Using Electronic Medical Records and Health Claim Data to Develop a Patient Engagement Score for Patients With Multiple Chronic Conditions: An Exploratory Study

Surachat Ngorsuraches, PhD[1] [ID], Semhar Michael, PhD[2],
Nabin Poudel, BPharm, MBA[1], Gemechis Djira, PhD[2],
Emily Griese, PhD[3,4], Arielle Selya, PhD[3,4],
and Patricia Da Rosa, DDS, MSc[5]

Abstract

The study objective was to (1) develop a statistical model that creates a novel patient engagement score (PES) from electronic medical records (EMR) and health claim data, and (2) validate this developed score using health-related outcomes and charges of patients with multiple chronic conditions (MCCs). This study used 2014-16 EMR and health claim data of patients with MCCs from Sanford Health. Patient engagement score was created based on selected patients' engagement behaviors using Gaussian finite mixture model. The PES was validated using multiple logistic and linear regression analyses to examine the associations between the PES and health-related outcomes, and hospital charges, respectively. Patient engagement score was generated from 5095 patient records and included low, medium, and high levels of patient engagement. The PES was a significant predictor for low-density lipoprotein, emergency department visit, hemoglobin A_{1c}, estimated glomerular filtration rate, hospitalization, and hospital charge. The PES derived from patient behaviors recorded in EMR and health claim data can potentially serve as a patient engagement measure. Further study is needed to refine and validate the newly developed score.

Keywords

patient engagement score (PES), patient activation measure (PAM), multiple chronic conditions (MCC), electronic health records (EMR), treatment outcome

Introduction

Controlling health care costs and preventing adverse health outcomes for patients with multiple (2 or more) chronic conditions (MCCs) continue to be a challenge since the management of MCCs are highly influenced by patient behaviors. Recently, patient engagement has become a major focus of health care reform as it has been shown to improve health outcomes and lower health care cost (1–7). Patient activation measure (PAM) is a validated instrument that has been widely used to measure the level of patient engagement in real-world practice. Higher PAM score indicates greater activation and PAM segments patients into 4 activation levels, including (1) the level that patients are passive and lack confidence, (2) the level that patients become aware, but they are still struggling, (3) the level that patients take action,

[1] Department of Health Outcomes Research and Policy, Auburn University, Harrison School of Pharmacy, Auburn, AL, USA
[2] Department of Mathematics and Statistics, Jerome J. Lohr College of Engineering, South Dakota State University, Brookings, SD, USA
[3] Department of Pediatrics, University of South Dakota Sanford School of Medicine, Sanford Research, Sioux Falls, SD, USA
[4] Behavioral Sciences Group, Sanford Research, Sioux Falls, SD, USA
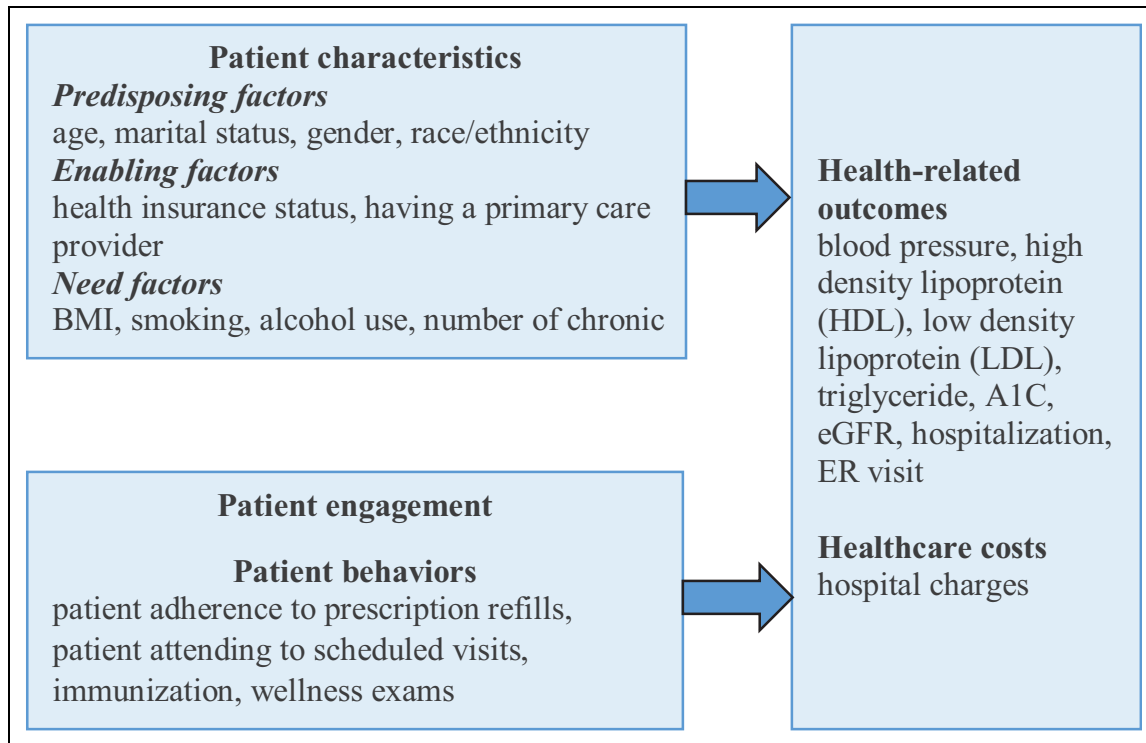[5] Population Health Evaluation Center, Office of Nursing Research, College of Nursing, South Dakota State University, Brookings, SD, USA

Corresponding Author:
Surachat Ngorsuraches, PhD, Department of Health Outcomes Research and Policy, Harrison School of Pharmacy, Auburn University, 4306a Walker Building, Auburn, AL 36849, USA.
Email: surachat@auburn.edu

Figure 1. Conceptual framework adapted from Andersen's model.

and (4) the level that patients can maintain behaviors (8). However, routinely assessing PAM is challenging since it requires patients to spend time and respond to the questions (9). In addition, while the level of patient engagement could change over time, especially for chronic conditions (9).

