0.135** (0.112, 0.163) *129%*
Model 10: Baseline + HRQOL + prior utilization	0.043** (0.032, 0.057) *169%*	0.062** (0.049, 0.078) *82%*	0.142** (0.116, 0.169) *141%*
Model 11: Baseline + conditions + prior utilization + HRQOL	0.049** (0.036, 0.062) *206%*	0.070** (0.056, 0.084) *106%*	0.146** (0.122, 0.173) *147%*

Note. $N = 3{,}363$ for validation sample. Confidence intervals of 95% in parentheses. Integrated discrimination improvement calculated as % improvement in discrimination slope over base model in italics.
*Indicates that discrimination slope differs from base model at $p < .05$; **Indicates that discrimination slope differs from base model at $p < .01$.

Table 5: Sensitivity, Specificity, and Predictive Values by Risk Threshold

	High ER			*Any Hospitalization*			*High Cost*		
	50th Percentile	*75th Percentile*	*90th Percentile*	*50th Percentile*	*75th Percentile*	*90th Percentile*	*50th Percentile*	*75th Percentile*	*90th Percentile*
Model 1: Baseline									
Sensitivity	0.632	0.390	0.186	0.721	0.474	0.189	0.749	0.490	0.237
Specificity	0.509	0.759	0.906	0.521	0.771	0.909	0.525	0.775	0.913
Positive predictive value	0.081	0.100	0.118	0.126	0.166	0.166	0.139	0.182	0.219
Negative predictive value	0.953	0.948	0.942	0.951	0.938	0.921	0.953	0.937	0.921
Model 9: Baseline + conditions + prior utilization									
Sensitivity	0.750	0.521	0.266	0.761	0.503	0.278	0.789	0.564	0.372
Specificity	0.516	0.768	0.911	0.525	0.774	0.917	0.530	0.782	0.927
Positive predictive value	0.096	0.133	0.169	0.134	0.176	0.244	0.147	0.210	0.344
Negative predictive value	0.968	0.959	0.948	0.958	0.942	0.930	0.961	0.946	0.935
Model 10: Baseline + HRQOL + prior utilization									
Sensitivity	0.769	0.507	0.262	0.773	0.490	0.255	0.783	0.585	0.375
Specificity	0.518	0.767	0.911	0.526	0.773	0.914	0.529	0.784	0.927
Positive predictive value	0.098	0.129	0.167	0.136	0.172	0.223	0.146	0.217	0.346
Negative predictive value	0.970	0.958	0.948	0.960	0.940	0.927	0.960	0.949	0.935
Model 11: Baseline + conditions + HRQOL + prior utilization									
Sensitivity	0.777	0.540	0.278	0.770	0.518	0.273	0.803	0.592	0.383
Specificity	0.519	0.769	0.912	0.526	0.775	0.916	0.531	0.785	0.929
Positive predictive value	0.099	0.138	0.178	0.135	0.182	0.239	0.149	0.220	0.356
Negative predictive value	0.972	0.961	0.949	0.960	0.944	0.929	0.963	0.949	0.936

Note. Risk threshold refers to the cutoff in predicted probability (50th, 75th, or 90th percentile) used to predict the outcome.

less favorably than specifications with the total number or list of conditions. Restricting the sample to years in which we are able to use the full set of conditions listed in Chaudhry, Jin, and Meltzer leads to no meaningful changes in performance.

In addition, we examine specifications exploring general self-rated health (GSRH) in isolation (contrasting with the specification in which it serves as a component of the HALex), as prior work suggests that the single-item GSRH score has comparable predictive performance to multi-item SRH measures (DeSalvo et al. 2005). In our context, we find that HALex considerably outperforms GSRH alone. We also test two additional measures—the presence of functional limitations and respondent reports of pain—that are absent in the HALex but are similar to additional component measures of the SF-12, a widely-used HRQOL index (Gandek et al. 1998). Their addition to HALex in the HRQOL predictor set does not appreciably change the domain's predictive performance, with the exception of pain improving predictive performance for ER utilization. This suggests that the shorter HALex index may perform comparably to the SF-12 and may even be preferable to states if it lessens respondent burden.

