

The responses below provide Tables and Figures (along with associated narrative response) that could not be uploaded directly into the PQM submission portal. This PDF is bookmarked for ease of review.

Section 2. Importance

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.

Table 1 - Performance Gap for Dataset 3

	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Mean 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
Entities	32	1	4	3	3	3	3	3	3	3	3	4	1
Total Persons	11367	53	915	1180	2136	967	1265	2074	862	643	748	577	1482

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

Section 4. Scientific Acceptability

4.1.3 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.

Table 2 - Description of clinician groups in the datasets used for reliability and validity testing

*	Total Clinician groups	Total Patients	Patients per clinician group - Mean	Patients per clinician group - Std Dev	Patients per clinician group - Min	Patients per clinician group - 25%	Patients per clinician group - 50%	Patients per clinician group - 75%	Patients per clinician group - Max
Dataset 1	13	2259	174	114	78	89	151	205	411
Dataset 2	45	6145	137	114	50	71	101	136	520
Dataset 3	32	11367	355	357	53	113	251	439	1482

*	Total Clinician groups	Total Patients	Patients per clinician group - Mean	Patients per clinician group – Std Dev	Patients per clinician group - Min	Patients per clinician group - 25%	Patients per clinician group - 50%	Patients per clinician group - 75%	Patients per clinician group - Max
Dataset 4	25	10022	400	385	53	123	298	440	1482

*Cells left intentionally blank

4.1.4 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 3 provides age and gender, and median household income¹ demographics for Dataset 3. It should be noted that income information is not available for 17.2% of patients.

Table 3 - Description of patients included in Dataset 3

Demographic	Group	Patient Count (% of Total Patients)	Proportion of Patients with 3 Point Increase (95% CI)
Gender	Male	3694 (32.5%)	0.39 (0.37-0.4)
	Female	7673 (67.5%)	0.41 (0.4-0.42)
Age	18to25	908 (8.0%)	0.42 (0.39-0.45)
	25to40	2602 (22.9%)	0.42 (0.4-0.44)
	40to50	2086 (18.4%)	0.41 (0.39-0.43)
	50to65	3566 (31.4%)	0.40 (0.38-0.41)
	65+	2205 (19.4%)	0.37 (0.35-0.39)
Median Household Income	\$15,000-\$24,999	7 (0.1%)	0.57 (0.18 - 0.9)
	\$25,000-\$34,999	134 (1.4%)	0.4 (0.31 - 0.48)
	\$35,000-\$49,999	1194 (12.7%)	0.4 (0.37 - 0.43)
	\$50,000-\$74,999	4857 (51.6%)	0.4 (0.39 - 0.41)
	\$75,000-\$99,999	1767 (18.8%)	0.42 (0.39 - 0.44)
	\$100,000-\$149,999	1357 (14.4%)	0.4 (0.38 - 0.43)
	\$150,000-\$199,999	97 (1.0%)	0.31 (0.22 - 0.41)
	Total	Patients	11367
	Clinician Groups	32	0.4 (0.39-0.41)

*Cell left intentionally blank

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.

Table 4 - Description of patients in validity analysis cohort – Dataset 4

Demographic	Group	Patient Count (% of Total Patients)	Proportion of Patients with 3 Point Increase (95% CI)	
Gender	Male	3310 (33.0%)	0.39 (0.37-0.4)	
	Female	6712 (67.0%)	0.41 (0.4-0.42)	
Age	18to25	791 (7.9%)	0.42 (0.39-0.46)	
	25to40	2237 (22.3%)	0.42 (0.4-0.44)	
	40to50	1854 (18.5%)	0.42 (0.39-0.44)	
	50to65	3166 (31.6%)	0.40 (0.38-0.41)	
	65+	1974 (19.7%)	0.37 (0.35-0.39)	
Median Household Income	\$15,000-\$24,999	6 (0.1%)	0.5 (0.12 - 0.88)	
	\$25,000-\$34,999	134 (1.6%)	0.4 (0.31 - 0.48)	
	\$35,000-\$49,999	1147 (14.1%)	0.4 (0.37 - 0.43)	
	\$50,000-\$74,999	4084 (50.1%)	0.4 (0.38 - 0.41)	
	\$75,000-\$99,999	1544 (18.9%)	0.42 (0.39 - 0.44)	
	\$100,000-\$149,999	1142 (14.0%)	0.41 (0.38 - 0.44)	
	\$150,000-\$199,999	96 (1.2%)	0.31 (0.22 - 0.42)	
	Total	Patients	10022	0.4 (0.39-0.41)
		Clinician Groups	25	0.4 (0.39-0.41)

*Cells left intentionally blank

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=010XX00US\\$860000](https://data.census.gov/table/ACSST5Y2022.S1901?q=median%20income&g=010XX00US$860000)).