Electronic medical record (EMR) and health claim data have been widely used for several purposes in the US health care. Some EMR and claim data reflect patient engagement behaviors in various health care activities, for example, immunizations, medication refills, and wellness exams. This study intended to explore the use of these behaviors available in the EMR and health claims from a large regional health care provider to determine the level of patient engagement. This might enable health care providers to routinely capture the level of patient engagement without burden to patients. Thus, the objectives of this study were to develop a statistical model that generated a novel patient engagement score (PES) from EMR and health claims of patients with MCCs, and to validate this developed score using health-related outcomes and charges.

Methods

We conducted a cross-sectional study since only the most recent records of the majority of data elements were available. Figure 1 shows the conceptual framework of this study. We built the framework from Andersen's behavioral model of health service utilization (10–13). The Andersen's model suggests that predisposing, enabling, and need factors could influence health-related outcomes (11). Based on various

studies (1–7,14–16) indicating that patient engagement improved health-related outcomes and lowered health care cost, patient engagement was included in the study framework. Also, previous studies reported that behavioral measures, for example, adherence to medication and self-care behavior, were linked to a high level of patient activation (17–19). This study proposed to construct a patient engagement measure from these behaviors.

Data Collection

This study obtained data through the Sanford Data Collaborative. The data comprised EMR data from Sanford Health, including 44 hospitals, as well as claim data over a 3-year (calendar years 2014-2016) period. All data were de-identified. The data elements included *International Classification of Diseases, Tenth Revision*, age, gender, race, ethnicity, marital status, primary insurance, primary care provider, tobacco use, smokeless cigarette use, alcohol use, body mass index (BMI), PAM score, number of immunizations, number of wellness exams, number of chronic conditions, systolic blood pressure (SBP), diastolic blood pressure (DBP), high-density lipoprotein (HDL), low-density lipoprotein (LDL), hemoglobin A_{1c}, number of emergency department (ED) visits, number of hospitalizations, estimated glomerular filtration rate (eGFR), number of missing scheduled visits, clinic charges, and hospital charges. Only cumulative values over the 3-year period were available for the number of immunizations, wellness exams, ED visits, hospitalizations, and missing scheduled visits. Late medication

refills were determined from 3 variables, including medication name, the number of day supply, and prescription date, using Sanford Health claim data. We assumed that the late medication refill occurred when the number of days supply was shorter than the number of days between 2 consecutive prescription refill dates. Then, the proportion of late refills was calculated for each patient. This study included patients who were aged at least 18 years old and had MCCs with one of the following primary diagnoses: hypertension, hyperlipidemia, ischemic heart disease, heart failure, diabetes, and chronic kidney disease; since these diseases were among highest chronic condition rates with highest rates of MCCs. This resulted in a total of 147 687 individual patient records.

Data Analysis

We conducted descriptive analyses of all data elements for both patients with PAM and PES scores. To explore the potential use of the EMR and health claim data to determine the level of patient engagement, 3 behavioral measures, including the proportion of late medication refills, the number of missing scheduled visits, and the number of immunizations over the 3-year period were purposely selected, since they likely reflected how much patients engaged in their health care. We then developed the PES score by using data from 5095 patients with complete records of these 3 behavioral measures. A finite mixture of Gaussian model was used to group patients based on the behavioral measures into a collection of nonhomogeneous distributions (clusters) (20,21). The model we assumed a G-component Gaussian mixture given by $f(x;\theta) = \sum_{g=1}^{G} \pi_g f_g(x, \mu_g, \Sigma_g)$, where the component density function φ_g was Gaussian distribution with mean vector μ_g and covariance matrix Σ_g. The mixing proportions π_g had restrictions $0 < \pi_g \leq 1$ and $\sum_{g=1}^{G} \pi_g = 1$. Here, θ contained $\{\pi_g, \mu_g, \Sigma_g, g = 1, \ldots, G\}$. Since, it was assumed that the data came from a mixture of multivariate normal distributions; hence, the 3 variables were log-transformed to have an approximately normal mixture distribution. The expectation-maximization algorithm was used to find the parameter estimates of the model at different numbers of groups. The iterative algorithm was stopped after the relative difference in likelihood value was below a given threshold. The optimal number of groups was identified using the Bayesian information criterion (BIC) (21). After obtaining the component parameters for the best model, maximum a-posteriori estimate was used by which an individual was assigned to a component with the highest posterior probability of belonging was used to assign individuals to clusters. To generate the PES score, the clusters were ranked based on the mean vector by which converting clusters labels to ordered categorical variable.