Finally, we estimate a specification that incorporates all predictor domains. We consider this final "kitchen sink" specification as exploratory, however, given that the number of predictors exceeds our chosen maximum limit (37 vs. 25). We find that our final specification with the top three domains performs comparably to the all-domain model.

DISCUSSION

Our findings offer Medicaid programs a promising strategy to prospectively identify high health care users when historical claims data are unavailable as will likely be the case for the ACA Medicaid expansion population. Specifically, we find that SRH measures provide substantial increases in predictive performance relative to the set of baseline administrative characteristics currently collected at application. Across performance indicators, prior health care utilization yields the single largest incremental gain in predictive performance relative to the baseline model among the six SRH domains. However, the addition of any one of the top three performing SRH domains in the study (prior utilization, health conditions, and HRQOL) improves the predictive performance of the baseline model from below to above the Hosmer–Lemeshow rule-of-thumb threshold of acceptability for one or more outcomes (i.e.,

c-statistic >0.70; Hosmer and Lemeshow 2000). When all three domains are included, the model achieves acceptable levels of performance for all outcomes. Importantly, the performance of this specification approaches if not quite meets that of published claims-based models (Weir, Aweh, and Clark 2008).

The collection and use of SRH measures for prospective case-finding is likely to prove appealing to state Medicaid programs anticipating large growth in enrollment under the ACA. Indeed, Iowa, Michigan, and Pennsylvania have recently proposed incorporating HRAs into their programs for the new adult expansion population (Iowa Department of Human Services 2013; Michigan Department of Community Health 2013; Rudowitz, Artiga, and Musumeci 2014). While no data collection or analytic exercise is costless, the SRH strategy that we explore does not require proprietary software or algorithms to implement. Validated SRH measures, such as those employed in this study, are widely available in the public domain (listed in Figure A2). Notably, the most predictive domain in this study includes just three measures—or potential questions on a Medicaid enrollment application. Importantly, the analytic demands of the predictive modeling deployed here are also likely to be within the capabilities of technical policy analysts in state health agencies.

There are several limitations to this study. We examine the utilization behavior of low-income adults meeting the eligibility criteria for optional state Medicaid expansions under the ACA. However, which individuals gain coverage will depend on existing eligibility criteria and state decisions regarding expanded eligibility. It is possible that the demand for health care and its predictors will differ between new Medicaid enrollees and the sample studied here. Furthermore, we excluded adults likely to be eligible for Medicaid under preexisting pathways (e.g., pregnancy and disability), as well as those with missing survey data, who may have different characteristics and health care needs than those captured here. In addition, the analysis is limited to the SRH measures available in the NHIS. Alternative measures within any one of the domains (e.g., mental health, health conditions) may perform better than those tested here. For instance, the aforementioned Wisconsin study (Leininger et al. 2014) found that self-reported depression was both prevalent and predictive of subsequent high ER use and high costs among newly eligible childless adult Medicaid beneficiaries. As already noted, serious, and comorbid, mental illnesses are frequent correlates of elevated health care use and spending. However, the only measure of mental health in the NHIS, the Kessler-6 scale, does not improve the predictive performance of our models. Developed to identify the presence of serious mental illness, prevalence

defined by the Kessler-6 is low in this population (i.e., 6.5 percent) and, for this reason, unlikely to influence the predictive performance of a population-based model. In addition, our selection of health conditions was informed by the CCI, but it may overlook other important chronic health conditions that influence health care utilization.

