CBE ID

2483

Title

Gains in Patient Activation Measure (PAM) Scores at 12 Months

Project

Primary Prevention

Endorsement Status

Endorsed with Conditions

E&M Committee Rationale/Justification

When the measure returns for maintenance, the committee would like to see:

- Progression on electronic health record (EHR) integration
- Evaluation of bias due to changes in the population over time

Is Under Review

No

Next Maintenance Cycle

Spring 2029

Previous Endorsement Cycle

Spring 2024

Steward

Insignia Health

1.0 New or Maintenance

Maintenance

1.3 Electronic Clinical Quality Measure (eCQM)

No

1.6 Measure Description

The measure is the percentage of patients who achieve a 3-point increase in their Patient Activation Measure® (PAM®) survey score within 12 months. The outcome measure demonstrates how a clinician group performed in providing best care to its patients by quantifying the proportion of patients who had at least a 3-point score change.

The PAM surveys the knowledge, skill, and confidence necessary for self-management on a 0-100

point scale that can be broken down into 4 levels from low activation to high activation. The 13 (or 10) item survey has strong measurement properties and is predictive of most health behaviors, many clinical outcomes, and patient experience. PAM® scores are also predictive of health care costs, with lower scores predictive of higher costs.

1.7 Composite Measure

No

1.7 Measure Type

Patient-reported Outcome Performance Measure (PRO-PM)

1.8 Level of Analysis

Clinician: Group/Practice

1.9 Care Setting

Clinician Office/Clinic

1.10 Measure Rationale

The Patient Activation Measure® (PAM®) is a 10- or 13- item questionnaire that assesses an

individual´s knowledge, skills and confidence for managing their health and health care. The measure

assesses individuals on a 0-100 scale that can be converted to one of four levels of activation, from low (1) to high (4). The PAM performance measure (PAM-PM) is the proportion of patients who achieve a 3-point change in the 0-100 scale score on the PAM from baseline to follow-up measurement (CBE #2483). A positive change would mean the patient is gaining in their ability to manage their health. The measure is not disease specific but has been successfully used with a wide variety of chronic conditions, as well as with people with no medical diagnosis.

The PAM is predictive of most health outcomes, including such diverse outcomes as how a patient fares after orthopedic surgery; remission of depression over time; the likelihood of hospital re-admission or

ambulatory care sensitive (ACS) utilization; the trajectory of a chronic disease over time; and even the

likelihood of a new chronic disease diagnosis in the coming year. PAM scores are also predictive of

health care costs, with lower scores predictive of higher costs.

The PAM is in use both in the US and internationally in research (including more than 850 peer-

reviewed

journal articles) as well as clinical settings. It has been translated into more than 30 languages. Because

researchers all over the world use PAM, we have been able to validate the instrument with people of

different racial and ethnic backgrounds, and with people from different socio-economic levels. The

measure has been shown to be valid and reliable in different clinical settings and under different

payment models.

1.13 Data Dictionary

Attached

1.14 Numerator

The numerator includes eligible patients whose PAM score increased by at least 3 points in a 6-12 month period.

1.14a Numerator Details

The numerator includes the count of eligible patients in a clinician group who had an increase of PAM survey score of at least 3 points, based on the difference between the baseline PAM survey and a second score taken between 6-12 months of the baseline.

The below rules should be applied to ensure appropriate patients are selected:

- All patients who have taken at least two valid PAM surveys between 180 and 365 days (6-12 months).
- The first survey should be set as the patient's baseline.
- The last survey administered between 180-365 days after the baseline survey should be selected as the follow up measurement.
- In the case where survey completion dates are not available, and data is organized by year at the patient level, one year may be joined to the following year.
- Patients with PAM level 1-3 in their baseline survey should only be included.

The following rules should be applied to ensure appropriate surveys are selected:

- Each PAM-13 survey has less than 4 missing responses; each PAM-10 survey has less than 3 missing responses.
- Surveys that have PAM scores between, but not including, 0 and 100 are included.
- Only surveys where answers come directly from the patient are included.

1.15 Denominator

The denominator includes eligible patients with two PAM scores no less than 6 months and not more than 12 months apart who were seen for a qualifying visit at least once during the performance period. Clinician groups would need to have two PAM scores on a minimum of 50 patients.

1.15a Denominator Details

Patients aged 14 and older with two PAM scores no less than 6 months and not more than 12 months apart who were seen for at least one qualifying visit at least once during the performance period. Qualifying visits include visits with CPT codes 99201-99205; 99212-99215; 99324-99337; 99341-99350; 99381-99387; 99391-99397; 99490; 99495-99496; 98966-98968, 98969-98972, 99421-99423, 99441-99443, 99444 .