We used multiple logistic regression analyses to examine the relationship between all outcomes (dependent variable)—SBP, DBP, HDL, LDL, A_{1c}, eGFR (normal vs not normal range), and ED visit and hospitalization (yes vs no) and independent variables—age, gender, race, marital status, primary health insurance, primary care provider, BMI, tobacco use, alcohol use, number of chronic conditions, PAM, or PES score. We used multiple linear regression analyses to examine the relationship between the same independent variables and hospital charges. However, since the hospital charge had a right screwed distribution, we log-transformed the charge before we conducted the analyses. A 2-sided P value of <.05 was considered statistically significant.

Results

Table 1 shows the descriptive statistics of all study variables for patients with PAM scores and patients with PES scores. Only 1442 patient records had PAM scores. The average ages of patients with PAM and PES scores were about 67 and 56 years old, respectively. Approximately 64% and 49% of the patients with PAM and PES scores were female, respectively. More than 90% of the patients were white and non-Hispanic or Latino for both groups. The majority of the patients with PAM scores and the patients with PES scores were married (55.3% and 64.1%, respectively), had primary care providers (94.2% and 94.1%, respectively), and were former or nonsmokers (89.7% and 82.0%, respectively). Although 65.9% of the patients with PAM scores had government-sponsored health insurance, 16.1% of the patients with PES scores had this type of insurance. Less than 40% of the patients with PAM scores, but 52.0% of the patients with PES scores, consumed alcoholic beverages. The average BMI (25.3) of the patients with PES scores were lower than those of the patients with PAM scores (32.3). Over 58% of the patients with PAM scores had the scores at either 3 or 4. Both groups had similar health-related outcomes. The average hospital charges for the patients with PAM scores, and the patients with PES scores were $111 268 and $66 986, respectively.

From the Gaussian model, a 3-component mixture was chosen. The best model based on BIC assumes the covariance matrices associated with the groups are symmetric and equal variance for all 3 variables within a group. By inspection of the associated behavioral characteristics, the groups were labeled as low (1), medium (2), and high levels of engagement (22). The mean number of immunizations was 4.632, 4.980, and 5.245, whereas the mean number of missing scheduled visits was 2.627, 2.548, and 1.100 when the patient engagement was at low, medium, and high levels, respectively. Similarly, the mean proportion of late medication refills was 3.568 when the patient engagement was at low level and 0 when the patient engagement was at medium and high levels.

Tables 2 and 3 show the associations between PAM/PES scores and health-related outcomes and hospital charges, respectively. After controlling for all variables, the PAM score was not a significant factor associated with any of the health-related outcomes, except hospital charges. Patients with the lowest level of PAM score had significantly higher

Table 1. Patients' Characteristics, Health Behavior, and Health-Related Outcomes for All Patients, Patients With PAM Scores, and Patients With PES Scores.

Variables	Patients with PAM scores	Patients with PES scores
Average age ± SD years	67.3 ± 12.9	55.8 ± 10.8
	N = 1442	N = 5095
Gender, # (%)		
Female	916 (63.5)	2527 (49.6)
Male	526 (36.5)	2568 (50.4)
Race, # (%)		
Black	19 (1.3)	122 (2.4)
American Indian or Alaskan Native	50 (3.5)	195 (3.8)
White	1348 (93.5)	4658 (91.4)
Asian	7 (0.5)	55 (1.1)
Native Hawaiian or Pacific Islander	2 (0.1)	7 (0.1)
Missing	16 (1.1)	58 (1.1)
Ethnicity, # (%)		
Hispanic or Latino	12 (0.8)	99 (1.9)
Not Hispanic or Latino	1405 (97.4)	4953 (97.2)
Missing	25 (1.7)	43 (0.8)
Marital status, # (%)		
Married	797 (55.3)	3266 (64.1)
Single	217 (15.0)	1168 (22.9)
Others	428 (29.7)	652 (12.8)
Missing	–	9 (0.2)
Primary insurance, # (%)		
Government sponsored	950 (65.9)	821 (16.1)
Nongovernment sponsored	470 (32.6)	3912 (76.8)
Missing	22 (1.5)	362 (7.1)
Primary care provider, # (%)		
Yes	1359 (94.2)	4793 (94.1)
No	83 (5.8)	302 (5.9)
Missing	–	–
Tobacco use, # (%)		
Current every day smoker	149 (10.3)	917 (18.0)
Former smoker	591 (41.0)	1766 (34.7)
Never smoker	702 (48.7)	2409 (47.3)
Missing	–	3 (0.1)
Smokeless cigarette use, # (%)		
Current every day smoker	10 (0.7)	143 (2.8)
Former smoker	60 (4.2)	298 (5.8)
Never smoker	1309 (90.8)	4434 (87.0)
Missing	63 (4.4)	220 (4.3)
Alcohol use, # (%)		
Yes	557 (38.6)	2651 (52.0)
No	835 (57.9)	2317 (45.5)
Missing	50 (3.5)	127 (2.5)
Average BMI ± SD	32.3 ± 6.7	25.3 ± 7.6
	N = 1399	N = 5070
PAM score, # (%)		
Level 1	153 (10.6)	8 (0.2)
Level 2	445 (30.9)	17 (0.3)
Level 3	331 (23.0)	24 (0.5)
Level 4	513 (35.6)	41 (0.8)
Missing	–	5005 (98.2)
Average # of immunizations ± SD	7.1 ± 4.9	6.2 ± 4.9
	N = 1274	N = 5095
Average # of wellness exams	1.3 ± 0.6	1.5 ± 0.8
	N = 3	N = 5095
Average % late medication refills ± SD	55.7 ± 22.4	54.1 ± 23.4

(continued)

Table I. (continued)

