Quality Payment
PROGRAM

Heart Failure

Measure Testing Form

January 2022

Contents

1.0 Introduction

This Measure Testing Form (MTF) provides a brief summary of the preliminary measure testing results as part of field testing five episode-based cost measures. Stakeholders may review these results, alongside other documentation, to provide feedback on the draft measure using the field testing survey. The testing results reflect the performance of the measure as specified at the time of field testing, which is part of the measure development process. Please see the Draft Cost Measure Methodology for a description of the measure specifications and the Draft Measure Codes List for the list of codes used to specify the measure.[1]

1.1 Project Title and Overview

The Centers for Medicare & Medicaid Services (CMS) has contracted with Acumen, LLC to develop care episode and patient condition groups for use in cost measures to meet the requirements of the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). The contract name is "Physician Cost Measures and Patient Relationship Codes (PCMP)." The contract number is 75FCMC18D0015, Task Order 75FCMC19F0004.

1.2 Measure Name

Heart Failure

1.3 Type of Measure

Cost/Resource Use

1.4 Data

The study period is January 1, 2019 through December 31, 2019. All episodes ending during the study period that meet inclusion and exclusion criteria are included in testing. The measure is calculated with Medicare Parts A, B, and D administrative claims data, the Long-Term Minimum Data Set, and the Medicare Enrollment Database. For testing purposes, other data sources are used, including the American Community Survey, Common Medicare Environment, and Uniform Data System.

Testing results are presented at a testing volume threshold of 20 episodes for clinician groups and individual practitioners. Clinician groups are identified by a Tax Identification Number (TIN). Individual clinicians are identified using a combination of a Tax Identification Number and National Provider Identifier (TIN-NPI).

[1] These documents will be available on the MACRA Feedback Page once field testing begins. https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback

2.0 Preliminary Testing Results

This section presents preliminary testing results based on the measure as specified for field testing. Section 2.1 provides an overview of the measure's coverage of beneficiaries and cost. Section 2.2 lists the most frequently attributed specialties. Sections 2.3 through 2.5 provide evidence of scientific acceptability of the measure. Section 2.6 presents empirical results of the risk adjustment and stratification methods used by this measure. Section 2.7 examines the impact of adding social risk factors to the measure's risk adjustment model. Lastly, Section 2.8 examines the impact of exclusion criteria used by the measure through their frequency and resource use patterns.

2.1 Measure Coverage

Table 1 shows the number of patients and the amount of Medicare Parts A and B cost covered by this measure. This measure has the potential to have high impact as heart failure is a common and costly condition for the Medicare population.

Table 1. Beneficiary and Cost Coverage

Coverage Metrics	Coverage at Volume Thresholds	
	1 Episode	20 Episodes
Number of Patients	1,245,901	1,141,630
Percentage of Parts A & B Costs – TIN	3.21%	2.87%
Percentage of Parts A & B Costs – TIN-NPI	3.21%	1.87%

2.2 Frequently Attributed Specialties

Table 2 shows the top 10 most frequently attributed specialties for this measure, using a 20-episode testing volume threshold. As intended, the measure includes a range of clinicians who provide outpatient management and care for heart failure, including specialists and primary care clinicians. These specialties are also consistent with input provided by stakeholders, including patient and family partners (PFPs), during the measure development process. PFPs identified cardiologists, primary care physicians, pulmonologists, and electrophysiologists among others as being part of their care team. The Draft Cost Measure Methodology contains more information about the billing codes used to attribute the measure to clinicians.

Table 2 also provides metrics to show the breakdown of the share of episodes covered by each specialty. Cardiologists make up around half of all clinicians who meet the testing volume threshold (51.6%). Internal medicine clinicians are the second most frequently attributed specialty, comprising 19.5% of all attributed clinicians.