1.15b Denominator Exclusions

Diagnosis of Dementia (ICD-10-CM): F01.5, F02.80, F02, F03.9, F10.27, F10.97, F13.97, F13.27, F18.17,

F18.27, F19.97, F19.17, F19.27, G31.0

OR

Diagnosis of Huntington's disease (ICD-10-CM): G10

OR

Diagnosis of Cognitive Impairment or Alzheimer’s disease (ICD-10-CM): A81.00, A81.09, G20.0, G30.0,

G30.1, G30.9, G31.01, G31.84, G40.909, I67.850, R41.0

1.15c Denominator Exclusions Details

The denominator exclusions are ICD-10 codes and do not require calculations.

1.16 Type of Score

Rate/proportion

1.17 Measure Score Interpretation

Better performance = Higher score

1.18 Calculation of Measure Score

Information for the measure calculation is collected via the PAM survey instrument, which is provided as an appendix to this application. The below steps should be completed to minimize bias and reduce workload burden on programs. Entities using the measure will be responsible for

identifying eligible cases using electronic/automated queries, fielding the survey in the appropriate timeframes, and receiving, cleaning, and summarizing survey data for group-level quality improvement.

- Identify the denominator
    1. Patient aged 14 or older?
        1. Yes – Eligible
        2. No – Not eligible and do not include in the denominator
    2. Patient was seen for at least one qualifying visit during the performance period, defined as a visit with CPT codes 99201-99205; 99212-99215; 99324-99337; 99341-99350; 99381-99387; 99391-99397; 99490; 99495-99496; 98966-98968, 98969-98972, 99421-99423, 99441-99443, 99444?
        1. Yes – Eligible
        2. No – Not eligible and do not include in the denominator
    3. Patient does not have a diagnosis of dementia, Huntington's disease, cognitive impairment or Alzheimer's disease, as defined by ICD-10-COM codes (F01.5, F02.80, F02, F03.9, F10.27, F10.97, F13.97, F13.27, F18.17, F18.27, F19.97, F19.17, F19.27, G31.0, G10, A81.00, A81.09, G20.0, G30.0, G30.1, G30.9, G31.01, G31.84, G40.909, I67.850, R41.0)?
        1. Yes – Eligible
        2. No – Not eligible and remove from denominator
- Identify eligibility for numerator within target respondent population
    1. Patient has taken at least two valid PAM surveys between 180-365 days (6-12 months)?
        1. Yes – Eligible
        2. No – Not eligible
    2. The baseline PAM survey score is between Level 1-3?
        1. Yes – Eligible
        2. No – Not eligible
    3. Each PAM-13 survey has less than 4 missing responses? OR Each PAM-10 survey has less than 3 missing responses?
        1. Yes – Eligible
        2. No – Not eligible
    4. PAM surveys have PAM scores between, but not including, 0 and 100?
        1. Yes – Eligible
        2. No – Not eligible
    5. Did the patient have an increase of PAM survey score of at least 3 points?
        1. Yes – Include in proportion who achieved measure
        2. No – Do not include.
- For those patients who meet the denominator criteria but have not achieved the 3 point increase, have denominator exceptions (Patients who are at PAM® Level 4 at baseline or patients who are flagged with extreme straight line response sets on the PAM®) been documented?
    1. Yes – remove from the denominator
    2. No – include in the denominator and performance is not met.

- Summarize performance on the measure
    1. Did the patient in the clinician group have an increase of PAM survey score of at least 3 points?
        1. Yes – Include in proportion who achieved measure
        2. No –performance is not met.

1.19 Measure Stratification Details

This measure is not intended to be stratified based on risk.

1.20 Types of Data Sources

Administrative Data, Electronic Health Records, Paper Patient Medical Records, Patient-Reported Data and/or Survey Data

1.21b Attach Data Collection Tool(s)

PAM Surveys.zip

1.22 Proxy Responses

No

1.23 Survey Respondent

Patient

1.24 Data Collection and Response Rate

Individual clinicians would need to have two PAM® scores on a minimum of 50 patients with two PAM® scores. We recommend groups attempt to survey each eligible patient to reach this minimum threshold.

The Patient Activation Measure® (PAM®) is a 10- or 13- item questionnaire that assesses an individual´s knowledge, skills and confidence for managing their health and health care. The survey assesses individuals on a 0-100 scale that can be converted to one of four levels of activation, from low (1) to high (4).

The PAM performance measure (PAM-PM) is the change on the 0-100 PAM scale score from baseline to follow-up measurement. To help prevent or mitigate against poor response rates that can be prevalent with PRO-PMs, there is evidence of flexibility in how the PAM is administered both from a methodological and linguistic perspective. Substantial research has shown that the PAM survey can be administered by an interviewer, self-administered on paper, and self-administered digitally with similar results.[1] The PAM is available in English, Spanish, and over 30 other languages.