This study's limitations suggest additional directions for future work. Ongoing work is needed to identify SRH measures, especially alternative mental health measures, which will increase the performance of predictive models. The inclusion of SRH measures in the predictive models yields large relative gains in our capacity to identify Medicaid-eligible adults that would likely become heavy health care users. However, in absolute terms, there remains considerable uncertainty in our capacity to predict high care use beneficiaries, a limitation that is shared with claims-based predictive modeling. At current levels of performance, these predictive models may function most effectively as initial screeners. Medicaid programs might obtain additional information from beneficiaries that screen "positive" to best allocate care management resources. Future research that explores two-stage prediction modeling in parallel with efforts to improve the performance of the initial model is warranted. In addition, examination of the performance of SRH-based screeners in the context of a decision analytic framework (see Leening et al. 2014) that examines the net benefits and costs associated with screening this population would be a valuable next step.

Finally, the marginal difference in the predictive performance between SRH and claims-based strategies is both uncertain and highly relevant to state Medicaid program decisions regarding the development and use of all-payer claims datasets. To date, there has been no direct comparison of the performance of SRH and claims-based approaches within a common Medicaid-eligible population. Medicaid programs may at some point have the luxury to choose between SRH and claims-based data for prospectively identifying high-cost beneficiaries. Research that quantifies the comparative predictive performance of SRH and claims-based alternatives is needed to inform these decisions.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: We gratefully acknowledge funding from the Robert Wood Johnson State Health Access Reform Evaluation (SHARE) program (Leininger); the Robert Wood Johnson Health and Society

Scholars Program (Wherry); and a Career Development Award from the National Institute of Mental Health K01MH092338 (Burns). Our findings do not necessarily represent those of either the Robert Wood Johnson Foundation or NIMH. We have no known conflicts of interest.

Disclosures: None.

Disclaimers: None.

NOTE

1. The guideline is to include no more than $p/10$ predictors in a multivariate prediction model for a binary outcome, with p representing the number of individuals in the model-building subsample in the less frequent outcome category. In our sample, the suggested maximum ranges from 24 predictors for the ER outcome to 32 predictors for the high-cost and inpatient outcomes.

REFERENCES

Advisory Board. 2013. Beyond the Pledges: Where the States Stand on Medicaid [accessed on July 30, 2013]. Available at http://www.advisory.com/DailyBriefing/Resources/Primers/MedicaidMap

America's Health Insurance Plans. 2005. Innovations in Medicaid Managed Care: Health Plan Programs to Improve the Health and Well-Being of Medicaid Beneficiaries [accessed on July 30, 2013]. Available at http://www.ahip.org/Innovations-in-Medicaid-Managed-Care-Report/

Andersen, R. M. 1995. "Revisiting the Behavioral Model and Access to Medical Care? Does It Matter?" *Journal of Health and Social Behavior* 36 (1): 1–10.

Ash, A. S., and R. P. Ellis. 2012. "Risk-Adjusted Payment and Performance Assessment for Primary Care." *Medical Care* 50 (8): 643–53.

Ash, A. S., Y. Zhao, R. P. Ellis, and M. S. Kramer. 2001. "Finding Future High-Cost Cases: Comparing Prior Cost versus Diagnosis-Based Methods." *Health Services Research* 36 (6 Pt 2): 194–206.

Austin, S. R., Y. N. Wong, R. G. Uzzo, J. R. Beck, and B. L. Egleston. 2013. "Why Summary Comorbidity Measures Such as the Charlson Comorbidity Index and Elixhauser Score Work." *Medical Care* 2013 May 23 epub ahead of print.

Babitsch, B., D. Gohl, and T. von Lengerke. 2012. "Re-Revisiting Andersen's Behavioral Model of Health Services Use: A Systematic Review of Studies from 1998-2011." *Psychosocial Medicine* 9: 2012; 9: Doc 11. Epub 2012, Oct 25.

Banerjee, R., J. Y. Ziegenfuss, and N. D. Shah. 2010. "Impact of Discontinuity in Health Insurance on Resource Utilization." *BMC Health Services Research* 10: 195.