Finally, Table 2 shows the percentage of each specialty that is covered by this measure. 29.3% of all cardiologists who billed at least one Part B physician/supplier claim in 2019 have at least 20 Heart Failure episodes, reflecting the high prevalence of the condition. This metric should be interpreted with caution as specialties vary in size; for instance, the share of internal medicine clinicians covered by the measure appears low (2.8%), but it is a very large specialty.

Table 2. Count and Attribution Frequency of the Top 10 Attributed Specialties at a Testing Volume of 20 Episodes

Rank	Specialty	Number of TIN-NPIs Attributed	Percent of Specialty Among All Attributed TIN-NPIs	Percent of All Episodes	Percent of Specialty (TIN-NPI) Attributed to Measure
1	Cardiology	10,232	51.6%	54.1%	29.3%
2	Internal Medicine	3,868	19.5%	14.8%	2.8%
3	Interventional Cardiology	1,632	8.3%	7.9%	26.5%
4	Family Practice	1,561	7.9%	5.7%	1.4%
5	Nurse Practitioner	1,242	6.3%	5.1%	0.6%
6	Cardiac Electrophysiology	1,227	6.2%	7.9%	35.2%
7	Physician Assistant	288	1.5%	1.1%	0.3%
8	Nephrology	285	1.4%	1.0%	2.4%
9	Pulmonary Disease	248	1.3%	0.8%	1.7%
10	Advanced Heart Failure and Transplant Cardiology	237	1.2%	1.1%	26.6%

2.3 Reliability

Reliability evaluates a measure's ability to consistently differentiate the performance of one clinician from another. The signal-to-noise ratio is used to estimate reliability, which indicates how much of the variation in the measure score is explained by differences among clinicians performance (i.e., signal) instead of differences within each clinician's performance (i.e., noise). Specifically, noise is the variation from one episode to another during the performance period for a particular clinician.

Table 3 shows reliability metrics at various testing volume thresholds. While higher thresholds yield higher reliability, it is at the cost of further reducing the number of clinicians and clinician groups eligible for the measure, which would reduce the potential impact of the measure. For the purposes of field testing, we used a 20-episode volume threshold (bolded in the table below); for simplicity, we use this threshold across all measures. If the measure is implemented in the Merit-based Incentive Payment System (MIPS) in the future, CMS will establish a case minimum through notice-and-comment rulemaking.

Table 3. Sample Size, Mean Reliability, and Proportion of Clinicians above Moderate Reliability at Various Testing Volume Thresholds

Case Volume Threshold	TIN			TIN-NPI		
	Number of TINs	Mean Reliability	Percent Above 0.4	Number TIN-NPIs	Mean Reliability	Percent Above 0.4
10	17,554	0.60	77.7%	41,553	0.52	70.1%
20	**10,667**	**0.71**	**94.1%**	**19,829**	**0.62**	**89.9%**
30	7,680	0.77	98.9%	11,793	0.68	97.5%

At the testing volume of 20 episodes, the Heart Failure measure has mean reliability of 0.71 at the TIN level, and 0.62 at the TIN-NPI level (Table 3). CMS generally considers 0.4 as the threshold indicating 'moderate' reliability and 0.7 as 'high' reliability, which is supported by previous work into reliability and the threshold was finalized in the CY 2022 Physician Fee

Schedule final rule.[2,3] Additionally, most TINs and TIN-NPIs meet or exceed the threshold of moderate reliability, at 94.1% and 89.9% respectively.

2.4 Validity

Validity evaluates whether the cost measure is able to quantify the construct that it aims to measure, which is the cost performance of clinicians. Validity is tested empirically by examining the association between the measure score and high-cost events that drive the measure score, such as downstream complications and consequences of care, and the correlation between cost and quality measures.

The measure score reflects the average risk-adjusted and specialty-adjusted cost of episodes that were attributed to a particular clinician or group. The risk-adjusted and specialty-adjusted cost neutralizes the effects of risk factors deemed to be outside of a clinician's influence (e.g., pre-existing conditions, age, or indicators of clinical severity) on the standardized cost[4] of clinically relevant services observed during a care episode. For the full list of factors that were risk adjusted and the specialty adjustment methodology for this measure, please see the Draft Cost Measure Methodology.

