NATIONAL QUALITY FORUM
Driving measurable health improvements together

Cost and Efficiency, Spring 2020 Cycle: CDP Report

TECHNICAL REPORT
FEBRUARY 22, 2021

This report is funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001.

http://www.qualityforum.org

Contents

Executive Summary

It is estimated that healthcare spending in the United States (U.S.) will increase its share of gross domestic product (GDP) from 17.9 percent in 2017 to 19.4 percent by 2027.[1] This level of healthcare spending and growth has the potential to increase federal deficits and debt further or crowd out spending for other important national priorities. These economic realities require performance measures that can accurately capture spending, particularly spending that is the result of inefficient or poor-quality care.

Reducing wasteful spending requires the coordination of multiple providers and care settings to ensure efficient, high quality patient transitions. Thus, cost and quality measures used together can help assess efficiency and value of care delivered and drive improvement in the U.S. healthcare system.

The National Quality Forum (NQF) Cost and Efficiency Standing Committee oversees NQF's portfolio of Cost and Efficiency measures. For the spring 2020 cycle, the Standing Committee deliberated on several overarching issues related to cost and efficiency measurement, specifically reliability thresholds, threats to validity, and the adjustment for social risk factors. The Standing Committee underscored the importance of aligning how patient risk is handled in payment with expected costs in performance measurement. The Standing Committee noted the importance of excluding clinically unrelated services, such as exclusion of certain downstream costs that are not associated with the measure focus or within control of the accountable unit. Additionally, it encouraged developers to examine the complex role of social risk factors in cost and resource use measures. These overarching issues were factored into the Standing Committee's evaluation and recommendations for all the measures under endorsement consideration.

For this project, the Standing Committee evaluated six newly submitted measures against NQF's measure evaluation criteria. The Standing Committee recommended three measures for endorsement and did not recommend three measures for endorsement. The Consensus Standards Approval Committee (CSAC) upheld the Standing Committee's recommendations.

Endorsed Measures:

- NQF #3561 Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities (Acumen, LLC)
- NQF #3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals (Acumen, LLC)
- NQF #3575 Total Per Capita Cost (TPCC) (Acumen, LLC)

Measures Not Endorsed:

- NQF #3563 Medicare Spending Per Beneficiary – Post Acute Care Measure for Skilled-Nursing Facilities (Acumen, LLC)
- NQF #3564 Medicare Spending Per Beneficiary – Post Acute Care Measure for Home Health Agencies (Abt Associates)
- NQF #3574 Medicare Spending Per Beneficiary (MSBP) Clinician (Acumen, LLC)

Brief summaries of the spring 2020 measures are included in the body of the report. Detailed summaries of the Standing Committee's discussion and ratings of the criteria for each measure are in Appendix A.

Introduction

In 2018, healthcare spending in the U.S. reached $3.6 trillion or approximately $11,172 per person.[1] This level of spending accounted for 17.7 percent of GDP. Forecasts from 2018 to 2027 estimate that healthcare spending will outpace GDP growth by 0.8 percent.[1] This increase will raise the healthcare share of GDP from 17.9 percent in 2017 to 19.4 percent by 2027.[1] Spending on the overall Medicare program is growing rapidly as well—from 15 percent of federal spending in 2018 to an expected 17 percent by 2027.[1]

Improving health system efficiency has the potential to simultaneously reduce the rate of cost growth and improve the quality of care provided. Cost measures are the building blocks to efficiency and value. It is important to note that cost and resource use measures should be used in the context of and reported with quality measures.

A key area where cost measurement continues to be a critical component to assess the efficiency of the healthcare system is post-acute care (PAC). PAC providers that include inpatient rehabilitation facilities (IRFs), long-term care hospitals (LTCHs), skilled nursing facilities (SNFs), and home health agencies (HHAs) offer rehabilitation and recuperation services to patients typically after an acute hospitalization. In 2018, the Medicare fee-for-service (FFS) program spent $58.6 billion on PAC services.[2] This level of spending can be attributed to a wide variety of causes, including important clinical care for patients but also high drugs costs and poor care coordination.[2]

NQF Portfolio of Performance Measures for Cost and Efficiency Project

The Cost and Efficiency Standing Committee (Appendix C) oversees NQF's portfolio of Cost and Efficiency measures (Appendix B), which includes both condition-specific and non-condition-specific measures. The Cost and Efficiency Standing Committee's charge is to assess cost and resource use measures and efficiency more broadly, including measures assessing the efficiency of healthcare delivery. The Standing Committee seeks to take a more holistic view of drivers of healthcare spending and identify sources of inefficiency and waste across the system. This portfolio contains 10 cost/resource use measures (see Table 1 below).

Table 1. NQF Cost and Efficiency Portfolio of Measures

	Cost/Resource Use
Condition-Specific	7
Non-condition-Specific	3
Total	10

Cost and Efficiency Measure Evaluation

On July 10, 2020, the Cost and Efficiency Standing Committee evaluated six new measures against NQF's Cost and Efficiency Evaluation Criteria (see Table 2 below).

Table 2. Cost and Efficiency Measure Evaluation Summary

	Maintenance	New	Total
Measures reviewed	0	6	6
Measures endorsed	0	3	3
Measures not recommended for endorsement	0	3	3
Reasons for not recommending		Importance – 0 Scientific Acceptability – 3 Overall Suitability – 0 Competing Measure – 0	

Comments Received Prior to Committee Evaluation

NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). In addition, NQF accepts comments for a continuous 16-week period during each evaluation cycle via an online tool located on the project webpage. For this evaluation cycle, the commenting period opened on May 11, 2020 and closed on September 14, 2020. The Cost and Efficiency project received 18 comments, which were shared with the Standing Committee prior to the measure evaluation meeting (Appendix F).

Comments Received After Committee Evaluation

The continuous 16-week public commenting period with NQF member support closed on September 14, 2020. Following the Standing Committee's evaluation of the measures under review, NQF received 17 comments pertaining to the draft report and to the measures under review. All comments for each measure under review have been summarized in Appendix A.

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support (either *support* or *do not support*) for each measure submitted for endorsement consideration to inform the Standing Committee's recommendations. Two NQF members provided their expression of non-support for NQF #3574 *Medicare Spending Per Beneficiary (MSBP) Clinician*, and NQF #3575 *Total Per Capita Cost (TPCC)*.

Overarching Issues

During the Standing Committee's discussion of the measures, several overarching issues emerged that were factored into the Standing Committee's ratings and recommendations for multiple measures and are not repeated in detail with each individual measure.

Reliability Thresholds and Variations by Case Volume

Variation in reliability due to the number of cases in practices or facilities was a point of discussion by the Standing Committee, as greater variance can be inherent in practices with lower case volume. For several of the measures reviewed this cycle, the Standing Committee raised concerns that the signal-to-noise or split-sample reliability statistics for practices with small case volumes may not be sufficient for the measure to be considered reliable. For several review cycles, the Standing Committee has recognized the challenge of achieving acceptable thresholds for measure score reliability statistics.

Threats to Validity

Validity refers to the correctness of measurement. Threats to validity include other aspects of the measure specifications, such as inappropriate exclusions, lack of appropriate risk adjustment or risk stratification for outcome and resource use measures, use of multiple data sources or methods that result in different scores and conclusions about resource use or quality, and systematic missing or "incorrect" data. Most importantly, a measure may be invalid because the measurement has not correctly captured the concept of cost or resource use.

For some of the measures reviewed this cycle, the Standing Committee raised concerns regarding various overarching threats to measure validity. For several PAC measures, the Standing Committee questioned whether the use of 30- or 60-days post discharge is an appropriate length of time in which PAC settings can influence downstream care decisions. The Standing Committee highlighted the need to empirically evaluate and validate if 30- or 60-days post discharge is the appropriate length of time to capture complications that can be attributed to the respective setting of care.

The Standing Committee also had concerns regarding the lack of alignment on how patient risk is handled in PAC-setting payment programs and the expected costs calculation in the measures. Specifically, there are certain variables that influence cost that are not included in the risk adjustment. The developer discussed this was done due to the concern that certain aspects of the payment system in PAC can be gameable, and there is more alignment in the expected episode cost and the payment program for long-term care hospitals and less for the other PAC measures. With respect to measures exclusions, the Standing Committee discussed approaches to truncation (a method of minimizing the effects of outliers in data) of low- and high- cost episodes and questioned the approach of how death is handled within the episode window. Additionally, the Standing Committee raised concerns regarding clinically unrelated services, highlighting the importance of excluding certain downstream costs/events not associated with the respective episode of care. Lastly, the Standing Committee also raised concerns with how well the risk adjustment models for several of the measures accurately capture patient risk, due to r-squared values ranging from 0.09 to 0.49. These issues were factored into the Standing Committee's evaluation and recommendations for several of the measures.

Social Risk Adjustment

Cost and efficiency measurement is influenced by the care received in a healthcare setting and patient clinical and social risk factors (e.g., age, race, ethnicity, gender, social relationships, residential and community context) since they typically measure the cost or resource use across time and over multiple providers and settings. While the developer did test several social risk factors for the risk adjustment model, none were included in the six measures under review. The Standing Committee noted the need to ensure that providers serving people with social risk factors are not penalized unfairly by a lack of social risk adjustment. While the Standing Committee noted that it is important to maximize the predictive value of a risk-adjustment model, understanding the role that social risk factors play in clinical cost episodes is critical. The impact of social risk factors in cost and efficiency measures is unique in that these factors may ultimately increase overall costs through poor transitions and hand-offs or, potentially, lower resource use because of access-to-care challenges. Each cost and efficiency measure should be examined on a case-by-case basis to understand the role of patient social risk factors in the measure.

Summary of Measure Evaluation

The following summaries of the measure evaluation highlight the major issues that the Standing Committee considered. Details of the Standing Committee's discussion and ratings of the criteria for each measure are included in Appendix A.

NQF #3561 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Inpatient Rehabilitation Facilities (Acumen, LLC): Endorsed

Description: The *Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facility* (MSPB-PAC IRF) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each IRF's efficiency relative to that of the national median IRF. Specifically, the measure assesses Medicare spending by the IRF and other healthcare providers during a Medicare Spending Per Benificiary (MSPB) episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each IRF divided by the episode-weighted median MSPB-PAC Amount across all IRFs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all IRFs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF admissions from October 1, 2015 through September 30, 2017.

Measure Type: Cost/Resource Use; **Level of Analysis**: Facility; **Setting of Care**: Post-Acute Care; **Data Source**: Assessment Data, Claims, Enrollment Data, Other

During the measure evaluation meeting on July 10, 2020, the Standing Committee did not vote on the recommendation for endorsement because the Standing Committee did not pass the measure on validity—a must-pass criterion.

Several Standing Committee members noted that the measure was developed as a part of the IMPACT Act and is a legislative requirement. The Standing Committee reviewed publicly reported measure score data provided by the developer for all U.S. providers under Medicare's inpatient rehabilitation facilities (IRF) Prospective Payment System (PPS) with 20 or more eligible episodes in the reporting period of 2016-2017. The Standing Committee acknowledged that the developer demonstrated significant variability in resource use across IRFs and agreed that this measure addresses a high resource use aspect of healthcare and that there is an opportunity for improvement. The Standing Committee passed the measure on the importance to measure and report criterion.

Prior to the Standing Committee meeting, this measure was reviewed by the NQF Scientific Method Panel (SMP). The SMP had raised concerns regarding outlier exclusions and recommended that the developer report the distribution of outlier exclusion across facilities to ensure that they don't concentrate in a limited number of facilities. The developers provided written responses and results that show excluded outliers are not concentrated on a small number of providers. The SMP was satisfied with the developer responses and ultimately passed the measure with a moderate rating for both reliability and validity. The SMP review and the measure developer responses are summarized in the SMP Spring 2020 SMP Discussion Guide. The Standing Committee noted that the developers demonstrated reliability using signal-to-noise analysis through the Adams' method. The Standing Committee reviewed the reliability results and agreed with the SMP, ultimately passing the measure on

reliability. The Standing Committee raised several concerns regarding the validity of the measure. The Standing Committee questioned the use of 30-days as the appropriate length of time that IRFs can influence downstream care decisions. The Standing Committee highlighted the need to empirically evaluate and validate if 30-days post discharge is an appropriate length of time to capture complications that can be attributed to IRF care. The Standing Committee raised concerns regarding the approach to truncation/winsorization of low- and high-cost episodes and questioned the approach to how death is handled within the episode window. The Standing Committee also questioned how well the model predicts downstream costs and raised concern regarding the lack of adjustment for social risk factors. The developer noted that there was limited impact of social risk factor effects under the current risk adjustment model. The Standing Committee highlighted that accounting for social risk factors associated with the outcome of interest that are outside provider's control reduces bias in measurement.

Lastly, the Standing Committee raised concerns that the expected costs were not aligned with how patient risk is accounted for in IRF payment programs. The developer noted that there is more alignment in the expected episode cost and the payment program for long-term care hospitals (NQF #3562) and less for IRF and the other PAC measures. The Standing Committee did not vote to pass the measure on the validity criterion, noting additional concerns beyond the SMP review related to the lack of social adjustment in the risk adjustment model, alignment of patient risk between expected costs and IRF payment programs, and additional threats to validity as stated above.

Following the measure evaluation meeting, the measure developer submitted a reconsideration request for this measure, asserting that the NQF evaluation criteria were not correctly applied and that there was inconsistent application of the evaluation criteria that led to a measure not being recommended for endorsement.

The Standing Committee reevaluated the measure during the October 1, 2020 post-comment web meeting. During the meeting discussion, the developer stated that NQF #3561 purposefully excludes some payment variables, noting that this was an explicit policy decision made by Centers for Medicare & Medicaid Services (CMS), as excessive spending in these settings has historically been driven by increased use of therapy and variability in coding of patient status on assessment instruments.

With respect to social risk adjustment, the developer mentioned that adjusting for social risk factors may mask disparities in care, creating a lower standard of care for beneficiaries with higher social risk, and this could allow for a higher rate of readmissions, complications, etc. among those with high social risk. The developer mentioned that this may be appropriate if such outcomes are outside of a provider's control, but the developer's empirical testing showed that poorer performance for high social risk individuals is closely tied to providers themselves, rather than individual beneficiaries, with especially strong effects in particular settings. The developer argued that due to the negative relationship between social factors and the measure scores, inclusion of certain factors may penalize providers for taking on beneficiaries with high social risk. Finally, the developer stated that including these factors in the model had a minimal impact on provider scores.

The Standing Committee did not achieve quorum during the post-comment web meeting. Therefore, a recording of the meeting and the offline voting surveys were distributed to members. The Standing Committee revoted to recommend the measure for endorsement.

The CSAC expressed no concerns with the Standing Committee's evaluation or recommendation and voted unanimously to endorse the measure.

NQF #3562 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Long-Term Care Hospitals (Acumen, LLC): Endorsed

Description: The *Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals* (MSPB-PAC LTCH) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each LTCH's efficiency relative to that of the national median LTCH. Specifically, the measure assesses Medicare spending by the LTCH and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each LTCH divided by the episode-weighted median MSPB-PAC Amount across all LTCH facilities. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all LTCHs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., LTCH admissions from October 1, 2015 through September 30, 2017.

Measure Type: Cost/Resource Use; **Level of Analysis**: Facility; **Setting of Care**: Post-Acute Care; **Data Source**: Assessment Data, Claims, Enrollment Data, Other

The Standing Committee recommended the measure for initial endorsement.

Several Standing Committee members noted that the measure was developed as a part of the IMPACT Act and is a legislative requirement. The Standing Committee acknowledged that the developer demonstrated significant variability in resource use in long-term care hospitals with opportunities for improvement. The Standing Committee did note that there is geographic variability in the availability of LTCH settings across the country. The Standing Committee generally agreed that this measure addresses a high resource use aspect of healthcare. They also acknowledged that the developer demonstrated variation in post-acute care spending to warrant measurement. The Standing Committee passed the measure on importance to measure and report criterion.

Prior to the Standing Committee meeting, this measure was reviewed by the NQF SMP. The SMP had raised concerns regarding outlier exclusions and recommended that the developer report the distribution of outlier exclusion across facilities to ensure that they don't concentrate in a limited number of facilities. The developers provided written responses and results that show excluded outliers are not concentrated on a small number of providers. The SMP was satisfied with the developer responses and ultimately passed the measure with a moderate rating for both reliability and validity. The SMP review and the measure developer responses are summarized in the SMP Spring 2020 SMP Discussion Guide. The Standing Committee noted that the developers demonstrated reliability using signal-to-noise analysis through the Adams' method. The Standing Committee reviewed the reliability results and agreed with the SMP that this measure was reliable and voted to pass the measure on reliability. The Standing Committee expressed some concerns related to the validity of the measure. It questioned the use of 30 days as the appropriate length of time that LTCHs can influence downstream care decisions. The Standing Committee highlighted the need to empirically evaluate and validate if 30 days post-discharge period is the appropriate length of time to capture complications that can be

attributed to LTCHs. The Standing Committee raised concerns that the calculation of expected cost is not aligned with LTCH payment programs. The developer noted that there is more alignment in the expected episode cost and the payment program for LTCH and less for IRF and the other PAC measures. With respect to risk adjustment, the Standing Committee acknowledged that the adjusted r-squared for this LTCH measure is higher than the other PAC measures (r-squared value of 0.4894). Additionally, the Standing Committee highlighted the importance of adjusting for social risk factors to reduce bias in measurement. Ultimately, the Standing Committee passed this measure on validity.

The Standing Committee regarded the measure as feasible with no concerns. The Standing Committee passed the measure on use and usability, acknowledging that this is a new measure and is publicly reported as part of the CMS LTCH Quality Reporting Program. The Standing Committee discussed related and competing measures during its post-comment web meeting on October 1, 2020.

The CSAC expressed no concerns with the Standing Committee's evaluation or recommendation and voted unanimously to endorse the measure.

NQF #3563 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Skilled Nursing Facilities (Acumen, LLC): Not Endorsed

Description: *The Medicare Spending Per Beneficiary – Post Acute Care Measure for Skilled Nursing Facilities* (MSPB-PAC SNF) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each SNF's efficiency relative to that of the national median SNF. Specifically, the measure assesses Medicare spending by the SNF and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each SNF divided by the episode-weighted median MSPB-PAC Amount across all SNFs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all SNFs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., SNF admissions from October 1, 2015 through September 30, 2017.

Measure Type: Cost/Resource Use; **Level of Analysis**: Facility; **Setting of Care**: Post-Acute Care; **Data Source**: Assessment Data, Claims, Enrollment Data, Other

During the measure evaluation meeting on July 10, 2020, the Standing Committee did not vote on the recommendation for endorsement at the meeting because the Standing Committee did not reach consensus on validity—a must-pass criterion.

The Standing Committee acknowledged that the measure was developed to address the resource use aspect of the IMPACT Act. The Standing Committee reviewed a range of data demonstrating high impact through differences in PAC payments across SNFs. The Standing Committee agreed that the range of performance demonstrated an opportunity for improvement in reducing the variability in spending. Overall, the Standing Committee passed the measure on the importance of the measure and report criterion.

Prior to the Standing Committee meeting, this measure was reviewed by the NQF SMP. The SMP had raised concerns regarding outlier exclusions and recommended that the developer report the distribution of outlier exclusion across facilities to ensure that they don't concentrate in a limited

number of facilities. The developers provided written responses and results that show excluded outliers are not concentrated on a small number of providers. The SMP was satisfied with the developer responses and ultimately passed the measure with a moderate rating for both reliability and validity. The SMP review and the measure developer responses are summarized in the SMP Spring 2020 SMP Discussion Guide.

For reliability, the Standing Committee reviewed the signal-to-noise analysis and split sample reliability testing conducted and did not raise any major concerns and passed the measure on this criterion. In terms of validity, the Standing Committee reviewed the exclusion of clinically unrelated services provided by the developer and commented on the importance of excluding certain downstream costs not associated with SNF care. The Standing Committee also had concerns that the calculation of expected cost was not aligned with SNF payment programs. The developer noted on the call that there is more alignment in the expected episode cost and the payment program for LTCH and less for the SNF and the other PAC measures. The Standing Committee also questioned how well the model predicts downstream costs (r-squared of 0.11) and raised concerns regarding the lack of including social risk factors. The developer indicated that though each of the social factors tested was statistically significant in the model, they did not improve the model fit and the adjusted r-squared values increased by less than 0.01. The Standing Committee stressed that inclusion of social risk factors should be about minimizing bias and may not always improve model fit. Due to these concerns, the Standing Committee did not reach consensus on validity.

The Standing Committee passed the measure on feasibility, acknowledging that the measure data are routinely collected and that this measure poses no additional data collection burden on providers. The Standing Committee acknowledged that this is a new measure and is publicly reported as part of the CMS SNF Quality Reporting Program. There was some concern that there is a lack of clarity on whether providers have enough information to target improvement, as there was no indication of areas of high or low spending by provider. However, the Standing Committee ultimately passed the measure on use and usability.

During the October 1, 2020 post-comment meeting, the Standing Committee considered the comments received and the developer's responses related to validity. The Standing Committee did not achieve quorum during the post-comment web meeting. Therefore, a recording of the meeting and the offline voting surveys were distributed to members of the Standing Committee. The Standing Committee revoted on the measure and did not pass the measure on validity. The developer submitted a reconsideration request to the CSAC for this measure. The developer cited three areas of concern related to the review of the measures, namely, inconsistency in the Standing Committee deliberations and process, specifically voting; misapplication of measure evaluation criteria and guidance; and transparency of the Standing Committee deliberation materials. Related to voting inconsistencies, the developer asserted that the Standing Committee did not consistently apply the voting processes established by NQF guidelines. The developer expressed concerns that the Standing Committee departed from the Consensus Development Process (CDP) and that the Standing Committee's votes are inconsistent in their application of overarching issues on the set of cost measures in the spring 2020 cycle and lack a clear rationale for these decisions. Lastly, the developer asserted that NQF did not follow the CDP's process for ensuring transparency, which they state created challenges for understanding reasons for decisions, and therefore, they could not then respond to the Standing Committee appropriately.

The Cost and Efficiency co-chairs, Cheryl Damberg and Sunny Jhamnani, acknowledged the perspective of the measure developer related to the evaluation of the cost measures reviewed by the Standing Committee during the spring 2020 review cycle. The co-chairs recognized that these measures address a

high resource use aspect of healthcare and that there is an opportunity for improvement in resource use outcomes, and because these measures are used for Medicare payment purposes, it is critical to make sure the measures are valid and can reliably differentiate performances.

The co-chairs reaffirmed that the Standing Committee was responsive to the developer's request for a careful review and reconsideration of measures. During the spring 2020 post-comment meetings, the co-chairs expressed that the Standing Committee remained focused on the criteria where consensus was not reached or where a recommendation was not given for endorsement and provided rationales for why the Standing Committee did not consider the measures to meet NQF criteria. CSAC members noted the complexity of this measure; however, they decided not to reconsider and voted to uphold the Cost and Efficiency Standing Committee's recommendation.

NQF #3564 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Home Health Agencies (Abt Associates): Not Endorsed

Description: The *Medicare Spending Per Beneficiary – Post Acute Care Measure for Home Health Agencies* (MSPB-PAC HH) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each home health (HH) agency's efficiency relative to that of the national median home health agency (HHA). Specifically, the measure assesses Medicare spending by the HHA and other healthcare providers during an MSPB-PAC HH episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each HHA divided by the episode-weighted median MSPB-PAC Amount across all HHAs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all HHAs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on CY 2016-2017 data; i.e., HHA admissions from January 1, 2016 through December 31, 2017.

Measure Type: Cost/Resource Use; **Level of Analysis**: Facility; **Setting of Care**: Home Care; **Data Source**: Assessment Data, Claims, Enrollment Data, Other

During the measure evaluation meeting on July 10, 2020, the Standing Committee did not vote on the recommendation for endorsement at the meeting because the Standing Committee did not reach consensus on reliability and validity—which are must-pass criteria.

The Standing Committee acknowledged that the measure was developed to address the resource use aspect of the IMPACT Act. The Standing Committee reviewed a range of data demonstrating high impact through differences in PAC payments across HHAs. The Standing Committee agreed that the range of performance demonstrated an opportunity for improvement in reducing the variability in spending and passed the measure on the importance of the measure and report criterion.

Prior to the Standing Committee meeting, this measure was reviewed by the NQF SMP. The SMP had raised concerns regarding outlier exclusions and recommended that the developer report the distribution of outlier exclusion across facilities to ensure that they don't concentrate on a limited number of facilities. The developers provided written responses and results that show excluded outliers are not concentrated on a small number of providers. The SMP was satisfied with the developer responses and ultimately passed the measure with a moderate rating for both reliability and validity. The SMP review and the measure developer responses are summarized in the SMP Spring 2020 SMP

Discussion Guide. The Standing Committee noted that this measure has been evaluated by the SMP, which gave passing ratings for both reliability and validity. For reliability, the Standing Committee reviewed the signal-to-noise analysis and split-sample reliability testing conducted by the developer. Some Standing Committee members raised concerns with some of the lower volume quartile reliability scores for the taxpayer identification number-national provider identifier (TIN-NPI) reporting level (0.63 and 0.57 for the S/N and split-sample, respectively), stating that they are low. Several Standing Committee members noted that it may be difficult to differentiate HHAs with smaller number of qualifying episodes. Due to these concerns, the Standing Committee did not reach consensus on reliability.

For validity, the Standing Committee questioned whether HHAs can sufficiently control costs that resulted after their care and questioned the developer's decision to utilize a 60-day episode period. The developer clarified that as the measure emphasized upstream interventions and coordination of care, the costs associated with the amount of care needed during hospitalization or emergency department use can be influenced by HHAs. The developer further clarified that HHA care tended to be long term and that the first 60 days of HHA care is a strong indicator of downstream outcomes. The Standing Committee also had concerns that the calculation of patient risk in expected cost was not aligned with HHA payment programs. The developer noted on the call that there is more alignment in the expected episode cost and the payment program for LTCH and less for the HHA and the other PAC measures. Lastly, the Standing Committee questioned how well the model predicts downstream costs (r-squared of 0.092) and raised concerns regarding the lack of including social risk factors. The developer indicated that though each of the social risk factors tested was statistically significant in the model, they did not improve the model fit. Additionally, the developer commented that the dual eligibility in the social risk factor testing carries a negative coefficient, which would lower expected cost. This would penalize providers for taking care of dual-eligible beneficiaries in certain episodes. The Standing Committee stressed that inclusion of social risk factors should be about minimizing bias and may not always improve model fit. Due to these concerns, the Standing Committee did not reach consensus on validity.

The Standing Committee agreed that this measure would be feasible as all were routinely collected and posed no additional data collection burden on providers. The Standing Committee passed this measure on use and usability, acknowledging that this is a new measure and that it is publicly reported as part of the CMS Home Health Quality Reporting Program.

During the October 1, 2020 post-comment meeting, the Standing Committee considered the comments received and the developer's responses, related to validity. The Standing Committee did not achieve quorum during the post-comment web meeting. Therefore, a recording of the meeting and the offline voting surveys were distributed to members of the Standing Committee. The Standing Committee revoted on the measure and did not pass it on validity.

The developer submitted a reconsideration request to the CSAC for this measure. The developer cited three areas of concerns related to the review of the measures, namely, inconsistency in Standing Committee deliberations and process, specifically voting; misapplication of measure evaluation criteria and guidance; and transparency of the Standing Committee deliberation materials. Related to voting inconsistencies, the developer asserted that the Standing Committee did not consistently apply the voting processes established by NQF guidelines. The developer expressed concerns that the Standing Committee departed from the CDP and that the Standing Committee's votes are inconsistent in their application of overarching issues on the set of cost measures in the spring 2020 cycle and lack a clear rationale for these decisions. Lastly, the developers asserted that NQF did not follow the CDP's process

for ensuring transparency, which they state created challenges for understanding reasons for decisions, and therefore, they could not then respond to the Standing Committee appropriately.

The Cost and Efficiency co-chairs, Cheryl Damberg and Sunny Jhamnani, acknowledged the perspective of the measure developers related to the evaluation of the cost measures reviewed by the Standing Committee during the spring 2020 review cycle. The co-chairs recognized that these measures address a high resource use aspect of healthcare and that there is an opportunity for improvement in resource use outcomes, and because these measures are used for Medicare payment purposes, it is critical to make sure the measures are valid and can reliably differentiate performances.

The co-chairs reaffirmed that the Standing Committee was responsive to the developers' request for a careful review and reconsideration of measures. During the spring 2020 post-comment meetings, the co-chairs expressed that the Standing Committee remained focused on the criteria where consensus was not reached or where a recommendation was not given for endorsement and provided rationales for why the Standing Committee did not think the measures met NQF criteria. CSAC members noted the complexity of this measure; however, they decided not to reconsider and voted to uphold the Cost and Efficiency Standing Committee's recommendation.

NQF #3574 Medicare Spending Per Beneficiary (MSPB) Clinician (Acumen, LLC): Not Endorsed

Description: The *MSPB Clinician* measure assesses the cost to Medicare for services by a clinician and other healthcare providers during an MSPB episode, which focuses on a patient's inpatient hospitalization. The MSPB episode spans from three days prior to the hospital stay ("index admission") through to 30 days following discharge from that hospital. The measure includes the costs of all services during the episode window, except for a limited list of services identified as being unlikely to be influenced by the clinician's care decisions and that are considered clinically unrelated to the management of care. The episode is attributed to the clinician(s) responsible for managing the beneficiary's care during the inpatient hospitalization. The *MSPB Clinician* measure score is a clinician's average risk-adjusted cost across all episodes attributed to the clinician. The beneficiary populations eligible for the *MSPB Clinician* measure include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.

Measure Type: Cost/Resource Use; **Level of Analysis**: Clinician : Group/Practice, Clinician : Individual; **Setting of Care**: Inpatient/Hospital; **Data Source**: Assessment Data, Claims, Enrollment Data, Other

During the measure evaluation meeting on July 10, 2020, the Standing Committee did not recommend the measure for initial endorsement because the Standing Committee did not pass the measure on validity—a must-pass criterion.

The Standing Committee acknowledged the 2017 MedPAC report cited by the developer indicating that inpatient hospital spending accounted for 22 percent of total Medicare spending in 2015 and represented the second largest Medicare spending category in 2015.[1] The Standing Committee reviewed data provided by the developer demonstrating that MPSB episodes have a range of cost performance at the taxpayer identification number (TIN) level and the TIN-NPI level. The Standing Committee agreed that there is an opportunity for improvement and ultimately passed the measure on the Importance to Measure and Report criterion.

Prior to the Standing Committee meeting, this measure was reviewed by the NQF SMP. The SMP had raised concerns regarding outlier exclusions and recommended that the developer report the distribution of outlier exclusion across facilities to ensure that they don't concentrate on a limited number of facilities. The developers provided written responses and results that show excluded outliers are not concentrated on a small number of providers. The SMP was satisfied with the developer responses and ultimately passed the measure with a moderate rating for both reliability and validity. The SMP review and the measure developer responses are summarized in the SMP Spring 2020 SMP Discussion Guide.

For reliability, the Standing Committee reviewed the signal-to-noise analysis and split sample reliability testing conducted by the developer. Some Standing Committee members raised concerns with some of the lower quartile reliability scores for the TIN-NPI reporting level (0.60), stating that they are low. The developer noted that this may be due to low participation in the CMS Merit-based Incentive Payment System (MIPS) for TIN-NPI, which declined from 2017 to 2018. However, given this concern by some of the Standing Committee members, the Standing Committee did not reach consensus on reliability. For validity, the Standing Committee reviewed the results for both face and empirical validity testing conducted by the developer. The Standing Committee did not have any concerns regarding the face validity but did raise several concerns regarding the empirical validity of the measure. Some Standing Committee members questioned the attribution to multiple clinicians and whether a care episode could be attributed to multiple clinician groups and multiple clinicians. The Standing Committee also questioned the validity of the time window of three and 30 days pre- and post-discharge for each episode, respectively, and that this might need to be more specific for certain medical conditions. The Standing Committee also questioned the strength of the correlations, noting that the correlation between predicted value and the six different clinical themes (e.g., PAC settings), and the correlation with the risk-adjusted value and the six different clinical themes, were low. Lastly, the Standing Committee questioned how well the model predicts downstream costs after a hospitalization and raised concerns regarding the lack of including social risk factors. The developer noted that testing demonstrated significance of the social risk factors, but inconsistent direction of the social risk factors and limited impact of social risk factor effects under the current risk adjustment model. The Standing Committee noted that risk adjustment should be focused on reducing bias and may not improving the model fit. The Standing Committee ultimately did not pass the measure on validity.