1. Greene, J., Speizer, H., & Witala, W. (2008). Telephone and Web: The Mixed-Mode

Challenge. *Health Services Research*, 43(1). 230-248

1.25 Data Source Details

The performance measure makes use of PAM survey data to quantify changes in patient activation over time. Administrative data, electronic health records, and/or paper patient medical records may also be used to help identify denominator exceptions and exclusions.

1.26 Minimum Sample Size

Individual clinicians would need to have two PAM® scores on a minimum of 50 patients with two PAM® scores in the measurement period.

The threshold for reliability is 50 cases. Additional parameters, such as a percentage threshold for re-administration may be advisable based on specific program design and intent.

2.1 Attach Logic Model

Section 2.1 Logic Model.pdf

2.2 Evidence of Measure Importance

The PAM measure is cross-cutting and is relevant to any clinical situation where the patient has a role to play. At this point, there are enough studies in the literature that multiple systematic reviews have been published based on existing PAM studies. For example, a recent systematic review examined utilization among patients with chronic illnesses. The authors were able to include ten studies and found that patients with low PAM scores were more likely to seek care in an emergency department.[1] Another systematic review examined which measures of self-management were most useful to clinicians. The authors reviewed ten studies and concluded that patient activation and Health Related Quality of Life (HRQOL) were the most useful to clinicians.[2] Finally, one systematic review examined the effectiveness of interventions to increase patient activation in patients with a chronic condition.[3] That study concluded that most of the examined interventions (including individual coaching, group intervention, telephone-based support, and motivational interviewing) resulted in significant improvements in patient activation and in patient behaviors. The authors further noted that tailoring care to the patient's PAM Level was the most effective approach to increasing activation.

When patients are appropriately supported by their clinical teams, they typically gain in their ability to self-manage. Research shows that when the least activated (measuring at levels 1 or 2) are provided appropriate support, they will typically gain 6-9 points on the 0-100 scale within 6 months. This level of change in PAM scores is significant in terms of changing patients' clinical trajectory, as well as their cost trajectory.[4]

1. 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(5):545-552. doi:10.1016/j.pec.2015.02.005
2. Newland P, Lorenz R, Oliver BJ. Patient activation in adults with chronic conditions: A systematic review. *J Health Psychol*. August 2020:135910532094779. doi:10.1177/1359105320947790
3. Cuevas H, Heitkemper E, Huang Y-C, Jang DE, García AA, Zuñiga JA. A systematic review and meta-analysis of patient activation in people living with chronic conditions. *Patient Educ Couns*. February 2021. doi:10.1016/J.PEC.2021.02.016
4. Lindsay A, Hibbard JH, Boothroyd DB, Glaseroff A, Asch SM. Patient Activation Changes as a Potential Signal for Changes in Health Care Costs: Cohort Study of US High-Cost Patients. *J Gen Intern Med*. 2018;33(12):2106-2112. doi:10.1007/s11606-018-4657-6

The above is a sampling of over 850 peer-reviewed publications that have demonstrated the relationship between PAM scores and aspects of healthcare structure, process, intervention, or service.

2.4 Performance Gap

Table 1 provides the distribution of mean performance (proportion of patients with 3-point change) by practice across 32 practices in Dataset 3. See description of Dataset 3 under Scientific Acceptability.

[Please see Supplemental Attachment for Table 1]

The overall performance for the clinician groups in Dataset 3 is 0.40, or on average, 40% of patients at a clinician group achieve a 3-point change in their PAM score. The range of performance ranges from 0.33 to 0.48 for deciles 1-10, which demonstrates a wide range of performance and an overall opportunity for improvement across clinician groups.

Additional performance gap data can be seen in Tables 6a-6c for datasets used for reliability testing.

[Please see Supplemental Attachment for Tables 6a, 6b, and 6c]

Table 1. Performance Scores by Decile

Performance Gap

	Overall	Minimum	Decile_1	Decile_2	Decile_3	Decile_4	Decile_5	Decile_6	Decile_7	Decile_8	Decile_9	Decile_10	Maximum
Mean Performance Score	0.4	0.3	0.33	0.37	0.38	0.38	0.4	0.42	0.43	0.44	0.45	0.48	0.5
N of Entities	32	1	4	3	3	3	3	3	3	3	3	4	1
N of Persons / Encounters / Episodes	11367	53	915	1180	2136	967	1265	2074	862	643	748	577	1482

2.6 Meaningfulness to Target Population

We are aware of a few studies that provide evidence that the target population values the PAM survey and finds it meaningful.