Variables	Patients with PAM scores	Patients with PES scores
	N = 125	N = 5095
Average # of chronic conditions	3.1 ± 1.2	2.5 ± 1.0
	N = 1442	N = 5095
Average SBP ± SD	127.0 ± 14.9	126.3 ± 14.6
	N = 1441	N = 5094
Average DBP ± SD	72.2 ± 11.0	75.9 ± 10.2
	N = 1441	N = 5092
Average HDL ± SD	47.0 ± 13.2	46.9 ± 13.3
	N = 1380	N = 4601
Average LDL ± SD	93.7 ± 32.1	99.4 ± 33.8
	N = 1367	N = 4559
Average A1C ± SD	7.0 ± 1.3	6.8 ± 1.3
	N = 996	N = 2879
Average # of ED visits ± SD	3.2 ± 2.5	2.5 ± 2.1
	N = 771	N = 2267
Average # of hospitalizations ± SD	2.2 ± 1.4	1.7 ± 1.1
	N = 547	N = 1662
Average eGFR ± SD	61.6 ± 17.0	69.6 ± 15.8
	N = 1259	N = 3619
Average # of missing scheduled visits ± SD	4.1 ± 3.8	3.3 ± 3.3
	N = 924	N = 5095
Average hospital charge ± SD	$111 268.0 ± $208 464.4	$66 986.5 ± $116 841.9
	N = 1396	N = 4379

Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; ED, emergency department; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; LDL, low-density lipoprotein; PAM, Patient Activation Measure; PES, patient engagement score; SBP, systolic blood pressure; SD, standard deviation.

hospital charges ($\hat{\beta}[SE] = 0.288$ [0.145], $P = .046$), compared to patients with the highest level of PAM score. The patients with medium PES score had significantly higher odds of having abnormal A_{1c} (odds ratio [OR] [95% CI] = 1.340 [1.110-1.617], $P = .002$), at least one ED visit over the 3-year period (OR [95% CI] = 1.264 [1.093-1.462], $P = .002$), at least one hospitalization over the 3-year period (OR [95% CI] = 1.323 [1.134-1.542], $P < .001$), and had higher hospital charges ($\hat{\beta}[SE] = 0.291$ [0.060], $P < .001$) than those patients with the highest PES score. For eGFR, patients with the lowest PES scores had significantly higher odds of having abnormal kidney function (OR [95% CI] = 1.278 [1.064-1.534], $P = .009$), compared to the patients with the highest PES score. For LDL, while the patients with the lowest PES scores had significantly higher odds of having abnormal LDL (OR [95% CI] = 1.182 [1.005-1.390], $P = .043$), the patients with medium PES scores had significantly lower odds (OR [95% CI] = 0.818 [0.702-0.955], $P = .011$), compared to the patients with the highest PES scores.

Discussions

In general, the patient characteristics, health behavior, and health-related outcomes of the patients with PAM scores, and the patients with PES scores were similar. However, it is noteworthy that the patients with PAM scores were likely younger, female, had nongovernment sponsored health insurance, consumed alcoholic beverages, and had slightly lower BMI. PAM score was a significant predictor for only hospital charges, whereas PES score was a significant predictor for several outcomes, for example, LDL, ED visit, A_{1c}, eGFR, hospitalization, and hospital charges.

Interestingly, the PAM score was a significant predictor for only hospital charges in this study. Several previous studies showed that patient activation, as a proxy of patient engagement, was a significant predictor of health-related outcomes (2,23-25). Besides the demographic differences, one of the reasons could be that the PAM score was not fully implemented at Sanford Health during the study period. The records showed that only 1% of all patients with MCCs had PAM scores. These results supported the rationale of this study, which was to develop another patient engagement measurement from existing data in order to reduce any patient burnout or PAM survey administration. In addition, the PAM score was criticized for its inability to fully capture the broad concept of patient engagement (26).

Based on 3 real behavioral measures (average percentage of late medication refills, average number of missing scheduled visits, and average number of immunizations) of patient engagement with their health care, PES was developed in this study. We used a χ^2 test to examine the association between PAM and PES scores from the same individuals and found no significant relationship ($P = .957$). However, the test was based on a small sample size (only 90 individual patients) who had both scores. Despite the lack of statistical

Table 2. Multiple Logistic Regression Results for the Association Between Either PAM or PES, and Health-Related Outcomes.