Following the measure evaluation meeting, the developers submitted a reconsideration request for this measure, which discussed the issues raised by the Standing Committee with respect to validity. The developer asserted that the NQF evaluation criteria were not correctly applied and that there was inconsistent application of the evaluation criteria that led to a measure not being recommended for endorsement.

During the post-comment meeting, the Standing Committee members discussed the issues related to the masking of disparities with absence of social risk factors in the model and the low correlation in empirical testing results. The developer further encouraged the Standing Committee consider the face validity testing that was performed on this measure. The developer stated that NQF guidance was utilized for face validity testing and mentioned that the NQF algorithm within NQF's Measure Evaluation Criteria depicts that if face validity is established for a measure, it is an acceptable method to establish validity for new submissions. NQF staff added that within the NQF Measure Evaluation Criteria, while face validity can be an acceptable form of validity for new measures, it is evaluated together with the

threats to validity, like that of the risk adjustment. The Standing Committee reevaluated the measure during the October 13, 2020 post-comment web meeting. Quorum was achieved, and the Standing Committee voted to not reconsider the measure.

Following the post-comment meeting, the developer submitted a reconsideration request to the CSAC for this measure. The developer cited three areas of concerns related to the review of the measures, namely, inconsistency in the Standing Committee deliberations and process, specifically voting; misapplication of measure evaluation criteria and guidance; and transparency of the Standing Committee deliberation materials. Related to voting inconsistencies, the developer asserted that the Standing Committee did not consistently apply the voting processes established by NQF guidelines. The developer expressed concerns that the Standing Committee departed from the CDP and that the Standing Committee's votes are inconsistent in their application of overarching issues on the set of cost measures in the spring 2020 cycle and lack a clear rationale for these decisions. Lastly, the developers asserted that NQF did not follow the CDP's process for ensuring transparency, which they state created challenges for understanding reasons for decisions, and therefore, they could not then respond to the Standing Committee appropriately.

The Cost and Efficiency co-chairs, Cheryl Damberg and Sunny Jhamnani, acknowledged the perspective of the measure developers related to the evaluation of the cost measures reviewed by the Standing Committee during the spring 2020 review cycle. The co-chairs recognized that these measures address a high resource use aspect of healthcare and that there is an opportunity for improvement in resource use outcomes, and because these measures are used for Medicare payment purposes, it is critical to make sure the measures are valid and can reliably differentiate performances.

The co-chairs reaffirmed that the Standing Committee was responsive to the developers' request for a careful review and reconsideration of measures. During the spring 2020 post-comment meetings, the co-chairs expressed that the Standing Committee remained focused on the criteria where consensus was not reached or where a recommendation was not given for endorsement and provided rationales for why the Standing Committee did not consider the measures to meet NQF criteria. CSAC members noted the complexity of this measure; however, they decided not to reconsider this measure and voted to uphold the Cost and Efficiency Standing Committee's recommendation.

NQF #3575 Total Per Capita Cost (TPCC) (Acumen, LLC): Endorsed

Description: The *Total Per Capita Cost* (TPCC) measure assesses the overall cost of care delivered to a beneficiary with a focus on the primary care they receive from their provider(s). The TPCC measure score is a clinician's average risk-adjusted and specialty-adjusted cost across all beneficiary months attributed to the clinician during a one year performance period. The measure is attributed to clinicians providing primary care management for the beneficiary, who are identified by their unique Taxpayer Identification Number and National Provider Identifier pair (TIN-NPI) and clinician groups, identified by their TIN number. Clinicians are attributed beneficiaries for one year, beginning from a combination of services indicate that a primary care relationship has begun. The resulting periods of attribution are then measured on a monthly level, assessing all Part A and Part B cost for the beneficiary for those months that occur during the performance period. The beneficiary populations eligible for the TPCC include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.

Measure Type: Cost/Resource Use; **Level of Analysis**: Clinician : Group/Practice, Clinician : Individual; **Setting of Care**: No Applicable Care Setting; **Data Source**: Assessment Data, Claims, Enrollment Data, Other

During the measure evaluation meeting on July 10, 2020, the Standing Committee did not vote on the recommendation for endorsement at the meeting because the Standing Committee did not reach consensus on validity—a must-pass criterion.

The Standing Committee acknowledged research cited by the developer that indicated how primary care management in certain settings, such as patient-centered medical homes, can reduce the total cost of care by reducing utilization of high-cost service. The Standing Committee reviewed data provided by the developer demonstrating that TPCC has a range of cost performance at the TIN level and the TIN-NPI level. The Standing Committee agreed that there is an opportunity for improvement and ultimately passed the measure on the importance to measure and report criterion.

Prior to the Standing Committee meeting, this measure was reviewed by the NQF SMP. The SMP had raised concerns regarding outlier exclusions and recommended that the developer report the distribution of outlier exclusion across facilities to ensure that they don't concentrate on a limited number of facilities. The developers provided written responses and results that show excluded outliers are not concentrated on a small number of providers. The SMP was satisfied with the developer responses and ultimately passed the measure with a moderate rating for both reliability and validity. The SMP review and the measure developer responses are summarized in the SMP Spring 2020 SMP Discussion Guide. For reliability, the Standing Committee reviewed the signal-to-noise analysis and split sample reliability testing conducted by the developer. The Standing Committee did not raise any major concerns and agreed with the SMP that the measure was reliable and passed the measure on this criterion.

For validity, the Standing Committee reviewed the results for both face and empirical validity testing conducted by the developer. The Standing Committee did not have any concerns regarding the face validity but did raise several concerns regarding the empirical validity of the measure. It questioned the strength of the correlations, noting that the correlation with risk- and specialty-adjusted cost were low to moderate. The Standing Committee also raised concerns regarding the lack of social factors in the risk adjustment model. The developer reported that inclusion of social factors in the model did not significantly change TIN or TIN-NPI performance scores on average. The Standing Committee noted that social risk adjustment should be focused on reducing bias and may not improving the model fit. The Standing Committee ultimately did not reach consensus on validity.

Moving to feasibility, the Standing Committee agreed that this measure would be feasible as all data were routinely collected and posed no additional data collection burden on providers. The Standing Committee also passed this measure on use and usability. They acknowledged that the measure is publicly reported as part of the CMS Quality Payment Program MIPS, and this measure will be implemented as part of MIPS beginning in the 2020 MIPS performance year and 2022 MIPS payment year.

During the October 1, 2020 post-comment web meeting, the Standing Committee did not achieve quorum. Therefore, a recording of the meeting and the offline voting surveys were distributed to

members of the Standing Committee. The Standing Committee revoted on the measure and recommended the measure for endorsement.

The CSAC expressed no concerns with the Standing Committee's evaluation or recommendation and voted unanimously to endorse the measure.

References

1 Medicare Payment Advisory Commission (MedPAC). Medicare Payment Advisory Commission's (MedPAC) March 2020 Report to the Congress: Medicare Payment Policy - Context for Medicare Payment Policy (Chp 1). March 2020. http://www.medpac.gov/docs/default-source/reports/mar20_medpac_ch1_sec.pdf?sfvrsn=0

2 MedPAC. "Medicare Payment Advisory Commission's (MedPAC) March 2020 Report to the Congress: Medicare Payment Policy - Improving Medicare Payment for Post-Acute Care." Report to the Congress: Medicare Payment Policy (Chp 7). March 2020. http://www.medpac.gov/docs/default-source/reports/mar20_medpac_ch7_sec.pdf?sfvrsn=0.

Appendix A: Details of Measure Evaluation

Rating Scale: H=High; **M**=Moderate; **L**=Low; **I**=Insufficient; **NA**=Not Applicable

Note: Vote totals may differ between measure criteria and between measures as Standing Committee members often have to join calls late or leave calls early. NQF ensures that quorum is maintained for all live voting. All voting outcomes are calculated using the number of Standing Committee members present for that vote as the denominator.

During the July 10, 2020 and October 1, 2020 web meetings, some Standing Committee members were unable to attend the entire meeting. There were early departures and late arrivals due to pre-existing conflicts, including those related to COVID-19. Quorum required for voting was only achieved for the entirety of #3561 deliberations on July 1, 2020, and for the second post-comment meeting on October 13, 2020 for the review of #3574. Quorum was not achieved for all other measure discussions. Therefore, the Standing Committee discussed all relevant criteria and voted after the meeting using an online voting tool.

Endorsed Measures

NQF #3561 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Inpatient Rehabilitation Facilities

Submission | Specifications

Description: The *Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facility (*MSPB-PAC IRF) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each IRF's efficiency relative to that of the national median IRF. Specifically, the measure assesses Medicare spending by the IRF and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each IRF divided by the episode-weighted median MSPB-PAC Amount across all IRFs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all IRFs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF admissions from October 1, 2015 through September 30, 2017.

Claims-based MSPB-PAC measures were developed in parallel for the IRF, long-term care hospital (LTCH), skilled nursing facility (SNF), and home health agency (HHA) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with IRF experience evaluated IRF claims and then gave direction on how to adjust for specific patient and case-mix characteristics.

The MSPB-PAC IRF measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the IRF Quality Reporting Program (QRP) and finalized in the FY 2017 IRF Prospective Payment System (PPS) Final Rule.[1] Public reporting for the measure began in Fall 2018 through the IRF Compare website (https://www.medicare.gov/inpatientrehabilitationfacilitycompare/) using FY 2016-2017 data.

Notes:

[1] Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 Federal Register, Vol. 81, No. 151. https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf

Numerator Statement: The numerator is the MSPB-PAC IRF Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider. This is then multiplied by the national average episode spending level for all IRF providers nationally.

NQF #3561 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Inpatient Rehabilitation Facilities

Denominator Statement: The denominator is the episode-weighted national median of the MSPB-PAC IRF Amounts for all IRFs nationally.

Exclusions: Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC IRF episodes because they are clinically unrelated to IRF care and/or because IRF providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given IRF provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration. Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include:

- Planned hospital admissions[1]
- Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants)
- Some routine screening and health care maintenance (e.g., colonoscopy and mammograms)
- Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis)

Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as:

1) Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories.

Rationale: This exclusion ensures that complete claims data are available for each provider.

2) Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated.

Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC IRF measure could potentially misrepresent a providers' resource use.

3) Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window.

Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC IRF measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims.

4) Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window.

Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data.

5) Any episode where the claim(s) constituting the attributed PAC provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill.

Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization.

6) Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory)

Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements.

Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers.

Notes:

[1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6

Adjustment/Stratification: Statistical risk model; Not applicable: the MSBP-PAC IRF measure is not stratified.

Level of Analysis: Facility

Setting of Care: Post-Acute Care

Type of Measure: Cost/Resource Use

Data Source: Assessment Data, Claims, Enrollment Data, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING July 10, 2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria.</u>

(1a. High Impact or High Resource Use, 1b. Opportunity for Improvement)

1a. High Impact or High Resource Use & 1b. Opportunity for Improvement: **H-3**; **M-14**; **L-1**; **I-0**

<u>Rationale</u>:

- The Standing Committee reviewed publicly reported measure score data provided by the developer for all U.S. providers under Medicare's IRF PPS with 20 or more eligible episodes in the reporting period of FY 2016-2017. The data from 1,147 IRFs with 20 or more episodes in the reporting period of FY 2016-2017, which include measure scores from 618,123 patient episodes, shows the mean of 1.00 with a standard deviation of 0.08. The Standing Committee also acknowledged the interquartile range of 0.10 (min: 0.74 and max: 1.47).
- The developer provided data that demonstrated that the differences in post-acute care payments are key drivers of variation in Medicare spending overall. They also provided citations demonstrating significant variability in IRF care and outcomes, links between facility characteristics and readmissions, and significant opportunities for improvement.
- Several Standing Committee members noted that the measure was developed as a part of the IMPACT Act and is a legislative requirement. The Standing Committee acknowledged that the developer demonstrated significant variability in resource use in inpatient rehabilitation facilities with opportunities for improvement.

- However, several Standing Committee members also challenged the assumption that lower resource use is necessarily better patient care and discussed the importance of pairing these measures with associated quality measures.
- Ultimately, the Standing Committee generally agreed that this measure addresses a high resource use aspect of healthcare and that the developer demonstrated variation in post-acute care spending to warrant measurement. The Standing Committee passed the measure on this criterion.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria.</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity

2a. Reliability: **Y-14; N-1**; 2b. Validity: **H-0**; **M-12**; **L-5**; **I-0**

<u>Rationale</u>:

Reliability

- The developers demonstrated reliability using signal-to-noise analysis through the Adams' method to examine the measure score's ability to capture between-facility differences versus random error. They reported mean reliability score of 0.86 and median of 0.89. The results demonstrated that on average, 86% of the variation in the risk adjusted MSPB amount was associated with systematic differences between facilities, with a range of 70 to 96% (on average) among the smallest and largest facility quartiles, respectively.
- The developers also assessed measure score reliability using split-sample method with intraclass correlation coefficients (ICCs) to examine agreement between two performance measure scores for a facility based on randomly split, independent subsets of IRF episodes. The reported mean score of 0.87 with 95% confidence interval of 0.85 to 0.88, with a range of 0.81 to 0.95 (on average) among the smallest and largest facility quartiles, respectively.
- This measure was reviewed by the NQF SMP. They voted to pass the measure on the reliability criterion (H-3; M-4; L-0; I-0).
- The Standing Committee raised concern with the low reliability of quartile 1 facilities, or those with 20-190 episodes. Quartile 1 had a mean reliability score of 0.70.
- The Standing Committee ultimately agreed that this measure is reliable and voted to pass the measure on this subcriterion.

Validity

- The developer sought to demonstrate validity through three separate empirical tests:

  1) Assessed how this measure correlates to resource /utilization such as hospitalization within the episode window and emergency room (ER) visits within the episode window;

  2) Correlated this measure with the Discharge to Community (DTC) rates for IRFs (measure endorsed by NQF #3479);

  3) Correlated this measure with *Percent of Residents or Patients with Pressure Ulcers That Are New or Worsened (Short-Stay)* (NQF #0678).
- The Standing Committee raised several concerns regarding the validity of the measure. First, the Standing Committee questioned the use of 30-days as the appropriate length of time that IRF can influence downstream care decisions. The Standing Committee highlighted the need to empirically evaluate and validate if 30-days after the discharge period is the appropriate length of time to capture complications that can be attributed to IRF care.
- The Standing Committee questioned the approach to truncation/winsorization of low- and high-cost episodes, commenting that doing winsorization at one spectrum and then excluding both low-end and high-end cases seems redundant. The developer commented that winsorization happens just for expected costs that are predicted from the risk adjustment model. The developer stated that very low values of expected costs will make providers' ratios for that episode look extremely high, so

winsorization is done at the low end. Exclusions are applied to outliers in the deviation from observed costs to expected costs, which can occur at both the high and low ends.

- The Standing Committee questioned the approach to why death is excluded within the episode window.
- The MSPB-PAC IRF risk adjustment model is adapted from the model used in the NQF-endorsed hospital MSPB measure (#2158), which is an adaptation of the standard CMS-HCC risk adjustment model. The MSPB-PAC IRF model uses a linear regression framework and a 90-day HCC lookback period. The risk adjustment model has a r-squared of 0.1595. The Standing Committee raised concern that the calculation of expected cost is not aligned with IRF payment programs.
- The Standing Committee raised concerns about the risk adjustment model (the adjusted r-squared value of 0.1595), specifically the lack of adjustment for social risk factors. The Standing Committee highlighted that accounting for social risk factors that are associated with the outcome of interest that are outside provider's control reduces bias in measurement.
- The Standing Committee also raised concern that expected costs are not aligned with how patient risk is accounted for in IRF payment programs.
- This measure was reviewed by the SMP and was voted to pass the measure on the validity criterion (H-1; M-6; L-1; I-0). However, during the measure evaluation meeting, the Standing Committee did not pass the measure on the validity criterion, noting additional concerns beyond the SMP review related to the lack of sociodemographic status (SDS) adjustment in the risk adjustment model, alignment of patient risk between expected costs and IRF payment programs, and additional threats to validity as stated above.
- During the post-comment period, the developer submitted a reconsideration request. The developer stated that NQF #3561, NQF #3563, and NQF #3564 purposefully exclude some payment variables, noting that this was an explicit policy decision by the Center for Medicare and Center for Clinical Standards and Quality as excessive spending in these settings has historically been driven by excessive use of therapy and variability in coding of patient status on assessment instruments.
- The developer mentioned that it did consider such variables and included a wide range of clinical factors, including IRF Rehabilitation Impairment Categories. However, the developer mentioned that inclusion of such variables would violate NQF guidance, which emphasizes that variables be resistant to gaming, and be present at the start of care, are not indicators of the care provided.
- With respect to social risk adjustment, the developer mentioned that adjusting for social risk factors may mask disparities in care, creating a lower standard of care for beneficiaries with higher social risk, and this could allow for a higher rate of readmissions, complications, etc., among those with high social risk.
- The developer mentioned that this may be appropriate if such outcomes are outside of a provider's control, but the developer's empirical testing showed that poorer performance for high social risk individuals is closely tied to providers themselves, rather than individual beneficiaries, with especially strong effects in particular settings.
- The developer argued that due to the negative relationship between social factors and the measure scores, inclusion of certain factors may penalize providers for taking on beneficiaries with high social risk. Finally, the developer stated that including these factors in the model had a minimal impact on provider scores.
- The Standing Committee did not have any further questions for the developer and passed the measure on validity.

3. Feasibility: H-6; **M-13**; **L-0**; **I-0**

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

NQF #3561 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Inpatient Rehabilitation Facilities

- The developer indicated that all data elements are coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims) and are defined in fields in a combination of electronic sources.
- The developer noted that the data used to calculate this measure are already collected as part of Medicare's payment process.
- During the post-comment call, the Standing Committee did not raise any concerns and passed the measure on feasibility.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-17; No Pass-2** *4b. Usability:* **H-0; M-16; L-3; I-0**

- The Standing Committee considered that this measure is publicly reported as part of the Center of Medicare & Medicaid Services' Inpatient Rehabilitation Facilities Quality Reporting Program (IRF QPS) on IRF Compare website.
- The Standing Committee also considered that the NQF Measures Application Partnership, which encouraged continued development, they did note concerns about the potential for unintended consequences (e.g., premature discharges).
- During the post-comment call, the Standing Committee did not raise any concerns and passed the measure on use and usability.

5. Related and Competing Measures

- The developer identified the following NQF endorsed measures as related measures:
    - NQF #2158 Medicare Spending Per Beneficiary (MSPB) – Hospital
- The developer stated that the MSPB-PAC measures are harmonized across PAC settings as well as with MSPB-Hospital.

6. Standing Committee Recommendation for Endorsement: **Y-16; N-3**

- The developer submitted a reconsideration request for this measure. The Standing Committee reevaluated the measure during the post-comment web meetings on October 1 and October 13, 2020 and revoted to recommend the measure for endorsement.

7. Public and Member Comment

- A commenter had doubts about the value of the measure and agreed it should not be endorsed. They stated that inpatient rehabilitation facilities' funding and utilization are controversial, but they have a modest volume and impact in comparison to Skilled Nursing Facilities and Long-Term Acute Care Hospitals and more controlled utilization.

8. Consensus Standards Approval Committee (CSAC) Vote: Y-8; N-3 (November 17-18, 2020: Approved for endorsement)

9. Appeals

- No appeals were received.

NQF #3562 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Long-Term Care Hospitals

Submission | Specifications

Description: *The Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals* (MSPB-PAC LTCH) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each LTCH's efficiency relative to that of the national median LTCH. Specifically, the measure assesses Medicare spending by the LTCH and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each LTCH divided by the episode-weighted median MSPB-PAC Amount across all LTCH facilities. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all LTCHs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., LTCH admissions from October 1, 2015 through September 30, 2017.

Claims-based MSPB-PAC measures were developed in parallel for the LTCH, inpatient rehabilitation facility (IRF), skilled nursing facility (SNF), and home health agency (HHA) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with LTCH expertise evaluated LTCH claims and then gave direction on how to adjust for specific patient and case-mix characteristics.

The MSPB-PAC LTCH measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the LTCH Quality Reporting Program (QRP) and finalized in the FY 2017 LTCH Prospective Payment System (PPS) Final Rule.[1] The measure entered into use on October 1, 2016. Public reporting for the measure began in Fall 2018 through the LTCH Compare website (https://www.medicare.gov/longtermcarehospitalcompare/) using FY 2016-2017 data.

Notes:

[1] Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2017 Rates. Federal Register, Vol. 81, No. 162. https://www.govinfo.gov/content/pkg/FR-2016-08-22/pdf/2016-18476.pdf

Numerator Statement: The numerator is the MSPB-PAC LTCH Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider, comparing Standard and Site Neutral episodes only with episodes of the same type. This is then multiplied by the national average episode spending level for all LTCH providers nationally.

Denominator Statement: The denominator is the episode-weighted national median of the MSPB-PAC LTCH Amounts for all LTCH facilities nationally.

Exclusions: Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC LTCH episodes, because they are clinically unrelated to LTCH care and/or because LTCH providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given LTCH provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration. Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include:

- Planned hospital admissions[1]

- Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants)
- Some routine screening and health care maintenance (e.g., colonoscopy and mammograms)
- Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis)

Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as:

1) Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories.

Rationale: This exclusion ensures that complete claims data are available for each provider.

2) Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated.

Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC LTCH measure could potentially misrepresent a providers' resource use.

3) Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window.

Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC LTCH measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims.

4) Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window.

Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data.

5) Any episode where the claim(s) constituting the attributed PAC provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill.

Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization.

6) Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory)

Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements.

Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers.

Notes:

[1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research &

NQF #3562 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Long-Term Care Hospitals

Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6

Adjustment/Stratification: Statistical risk model; The MSPB-PAC LTCH measure is stratified by standard and site neutral payment rate admissions. An MSPB-PAC LTCH Standard episode is triggered by a standard payment rate claim, while an MSPB-PAC LTCH Site Neutral episode is triggered by a site neutral payment rate claim. Risk adjustment is then performed separately for MSPB-PAC LTCH Standard and Site Neutral cases. Thus, LTCH Standard and Site Neutral episodes are compared only with LTCH Standard and Site Neutral episodes, respectively, to ensure that the measure is making fair comparisons between clinically similar beneficiaries.

Level of Analysis: Facility

Setting of Care: Post-Acute Care

Type of Measure: Cost/Resource Use

Data Source: Assessment Data, Claims, Enrollment Data, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING July 10, 2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria.</u>

(1a. High Impact or High Resource Use, 1b. Opportunity for Improvement)

1a. High Impact or High Resource Use & 1b. Opportunity for Improvement: H-2; M-14; L-0; I-0;

<u>Rationale</u>:

- The Standing Committee reviewed publicly reported measure score data provided by the developer for all U.S. providers under Medicare's LTCH PPS with 20 or more eligible episodes in the reporting period of 2016-2017. The developer reported data from 422 LTCHs with 20 or more episodes, which include measure scores from 153,864 patient episodes and showed the mean of 1.00 with a standard deviation of 0.08. The Standing Committee also acknowledged the interquartile range of 0.09 (min: 0.76 and max: 1.50).
- The Standing Committee also reviewed data that showed that the differences in post-acute care payments are key drivers of variation in Medicare spending overall, and a number of studies demonstrating relationships between facility characteristics and resource use, links between LTCHs' financial incentives and strategic discharge of patients from facilities, and significant opportunities for improvement.
- Several Standing Committee members noted that the measure was developed as a part of the IMPACT Act and is a legislative requirement. The Standing Committee acknowledged that the developer demonstrated significant variability in resource use in long-term care hospitals with opportunities for improvement. The Standing Committee did note that there is geographic variability in the availability of long-term acute care settings across the country.
- Similar to measure #3561, several Standing Committee members also challenged the assumption that lower resource use is necessarily better patient care and discussed the importance of pairing these measures with associated quality measures. They noted that this is an area for improvement, the four MSBP cost measures presented (#3561, #3562, #3563, and #3564) are for each of the different post-acute care settings. They recommended that a more effective approach may be to allow providers to identify the appropriate lower cost and appropriate post-acute care settings to send a patient rather than trying to control costs within a specific setting.
- Ultimately, the Standing Committee generally agreed that this measure addresses a high resource use aspect of healthcare. They also acknowledged that the developer demonstrated variation in post-acute care spending to warrant measurement. The Standing Committee passed the measure on this criterion.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria.</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity

2a. Reliability: **Y-15; N-0**; 2b. Validity: **H-0**; **M-10**; **L-6**; **I-0**

NQF #3562 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Long-Term Care Hospitals

Rationale:

Reliability

- The developers demonstrated reliability using signal-to-noise analysis through Adams' method to examine the measure score's ability to capture between-facility differences versus random error. They reported mean reliability score of 0.87 and median of 0.90. The results demonstrated that on average, 87% of the variation in the risk adjusted MSPB amount was associated with systematic differences between facilities, with a range of 75% to 94% (on average) among the smallest and largest facility quartiles, respectively.
- The developers also assessed measure score reliability using a split-sample method with intraclass ICCs to examine agreement between two performance measure scores for a facility based on randomly-split, independent subsets of LTCH episodes. They reported mean score of 0.86 with 95% confidence interval of 0.84-0.89, with a range of 0.86 to 0.90 (on average) among the smallest and largest facility quartiles, respectively.
- This measure was reviewed by the NQF SMP. The SMP voted to pass the measure on the reliability criterion (H-5; M-2; L-0; I-0).
- The Standing Committee ultimately agreed that this measure is reliable and voted to pass the measure on reliability.

Validity

- The developer sought to demonstrate validity through three separate empirical tests: 1) Assessed how this measure correlates to resource/utilization such as hospitalization within the episode window and ER visits within the episode window; 2) Correlated this measure with the *Discharge to Community (DTC) rates for LTCH* (measure endorsed by NQF #3480); 3) Correlated this measure with the *Percent of Residents or Patients with Pressure Ulcers That Are New or Worsened (Short-Stay*) *Measure* (NQF #0678); *Catheter-associated Urinary Tract Infection (CAUTI) Outcome Measure* (NQF #0138); *Central line-associated Bloodstream Infection (CLABSI) Outcome Measure* (NQF #0139); and *Facility-wide Inpatient Hospital-onset Clostridium difficile Infection (CDI) Outcome Measure* (NQF #1717).
- The Standing Committee questioned the use of 30 days as the appropriate length of time that LTCH can influence downstream care decisions. The Standing Committee highlighted the need to empirically evaluate and validate if 30-days after the discharge period is the appropriate length of time to capture complications that can be attributed to LTAC care. The Standing Committee also questioned the approach to how death is handled within the episode window. The Standing Committee also raised concerns regarding the approach to truncation or winsorization of low and high-cost episodes.
- The MSPB-PAC LTCH risk adjustment model is adapted from the model used in the NQF-endorsed hospital MSPB measure (NQF #2158), which is an adaptation of the standard CMS-HCC risk adjustment model. The MSPB-PAC LTCH model uses a linear regression framework and a 90-day HCC lookback period. The risk adjustment model has a r-squared of 0.4894. The Standing Committee raised concern that the calculation of expected cost is not aligned with LTCH payment programs. The developer noted on the call that there is more alignment in the expected episode cost and the payment program for LTCH and less for IRF and the other PAC measures.
- The Standing Committee raised additional concerns about the risk adjustment model (overall adjusted r-squared value of 0.4894), specifically the lack of adjustment for social risk factors. The Standing Committee acknowledged that the adjusted r-squared for this LTCH measure is higher than the other PAC measures. The Standing Committee highlighted that accounting for social risk factors that are associated with the outcome of interest that are outside provider's control and reduces bias in measurement.
- This measure was reviewed by the SMP, which voted to pass the measure on the validity criterion (H-2; M-3; L-2; I-0).
- The Standing Committee ultimately passed the measure on this subcriterion. However, validity concerns were raised by several Standing Committee members regarding alignment of how patient risk

is handled in this measure compared to the payment program and the predictive value of the risk adjustment model for this measure was higher than the other PAC measures reviewed.

3. Feasibility: H-7; M-11; L-0; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee expressed no concerns with the feasibility of this measure given that all data elements are in defined fields and administrative claims data can be accessed electronically.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-18; No Pass-0** *4b. Usability:* **H-1; M-11; L-5; I-1**

Rationale:

- The Standing Committee acknowledged that this is a new measure and is publicly reported as part of the Center of Medicare & Medicaid Services' LTCH QRP. The data are publicly reported on LTCH Compare website, which in addition to tracking quality of care, are intended to help consumers make informed decisions when selecting healthcare providers.
- This measure was reviewed by Measure Applications Partnership (MAP) Post-Acute Care Workgroup in the 2015-2016 cycle. The Workgroup's finalized recommendation of "encourage continued development" was released in February 2016. Although the MAP encouraged continued development, it did note a number of concerns about the potential for unintended consequences. In particular, the group raised concerns about issues of premature discharges and ability to make comparisons across providers. The group noted this could put a tremendous burden on family caregivers who may have to care for a patient they are not fully able to support. Members noted the need to consider risk adjustment for severity and socioeconomic status and urged CMS to incorporate functional status assessments into risk adjustment models to promote improvements. MAP requested consideration in the finalization of specifications to ensure costs are not double counted between care settings and recommended submission to NQF for endorsement. The MAP noted the measure was never fully specified before the post-acute care/long-term care workgroup deliberations and the current specifications were released in mid-January with public comment period closing Jan 27. It was noted that the measures double count costs between providers and is inconsistent with the IMPACT Act to develop comparable resource measures of PAC providers. While the MAP's final decision was to recommend continued development, there was a level of discomfort in this decision expressed by a number of members.
- The Standing Committee noted that there was no indication of areas of high spending or low spending by providers associated with IRF, so it was unclear if providers had enough information to target improvement.
- Ultimately, the Standing Committee passed this measure on use and usability criteria.

5. Related and Competing Measures

- This measure is related to the following measure:
  - NQF #2158 Medicare Spending Per Beneficiary (MSPB) – Hospital
- The developer stated that the MSPB-PAC IRF measure was harmonized with MSPB-Hospital.

6. Standing Committee Recommendation for Endorsement: **Y-10**; **N-6**

7. Public and Member Comment

NQF #3562 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Long-Term Care Hospitals

- A commenter stated that Long-Term Acute Care Hospitals' (LTACH) funding and utilization are controversial. Though they supported the Standing Committee's endorsement of the measure, they believed that high LTACH utilization does not necessarily correlate with higher quality or better outcomes and suspected that there was substantial regional variation.

8. Consensus Standards Approval Committee (CSAC) Vote: Y-8; N-3 (November 17-18, 2020: Approved for endorsement)

9. Appeals

- No appeals were received.

NQF #3575 Total Per Capita Cost (TPCC)

Submission | Specifications

Description: The *Total Per Capita Cost* (TPCC) measure assesses the overall cost of care delivered to a beneficiary with a focus on the primary care they receive from their provider(s). The TPCC measure score is a clinician's average risk-adjusted and specialty-adjusted cost across all beneficiary months attributed to the clinician during a one year performance period.

The measure is attributed to clinicians providing primary care management for the beneficiary, who are identified by their unique Taxpayer Identification Number and National Provider Identifier pair (TIN-NPI) and clinician groups, identified by their TIN number. Clinicians are attributed beneficiaries for one year, beginning from a combination of services indicate that a primary care relationship has begun. The resulting periods of attribution are then measured on a monthly level, assessing all Part A and Part B cost for the beneficiary for those months that occur during the performance period. The beneficiary populations eligible for the TPCC include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.

Numerator Statement: N/A

Denominator Statement: N/A

Exclusions: Included population:

The beneficiary population eligible for the TPCC measure consists of Medicare beneficiaries enrolled in Medicare Parts A and B for whom the measure identifies as having a primary care relationship with a clinician. To be included, the beneficiary must have at one of his or her beneficiary month occurring during the performance period.

Exclusions:

Several steps in the construction of the TPCC measure ensure comparability by fostering comparability in the beneficiary population captured and clinician population measured. These are detailed in Section S.7.2.