The first is the Patient Engagement Scale (PES) Consumer Testing Analysis**,** prepared for Pfizer Health Solutions, 2006. For this study, investigators conducted structured interviews with 48 patients to assess their perceptions of patient activation and their experience being administered the PAM. Respondents interviewed for this study included patients with diabetes, asthma, COPD, and cardiovascular disease with varying levels of literacy. **94% of respondents reported a positive perception of the PAM.**

Other studies that address perceived satisfaction and meaningfulness of the PAM include the following. More activated patients or members (as assessed by the PAM survey) are more interested in and able to participate in and experience satisfaction with their shared decision making:

Smith, S, Pandit A. Rush S. et al The role of patient activation in preferences for shared decision making: Results from a national survey of U.S. Adults. *Journal of Health Communication (2016) 21(1) 67-75*

Poon BY, Shortell S, Rodriguez H. Patient Activation as a Pathway to Shared Decision-making for Adults with Diabetes or Cardiovascular Disease. *Journal of General Internal Medicine*. October 2019. https://doi.org/10.1007/s11606-019-05351-6

Kidd, L. Better patient activation is a precursor to engagement in shared decision making. *Evidence-Based Nursing,* 24(2): 2021. https://doi.org/10.1136/ebnurs-2019-103241

3.1 Contributions Towards Closing Care Gaps

PAM does not assess a static trait; in fact, once a patient's care team is aware of their PAM level, teams can intervene to improve patient activation, which in turn helps to reduce health disparities. Through appropriate interventions, like tailored support, anyone (regardless of their starting PAM level) can become more activated and have better health outcomes.

Unlike other predictive tools used in healthcare, PAM is squarely patient-focused and doesn't rely on proxy methods of assessing risk, such as claims data which may overemphasize those patients who access care vs unhealthy participants who are under-utilizers.[1] Predictive tools based on claims can also focus precious care management resources on highly activated patients with high-cost conditions vs low activated patients whose outcomes and costs can be impacted.

We have specifically looked at the available evidence on the impact of racial/ethnic or socioeconomic factors on an individual's activation level. The evidence suggests that factors such as education, income, ethnicity, and gender account for less than 10% of the variance in PAM scores (vs. explaining upwards of 25%+ variance in concepts like health literacy).[2]

In all populations, regardless of demographic characteristics, when interventions are tailored to a specific patient's PAM level, we see that lower activated patients are able to achieve higher score changes. When PAM is used in Medicaid, duals and uninsured populations, we still see a full range of PAM levels in each group. When resources are focused on the low activated patients within the group, we see high score change and meaningful reductions in utilization and costs.

Evidence from the successful use of PAM in underserved communities suggests improving patient activation has the potential to reduce the impact of health disparities, such as racial or ethnic minority status and low income, on health outcomes.

1. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447-453. doi: 10.1126/science.aax2342. PMID: 31649194.
2. Unpublished internal data.

4.1 Feasibility Assessment

Given that the PRO-PM is based on the PAM survey, which asks about a patient's knowledge, skills, and confidence regarding their ability to self-manage their healthcare, we consider the risks

associated with the quality measure to be relatively low.

The PAM survey has been demonstrated to show equivalent results whether administered by an interviewer, self-directed by a patient using a paper form, or through digital means. Adding to that flexibility is the fact that the measure is available in English, Spanish, and over 30 other languages.

We acknowledge that there is the potential for some missing data, given that the measure requires administration of a patient-reported outcome measure that is not generally part of routine care. However, we have seen, for example, the ability to collect high volume, high quality PAM data both on the Phreesia platform and within the context of programs like the CMMI Kidney Care Choices (KCC) model.

Across the various implementations of the PAM-PM since the last CBE review, we have neither observed nor been told of any unexpected findings or unintended impact on patients.

4.3 Feasibility Informed Final Measure

Our experience with the PAM survey and the PAM Performance Measure, along with the signficant published literature on the survey, allowed us to align our final measure specifications with those learnings. This allowed for flexibility in terms of how the survey is administered while also focusing on patients who provided reliable and valid scores, that are likely to improve with intervention.

4.4 Proprietary Information

Proprietary measure or components with fees

4.4a Fees, Licensing, or Other Requirements

Patient Activation Measure® and PAM® are registered trademarks of Insignia Health, LLC, a Phreesia company. Copyright © 2003-2024, University of Oregon. All Rights Reserved.