Bertakis, K. D., and R. Azari. 2006. "The Influence of Obesity, Alcohol Abuse, and Smoking on Utilization of Health Care Services." *Family Medicine* 38 (6): 427–34.

Bodenheimer, T., and R. Berry-Millett. 2009. *Care Management of Patients with Complex Health Care Needs*. Research Synthesis Report No. 19. Princeton, NJ: Robert Wood Johnson Foundation.

Brown, R. 2009. The Promise of Care Coordination: Models that Decrease Hospitalizations and Improve Outcomes for Medicare Beneficiaries with Chronic Illnesses. Commissioned by the National Coalition on Care Coordination [accessed on July 30, 2013]. Available at http://www.nyam.org/social-work-leadership-institute/docs/N3C-Promise-of-Care-Coordination.pdf

Center for Medicare and Medicaid Services. 2012. Data Analysis Considerations to Inform Medicaid Health Home Program Design. Technical Assistance Brief [accessed on July 30, 2013]. Available at http://www.medicaid.gov/State-Resource-Center/Medicaid-State-Technical-Assistance/Health-Homes-Technical-Assistance/Downloads/HH-Data-Analysis-Brief-6_11_12.pdf

Centers for Medicare and Medicaid Services. 2010. "Re: Health Homes for Enrollees with Chronic Conditions." State Medicaid Director Letter #10-024, November 16 [accessed on August 13, 2013]. Available at http://downloads.cms.gov/cms-gov/archived-downloads/SMDL/downloads/SMD10024.pdf

Charlson, M. E., R. E. Charlson, J. C. Peterson, S. S. Marinopoulos, W. M. Briggs, and J. P. Hollenberg. 2008. "The Charlson Comorbidity Index Is Adapted to Predict Costs of Chronic Disease in Primary Care Patients." *Journal of Clinical Epidemiology* 61: 1234–40.

Chaudhry, S., L. Jin, and D. Meltzer. 2005. "Use of a Self-Generated Charlson Comorbidity Index for Predicting Mortality." *Medical Care* 43 (6): 607–15.

Cohen, S. B., T. Ezzati-Rice, and W. Yu. 2006. "The Utility of Extended Longitudinal Profiles in Predicting Future Health Care Expenditures." *Medical Care* 44 (5 Suppl): I-45–53.

Congressional Budget Office. 2013. Table 1. CBO's May 2013 Estimate of the Effects of the Affordable Care Act on Health Insurance Coverage [accessed on July 30, 2013]. Available at http://www.cbo.gov/sites/default/files/cbofiles/attachments/44190_EffectsAffordableCareActHealthInsuranceCoverage_2.pdf

Delaware Division of Medicaid and Medical Assistance. 2010. Delaware Medicaid and CHIP Managed Care Quality Strategy [accessed on July 30, 2013]. Available at http://regulations.delaware.gov/register/november2010/proposed/CHIP.pdf

DeSalvo, K. B., V. S. Fan, M. B. McDonell, and S. D. Fihn. 2005. "Predicting Mortality and Healthcare Utilization with a Single Question." *Health Services Research* 40 (4): 1234–46.

DeSalvo, K. B., T. M. Jones, J. Peabody, J. McDonald, S. Fihn, V. Fan, J. He, and P. Muntner. 2009. "Health Care Expenditure Prediction with a Single Item, Self-Rated Health Measure." *Health Services Research* 47 (4): 440–7.

Egede, L. E., D. Zheng, and K. Simpson. 2002. "Comorbid Depression Is Associated with Increased Health Care Use and Expenditures in Individuals with Diabetes." *Diabetes Care* 25 (3): 464–70.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. 1998. "Comorbidity Measures for Use with Administrative Data." *Medical Care* 36 (1): 8–27.

Erickson, P., R. Wilson, and I. Shannon. 1995. "Years of Healthy Life." *Healthy People 2000 Statistical Notes* 7: 1–15.