In keeping with the measure intent to capture the overall costs of care for beneficiaries receiving primary care services, there are a limited set of exclusions primarily to ensure that, as part of data processing, sufficient data are available to accurately determine resource use and calculate risk adjustment for each beneficiary. These exclusions, along with their rationales, are listed below:

•The beneficiary was not continuously enrolled in Medicare Parts A and B unless partial enrollment was the result of either new enrollment or death only. These beneficiaries may have gaps in their Medicare claim records when benefits are covered by other payers.

•The beneficiary resides outside the United States or its territories during the performance period. Differences in reimbursement policy for healthcare services provided outside the U.S. can lead to unfair comparisons of cost.

•The beneficiary receives benefits from the Railroad Retirement Board (RRB). Beneficiaries covered by the RRB may have healthcare benefits normally covered by Medicare paid by the RRB, which may bias the observed cost for these beneficiaries.

To ensure the clinicians attributed the measure are within the intended scope of primary care management, exclusions of clinicians are used to ensure comparability. Clinicians who would not reasonably be responsible for providing primary care are excluded from attribution of the revised TPCC measure using their CMS HCFA specialty designation assigned on Part B physician/supplier claims. This exclusion aims to keep primary care specialists and internal medicine subspecialists who frequently manage patients with chronic conditions falling in their areas of specialty. Additionally, clinicians are characterized by their Part B billing behavior and excluded from attribution if found meeting a threshold of billing for the following service categories; 10-day or 90-day global surgery services, anesthesia services, therapeutic radiation services, chemotherapy services. The methodology and clinical logic for exclusions of clinicians from attribution is further detailed in Section S.8.2

Data truncation is applied to risk-adjusted beneficiary monthly costs for outlier values through winsorization on the right tail. Monthly costs at the 99th percentile are assigned to all attributable beneficiary months with costs above the 99th percentile. Winsorization aims to limit the effects of extreme values on expected costs. Winsorization is a statistical transformation that limits extreme values in data to reduce the effect of possible outliers. The risk adjustment approach is detailed in Section S.7.2 and in S.9.3.

Adjustment/Stratification: Stratification by risk category/subgroup; Differences in patient case mix are accounted for by using separate risk adjustment models for the following types of beneficiaries, as discussed in Section S.7.2:

1) Beneficiaries without ESRD

1a) Beneficiaries with fewer than 12 months of Medicare medical history

2a) Beneficiaries with at least 12 months of Medicare medical history

3a) Beneficiaries in long-term institutional care settings

2) Beneficiaries with ESRD receiving dialysis

2a) Beneficiaries with fewer than 12 months of Medicare medical history

2b) Beneficiaries with at least 12 months of Medicare medical history

This stratification accounts for the very different patient clinical profiles for patients with ESRD receiving dialysis and patients without ESRD, as well as maximizes the availability of Medicare claims history to be able to construct indicator variables for clinical conditions.

The TPCC measure uses the CMS-HCC V22 risk adjustment models for new enrollee, community, and long-term institutional beneficiaries without ESRD. A beneficiary month is measured under the new enrollee model if they do not have a full one-year lookback of Medicare claims data as of the start of a beneficiary month. As a result, the model is derived primarily from beneficiary enrollment data. This model adjusts for gender, age, dual Medicare and Medicaid enrollment, and whether the beneficiary was originally entitled to Medicare due to disability through a series of interacted covariates. Beneficiaries with sufficient Medicare claims history are measured under the community or the institutional model if they are institutionalized in a long term care facility. In both models, severity of illness is measured using HCCs and disease interactions. 79 HCCs are accounted for under CMS-HCC V22 model for beneficiaries classified as community enrollees and long-term institutional enrollees while the exact number and types of disease interaction can vary. Both models interact beneficiary age with gender. In addition, the community model interacts dual enrollment status, gender, and the indicator for whether the beneficiary was originally entitled to Medicare due to disability, while the institutional model adjusts for disability as the original reason for Medicare enrollment and dual enrollment status independently.

For ESRD beneficiaries receiving dialysis, the TPCC measure utilizes the CMS-ESRD V21 risk adjustment models. Differentiated models are implemented for dialysis new enrollees and dialysis community enrollees. Similar to

the CMS-HCC V22, enrollees are classified as new enrollees if they were not continuously enrolled in Parts A and B for the one-year lookback period prior to each beneficiary month. As a result of this, the model primarily uses information from the beneficiary's enrollment data. This model adjusts for gender, age, dual enrollment status, and whether the beneficiary was originally entitled to Medicare due to disability through a series of interacted covariates. In addition to accounting for these patient characteristics, the dialysis community model also risk adjusts for medical severity using 87 HCCs and additional disease interactions.

The CMS-ESRD V21 and CMS-HCC V22 models both generate a risk score for each beneficiary that summarizes the beneficiary's expected cost of care relative to other beneficiaries. Risk scores for ESRD beneficiaries are normalized to enable comparison with the HCC V22 risk scores. This is achieved by multiplying ESRD risk scores by the mean annual Medicare spending for the ESRD population applied in the CMS-ESRD V21 model and dividing by the mean annual Medicare spending for the total Medicare population applied in the CMS-HCC V22 model, effectively renormalizing ESRD risk score values to the equivalent scale of the HCC models. A risk score equal to one indicates risk associated with expenditures for the average beneficiary nationwide. Risk scores below or above one indicate below and above average risk, respectively.

The complete list of risk adjustment variables for each model are listed in the Measure Codes List linked in Section S.1 in the tab titled HCC_Risk_Adjust.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: No Applicable Care Setting

Type of Measure: Cost/Resource Use

Data Source: Assessment Data, Claims, Enrollment Data, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING July 10, 2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria.</u>

(1a. High Impact or High Resource Use, 1b. Opportunity for Improvement)

1a. High Impact or High Resource Use & 1b. Opportunity for Improvement: **H-6**; **M-11**; **L-0**; **I-0**;

<u>Rationale</u>:

- The Standing Committee reviewed data provided by the developer demonstrating that TPCC has a range of cost performance at the TIN level and the TIN-NPI level. Specifically, the interquartile range (IQR) of performance for TIN level scores is $255 and mean performance of $1,109 for 74,191 clinician group practices. The IQR for TIN-NPI is $297 and mean performance of $1,169 for 335,480 practitioners.
- The developer cited research that shows how primary care management in certain settings, such as patient-centered medical homes, can reduce the total cost of care by reducing utilization of high-cost service.
- The Standing Committee agreed that there is an opportunity for improvement and ultimately passed the measure on the importance to measure and report criterion.

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria.</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity

2a. Reliability: **H-0**; **M-14**; **L-1**; **I-0**; 2b. Validity: **H-0**; **M-10**; **L-5**; **I-1**

<u>Rationale</u>:

Reliability

- The Standing Committee reviewed the signal-to-noise analysis and split sample reliability testing conducted by the developer.
- The performance measure score reliability testing was based on 74,191 clinician groups and 335,480 individual clinicians with 20 or more episodes in the measurement period of 2017-2018.

NQF #3575 Total Per Capita Cost (TPCC)

- The developer reported that the mean reliability score for all clinician group practices was 0.84 with range of 0.77 (25th percentile) to 0.95 (75th percentile). For the 335,480 individual practitioners, the mean reliability was slightly higher at 0.88 with range of 0.83 (25th percentile) to 0.95 (75th percentile). When examined by clinician group size, the average reliability score ranged from 0.81 (one clinician) to 0.94 (21+ clinicians).
- The developer reported that the interclass correlation coefficient for the overall sample was 0.76 with 95% confidence interval of 0.75-0.77. The ICC for 68,413 clinician groups as measured by Pearson correlation coefficient was 0.76 and for 265,106 individual practitioners was 0.64.
- This measure was reviewed by the SMP, which passed the measure on the reliability criterion (H-1; M-4; L-3; I-0).
- The Standing Committee generally agreed with the SMP that the measure was reliable and passed the measure on this criterion.

Validity

- The developer conducted both face and empirical validity testing for this measure. The developer reported that 80% (12 out of 15) of the experts agreed that the scores from the measure as specified after comprehensive reevaluation would provide an accurate reflection of cost effectiveness.
- Four clinical themes were created around inpatient service, PAC, emergency services not included in an admission, and outpatient evaluation and management (E&M) services, procedures, and therapy.
- The Standing Committee reviewed the empirical validity testing data showing a positive relationship between the measure and known indicators of resource or service utilization.
    - The mean of beneficiary's average risk- and specialty-adjusted monthly cost for a beneficiary during the measurement period is $1,187.
    - The mean of beneficiary's average risk- and specialty-adjusted monthly cost for beneficiaries with services relating to acute inpatient admissions is $2,647, compared with $866 for a beneficiary without acute inpatient admissions.
    - The mean of beneficiary's average risk- and specialty-adjusted monthly cost with services relating to PAC is $2,427 compared with $996 for a beneficiary without PAC.
- The Standing Committee reviewed the developer's findings.
    - The correlation with risk- and specialty-adjusted cost were low to moderate.
    - At both the TIN and TIN-NPI levels, there was a moderate correlation between the Skilled Nursing Facility service category and risk-adjusted cost (0.54); low correlation between Outpatient E&M Services, Procedures, and Therapy and risk-adjusted cost (0.45) and Acute Inpatient Services (0.38); and very low for the Home Health category (0.11) and Non-Hospital Admission Emergency Services (0.15).
- The Standing Committee noted that the developer reported 15.3% of episodes were excluded because of one or more exclusion criteria.
- The developer reported results showing that dual enrollment is associated with higher cost. The addition of the Agency for Healthcare Research and Quality Socioeconomic Status (AHRQ SES) index was significant and negative in value for the community, institutional, and new enrollee models, but not found to be significant in either the dialysis or new enrollee dialysis models. The developer reported that inclusion of SES in the model did not significantly change TIN or TIN-NPI performance scores on average.
- The Standing Committee raised concerns about the developer's exclusion of social risk factors in the overall risk adjustment model, given that they were statistically significant.
- Though the Scientific Methods Panel passed this measure on validity (H-0; M-5; L-3; I-0), the Standing Committee was unable to come to a consensus on this sub-criterion due to the concerns to validity raised above during the measure evaluation web meeting.

NQF #3575 Total Per Capita Cost (TPCC)

- The Standing Committee revoted on this criterion during the post-comment web meeting and concluded that the measure meets the scientific acceptability criterion.

3. Feasibility: H-6; M-11; L-0; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee acknowledged that all data elements are in defined fields in a combination of electronic sources coded by someone other than person obtaining original information. Since these data are routinely collected, this measure poses no additional data collection burden on providers.
- The Standing Committee agreed that this measure would be feasible and passed it on feasibility.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-16; No Pass-1** *4b. Usability:* **H-0; M-16; L-1; I-0**

Rationale:

- The Standing Committee acknowledged that the measure is publicly reported as part of the Centers for Medicare & Medicaid Services' Quality Payment Program MIPS.
- The developer stated that this measure will be implemented as part of MIPS beginning in the 2020 MIPS performance year and 2022 MIPS payment year.
- The developer noted that the overarching feedback that was received on measure performance and implementation from the measured entities and others included comments that (i) the revised specifications made several improvements to the measure; (ii) while the field test reports and other supplementary materials were helpful, the complexity of these documents was a challenge to some stakeholders, and; (iii) general feedback on the measure's attribution methodology, candidate events, and specialty adjustment.
- The MAP Coordinating Committee reviewed this for 2018-2019 measures under consideration. The MAP did not support this measure for rulemaking with the potential for mitigation. Mitigating factors include greater transparency around the attribution model and testing results.
- The MAP also noted a need to better understand how this measure handles the issue of small numbers and evaluate if there is a need to include social risk factors in the measure's risk adjustment model.
- Finally, the MAP noted the desire to avoid double counting clinician costs in the total cost measures and the episode-based cost measures and for CMS to consider consolidating the MSPB and TPCC measures to avoid overlap.
- The developer stated that there are no unexpected findings during the development and testing for the measure.
- Overall, the Standing Committee passed this measure on use and usability.

5. Related and Competing Measures

- This measure is related to the following measures:
    - NQF #1604 Total Cost of Care Population-based PMPM Index
- The measure developer indicates that this measure is not harmonized.
- The developer stated that #1604 is tested and endorsed for a population of patients less than 65 years of age, while TPCC was developed and tested on the Medicare population, affecting the appropriate intended use of each respective measure.
- NQF staff noted that related and competing measures would be discussed during the spring 2020 post-comment meeting on October 1, 2020.

NQF #3575 Total Per Capita Cost (TPCC)

6. Standing Committee Recommendation for Endorsement: **Y-12; N-6**

- The Standing Committee revoted on this measure during the post-comment web meeting and recommended it for endorsement.

7. Public and Member Comment

- Similar concerns to NQF #3574 were raised by commenters for this measure regarding measure specification, attribution at the individual clinician level, rare correlation with better outcomes, exclusion of patients who died in the overall model, the lack of correlation between cost and quality measures, and scientific acceptability. The commenters also mentioned that they were unsure the developer showed that the measure correlates to any one quality measure within the MIPS program and requested the Standing Committee discuss whether the results of the attribution and validity in the measure could lead to negative unintended consequences. They were also concerned with the lack of information on reliability results below the 25th percentile, particularly in light of the reference within the response of 2a2.3 that CMS generally considers 0.4 to be the threshold for moderate reliability and 100% of practices and clinicians with at least 20 episodes meet it.
- A commenter stated the attribution methodology is a significant threat to the validity of the measure. It was acknowledged that the TPCC eliminates the problem of attributing costs that occurred before the clinician ever saw the patient. However, the current approach could attribute the measure to practices and clinicians that billed E&M claims lower than desirable percentages. There were concerns that the attribution methodology assumes that a primary care relationship exists if two things happen within three days or three months and not otherwise. This would lead to significant problems when considering best practices in care. In addition, an oncologist will not know if they qualify for the TPCC measure, as the exemption is applied retrospectively based on a measurement of candidate events for which the oncologist bills for chemotherapy or radiation therapy services.
- A commenter also stated that within the attribution methodology, there is not an end to the clinician's primary care responsibility for the patient when a Medicare beneficiary switches to a new clinician. TPCC assigns responsibility for all Medicare Part A and B costs for 12 months after attribution. This would result in attribution to multiple clinicians, as patients switch providers. This would be inappropriate as only one clinician would be coordinating the patient's care and the other will not be aware of any services provided.
- There was a request that all medical and radiation oncologists be excluded from the TPCC measure. It was recommended radiation therapy be excluded from post-trigger inpatient and outpatient components.
- Commenters believed that the concerns outlined by the Standing Committee during the initial review along with deficiencies in the attribution methodology should result in the measure not achieving a recommendation for endorsement. It was, overall, urged that the Standing Committee should not endorse this measure.

8. Consensus Standards Approval Committee (CSAC) Vote: Y-11; N-0 (November 17-18, 2020: Approved for endorsement)

9. Appeals

- No appeals were received.

8. CSAC Review (November 17-18, 2020): Y-11; N-0

9. CSAC Decision: Approved for continued endorsement

10. Appeals: No appeals were received.

Measures Not Endorsed

NQF #3563 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Skilled Nursing Facilities

Submission | Specifications

Description: The *Medicare Spending Per Beneficiary – Post Acute Care Measure for Skilled Nursing Facilities* (MSPB-PAC SNF) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each SNF's efficiency relative to that of the national median SNF. Specifically, the measure assesses Medicare spending by the SNF and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each SNF divided by the episode-weighted median MSPB-PAC Amount across all SNFs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all SNFs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., SNF admissions from October 1, 2015 through September 30, 2017.

Claims-based MSPB-PAC measures were developed in parallel for the SNF, long-term care hospital (LTCH), inpatient rehabilitation facility (IRF), and home health agency (HHA) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with SNF experience evaluated SNF claims and then gave direction on how to adjust for specific patient and case-mix characteristics.

The MSPB-PAC SNF measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the SNF Quality Reporting Program (QRP) and finalized in the FY 2017 SNF Prospective Payment System (PPS) Final Rule.[1] Public reporting for the measure began in Fall 2018 through the Nursing Home Compare website (https://www.medicare.gov/nursinghomecompare/search.html?) using FY 2017 data.

Notes:

[1] Medicare Program; Prospective Payment System and Consolidated Billing for Skilled Nursing Facilities for FY 2017, SNF Value-Based Purchasing Program, SNF Quality Reporting Program, and SNF Payment Models Research; Final Rule. Federal Register, Vol. 81, No. 151. https://www.govinfo.gov/content/pkg/FR-2016-08-05/pdf/2016-18113.pdf

Numerator Statement: The numerator is the MSPB-PAC SNF Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider. This is then multiplied by the national average episode spending level for all SNF providers nationally.

Denominator Statement: The denominator is the episode-weighted national median of the MSPB-PAC SNF Amounts for all SNFs nationally.

Exclusions: Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC SNF episodes because they are clinically unrelated to SNF care and/or because SNF providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given SNF provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration. Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include:

- Planned hospital admissions[1]

- Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants)
- Some routine screening and health care maintenance (e.g., colonoscopy and mammograms)
- Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis)

Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as:

1) Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories.

Rationale: This exclusion ensures that complete claims data are available for each provider.

2) Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated.

Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC SNF measure could potentially misrepresent a provider's resource use.

3) Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window.

Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC SNF measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims.

4) Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window.

Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data.

5) Any episode where the claim(s) constituting the attributed PAC provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill.

Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization.

6) Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory)

Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements.

Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers.

Notes:

[1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research &

Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6

Adjustment/Stratification: Statistical risk model; Not applicable: the MSBP-PAC SNF measure is not stratified.

Level of Analysis: Facility

Setting of Care: Post-Acute Care

Type of Measure: Cost/Resource Use

Data Source: Assessment Data, Claims, Enrollment Data, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING July 10, 2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria.</u>

(1a. High Impact or High Resource Use, 1b. Opportunity for Improvement)

1a. High Impact or High Resource Use & 1b. Opportunity for Improvement: **H-4**; **M-13**; **L-0**; **I-0**;

<u>Rationale</u>:

- The Standing Committee discussed the development of this measure to address the resource use aspect of the IMPACT Act to allow for a better ability to measure resource use and efficiency of care to improve outcomes and align incentives and care coordination across PAC providers.
- The Standing Committee reviewed MSPB-PAC SNF measure scores for all U.S. providers paid under Medicare's SNF PPS with 20 or more eligible episodes in the reporting period of 2016-2017. There was a total of 14,903 SNFs with 20 or more episodes, resulting in 3,017,578 patient episodes.
- The Standing Committee acknowledged mean performance on this measure was 1.03, with a standard deviation of 0.24. The performance scores ranged from a minimum of 0.22 to a maximum of 2.09, with an interquartile range of 0.31. The Standing Committee agreed that this range of performance demonstrated an opportunity for improvement in reducing the variability in spending.
- Overall, the Standing Committee agreed with the importance of the measure focus and there is an opportunity to improve performance.

2. Scientific Acceptability of Measure Properties: <u>The measure does not meet the Scientific Acceptability criteria.</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity

2a. Reliability: **H-3**; **M-11**; **L-1**; **I-0**; 2b. Validity: **H-0**; **M-9**; **L-6**; **I-1**

<u>Rationale</u>:

Reliability

- The developer presented reliability scores for all 14,903 SNFs from FY 2016-2017 that averaged 0.92 with a median of 0.94 from a signal-to-noise analysis and a split sample ICC of 0.93 from split sample reliability testing.
- Facilities with a greater number of episodes were correlated with an increased reliability score. When examined by case volume quartiles, the average reliability score ranged from 0.84 (Quartile 1: 20-75 episodes) to 0.98 (Quartile 4: 258-3,104 episodes). The ICC for the overall sample was 0.93 with 95 percent confidence interval of 0.93-0.934.
- This measure was reviewed by the NQF SMP. The SMP voted to pass the measure on the reliability criterion.
- The Standing Committee ultimately agreed that this measure is reliable and voted to pass the measure on this criterion (H-5; M-3; L-0; I-0).

Validity

- The Standing Committee reviewed the empirical validity testing conducted by the measure developer. Specifically, the developer evaluated the MSPB-PAC SNF measure and examined correlation with other

known indicators of resource utilization, specifically hospital admissions and ER visits, DTC rates, and percent of residents or patients with pressure ulcers that are new or worsened (short-stay).

- The developer found a positive correlation between MSPB-PAC SNF measure and hospital admissions and ER visits. The developer found a moderate negative relationship between MSPB-PAC SNF measure and DTC measure scores and no association between the measure scores and pressure ulcers measures scores. All relationships are consistent with the developer's hypotheses.
- The Standing Committee raised several threats to validity including exclusions, alignment of patient risk with SNF payment programs, and the risk adjustment approach. The Standing Committee reviewed the exclusion of clinically unrelated services provided by the developer and noted downstream costs not associated with SNF care should be included in the list of clinically unrelated services.
- The MSPB-PAC SNF risk adjustment model is adapted from the model used in the NQF-endorsed hospital MSPB measure (#2158), which is an adaptation of the standard CMS-HCC risk adjustment model. The MSPB-PAC SNF model uses a linear regression framework and a 90-day HCC lookback period. The risk adjustment model has a r-squared of 0.11. The Standing Committee raised concerns that the calculation of expected cost is not aligned with SNF payment programs.
- The developer tested the impact of including social risk factors using stepwise regression models and testing the final models with and without social risk factors. Though each of the social factors tested remained statistically significant in the multivariate models individually and when added together with the other social factors in the model, the developer did not include them in the model, noting that the addition of the social factors did not improve the model fit and the adjusted r-squared values increased by less than 0.01.
- This measure was reviewed by the SMP, which voted to pass the measure on the validity criterion.
- However, the Standing Committee raised concerns about the exclusion of social factors in the risk adjustment model. They noted that inclusion of risk factors should minimize bias and may not always improve model fit.
- The Standing Committee shared concerns by the SMP on whether the measure should exclude outliers. The developer provided information in their submission materials regarding the exclusions of certain clinically unrelated services. The Standing Committee underscored the importance that by not excluding certain clinically unrelated services, costs associated with these services could lead to low-volume providers being penalized by random events unrelated to their care.
- Though the SMP passed this measure on validity, the Standing Committee did not reach consensus on validity during the measure evaluation meeting (H-2; M-4; L-1; I-1).
- The Standing Committee revoted on this criterion during the post-comment web meeting and voted to not pass the measure on validity.

3. Feasibility: H-5; M-11; L-1; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- The developer stated that all data elements are in defined fields in a combination of electronic sources, coded by someone other than person obtaining original information. Since these data are routinely collected, this measure poses no additional data collection burden on providers.
- The developer also stated that this measure uses data from the Minimum Data Set (MDS), which does not pose any additional burden on providers, as the submission of MDS is part of the federally mandated process for clinical assessment of all residents in Medicare and Medicaid certified nursing homes.
- The Standing Committee agreed that this measure would be feasible and passed it on feasibility.

4. Use and Usability

NQF #3563 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Skilled Nursing Facilities

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-16; No Pass-1** *4b. Usability:* **H-0; M-11; L-5; I-1**

Rationale:

- The developer indicated that this measure is publicly reported as part of the Centers for Medicare & Medicaid Services' SNF Quality Reporting Program.
- The developer addressed all comments during development and implementation by either revising the measure or by providing the rationale why revisions are not necessary or appropriate before finalizing the measure in the FY 2017 SNF PPS Final Rule.
- The NQF MAP Coordinating Committee provided a recommendation of "encourage continued development" in February 2016. MAP members found importance in balancing cost measures with quality and access, even though there were concerns about the ability to make comparisons across providers and premature discharges. Members noted the need to consider these factors:
    - Risk adjustment for severity and socioeconomic status and urging CMS to incorporate functional status assessments into risk adjustment models to promote improvements
    - The finalization of specifications to ensure costs are not double counted between care settings
    - Recommended submission to NQF for endorsement
- The developer stated that there were no unexpected findings during the development and testing for the measure.
- The Standing Committee expressed a lack of clarity on whether providers have enough information to target improvement, as there was no indication of areas of high or low spending by provider associated with SNF.
- Overall, the Standing Committee passed this measure on use and usability.

5. Related and Competing Measures

- No related or competing measures were noted.

6. Standing Committee Recommendation for Endorsement: **Y-11; N-7**

- The Standing Committee revoted on this measure during the post-comment web meeting and did not pass the measure on the validity criterion.

7. Public and Member Comment

- A commenter expressed nonsupport for the measure, as they stated post-acute SNF utilization is not necessarily meaningful in and of itself.

8. Consensus Standards Approval Committee (CSAC) Vote: Y-8; N-3 (November 17-18, 2020: Not approved for endorsement)

- The developer submitted a reconsideration request for this measure. The CSAC voted to not reconsider this measure (Y-5; N-6).
- The CSAC upheld the Standing Committee's decision not to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

NQF #3564 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Home Health Agencies

Submission | Specifications

Description: *The Medicare Spending Per Beneficiary – Post Acute Care Measure for Home Health Agencies* (MSPB-PAC HH) was developed to address the resource use domain of the Improving Medicare Post-Acute Care

NQF #3564 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Home Health Agencies

Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each home health (HH) agency's efficiency relative to that of the national median home health agency (HHA). Specifically, the measure assesses Medicare spending by the HHA and other healthcare providers during an MSPB-PAC HH episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each HHA divided by the episode-weighted median MSPB-PAC Amount across all HHAs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all HHAs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on CY 2016-2017 data; i.e., HHA admissions from January 1, 2016 through December 31, 2017.

Claims-based MSPB-PAC measures were developed in parallel for the HH, inpatient rehabilitation facility (IRF), long-term care hospital (LTCH), and skilled nursing facility (SNF) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with HH experience evaluated HH claims and then gave direction on how to adjust for specific patient and case-mix characteristics.

The MSPB-PAC HH measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the HHA Quality Reporting Program (QRP) and finalized in the CY 2017 Home Health Prospective Payment System Rate Update; Home Health Value-Based Purchasing Model; and Home Health Quality Reporting Requirements.[1] Public reporting for the measure began in Fall 2018 through the Home Health Compare website (https://www.medicare.gov/homehealthcompare/search.html) using CY 2017 data.

Notes:

[1] Medicare and Medicaid Programs; CY 2017 Home Health Prospective Payment System Rate Update; Home Health Value-Based Purchasing Model; and Home Health Quality Reporting Requirements. Federal Register, Vol. 81, No. 213. https://www.govinfo.gov/content/pkg/FR-2016-11-03/pdf/2016-26290.pdf

Numerator Statement: The numerator is the MSPB-PAC HH Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider. This is then multiplied by the national average episode spending level for all HH providers nationally.

Denominator Statement: The denominator is the episode-weighted national median of the MSPB-PAC HH Amounts for all HHAs nationally.

Exclusions: Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC HH episodes because they are clinically unrelated to HH care and/or because HH providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given HHA's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration. Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include:

- Planned hospital admissions[1]
- Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants)
- Some routine screening and health care maintenance (e.g., colonoscopy and mammograms)
- Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis)

Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as:

1) Any episode that results from a Request for Anticipated Payment (RAP)

Rationale: HHA requests for anticipated payment claims are interim claims that do not reflect the final payment made by Medicare for the services.

2) Any episode that is triggered by an HH claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories.

Rationale: This exclusion ensures that complete claims data are available for each provider.

3) Any episode where the claim(s) constituting the attributed HH provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated.

Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC HH measure could potentially misrepresent a providers' resource use.

4) Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window.

Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC HH measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims.

5) Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window.

Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data.

6) Any episode where the claim(s) constituting the attributed HH provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill.

Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization.

7) Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory)

Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements.

Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers.

Notes:

[1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6

NQF #3564 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Home Health Agencies

Adjustment/Stratification: Statistical risk model; The MSPB-PAC HH measure is stratified by Standard, LUPA, and PEP claims types. Risk adjustment is then performed separately for MSPB-PAC HH Standard, LUPA, and PEP cases. Thus, HH Standard, LUPA and PEP episodes are compared only with HH Standard, LUPA and PEP episodes, respectively, to ensure that the measure is making fair comparisons between clinically similar beneficiaries.

Level of Analysis: Facility

Setting of Care: Home Care

Type of Measure: Cost/Resource Use

Data Source: Assessment Data, Claims, Enrollment Data, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING July 10, 2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria.</u>

(1a. High Impact or High Resource Use, 1b. Opportunity for Improvement)

1a. High Impact or High Resource Use & 1b. Opportunity for Improvement: **H-1**; **M-14**; **L-2**; **I-0**;

<u>Rationale</u>:

- The Standing Committee acknowledged that this measure was developed to address the resource use aspect of the IMPACT Act to allow for a better ability to measure resource use and efficiency of care to improve outcomes and align incentives and care coordination across PAC providers.
- The Standing Committee agreed that the measure demonstrated high impact, noting variability in (HHA) care and patient outcomes.
- The Standing Committee reviewed the measure scores reported publicly for all U.S. providers paid under Medicare's Home Health Prospective Payment System with 20 or more eligible episodes in the reporting period 2016-2017. There was a total of 10,470 HHAs with 20 or more episodes in the reporting period. These scores represent 10,321,802 patient episodes, after all exclusions were applied. The scores variability included a mean of 0.96, standard deviation: 0.15, with a minimum of 0.31 and maximum of 2.44, and an interquartile range of 0.18.
- Several Standing Committee members questioned the extent that HHAs can prevent hospitalizations or j (ED) visits given their scope of practice. The developer emphasized that the measure is aligned with similar home health measures for quality improvement in NQF's portfolio with emphasis on care coordination.
- Overall, the Standing Committee agreed with the importance of the measure focus with an opportunity to improve performance.

2. Scientific Acceptability of Measure Properties: <u>The measure does not meet the Scientific Acceptability criteria.</u> (2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity

2a. Reliability: **H-1**; **M-13**; **L-2**; **I-0**; 2b. Validity: **H-0**; **M-9**; **L-7**; **I-0**

<u>Rationale</u>:

Reliability

- The developer reported a mean reliability score for 10,470 HHAs from 2016-2017 was 0.84 with median of 0.90. When examined by facility size, the average reliability score ranged from 0.63 (Q1) to 0.97 (Q4). The ICC for the overall sample was 0.76 with 95 percent confidence interval of 0.75-0.77. The ICC was lowest for Q1 (0.57) and highest in Q4 (0.94).
- This measure was reviewed by the SMP, which passed the measure on reliability (H-3; M-3; L-1; I-1).
- However, several Standing Committee members raised concerns that the reliability statistics for low volume providers were too low for acceptable reliability (0.57). Several Standing Committee members noted that it may be difficult to differentiate HHAs with smaller number of qualifying episodes.
- Due to these concerns, the Standing Committee was unable to come to a consensus on this subcriterion.

NQF #3564 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Home Health Agencies

- During the post comment meeting, the Standing Committee revoted to pass the measure on reliability.

Validity

- The Standing Committee reviewed the empirical validity testing data showing a positive relationship between MSPB and known indicators of resource or service utilization.
    - The mean observed-to-expected cost ratio for episodes without a hospital admission is 0.68, compared with 2.31 for episodes with at least one hospital admission during the episode period (p-value<0.0001).
    - The mean observed-to-expected cost ratio for episodes without an ER visit is 0.89, compared to 1.39 for episodes with at least one ER visits (p-value<0.0001). They also observed a positive relationship between the mean observed-to-expected cost ratio and the number of hospitalizations/ER visits as hypothesized.
- The Standing Committee reviewed the developer's findings, including the following:
    - a small but significant negative association between the measure scores and the DTC) measure scores as hypothesized and a very small but statistically significant correlation (Pearson -0.240; Spearman -0.250) between the measure scores and DTC measure scores
    - a small positive correlation between the measure scores and Acute Care Hospitalization (ACH) scores (Pearson 0.298; Spearman 0.305)
    - a small but significant positive correlation between the measure scores and the various functional improvement scores as hypothesized (Pearson correlations ranging from 0.075 to 0.163; Spearman ranged from 0.041 to 0.152)
- The Standing Committee noted that the developer reported 19.8% of episodes were excluded because of one or more exclusion criteria.
- This measure was reviewed by the SMP, which passed the measure on validity.
- Several Standing Committee members raised concerned that the developer reported a low overall risk adjustment r-squared of 0.092.
- The Standing Committee raised concerns regarding the developer's exclusion of social risk factors in the overall risk adjustment model, given that these factors were statistically significant. The developer noted that the dual eligibility in the social risk factor testing actually carries a negative coefficient, which would lower expected cost. The developer also noted that this would penalize providers for taking care of dual-eligible beneficiaries' certain episodes.
- The Standing Committee also raised concerns that approach to characterizing patient risk for the expected cost is not aligned with the approach to handling payment for HHAs.
- The Standing Committee was concerned that HHAs may not be able to control costs that resulted after their care and questioned the developer's decision to utilize a 60-day episode period. The developer clarified that as the measure emphasized upstream intervention and coordination of care, the costs associated with the amount of care needed during hospitalization or ED can be influenced by HH. The developer clarified that though home healthcare tended to be long term, the first 60 days of HHA care is a strong indicator of downstream outcomes.
- Though the SMP passed this measure on validity (H-3; M-3; L-1; I-1), the Standing Committee was unable to come to a consensus on this sub-criterion due to the threats to validity raised above during the measure evaluation meeting.
- The Standing Committee revoted on this criterion during the post-comment web meeting and voted to not pass the measure on validity.