The Patient Activation Measure® (PAM®) was developed and is owned by the University of Oregon. A license is required for all commercial and non-commercial uses of PAM and related assessments (*e.g.*, Parent PAM, Caregiver PAM, language translations of PAM). This license is executed with Insignia Health, a Phreesia company, which licenses the PAM survey and scoring resources on

behalf of the University of Oregon.

The PAM is a component of the “Gains in Patient Activation Measure (PAM) Scores at 12 Months” performance measure. Currently there are two permitted uses of the PAM for performance measure reporting, for the Centers for Medicare and Medicaid Services (CMS) Merit-Based Incentive Payment System (MIPS) and for select, Centers for Medicare and Medicaid Innovation (CMMI) model participants. For MIPS, we have made an online version of the PAM survey and scoring tool available at https://www.phreesia.com/mips via a no cost licensing agreement, to Medicare providers for the permitted use of submitting the measure to MIPS. The second currently permitted use is within CMMI. In that instance the voluntary participants in select CMMI models have access to the PAM through a CMMI contract covering survey instrument administration, training, analysis and scoring support. More information is publicly available on SAM.gov.

5.1.1 Data Used for Testing

Four datasets were used to empirically test the reliability and validity of the PAM-PM measure score at the clinician-group level. Below are descriptions of the time period of the datasets and how the accountable entity is defined. Testing methods and results will refer to the dataset number indicated below to facilitate review.

1. **Dataset 1** –January 2022 to July 2023 PAM measure score results from the CMS Kidney Care First (KCF) program. The accountable entities in this dataset are physician group practices.
2. **Dataset 2** – January 2022 to May 2023 PAM measure score results from a kidney care management group. The accountable entities in this dataset are physician practices at nephrology practices.
3. **Dataset 3** – November 2022 to December 2023 PAM measure score results from Phreesia. The accountable entities in this dataset are primary care clinician groups that had Phreesia collect PAM data during this time period. Clinician groups were validated from the NPPES NPI Registry[1] and mapped to the CMS provider taxonomy[2]. The dataset is used for empirical testing of reliability and risk adjustment justification.

4. **Dataset 4**- Consists of clinician groups from Dataset 3 where net promoter scores (NPS) on providers were collected during the same time period. This dataset is used for validity testing of the measure.
5. **Dataset 5** – April 2021 to August 2022 PAM measure score results from a yearly health and wellness survey study. This dataset consists of patient level information used for validity testing of the risk adjustment model. Specifically, a patient-level analysis was conducted to empirically test the presence of differences in performance scores across multiple patient-level socio-demographic variables.

1. CMS NPPES NPI Registry Downloadable File. https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplificat...;
2. Find Your Taxonomy Code. https://www.cms.gov/medicare/enrollment-renewal/providers-suppliers/hea...;

5.1.2 Differences in Data

Multiple datasets were used in the empirical testing for reliability and validity of the measure score results. A detailed description of the multiple datasets is included earlier in the Section 4.1.1 (Data Used for Testing).

5.1.3 Characteristics of Measured Entities

Table 2 outlines the number of clinician groups used for scientific acceptability testing from each data source and descriptive statistics on the number of patients per clinician group.

[Please see Supplemental Attachment for Table 2]

5.1.4 Characteristics of Units of the Eligible Population

All accountable entity level analyses were performed using patients included in the measure specification. Patient-level identifiable demographic data was unavailable for Dataset 1 and Dataset 2. Table 1b provides age and gender, and median household income[1] demographics for Dataset 3. It should be noted that income information is not available for 17.2% of patients.

[Please see Supplemental Attachment for Table 3]

Table 4 shows the demographics of the patients used for Dataset 4. It should be noted that income information is not available for 18.6% of patients.

[Please see Supplemental Attachment for Table 4]

1. Median household income is proxied by mapping the patient's location, where available, to the S1901: Income in the Past 12 Months (in 2022 Inflation-Adjusted Dollars) dataset from the United States Census Bureau. (Source: https://data.census.gov/table/ACSST5Y2022.S1901?q=median%20income&g=010...;

5.2.1 Level(s) of Reliability Testing Conducted

Accountable entity level (i.e., measure score) (e.g., signal-to-noise analysis)

5.2.2 Method(s) of Reliability Testing

Reliability was calculated using established reliability testing methods described in *The Reliability of Provider Profiling: A Tutorial (2009).* A beta-binomial approach is used for estimating the reliability of a simple pass/fail rate measure and calculates the ability for the measure to distinguish performance between clinician groups. [1]

The beta-binomial is computed on all clinician groups, and mean reliability of clinician groups was calculated for the clinician groups for the measure specification range.