Table 4 shows that downstream high-cost events such as emergency department visits or acute inpatient stay are correlated with the cost measure. Specifically, providers with higher frequency of high-cost events, indicated by the quartiles of the high-cost event frequency, have higher mean scores at both the TIN and TIN-NPI reporting levels. Providers with the highest frequency of high-cost events, at Q3 or Q4, also have mean scores that are higher than the overall mean score for all providers that indicate lower-than-average performance.

The results show that the measure score is not impacted until a provider has substantially more high-cost events than their peers. Specifically, providers with lowest frequency of high-cost events, either 0% or at Q1 or Q2, had lower mean score than the overall mean score for all providers. In other words, the measure differentiates performance based the relative frequency of high-cost events compared to peers instead of the simple presence of high-cost events during a performance period.

---

[2] Mathematica, Inc., "Memorandum: Reporting Period and Reliability of AHRQ, CMS 30-Day and HAC Quality Measures – Revised," http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/Downloads/HVBP_Measure_Reliability-.pdf.

[3] CMS, "Medicare Program; CY 2022 Payment Policies Under the Physician Fee Schedule and Other Changes to Part B Payment Policies; Medicare Shared Savings Program Requirements; Provider Enrollment Regulation Updates; and Provider and Supplier Prepayment and Post-Payment Medical Review Requirements," 86 FR 64996-66031.

[4] Claim payments are standardized to account for differences in Medicare payments for the same service(s) across Medicare providers. Payment standardized costs remove the effect of differences in Medicare payment among health-care providers that are the result of differences in regional health-care provider expenses measured by hospital wage indexes and geographic price cost indexes (GPCIs) or other payment adjustments, such as those for teaching hospitals. For more information, please refer to the "CMS Price (Payment) Standardization - Basics" and "CMS Price (Payment) Standardization - Detailed Methods" documents posted on the Payment Standardization webpage. (https://resdac.org/articles/cms-price-payment-standardization-overview)

Table 4: Mean and Standard Deviation of Score Stratified by the Frequency of Observing High-Cost Events

High-Cost Events	TIN			TIN-NPI		
	Quartiles (Frequency Range)	Mean Score	Score Standard Deviation	Quartiles (Frequency Range)	Mean Score	Score Standard Deviation
All Episodes	N/A	$12,552	$3,458	N/A	$12,427	$3,617
Emergency Department	0%	*	*	0%	$5,867	*
	Q1: (3.0% - 34.6%)	$11,185	$3,572	Q1: (3% - 31.8%)	$11,214	$3,514
	Q2: (34.6% - 41.7%)	$12,264	$3,036	Q2: (31.8% - 38.9%)	$12,045	$3,367
	Q3: (41.7% - 48.2%)	$12,840	$3,076	Q3: (38.9% - 46.2%)	$12,768	$3,411
	Q4: (48.2% - 90.9%)	$13,908	$3,543	Q4: (46.2% - 86.4%)	$13,683	$3,701
Acute Inpatient Stay	0%	$6,018	$3,086	0%	$6,873	$2,653
	Q1: (1.8% - 18.8%)	$10,136	$3,087	Q1: (1.5% - 15.1%)	$10,008	$2,920
	Q2: (18.8% - 24.1%)	$12,056	$2,676	Q2: (15.1% - 20.6%)	$11,687	$2,830
	Q3: (24.1% - 30.3%)	$13,099	$2,787	Q3: (20.6% - 27.1%)	$13,058	$3,020
	Q4: (30.3% - 84.4%)	$14,975	$3,280	Q4: (27.1% - 70.0%)	$15,060	$3,533

Note: Cells with '*' was not possible to be calculated due to insufficient sample size