3. Feasibility: H-6; M-11; L-0; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

NQF #3564 Medicare Spending Per Beneficiary – Post-Acute Care Measure for Home Health Agencies

Rationale:

- The Standing Committee acknowledged that all data elements are in defined fields in a combination of electronic sources and coded by someone other than person obtaining original information. Since these data are routinely collected, this measure poses no additional data collection burden on providers.
- The developer also stated that this measure uses data from the MDS, which does not pose any additional burden on providers, as the submission of MDS is part of the federally mandated process for clinical assessment of all residents in Medicare and Medicaid certified nursing homes.
- The Standing Committee agreed that this measure would be feasible and passed it on feasibility.

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Pass-16; No Pass-1** *4b. Usability:* **H-0; M-11; L-6; I-0**

Rationale:

- The Standing Committee acknowledged that the measure is publicly reported as part of the Centers for Medicare & Medicaid Services' Home Health Quality Reporting Program, and confidential feedback reports on the MSPB-PAC HH measure were provided to all active HH providers under the HH QRP starting in January 2018.They addressed comments received by either revising the measure or by providing the rationale why revisions are not necessary or appropriate, before finalizing the measure in the CY 2017 HH PPS final rule.
- The MAP Coordinating Committee considered these comments alongside the Workgroup recommendation and finalized the recommendation of "encourage continued development" in February 2016. There were concerns about the ability to make comparisons across providers and premature discharges. Members also noted the need to consider risk adjustment for severity and socioeconomic status and the finalization of specifications to ensure costs are not double counted between care settings. They urged CMS to incorporate functional status assessments into risk adjustment models to promote improvements. It was noted that the measures double count costs between providers and is inconsistent with the IMPACT Act in developing comparable resource measures of PAC providers.
- The developer stated that there are no unexpected findings during the development and testing for the measure.
- Overall, the Standing Committee passed this measure on use and usability.
- **5. Related and Competing Measures** No related or competing measures were noted.

6. Standing Recommendation for Endorsement: **Y-10; N-8**

- The Standing Committee revoted on this measure during the post-comment web meeting and did not reach consensus on the scientific acceptability criterion.

7. Public and Member Comment

- A commenter expressed non-support for the measure, as they stated post-acute care HHA utilization is not necessarily meaningful in and of itself.

8. Consensus Standards Approval Committee (CSAC) Vote: Y-9; N-2 (November 17-18, 2020: Not approved for endorsement)

- The developer submitted a reconsideration request for this measure. The CSAC voted to not reconsider this measure (Y-5; N-6).
- The CSAC upheld the Standing Committee's decision not to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

NQF #3574 Medicare Spending Per Beneficiary (MSPB) Clinician

Submission | **Specifications**

Description: The *MSPB Clinician* measure assesses the cost to Medicare for services by a clinician and other healthcare providers during an MSPB episode, which focuses on a patient's inpatient hospitalization. The MSPB episode spans from 3 days prior to the hospital stay ("index admission") through to 30 days following discharge from that hospital. The measure includes the costs of all services during the episode window, except for a limited list of services identified as being unlikely to be influenced by the clinician's care decisions and that are considered clinically unrelated to the management of care. The episode is attributed to the clinician(s) responsible for managing the beneficiary's care during the inpatient hospitalization. The MSPB Clinician measure score is a clinician's average risk-adjusted cost across all episodes attributed to the clinician. The beneficiary populations eligible for the MSPB Clinician measure include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.

Numerator Statement: N/A

Denominator Statement: N/A

Exclusions: Included population:

The beneficiary population eligible for the MSPB Clinician measure calculation consists of Medicare beneficiaries enrolled in Medicare Parts A and B who had an index admission to an inpatient hospital. To be included, the beneficiary must have an episode ending during the performance period.

Exclusions:

Several steps in the construction of the MSPB Clinician measure ensure comparability of the MSPB Clinician measure by fostering comparability in the service profiles and population captured by the measure, as discussed in Section S.7.2.

The measure excludes services that are clinically unrelated to clinician care management or the index hospitalization furthers the comparability of services captured by measure by limiting service variation to services that are likely to be influenced by clinician care management and related to the index admission. This is Step 3 of the measure construction methodology.

The measure excludes select episodes, detailed in Step 4 of the measure construction methodology, furthers the comparability of the Medicare beneficiary population studied by excluding episodes if any of the following conditions are met:

•Beneficiary has a primary payer other than Medicare during the episode window or in the 90-day lookback period

•Beneficiary was not enrolled in Medicare Parts A and B, or was enrolled in Part C, during the 90-day lookback period and episode window

•The beneficiary's death occurred during the episode.

•The index admission for the episode did not occur in either a subsection (d) hospital paid under the Inpatient Prospective Payment System (IPPS) or in an acute hospital in Maryland.

•The index admission for the episode is involved in an acute-to-acute hospital transfer (i.e., the admission ends in a hospital transfer or begins because of a hospital transfer).

•The index admission inpatient claim indicates a $0 actual payment or a $0 standardized payment.

The rationale and testing results for these exclusions are contained in the testing attachment, Section 2b2.

The MSPB Clinician measure applies risk adjustment, statistical exclusions, and renormalization to further ensure comparability, described in Step 5 of the construction methodology. The risk adjustment approach accounts for patient level variation prior to the index hospitalization and the severity of the index hospitalization. Statistical exclusions and renormalizations are engaged during measure construction after excluding outlier episodes to ensure that distributions resulting from outlier exclusions remain true to population averages.

As with the CMS-HCC model, the risk adjustment approach for this measure uses an ordinary least squares linear regression model. The predicted, or expected, cost is winsorized at 0.5th percentile to make sure episodes with unusually small predicted cost, which would lead to abnormally large O/E ratios, do not dominate

certain clinicians' final score. The winsorized expected costs are renormalized to ensure the average expected episode cost is the same before and after winsorizing. Then, extremely low- or high-cost outlier episodes with residuals below the 1st percentile or above the 99th percentile are excluded to reduce the effect of these episodes that deviate the most from their expected values in absolute terms. The expected cost after excluding these outliers is again renormalized to ensure that average expected costs are the same after outlier removal.

Adjustment/Stratification: Stratification by risk category/subgroup; The MSPB Clinician measure is stratified by MDC, which are mutually exclusive groups of MS-DRGs that correspond to an organ system (e.g., diseases and disorders of the digestive system) or cause of admission (e.g., burns). There are 25 MDCs (numbered 01-25), and a Pre-MDC group for extremely resource intensive MS-DRGs. Unlike MS-DRGs within the numbered MDCs which are determined largely by principal diagnosis, MS-DRGs within the Pre-MDC group are determined by Operating Room procedures (e.g., organ transplant). By running the risk adjustment model described in Section S.7.2 separately for episodes within each MDC determined by the MS-DRG of the index admission, the MSPB Clinician measure accounts for differences in resource use due to the nature of the reason for hospitalization. This helps ensure that the cost measure is fairly comparing clinicians for their patient case-mix, while preserving clinically meaningful distinctions in the beneficiary population within each MDC.

Level of Analysis: Clinician : Group/Practice, Clinician : Individual

Setting of Care: Inpatient/Hospital

Type of Measure: Cost/Resource Use

Data Source: Assessment Data, Claims, Enrollment Data, Other

Measure Steward: Centers for Medicare & Medicaid Services

STANDING COMMITTEE MEETING July 10, 2020

1. Importance to Measure and Report: <u>The measure meets the Importance criteria.</u>

(1a. High Impact or High Resource Use, 1b. Opportunity for Improvement)

1a. High Impact or High Resource Use & 1b. Opportunity for Improvement: **H-3**; **M-13**; **L-1**; **I-0**;

<u>Rationale</u>:

- The Standing Committee reviewed data provided by the developer demonstrating that MPSB episodes have a range of cost performance at the TIN level and the TIN-NPI level. Specifically, the interquartile range of performance for TIN level scores is $2,049 and mean performance of $19,194 for 19,213 group practices. The interquartile range of performance for TIN-NPI is $2,335, and mean performance of $19,741 for 126,628 practitioners.
- The Standing Committee acknowledged the 2017 MedPAC report cited by the developer indicating that inpatient hospital spending accounted for 22% of total Medicare spending in 2015 and represented the second largest Medicare spending category in 2015.
- The Standing Committee agreed that there is an opportunity for improvement and ultimately passed the measure on this criterion.

2. Scientific Acceptability of Measure Properties: <u>The measure does not meet the Scientific Acceptability criteria.</u>

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity

2a. Reliability: **H-0**; **M-9**; **L-6**; **I-0**; 2b. Validity: **H-0**; **M-5**; **L-10**; **I-0**

<u>Rationale</u>:

Reliability

- The Standing Committee reviewed the signal-to-noise analysis and split sample reliability testing conducted by the developer.
- Reliability scores were a mean of 0.78 and standard deviation of 0.13 for 19,213 TIN's, and a mean of 0.70 with a standard deviation of 0.11 for 126,628 TIN-NPI's.
- Split sample intraclass correlation coefficients were 0.66 for TIN and 0.60 for TIN-NPI.

NQF #3574 Medicare Spending Per Beneficiary (MSPB) Clinician

- Some Standing Committee members raised concerns with the reliability scores for the TIN-NPI reporting level (0.60), stating that they are low.
- The developer noted that this may be due to low participation in MIPS for TIN-NPI and declining from 2017 to 2018.
- This measure was reviewed by the Scientific Methods Panel (SMP) who passed the measure on the reliability criterion (H-1; M-4; L-3; I-0).
- However, the Standing Committee did not reach consensus on reliability.

Validity

- The Standing Committee reviewed the face validity and empirical validity testing conducted by the developer.
- Face validity comprised of administering a structured process for gathering detailed input from recognized clinician experts on inpatient care. The developer convened multiple expert panels to inform the face validity of the measure at different time points: a TEP, the MSPB service refinement group, and stakeholder feedback from national field testing.
- The Standing Committee noted that 14/15 (93%) TEP members convened by the developer agreed that the scores from the measure as specified after comprehensive reevaluation would provide an accurate reflection of cost effectiveness.
- For establishing empirical validity, the developer sought to confirm the expectation that the measure captures variation in service utilization by examining differences in risk-adjusted cost for known indicators of resource or service utilization (e.g., inpatient readmissions or post-acute care) through the ratio of O/E cost ratio.
- The mean O/E cost ratio for episodes with downstream acute readmission was 1.58, compared with 0.91 for episodes without downstream acute readmission. The mean O/E cost ratio for episodes with post-acute care (PAC) is 1.20, while for episodes without PAC is 0.80, as hypothesized.
- This measure was reviewed by the SMP, which passed the measure on validity (H-0; M-5; L-3; I-0).
- However, similar to the SMP concerns, some Standing Committee members raised concern about the attribution to multiple clinicians and whether a care episode could be attributed to multiple clinician groups and multiple clinicians.
- The developer noted that for medical diagnosis related groups (DRGs), the episode would be attributed to the TIN that meets the 30% threshold of E&M codes and also to any individual clinicians who are involved in that case and billing at least one E&M within the TIN.
- The Standing Committee questioned the validity of the time window of three and 30 days pre and post discharge, respectively, for each episode DRG and that this might need to be more specific for certain medical conditions.
- Some SMP members questioned the strength of the correlations, noting that the correlation between predicted value and six different clinical themes (e.g., PAC settings) was low (< 0.10) in all cases except PAC IRF/LTCH, and that the correlation with risk adjusted value and six different clinical themes was also not high—and was negative (-0.18) with PAC Home Health.
- The Standing Committee raised similar concerns regarding the correlation with SNF costs and how well the model is doing on predicting downstream costs after a hospitalization.
- The Standing Committee also raised concerns regarding the lack of including social factors within the risk adjustment model.
- The developer tested the impact of including social risk factors using T-tests and F-tests of variable coefficients and p-values, testing with stepwise regression models, and testing the final models with and without social risk factors.
- The developer noted that testing demonstrated significance of the social factors, but inconsistent direction of the social risk factors and limited impact of social risk factor effects under the current risk adjustment model.
- The Standing Committee noted that risk adjustment should be focused on reducing bias and may not always improve model fit.

NQF #3574 Medicare Spending Per Beneficiary (MSPB) Clinician

- Weighing all of the validity sub-criteria, the Standing Committee ultimately did not pass the measure on validity.

3. Feasibility: Vote not taken

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

4. Use and Usability

4a. Use; 4a1. Accountability and transparency; 4a2. Feedback on the measure by those being measured and others; 4b. Usability; 4b1. Improvement; 4b2. The benefits to patients outweigh evidence of unintended negative consequences to patients)

4a. Use: **Vote not taken** *4b. Usability:* **Vote not taken**

5. Related and Competing Measures No related or competing measures were noted.

6. Standing Committee Recommendation for Endorsement: **Y-12; N-5**

- The developer submitted a reconsideration request for this measure. The Standing Committee voted to not reconsider the measure during the post-comment web meeting.

7. Public and Member Comment

- A commenter expressed concerns with the measure specifications and reliability and attribution at the individual clinician level. They disagreed with the measure's attribution of costs to providers, like primary care physicians, for care they did not provide and who have limited control over many of those costs. They noted that primary care services represent a very small portion of overall costs. The commenter also had concerns about the impact of excluding patients who died on the overall model and the lack of correlation between cost and quality measures, particularly patient outcomes. Another commenter agreed with the Standing Committee's concerns on the scientific acceptability of the measure, expressing the need for the developer to demonstrate reliable and valid results to allow users to make meaningful distinctions in care costs. Commenters were also concerned with the lack of information on reliability results below the 25th percentile, particularly in light of the reference within the response of 2a2.3 that CMS generally considers 0.4 to be the threshold for moderate reliability and 100% of practices and clinicians with at least 20 episodes meet it.
- It was stated that the higher Medicare Spending Per Beneficiary rarely correlates with better outcomes, but this is very difficult to sort out at the clinician level. A member voiced concerns about the necessity of the TPCC and MSPB measures, as many of the beneficiaries captured in the episode-based measures will also be included in either or both the MSPB and TPCC measures. This would result in a beneficiary potentially being attributed to multiple providers within and across multiple measures, which could magnify the impact on cost measures of any individual beneficiary and complicate differences in cost and value.
- Commenters requested information and testing to demonstrate that the measure's use in Merit Incentive Payment System would yield reliable and valid results and enable end users to make meaningful distinctions on the costs associated with the care provided to patients. Commenters supported the Standing Committee's decision not to endorse this measure. They stated that outside of an accountable care organization (ACO) setting or other risk-sharing arrangement that covers all care provided to a population, the measure attributes costs to providers for care they did not provide and who have limited control over many of those costs. Concerns were shared that the measure did not provide insight into which treatments were most effective in providing high quality, low-cost care. Episode-based cost measures were brought up as a better approach to evaluating value. It was also recommended radiation therapy be excluded from post-trigger inpatient and outpatient components.

NQF #3574 Medicare Spending Per Beneficiary (MSPB) Clinician

8. Consensus Standards Approval Committee (CSAC) Vote: Y-2; N-2 (November 17-18, 2020: Not approved for continued endorsement)

- The developer submitted a reconsideration request for this measure. The CSAC voted to not reconsider this measure (Y-5; N-6).
- The CSAC upheld the Standing Committee's decision not to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

Appendix B: Cost and Efficiency Portfolio—Use in Federal Programs[1]

NQF #	Title	Federal Programs: Finalized or Implemented as of June 22, 2020
1598	Total Resource Use Population-Based PMPM Index	None
1604	Total Cost of Care Population-Based PMPM Index	None
2158	Medicare Spending Per Beneficiary (MSPB) – Hospital	Hospital Value-Based Purchasing (Implemented)
2431	Hospital-Level, Risk-Standardized Payment Associated With a 30-Day Episode of Care for Acute Myocardial Infarction (AMI)	Hospital Inpatient Quality Reporting (Implemented)
2436	Hospital-Level, Risk-Standardized Payment Associated With a 30-Day Episode of Care for Heart Failure	Hospital Inpatient Quality Reporting (Implemented)
2579	Hospital-Level, Risk-Standardized Payment Associated With a 30-Day Episode of Care for Pneumonia	Hospital Inpatient Quality Reporting (Implemented)
3474	Hospital-Level, Risk Standardized Payment Elective for THA/TKA	Hospital Inpatient Quality Reporting (Implemented)
3509	Routine Cataract Removal With Intraocular Lens (IOL) Implantation	None
3510	Screening/Surveillance Colonoscopy	None
3512	Knee Arthroplasty	None

[1] Per CMS Measures Inventory Tool as of 02/19/2021

Appendix C: Cost and Efficiency Standing Committee and NQF Staff

STANDING COMMITTEE

Sunny Jhamnani , MD (Co-Chair)
Yale University
New Haven, Connecticut

Cheryl Damberg, PhD (Co-Chair)
The RAND Corporation
Santa Monica, California

Kristine Martin Anderson, MBA
Booz Allen Hamilton
Rockville, Maryland

Robert Bailey, MD
Johnson & Johnson Health Care Systems, Inc.
Titusville, NJ

Bijan Borah, MSc, PhD
Mayo Clinic, College of Medicine
Rochester, MN

John Brooks, PhD
University of South Carolina
South Carolina

Cory Byrd
Humana, Inc.
Louisville, KY

Amy Chin, MS
Greater New York Hospital Association
New York City, NY

Lindsay Erickson, MPH
Integrated Healthcare Association
Oakland, CA

Troy Fiesinger, MD, FAAFP
Village Family Practice
Houston, TX

Emma Hoo
Pacific Business Group on Health
San Francisco, CA

Sean Hopkins, BS
New Jersey Hospital Association
Princeton, NJ

Rachael Howe, MS, BSN, RN
3M HIS
Murray, Utah

Donald Klitgaard, MD, FAAFP
MedLink Advantage
Avoca, IA

Lisa Latts, MD, MSPH, MBA, FACP
Colorado Department of Health Care Policy & Financing
Denver, Colorado

Jason Lott, MD, MHS, MSHP, FAAD
Bayer US LLC
Whippany, New Jersey

Alefiyah Mesiwala, MD, MPH
UPMC Health Plan
Pittsburgh, PA

Jack Needleman, PhD
UCLA Fielding School of Public Health
Los Angeles, California

Janis Orlowski, MD, MACP
Association of American Medical Colleges
Washington, DC

John Ratliff, MD, FACS, FAANS
Stanford University Medical Center
Stanford, California

Srinivas Sridhara, PhD, MHS
Optum Enterprise Analytics
Washington, DC

Mahil Senathirajah, MBA
IBM Watson Health
Santa Barbara, CA

Danny van Leeuwen, RN, MPH,
Health Hats
Arlington, VA

NQF STAFF
Sheri Winsper, MPH
Senior Vice President, Quality Measurement

Apryl Clark, MHSA
Acting Vice President, Quality Measurement

Michael Katherine Haynie
Senior Managing Director, Quality Measurement

Matthew Pickering, PharmD
Senior Director, Quality Measurement

Janaki Panchal, MSPH
Manager, Quality Measurement

Funmilayo Idaomi
Analyst, Quality Measurement

Yemsrach Kidane, PMP
Project Manager, Quality Measurement

Taroon Amin, PhD, MPH
Consultant

Appendix D: Measure Specifications

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
Steward	Centers for Medicare & Medicaid Services
Description	The Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facility (MSPB-PAC IRF) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each IRF's efficiency relative to that of the national median IRF. Specifically, the measure assesses Medicare spending by the IRF and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each IRF divided by the episode-weighted median MSPB-PAC Amount across all IRFs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all IRFs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF admissions from October 1, 2015 through September 30, 2017. Claims-based MSPB-PAC measures were developed in parallel for the IRF, long-term care hospital (LTCH), skilled nursing facility (SNF), and home health agency (HHA) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with IRF experience evaluated IRF claims and then gave direction on how to adjust for specific patient and case-mix characteristics. The MSPB-PAC IRF measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the IRF Quality Reporting Program (QRP) and finalized in the FY 2017 IRF Prospective Payment System (PPS) Final Rule.[1] Public reporting for the measure began in Fall 2018 through the IRF Compare website (https://www.medicare.gov/inpatientrehabilitationfacilitycompare/) using FY 2016-2017 data. Notes: [1] Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 Federal Register, Vol. 81, No. 151. https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf
Type	Cost/Resource Use
Data Source	Assessment Data, Claims, Enrollment Data, Other This measure is based on Medicare FFS administrative claims and uses data from the Medicare enrollment database and Minimum Data Set (MDS). The enrollment database provides information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A and Part B coverage, and periods in the Medicare FFS program. The MDS is used to construct a risk adjustment variable, indicating beneficiaries who have been institutionalized for at least 90 days in a given year. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, and revenue center codes. The Medicare FFS claims data files are used to identify Medicare services from IRFs and other settings (e.g., the outpatient setting) within the episode window. No data beyond the claims submitted in the normal course of business are required from providers for the calculation of this measure.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
	This measure submission is based on FY 2016-2017 data, which were the most recent data available at the time of our analyses. We used the data sources listed below to develop the analytic file for measure specification and testing: •Medicare Fee-For-Services claims and enrollment data: We accessed inpatient, outpatient, carrier, skilled nursing facility, home health, durable medical equipment, and hospice claims through the Centers for Medicare & Medicaid Services (CMS) Common Working File (CWF). The data dictionary for all Medicare FFS claims, demographic, and enrollment data are available at: https://www.resdac.org/cms-data?tid%5B%5D=4931&tid_1%5B%5D=1&=Find+Data+Files. General information about the CWF is available at: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/clm104c27.pdf. •Minimum Data Set (MDS): Acumen obtains the MDS through the Quality Improvement and Evaluation System (QIES). The data dictionary for the MDS data is available at: https://www.resdac.org/cms-data/files/mds-3.0/data-documentation. We used two mappings to group diagnosis and procedure codes for use in identifying clinical events, implementing exclusions and applying risk adjustment: •Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings for Services and Procedures: Software is available for download at: https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp •CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes: We used the Version 22 CMS-HCC mapping, which is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. We used five additional data sources for measure testing purposes only and not for measure specification: •2017 American Community Survey (ACS) 5-year estimate: We used the ACS to obtain the ZIP Code Tabulation Area (ZCTA) level measures needed to compute the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score for use in social risk factor testing. This information is downloadable at the US Census website: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t. •Rural-Urban Continuum Codes 2013: We used this data source to construct rural-urban identifiers for social risk factor testing. These codes include county FIPS indicators, which are then merged onto our episode file. More information on this data source can be found at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. •Provider of Services Current Files (POS File): We used this data source to describe the characteristics of IRFs included in measure specification and testing, such as census region, ownership type, and rurality, as reported in Table 1. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html. •IRF Compare data: We used this data source to examine the relationship between MSPB and assessment-based quality measures. The IRF Compare data include publicly reported IRF quality measures. The data are available at https://data.medicare.gov/data/inpatient-rehabilitation-facility-compare

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
	•Common Medicare Environment (CME) database: We extracted patient-level dual eligibility information from the CME database for social risk factor testing. CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME database receives information from the EDB and also contains additional information not available in the EDB. A description of the CME is available at: https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. This measure is based on Medicare FFS administrative claims and uses data from the Medicare enrollment database and Minimum Data Set (MDS). The enrollment database provides information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A and Part B coverage, and periods in the Medicare FFS program. The MDS is used to construct a risk adjustment variable, indicating beneficiaries who have been institutionalized for at least 90 days in a given year. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, and revenue center codes. The Medicare FFS claims data files are used to identify Medicare services from IRFs and other settings (e.g., the outpatient setting) within the episode window. No data beyond the claims submitted in the normal course of business are required from providers for the calculation of this measure. This measure submission is based on FY 2016-2017 data, which were the most recent data available at the time of our analyses. We used the data sources listed below to develop the analytic file for measure specification and testing: •Medicare Fee-For-Services claims and enrollment data: We accessed inpatient, outpatient, carrier, skilled nursing facility, home health, durable medical equipment, and hospice claims through the Centers for Medicare & Medicaid Services (CMS) Common Working File (CWF). The data dictionary for all Medicare FFS claims, demographic, and enrollment data are available at: https://www.resdac.org/cms-data?tid%5B%5D=4931&tid_1%5B%5D=1&=Find+Data+Files. General information about the CWF is available at: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/clm104c27.pdf. •Minimum Data Set (MDS): Acumen obtains the MDS through the Quality Improvement and Evaluation System (QIES). The data dictionary for the MDS data is available at: https://www.resdac.org/cms-data/files/mds-3.0/data-documentation. We used two mappings to group diagnosis and procedure codes for use in identifying clinical events, implementing exclusions and applying risk adjustment: •Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings for Services and Procedures: Software is available for download at: https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp •CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes: We used the Version 22 CMS-HCC mapping, which is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. We used five additional data sources for measure testing purposes only and not for measure specification:

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
	•2017 American Community Survey (ACS) 5-year estimate: We used the ACS to obtain the ZIP Code Tabulation Area (ZCTA) level measures needed to compute the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score for use in social risk factor testing. This information is downloadable at the US Census website: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t. •Rural-Urban Continuum Codes 2013: We used this data source to construct rural-urban identifiers for social risk factor testing. These codes include county FIPS indicators, which are then merged onto our episode file. More information on this data source can be found at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. •Provider of Services Current Files (POS File): We used this data source to describe the characteristics of IRFs included in measure specification and testing, such as census region, ownership type, and rurality, as reported in Table 1. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html. •IRF Compare data: We used this data source to examine the relationship between MSPB and assessment-based quality measures. The IRF Compare data include publicly reported IRF quality measures. The data are available at https://data.medicare.gov/data/inpatient-rehabilitation-facility-compare •Common Medicare Environment (CME) database: We extracted patient-level dual eligibility information from the CME database for social risk factor testing. CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME database receives information from the EDB and also contains additional information not available in the EDB. A description of the CME is available at: https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. Data dictionary URL; Code table attachment Data dictionary URL; Code table attachment
Level	Facility
Setting	Post-Acute Care
Numerator Statement	The numerator is the MSPB-PAC IRF Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider. This is then multiplied by the national average episode spending level for all IRF providers nationally.
Numerator Details	N/A
Denominator Statement	The denominator is the episode-weighted national median of the MSPB-PAC IRF Amounts for all IRFs nationally.
Denominator Details	N/A
Exclusions	Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC IRF episodes because they are clinically unrelated to IRF care and/or because IRF providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given IRF provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was

| --- | --- |
| | developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration. |
| Exclusion Details | Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include:
•Planned hospital admissions[1]
•Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants)
•Some routine screening and health care maintenance (e.g., colonoscopy and mammograms)
•Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis)
Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as:
1)Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories.
Rationale: This exclusion ensures that complete claims data are available for each provider.
2)Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated.
Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC IRF measure could potentially misrepresent a providers' resource use.
3)Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window.
Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC IRF measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims.
4)Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window.
Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data. |

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
	5)Any episode where the claim(s) constituting the attributed PAC provider's treatment includes at least one related condition code indicating that it is not a prospective payment system bill. Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization. 6)Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory) Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements. Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers. Notes: [1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6
Risk Adjustment	Statistical risk model The detailed steps to computing the measure score are described in section S.7.2. Risk-adjustment is applied in "Step 3: Calculate Predicted Episode Payments." The purpose of risk adjustment is to compensate for patient health circumstances and demographic factors that affect resource use but are beyond the influence of the attributed provider. The MSPB-PAC IRF measure risk adjustment model is adapted from the model used in the NQF-endorsed MSPB-Hospital measure, which itself is an adaptation of the standard CMS-HCC risk-adjustment model.[1,2] The MSPB-PAC IRF model uses a linear regression framework and a 90-day HCC lookback period. The risk adjustment model is estimated on all MSPB-PAC IRF episodes that meet the exclusion criteria. Each provider's MSPB-PAC IRF measure score is calculated as a provider's average MSPB-PAC Amount divided by the median MSPB-PAC Amount across all providers. A provider's MSPB-PAC IRF Amount is defined as the sum of standardized, risk-adjusted spending across all of a provider's eligible episodes divided by the number of episodes for that provider. Below is a description of the risk adjustment variables. Risk-Adjustment Variables The following beneficiary health status indicators are included as covariates in each MSPB-PAC IRF risk adjustment model and to the greatest extent possible are consistent across PAC settings (see Appendix C of the Measure Specifications document provided in section S.1 for a comprehensive list of independent variables used in the risk adjustment model): •70 HCCs •11 HCC interactions •11 brackets for age at the start of the episode •Original entitlement to Medicare through disability

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
	•ESRD status •Long-term care institutionalization at start of episode.[3] •Six clinical case-mix categories reflecting recent prior care (described further below).[4] •Hospice utilization during the episode •Prior acute ICU utilization day categories •Prior acute length of stay categories •Rehabilitation Impairment Categories (RICs) The clinical case-mix category variables used in the MSPB-PAC IRF risk adjustment model are included to account for differences in intensity and type of care received by beneficiaries prior to the start of an MSPB-PAC IRF episode. See section S.7.5 for more details on the methodology of assigning clinical case-mix categories to each episode. Notes: [1] QualityNet, "CMS Price (Payment) Standardization – Detailed Methods" (Revised April 2019) https://www.qualitynet.org/inpatient/measures/payment-standardization [2] CMS, "Medicare Risk Adjustment Information" (2016) https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html [3] Identifies beneficiaries who have been institutionalized for at least 90 days in a given year. The indicator is based on 90-day assessments from the Minimum Data Set (MDS) and is calculated based on CMS' definition of institutionalized individuals. [4] There are 7 case-mix categories as described above, but one category is removed to prevent collinearity. Statistical risk model
Stratification	Not applicable: the MSBP-PAC IRF measure is not stratified.
Type Score	Ratio An MSPB-PAC IRF measure score of 1 indicates that an IRF had an average MSPB-PAC Amount (i.e., risk-adjusted spending level) which is equal to the national episode-weighted median MSPB-PAC Amount across all IRFs during a given performance period. An MSPB-PAC IRF measure score of greater than 1 indicates that an IRF had higher average risk-adjusted spending levels compared to those of the national median IRF. For example, a measure score of 1.1 indicates that the IRF had average risk-adjusted spending levels that are 10 percent higher than the median IRF. On the other hand, an MSPB-PAC IRF measure score of less than 1 indicates that an IRF had lower average risk-adjusted spending levels compared to those of the median IRF. For example, a measure score of 0.9 indicates that the IRF had average risk-adjusted spending levels that are 10 percent lower than the median IRF.
Algorithm	In order to create a resource use measure that is clinically valid, there were multiple steps involved in excluding the least clinically relevant codes. Using an episode window, we organized claims into clinically meaningful service categories or settings. For example, Medicare Severity-Diagnosis Related Groups (MS-DRGs) noted after an IRF discharge were evaluated as medical or surgical admissions post-discharge. Clinical Classifications Software (CCS) and Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) services were organized into outpatient services, emergency department (ER) services, and durable medical equipment claims and evaluated for their relevance or relatedness to IRF care. Extensive clinical review was performed by clinicians with experience providing care in IRF settings, as well as in collaboration with Medical Officers at CMS. The hospitalizations and

| --- | --- |
| | outpatient services least clinically related to the IRF care were excluded from resource use calculation. For instance, it was not felt that an IRF could influence a beneficiary's rehospitalization for nervous system neoplasms (DRG 054), post-discharge outpatient services for kidney transplant (CCS 105), or routine fecal occult blood testing (CPT 82270). Therefore, these types of services were excluded. Services were only added to the exclusions list if there was consensus across IRF and CMS clinicians. Please see section S.9.1 for overall clinical consensus regarding the types of exclusions.