1. Adams JL. *The Reliability of Provider Profiling: A Tutorial*. RAND Corporation; 2009. Accessed May 10, 2023. https://www.rand.org/pubs/technical_reports/TR653.html

5.2.3 Reliability Testing Results

Table 5 displays the mean reliability of the beta-binomial at the clinician group level across the three data sources using the minimum sample size of 50 patients, as outlined in the measure specification.

[Please see Supplemental Attachment for Table 5]

Tables 6a-6c show reliability results for each dataset.

[Please see Supplemental Attachment for Tables 6a, 6b, and 6c]

5.2.4 Interpretation of Reliability Results

Across all three datasets the mean reliability of the beta-binomial exceeds 0.8. This value demonstrates high reliability of the measure score, and this is further supported by the consistent highly reliable results across multiple datasets.

5.3.1 Level(s) of Validity Testing Conducted

Accountable entity level (i.e., measure score) (e.g., criterion validity)

5.3.3 Method(s) of Validity Testing

Empirical validity testing was conducted to test the pathway defined in the logic model that increased PAM scores lead to the improved health behaviors, navigation, and communication, which in turn leads to improved patient satisfaction with care. The hypothesis tested was the following: as average measure performance improves at the accountable entity, average ratings of patient satisfaction with care at the accountable entity also improves.

Patient satisfaction can be operationalized in healthcare settings using a Net Promoter Score. [1]

Specifically, the calculation of Net Promoter Score is as follows:

- A patient is asked on a 1-10 scale of their likelihood to recommend a practice.
- Answers 9-10 are considered promotors, 7-8 are considered neutral, and 6 or less are considered detractors:
- NPS = (count of promotors - count of detractors) / total count of respondents.
- The average net promotor score was calculated for each practice for all patients who responded to the likelihood to recommend question during the study period of Dataset 4.

Testing was conducted by performing an Ordinary Least Squares (OLS) regression analysis using Dataset 4. The regression coefficients that return from the analysis, in addition to calculating Pearson's correlation coefficient, give indication to the direction and magnitude of the association between the average measure performance and NPS rating.

1. Healthcare Net Promoter Score: Definition, Formula, & Benchmarks. https://www.driveresearch.com/market-research-company-blog/healthcare-n...;

5.3.4 Validity Testing Results

OLS regression results show a statistically significant positive association between the

performance measure and patient satisfaction with coefficient effect size 1.13 (0.1->2.16) and p-value 0.03. The Pearson's correlation coefficient showed a similar statistically significant positive association between the PAM-PM measure and patient satisfaction of 0.43 with p-value 0.03.

The PAM-PM was not tested with denominator exclusions; however, we note that the self-report of individuals with significant cognitive concerns may be of questionable reliability and validity and for that clinical reason, the measure includes these denominator exclusions.

5.3.5 Interpretation of Validity Results

These results demonstrate a statistically significant correlation between PAM-PM and patient satisfaction in the hypothesized direction. Clinician groups that achieve a greater proportion of patients who improve their PAM scores by at least 3 points also have a greater proportion of patients who indicate they are satisfied with care from their providers.

A graphical representation of the regression results can be found in Figure 2. The plot shows on the x-axis the PAM-PM performance measure scores, and the y-axis shows patient satisfaction operationalized as the average net promoter score for each clinician group. Each point on the plot represents a clinician group in the validity analysis sample. The regression trendline trending upward indicates that a better performance on the PAM-PM is associated with better patient satisfaction.

[Please see Supplemental Attachment for Figure 2]

5.3.2 Type of Accountable Entity Level Validity Testing Conducted (derived)

Empirical validity testing at the accountable entity-level (e.g., criterion validity, construct validity, known groups analysis)

5.4.1 Methods Used to Address Risk Factors

Other, No risk adjustment or stratification

5.4.1a Other Methods Used

We used several methods to assess potential risk factors, including a random effects logistic regression model (Dataset 3) to predict the probability of meeting measure performance. We also looked at potential patient-level differences looking at fixed effects in the regression model (Dataset 3), with additional review of chi-square results from Dataset 5. More details on the testing and results are available in section 4.4.1b.

5.4.1b Rationale For No Adjustment or Stratification

Empirical testing to determine whether to consider risk adjustment or stratification for the measure was performed using Dataset 3 and Dataset 5. One consideration was to assess whether measure performance is affected by the case-mix of patients across available socioeconomic (SES) factors, while controlling for any accountable entity level effects. It is important to note that no accountable entity level characteristics and/or patient level clinical features were available for consideration in our analyses.

To do this, a random effects logistic regression model was built using Dataset 3 to predict the probability of meeting measure performance.[1] The limited SES factors available for analysis (Age Group, Gender, and Median Household Income) were considered as predictors in the model as fixed effects, plus a random intercept to account for the effects of the accountable entities. A variance component analysis from the model concluded that there is limited contribution from the accountable entities to the total variability in the performance outcome. Given lack of significant between-group variance, decision would be to not control for this factor in a potential risk adjustment model.