Finkelstein, E. A., I. C. Fiebelkorn, and G. Wang. 2003. "National Medical Sepdning Attributable to Overweight and Obesity: How Much, and Who's Paying?" *Health Affairs* 22 (w3): 219–26.

Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz. 2009. "Annual Medical Spending Attributable to Obesity: Payer-and Service-Specific Estimates." *Health Affairs* 28 (5): w822–31.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, and K. Baicker, and the Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127(3): 1057–106.

Fleishman, J. A., and J. W. Cohen. 2010. "Using Information on Clinical Conditions to Predict High-Cost Patients." *Health Services Research* 45 (2): 532–52.

Fleishman, J. A., J. W. Cohen, W. G. Manning, and M. Kosinski. 2006. "Using the SF-12 Health Status Measure to Improve Predictions of Medical Expenditures." *Medical Care* 44 (5 Suppl): I54–63.

Gandek, B., J. E. Ware, N. K. Aaronson, G. Apolone, J. B. Bjorner, J. E. Brazier, M. Bullinger, S. Kaasa, A. Leplege, L. Prieto, and M. Sullivan. 1998. "Cross-Validation of Item Selection and Scoring for the SF-12 Health Survey in Nine Countries." *Journal of Clinical Epidemiology* 51 (11): 1171–8.

Gill, S. C., P. Butterworth, B. Rodgers, and A. Mackinnon. 2007. "Validity of the Mental Health Component Scale of the 12-item Short-Form Health Survey (MCS-12) as Measure of Common Mental Health Disorders in the General Population." *Psychiatry Research* 152 (1): 63–71.

Goetzel, R. Z., P. Staley, L. Ogden, P. Stange, J. Fox, J. Spangler, M. Tabrizi, M. Beckowski, N. Kowlessar, R. E. Glasgow, M. V. Taylor, and C. Richards. 2011. A Framework for Patient-Centered Health Risk Assessments—Providing Health Promotion and Disease Prevention Services to Medicare Beneficiaries. Centers for Disease Control and Prevention [accessed on July 30, 2013]. Available at http://www.cdc.gov/policy/opth/hra/

Harrell, F. E., K. L. Lee, and D. B. Mark. 1996. "Tutorial in Biostatistics. Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors." *Statistics in Medicine* 15 (4): 361–87.

Hilltop Institute (as the Center for Health Program Development and Management) and Actuarial Research Corporation. 2003. *A Guide to Implementing a Health-Based Risk-Adjusted Payment System for Medicaid Managed Care Programs.* Submitted to the Centers for Medicare and Medicaid Services.

Himelhoch, S., W. E. Weller, A. W. Wu, G. F. Anderson, and L. A. Cooper. 2004. "Chronic Medical Illness, Depression, and Use of Acute Medical Services among Medicare Beneficiaries." *Medical Care* 42 (6): 512–21.

Hosmer, D. W., and S. Lemeshow. 2000. *Applied Logistic Regression*, 2d Edition. New York: John Wiley & Sons Inc.

Iowa Department of Human Services. 2013. Iowa Wellness Plan 1115 Waiver Application. Waiver proposal submitted to the Centers for Medicare and Medicaid Services, August.

Kaiser Family Foundation. 2011. Medicaid's New Health Home Option [accessed on July 30, 2013]. Available at http://kaiserfamilyfoundation.files.wordpress.com/2013/01/8136.pdf

Kessler, R. C., P. R. Barker, L. J. Colpe, J. F. Epstein, J. C. Gfroerer, E. Hiripi, M. J. Howes, S. T. Normand, R. W. Manderscheid, E. E. Waters, and A. M. Zaslavsky. 2003. "Screening for Serious Mental Illness in the General Population." *Archives of General Psychiatry* 60: 184–9.

Kim, S. E., C. Michalopoulous, R. M. Kwong, A. Warren, and M. S. Manno. 2013. "Telephone Care Management's Effectiveness in Coordinating Care for Medicaid Beneficiaries in Managed Care: A Randomized Controlled Study." *Health Services Research*. doi:10.1111/1475-6773.12060.