We also examined the correlation between the cost measure and related quality measures to test the relationship between these different metrics. While there are important limitations to this analysis, it can provide some indication of whether there is variation in cost with different levels of quality. In brief, we note the following key points for interpreting the strength and direction of correlations:

- A strong inverse correlation – good performance on cost with poor quality performance – would indicate that variation in cost is solely reflective of variation in quality. This suggests that care stinting could be a concern.
- A weak correlation between cost and quality in either a positive or negative direction indicates variation in cost at any given level of quality. This suggests that cost performance can be improved without negatively impacting quality.

In general, the direction of correlations indicate the following:

- Positive correlations with quality measures indicate that clinicians providing better quality care on that particular metric tend to also have lower costs. That is, clinicians who have high rates of performing specific quality actions (as measured through process measures) or achieve better patient health outcomes (as measured through outcomes measures) tend to have lower costs of care. As such, these associations could represent ways to lower costs while also providing high-quality care
- A negative correlation between a cost and quality measure does not indicate an absence of cost improvement potential consistent with high-value care. This is because other

approaches to improving cost performance (e.g., patient education) may not be captured by the selected quality measure.

There are several key considerations regarding the conceptual relationship between measures and data limitations when interpreting the results. The extent to which correlations can provide meaningful information depends on what is being measured. Ideally, measures should apply to the same care provided for the same patient cohort for the same time horizon. Correlations with a quality measure that focuses on outcomes for the same patient cohort may be more informative than a broadly applicable process measure that applies across a wide range of conditions. Clinicians select only 6 MIPS quality measures to report and are generally required to only report one outcome or high-priority measure. This selective reporting likely biases the observed sample.

Table 5 shows the correlation between the Heart Failure cost measure and related MIPS quality measures. To aid interpretability, the direction of all non-inverse quality measures is inverted so that a lower score indicates better performance for all cost and quality measures. Quality measures are selected based on their clinical proximity to the cost measure, such as assessing quality actions related to a similar patient cohort, and the number of clinicians with both cost and quality measures.

There are very low numbers of TINs and TIN-NPIs that have both the quality and cost measure, providing a limited sample for this analysis. There are 10,667 TINs and 19,829 TIN-NPIs that meet the testing volume threshold for Heart Failure; across all the quality measure pairs tested, the highest number of overlapping providers was 1,075. In general and for the only 2 statistically significant results, the correlations are weakly positive, suggesting that cost performance can be improved without negatively impacting quality. There are no strong inverse correlations, suggesting that care stinting is not a concern within the limitations of this analysis.

Table 5. Correlation between Quality Measures and Cost Measure

Related Quality Measure (Type)	TIN or TIN-NPI	Number of Entities with Both Cost and Quality Measure	Pearson Correlation	P-value
Q005: Heart Failure (HF): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD) (Process)	TIN	225	0.066	0.321
	TIN-NPI	996	0.043	0.178
Q008: Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD) (Process)	TIN	239	0.080	0.216
	TIN-NPI	1,075	0.029	0.346
Q377: Functional Status Assessments for Congestive Heart Failure (Process)	TIN	72	0.069	0.563
	TIN-NPI	222	-0.009	0.892
Q438: Statin Therapy for the Prevention and Treatment of Cardiovascular Disease (Process)	TIN	236	0.228	< 0.001*
	TIN-NPI	856	0.044	0.203
Q458: All-cause Hospital Readmission^ (Outcome)	TIN	697	0.162	0.000*

*P-value <0.05 indicates statistical significance
^Replaced from MIPS CY 2021 onwards with a re-specified version: Q479 Hospital-Wide, 30-Day, All-Cause Unplanned Readmission (HWR) Rate for MIPS Groups

2.5 Performance Gap

Table 6 shows the distribution of the measure score for TIN and TIN-NPI levels. There are substantial variation observed in the measure score in both TIN and TIN-NPI levels, indicated by the interquartile ranges, standard deviations, and coefficients of variation. The 90th percentile of score is approximately double the 10th percentile at both TIN and TIN-NPI levels. The results highlight an opportunity for improvement by closing the gap between the most and least efficient providers.