	Odds ratio (95% CI), P value			
	Systolic blood pressure[a]		Diastolic blood pressure[a]	
Variable	PAM	PES	PAM	PES
Age	1.008 (0.991-1.025), .361	**1.010 (1.001-1.020), .039**	**0.965 (0.942-0.990), .005**	**0.969 (0.958-0.980), <.001**
Gender				
Female	0.966 (0.666-1.401), .855	**1.253 (1.036-1.514), .020**	0.586 (0.326-1.053), .074	**1.642 (1.267-2.129), <.001**
Male[b]				
Race				
White	1.257 (0.609-2.596), .536	**1.438 (1.026), .035**	1.515 (0.569-4.040), .406	1.173 (0.757-1.817), .475
Nonwhite[b]				
Marital status				
Married	1.041 (0.726-1.494), .827	1.332 (1.092-1.623), .005	1.207 (0.659-2.209), .543	1.282 (0.981-1.674), .069
Single & others[b]				
Health insurance				
Government sponsored	0.751 (0.495-1.140), .179	1.191 (0.934-1.517), .158	0.573 (0.298-1.102), .095	1.128 (0.791-1.607), .506
Nongovernment sponsored[b]				
Primary care provider				
No	1.237 (0.591-2.588), .572	1.323 (0.920-1.903), .131	1.601 (0.585-4.383), .359	1.319 (0.818-2.127), .257
Yes[b]				
BMI	1.003 (0.977-1.030), .806	**1.022 (1.008-1.037), .002**	0.974 (0.932-1.019), .252	**1.020 (1.002-1.039), .032**
Tobacco use				
Former smoker & Nonsmoker	0.889 (0.475-1.665), .714	**1.501 (1.195-1.886), <.001**	1.480 (0.685-3.197), .319	**1.920 (1.448-2.545), <.001**
Smoker[b]				
Alcohol				
No	**1.530 (1.053-2.224), .026**	0.992 (0.822-1.196), .931	0.984 (0.545-1.777), .957	0.992 (0.769-1.280), .949
Yes[b]				
Number of chronic conditions	1.047 (0.983-1.114), .153	0.982 (0.928-1.038), .520	1.034 (0.925-1.156), .553	0.919 (0.841-1.005), .064
PAM or PES	1 0.735 (0.390-1.386), .342	1 0.954 (0.756-1.205), .695	1 0.981 (0.348-2.768), .971	1 1.076 (0.792-1.464), .638
	2 0.936 (0.618-1.416), .753	2 0.931 (0.747-1.160), .524	2 1.188 (0.601-2.349), .619	2 1.057 (0.783-1.427), .718
	3 0.846 (0.531-1.347), .480	3[b]	3 0.720 (0.318-1.634), .432	3[b]
	4[b]		4[b]	
N	1316	4465	1316	4465
Nagelkerke R^2	0.018	0.021	0.082	0.021
-2 Log likelihood	938.017	3197.349	407.037	1958.338
Model χ^2 (P value)	11.998 (.528)	48.080 (<.001)	30.916 (.003)	96.444 (<.001)

(continued)

Table 2. (continued)

	Odds ratio (95% CI), P value			
	HDL[a]		LDL[a]	
	PAM	PES	PAM	PES
	0.980 (0.986-0.993), .002	**0.986 (0.979-0.994), <.001**	**0.979 (0.967-0.991), .001**	**0.986 (0.979-0.993), <.001**
	0.246 (0.187-0.324), <.001	**3.687 (3.160-4.301), <.001**	**1.657 (1.270-2.162), <.001**	**0.584 (0.512-0.666), <.001**
	1.102 (0.617-1.968), .744	0.875 (0.647-1.184), .388	0.753 (0.424-1.337), .333	0.919 (0.699-1.208), .543
	1.136 (0.859-1.503), .372	1.055 (0.898-1.240), .514	0.854 (0.660-1.105), .230	0.866 (0.750-1.000), .050
	1.358 (0.982-1.879), .065	1.070 (0.873-1.311), .513	**0.737 (0.548-0.992), .044**	0.865 (0.719-1.041), .126
	1.229 (0.654-2.310), .522	1.347 (0.957-1.894), .087	1.340 (0.742-2.417), .332	**0.700 (0.498-0.982), .039**
	1.039 (1.018-1.060), <.001	**1.052 (1.040-1.064), <.001**	**0.973 (0.954-0.992), .005**	0.994 (0.984-1.004), .238
	1.821 (1.191-2.785), .006	**1.696 (1.403-2.050), <.001**	**0.604 (0.392-0.932), .023**	1.122 (0.940-1.340), .203
	1.293 (0.982-1.703), .067	**1.474 (1.274-1.707), <.001**	0.984 (0.765-1.267), .901	0.911 (0.799-1.039), .164
	1.114 (1.059-1.171), <.001	**1.157 (1.110-1.206), <.0001**	**0.844 (0.795-0.896), <.001**	**0.839 (0.803-0.877), <.001**
1	1.306 (0.833-2.049), .245	1.093 (0.911-1.311), .339	0.994 (0.639-1.544), .977	1.182 (1.005-1.390), .043
2	1.289 (0.933-1.781), .124	1.178 (0.994-1.396), .059	0.791 (0.586-1.069), .127	0.818 (0.702-0.955), .011
3	1.080 (0.759-1.536), .669	3[b]	1.001 (0.727-1.377), .997	3[b]
4[b]			4[b]	
	1263	4060	1253	4024
	0.201	0.176	0.124	0.079
	1380.554	4482.109	1535.516	5306.996
	195.042 (<0.001)	540.820 (<0.001)	119.396 (<0.001)	245.013 (<0.001)