4.2 Reliability

4.2.3 [If reliability testing was conducted] 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.

Table 5 - Description of reliability analysis using beta binomial

Metric	Dataset 1	Dataset 2	Dataset 3
Patient count	2259	6145	11367
Clinician group count	13	45	32
Mean reliability of the beta-binomial	0.96	0.89	0.81

Table 2 [If accountable entity-level testing was conducted, i.e., if 4.2.1 includes “Accountable Entity-Level”] Accountable Entity-Level Reliability Testing Results

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

Table 6a - Accountable Entity-Level Reliability Testing Results – Dataset 1

*	Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max
Reliability	0.96	0.92	0.93	0.93	0.94	0.95	0.96	0.96	0.96	0.97	0.98	0.98	0.985
Mean Performance Score	0.54	0.32	0.49	0.46	0.54	0.32	0.70	0.62	0.59	0.41	0.57	0.52	0.85
N of Entities	13	1	2	1	1	1	2	1	1	1	1	2	1
N of Persons / Encounters / Episodes	2259	78	162	89	95	108	229	154	160	205	251	806	411

* Cell left intentionally blank

Table 6b - Accountable Entity-Level Reliability Testing Results – Dataset 2

* Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max	
Reliability	0.89	0.79	0.8	0.83	0.85	0.87	0.89	0.9	0.91	0.92	0.94	0.97	0.97
Mean Performance Score	0.55	0.35	0.56	0.53	0.48	0.58	0.58	0.49	0.47	0.65	0.57	0.55	0.91
N of Entities	45	1	5	4	5	4	5	4	4	5	4	5	1
N of Persons / Encounters / Episodes	6145	50	263	262	386	335	490	467	522	559	733	2128	520

* Cell left intentionally blank

Table 6c - Accountable Entity-Level Reliability Testing Results – Dataset 3

* Overall	Min	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Max	
Reliability	0.81	0.55	0.56	0.66	0.71	0.77	0.83	0.86	0.89	0.91	0.92	0.96	0.97
Mean Performance Score	0.4	0.3	0.41	0.39	0.48	0.4	0.4	0.42	0.38	0.41	0.38	0.4	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	229	257	342	451	671	825	1078	1316	1571	4627	1482

* Cell left intentionally blank

4.3 Validity

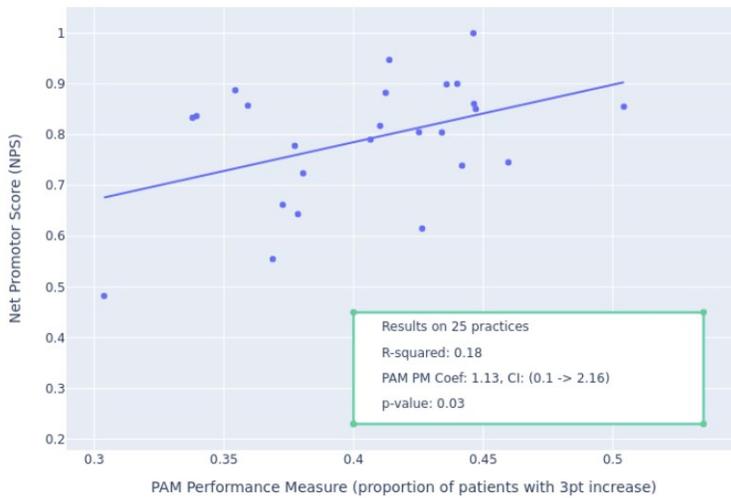
4.3.5 [If validity testing was conducted] Interpretation of Validity Results *

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.

Figure 2 - Clinician Group PAM Performance Measures Association with Net Promotor Score - Dataset 4

Clinician Group PAM Performance Measures Association with Net Promotor Score for Practices with at Least 50 Patients



4.4 Risk Adjustment

4.4.1 Methods Used to Address Risk Factors *

- Statistical risk adjustment model with risk factors
- Stratification by risk factor category
- Other

4.4.1a Describe other method(s) used

- No risk adjustment or stratification.

4.4.1b [If Measure Type is outcome or cost/resource]

Provide a rationale for why there is no need to address differences in patient characteristics (i.e., case mix) to achieve fair comparisons across measured entities for your outcome or resource measure.