To account for the association between clinical severity and resource use, we risk adjust the total observed episode spending (described in section S.12) using CMS-HCC indicators and interactions between selected comorbidities. Diagnosis codes on claims that occur during the 90-day period prior to the start of an MSPB-PAC IRF episode (90-day "look back") are used to create HCC indicators. The MSPB-PAC IRF measure accounts for comorbid conditions and interactions by broadly following the CMS-HCC risk adjustment methodology, which is derived from Medicare Part A and B claims and is used in the Medicare Advantage (MA) program. For example, the measure accounts for interactions between disability and selected HCC groups (e.g., Cystic Fibrosis, Severe Hematological Disorders, Opportunistic Infections, among others). Given the fact that beneficiaries often have more than one comorbidity, the model also includes commonly observed paired condition interactions, (e.g., chronic obstructive pulmonary disease [COPD] and congestive heart failure [CHF]) and commonly observed triple-interactions (e.g., diabetes mellitus, congestive heart failure, and renal failure). The full list of variables used in the risk adjustment model can be found in the Measure Specifications document provided in section S.1.

In addition to comorbidities, the MSPB-PAC IRF measure utilizes clinical case-mix categories to create clinically meaningful subgroups that influence the type of services a beneficiary will receive in an IRF. To create these subgroups, information was derived from the institutional claim of the most recent hospitalization. The clinical case-mix category variables used in the MSPB-PAC IRF risk-adjustment model are included to account for differences in intensity and type of care received by beneficiaries prior to the start of an MSPB-PAC IRF episode. Taking the most recent institutional claim (by end date) in the 60 days prior to the start of an MSPB-PAC IRF episode, the episode is assigned to one of the following mutually exclusive and exhaustive clinical case-mix categories:
1) Prior Acute Surgical IP – Orthopedic – beneficiaries who have most recently undergone orthopedic surgery in an acute inpatient hospital
2) Prior Acute Surgical IP – Non-Orthopedic – beneficiaries who have most recently undergone a non-orthopedic surgery in an acute inpatient hospital
3) Prior Acute Medical IP with ICU – beneficiaries who have most recently stayed in an acute inpatient hospital for non-surgical reasons and had a stay in the ICU
4) Prior Acute Medical IP without ICU – beneficiaries who have most recently stayed in an acute inpatient hospital for non-surgical reasons but did not have a stay in the ICU
5) Prior PAC - Institutional – beneficiaries who are continuing PAC from an institutional PAC setting (i.e., coming from an LTCH, IRF, or SNF)
6) Prior PAC - HHA – beneficiaries who are continuing PAC from a HHA
7) Community – all other beneficiaries
Finally, the MSPB-PAC IRF measure uses RICs from the IRF admission. A full list of the RICs used in the risk adjustment model is included in Appendix C of the Measure Specifications document provided in section S.1. |

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities
	To simplify the clinical logic and avoid the issue of attributing claims to MSPB-PAC IRF episodes in the case of concurrent clinical events, all claims that begin within the episode window (treatment period and associated services period) are included in the MSPB-PAC IRF measure. An MSPB-PAC IRF episode is assigned to the rehabilitation facility of the index admission. A new episode may begin during the associated services period of a previous MSPB-PAC IRF episode in the 30 days post-discharge from the IRF.

	3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals
Steward	Centers for Medicare & Medicaid Services
Description	The Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals (MSPB-PAC LTCH) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each LTCH's efficiency relative to that of the national median LTCH. Specifically, the measure assesses Medicare spending by the LTCH and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each LTCH divided by the episode-weighted median MSPB-PAC Amount across all LTCH facilities. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all LTCHs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., LTCH admissions from October 1, 2015 through September 30, 2017. Claims-based MSPB-PAC measures were developed in parallel for the LTCH, inpatient rehabilitation facility (IRF), skilled nursing facility (SNF), and home health agency (HHA) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with LTCH expertise evaluated LTCH claims and then gave direction on how to adjust for specific patient and case-mix characteristics. The MSPB-PAC LTCH measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the LTCH Quality Reporting Program (QRP) and finalized in the FY 2017 LTCH Prospective Payment System (PPS) Final Rule.[1] The measure entered into use on October 1, 2016. Public reporting for the measure began in Fall 2018 through the LTCH Compare website (https://www.medicare.gov/longtermcarehospitalcompare/) using FY 2016-2017 data. Notes: [1] Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2017 Rates. Federal Register, Vol. 81, No. 162. https://www.govinfo.gov/content/pkg/FR-2016-08-22/pdf/2016-18476.pdf
Type	Cost/Resource Use

Data Source	Assessment Data, Claims, Enrollment Data, Other This measure is based on Medicare FFS administrative claims and uses data from the Medicare enrollment database and Minimum Data Set (MDS). The enrollment database provides information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A and Part B coverage, and periods in the Medicare FFS program. The MDS is used to construct a risk adjustment variable, indicating beneficiaries who have been institutionalized for at least 90 days in a given year. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, and revenue center codes. The Medicare FFS claims data files are used to identify Medicare services from LTCH and other settings (e.g., the outpatient setting) within the episode window. No data beyond the claims submitted in the normal course of business are required from providers for the calculation of this measure. This measure submission is based on FY 2016-2017 data, which were the most recent data available at the time of our analyses. We used the data sources listed below to develop the analytic file for measure specification and testing: • Medicare Fee-For-Services claims and enrollment data: We access inpatient, outpatient, carrier, skilled nursing facility, home health, durable medical equipment, and hospice claims through the Centers for Medicare & Medicaid Services (CMS) Common Working File (CWF). The data dictionary for all Medicare FFS claims, demographic, and enrollment data are available at: https://www.resdac.org/cms-data?tid%5B%5D=4931&tid_1%5B%5D=1&=Find+Data+Files. General information about the CWF is available at: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/clm104c27.pdf. • Minimum Data Set (MDS): Acumen obtains the MDS through the Quality Improvement and Evaluation System (QIES). The data dictionary for the MDS data is available at: https://www.resdac.org/cms-data/files/mds-3.0/data-documentation. We used two mappings to group diagnosis and procedure codes for use in identifying clinical events, implementing exclusions and applying risk adjustment: • Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings for Services and Procedures: Software is available for download at: https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp • CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes: We used the Version 22 CMS-HCC mapping, which is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. We used five additional data sources for measure testing purposes only and not for measure specification: • 2017 American Community Survey (ACS) 5-year estimate: We used the ACS to obtain the ZIP Code Tabulation Area (ZCTA) level measures needed to compute the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score for use in social risk factor testing. This • http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t. • Rural-Urban Continuum Codes 2013: We used this data source to construct rural-urban identifiers for social risk factor testing. These codes include county FIPS indicators, which are then merged onto our episode file. More information on this data source can be found at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. • Provider of Services Current Files (POS File): We used this data source to describe the characteristics of LTCH facilities included in measure specification and testing, such as census region, ownership type, and rurality, as reported in Table 1. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the

	3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals
	name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html. • LTCH Compare data: We used this data source to examine the relationship between MSPB and assessment-based quality measures. The LTCH Compare data include publicly reported LTCH quality measures. The data are available at https://data.medicare.gov/data/long-term-care-hospital-compare. • Common Medicare Environment (CME) database: We extracted patient-level dual eligibility information from the CME database for social risk factor testing. CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME database receives information from the EDB and also contains additional information not available in the EDB. A description of the CME is available at: https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf.
Level Setting	Facility Post-Acute Care
Numerator Statement	The numerator is the MSPB-PAC LTCH Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider, comparing Standard and Site Neutral episodes only with episodes of the same type. This is then multiplied by the national average episode spending level for all LTCH providers nationally.
Numerator Details	N/A
Denominator Statement Denominator Details	The denominator is the episode-weighted national median of the MSPB-PAC LTCH Amounts for all LTCH facilities nationally. N/A
Exclusions	Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC LTCH episodes because they are clinically unrelated to LTCH care and/or because LTCH providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given LTCH provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration.

Exclusion details	Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include: • Planned hospital admissions[1] • Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants) • Some routine screening and health care maintenance (e.g., colonoscopy and mammograms) • Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis) Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as: 1) Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories. Rationale: This exclusion ensures that complete claims data are available for each provider. 2) Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated. Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC LTCH measure could potentially misrepresent a providers' resource use. 3) Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window. Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC LTCH measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims. 4) Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window. Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data. 5) Any episode where the claim(s) constituting the attributed PAC provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill. Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization. 6) Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory)

	3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals
	Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements. Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers. Notes: [1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6

	3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals
Risk Adjustment	Statistical risk model To account for the association between clinical severity and resource use, we risk adjust the total observed episode spending (described in section S.12) using CMS-HCC indicators and interactions between selected comorbidities. The MSPB-PAC LTCH measure accounts for comorbid conditions and interactions by broadly following the CMS-HCC risk adjustment methodology, which is derived from Medicare Part A and B claims and is used in the Medicare Advantage (MA) program. Diagnosis codes on claims that occur during the 90-day period prior to the start of an MSPB-PAC LTCH episode (90-day "look back") are used to create HCC indicators. For example, the measure accounts for interactions disability status and selected HCC groups (e.g., Cystic Fibrosis, Severe Hematological Disorders, Opportunistic Infections, among others). Given the fact that beneficiaries often have more than one comorbidity, the model also includes commonly observed paired condition interactions, (e.g., chronic obstructive pulmonary disease [COPD] and congestive heart failure [CHF]) and commonly observed triple-interactions (e.g., diabetes mellitus, congestive heart failure, and renal failure). The full list of variables used in the risk adjustment model can be found in the Measure Specifications document provided in section S.1. In addition to comorbidities, the MSPB-PAC LTCH measure utilizes clinical case-mix categories to create clinically meaningful subgroups that influence the type of services a beneficiary will receive in an LTCH. To create these subgroups, information was derived from the institutional claim of the most recent hospitalization. The clinical case-mix category variables used in the MSPB-PAC LTCH risk-adjustment model are included to account for differences in intensity and type of care received by beneficiaries prior to the start of an MSPB-PAC LTCH episode. Taking the most recent institutional claim (by end date) in the 60 days prior to the start of an MSPB-PAC LTCH episode, the episode is assigned to one of the following mutually exclusive and exhaustive clinical case-mix categories: 1) Prior Acute Surgical IP – Orthopedic – beneficiaries who have most recently undergone orthopedic surgery in an acute inpatient hospital 2) Prior Acute Surgical IP – Non-Orthopedic – beneficiaries who have most recently undergone a non-orthopedic surgery in an acute inpatient hospital 3) Prior Acute Medical IP with ICU – beneficiaries who have most recently stayed in an acute inpatient hospital for non-surgical reasons and had a stay in the ICU 4) Prior Acute Medical IP without ICU – beneficiaries who have most recently stayed in an acute inpatient hospital for non-surgical reasons but did not have a stay in the ICU 5) Prior PAC - Institutional – beneficiaries who are continuing PAC from an institutional PAC setting (i.e., coming from an LTCH, IRF, or SNF)[1] 6) Prior PAC - HHA – beneficiaries who are continuing PAC from a HHA[1] 7) Community – all other beneficiaries[1] Finally, the MSPB-PAC LTCH Measure includes variables for MS-LTC-DRGs from the LTCH admission.

	3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals
Stratification	The MSPB-PAC LTCH measure is stratified by standard and site neutral payment rate admissions. An MSPB-PAC LTCH Standard episode is triggered by a standard payment rate claim, while an MSPB-PAC LTCH Site Neutral episode is triggered by a site neutral payment rate claim. Risk adjustment is then performed separately for MSPB-PAC LTCH Standard and Site Neutral cases. Thus, LTCH Standard and Site Neutral episodes are compared only with LTCH Standard and Site Neutral episodes, respectively, to ensure that the measure is making fair comparisons between clinically similar beneficiaries.
Type Score	Ratio An MSPB-PAC LTCH measure score of 1 indicates that an LTCH had an average MSPB-PAC Amount (i.e., risk-adjusted spending level) which is equal to the national episode-weighted median MSPB-PAC Amount across all LTCH facilities during a given performance period. An MSPB-PAC LTCH measure score of greater than 1 indicates that an LTCH had higher average risk-adjusted spending levels compared to those of the national median LTCH. For example, a measure score of 1.1 indicates that the LTCH had average risk-adjusted spending levels that are 10 percent higher than the median LTCH. On the other hand, an MSPB-PAC LTCH measure score of less than 1 indicates that an LTCH had lower average risk-adjusted spending levels compared to those of the median LTCH. For example, a measure score of 0.9 indicates that the LTCH had average risk-adjusted spending levels that are 10 percent lower than the median LTCH.

	3562 Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals
Algorithm	Grouping methodology: The grouping methodology includes all Medicare Part A and B services delivered to a beneficiary during the treatment period (from admission to the LTCH through to discharge from the LTCH) and associated services period (from admission to the LTCH through to 30 days after discharge from the LTCH). To simplify the clinical logic and avoid the issue of attributing claims to MSPB episodes in the case of concurrent clinical events, all claims that begin within the episode window (treatment period and associated services period) are included in the MSPB-PAC LTCH measure. In order to create a resource use measure that is clinically valid, there were multiple steps involved in excluding the least clinically relevant codes. Using an episode window, we organized claims into clinically meaningful service categories or settings. For example, Medicare Severity-Diagnosis Related Groups (MS-DRGs) noted after an LTCH discharge were evaluated as medical or surgical admissions post-discharge. Clinical Classifications Software (CCS) and Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) services were organized into outpatient services, emergency department (ER) services, and durable medical equipment claims and evaluated for their relevance or relatedness to LTCH care. Extensive clinical review was performed by clinicians with experience and expertise in LTCH, as well as in collaboration with Medical Officers at CMS. The inpatient, outpatient, Part B physician and supplier, and DMEPOS services least clinically related to the LTCH care were excluded from the measure. For instance, services related to the routine management of preexisting chronic conditions (e.g., dialysis for ESRD, treatment for preexisting cancers, and treatment for organ transplants) were felt to be clinically unrelated to the scope of the type of care that LTCHs provide. Therefore, these types of services were excluded. Services were excluded if there was consensus across clinicians from the measure developer, external clinical experts including TEP members, and CMS medical officers. Please see section S.9.1 for overall clinical consensus regarding the types of exclusions. Attribution algorithm: An MSPB-PAC LTCH episode is assigned to the facility of the index admission. A new episode may begin during the associated services period of a previous MSPB-PAC LTCH episode in the 30 days post-discharge from the LTCH.

	3575 Total Per Capita Cost (TPCC)
Steward	Centers for Medicare & Medicaid Services
Description	The Total Per Capita Cost (TPCC) measure assesses the overall cost of care delivered to a beneficiary with a focus on the primary care they receive from their provider(s). The TPCC measure score is a clinician's average risk-adjusted and specialty-adjusted cost across all beneficiary months attributed to the clinician during a one year performance period. The measure is attributed to clinicians providing primary care management for the beneficiary, who are identified by their unique Taxpayer Identification Number and National Provider Identifier pair (TIN-NPI) and clinician groups, identified by their TIN number.

	Clinicians are attributed beneficiaries for one year, beginning from a combination of services indicate that a primary care relationship has begun. The resulting periods of attribution are then measured on a monthly level, assessing all Part A and Part B cost for the beneficiary for those months that occur during the performance period. The beneficiary populations eligible for the TPCC include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.
Type	Cost/Resource Use
Data Source	Assessment Data, Claims, Enrollment Data, Other Medicare Part A and Part B claims data: TPCC uses Part A and B claims data to attribute beneficiaries to clinicians, calculate beneficiary's costs, and construct risk adjustors. CMS Office of Information Systems (OIS) maintains a detailed Medicare Claims Processing Manual available at the following URL: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Internet-Only-Manuals-IOMs-Items/CMS018912. Medicare Enrollment Database (EDB): This is used to determine beneficiary-level exclusions and supplemental risk adjustors, specifically Medicare Parts A, B, and C enrollment; other primary payers; disability status; sex; end-stage renal disease (ESRD); beneficiary birth dates; and beneficiary death dates. Common Medicare Environment (CME) database: This is used to determine beneficiary's dual status. https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. Minimum Data Set (MDS): The MDS is used to identify beneficiaries that should be risk adjusted through the CMS-HCC v22 institutional model. https://www.resdac.org/cms-data/files/mds-3.0. For measure testing purposes, data from the American Census, American Community Survey (ACS) is used in the analyses evaluating patient cohorts and social risk factors in risk adjustment. https://www.census.gov/programs-surveys/acs/technical-documentation/summary-file-documentation.html.
Level	Clinician : Group/Practice, Clinician : Individual
Setting	No Applicable Care Setting
Numerator Statement	N/A
Numerator Details	N/A
Denominator Statement	N/A
Denominator Details	N/A
Exclusions	Several steps in the construction of the TPCC measure ensure comparability by fostering comparability in the beneficiary population captured and clinician population measured. These are detailed in Section S.7.2. In keeping with the measure intent to capture the overall costs of care for beneficiaries receiving primary care services, there are a limited set of exclusions primarily to ensure that, as part of data processing, sufficient data are available to accurately determine resource use and calculate risk adjustment for each beneficiary. These exclusions, along with their rationales, are listed below.
Exclusion details	•The beneficiary was not continuously enrolled in Medicare Parts A and B unless partial enrollment was the result of either new enrollment or death only. These beneficiaries may have gaps in their Medicare claim records when benefits are covered by other payers. •The beneficiary resides outside the United States or its territories during the performance period. Differences in reimbursement policy for healthcare services provided outside the U.S. can lead to unfair comparisons of cost.

	•The beneficiary receives benefits from the RRRB. Beneficiaries covered by the RRB may have healthcare benefits normally covered by Medicare paid by the RRB, which may bias the observed cost for these beneficiaries. To ensure the clinicians attributed the measure are within the intended scope of primary care management, exclusions of clinicians are used to ensure comparability. Clinicians who would not reasonably be responsible for providing primary care are excluded from attribution of the revised TPCC measure using their CMS HCFA specialty designation assigned on Part B physician/supplier claims. This exclusion aims to keep primary care specialists and internal medicine subspecialists who frequently manage patients with chronic conditions falling in their areas of specialty. Additionally, clinicians are characterized by their Part B billing behavior and excluded from attribution if found meeting a threshold of billing for the following service categories; 10-day or 90-day global surgery services, anesthesia services, therapeutic radiation services, chemotherapy services. The methodology and clinical logic for exclusions of clinicians from attribution is further detailed in Section S.8.2 Data truncation is applied to risk-adjusted beneficiary monthly costs for outlier values through winsorization on the right tail. Monthly costs at the 99th percentile are assigned to all attributable beneficiary months with costs above the 99th percentile. Winsorization aims to limit the effects of extreme values on expected costs. Winsorization is a statistical transformation that limits extreme values in data to reduce the effect of possible outliers.
Risk Adjustment	Stratification by risk category/subgroup Beneficiary cost may differ across clinicians for reasons unrelated to the attributed clinicians' treatment and outside of their control. Risk adjustment accounts for case-mix of patients and other non-clinical characteristics that influence complexity of case-mix and is defined by a patient's claims found one year prior the start of a respective beneficiary month. The CMS Hierarchical Condition Category Version 22 (CMS-HCC V22) 2016 Risk Adjustment models are used for beneficiaries without End Stage Renal Disease (ESRD). Specifically, •The new enrollee model is used for beneficiaries that have fewer than 12 months of Medicare medical history. The model accounts for each beneficiary's age, sex, disability status, original reason for Medicare entitlement (age or disability), and Medicaid eligibility. •The community model is used for beneficiaries that have least 12 months of Medicare medical history. The model includes the same demographic information as the new enrollee model but also accounts for clinical conditions as measured by HCCs. •The institutional model is used for beneficiaries who were in long-term institutional settings. The model includes demographic variables, clinical conditions as measured by HCCs, and various interaction terms. The CMS-ESRD Version 21 (CMS-ESRD V21) 2016 Risk Adjustment models are used for ESRD beneficiaries receiving dialysis. Specifically, •The dialysis new enrollee model is used for ESRD beneficiaries that have fewer than 12 months of Medicare medical history. The model accounts for each beneficiary's age, sex, disability status, original reason for Medicare entitlement (age or disability), Medicaid eligibility, and ESRD. •The dialysis community model is used for ESRD beneficiaries that have at least 12 months of Medicare medical history. The model includes the same demographic information as the new enrollee model but also accounts for clinical conditions as measured by HCCs. The "HCC_Risk_Adjust" tab of the Measure Codes List file lists all variables included in the CMS-ESRD V21 and the CMS-HCC V22 risk adjustment models. The downloadable file is linked in Section S.1. The standardized risk scores from the CMS-ESRD V21 and CMS-HCC V22 models are generated for each beneficiary's month that summarizes the beneficiary's expected cost of

	care relative to other beneficiaries. Risk scores for ESRD beneficiaries are normalized to be on a comparable scale with the HCC V22 risk scores. A risk score equal to 1 indicates risk associated with expenditures for the average beneficiary nationwide. A risk score greater than 1 indicates above average risk, while a risk score less than 1 indicates below average risk. The risk-adjusted monthly cost for each attributed month is calculated according to the following steps: •Calculate CMS risk score for each beneficiary month using diagnostic data from the year prior to the month. This risk score is normalized by dividing by the average risk score for all beneficiary months. •Divide observed costs for each beneficiary month by the normalized risk score to obtain risk-adjusted monthly costs. •Winsorize risk-adjusted monthly costs at the 99th percentile by assigning the 99th percentile of monthly costs to all attributable beneficiary months with costs above the 99th percentile. •Normalize monthly costs to account for differences in expected costs based on the number of clinician groups to which a beneficiary is attributed in a given month. The normalization factor is the inverse cube root of the number of attributed clinician groups for that beneficiary month. The specialty adjustment for the TPCC measure is a cost adjustment applied to account for the fact that costs vary across specialties and across TINs with varying specialty compositions. The specialty adjustment at the TIN and TIN-NPI levels is calculated as follows: 1) Calculate the average risk-adjusted monthly cost for each TIN and TIN-NPI by averaging risk-adjusted monthly cost across all attributed beneficiary months. 2) Calculate the national specialty-specific expected cost for each specialty as the weighted average of TIN/TIN-NPI's risk-adjusted monthly cost. 2a) Define the weight for each TIN/TIN-NPI as the percentage of clinicians with that specialty multiplied by the total number of beneficiary months attributed to the TIN/TIN-NPI multiplied by the number of clinicians with that specialty. 2b) There will only be one specialty designation for a TIN-NPI. Therefore, the percentage of clinicians with a specialty and number of clinicians with a specialty will always be equal to 1. 3) Calculate the specialty-adjustment factor for each TIN or TIN-NPI as follows: 3a) Multiply the national specialty-specific expected cost for each specialty by the respective specialty's share of Part B payment within a TIN or TIN-NPI. 3b) Sum the weighted share of national specialty-specific expected cost calculated in the previous step across all the specialties under a given TIN or TIN-NPI.
Stratification	Differences in patient case mix are accounted for by using separate risk adjustment models for the following types of beneficiaries, as discussed in Section S.7.2: 1) Beneficiaries without ESRD 1a) Beneficiaries with fewer than 12 months of Medicare medical history 2a) Beneficiaries with at least 12 months of Medicare medical history 3a) Beneficiaries in long-term institutional care settings 2) Beneficiaries with ESRD receiving dialysis 2a) Beneficiaries with fewer than 12 months of Medicare medical history 2b) Beneficiaries with at least 12 months of Medicare medical history This stratification accounts for the very different patient clinical profiles for patients with ESRD receiving dialysis and patients without ESRD, as well as maximizes the availability of Medicare claims history to be able to construct indicator variables for clinical conditions.

	The TPCC measure uses the CMS-HCC V22 risk adjustment models for new enrollee, community, and long-term institutional beneficiaries without ESRD. A beneficiary month is measured under the new enrollee model if they do not have a full one-year lookback of Medicare claims data as of the start of a beneficiary month. As a result, the model is derived primarily from beneficiary enrollment data. This model adjusts for gender, age, dual Medicare and Medicaid enrollment, and whether the beneficiary was originally entitled to Medicare due to disability through a series of interacted covariates. Beneficiaries with sufficient Medicare claims history are measured under the community or the institutional model if they are institutionalized in a long term care facility. In both models, severity of illness is measured using HCCs and disease interactions. 79 HCCs are accounted for under CMS-HCC V22 model for beneficiaries classified as community enrollees and long-term institutional enrollees while the exact number and types of disease interaction can vary. Both models interact beneficiary age with gender. In addition, the community model interacts dual enrollment status, gender, and the indicator for whether the beneficiary was originally entitled to Medicare due to disability, while the institutional model adjusts for disability as the original reason for Medicare enrollment and dual enrollment status independently. For ESRD beneficiaries receiving dialysis, the TPCC measure utilizes the CMS-ESRD V21 risk adjustment models. Differentiated models are implemented for dialysis new enrollees and dialysis community enrollees. Similar to the CMS-HCC V22, enrollees are classified as new enrollees if they were not continuously enrolled in Parts A and B for the one-year lookback period prior to each beneficiary month. As a result of this, the model primarily uses information from the beneficiary's enrollment data. This model adjusts for gender, age, dual enrollment status, and whether the beneficiary was originally entitled to Medicare due to disability through a series of interacted covariates. In addition to accounting for these patient characteristics, the dialysis community model also risk adjusts for medical severity using 87 HCCs and additional disease interactions. The CMS-ESRD V21 and CMS-HCC V22 models both generate a risk score for each beneficiary that summarizes the beneficiary's expected cost of care relative to other beneficiaries. Risk scores for ESRD beneficiaries are normalized to enable comparison with the HCC V22 risk scores. This is achieved by multiplying ESRD risk scores by the mean annual Medicare spending for the ESRD population applied in the CMS-ESRD V21 model and dividing by the mean annual Medicare spending for the total Medicare population applied in the CMS-HCC V22 model, effectively renormalizing ESRD risk score values to the equivalent scale of the HCC models. A risk score equal to one indicates risk associated with expenditures for the average beneficiary nationwide. Risk scores below or above one indicate below and above average risk, respectively. The complete list of risk adjustment variables for each model are listed in the Measure Codes List linked in Section S.1 in the tab titled HCC_Risk_Adjust.
Type Score	Continuous variable The TPCC measure score is the average payment-standardized, risk-adjusted, and specialty-adjusted monthly cost across all beneficiary months in the performance period attributed to a clinician or clinician group. A lower measure score indicates that the observed episode costs are lower than or similar to expected costs for the care provided for the particular patients included in the calculation. A higher measure score indicates that the observed episode costs are higher than expected for the care provided for the particular patients included in the calculation.
Algorithm	As described in Section S.7.2, to account for the clinical severity of patients, one of five separate risk adjustment models are applied based on the patients characteristics observed in the year prior to the beneficiary month being measured. For non-ESRD patients, the three models are the new enrollee model, community model, and institutional model from CMS' Hierarchical Condition Category Version 22 (CMS-HCC V22). For ESRD patients, the two models are the dialysis new enrollee model and dialysis community model from CMS' ESRD Version 21 (CMS-ESRD V21). Each model includes beneficiary demographic and enrollment information such as age, gender, disability, and dual enrollment status. Both the new enrollee model and dialysis new enrollee models are limited to these factors as the patient

	does not have sufficient Medicare claims history for further evaluation. The remaining models (community model, institutional model, and dialysis community) include either 79 (CMS-HCC V22) or 87 (CMS-ESRD V21) hierarchical condition categories to characterize the patient severity and comorbidities. The indicators used for risk adjustment and the methodology are detailed in the Measure Information Form linked in Section S.1. The start of a primary care relationship between a clinician and beneficiary is identified by the occurrence of two Part B Physician/Supplier (Carrier) claims with particular CPT/HCPCS services billed in close proximity. There are two different sets of CPT/HCPCS codes used: E&M primary care services and primary care services. E&M primary care services are a specific set of evaluation and management codes for physician visits in the outpatient setting, physician office, nursing facility, or assisted living. Primary care services are a broader list of services related to routine primary care that generally fall into the following categories: Durable Medical Equipment (DME) and Supplies, Electrocardiogram, Laboratory - Chemistry and Hematology, Other Diagnostic Procedures (Interview, Evaluation, Consultation), Other Diagnostic Radiology and Related Techniques, Prophylactic Vaccinations and Inoculations, Routine Chest X-ray, Clinical Labs, and Preventive Services The codes used to attribute beneficiaries to clinicians are listed in the tabs titled E&M_Prim_Care and Prim_Care_Services within the Measure Codes List linked in Section S.1. Clinicians who would not reasonably be responsible for providing primary care are excluded from attribution of the revised TPCC measure using their CMS HCFA specialty designation assigned on Part B physician/supplier claims. This exclusion aims to keep primary care specialists and internal medicine subspecialists who frequently manage patients with chronic conditions falling in their areas of specialty. The excluded specialties list contains 56 specialties that fall into the following broad categories: •Surgical sub-specialties •Non-physicians without chronic management of significant medical conditions •Internal medicine sub-specialties with additional highly procedural sub-specialization •Internal medicine specialties that practice primarily inpatient care without chronic care management •Pediatricians who do not typically practice adult medicine The codes used to exclude clinicians from attribution base on their CMS HCFA specialty are listed in in the tab titled Eligible_Clinicians within the Measure Codes List linked in Section S.1. Additionally, TIN-NPI are removed from attribution if a clinician met any of the following four service category thresholds for the same beneficiary by billing the specified CPT/HCPCS within +/-180 days of the candidate event on Part B physician/supplier claims: •At least 15 percent of the clinician's attributable events are comprised of 10-day or 90-day global surgery services. •At least 5 percent of the clinician's attributable events are comprised of anesthesia services. •At least 5 percent of the clinician's attributable events are comprised of therapeutic radiation services. •At least 10 percent of the clinician's attributable events are comprised of chemotherapy services.