Table 6. Distribution of the Measure Score

Metric	TIN	TIN-NPI
Mean Score	$12,552	$12,427
Score Interquartile Range (IQR)	$3,912	$4,506
Standard Deviation	$3,458	$3,617
Coefficient of Variation	0.28	0.29
Score Percentile		
10th	$8,625	$8,234
25th	$10,383	$9,953
50th	$12,176	$12,020
75th	$14,295	$14,458
90th	$16,999	$17,098

2.6 Risk Adjustment and Stratification

2.6.1 Discrimination

Discrimination is a statistical criterion that evaluates the measure's ability to distinguish high-cost episodes from low-cost episodes, or the ability to explain the variance in cost of individual episodes. The amount of variance explained is estimated by the R-squared metric with the range between 0 and 1. The R-squared for the measure is 0.100, and 0.100 after adjusting for the model's complexity based on the number of risk adjustors used. In other words, 10% of the variation in the actual observed cost of episodes is explained by the risk adjustment model.

The remaining unexplained variance is due to variation in factors that are not adjusted for by the measure, such as the clinician's performance. The objective of a cost measure is to evaluate and differentiate the performance of clinicians. Therefore, achieving high explained variance is not essential because not all of the variation in cost of care should be adjusted. In collaboration with the members of the Heart Failure Clinician Expert Workgroup, this measure only adjusts for factors that are deemed to be outside of the reasonable influence of clinicians. Please see the Draft Heart Failure Measure Methodology for more information on the full list of risk adjustors.

2.6.2 Calibration

Calibration evaluates the consistency of the measure in estimating episode cost across the full range of resource use patterns in the population. Calibration is estimated by the average predictive ratios across groups within the population, specifically groups are partitioned by deciles of expected episode cost. The predictive ratio is calculated using the formula of average expected cost / average observed cost for all episodes in each decile. A well-calibrated measure should have predictive ratios close to 1.00 across all deciles. In other words, such

results show that the measure is consistent because it does not under- or over-predict cost throughout the range of resource use patterns in the population.

Table 7 shows that that the risk adjustment model is consistent, with the average predictive ratios observed to be close to 1.00 across all deciles, with the range between 0.95 and 1.01. The largest deviation is at the 1st decile where the average predictive ratio is 0.95, which is a larger deviation than the remaining deciles, but the monetary magnitude of the deviation is only a few hundred dollars. Overall, the risk adjustment model does not over- or under-predict cost across the full range of resource use patterns in the population.

Table 7. Predictive Ratio by Decile of Predicted Episode Cost

Decile	Average Predictive Ratio
Decile 1	0.95
Decile 2	0.99
Decile 3	1.00
Decile 4	1.00
Decile 5	1.01
Decile 6	1.01
Decile 7	1.01
Decile 8	1.01
Decile 9	1.00
Decile 10	0.99

2.7 Social Risk Factor Analysis

Beyond clinical characteristics of patients, the cost of care may be influenced by non-clinical factors related to a patient's social risk factors (SRFs), such as race, income, education, and employment. At the program level, MIPS adjusts for SRFs using the MIPS Complex Patient Bonus to ensure clinicians or groups treating more complex patients are not disadvantaged.[5]

To examine the extent that SRFs can mask the underlying performance of clinicians, this analysis added SRFs to the base risk adjustment model to examine their potential impact to clinicians' scores and goodness of fit of the risk adjustment model. The base risk adjustment model includes the standard set of risk adjustors from the CMS-Hierarchical Condition Categories (HCC) version 22 in 2016, disability status, End-Stage Renal Disease (ESRD) status, comorbidity interaction variables, recent long-term care use, HCC count, and measure-specific clinical risk adjustors. For the full list of factors that were risk adjusted for this measure, please see the Draft Cost Measure Methodology.