	Odds ratio (95% CI), P value			
	ED visit[a]		AIC[a]	
	PAM	PES	PAM	PES
	0.999 (0.986-1.011), .815	**0.985 (0.978-0.991), <.001**	**0.985 (0.971-0.999), .033**	**0.980 (0.971-0.988), <.001**
	0.989 (0.762-1.284), .935	1.098 (0.969-1.244), .144	0.972 (0.724-1.305), .849	0.986 (0.835-1.165), .872
	1.134 (0.636-2.020), .670	1.238 (0.965-1.589), .093	1.647 (0.863-3.143), .130	**1.369 (1.001-1.872), .049**
	1.631 (1.266-2.101), <.001	**1.192 (1.042-1.363), .011**	1.115 (0.835-1.490), .459	1.007 (0.844-1.202), .934
	1.669 (1.242-2.444), .001	**1.440 (1.212-1.709), <.001**	1.019 (0.721-1.441), .914	0.862 (0.689-1.079), .196
	1.004 (0.575-1.755), .988	0.841 (0.638-1.110), .221	0.974 (0.487-1.948), .941	0.985 (0.6540-1.482), .941
	0.999 (0.980-1.018), .887	**0.988 (0.979-0.997), .012**	**1.037 (1.015-1.060), .001**	**1.016 (1.003-1.029), .017**
	1.559 (1.013-2.398), .044	**1.305 (1.107-1.539), .002**	0.971 (0.601-1.568), .904	1.113 (0.891-1.390), .346
	1.794 (1.401-2.297), <.001	**1.274 (1.125-1.443), <.001**	1.324 (0.992-1.768), .057	1.118 (0.950-1.317), .180
	1.302 (1.219-1.389), <.001	1.209 (1.161-1.260)	**1.095 (1.035-1.159), .002**	**1.177 (1.121-1.235), <.001**
1	1.035 (0.660-1.623), .880	1.071 (0.917-1.250), .386	0.882 (0.550-1.416), .604	1.015 (0.826-1.246), .888
2	0.823 (0.611-1.108), .198	**1.264 (1.093-1.462), .002**	1.292 (0.919-1.817), .141	**1.340 (1.110-1.617), .002**
3	1.049 (0.761-1.446), .772	3[b]	1.122 (0.775-1.623), .542	3[b]
4[b]			4[b]	
	1273	4416	901	2507
	0.185	0.061	0.065	0.055
	1558.865	5862.083	1179.358	3368.688
	188.875 (<0.001)	206.639 (<0.001)	44.267 (<0.001)	105.714 (<0.001)

(continued)

Table 2. (continued)

	Odds ratio (95% CI), P value			
	eGFR[a]		Hospitalization[a]	
	PAM	PES	PAM	PES
	1.072 (1.053-1.090), <.001	**1.062 (1.054-1.070), <.001**	**1.018 (1.005-1.031), .008**	1.001 (0.994-1.008), .737
	3.063 (2.078-4.515), <.001	**0.542 (0.467-0.629), <.001**	0.761 (0.576-1.004), .054	**1.236 (1.080-1.413), .002**
	1.110 (0.534-2.310), .779	**0.725 (0.552-0.953), .021**	1.039 (0.583-1.851), .897	0.942 (0.723-1.228), .657
	0.944 (0.629-1.418), .782	0.979 (0.836-1.147), .796	**1.371 (1.050-1.790), .020**	1.087 (0.941-1.256), .255
	0.896 (0.581-1.382), .619	1.069 (0.862-1.326), .543	**1.570 (1.142-2.158), .005**	**1.647 (1.382-1.963), <.001**
	2.063 (0.793-5.372), .138	**1.497 (1.053-2.128), .025**	**1.852 (1.034-3.314), .038**	**1.386 (1.049-1.832), .022**
	1.019 (0.990-1.048), .209	**0.988 (0.977-0.999), .027**	1.003 (0.983-1.023), .790	**0.989 (0.979-0.999), .035**
	0.702 (0.422-1.169), .174	**0.705 (0.587-0.845), <.001**	**1.695 (1.100-2.611), .017**	1.101 (0.9240-1.311), .282
	0.567 (0.381-0.845), .005	1.017 (0.879-1.177), .820	**1.444 (1.107-1.884), .007**	**1.424 (1.247-1.626), <.001**
	1.129 (1.029-1.239), .011	**1.063 (1.014-1.534), .011**	**1.434 (1.340-1.534), <.001**	**1.270 (1.218-1.325), <.001**
	1 0.570 (0.312-1.041), .067	1 **1.278 (1.064-1.534), .009**	1 1.052 (0.671-1.649), .825	1 0.960 (0.811-1.137), .638
	2 1.105 (0.686-1.780), .681	2 1.027 (0.867-1.217), .754	2 0.896 (0.653-1.228), .493	2 **1.323 (1.134-1.542), <.001**
	3 1.013 (0.613-1.674), .961	3[b]	3 1.059 (0.753-1.489), .742	3[b]
	4[b]		4[b]	
	1305	4298	1257	4434
	0.229	0.144	0.252	0.093
	775.735	4570.752	1435.397	5294.774
	163.454 (<0.001)	449.968 (<0.001)	259.116 (<0.001)	304.429 (<0.001)

Abbreviations: BMI, body mass index; ED, emergency department; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; LDL, low-density lipoprotein; PAM, patient activation measure; PES, patient engagement score.

[a] Coded as binary variable (normal vs not normal)

[b] Reference group.

Table 3. Linear Regression Results for the Association Between PAM/PES and Hospital Charges.