Appendix E: Related and Competing Measures

Comparison of NQF #3561 and NQF #2158

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
Steward	Centers for Medicare & Medicaid Services	Centers for Medicare & Medicaid Services
Description	The Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facility (MSPB-PAC IRF) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each IRF's efficiency relative to that of the national median IRF. Specifically, the measure assesses Medicare spending by the IRF and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each IRF divided by the episode-weighted median MSPB-PAC Amount across all IRFs. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all IRFs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., IRF admissions from October 1, 2015 through September 30, 2017. Claims-based MSPB-PAC measures were developed in parallel for the IRF, long-term care hospital (LTCH), skilled nursing facility (SNF), and home health agency (HHA) settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with IRF experience evaluated IRF claims and then gave direction on how to adjust for specific patient and case-mix characteristics.	The Medicare Spending Per Beneficiary (MSPB) - Hospital measure evaluates hospitals' risk-adjusted episode costs relative to the risk-adjusted episode costs of the national median hospital. Specifically, the MSPB-Hospital measure assesses the cost to Medicare for services performed by hospitals and other healthcare providers during an MSPB-Hospital episode, which is comprised of the periods immediately prior to, during, and following a patient's hospital stay. The MSPB-Hospital measure is not condition specific and uses standardized prices when measuring costs. Beneficiary populations eligible for the MSPB-Hospital calculation include Medicare beneficiaries enrolled in Medicare Parts A and B who were discharged from short-term acute Inpatient Prospective Payment System (IPPS) hospitals during the period of performance.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	The MSPB-PAC IRF measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the IRF Quality Reporting Program (QRP) and finalized in the FY 2017 IRF Prospective Payment System (PPS) Final Rule.[1] Public reporting for the measure began in Fall 2018 through the IRF Compare website (https://www.medicare.gov/inpatientrehabilitationfacilitycompare/) using FY 2016-2017 data. Notes: [1] Medicare Program; Inpatient Rehabilitation Facility Prospective Payment System for Federal Fiscal Year 2017 Federal Register, Vol. 81, No. 151. https://www.gpo.gov/fdsys/pkg/FR-2016-08-05/pdf/2016-18196.pdf	
Type	Cost/Resource Use	Cost/Resource Use
Data Source	Assessment Data, Claims, Enrollment Data, Other This measure is based on Medicare FFS administrative claims and uses data from the Medicare enrollment database and Minimum Data Set (MDS). The enrollment database provides information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A and Part B coverage, and periods in the Medicare FFS program. The MDS is used to construct a risk adjustment variable, indicating beneficiaries who have been institutionalized for at least 90 days in a given year. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, and revenue center codes. The Medicare FFS claims data files are used to identify Medicare services from IRFs and other settings (e.g., the outpatient setting) within the episode window. No data beyond the claims submitted in the normal course of business are required from providers for the calculation of this measure. This measure submission is based on FY 2016-2017 data, which were the most recent data available at the time of our analyses.	Claims, Other The MSPB-Hospital measure uses Medicare Part A and Part B claims data, which is maintained by CMS' Office of Information System (OIS). Data from the Medicare Enrollment Database (EDB) are used to predict costs of episodes and determine beneficiary-level exclusions, specifically to determine the following: Medicare Parts A, B, and C enrollment; primary payer; disability status; end-stage renal disease (ESRD); beneficiary birth dates; and beneficiary death dates. The risk adjustment model also accounts for expected differences in payment for services provided to beneficiaries in long term care, and that information comes from the Minimum Data Set (MDS). The MDS is used to create the Long Term Care Indicator variable in risk adjustment (denoted as LTC_Indicator). Data from the American Community Survey (ACS) is in the analyses performed to evaluate including SES/SDS in risk adjustment (see Testing Attachment Section 2b4). Claims, Other Data dictionary URL The MSPB-Hospital measure relies on Medicare claims data. The Research Data Assistance Center (ResDAC) maintains an updated Medicare claims data dictionary available at the following URL: http://www.resdac.org/cms-data/file-family/Medicare-Claims.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	We used the data sources listed below to develop the analytic file for measure specification and testing: • Medicare Fee-For-Services claims and enrollment data: We accessed inpatient, outpatient, carrier, skilled nursing facility, home health, durable medical equipment, and hospice claims through the Centers for Medicare & Medicaid Services (CMS) Common Working File (CWF). The data dictionary for all Medicare FFS claims, demographic, and enrollment data are available at: https://www.resdac.org/cms-data?tid%5B%5D=4931&tid_1%5B%5D=1&=Find+Data+Files. General information about the CWF is available at: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/clm104c27.pdf. • Minimum Data Set (MDS): Acumen obtains the MDS through the Quality Improvement and Evaluation System (QIES). The data dictionary for the MDS data is available at: https://www.resdac.org/cms-data/files/mds-3.0/data-documentation. We used two mappings to group diagnosis and procedure codes for use in identifying clinical events, implementing exclusions and applying risk adjustment: • Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings for Services and Procedures: Software is available for download at: https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp • CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes: We used the Version 22 CMS-HCC mapping, which is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. We used five additional data sources for measure testing purposes only and not for measure specification: • 2017 American Community Survey (ACS) 5-year estimate: We used the ACS to obtain the ZIP Code Tabulation	

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	Area (ZCTA) level measures needed to compute the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score for use in social risk factor testing. This information is downloadable at the US Census website: http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t. • Rural-Urban Continuum Codes 2013: We used this data source to construct rural-urban identifiers for social risk factor testing. These codes include county FIPS indicators, which are then merged onto our episode file. More information on this data source can be found at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. • Provider of Services Current Files (POS File): We used this data source to describe the characteristics of IRFs included in measure specification and testing, such as census region, ownership type, and rurality, as reported in Table 1. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html. • IRF Compare data: We used this data source to examine the relationship between MSPB and assessment-based quality measures. The IRF Compare data include publicly reported IRF quality measures. The data are available at https://data.medicare.gov/data/inpatient-rehabilitation-facility-compare • Common Medicare Environment (CME) database: We extracted patient-level dual eligibility information from the CME database for social risk factor testing. CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME	

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	database receives information from the EDB and also contains additional information not available in the EDB. A description of the CME is available at: https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. Assessment Data, Claims, Enrollment Data, Other Data dictionary URL; Code table attachment Data dictionary URL; Code table attachment	
Level	Facility	Facility
Setting	Post-Acute Care	Inpatient/Hospital
Numerator Statement	The numerator is the MSPB-PAC IRF Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider. This is then multiplied by the national average episode spending level for all IRF providers nationally.	The numerator for a hospital's MSPB-Hospital measure is the hospital's MSPB-Hospital amount, which is the average spending level for the hospital's MSPB-Hospital episodes divided by the average expected episode spending level for the hospital's episodes, multiplied by the average spending over all episodes across all hospitals nationally. An MSPB-Hospital episode includes all Medicare Part A and Part B claims with a start date falling between 3 days prior to an Inpatient Prospective Payment System (IPPS) hospital admission (also known as the "index admission" for the episode) through 30 days post-hospital discharge.
Numerator Details	N/A	N/A
Denominator Statement	The denominator is the episode-weighted national median of the MSPB-PAC IRF Amounts for all IRFs nationally.	The denominator for a hospital's MSPB-Hospital measure is the episode-weighted median MSPB-Hospital amount across all episodes nationally.
Denominator Details	N/A	N/A
Exclusions	Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC IRF episodes because they are clinically unrelated to IRF care and/or because IRF providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given IRF provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by	Populations excluded from the MSPB-Hospital calculation are any episodes where at any time 90 days before or during the episode, the beneficiary is enrolled in a Medicare Advantage plan or Medicare is the secondary payer. Episodes where the beneficiary becomes deceased during the episode are also excluded. Regarding beneficiaries whose primary insurance becomes Medicaid during an episode due to exhaustion of Medicare Part A benefits, the beneficiaries themselves are not excluded. Rather, Medicaid payments made for services rendered to these beneficiaries are excluded, while all Medicare Part A

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration.	payments made before benefits are exhausted and all Medicare Part B payments made during the episode are included. We believe this is the most appropriate method for addressing benefits exhaust episodes, because these beneficiaries represent high resource use cases that should be included in a hospital's measure. In addition, this removes the potential for hospitals to exhaust a beneficiary's Part A benefits to exclude high resource use episodes from their measure. Further, any episode in which the index admission inpatient claim has a $0 actual payment or a $0 standardized payment is excluded. In addition, acute-to-acute transfers (where a transfer is defined based on the claim discharge code) are not considered index admissions. In other words, these cases do not generate new MSPB-Hospital episodes; neither the hospital which transfers a patient to another subsection (d) hospital, nor the receiving subsection (d) hospital will have an index admission or associated MSPB-Hospital episode attributed to them. This exclusion addresses stakeholder concerns that neither the admitting nor receiving hospital is fully able to coordinate care. Stakeholders stated that it was inappropriate to hold the initially-admitting hospital accountable for services rendered by the receiving hospital. In addition, stakeholders expressed concern with holding the receiving hospital accountable for any issues that arose as a result of the initially-admitting hospital's care and/or follow up care rendered near the beneficiary's home, where the receiving hospital may not be in an ideal place to coordinate that care. Admissions to hospitals that Medicare does not reimburse through the IPPS system (e.g., cancer hospitals, critical access hospitals, hospitals in Maryland) are not considered index admissions and are therefore not eligible to begin an MSPB-Hospital episode. If an acute-to-acute hospital transfer or a hospitalization in a PPS-exempt hospital type happens during the 30-day window following an included index admission, however, it will be counted in the measure. This is because the MSPB-Hospital measure includes all claims and services that occur 30 days after discharge from the index hospital; an episode includes the 30 days after a hospital discharge to emphasize the importance of care transitions and care coordination in improving patient care.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
Exclusion Details	Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include: •Planned hospital admissions[1] •Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants) •Some routine screening and health care maintenance (e.g., colonoscopy and mammograms) •Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis) Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as: 1)Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories. Rationale: This exclusion ensures that complete claims data are available for each provider. 2)Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated. Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC IRF measure could potentially misrepresent a providers' resource use. 3)Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day	The following lists details the exclusions made to all episodes of care for which full data are not available or for which Medicare spending by itself cannot reasonably be considered a signal of efficiency: • [I] Any episodes without all observable claims or a complete episode window (i.e., episodes in which Medicare is the secondary payer, episodes in which the beneficiary is enrolled in a Medicare Advantage plan, episodes in which the beneficiary is enrolled only in Medicare Part A, episodes in which the beneficiary becomes deceased). Episodes in which the beneficiary is enrolled only in Medicare Part A, for example, are excluded because these beneficiaries may receive services not observed in the data. Similarly, episodes in which the beneficiary dies at any point during the episode. Episodes in which the patient dies are—by definition—truncated episodes and do not have a complete episode window. Episodes in which the patient dies were identified as an index hospitalization with death discharge code (STUS_CD "20" "41") or if a beneficiary's death was within an MSPB-Hospital episode. Including episodes without all observable claims or a complete episode window could potentially make hospitals seem efficient not due to any action of their own, but because the data are missing services that would be included in the MSPB-Hospital measure calculation. Episodes where Medicare is the secondary payer: if a beneficiary was the primary payer any time during the MSPB-Hospital episode, the beneficiary was excluded (i.e., if bene_prmry_pyr_entlmt_strt_dt (start date of primary payer enrollment) bene_prmry_pyr_entlmt_end_dt (end date of primary payer enrollment) fell within the episode). • [II] Regarding beneficiaries whose primary insurance becomes Medicaid during an episode due to exhaustion of Medicare Part A benefits, these beneficiaries are not excluded. Rather, Medicaid payments made for services rendered to these beneficiaries are excluded; all Medicare Part A payments made before benefits are exhausted and all Medicare Part B payments made during the episode are included. The MSPB-Hospital measure is calculated using only Medicare Part A and Part B claims; as a result no Medicaid claims are included in the MSPB-Hospital measure calculation.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window. Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC IRF measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims. 4)Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window. Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data. 5)Any episode where the claim(s) constituting the attributed PAC provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill. Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization. 6)Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory) Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements. Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the	• [III] Any episode in which the index admission inpatient claim has a $0 actual payment or a $0 standardized payment is excluded. $0 inpatient admissions may represent errors in the data, or payment corrections rather than actual services rendered. Only when the Claim Payment amount (pmt_amt) for the IP stay is greater than 0 OR standard_allowed_amt is greater than 0 is the amount included in the MSPB-Hospital measure calculation. • [IV] Due to the uncertainty surrounding attributing episodes to hospitals in cases where the patient was transferred between acute hospitals during the index admission, acute-to-acute transfers during the index admission (where a transfer is defined based on the claim discharge code) are not considered index admissions for the purposes of the MSPB-Hospital measure. In other words, these cases will not generate new MSPB-Hospital episodes; neither the hospital which transfers a patient to another short-term acute hospital, nor the receiving short-term acute hospital will have an index admission attributed to them. This exclusion avoids assigning responsibility to an MSPB-Hospital episode in a case where multiple hospitals treat the patient during the index admission. • [V] Cancer hospitals, MD Hospitals (provider variable starting with "21"), emergency hospitals (provider variable last position "E" OR "F"), and veteran's hospital (provider variable position "V") are also excluded. • [VI] In response to stakeholder comments, the FY 2012 IPPS Final Rule states that the MSPB-Hospital measure will "exclude statistical outliers from the calculation" (76 FR 51626: www.gpo.gov/fdsys/pkg/FR-2011-08-18/pdf/2011-19719.pdf). To mitigate the effect of high-cost outliers on each hospital's MSPB-Hospital measure score, MSPB-Hospital episodes whose relative scores fall above the 99th percentile or below the 1st percentile of the distribution of residuals are excluded from the MSPB-Hospital calculation. Excluding outliers based on residuals eliminates the episodes that deviate most from their predicted values in absolute terms.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers. Notes: [1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6	
Risk Adjustment	Statistical risk model The detailed steps to computing the measure score are described in section S.7.2. Risk-adjustment is applied in “Step 3: Calculate Predicted Episode Payments.” The purpose of risk adjustment is to compensate for patient health circumstances and demographic factors that affect resource use but are beyond the influence of the attributed provider. The MSPB-PAC IRF measure risk adjustment model is adapted from the model used in the NQF-endorsed MSPB-Hospital measure, which itself is an adaptation of the standard CMS-HCC risk-adjustment model.[1,2] The MSPB-PAC IRF model uses a linear regression framework and a 90-day HCC lookback period. The risk adjustment model is estimated on all MSPB-PAC IRF episodes that meet the exclusion criteria. Each provider’s MSPB-PAC IRF measure score is calculated as a provider’s average MSPB-PAC Amount divided by the median MSPB-PAC Amount across all providers. A provider’s MSPB-PAC IRF Amount is defined as the	Statistical risk model

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	sum of standardized, risk-adjusted spending across all of a provider's eligible episodes divided by the number of episodes for that provider. Below is a description of the risk adjustment variables. Risk-Adjustment Variables The following beneficiary health status indicators are included as covariates in each MSPB-PAC IRF risk adjustment model and to the greatest extent possible are consistent across PAC settings (see Appendix C of the Measure Specifications document provided in section S.1 for a comprehensive list of independent variables used in the risk adjustment model): •70 HCCs •11 HCC interactions •11 brackets for age at the start of the episode •Original entitlement to Medicare through disability •ESRD status •Long-term care institutionalization at start of episode.[3] •Six clinical case-mix categories reflecting recent prior care (described further below).[4] •Hospice utilization during the episode •Prior acute ICU utilization day categories •Prior acute length of stay categories •Rehabilitation Impairment Categories (RICs) The clinical case-mix category variables used in the MSPB-PAC IRF risk adjustment model are included to account for differences in intensity and type of care received by beneficiaries prior to the start of an MSPB-PAC IRF episode. See section S.7.5 for more details on the methodology of assigning clinical case-mix categories to each episode. Notes: [1] QualityNet, "CMS Price (Payment) Standardization – Detailed Methods" (Revised April 2019)	

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	https://www.qualitynet.org/inpatient/measures/payment-standardization [2] CMS, "Medicare Risk Adjustment Information" (2016) https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html [3] Identifies beneficiaries who have been institutionalized for at least 90 days in a given year. The indicator is based on 90-day assessments from the Minimum Data Set (MDS) and is calculated based on CMS' definition of institutionalized individuals. [4] There are 7 case-mix categories as described above, but one category is removed to prevent collinearity. Statistical risk model	
Stratification	Not applicable: the MSBP-PAC IRF measure is not stratified.	While the measure results are not stratified, expected costs for episodes are determined by using a separate risk adjustment model for episodes within each MDC. MDCs are aggregations of Diagnosis Related Groups (MS-DRG), which CMS uses to classify acute inpatient admissions. The MS-DRG/MDC crosswalk is available for order here: http://solutions9.3m.com/wps/portal/!ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz94NS8-NBg_Qj9KLP4IC8Py1BTI2MD9zAvFwMjYzMzCxNHd2OTACP9ggxHRQBm3gTM/
Type Score	Ratio An MSPB-PAC IRF measure score of 1 indicates that an IRF had an average MSPB-PAC Amount (i.e., risk-adjusted spending level) which is equal to the national episode-weighted median MSPB-PAC Amount across all IRFs during a given performance period. An MSPB-PAC IRF measure score of greater than 1 indicates that an IRF had higher average risk-adjusted spending levels compared to those of the national median IRF. For example, a measure score of 1.1 indicates that the IRF had average risk-adjusted spending levels that are 10 percent higher than the median IRF. On the other hand, an MSPB-PAC IRF measure score of less than 1 indicates that an IRF had lower average risk-adjusted spending levels compared to those of the median IRF. For example, a measure score of 0.9 indicates that	Ratio; Attachment An MSPB-Hospital measure that is less than 1 indicates that a given hospital's MSPB-Hospital amount (i.e. risk-adjusted spending) is less than the national episode-weighted median MSPB-Hospital amount across all hospitals during a given performance period. We note that results of the MSPB-Hospital measure alone do not necessarily reflect the quality of care provided by hospitals. Accordingly, lower MSPB-Hospital measure across performance periods (i.e., lower Medicare spending per beneficiary) in isolation should not be interpreted as better care. The MSPB-Hospital measure is most meaningful when presented in the context of other quality measures, which are part of the Hospital Value-Based Purchasing (VBP) Program. As part of the Hospital VBP Program, the MSPB-Hospital measure is aligned with current quality of care measures to facilitate

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	the IRF had average risk-adjusted spending levels that are 10 percent lower than the median IRF.	profiling hospital value (payments and quality). Improvement on this measure for a hospital would be observed as a lower MSPB-Hospital measure value across performance periods.
Algorithm	In order to create a resource use measure that is clinically valid, there were multiple steps involved in excluding the least clinically relevant codes. Using an episode window, we organized claims into clinically meaningful service categories or settings. For example, Medicare Severity-Diagnosis Related Groups (MS-DRGs) noted after an IRF discharge were evaluated as medical or surgical admissions post-discharge. Clinical Classifications Software (CCS) and Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) services were organized into outpatient services, emergency department (ER) services, and durable medical equipment claims and evaluated for their relevance or relatedness to IRF care. Extensive clinical review was performed by clinicians with experience providing care in IRF settings, as well as in collaboration with Medical Officers at CMS. The hospitalizations and outpatient services least clinically related to the IRF care were excluded from resource use calculation. For instance, it was not felt that an IRF could influence a beneficiary's rehospitalization for nervous system neoplasms (DRG 054), post-discharge outpatient services for kidney transplant (CCS 105), or routine fecal occult blood testing (CPT 82270). Therefore, these types of services were excluded. Services were only added to the exclusions list if there was consensus across IRF and CMS clinicians. Please see section S.9.1 for overall clinical consensus regarding the types of exclusions. To account for the association between clinical severity and resource use, we risk adjust the total observed episode spending (described in section S.12) using CMS-HCC indicators and interactions between selected comorbidities. Diagnosis codes on claims that occur during the 90-day period prior to the	Objective: The MSPB-Hospital measure aims to improve care coordination in the period between 3 days prior to an acute inpatient hospital admission through the period 30 days after discharge. The MSPB-Hospital measure recognizes lower costs associated with a reduction in unnecessary services, preventable complications, readmissions, and shifting post-acute care from more expensive to less expensive services when appropriate. Grouping methodology: The MSPB-Hospital measure evaluates resource use through the unit of MSPB-Hospital episodes. The MSPB-Hospital episodes are constructed by including all Medicare Part A and Part B claims with a start date falling between 3 days prior to an acute inpatient hospital admission through the period 30 days after discharge. Any episodes where at any time during the episode the beneficiary is enrolled in a Medicare Advantage plan, the beneficiary becomes deceased, or Medicare is the secondary payer will be excluded from the MSPB-Hospital calculation. Regarding beneficiaries whose primary insurance becomes Medicaid during an episode due to exhaustion of Medicare Part A benefits, Medicaid payments made for services rendered to these beneficiaries are excluded; however, all Medicare Part A payments made before benefits are exhausted and all Medicare Part B payments made during the episode are included. Cost Calculation: The MSPB-Hospital amount includes the cost of services performed by hospitals and other healthcare providers during an MSPB-Hospital episode, which is comprised of the period 3 days prior to an inpatient PPS hospital admission (index admission) through 30 days post-hospital discharge. All costs are payment standardized to control for geographic variation in Medicare reimbursement rates. All costs are risk adjusted to account for age and severity of illness.

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	start of an MSPB-PAC IRF episode (90-day "look back") are used to create HCC indicators. The MSPB-PAC IRF measure accounts for comorbid conditions and interactions by broadly following the CMS-HCC risk adjustment methodology, which is derived from Medicare Part A and B claims and is used in the Medicare Advantage (MA) program. For example, the measure accounts for interactions between disability and selected HCC groups (e.g., Cystic Fibrosis, Severe Hematological Disorders, Opportunistic Infections, among others). Given the fact that beneficiaries often have more than one comorbidity, the model also includes commonly observed paired condition interactions, (e.g., chronic obstructive pulmonary disease [COPD] and congestive heart failure [CHF]) and commonly observed triple-interactions (e.g., diabetes mellitus, congestive heart failure, and renal failure). The full list of variables used in the risk adjustment model can be found in the Measure Specifications document provided in section S.1. In addition to comorbidities, the MSPB-PAC IRF measure utilizes clinical case-mix categories to create clinically meaningful subgroups that influence the type of services a beneficiary will receive in an IRF. To create these subgroups, information was derived from the institutional claim of the most recent hospitalization. The clinical case-mix category variables used in the MSPB-PAC IRF risk-adjustment model are included to account for differences in intensity and type of care received by beneficiaries prior to the start of an MSPB-PAC IRF episode. Taking the most recent institutional claim (by end date) in the 60 days prior to the start of an MSPB-PAC IRF episode, the episode is assigned to one of the following mutually exclusive and exhaustive clinical case-mix categories: 1) Prior Acute Surgical IP – Orthopedic – beneficiaries who have most recently undergone orthopedic surgery in an acute inpatient hospital	

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	2) Prior Acute Surgical IP – Non-Orthopedic – beneficiaries who have most recently undergone a non-orthopedic surgery in an acute inpatient hospital 3) Prior Acute Medical IP with ICU – beneficiaries who have most recently stayed in an acute inpatient hospital for non-surgical reasons and had a stay in the ICU 4) Prior Acute Medical IP without ICU – beneficiaries who have most recently stayed in an acute inpatient hospital for non-surgical reasons but did not have a stay in the ICU 5) Prior PAC - Institutional – beneficiaries who are continuing PAC from an institutional PAC setting (i.e., coming from an LTCH, IRF, or SNF) 6) Prior PAC - HHA – beneficiaries who are continuing PAC from a HHA 7) Community – all other beneficiaries Finally, the MSPB-PAC IRF measure uses RICs from the IRF admission. A full list of the RICs used in the risk adjustment model is included in Appendix C of the Measure Specifications document provided in section S.1. To simplify the clinical logic and avoid the issue of attributing claims to MSPB-PAC IRF episodes in the case of concurrent clinical events, all claims that begin within the episode window (treatment period and associated services period) are included in the MSPB-PAC IRF measure. An MSPB-PAC IRF episode is assigned to the rehabilitation facility of the index admission. A new episode may begin during the associated services period of a previous MSPB-PAC IRF episode in the 30 days post-discharge from the IRF.	
		5.1 Identified measures: 5a.1 Are specs completely harmonized? 5a.2 If not completely harmonized, identify difference, rationale, impact:

	3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
		5b.1 If competing, why superior or rationale for additive value: The MSPB-Hospital measure evaluates hospitals' efficiency relative to the efficiency of the median hospital. The target population is Medicare beneficiaries enrolled in Medicare Parts A and B who were discharged from short-term acute hospitals. There are currently no NQF-endorsed measures that address both this same measure focus AND this same target population.

Comparison of NQF #3562 and NQF #2158

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
Steward	Centers for Medicare & Medicaid Services	Centers for Medicare & Medicaid Services
Description	The Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals (MSPB-PAC LTCH) was developed to address the resource use domain of the Improving Medicare Post-Acute Care Transformation Act of 2014 (IMPACT Act). This resource use measure is intended to evaluate each LTCH's efficiency relative to that of the national median LTCH. Specifically, the measure assesses Medicare spending by the LTCH and other healthcare providers during an MSPB episode. The measure reports the ratio of the payment-standardized, risk-adjusted MSPB-PAC Amount for each LTCH divided by the episode-weighted median MSPB-PAC Amount across all LTCH facilities. The MSPB-PAC Amount is the ratio of the observed episode spending to the expected episode spending, multiplied by the national average episode spending for all LTCHs. The measure is calculated using two consecutive years of Medicare Fee-for-Service (FFS) claims data and was developed using calendar year (CY) 2015-2016 data. This submission is based on fiscal year (FY) 2016-2017 data; i.e., LTCH admissions from October 1, 2015 through September 30, 2017. Claims-based MSPB-PAC measures were developed in parallel for the LTCH, inpatient rehabilitation facility (IRF), skilled nursing facility (SNF), and home health agency (HHA) settings to meet	The Medicare Spending Per Beneficiary (MSPB) - Hospital measure evaluates hospitals' risk-adjusted episode costs relative to the risk-adjusted episode costs of the national median hospital. Specifically, the MSPB-Hospital measure assesses the cost to Medicare for services performed by hospitals and other healthcare providers during an MSPB-Hospital episode, which is comprised of the periods immediately prior to, during, and following a patient's hospital stay. The MSPB-Hospital measure is not condition specific and uses standardized prices when measuring costs. Beneficiary populations eligible for the MSPB-Hospital calculation include Medicare beneficiaries enrolled in Medicare Parts A and B who were discharged from short-term acute Inpatient Prospective Payment System (IPPS) hospitals during the period of performance.

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	the mandate of the IMPACT Act. To align with the goals of standardized assessment across all settings in PAC, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. Clinically meaningful case-mix considerations were evaluated at the level of each setting. For example, clinicians with LTCH expertise evaluated LTCH claims and then gave direction on how to adjust for specific patient and case-mix characteristics. The MSPB-PAC LTCH measure was adopted by the Centers for Medicare & Medicaid Services (CMS) for the LTCH Quality Reporting Program (QRP) and finalized in the FY 2017 LTCH Prospective Payment System (PPS) Final Rule.[1] The measure entered into use on October 1, 2016. Public reporting for the measure began in Fall 2018 through the LTCH Compare website (https://www.medicare.gov/longtermcarehospitalcompare/) using FY 2016-2017 data. Notes: [1] Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2017 Rates. Federal Register, Vol. 81, No. 162. https://www.govinfo.gov/content/pkg/FR-2016-08-22/pdf/2016-18476.pdf	
Type	Cost/Resource Use	Cost/Resource Use
Data Source	Assessment Data, Claims, Enrollment Data, Other This measure is based on Medicare FFS administrative claims and uses data from the Medicare enrollment database and Minimum Data Set (MDS). The enrollment database provides information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A and Part B coverage, and periods in the Medicare FFS program. The MDS is used to construct a risk adjustment variable, indicating beneficiaries who have been institutionalized for at least 90 days in a given year. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses,	Claims, Other The MSPB-Hospital measure uses Medicare Part A and Part B claims data, which is maintained by CMS' Office of Information System (OIS). Data from the Medicare Enrollment Database (EDB) are used to predict costs of episodes and determine beneficiary-level exclusions, specifically to determine the following: Medicare Parts A, B, and C enrollment; primary payer; disability status; end-stage renal disease (ESRD); beneficiary birth dates; and beneficiary death dates. The risk adjustment model also accounts for expected differences in payment for services provided to beneficiaries in long term care, and that information comes from the Minimum Data Set (MDS). The MDS is used to create the Long Term Care Indicator variable in risk adjustment (denoted as LTC_Indicator).

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	procedures, and revenue center codes. The Medicare FFS claims data files are used to identify Medicare services from LTCH and other settings (e.g., the outpatient setting) within the episode window. No data beyond the claims submitted in the normal course of business are required from providers for the calculation of this measure. This measure submission is based on FY 2016-2017 data, which were the most recent data available at the time of our analyses. We used the data sources listed below to develop the analytic file for measure specification and testing: • Medicare Fee-For-Services claims and enrollment data: We access inpatient, outpatient, carrier, skilled nursing facility, home health, durable medical equipment, and hospice claims through the Centers for Medicare & Medicaid Services (CMS) Common Working File (CWF). The data dictionary for all Medicare FFS claims, demographic, and enrollment data are available at: https://www.resdac.org/cms-data?tid%5B%5D=4931&tid_1%5B%5D=1&=Find+Data+Files. General information about the CWF is available at: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/clm104c27.pdf. • Minimum Data Set (MDS): Acumen obtains the MDS through the Quality Improvement and Evaluation System (QIES). The data dictionary for the MDS data is available at: https://www.resdac.org/cms-data/files/mds-3.0/data-documentation. We used two mappings to group diagnosis and procedure codes for use in identifying clinical events, implementing exclusions and applying risk adjustment: • Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings for Services and Procedures: Software is available for download at: https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp • CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes: We used the Version 22 CMS-HCC	Data from the American Community Survey (ACS) is in the analyses performed to evaluate including SES/SDS in risk adjustment (see Testing Attachment Section 2b4). Claims, Other Data dictionary URL The MSPB-Hospital measure relies on Medicare claims data. The Research Data Assistance Center (ResDAC) maintains an updated Medicare claims data dictionary available at the following URL: http://www.resdac.org/cms-data/file-family/Medicare-Claims.

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	mapping, which is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. We used five additional data sources for measure testing purposes only and not for measure specification: • 2017 American Community Survey (ACS) 5-year estimate: We used the ACS to obtain the ZIP Code Tabulation Area (ZCTA) level measures needed to compute the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score for use in social risk factor testing. This • http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t. • Rural-Urban Continuum Codes 2013: We used this data source to construct rural-urban identifiers for social risk factor testing. These codes include county FIPS indicators, which are then merged onto our episode file. More information on this data source can be found at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. • Provider of Services Current Files (POS File): We used this data source to describe the characteristics of LTCH facilities included in measure specification and testing, such as census region, ownership type, and rurality, as reported in Table 1. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html. • LTCH Compare data: We used this data source to examine the relationship between MSPB and assessment-based quality measures. The LTCH Compare data include publicly reported LTCH quality measures. The data are available at https://data.medicare.gov/data /long-term-care-hospital-compare.	

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	• Common Medicare Environment (CME) database: We extracted patient-level dual eligibility information from the CME database for social risk factor testing. CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME database receives information from the EDB and also contains additional information not available in the EDB. A description of the CME is available at: https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. This measure is based on Medicare FFS administrative claims and uses data from the Medicare enrollment database and Minimum Data Set (MDS). The enrollment database provides information such as date of birth, date of death, sex, reasons for Medicare eligibility, periods of Part A and Part B coverage, and periods in the Medicare FFS program. The MDS is used to construct a risk adjustment variable, indicating beneficiaries who have been institutionalized for at least 90 days in a given year. The data elements from the Medicare FFS claims are those basic to the operation of the Medicare payment systems and include data such as date of admission, date of discharge, diagnoses, procedures, and revenue center codes. The Medicare FFS claims data files are used to identify Medicare services from LTCH and other settings (e.g., the outpatient setting) within the episode window. No data beyond the claims submitted in the normal course of business are required from providers for the calculation of this measure. This measure submission is based on FY 2016-2017 data, which were the most recent data available at the time of our analyses. We used the data sources listed below to develop the analytic file for measure specification and testing: • Medicare Fee-For-Services claims and enrollment data: We access inpatient, outpatient, carrier, skilled nursing facility, home health, durable medical equipment, and hospice claims through the Centers for Medicare & Medicaid Services (CMS)	

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	Common Working File (CWF). The data dictionary for all Medicare FFS claims, demographic, and enrollment data are available at: https://www.resdac.org/cms-data?tid%5B%5D=4931&tid_1%5B%5D=1&=Find+Data+Files. General information about the CWF is available at: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/clm104c27.pdf. • Minimum Data Set (MDS): Acumen obtains the MDS through the Quality Improvement and Evaluation System (QIES). The data dictionary for the MDS data is available at: https://www.resdac.org/cms-data/files/mds-3.0/data-documentation. We used two mappings to group diagnosis and procedure codes for use in identifying clinical events, implementing exclusions and applying risk adjustment: • Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) groupings for Services and Procedures: Software is available for download at: https://www.hcup-us.ahrq.gov/toolssoftware/ccs_svcsproc/ccssvcproc.jsp • CMS-Hierarchical Condition Category (HCC) mappings of ICD-9 and ICD-10 codes: We used the Version 22 CMS-HCC mapping, which is included in the software available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. We used five additional data sources for measure testing purposes only and not for measure specification: • 2017 American Community Survey (ACS) 5-year estimate: We used the ACS to obtain the ZIP Code Tabulation Area (ZCTA) level measures needed to compute the Agency for Healthcare Research and Quality (AHRQ) Socioeconomic Status (SES) index score for use in social risk factor testing. This • http://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t.	

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	• Rural-Urban Continuum Codes 2013: We used this data source to construct rural-urban identifiers for social risk factor testing. These codes include county FIPS indicators, which are then merged onto our episode file. More information on this data source can be found at: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. • Provider of Services Current Files (POS File): We used this data source to describe the characteristics of LTCH facilities included in measure specification and testing, such as census region, ownership type, and rurality, as reported in Table 1. The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the CMS Regional Offices. General information about the POS Files is available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/index.html. • LTCH Compare data: We used this data source to examine the relationship between MSPB and assessment-based quality measures. The LTCH Compare data include publicly reported LTCH quality measures. The data are available at https://data.medicare.gov/data /long-term-care-hospital-compare. • Common Medicare Environment (CME) database: We extracted patient-level dual eligibility information from the CME database for social risk factor testing. CMS has designated the CME database as the single, enterprise-wide authoritative source for Medicare beneficiary enrollment and demographic data. The CME database integrates and standardizes different types of beneficiary data from CMS legacy systems. The CME database receives information from the EDB and also contains additional information not available in the EDB. A description of the CME is available at: https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf.	