The base model was compared against 3 models that included additional SRFs from the American Community Survey, Common Medicare Environment described below:

1. Dual Medicare-Medicaid eligibility status:
    - Base model
    - Medicare-Medicaid eligibility status: full, partial, or non-dual status
2. All Socioeconomic Status (SES) Variables Model:

[5] https://qpp-cm-prod-content.s3.amazonaws.com/uploads/966/QPP%20COVID-19%20Response%20Fact%20Sheet.pdf

- Base model
- Dual Medicare-Medicaid eligibility status: full, partial, or non-dual status
- Sex: female or male
- Race: Asian, Black, Hispanic, North American Native, White, and other
- Neighborhood Income Level: low income (median income < 33rd percentile nationally), medium income (median income between 33rd and 66th percentiles), and high income (median income > 66th percentile)
- Neighborhood Education Rate: majority less than high school, high school, or greater than high school
- Neighborhood Employment Rate: less than or greater than 10%

3. Agency for Healthcare Research and Quality (AHRQ) SES Index Model:
   - Base model
   - Dual Medicare-Medicaid eligibility status
   - Sex: female or male
   - Race: Asian, Black, Hispanic, North American Native, White, and other
   - AHRQ SES Index Score: calculated using the AHRQ's scoring algorithm, based on data from the American Community Survey

Table 8 shows the percent of providers who would see their performance shifted, based on their percentile ranking, using the new SRF models. There are marginal changes to the performance at both TIN and TIN-NPI reporting levels compared to the base model. Across the 3 models, over 90% TINs and TIN-NPIs would experience small and likely inconsequential swing in their performance, of less than 5% compared to the base model. Additionally, the performance ranking has Spearman correlation values of greater than 0.99 between the base model and each of the SRF models, which also suggests that adjusting for SRF variables does not fundamentally change how the measure evaluates provider performance.

Table 8. Clinicians' Performance Shift Measured by the Change in the Average Ratio of Observed-to-Expected Cost

Model	Reporting Level	Proportion of Clinicians Affected at Various Levels of Performance Shift						
		-10% or more	-9% to -6%	-5% to -2%	-1% to 1%	2% to 5%	6% to 9%	10% or more
Model 1	TIN	1.0%	2.4%	12.4%	64.4%	19.4%	0.2%	0.1%
	TIN-NPI	0.6%	1.6%	11.2%	72.4%	13.9%	0.2%	0.1%
Model 2	TIN	2.7%	4.0%	14.5%	46.6%	30.1%	1.8%	0.3%
	TIN-NPI	1.7%	3.6%	14.3%	53.3%	25.9%	1.0%	0.2%
Model 3	TIN	2.6%	3.8%	14.4%	47.7%	29.7%	1.5%	0.3%
	TIN-NPI	1.6%	3.5%	14.0%	54.4%	25.5%	0.9%	0.2%

2.8 Impact of Exclusions

Table 9 displays descriptive statistics of all episodes meeting the measure's triggering logic, excluded episodes, and final reportable episodes at both TIN and TIN-NPI levels. These exclusion criteria ensure that the reportable episode populations are more homogenous and comparable than all episodes meeting triggering logic. It is worth noting that only the observed cost is shown, which has not been risk adjusted for using our risk adjustment model. Therefore, the differences in cost may appear much smaller after risk adjustment than as-is.

Overall, exclusion criteria decrease the distribution of observed cost of all episodes meeting trigger logic, from the mean of $19,885 to $12,266 at the TIN level and $11,142 at the TIN-NPI level. All of the exclusion criteria have higher mean observed cost than all episodes meeting triggering logic.