	Ln (hospital charge)[a] Coefficient (SE), *P* value			
Variable	PAM		PES	
Age	0.004 (0.004), .294		**0.012 (0.003), <.001**	
Gender				
Female	−0.063 (0.088), .476		**−0.241 (0.051), <.001**	
Male[b]				
Race				
White	−0.089 (0.189), .638		0.124 (0.104), .230	
Nonwhite[b]				
Marital status				
Married	**−0.210 (0.086), .015**		**−0.144 (0.056), .010**	
Single & others[b]				
Health insurance				
Government sponsored	**−0.686 (0.102), <.001**		**−0.327 (0.068), <.001**	
Nongovernment sponsored[b]				
Primary care provider				
No	−0.141 (0.191), .461		**−0.319 (0.114), .005**	
Yes[b]				
BMI	−0.006 (0.006), .377		−0.005 (0.004), .223	
Tobacco use				
Former smoker & Nonsmoker	−0.210 (0.141), .137		−0.058 (0.068), .399	
Smoker[b]				
Alcohol				
No	**−0.303 (0.085), <.001**		**−0.251 (0.051), <.001**	
Yes[b]				
Number of chronic conditions	**0.237 (0.016), <.001**		**0.211 (0.015), <.001**	
PAM or PES	1[c]	**0.288 (0.145), .046**	1[c]	−0.075 (0.064), .241
	2[c]	0.036 (0.101), .719	2[c]	**0.291 (0.060), <.001**
	3[c]	0.169 (0.109), .123		
N	1281		3857	
R^2	0.251		0.115	

Abbreviations: BMI, body mass index; PAM, patient activation measure; PES, patient engagement score.
[a]Logarithmic transformation.
[b]Reference group.
[c]Dummy code (PES score 3 is reference group).

association, the PES score was a significant predictor for several outcomes, such as LDL, ED visit, A$_{1c}$, eGFR, hospitalization, and hospital charges. These results showed evidence that the PES score captured the engagement level of the patients with MCCs and their health-related outcomes. To our knowledge, this was the first study that used existing EMR and health claim data to define the level of patient engagement. These results provided an example for other health care settings to generate PES scores for their patients with MCCs and use them to make decisions or recommendations without the need to administer any surveys, which likely cause patient burnout. Furthermore, the PES score can be used to motivate patients to improve their engagement (eg, their medication adherence), to improve patient experience of care, to make decisions, and to make a recommendation that creates an environment promoting and encourage engagement. However, some relationships between the PES scores and those outcomes in this study indicated room for improvement in the current PES score. For instance, only the

patients with medium PES scores had significantly higher odds of having abnormal A$_{1c}$, ED visit, hospitalization, and had higher hospital charge, relative to those with the highest PES scores, while those patients with the lowest PES score did not; which was a counterintuitive finding. Similarly, the findings related to LDL were counterintuitive.

This project suffered from various limitations. First, this study used secondary data from EMR and health claims. Therefore, all study variables were limited. Second, the PES score was based on only 3 behavior variables due to data limitation. For instance, data measuring other dimensions of patient engagement were not available. Having access to other potential behaviors, including patient access to EMR, rescheduling a visit, and patient-led communication (eg, emails via EMR, reminders, and phone calls), would improve the performance of the PES score. Third, although all 3 variables were captured across the study period, most of the health-related outcomes in this study were based on cross-sectional data. Fourth, an inconsistency between the

prescribing dates and the number of day supplies was observed. The late medication refill was assumed to occur when the number of day supply was shorter than the number of days between 2 consecutive prescription dates. This assumption could be too restrictive because patients might still possess some medications from their last refills to cover those late few days. In other words, the late medication refills in this study might be overestimating the actual late refills. Fifth, since the PES was based on EMR and health claims, it could not be used to assess patient engagement in real time for new patients. Last, to assess patient engagement without engaging directly with the patient might result in an incomplete picture and a lack of understanding of other factors that could be driving health outcomes.

Conclusions

In our study, the PES score derived from patient behaviors in EMR and health claim data drastically outperformed the PAM score, which was based on patient survey. It was a significant predictor for various health-related outcomes, such as LDL, ED visit, A_{1c}, eGFR, hospitalization, and hospital charge. Further studies are needed to improve the PES score, and to incorporate its use into clinical settings to promote patient engagement and subsequently improve the care management of patients with MCCs.

Authors' Note

S.N. contributed to conceptualization; S.N., S.M., G.D., P.D.R., and E.G. contributed to methodology; S.N. and S.M. contributed to Formal analysis and investigation; N.P. and S.N. contributed to writing and original draft preparation; N.P., S.N., S.M., A.S., G.D., E.G., and P.D.R. contributed to writing and review and editing; S.N., E.G., and A.S. contributed to funding acquisition, resources, supervision. The data that support the findings of this study are available from Sanford Health, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Sanford Health. Ethical approval to report this case was obtained from institutional review boards at South Dakota State University and Sanford Health approved the study protocol before the study started. This article does not contain any studies with human or animal subjects. Informed consent for patient information to be published in this article was not obtained because secondary data was used for the analysis. Arielle Selya is also affiliated with Pinney Associates, Pittsburgh, PA, USA

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Emily Griese and Ariella Selya received funding from the National Institutes of General Medical Sciences (NIGMS) of the NIH, grant number 1P20GM121341.

ORCID iD

Surachat Ngorsuraches, PhD https://orcid.org/0000-0002-2327-224X

References

1. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood). 2015;34:431-7. doi:10.1377/hlthaff.2014.0452

2. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood). 2013;32:207-14. doi:10.1377/hlthaff.2012.1061

3. World Health Organization. Patient Engagement: Technical Series on Safer Primary Care; 2016.

4. Robert Wood Johnson Foundation. Health policy brief. Health Aff. 2013. doi:10.1377/hpb2013.4

5. Krist AH, Tong ST, Aycock RA, Longo DR. Engaging patients in decision-making and behavior change to promote prevention. Stud Health Technol Inform. 2017;240:284-302. doi:10.3233/978-1-61499-790-0-284

6. Sharma AE, Rivadeneira NA, Barr-Walker J, Stern RJ, Johnson AK, Sarkar U. Patient engagement in health care safety: an overview of mixed-quality evidence. Health Aff (Millwood). 2018;37:1813-20. doi:10.1377/hlthaff.2018.0716