Given this, the focus shifted to assessing patient-level differences. No statistically significant differences were concluded from the Gender and Median Household Income fixed effects in the regression model. However, the p-values did indicate small differences in measure performance amongst older age groups, specifically the 65+ age group. Additional patient level testing performed using Dataset 5, where similar SES factors were available (Age Group, Gender, Income Range, as well as Education Level) however came away with contradictory results. Empirical testing of the measure score in Dataset 5 using Chi-square tests showed no statistically significant differences in measure scores across all SES factors available. We present these results in Table 7.

[Please see Supplemental Attachment for Table 7]

Interpretation of results:

The results from Dataset 3 and Dataset 5 are mixed in terms of the conclusions that can be drawn from the effects of patient level differences. or specific SES factors in the form of a risk adjustment model, the decision remained to not risk adjust this measure. Regardless of their sociodemographic characteristics, the conceptual model rationale indicates that patients are able to report improved activation scores. This finding is consistent with a body of literature on the relationship of socio-contextual factors and the PAM survey itself. As the measure is expanded in its use, we will continue to monitor the impact of socio-contextual factors in assessments of accountable entity performance.

1. Bouwmeester, W., Twisk, J.W., Kappen, T.H. *et al.* Prediction models for clustered data: comparison of a random intercept and standard regression model. *BMC Med Res Methodol* 13, 19 (2013). https://doi.org/10.1186/1471-2288-13-19

5.4.2 Conceptual Model Rationale

We have attached a conceptual model (Figure 3) that illustrates our rationale for not risk-adjusting the PAM-PM, that builds on our overall measure logic model.

Based on prior research, we know that certain clinical and/or socioeconomic status categories may impact baseline levels of patient activation; however, activation is not a static trait; in fact, once a patient's care team is aware of their PAM level, teams can intervene to improve patient activation. Through appropriate interventions, like tailored support, anyone (regardless of their starting PAM level) can become more activated and have better health outcomes. We anticipate variability in accountable entities' ability and skill in delivering those activation interventions, related to resources, intervention competence, potential biases/discrimination, and other organizational factors.

In all populations, regardless of demographic characteristics, when interventions are tailored to a specific patient's PAM level, we see that lower activated patients are able to achieve higher score changes. When PAM is used in Medicaid, duals and uninsured populations, we still see a full range of PAM levels in each group. When resources are focused on the low activated patients within the group, we see increase in scores and meaningful reductions in utilization and costs. Said another way, while we see a full range of PAM scores across patient groups, and while there may be differences in baseline PAM scores, what we find is that the ability to improve in PAM-assessed activation is not bound by demographic and health factors. Patients, when appropriately supported, can improve their PAM scores. The resulting changes in activation have been shown to lead to improved clinical outcomes, decreased healthcare utilization, decreased healthcare costs, and improved patient satisfaction with care through improved health behaviors, navigation, and communication.

5.4.2a Attach Conceptual Model

Section 4.4.2a Conceptual Model - Figure 3 and Narrative.pdf

5.4.3 Variable Distribution Across Measured Entities

Table 8 provides descriptive statistics from Dataset 3 of how risk variables considered in the

analyses are distributed across the measured entities.

[Please see Supplemental Attachment for Table 8]

5.4.4 Risk/Case-Mix Adjustment Modeling and/or Stratification Results

N/A

5.4.6 Interpretation of Risk/Case-mix Factor Findings

N/A

5.4.7 Final Approach to Address Risk Factors

No risk adjustment or stratification

6.1.1 Current Status

In use

6.1.3 Current Use(s)

Payment Program, Quality Improvement (Internal to the specific organization)

6.1.3 Program Details

Name of the program and sponsor
Merit-based Incentive Payment System (MIPS)
URL of the program
https://qpp.cms.gov/
Purpose of the program
MIPS encourages improvement in clinical practice and supports advances in technology that allow for easy exchange of information.
Geographic area and percentage of accountable entities and patients included
MIPS eligible providers may earn performance-based payment adjustments for the services provided to Medicare patients in the USA.
Applicable level of analysis and care setting

Clinician; office/ambulatory care

6.2.1 Actions of Measured Entities to Improve Performance

Providers are using the PAM survey score to help them understand how to be most effective in supporting a patient. If the patient scores low, the clinical team recognizes that the patient is likely overwhelmed with the task of managing his or her health, has low confidence, and/or is likely have had several disappointments when trying to manage his or her health and consequently feels discouraged. The provider then understands that it is important to avoid

further overwhelming the patient with too much information and too many suggested changes. At the same time, it is important to give the patient an opportunity to experience a success such as suggesting a small step that the patient might take toward improving his or her health. The step may be quite small and thus not in itself, clinically meaningful, but if the patient is successful, it can feel like a win that, in turn, increases a patient's motivation to take further steps.