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	Data dictionary URL; Code table attachment Data dictionary URL; Code table attachment	
Level	Facility	Facility
Setting	Post-Acute Care	Inpatient/Hospital
Numerator Statement	The numerator is the MSPB-PAC LTCH Amount, or the average risk-adjusted episode spending across all episodes for the attributed provider, comparing Standard and Site Neutral episodes only with episodes of the same type. This is then multiplied by the national average episode spending level for all LTCH providers nationally.	The numerator for a hospital's MSPB-Hospital measure is the hospital's MSPB-Hospital amount, which is the average spending level for the hospital's MSPB-Hospital episodes divided by the average expected episode spending level for the hospital's episodes, multiplied by the average spending over all episodes across all hospitals nationally. An MSPB-Hospital episode includes all Medicare Part A and Part B claims with a start date falling between 3 days prior to an Inpatient Prospective Payment System (IPPS) hospital admission (also known as the "index admission" for the episode) through 30 days post-hospital discharge.
Numerator Details	N/A	N/A
Denominator Statement	The denominator is the episode-weighted national median of the MSPB-PAC LTCH Amounts for all LTCH facilities nationally.	The denominator for a hospital's MSPB-Hospital measure is the episode-weighted median MSPB-Hospital amount across all episodes nationally.
Denominator Details	N/A	N/A
Exclusions	Exclusion of clinically unrelated services. Certain services are excluded from the MSPB-PAC LTCH episodes because they are clinically unrelated to LTCH care and/or because LTCH providers may have limited influence over certain Medicare services delivered by other providers during the episode window. These limited service-level exclusions are not counted towards a given LTCH provider's Medicare spending to ensure that beneficiaries with certain conditions and complex care needs receive the necessary care. The list of excluded services was developed by obtaining consensus on the exclusion of each service from CMS clinicians, eight independently contracted clinicians (including two TEP members) with expertise in each of the PAC settings, and the measure developer's clinicians. Feedback from the TEP provided through the in-person meeting and follow-up email survey was also taken into consideration.	Populations excluded from the MSPB-Hospital calculation are any episodes where at any time 90 days before or during the episode, the beneficiary is enrolled in a Medicare Advantage plan or Medicare is the secondary payer. Episodes where the beneficiary becomes deceased during the episode are also excluded. Regarding beneficiaries whose primary insurance becomes Medicaid during an episode due to exhaustion of Medicare Part A benefits, the beneficiaries themselves are not excluded. Rather, Medicaid payments made for services rendered to these beneficiaries are excluded, while all Medicare Part A payments made before benefits are exhausted and all Medicare Part B payments made during the episode are included. We believe this is the most appropriate method for addressing benefits exhaust episodes, because these beneficiaries represent high resource use cases that should be included in a hospital's measure. In addition, this

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
		removes the potential for hospitals to exhaust a beneficiary's Part A benefits to exclude high resource use episodes from their measure. Further, any episode in which the index admission inpatient claim has a $0 actual payment or a $0 standardized payment is excluded. In addition, acute-to-acute transfers (where a transfer is defined based on the claim discharge code) are not considered index admissions. In other words, these cases do not generate new MSPB-Hospital episodes; neither the hospital which transfers a patient to another subsection (d) hospital, nor the receiving subsection (d) hospital will have an index admission or associated MSPB-Hospital episode attributed to them. This exclusion addresses stakeholder concerns that neither the admitting nor receiving hospital is fully able to coordinate care. Stakeholders stated that it was inappropriate to hold the initially-admitting hospital accountable for services rendered by the receiving hospital. In addition, stakeholders expressed concern with holding the receiving hospital accountable for any issues that arose as a result of the initially-admitting hospital's care and/or follow up care rendered near the beneficiary's home, where the receiving hospital may not be in an ideal place to coordinate that care. Admissions to hospitals that Medicare does not reimburse through the IPPS system (e.g., cancer hospitals, critical access hospitals, hospitals in Maryland) are not considered index admissions and are therefore not eligible to begin an MSPB-Hospital episode. If an acute-to-acute hospital transfer or a hospitalization in a PPS-exempt hospital type happens during the 30-day window following an included index admission, however, it will be counted in the measure. This is because the MSPB-Hospital measure includes all claims and services that occur 30 days after discharge from the index hospital; an episode includes the 30 days after a hospital discharge to emphasize the importance of care transitions and care coordination in improving patient care.
Exclusion Details	Additional information on the process for developing the list of clinically unrelated services is available in Appendix D of the Measure Specifications document provided in section S.1. The specialties of the non-CMS clinicians with whom we consulted during the measure development process are provided in Appendix F of the Measure Specifications document provided in	The following lists details the exclusions made to all episodes of care for which full data are not available or for which Medicare spending by itself cannot reasonably be considered a signal of efficiency: • [l] Any episodes without all observable claims or a complete episode window (i.e., episodes in which Medicare is the secondary payer, episodes in which the beneficiary is enrolled in a Medicare Advantage plan, episodes in which the beneficiary is enrolled only in

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	section S.1. Services that were determined by clinical consensus to be outside of the control of PAC providers include: • Planned hospital admissions[1] • Routine management of certain preexisting chronic conditions (e.g., dialysis for end-stage renal disease (ESRD), enzyme treatments for genetic conditions, treatment for preexisting cancers, and treatment for organ transplants) • Some routine screening and health care maintenance (e.g., colonoscopy and mammograms) • Immune modulating medications (e.g., immunosuppressants for organ transplant or rheumatoid arthritis) Other Exclusions. Once clinically unrelated services are excluded at the claim line level, we exclude episodes based on several other characteristics, such as: 1) Any episode that is triggered by a PAC claim outside the 50 states, D.C., Puerto Rico, and U.S. Territories. Rationale: This exclusion ensures that complete claims data are available for each provider. 2) Any episode where the claim(s) constituting the attributed PAC provider's treatment have a standard allowed amount of zero or where the standard allowed amount cannot be calculated. Rationale: Episodes where the claim(s) constituting the attributed PAC provider's treatment are zero or have unknown allowed payment do not reflect the cost to Medicare. Including these episodes in the calculation of MSPB-PAC LTCH measure could potentially misrepresent a providers' resource use. 3) Any episode in which a patient is not enrolled in Medicare FFS for the entirety of a 90-day lookback period (i.e., a 90-day period prior to the episode trigger) plus episode window (including where a beneficiary dies) or is enrolled in Part C for any part of the lookback period plus episode window. Rationale: Episodes meeting this criteria do not have complete claims information that is needed for risk-adjustment and the measure calculation as there may be other claims (e.g., for	Medicare Part A, episodes in which the beneficiary becomes deceased). Episodes in which the beneficiary is enrolled only in Medicare Part A, for example, are excluded because these beneficiaries may receive services not observed in the data. Similarly, episodes in which the beneficiary dies at any point during the episode. Episodes in which the patient dies are—by definition—truncated episodes and do not have a complete episode window. Episodes in which the patient dies were identified as an index hospitalization with death discharge code (STUS_CD "20" "41") or if a beneficiary's death was within an MSPB-Hospital episode. Including episodes without all observable claims or a complete episode window could potentially make hospitals seem efficient not due to any action of their own, but because the data are missing services that would be included in the MSPB-Hospital measure calculation. Episodes where Medicare is the secondary payer: if a beneficiary was the primary payer any time during the MSPB-Hospital episode, the beneficiary was excluded (i.e., if bene_prmry_pyr_entlmt_strt_dt (start date of primary payer enrollment) bene_prmry_pyr_entlmt_end_dt (end date of primary payer enrollment) fell within the episode). • [II] Regarding beneficiaries whose primary insurance becomes Medicaid during an episode due to exhaustion of Medicare Part A benefits, these beneficiaries are not excluded. Rather, Medicaid payments made for services rendered to these beneficiaries are excluded; all Medicare Part A payments made before benefits are exhausted and all Medicare Part B payments made during the episode are included. The MSPB-Hospital measure is calculated using only Medicare Part A and Part B claims; as a result no Medicaid claims are included in the MSPB-Hospital measure calculation. • [III] Any episode in which the index admission inpatient claim has a $0 actual payment or a $0 standardized payment is excluded. $0 inpatient admissions may represent errors in the data, or payment corrections rather than actual services rendered. Only when the Claim Payment amount (pmt_amt) for the IP stay is greater than 0 OR standard_allowed_amt is greater than 0 is the amount included in the MSPB-Hospital measure calculation.

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	services provided under Medicare Advantage [Part C]) that we do not observe in the Medicare Part A and B claims data. Similarly, episodes in which the patient dies are, by definition, truncated episodes and do not have a complete episode window. Including these episodes in the MSPB-PAC LTCH measure could potentially misrepresent a provider's resource use. This exclusion also allows us to faithfully construct Hierarchical Condition Categories (HCCs) for each episode by scanning the lookback period prior to its start without missing claims. 4) Any episode in which a patient has a primary payer other than Medicare for any part of the 90-day lookback period plus episode window. Rationale: When a patient has a primary payer other than Medicare, complete claims data may not be observable. These episodes are removed to ensure that the measures are accurately calculated using complete data. 5) Any episode where the claim(s) constituting the attributed PAC provider's treatment include at least one related condition code indicating that it is not a prospective payment system bill. Rationale: Claims that are not a prospective payment system bill may not report sufficient information to allow for payment standardization. 6) Any episode with problematic claims data (e.g., anomalous records for stays that overlap wholly or in part, or are otherwise erroneous or contradictory) Rationale: The episode with the most recent processing date is kept to ensure the accuracy of data elements. Finally, as part of the measure construction process described in section S.7.2, episodes with residuals below the 1st or above the 99th percentile of the residual distribution are excluded, reducing the impact of high- and low-payment outliers. Notes: [1] The lists of clinically unrelated services built off the planned readmissions algorithm developed by the Yale New Haven	• [IV] Due to the uncertainty surrounding attributing episodes to hospitals in cases where the patient was transferred between acute hospitals during the index admission, acute-to-acute transfers during the index admission (where a transfer is defined based on the claim discharge code) are not considered index admissions for the purposes of the MSPB-Hospital measure. In other words, these cases will not generate new MSPB-Hospital episodes; neither the hospital which transfers a patient to another short-term acute hospital, nor the receiving short-term acute hospital will have an index admission attributed to them. This exclusion avoids assigning responsibility to an MSPB-Hospital episode in a case where multiple hospitals treat the patient during the index admission. • [V] Cancer hospitals, MD Hospitals (provider variable starting with "21"), emergency hospitals (provider variable last position "E" OR "F"), and veteran's hospital (provider variable position "V") are also excluded. • [VI] In response to stakeholder comments, the FY 2012 IPPS Final Rule states that the MSPB-Hospital measure will "exclude statistical outliers from the calculation" (76 FR 51626: www.gpo.gov/fdsys/pkg/FR-2011-08-18/pdf/2011-19719.pdf). To mitigate the effect of high-cost outliers on each hospital's MSPB-Hospital measure score, MSPB-Hospital episodes whose relative scores fall above the 99th percentile or below the 1st percentile of the distribution of residuals are excluded from the MSPB-Hospital calculation. Excluding outliers based on residuals eliminates the episodes that deviate most from their predicted values in absolute terms.

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	Health Services Corporation/Center for Outcomes Research & Evaluation, as well as the expansions to the Yale algorithm by RTI. Clinicians reviewed the list of exclusions from that algorithm in the context of PAC treatment. During the review process, clinicians reviewed admissions observed in MSPB-PAC episodes and created exclusions that overlap with the Yale algorithm. Details on the Yale and RTI algorithms are available here: "Hospital-Wide All-Cause Unplanned Readmission Measure - Version 4.0," in 2015 Measure Updates and Specifications Report, ed. Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2015). 10-11. Laura Smith, West, S., Coots, L., Ingber, M., "Skilled Nursing Facility Readmission Measure (SNFRM) NQF #2510: All-Cause Risk-Standardized Readmission Measure," (Centers for Medicare & Medicaid Services, 2015). 5-6	
Risk Adjustment	Statistical risk model The detailed steps to computing the measure score are described in section S.7.2. Risk-adjustment is applied in "Step 3: Calculate Predicted Episode Payments." The purpose of risk adjustment is to compensate for patient health circumstances and demographic factors that affect resource use but are beyond the influence of the attributed provider. The MSPB-PAC LTCH measure risk adjustment model is adapted from the model used in the NQF-endorsed MSPB-Hospital measure, which itself is an adaptation of the standard CMS-HCC risk-adjustment model.[1,2] The MSPB-PAC LTCH model uses a linear regression framework and a 90-day HCC lookback period. The risk adjustment model is estimated on all MSPB-PAC LTCH episodes that meet the exclusion criteria. The model is estimated separately for Standard and Site Neutral episodes (see section S.7.2 for description of episode types). LTCH episodes are only compared to episodes of the same type (i.e., Standard episodes are only compared to Standard episodes, and Site Neutral episodes to Site Neutral episodes). This ensures that comparisons are fair, meaningful, and reflective of payment policy differences within particular LTCH settings.	Statistical risk model

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	Each provider's MSPB-PAC LTCH measure score is calculated as a provider's average MSPB-PAC Amount divided by the median MSPB-PAC Amount across all providers. A provider's MSPB-PAC LTCH Amount is defined as the sum of standardized, risk-adjusted spending across all of a provider's eligible episodes divided by the number of episodes for that provider. Below is a description of the risk adjustment variables. Risk-Adjustment Variables The following beneficiary health status indicators are included as covariates in each MSPB-PAC LTCH risk adjustment model and to the greatest extent possible are consistent across PAC settings (see Appendix C of the Measure Specifications document provided in section S.1 for a comprehensive list of independent variables used in the risk adjustment model): • 70 HCCs • 11 HCC interactions • 11 brackets for age at the start of the episode • Original entitlement to Medicare through disability • ESRD status • Long-term care institutionalization at start of episode.[3] • Six clinical case-mix categories reflecting recent prior care (described further below).[4] • Hospice utilization during the episode • Prior acute ICU utilization day categories • Prior acute length of stay categories • Medicare Severity-Long-Term Care Diagnosis-Related Groups (MS-LTC-DRGs) The clinical case-mix category variables used in the MSPB-PAC LTCH risk adjustment model are included to account for differences in intensity and type of care received by beneficiaries prior to the start of an MSPB-PAC LTCH episode. See section S.7.5 for more details on the methodology of assigning clinical case-mix categories to each episode. Notes:	

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	[1] QualityNet, "Measure Methodology Reports: Medicare Spending Per Beneficiary (MSPB) Measure," (2015). http://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier4&cid=1228772057350 [2] CMS, "Medicare Risk Adjustment Information" (2016) https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html [3] Identifies beneficiaries who have been institutionalized for at least 90 days in a given year. The indicator is based on 90-day assessments from the Minimum Data Set (MDS) and is calculated based on CMS' definition of institutionalized individuals. [4] There are 7 case-mix categories as described above, but one category is removed to prevent collinearity. Statistical risk model	
Stratification	The MSPB-PAC LTCH measure is stratified by standard and site neutral payment rate admissions. An MSPB-PAC LTCH Standard episode is triggered by a standard payment rate claim, while an MSPB-PAC LTCH Site Neutral episode is triggered by a site neutral payment rate claim. Risk adjustment is then performed separately for MSPB-PAC LTCH Standard and Site Neutral cases. Thus, LTCH Standard and Site Neutral episodes are compared only with LTCH Standard and Site Neutral episodes, respectively, to ensure that the measure is making fair comparisons between clinically similar beneficiaries.	While the measure results are not stratified, expected costs for episodes are determined by using a separate risk adjustment model for episodes within each MDC. MDCs are aggregations of Diagnosis Related Groups (MS-DRG), which CMS uses to classify acute inpatient admissions. The MS-DRG/MDC crosswalk is available for order here: http://solutions9.3m.com/wps/portal/!ut/p/c1/04_SB8K8xLLM9MSSzPy8xBz94NS8-NBg_Qj9KLP4IC8Py1BTI2MD9zAvFwMjYzMzCxNHd2OTACP9ggxHRQBm3gTM/
Type Score	Ratio An MSPB-PAC LTCH measure score of 1 indicates that an LTCH had an average MSPB-PAC Amount (i.e., risk-adjusted spending level) which is equal to the national episode-weighted median MSPB-PAC Amount across all LTCH facilities during a given performance period. An MSPB-PAC LTCH measure score of greater than 1 indicates that an LTCH had higher average risk-adjusted spending levels compared to those of the national median LTCH. For example, a measure score of 1.1 indicates that the LTCH had average risk-adjusted spending levels that are 10 percent higher than the median LTCH. On the other hand, an MSPB-PAC LTCH measure score of less than 1 indicates that an	Ratio; Attachment An MSPB-Hospital measure that is less than 1 indicates that a given hospital's MSPB-Hospital amount (i.e. risk-adjusted spending) is less than the national episode-weighted median MSPB-Hospital amount across all hospitals during a given performance period. We note that results of the MSPB-Hospital measure alone do not necessarily reflect the quality of care provided by hospitals. Accordingly, lower MSPB-Hospital measure across performance periods (i.e., lower Medicare spending per beneficiary) in isolation should not be interpreted as better care. The MSPB-Hospital measure is most meaningful when presented in the context of other quality measures, which are part of the Hospital Value-Based Purchasing (VBP)

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	LTCH had lower average risk-adjusted spending levels compared to those of the median LTCH. For example, a measure score of 0.9 indicates that the LTCH had average risk-adjusted spending levels that are 10 percent lower than the median LTCH.	Program. As part of the Hospital VBP Program, the MSPB-Hospital measure is aligned with current quality of care measures to facilitate profiling hospital value (payments and quality). Improvement on this measure for a hospital would be observed as a lower MSPB-Hospital measure value across performance periods.
Algorithm	Grouping methodology: The grouping methodology includes all Medicare Part A and B services delivered to a beneficiary during the treatment period (from admission to the LTCH through to discharge from the LTCH) and associated services period (from admission to the LTCH through to 30 days after discharge from the LTCH). To simplify the clinical logic and avoid the issue of attributing claims to MSPB episodes in the case of concurrent clinical events, all claims that begin within the episode window (treatment period and associated services period) are included in the MSPB-PAC LTCH measure. In order to create a resource use measure that is clinically valid, there were multiple steps involved in excluding the least clinically relevant codes. Using an episode window, we organized claims into clinically meaningful service categories or settings. For example, Medicare Severity-Diagnosis Related Groups (MS-DRGs) noted after an LTCH discharge were evaluated as medical or surgical admissions post-discharge. Clinical Classifications Software (CCS) and Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) services were organized into outpatient services, emergency department (ER) services, and durable medical equipment claims and evaluated for their relevance or relatedness to LTCH care. Extensive clinical review was performed by clinicians with experience and expertise in LTCH, as well as in collaboration with Medical Officers at CMS. The inpatient, outpatient, Part B physician and supplier, and DMEPOS services least clinically related to the LTCH care were excluded from the measure. For instance, services related to the routine management of	Objective: The MSPB-Hospital measure aims to improve care coordination in the period between 3 days prior to an acute inpatient hospital admission through the period 30 days after discharge. The MSPB-Hospital measure recognizes lower costs associated with a reduction in unnecessary services, preventable complications, readmissions, and shifting post-acute care from more expensive to less expensive services when appropriate. Grouping methodology: The MSPB-Hospital measure evaluates resource use through the unit of MSPB-Hospital episodes. The MSPB-Hospital episodes are constructed by including all Medicare Part A and Part B claims with a start date falling between 3 days prior to an acute inpatient hospital admission through the period 30 days after discharge. Any episodes where at any time during the episode the beneficiary is enrolled in a Medicare Advantage plan, the beneficiary becomes deceased, or Medicare is the secondary payer will be excluded from the MSPB-Hospital calculation. Regarding beneficiaries whose primary insurance becomes Medicaid during an episode due to exhaustion of Medicare Part A benefits, Medicaid payments made for services rendered to these beneficiaries are excluded; however, all Medicare Part A payments made before benefits are exhausted and all Medicare Part B payments made during the episode are included. Cost Calculation: The MSPB-Hospital amount includes the cost of services performed by hospitals and other healthcare providers during an MSPB-Hospital episode, which is comprised of the period 3 days prior to an inpatient PPS hospital admission (index admission) through 30 days post-hospital discharge. All costs are payment standardized to

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	preexisting chronic conditions (e.g., dialysis for ESRD, treatment for preexisting cancers, and treatment for organ transplants) were felt to be clinically unrelated to the scope of the type of care that LTCHs provide. Therefore, these types of services were excluded. Services were excluded if there was consensus across clinicians from the measure developer, external clinical experts including TEP members, and CMS medical officers. Please see section S.9.1 for overall clinical consensus regarding the types of exclusions. Attribution algorithm: An MSPB-PAC LTCH episode is assigned to the facility of the index admission. A new episode may begin during the associated services period of a previous MSPB-PAC LTCH episode in the 30 days post-discharge from the LTCH.	control for geographic variation in Medicare reimbursement rates. All costs are risk adjusted to account for age and severity of illness.
Submission items	5.1 Identified measures: 2158 : Medicare Spending Per Beneficiary (MSPB) - Hospital 5a.1 Are specs completely harmonized? Yes 5a.2 If not completely harmonized, identify difference, rationale, impact: 2158 : Medicare Spending Per Beneficiary (MSPB) - Hospital 5b.1 If competing, why superior or rationale for additive value: Not applicable. There are currently no measures that address both the same measure focus AND the same target population. MSPB-PAC measures are harmonized across PAC settings as well as with MSPB-Hospital. MSPB-PAC measures were developed in parallel for all PAC settings to meet the mandate of the IMPACT Act. To align with the goals of standardized assessment across PAC settings, these measures were conceptualized uniformly across the four settings in terms of the construction logic, the approach to risk adjustment, and measure calculation. The measures mirror the general construction of MSPB-Hospital. Aligning the MSPB-Hospital and MSPB-PAC measures in this way	5.1 Identified measures: 5a.1 Are specs completely harmonized? 5a.2 If not completely harmonized, identify difference, rationale, impact: 5b.1 If competing, why superior or rationale for additive value: The MSPB-Hospital measure evaluates hospitals' efficiency relative to the efficiency of the median hospital. The target population is Medicare beneficiaries enrolled in Medicare Parts A and B who were discharged from short-term acute hospitals. There are currently no NQF-endorsed measures that address both this same measure focus AND this same target population.

	3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	2158: Medicare Spending Per Beneficiary (MSPB) - Hospital
	creates continuous accountability and aligns incentives to improve care planning and coordination across inpatient and PAC settings.	

Comparison of NQF #3575 and NQF #1604

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
Steward	Centers for Medicare & Medicaid Services	HealthPartners
Description	The Total Per Capita Cost (TPCC) measure assesses the overall cost of care delivered to a beneficiary with a focus on the primary care they receive from their provider(s). The TPCC measure score is a clinician's average risk-adjusted and specialty-adjusted cost across all beneficiary months attributed to the clinician during a one year performance period. The measure is attributed to clinicians providing primary care management for the beneficiary, who are identified by their unique Taxpayer Identification Number and National Provider Identifier pair (TIN-NPI) and clinician groups, identified by their TIN number. Clinicians are attributed beneficiaries for one year, beginning from a combination of services indicate that a primary care relationship has begun. The resulting periods of attribution are then measured on a monthly level, assessing all Part A and Part B cost for the beneficiary for those months that occur during the performance period. The beneficiary populations eligible for the TPCC include Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period.	Total Cost of Care reflects a mix of complicated factors such as patient illness burden, service utilization and negotiated prices. Total Cost Index (TCI) is a measure of a primary care provider's risk adjusted cost effectiveness at managing the population they care for. TCI includes all costs associated with treating members including professional, facility inpatient and outpatient, pharmacy, lab, radiology, ancillary and behavioral health services. A Total Cost Index when viewed together with the Total Resource Use measure (NQF-endorsed #1598) provides a more complete picture of population based drivers of health care costs.
Type	Cost/Resource Use	Cost/Resource Use
Data Source	Assessment Data, Claims, Enrollment Data, Other Medicare Part A and Part B claims data: TPCC uses Part A and B claims data to attribute beneficiaries to clinicians, calculate beneficiary's costs, and construct risk adjustors. CMS Office of Information Systems	Claims Use administrative claims data base Risk Adjustment Tool, Johns Hopkins ACG System Use administrative claims data base

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	(OIS) maintains a detailed Medicare Claims Processing Manual available at the following URL: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Internet-Only-Manuals-IOMs-Items/CMS018912. Medicare Enrollment Database (EDB): This is used to determine beneficiary-level exclusions and supplemental risk adjustors, specifically Medicare Parts A, B, and C enrollment; other primary payers; disability status; sex; end-stage renal disease (ESRD); beneficiary birth dates; and beneficiary death dates. Common Medicare Environment (CME) database: This is used to determine beneficiary's dual status. https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. Minimum Data Set (MDS): The MDS is used to identify beneficiaries that should be risk adjusted through the CMS-HCC v22 institutional model. https://www.resdac.org/cms-data/files/mds-3.0. For measure testing purposes, data from the American Census, American Community Survey (ACS) is used in the analyses evaluating patient cohorts and social risk factors in risk adjustment. https://www.census.gov/programs-surveys/acs/technical-documentation/summary-file-documentation.html. Medicare Part A and Part B claims data: TPCC uses Part A and B claims data to attribute beneficiaries to clinicians, calculate beneficiary's costs, and construct risk adjustors. CMS Office of Information Systems (OIS) maintains a detailed Medicare Claims Processing Manual available at the following URL: https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Internet-Only-Manuals-IOMs-Items/CMS018912. Medicare Enrollment Database (EDB): This is used to determine beneficiary-level exclusions and supplemental risk adjustors, specifically Medicare Parts A, B, and C enrollment; other primary payers; disability status; sex; end-stage renal disease (ESRD); beneficiary birth dates; and beneficiary death dates.	Risk Adjustment Tool, Johns Hopkins ACG System

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	Common Medicare Environment (CME) database: This is used to determine beneficiary's dual status. https://www.ccwdata.org/documents/10280/19002256/medicare-enrollment-impact-of-conversion-from-edb-to-cme.pdf. Minimum Data Set (MDS): The MDS is used to identify beneficiaries that should be risk adjusted through the CMS-HCC v22 institutional model. https://www.resdac.org/cms-data/files/mds-3.0. For measure testing purposes, data from the American Census, American Community Survey (ACS) is used in the analyses evaluating patient cohorts and social risk factors in risk adjustment. https://www.census.gov/programs-surveys/acs/technical-documentation/summary-file-documentation.html. Data dictionary URL; Code table attachment Data dictionary URL; Code table attachment	
Level Setting	Clinician : Group/Practice, Clinician : Individual No Applicable Care Setting	Population : Community, County or City, Clinician : Group/Practice Emergency Department and Services, Home Care, Inpatient/Hospital, Other, Outpatient Services, Post-Acute Care All care settings
Numerator Statement	N/A	Numerator: Total PMPM = (Total Medical Cost / Medical Member Months) + (Total Pharmacy Cost / Pharmacy Member Months)
Numerator Details Denominator Statement	N/A N/A	N/A Denominator: Average Risk Score - the medical claims data is submitted through the Johns Hopkins ACG Risk Grouper which generates a relative risk score for each member. That risk score is then multiplied by the number of months a member has been enrolled creating a risk weight. The risk weights are then summed to the desired level of measurement (e.g., provider group) and divided by the total sum of the desired level's member months creating a member month weighted Average Risk Score.
Denominator Details Exclusions	N/A Several steps in the construction of the TPCC measure ensure comparability by fostering comparability in the beneficiary	ACG Adjusted PMPM = Total PMPM / ACG Risk Score TCI = Provider ACG Adjusted PMPM / Peer Group ACG Adjusted PMPM

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
Exclusion Details	population captured and clinician population measured. These are detailed in Section S.7.2. In keeping with the measure intent to capture the overall costs of care for beneficiaries receiving primary care services, there are a limited set of exclusions primarily to ensure that, as part of data processing, sufficient data are available to accurately determine resource use and calculate risk adjustment for each beneficiary. These exclusions, along with their rationales, are listed below. •The beneficiary was not continuously enrolled in Medicare Parts A and B unless partial enrollment was the result of either new enrollment or death only. These beneficiaries may have gaps in their Medicare claim records when benefits are covered by other payers. •The beneficiary resides outside the United States or its territories during the performance period. Differences in reimbursement policy for healthcare services provided outside the U.S. can lead to unfair comparisons of cost. •The beneficiary receives benefits from the Railroad Retirement Board (RRB). Beneficiaries covered by the RRB may have healthcare benefits normally covered by Medicare paid by the RRB, which may bias the observed cost for these beneficiaries. To ensure the clinicians attributed the measure are within the intended scope of primary care management, exclusions of clinicians are used to ensure comparability. Clinicians who would not reasonably be responsible for providing primary care are excluded from attribution of the revised TPCC measure using their CMS HCFA specialty designation assigned on Part B physician/supplier claims. This exclusion aims to keep primary care specialists and internal medicine subspecialists who frequently manage patients with chronic conditions falling in their areas of specialty. Additionally, clinicians are characterized by their Part B billing behavior and excluded from attribution if found meeting a threshold of billing for the following service categories; 10-day or 90-day global surgery services, anesthesia services, therapeutic radiation services, chemotherapy services. The methodology and clinical logic for exclusions of clinicians from attribution is further detailed in Section S.8.2	

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	Data truncation is applied to risk-adjusted beneficiary monthly costs for outlier values through winsorization on the right tail. Monthly costs at the 99th percentile are assigned to all attributable beneficiary months with costs above the 99th percentile. Winsorization aims to limit the effects of extreme values on expected costs. Winsorization is a statistical transformation that limits extreme values in data to reduce the effect of possible outliers.	
Risk Adjustment	Stratification by risk category/subgroup As described in Section S.7.2, the TPCC measure is calculated for each clinician and clinician group practice by averaging the risk-adjusted and specialty-adjusted cost across the beneficiary months attributed. Adjustments to observed monthly costs are calculated as follows: 1) Divide observed costs for each beneficiary month by the normalized risk score to obtain risk-adjusted monthly costs. 2) Winsorize risk-adjusted monthly costs at the 99th percentile by assigning the 99th percentile of monthly costs to all attributable beneficiary months with costs above the 99th percentile. 3) Normalize monthly costs to account for differences in expected costs based on the number of clinician groups to which a beneficiary is attributed in a given month. The normalization factor is the inverse cube root of the number of attributed clinician groups for that beneficiary month. 4) Calculate the average risk-adjusted monthly cost for each TIN and TIN-NPI by averaging risk-adjusted monthly cost across all attributed beneficiary months. 5) Calculate the national specialty-specific expected cost for each specialty as the weighted average of TIN/TIN-NPI's risk-adjusted monthly cost. 5a) Define the weight for each TIN/TIN-NPI as the percentage of clinicians with that specialty multiplied by the total number of beneficiary months attributed to the TIN/TIN-NPI multiplied by the number of clinicians with that specialty. 6) Calculate the specialty-adjustment factor for each TIN or TIN-NPI as follows:	Statistical risk model

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	6a) Multiply the national specialty-specific expected cost for each specialty by the respective specialty's share of Part B payment within a TIN or TIN-NPI and sum the weighted share of national specialty-specific expected cost calculated in the previous step across all the specialties under a given TIN or TIN-NPI. 7) Calculate final risk-adjusted, specialty-adjusted cost measure by dividing each TIN and TIN-NPI's average risk-adjusted monthly cost by their specialty-adjustment factor and multiply this ratio by the average non-risk-adjusted, winsorized observed cost across the total population of attributed beneficiary months.	
Stratification	Stratification by risk category/subgroup Differences in patient case mix are accounted for by using separate risk adjustment models for the following types of beneficiaries, as discussed in Section S.7.2: 1) Beneficiaries without ESRD 1a) Beneficiaries with fewer than 12 months of Medicare medical history 2a) Beneficiaries with at least 12 months of Medicare medical history 3a) Beneficiaries in long-term institutional care settings 2) Beneficiaries with ESRD receiving dialysis 2a) Beneficiaries with fewer than 12 months of Medicare medical history 2b) Beneficiaries with at least 12 months of Medicare medical history This stratification accounts for the very different patient clinical profiles for patients with ESRD receiving dialysis and patients without ESRD, as well as maximizes the availability of Medicare claims history to be able to construct indicator variables for clinical conditions. The TPCC measure uses the CMS-HCC V22 risk adjustment models for new enrollee, community, and long-term institutional beneficiaries without ESRD. A beneficiary month is measured under the new enrollee model if they do not have a full one-year lookback of Medicare claims data as of the start of a beneficiary month. As a result, the model is derived primarily from	Measures are adjusted for clinical risk and limited to the commercial population.