Kling, J. R., J. B. Liebman, and L. F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Knutson, D., M. Bella, and K. Llanos. 2009. Predictive Modeling: A Guide for State Medicaid Purchasers. Center for Health Care Strategies, Inc. [accessed on July 30, 2013]. Available at http://www.chcs.org/usr_doc/Predictive_Modeling_Guide.pdf

Kronick, R., T. Gilmer, T. Dreyfus, and L. Lee. 2000. "Improving Health-Based Payment for Medicaid Beneficiaries: CDPS." *Health Care Financing Review* 21 (3): 29–64.

Leening, M. J. G., M. M. Vedder, J. C. M. Witteman, M. J. Pencina, and E. W. Steyerberg. 2014. "Net Reclassification Improvement: Computation, Interpretation, and Controversies." *Annals of Internal Medicine* 160: 122–31.

Leininger, L. J., D. Friedsam, K. Voskuil, and T. DeLeire. 2014. "Predicting High-Need Cases among New Medicaid Enrollees." *American Journal of Managed Care*. Forthcoming.

Li, Y., L. G. Glance, X. Cai, and D. B. Mukamei. 2008. "Mental Illness and Hospitalization for Ambulatory Care Sensitive Medical Conditions." *Medical Care* 46 (12): 1249–56.

Manning, W. G., Jr, J. P. Newhouse, and J. E. Ware, Jr. 1982. "The Status of Health in Demand Estimation; or, Beyond Excellent, Good, Fair and Poor." In *Economic Aspects of Health*, edited by V. R. Fuchs, pp. 141–84. Chicago, IL: University of Chicago Press.

Maryland Department of Health and Mental Hygiene. 2013. Health Service Needs Information [accessed on July 30, 2013]. Available at https://mmcp.dhmh.maryland.gov/docs/hra-form-rev0699.PDF

Michigan Department of Community Health. 2013. Healthy Michigan Plan. Waiver proposal submitted to the Centers for Medicare and Medicaid Services, November.

Miller, V. P., C. Ernst, and F. Collin. 1999. "Smoking-Attributable Medical Care Costs in the USA." *Social Science and Medicine* 48 (3): 375–91.

Naessens, J. M., M. A. Baird, H. K. Van Houten, D. J. Vanness, and C. R. Campbell. 2005. "Predicting Persistently High Primary Care Use." *Annals of Family Medicine* 3 (4): 324–30.

Oremus, M., A. Hammill, and P. Raina. 2011. Health Risk Appraisal: Technology Assessment Report. Agency for Healthcare Research and Quality Technology Assessment Program [accessed on July 30, 2013]. Available at http://www.cms.gov/Medicare/Coverage/DeterminationProcess/downloads/id79ta.pdf

Pencina, M. J., R. B. D'Agostino, K. M. Pencina, A. C. J. W. Janssens, and P. Greenland. 2012. "Interpreting Incremental Value of Markers Added to Risk Prediction Models." *American Journal of Epidemiology* 176 (6): 473–81.

Perrin, N. A., M. Stiefel, D. M. Mosen, A. Bauck, E. Shuster, and E. M. Dirks. 2011. "Self-Reported Health and Functional Status Information Improves Prediction of Inpatient Admissions and Costs." *American Journal of Managed Care* 17 (12): e472–8.

Polen, M. R., C. A. Green, D. K. Freeborn, J. P. Mullooly, and F. Lynch. 2001. "Drinking Patterns, Health Care Utilization, and Costs among HMO Primary Care Patients." *The Journal of Behavioral Health Services and Research* 28 (4): 378–99.