In a recently published meta-analysis[1] summarizing the evidence about what works best for supporting gains in patient activation, the approach of tailoring goals to the patient's activation level was identified as an effective strategy. This is different than usual care management, as the focus is on meeting the patient where they are, rather than a one-size-fits-all approach.

There are specific steps the clinical team can take to support gains in activation among their patients. Those steps involve tailoring suggested action steps to the patient's level of activation (described above), problem solving with the patient to overcome barriers, giving encouragement, and showing caring and concern. The work of supporting gains in activation can be carried out mostly by a trained medical assistant. It does not have to be a clinical person doing this work. The clinician only needs to understand in a general way what the approach is, and support it in their interactions with the patient.

The PAM can also help reduce the burden on providers in another way; by helping them be more targeted and efficient in how they deploy their resources (e.g. care managers, social workers, nurses) in managing their patient population. For example, because more activated patients are more ready to use self-management resources, the provider can just push out digital tools to that segment of their patient population. They can then save their limited people resource to support less activated patients. Instead of using a one size fits all approach to counseling and educating patients, this more targeted approach will yield better results with the expenditure of less time resources.

1. Cuevas H, Heitkemper E, Huang Y-C, Jang DE, García AA, Zuñiga JA. A systematic review and meta-analysis of patient activation in people living with chronic conditions. *Patient Educ Couns*. February 2021. doi:10.1016/J.PEC.2021.02.016

6.2.2 Feedback on Measure Performance

The PAM-PM has primarily been tested in large ambulatory primary care practices and subsets of health plan populations. In our review of the available research in which two PAMs have been administered, the attrition rate is relatively low. Some studies show re-assessment rates in the order of 55%, but many studies show re-assessment rates in the 75-85% range. We recognize that different use cases or programs may mandate or expect different levels of data capture and our own analyses suggest a minimum of 50 cases is enough to establish a reliable and valid measure.

We also know that there do not appear to be any mode effects when administering the PAM, and so there is flexibility to administer the survey on paper, digitally, or via an interviewer.

In the CMMI KCC Model, we have implemented a monthly opportunity for groups to provide us feedback on the PAM and its implementation, which has provided us with an opportunity to provide continuous education on the PAM while also hearing from end users on implementation strategies.

6.2.3 Consideration of Measure Feedback

When we have implemented PAM-PM, at scale, we have sought out and been open to feedback from all stakeholders involved in the project. For example, in the KCC program, we have solicited input on the measure at regular meetings with stakeholders. Every piece of information that we have received has been carefully considered. When possible, we have conducted additional testing to assess the impact of potential changes to the measure specification on reliability and validity. We also anticipate receiving feedback on the measure from our inclusion in the MIPS program, which we will review carefully to the extent it suggests any modifications may be necessary or appropriate.

6.2.4 Progress on Improvement

Dataset 1 was used to study progress on improvement of measure performance across measured entities. Patients were evenly split into two groups based on the date of their baseline PAM survey, with Group 1 representing patients who took their baseline PAM survey earlier in time. The mean PAM-PM score across the entities was taken for both groups. We expected patients in the second group to show improvements in the measure, compared to those in the first group, due to measured entities having more experience activating patients in the cohort.

Table 9 provides the summary of the results across the two patient groups. The results show an increase in the mean score across the two groups, which we interpret to be a positive trend in performance of the measure.

[Please see Supplemental Attachment for Table 9]

6.2.5 Unexpected Findings

Given that the PRO-PM is based on the PAM survey, which asks about a patient's knowledge, skills, and confidence regarding their ability to self-manage their healthcare, we consider the risks associated with the quality measure to be relatively low. In fact, across the various implementations of the PAM-PM since the last CBE review, we have neither observed nor been

told of any unexpected findings or unintended impact on patients.

7.1 Supplemental Attachment

SUPPLEMENTAL ATTACHMENT - CBE #2483 - Gains in Patient Activation Meaure (PAM) Scores at 12 Months (REVISED 05-08-2024).pdf

Developer POC email

measuresteward@phreesia.com

Measure Developer POC

Hilary Hatch
Insignia, LLC, a wholly owned subsidiary of Phreesia
1521 Concord Pike, Suite 301 PMB 221
Wilmington, DE 19803
United States

The measure developer is different from the measure steward

No

Steward Address

United States

Steward Organization

Insignia Health