Episodes shorter than one year are excluded to ensure sufficient observation of chronic care that is often intermittent and sparse over a long period of time. Since the cost measure scales the episode cost to one year to aid comparison across episodes of different lengths, a shorter episode may artificially appear more expensive because the cost is distributed over fewer days, as observed to have a mean of $60,589. Although these episodes are excluded during the performance period being examined, they are likely to be included in the following performance period once the episode length is longer than one year.

Episodes with beneficiary death during an episode also have higher mean observed cost than all episodes meeting triggering logic, at $45,480. Similar to the minimum length criterion, excluding these episodes ensures that the truncated observable window does not artificially make the scaled cost to appear higher.

Episodes that are not reliably predicted by the risk adjustment model are excluded because their observed costs deviate substantially from the projected cost for a given patient risk profile. Table 9 shows substantial differences between these episodes and all episodes meeting triggering logic, with mean observed cost of $37,000 versus $19,885.

At the TIN-NPI level, some episodes are excluded because they cannot be attributed to an individual clinician. As such, they cannot be used in the measure for TIN-NPI reporting.

Episodes with the following comorbidities were also excluded: amyloidosis, congenital heart disease, high output heart failure, hypertrophic cardiomyopathy, other infiltrative disease, and peripartum cardiomyopathy. These episodes are excluded with input from stakeholders during the development process to ensure that the patient cohort is clinically homogenous and comparable. While these groups constitute a small number of episodes, their mean observed cost are higher than all episodes meeting triggering logic.

The largest exclusions come from applying the testing volume threshold to ensure a sufficient sample size for the measure.

Table 9: Cost Statistics for Measure Exclusions

Exclusion Criteria	Episodes		Observed Episode Cost					
				Percentile				
	Count	Percent of All Episodes Meeting Trigger Logic	Mean	10th	25th	50th	75th	90th
All Episodes Meeting Triggering Logic	2,264,500	100.0%	$19,885	$1,285	$2,893	$8,385	$23,356	$48,150
Episode Length Less Than One Year	246,662	10.9%	$60,589	$4,741	$11,923	$30,277	$71,939	$140,713
Beneficiary Death in Episode	454,702	20.1%	$45,480	$3,800	$9,917	$24,665	$53,838	$103,856
Outlier	35,052	1.6%	$37,000	$1,545	$3,036	$38,084	$73,432	$73,432
No Attributed NPI	302,198	13.4%	$27,084	$1,929	$4,859	$13,559	$32,609	$63,553
Amyloidosis	6,708	0.3%	$37,918	$2,614	$6,626	$20,161	$49,482	$88,383
Congenital Heart Disease	34,255	1.5%	$29,353	$2,053	$5,030	$13,456	$33,025	$67,359
High Output Heart Failure	503	< 0.1%	$32,246	$2,160	$6,817	$17,078	$39,754	$69,933
Hypertrophic Cardiomyopathy	16,512	0.7%	$28,197	$1,872	$4,490	$12,171	$32,073	$65,651
Other Infiltrative Disease	12,282	0.5%	$28,086	$1,807	$4,184	$11,779	$31,771	$66,314
Peripartum Cardiomyopathy	261	< 0.1%	$28,247	$1,374	$2,688	$10,252	$28,882	$64,968
TIN does not Meet Testing Volume Threshold	241,926	10.7%	$22,577	$1,404	$3,436	$10,146	$26,487	$55,005
TIN-NPI does not Meet Testing Volume Threshold	910,670	40.2%	$21,409	$1,317	$3,142	$9,196	$24,693	$51,528
Reportable Episodes (if all clinicians reported as TIN at the testing volume threshold)	**1,546,044**	**68.3%**	**$12,266**	**$1,126**	**$2,289**	**$6,163**	**$16,429**	**$32,862**
Reportable Episodes (if all clinicians reported as TIN-NPI at the testing volume threshold)	**853,265**	**37.7%**	**$11,142**	**$1,064**	**$2,064**	**$5,424**	**$14,517**	**$30,260**