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 performance¹ 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.

Table 7 - Results of validity testing of the risk adjustment model - Dataset 5

Characteristic	Category	Patient Count	Proportion of Patients with 3 Point Increase with CI	Chi-square value	Chi-square p value	Interpretation of Results
Age Group	18to25	25 (0.91%)	0.36 (0.18 - 0.575)	3.51	>0.05	No statistically significant differences
	25to40	224 (8.12%)	0.39 (0.328 - 0.46)	*	*	*
	40to50	282 (10.22%)	0.36 (0.302 - 0.417)	*	*	*
	50to65	1006 (36.46%)	0.33 (0.305 - 0.364)	*	*	*
	65+	1222 (44.29%)	0.34 (0.308 - 0.362)	*	*	*
Gender	Female	1535 (55.64%)	0.34 (0.318 - 0.366)	0.231	>0.05	No statistically significant differences
	Male	1224 (44.36%)	0.34 (0.315 - 0.369)	*	*	*
Income Range	Less than \$15,000	99 (3.59%)	0.34 (0.251 - 0.446)	5.87	>0.05	No statistically significant differences

Characteristic	Category	Patient Count	Proportion of Patients with 3 Point Increase with CI	Chi-square value	Chi-square p value	Interpretation of Results
*	15,000to24,999	183 (6.63%)	0.31 (0.245 - 0.384)	*	*	*
*	25,000to34,999	214 (7.76%)	0.30 (0.239 - 0.365)	*	*	*
*	35,000to49,999	319 (11.56%)	0.35 (0.296 - 0.403)	*	*	*
*	50,000to74,999	545 (19.75%)	0.34 (0.303 - 0.385)	*	*	*
*	75,000to99,999	451 (16.35%)	0.34 (0.293 - 0.383)	*	*	*
*	100,000to124,999	318 (11.53%)	0.37 (0.321 - 0.43)	*	*	*
*	125,000to149,999	203 (7.36%)	0.38 (0.312 - 0.45)	*	*	*
*	150,000to199,999	194 (7.03%)	0.35 (0.279 - 0.417)	*	*	*
*	\$200,000 or more	134 (4.86%)	0.31 (0.236 - 0.399)	*	*	*
*	Decline to answer	99 (3.59%)	0.33 (0.242 - 0.435)	*	*	*
Education Level	Less than high school	2 (0.07%)	1.0 (0.158 - 1.0)	6.7	>0.05	No significant differences
*	Completed some high school	34 (1.23%)	0.38 (0.222 - 0.564)	*	*	*
*	High school graduate or equivalent (e.g., GED)	396 (14.35%)	0.33 (0.282 - 0.377)	*	*	*
*	Completed some college, but no degree	523 (18.96%)	0.33 (0.289 - 0.371)	*	*	*
*	College graduate (e.g., B.A., A.B., B.S.)	737 (26.71%)	0.35 (0.313 - 0.383)	*	*	*
*	Completed some graduate school, but no degree	168 (6.09%)	0.37 (0.296 - 0.447)	*	*	*
*	Associate's degree	329 (11.92%)	0.33 (0.284 - 0.388)	*	*	*
*	Completed graduate school (e.g., M.S., M.D., Ph.D.)	568 (20.59%)	0.35 (0.309 - 0.389)	*	*	*

Demographic	Group	Number of entities	Mean	STD	Min	25%	50%	75%	Max
*	25000-34999	32	0.01	0.03	0.00	0.00	0.00	0.00	0.12
*	35000-49999	32	0.07	0.12	0.00	0.00	0.02	0.08	0.53
*	50000-74999	32	0.40	0.31	0.00	0.08	0.39	0.67	0.92
*	75000-99999	32	0.20	0.18	0.00	0.09	0.14	0.26	0.72
*	100000-149999	32	0.19	0.26	0.00	0.00	0.04	0.24	0.80
*	150000-199999	32	0.01	0.03	0.00	0.00	0.00	0.01	0.11

Section 6. Use & Usability

6.2 Usability

6.2.4 [If maintenance review OR Current Status = Yes (6.1.1)] 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.

Table 9 - Performance Trend of PAM-PM – Dataset 1

Group	Minimum Baseline PAM	Number of Patients	Number of Entities	Mean Score	STD Score
1	January 2022	1130	13	0.52	0.16
2	May 2022	1129	11	0.56	0.12

We will continue to evaluate progress on improvement as we collect more data using the measure.