7. Cunningham P. Patient engagement during medical visits and smoking cessation counseling. JAMA Intern Med. 2014;174:1291-8. doi:10.1001/jamainternmed.2014.2170

8. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res. 2005;40:1918-30. doi:10.1111/j.1475-6773.2005.00438.x

9. Kinney RL, Lemon SC, Person SD, Pagoto SL, Saczynski JS. The association between patient activation and medication adherence, hospitalization, and emergency room utilization in patients with chronic illnesses: a systematic review. Patient Educ Couns. 2015;98:545-52. doi:10.1016/j.pec.2015.02.005

10. Babitsch B, Gohl D, Lengerke TV. Re-revisiting Andersen's Behavioral Model of health services use: a systematic review of studies from 1998-2011. Psychosoc Med. 2012;9:Doc11.

11. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav. 1995;36:1-10.

12. Andersen RM. National health surveys and the behavioral model of health services use. Med Care 2008;46:647-53. doi:10.1097/MLR.0b013e31817a835d

13. Bradley EH, McGraw SA, Curry L, Buckser A, King KL, Kasl SV, et al. Expanding the Andersen model: the role of psychosocial factors in long-term care use. Health Serv Res. 2002;37:1221-42. doi:10.1111/1475-6773.01053

14. Aung E, Donald M, Coll JR, Williams GM, Doi SA. Association between patient activation and patient-assessed quality of care in type 2 diabetes: results of a longitudinal study. Health Expect. 2016;19:356-66. 2015/03/17. doi:10.1111/hex.12359

15. Remmers C, Hibbar J, Mosen DM, Wagenfield M, Hoye RE, Jones C. Is Patient activation associated with future health outcomes and healthcare utilization among patients with diabetes? J Ambul Care Manage. 2009;32:320-7.

16. Skolasky RL, Mackenzie EJ, Wegener ST, Riley III LH. Patient activation and functional recovery in persons undergoing spine surgery. J Bone Joint Surg Am. 2011;93: 1665-71. doi:10.2106/JBJS.J.00855

17. Woehrle H, Arzt M, Graml A, Fietze I, Young P, Teschler H, et al. Effect of a patient engagement tool on positive airway pressure adherence: analysis of a German healthcare provider database. Sleep Med. 2018;41:20-6. doi:10.1016/j.sleep.2017.07.026

18. Graffigna G, Barello S, Bonanomi A. The role of patient health engagement model (PHE-model) in affecting patient activation and medication adherence: a structural equation model. PLoS One. 2017;12:e0179865. doi:10.1371/journal.pone.0179865

19. Simmon LA, Wolever RQ, Bechard EM, Snyderman R. Patient engagement as a risk factor in personalized health care: a systematic review of the literature on chronic disease. Genome Med. 2014;6:1-13.

20. McLachlan G, Peel D. Finite Mixture Models. John Wiley & Sons, Inc; 2000.

21. Fraley C, Raftery AE, Murphy TB, Scrucca L.Mclust version 4 for R: normal mixture modeling for model-based clustering, classification, and density estimation. Report no. 597. Department of Statistics University of Washington; 2012.

22. Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc. 2002;97: 611-31. doi:10.1198/016214502760047131

23. Hibbard JH, Tusler M. Assessing activation stage and employing a "next steps" approach to supporting patient self-management. J Ambul Care Manage. 2007;30:2-8. doi:10.1097/00004479-200701000-00002

24. Mosen DM, Schmittdiel J, Hibbard J, Sobel D, Remmers C, Bellows J. Is patient activation associated with outcomes of care for adults with chronic conditions? J Ambul Care Manage. 2007;30:21-9. doi:10.1097/00004479-200701000-00005

25. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. J Gen Intern Med. 2012;27:520-6. 2011/12/01. doi:10.1007/s11606-011-1931-2

26. Graffigna G, Barello S, Bonanomi A, Lozza E. Measuring patient engagement: development and psychometric properties of the patient health engagement (PHE) scale. Front Psychol. 2015;6:274. doi:10.3389/fpsyg.2015.00274

Author Biographies

Surachat Ngorsuraches' research interest is pharmaceutical economics and policy. He has also worked on how to engage patients in research and value assessment framework.

Semhar Michael is an assistant professor at the Mathematics and Statistics Department, South Dakota State University. Her research is in the area of computational statistics with a focus on finite mixture modeling and its applications.

Nabin Poudel is a doctoral student at Harrison School of Pharmacy, Auburn University. His research areas include pharmaceutical economics and policy, patient's unmet needs, and quality of life.

Gemechis Djira is an associate professor of biostatistics at South Dakota State University. His research interests are simultaneous inferences and statistical consulting.

Emily Griese is the vice president of Population Health and Clinical Operations at Sanford Health Plan. She holds appointments as an assistant professor at the Sanford School of Medicine, University of South Dakota and as an assistant scientist in the Behavioral Sciences Group at Sanford Research.

Arielle Selya is a behavioral health researcher specializing in substance use and addiction, particularly nicotine use among adolescents. She also is active in health services research, and focuses on big data analysis administrative health records and translating research into improvements in healthcare delivery.

Patricia Da Rosa is a research associate at South Dakota State University. Her main research interest is in understanding and addressing health disparities in public health. She is currently involved in several research projects including cancer, oral health, tobacco, and health services.