Rawlings-Sekunda, J., D. Curtis, and N. Kaye. 2001. Emerging Practices in Medicaid Primary Care Case Management Programs. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services [accessed on July 30, 2013]. Available at http://aspe.hhs.gov/health/reports/pccm/index.htm#top

Rice, D. P., C. Conell, C. Weisner, E. M. Hunkeler, B. Fireman, and T. Hu. 2000. "Alchol Drinking Patterns and Medical Care Use in an HMO Setting." *The Journal of Behavioral Health Services and Research* 27 (1): 3–16.

Rothstein, M. A., and H. L. Harrell. 2009. "Health Risk Reduction Programs in Employer-Sponsored Health Plans: Part I – Efficacy." *Journal of Occupational and Environmental Medicine* 51 (8): 943–50.

Rudowitz, R., S. Artiga, and M. B. Musumeci. 2014. The ACA and Recent Section 1155 Medicaid Demonstration Waivers. Kaiser Family Foundation. February 2014 [accessed on March 17, 2014]. Available at http://kaiserfamilyfoundation.files.wordpress.com/2014/02/8551-the-aca-and-recent-section-1115-medicaid-demonstration-waivers.pdf

Sommers, B. D., and S. Rosenbaum. 2011. "Issues in Health Reform: How Changes in Eligibility May Move Billions Back and Forth between Medicaid and the Insurance Exchanges." *Health Affairs* 30 (2): 228–36.

Stanek, M., and M. Takach. 2010. Evaluating the Patient-Centered Medical Home: Potential and Limitations of Claims-Based Data. National Academy of State Health Policy [accessed on July 30, 2013]. Available at http://nashp.org/sites/default/files/EvalMedHomesFinal-9-10-10.pdf

Strine, T. W., R. Kobau, D. P. Chapman, D. J. Thurman, P. Price, and L. S. Balluz. 2005. "Psychological Distress, Comorbidities, and Health Behaviors among U.S. Adults with Seizures: Results from the 2002 National Health Interview Survey." *Epilepsia* 46 (7): 1133–9.

Sturm, R. 2002. "The Effects of Obesity, Smoking, and Drinking on Medical Problems and Costs." *Health Affairs* 21 (2): 245–53.

Thorpe, K. E., C. S. Florence, D. H. Howard, and P. Joski. 2004. "The Impact of Obesity on Rising Medical Spending." *Health Affairs* 23 (w4): 219–26.

Verdier, J. M., V. Byrd, and C. Stone. 2009. Enhanced Primary Care Case Management Programs in Medicaid: Issues and Options for States. Center for Health Care Strategies, Inc. Resource Paper [accessed on July 30, 2013]. Available at http://www.chcs.org/usr_doc/EPCCM_Full_Report.pdf

Washington State Health Insurance Pool. 2012. "Standard Health Questionnaire for Washington State" [accessed September 18, 2013]. Available at https://www.wship.org/Docs/SHQ%20Revised%2004%203%202012%20Final%20Post.pdf

Weiner, J. P., B. H. Starfield, D. M. Steinwachs, and L. M. Mumford. 1991. "Development and Application of a Population-Oriented Measure of Ambulatory Care Case-Mix." *Medical Care* 29 (5): 452–72.

Weir, S., G. Aweh, and R. E. Clark. 2008. "Case Selection for a Medicaid Chronic Care Management Program." *Health Care Financing Review* 30 (1): 61–74.

Wisconsin Department of Health Services. 2008. Health Insurance for Childless Adults. Waiver Proposal submitted to the Centers for Medicare and Medicaid Services.

Zarkin, G. A., J. W. Bray, T. F. Babor, and J. C. Higgins-Biddle. 2004. "Alcohol Drinking Patterns and Health Care Utilization in a Managed Care Organization." *Health Services Research* 39 (3): 553–69.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Figure A1. Timing of Data Collection.

Figure A2. Inventory of Survey Questions for Top Performing Domains.

Table A1. Details on Construction of Predictor Measures.

Table A2. Discriminative Ability for Insured Sample by Type of Predictor: *c*-Statistic.