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	beneficiary enrollment data. This model adjusts for gender, age, dual Medicare and Medicaid enrollment, and whether the beneficiary was originally entitled to Medicare due to disability through a series of interacted covariates. Beneficiaries with sufficient Medicare claims history are measured under the community or the institutional model if they are institutionalized in a long term care facility. In both models, severity of illness is measured using HCCs and disease interactions. 79 HCCs are accounted for under CMS-HCC V22 model for beneficiaries classified as community enrollees and long-term institutional enrollees while the exact number and types of disease interaction can vary. Both models interact beneficiary age with gender. In addition, the community model interacts dual enrollment status, gender, and the indicator for whether the beneficiary was originally entitled to Medicare due to disability, while the institutional model adjusts for disability as the original reason for Medicare enrollment and dual enrollment status independently. For ESRD beneficiaries receiving dialysis, the TPCC measure utilizes the CMS-ESRD V21 risk adjustment models. Differentiated models are implemented for dialysis new enrollees and dialysis community enrollees. Similar to the CMS-HCC V22, enrollees are classified as new enrollees if they were not continuously enrolled in Parts A and B for the one-year lookback period prior to each beneficiary month. As a result of this, the model primarily uses information from the beneficiary's enrollment data. This model adjusts for gender, age, dual enrollment status, and whether the beneficiary was originally entitled to Medicare due to disability through a series of interacted covariates. In addition to accounting for these patient characteristics, the dialysis community model also risk adjusts for medical severity using 87 HCCs and additional disease interactions. The CMS-ESRD V21 and CMS-HCC V22 models both generate a risk score for each beneficiary that summarizes the beneficiary's expected cost of care relative to other beneficiaries. Risk scores for ESRD beneficiaries are normalized to enable comparison with the HCC V22 risk scores. This is achieved by multiplying ESRD risk scores by the mean annual Medicare spending for the ESRD	

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	population applied in the CMS-ESRD V21 model and dividing by the mean annual Medicare spending for the total Medicare population applied in the CMS-HCC V22 model, effectively renormalizing ESRD risk score values to the equivalent scale of the HCC models. A risk score equal to one indicates risk associated with expenditures for the average beneficiary nationwide. Risk scores below or above one indicate below and above average risk, respectively. The complete list of risk adjustment variables for each model are listed in the Measure Codes List linked in Section S.1 in the tab titled HCC_Risk_Adjust.	
Type Score	Continuous variable The TPCC measure score is the average payment-standardized, risk-adjusted, and specialty-adjusted monthly cost across all beneficiary months in the performance period attributed to a clinician or clinician group. A lower measure score indicates that the observed episode costs are lower than or similar to expected costs for the care provided for the particular patients included in the calculation. A higher measure score indicates that the observed episode costs are higher than expected for the care provided for the particular patients included in the calculation.	Ratio; Other (specify): https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/dev_057910.pdf see page 9 A provider Total Cost Index (TCI) of 1.10 equates to 10% higher paid risk adjusted PMPM. Similarly, a provider TCI score of 0.90 equates to 10% less paid risk adjusted PMPM. A score of 1.0 is equivalent to the peer group average.
Algorithm	As described in Section S.7.2, to account for the clinical severity of patients, one of five separate risk adjustment models are applied based on the patients characteristics observed in the year prior to the beneficiary month being measured. For non-ESRD patients, the three models are the new enrollee model, community model, and institutional model from CMS' Hierarchical Condition Category Version 22 (CMS-HCC V22). For ESRD patients, the two models are the dialysis new enrollee model and dialysis community model from CMS' ESRD Version 21 (CMS-ESRD V21). Each model includes beneficiary demographic and enrollment information such as age, gender, disability, and dual enrollment status. Both the new enrollee model and dialysis new enrollee models are limited to these factors as the patient does not have sufficient Medicare claims history for further evaluation. The remaining models (community model, institutional model, and dialysis community) include either 79 (CMS-HCC V22) or 87 (CMS-ESRD V21) hierarchical condition	Not applicable. This is a population-based measure that applies to all care settings and conditions.

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	categories to characterize the patient severity and comorbidities. The indicators used for risk adjustment and the methodology are detailed in the Measure Information Form linked in Section S.1. The start of a primary care relationship between a clinician and beneficiary is identified by the occurrence of two Part B Physician/Supplier (Carrier) claims with particular CPT/HCPCS services billed in close proximity. There are two different sets of CPT/HCPCS codes used: E&M primary care services and primary care services. E&M primary care services are a specific set of evaluation and management codes for physician visits in the outpatient setting, physician office, nursing facility, or assisted living. Primary care services are a broader list of services related to routine primary care that generally fall into the following categories: Durable Medical Equipment (DME) and Supplies, Electrocardiogram, Laboratory - Chemistry and Hematology, Other Diagnostic Procedures (Interview, Evaluation, Consultation), Other Diagnostic Radiology and Related Techniques, Prophylactic Vaccinations and Inoculations, Routine Chest X-ray, Clinical Labs, and Preventive Services The codes used to attribute beneficiaries to clinicians are listed in the tabs titled E&M_Prim_Care and Prim_Care_Services within the Measure Codes List linked in Section S.1. Clinicians who would not reasonably be responsible for providing primary care are excluded from attribution of the revised TPCC measure using their CMS HCFA specialty designation assigned on Part B physician/supplier claims. This exclusion aims to keep primary care specialists and internal medicine subspecialists who frequently manage patients with chronic conditions falling in their areas of specialty. The excluded specialties list contains 56 specialties that fall into the following broad categories: •Surgical sub-specialties •Non-physicians without chronic management of significant medical conditions	

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	•Internal medicine sub-specialties with additional highly procedural sub-specialization •Internal medicine specialties that practice primarily inpatient care without chronic care management •Pediatricians who do not typically practice adult medicine The codes used to exclude clinicians from attribution base on their CMS HCFA specialty are listed in in the tab titled Eligible_Clinicians within the Measure Codes List linked in Section S.1. Additionally, TIN-NPI are removed from attribution if a clinician met any of the following four service category thresholds for the same beneficiary by billing the specified CPT/HCPCS within +/-180 days of the candidate event on Part B physician/supplier claims: •At least 15 percent of the clinician's attributable events are comprised of 10-day or 90-day global surgery services. •At least 5 percent of the clinician's attributable events are comprised of anesthesia services. •At least 5 percent of the clinician's attributable events are comprised of therapeutic radiation services. •At least 10 percent of the clinician's attributable events are comprised of chemotherapy services.	
Submission items	5.1 Identified measures: 1604 : Total Cost of Care Population-based PMPM Index 5a.1 Are specs completely harmonized? No 5a.2 If not completely harmonized, identify difference, rationale, impact: 1604 : Total Cost of Care Population-based PMPM Index 5b.1 If competing, why superior or rationale for additive value: N/A.	5.1 Identified measures: 5a.1 Are specs completely harmonized? 5a.2 If not completely harmonized, identify difference, rationale, impact: 5b.1 If competing, why superior or rationale for additive value: Not applicable

	3575: Total Per Capita Cost (TPCC)	1604: Total Cost of Care Population-based PMPM Index
	There are no competing NQF-endorsed or non-NQF-endorsed cost measures that address the same measure focus and target population.	

Appendix F: Pre-Evaluation Comments

Comments received as of June 26, 2020.

Topic	Commenter	Comment
3561: Medicare Spending Per Beneficiary – Post Acute Care Measure for Inpatient Rehabilitation Facilities	Submitted by Federation of American Hospitals.	The Federation of American Hospitals (FAH) appreciates the opportunity to comment on this measure prior to the Standing Committee's evaluation. The FAH requests that the committee carefully consider whether the measure as specified produces performance scores that are reliable and valid for facility-level reporting. Specifically, the FAH is concerned to see that reliability at the 25th percentile for 20-190 episodes was 0.64, which leads us to question what result was produced at the minimum level. We believe that the results currently provided indicate that the measure as specified may not produce scores that yield acceptable minimum thresholds for reliability. The scientific acceptability of the measure is further called into question on review of the risk model's fit with the overall adjusted R-squared as 0.1595. While the developer provides some explanation on why the result is low, the FAH does not believe that the reasons for this result are adequately addressed and risk adjustment must be improved prior to endorsement. In addition, while the FAH appreciates that social risk factors were reviewed, we believe that the risk adjustment approach should not consider the identification and testing of social risk factors as supplementary to clinical risk factors. This approach was identified as a concern by the NQF Disparities Standing Committee and developers must begin to include these factors within the testing of the model rather than the approach of "adding on" factors after the model is developed. This type of analysis would assist facilities and others in understanding how their inclusion could impact the model and provide additional information for groups examining this issue such as the NQF and Office of the Assistant Secretary for Planning and Evaluation. As a result, the FAH believes that this measure lacks sufficient information on the potential impact these social risk variables have on the risk adjustment model. Furthermore, while the developer provides information on the changes in performance scores that result from social risk factor adjustment (tables 2b3.4b_4a and 2b3.4b_4b), it is not necessarily clear on the degree to which these changes would result in a facility's score being statistically significantly different from the national average. If the interpretation of the results under meaningful differences leads to a "conclusion that even small difference between facility scores can be treated as meaningful"(response to 2b4.3 in the testing form), to what extent would changes in performance as a result of adjustment

Topic	Commenter	Comment
		for social risk factors also lead to different but meaningful results?
3562: Medicare Spending Per Beneficiary – Post Acute Care Measure for Long-Term Care Hospitals	Submitted by Federation of American Hospitals.	The Federation of American Hospitals (FAH) appreciates the opportunity to comment on this measure prior to the Standing Committee's evaluation. The FAH requests that the committee carefully consider whether the measure as specified produces performance scores that are reliable and valid for facility-level reporting. Specifically, the FAH is concerned to see that reliability at the 25th percentile for 22-197 episodes was 0.70, which leads us to question what result was produced at the minimum level. We believe that the results currently provided indicate that the measure as specified may not produce scores that yield acceptable minimum thresholds for reliability. The scientific acceptability of the measure is further called into question on review of the risk model's fit with the overall adjusted R-squared as 0.4894. While the developer provides some explanation on why the result is low, the FAH does not believe that the reasons for this result are adequately addressed and risk adjustment must be improved prior to endorsement. The FAH appreciates that social risk factors were reviewed, and believes that the risk adjustment approach should not consider the identification and testing of social risk factors as supplementary to clinical risk factors. This approach was identified as a concern by the NQF Disparities Standing Committee and developers must begin to include these factors within the testing of the model rather than the approach of "adding on" factors after the model is developed. This type of analysis would assist facilities and others in understanding how their inclusion could impact the model and provide additional information for groups examining this issue such as the NQF and Office of the Assistant Secretary for Planning and Evaluation. As a result, the FAH believes that this measure lacks sufficient information on the potential impact these social risk variables have on the risk adjustment model. Furthermore, while the developer provides information on the changes in performance scores that result from social risk factor adjustment (tables 2b3.4b_4a and 2b3.4b_4b), it is not necessarily clear on the degree to which these changes would result in a facility's score being statistically significantly different from the national average. If the interpretation of the results under meaningful differences leads to a "conclusion that even small difference between facility scores can be treated as meaningful"(response to 2b4.3 in the testing form), to what extent would changes in performance as a result of adjustment for social risk factors also lead to different but meaningful results?

Topic	Commenter	Comment
3563: Medicare Spending Per Beneficiary – Post Acute Care Measure for Skilled Nursing Facilities	Submitted by Federation of American Hospitals.	The Federation of American Hospitals (FAH) appreciates the opportunity to comment on this measure prior to the Standing Committee's evaluation. The FAH requests that the committee carefully consider whether the measure as specified produces performance scores that are reliable and valid for facility-level reporting. Specifically, the FAH is concerned to see that reliability at the 25th percentile for 20-75 episodes was 0.79, which leads us to question what result was produced at the minimum level. We believe that the results currently provided indicate that the measure as specified may not produce scores that yield acceptable minimum thresholds for reliability. The scientific acceptability of the measure is further called into question on review of the risk model's fit with the overall adjusted R-squared as 0.1157. While the developer provides some explanation on why the result is low, the FAH does not believe that the reasons for this result are adequately addressed and risk adjustment must be improved prior to endorsement. The FAH appreciates that social risk factors were reviewed, and believes that the risk adjustment approach should not consider the identification and testing of social risk factors as supplementary to clinical risk factors. This approach was identified as a concern by the NQF Disparities Standing Committee and developers must begin to include these factors within the testing of the model rather than the approach of "adding on" factors after the model is developed. This type of analysis would assist facilities and others in understanding how their inclusion could impact the model and provide additional information for groups examining this issue such as the NQF and Office of the Assistant Secretary for Planning and Evaluation. As a result, the FAH believes that this measure lacks sufficient information on the potential impact these social risk variables have on the risk adjustment model. Furthermore, while the developer provides information on the changes in performance scores that result from social risk factor adjustment (tables 2b3.4b_4a and 2b3.4b_4b), it is not necessarily clear on the degree to which these changes would result in a facility's score being statistically significantly different from the national average. If the interpretation of the results under meaningful differences leads to a "conclusion that even small difference between facility scores can be treated as meaningful"(response to 2b4.3 in the testing form), to what extent would changes in performance as a result of adjustment for social risk factors also lead to different but meaningful results?
3574: Medicare Spending Per	Submitted by American	The American Medical Association (AMA) appreciates the opportunity to comment on this measure and request that the

Topic	Commenter	Comment
Beneficiary (MSPB) Clinician	Medical Association.	Standing Committee carefully consider our comments on its scientific acceptability during this evaluation. The Centers for Medicare & Medicaid Services (CMS) developed this measure specifically for use in the Merit-based Incentive Payment System (MIPS) and we believe that the information and testing provided should demonstrate that its use in MIPS will yield reliable and valid results and enable end users to make meaningful distinctions in the costs associated with the care provided to these patients. The AMA is concerned that the testing results provided, particularly for measure score reliability, empirical validity and the risk adjustment approach, do not provide the information needed to ensure that MSPB Clinician produces the desired results. Regarding the measure score reliability, we are concerned with the lack of information on reliability results below the 25th percentile, particularly in light of the reference within the response of 2a2.3 that CMS generally considers 0.4 to be the threshold for moderate reliability and 100% of practices and clinicians with at least 35 episodes meet it. The AMA believes that the minimum acceptable thresholds should be 0.7 and the measure as specified does not. The AMA strongly supports the tenet that cost must be assessed within the context of the quality of care provided; yet, the developer was unable to demonstrate that this measure correlates to any one quality measure within the MIPS program and differs from what they were able to complete for other MSPB measures currently under review (3561, 3562, 3563, and 3564). We are very troubled that the testing did not include an assessment of MSPB Clinician with a measure such as the claims-based All-Cause Hospital Readmissions (#458) since it was also reported in 2017 and to our knowledge CMS attributed performance to practices for which this cost measure could also apply in that same year. While we acknowledge that the lack of alignment of attribution models creates challenges to complete these analyses, we believe that CMS could solve this issue since the agency serves as the steward for many of the claims-based measure. Regardless, the AMA does not believe that cost measures against which no quality measure can be assessed should achieve endorsement. The AMA does not believe that the current risk adjustment model is adequate due to R-squared results ranging from 0.09 to 0.64 across the groupings nor is the measure adequately tested and adjusted for social risk factors. It is unclear to us why the developer would test social risk factors after adjusting for clinical risk factors rather than assessing the impact of both clinical and social risk factors in the model at the same time. These variations in how risk adjustment factors are examined could also impact how each variable (clinical or social) perform in the model and

Topic	Commenter	Comment
		remain unanswered questions. In addition, while the developer believes that the small differences in measure results "can be interpreted as meaningful" (response in 2b4.2 in the testing form), it is not clear why this same reasoning was not applied for those clinicians and practices for whom inclusion of social risk factors in the models changed the ratios nor is it clear how these same factors would affect a change in performance across the 10 deciles used in the MIPS benchmarking methodology. In addition, the AMA questions whether the information provided in Section 2b4. Identification of Statistically Significant and Meaningful Differences in Performanceis truly useful for accountability and informing patients of the cost of care provided byphysicians and practices. Specifically that the testing does not directly address whether the costs attributed to physicians and practices enable us to distinguish low versus high performers. Since this measure was specifically developed for use in MIPS, analyses of the performance scores using the finalized benchmarking methodology across 10 deciles would provide valuable information on whether the differences in costs between physicians and practices could be considered useful and meaningful. We do not believe that stratifying scores by characteristics such as region, risk score, or the number of episodes attributed satisfactorily answers this question. The AMA requests that these gaps in testing be addressed prior to endorsement of this measure. We appreciate the Committee's consideration of our comments.
3574: Medicare Spending Per Beneficiary (MSPB) Clinician	Submitted by American Association of Neurological Surgeons.	The American Association of Neurological Surgeons (AANS) thanks the NQF for the opportunity to share input on the MSPB Clinician measure (#3574), which was developed and recently revised for use under the Merit-Based Incentive Payment System (MIPS). As noted by the American Medical Association's (AMA) more in-depth analysis, the information and testing provided—particularly for measure score reliability, empirical validity and the risk-adjustment approach— do not demonstrate that the use of this measure under MIPS will yield reliable or valid results or enable us to distinguish low versus high performers. As a result, end users will not be able to make meaningful distinctions regarding the costs associated with the care provided to these patients. Furthermore, the AANS has long voiced concerns about this measure's failure to evaluate cost in the context of quality. The revised version of this measure still does not correlate to any one quality measures used in MIPS.
3574: Medicare Spending Per Beneficiary (MSPB) Clinician	Submitted by American College of Physicians.	The ACP appreciates the opportunity to comment in advance of the NQF Cost and Efficiency Standing Committee's review of several measures submitted for endorsement consideration during the Spring 2020 cycle. The Medicare Spending per Beneficiary (MSPB) measure represents an important move towards cost assessment in pay-

Topic	Commenter	Comment
		for-performance programs. However, the methods that policymakers and measure developers apply to assessing episode-based costs is critical to the success of this initiative. In this regard, several inherent limitations to the measure exist. The Centers for Medicare & Medicaid Services (CMS) should consider addressing the concerns listed below in the interest of enhancing the validity of the measure. The Performance Measure Committee (PMC) of the ACP prefers that all cost measures be attributed to the level of the group/practice or higher for the following reasons: • If health plan administrators and government payers intend to create individual cost profiles to generate incentives to decrease health care costs, it is important that these profiles provide insights into which care management interventions are most effective in reducing costs year-over-year, even if what is measured does not encompass the totality of the cost to Medicare for the items and services provided to a patient during an episode of care. Measuring what is actionable could build trust with clinicians, feed a cycle of participation, and discourage dysfunctional behaviors such as avoiding attribution. Stratifying and comparing results based on costs related to 1) services that are under the direct control of the individual clinician, 2) indirect costs, and 3) services under the control of the facility could help to mitigate this concern by identifying behaviors that correspond with opportunities for improvement. • While improvements have been made to the attribution model, revisions do not address the possibility of multiple clinicians being held accountable for the total costs associated with a single episode. CMS attributes each MSPB episode to the Taxpayer Identification Number-National Provider Identifier (TIN-NPI) responsible for 30% of Part B Physician/Supplier services during the index admission. • According to this model, multiple clinicians could be accountable for the total costs associated with a single episode of care. While we generally support the attribution model at the facility, system, and health plan levels, we caution that attributing patient costs to individual clinicians can be technically challenging. Healthcare costs are influenced not only by the actions of one clinician but often by the actions of multiple clinicians as well as a patient's social, economic, and environmental factors. It is difficult to determine the relative influence that an individual clinician has on a patient's expenses. Understanding who is responsible is essential to driving improvements in care as well as for securing long-term buy-in from clinicians and facilitating

Topic	Commenter	Comment
		the ability of value-based purchasing programs to influence clinician behavior. The current model does not speak to the care coordination system that most clinicians would likely endorse. For example, Accountable Care Organizations that build on the value-based purchasing framework to enhance care coordination and promote responsibility for clinical and efficiency outcomes. Additional areas of concern are as follows: • We are unable to assess the benefit of assessing costs (e.g., if it helps to improve outcomes at lower costs) without assessing the evidence to support this claim. We recommend that NQF require that measure developers document the evidence base for cost/resource use measures and that it at least aligns with what is required for outcome measures (i.e., Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service). • The implications of the risk-adjustment model as currently specified are unclear. The model estimates expected episode costs in recognition of the different levels of care beneficiaries may require due to comorbidities, disability, age and other risk factors. This model is not sufficient to control for all significant social determinants of health (SDOH) that may influence the clinical health status of patients as well as the outcome of acute admissions. The Centers for Medicare & Medicaid Services (CMS) should consider revising the risk-adjustment model to include SDOH that are most likely to influence the clinical health status of the denominator population under consideration. Aligning the model for risk-adjustment with more robust methods for statistical analyses that consider all factors that are independently and significantly associated with outcomes and that vary across measurement participant (e.g., the Society for Thoracic Surgeons Adult Cardiac Surgery Risk Model) could enhance individual clinician acceptance of outcomes measures and helps to mitigate risk aversion. • The 30% threshold is too low to attribute episode-based care to an individual clinician. CMS should consider increasing the attribution threshold to an evidence-based percentage that represents the majority of services during hospitalization. • The 30-day episode window is arbitrary. Recent literature suggests that shorter intervals of seven or fewer days might improve the accuracy and equity of episode-based costs to Medicare as a measure of facility quality for public accountability.

Topic	Commenter	Comment
		• While we note that the current use of this measure requires that clinicians and clinician groups meet a 35-episode case minimum which is referenced in a few sections of the submission form, we would recommend that this minimum requirement be included in the technical measure specifications - either in the denominator requirements or exclusions. This is particularly important given that the measure's reported reliability results rely on a minimum volume threshold of 35 episodes. • Maximizing transparency could build trust with clinicians and feed a cycle of participation. CMS should consider establishing a premortem approach for evaluating the impact of performance measures to combat the unintended consequences of implementation and correctly identify reasons for future outcomes. While this measure aims to reduce low-value care, implementation may result in consequences directly contrary to the spirit of the measure. The measure specifies "episodes of care for a beneficiary if the beneficiary dies during the episode" as exclusion criteria. Therefore, the measure rewards clinicians for expending minimal resources on patients in stable conditions, while disregarding mortality rates, and penalizes clinicians for disbursing sufficient resources to maintain the stability of medically complex patients during an episode of care.
3574: Medicare Spending Per Beneficiary (MSPB) Clinician	Submitted by American Society of Clinical Oncology.	The American Society of Clinical Oncology (ASCO) appreciates the opportunity to submit comments to the National Quality Forum (NQF) Cost and Efficiency Technical Advisory Panel. Following are our general comments on the Medicare Spending per Beneficiary (MSPB) and Total per Capita Cost (TPCC) measures. ASCO is the national organization representing nearly 45,000 physicians and other health care professionals specializing in cancer treatment, diagnosis, and prevention. ASCO members are also dedicated to conducting research that leads to improved patient outcomes, and we are committed to ensuring that evidence-based practices for the prevention, diagnosis, and treatment of cancer are available to all Americans, including Medicare beneficiaries. • Given the growing number of episode-based cost measures, and continued work on their development, ASCO would encourage the NQF and CMS to consider whether the TPCC and MSPB measures still serve a purpose, as many of the beneficiaries captured in the episode-based measures will also be included in either or both the MSPB and TPCC measures. With the measures as proposed, a beneficiary could potentially be attributed to multiple providers within and across multiple measures. First, this could magnify the impact on cost

Topic	Commenter	Comment
		measures of any individual beneficiary and second, could complicate any true differences in cost and value. CMS developed these measures specifically for use in the Merit-based Incentive Payment System (MIPS) and we believe that the measure and attribution should demonstrate that its use in MIPS will not just yield reliable and valid results, but most importantly, enable end users to make meaningful distinctions in the costs associated with the care provided to these patients.
3574: Medicare Spending Per Beneficiary (MSPB) Clinician	Submitted by Infectious Diseases Society of America.	IDSA appreciates the opportunity to provide comments to the NQF Cost and Efficiency Standing Committee. IDSA agrees with the findings of the American Medical Association's (AMA) more detailed analysis of the MIPS Medicare Spending per Beneficiary (MSPB) and Total per Capita Cost (TPCC) measures. We continue to have concerns about the ability of these measures to accurately and reliably distinguish performance among clinicians, the ongoing failure of these cost measures to link to relevant quality measures under MIPS, and the ongoing failure of these measures to produce meaningful and comprehendible information that clinicians can use to enhance patient care and value. We are also concerned that ID physicians may be held accountable simultaneously for both cost measures under MIPS. While recent revisions to these measures were intended to avoid this situation, many members of our specialty work in both inpatient and outpatient settings. As a result, they may be captured under the MSPB measure under the medical E/M attribution rule, but also under the TPCC measure since the ID specialty is not specifically excluded from this measure.
3574: Medicare Spending Per Beneficiary (MSPB) Clinician	Submitted by American Psychiatric Association.	The American Psychiatric Association appreciates the opportunity to submit comments for the Cost and Efficiency Standing Committee's review. APA continues to have serious concerns about the Medicare Spending Per Beneficiary (MSPB) measure, and concurs with the American Medical Association (AMA)'s more detailed analysis of the measure. It is not clear that clinicians can control the costs that are attributed to them as part of this measure, particularly those costs that are incurred after hospital discharge. In addition, the developer notes that they are unable to adequately test the relationship between performance on the cost measure and performance on conceptually-related quality measures, including patient outcomes. This is a relationship that should be explored more thoroughly before implementation of the measures, to guard against unintended consequences or mis-alignment of incentives for healthcare providers.
3574: Medicare Spending Per Beneficiary	Submitted by American Academy of	The American Academy of Neurology (AAN) appreciates the opportunity to comment on this measure and hopes the Cost and Efficiency Technical Advisory Panel takes these comments into consideration when deliberating on the measure.

Topic	Commenter	Comment
(MSPB) Clinician	Neurology (AAN).	The AAN echoes the American Medical Association's concerns related to the measure score reliability, empirical validity and risk adjustment methodology for this measure. While The Centers for Medicare & Medicaid Services (CMS) developed the Medicare Spending Per Beneficiary (MSPB) Clinician measure for use in the Merit-based Incentive Payment System (MIPS) to meaningfully and reliably distinguish individual and groups by measuring costs associated with the care provided to patients, the measure testing results are unclear and fail to demonstrate reliable and valid results that support moving forward with measure endorsement and implementation at this time. Our top concerns based on the measure testing results include: • An inadequate moderate reliability threshold • A lack of correlation to quality measures used in MIPS; cost of care assessments should be rooted within the context of quality measure assessment, which this measure fails to do • An inadequate and unclear risk adjustment model, including lack of appropriate testing and adjustment for social risk factors With these concerns in mind, the AAN does not support the measure based on the testing results provided and these gaps should be addressed before endorsement and implementation.
3575: Total Per Capita Cost (TPCC)	Submitted by Federation of American Hospitals.	The Federation of American Hospitals (FAH) appreciates the opportunity to comment on this measure prior to the Standing Committee's evaluation. The FAH requests that the committee carefully consider whether the measure as specified produces performance scores that are reliable and valid for reporting at either the clinician or practice levels. While reliability at the 25th percentile for at least 20 episodes for practices was 0.77 and 0.83 for clinicians, the FAH questions what result was produced at the minimum level for either reporting group. The FAH was particularly concerned to review the additional explanation provided in section 2a2.3 in the testing form that "100 percent of TINs and TIN-NPIs at the reporting case minimum have reliability greater than or equal to 0.4, the standard that CMS generally considers as the threshold for 'moderate' reliability", which should not be considered an acceptable minimum threshold. The FAH believes that additional information regarding the minimum result is needed to ensure that the measure as specified produces scores that achieve an acceptable minimum threshold for reliability. The FAH is extremely troubled by the lack of any validity testing demonstrating the presence or absence of correlations of this cost measure to quality measures. We found the rationales outlining the inability of the developer to identify appropriate quality measures to be weak since QPP#458, All-cause Hospital Readmission, which is also a claims-based measure, was

Topic	Commenter	Comment
		attributed to groups in 2017 and we assume that CMS would be able to enable matching of groups to whom this measure and QPP#458 applied. In addition, it is concerning for a cost measure to be considered for endorsement when the developer is unable to identify applicable quality measures and reports that the quality measures used within the program do not enable this type of analysis; yet, points are achieved and an overall score is derived in part from the quality and cost categories in MIPS. Due to the importance of understanding costs as they relate to quality, it seems imprudent to consider endorsement of a measure for which its association to quality cannot be demonstrated. The FAH was further concerned with the results from the analysis of the measure's attribution methodology. Similar to our questions on the reliability testing, the 25th percentile showing that just under 29% and 35% of E&M claims were billed by the attributed clinician and practice, respectively. These results lead us to ask what share of claims occurred at lower levels. The other analysis to determine the extent to which nurse practitioners and physician assistants to whom this measure was attributed also shows that inappropriate attributions are occurring to 5.5% of practices. While the developer views this result to be acceptable, the specialties in which this occurs includes those that are intentionally excluded within the specifications. We believe that these results impact the validity of the measure and could result in negative unintended consequences in attributing costs to clinicians and practices who do not provide primary care services. In addition, while the FAH appreciates that social risk factors were reviewed, and believes that the risk adjustment approach should not consider the identification and testing of social risk factors as supplementary to clinical risk factors. This approach was identified as a concern by the NQF Disparities Standing Committee and developers must begin to include these factors within the testing of the model rather than the approach of "adding on" factors after the model is developed. This type of analysis would assist facilities and others in understanding how their inclusion could impact the model and provide additional information for groups examining this issue such as the NQF and Office of the Assistant Secretary for Planning and Evaluation. As a result, the FAH believes that this measure lacks sufficient information on the potential impact these social risk variables have on the risk adjustment model. Furthermore, while the developer provides information on the differences in observed to expected cost ratios that result from social risk factor adjustment (table 2b3.4b.c), it is not necessarily clear on the degree to which these changes would result in a practice or clinician's result changing. Specifically, what is not answered is whether the addition of social risk factors to the model would lead to a clinician or practice earning higher or

Topic	Commenter	Comment
		lower points in the benchmarks currently used for this measure within MIPS. If the interpretation of the results under meaningful differences leads to statements that "small differences in scores can be interpreted as meaningful"(response to 2b4.3 in the testing form), to what extent would changes in performance as a result of adjustment for social risk factors also lead to different but meaningful results?
3575: Total Per Capita Cost (TPCC)	Submitted by American Society of Clinical Oncology	The American Society of Clinical Oncology (ASCO) appreciates the opportunity to submit comments to the National Quality Forum (NQF) Cost and Efficiency Technical Advisory Panel. Following are our general comments on the Medicare Spending per Beneficiary (MSPB) and Total per Capita Cost (TPCC) measures. ASCO is the national organization representing nearly 45,000 physicians and other health care professionals specializing in cancer treatment, diagnosis, and prevention. ASCO members are also dedicated to conducting research that leads to improved patient outcomes, and we are committed to ensuring that evidence-based practices for the prevention, diagnosis, and treatment of cancer are available to all Americans, including Medicare beneficiaries. Given the growing number of episode-based cost measures, and continued work on their development, ASCO would encourage the NQF and CMS to consider whether the TPCC and MSPB measures still serve a purpose, as many of the beneficiaries captured in the episode-based measures will also be included in either or both the MSPB and TPCC measures. With the measures as proposed, a beneficiary could potentially be attributed to multiple providers within and across multiple measures. First, this could magnify the impact on cost measures of any individual beneficiary and second, could complicate any true differences in cost and value. CMS developed these measures specifically for use in the Merit-based Incentive Payment System (MIPS) and we believe that the measure and attribution should demonstrate that its use in MIPS will not just yield reliable and valid results, but most importantly, enable end users to make meaningful distinctions in the costs associated with the care provided to these patients. ASCO requests that the Standing Committee evaluate whether the attribution methodology is valid and does not lead to negative unintended consequences. While the TPCC eliminates the problem of attributing costs that occurred before the clinician ever saw the patient, which ASCO supports ,the attribution methodology assumes that a primary care relationship exists if two things happen within three days or three months, and not otherwise, leading to problems as identified in the following examples: • A cancer survivor receives a twelve-month follow-up exam from their oncologist, along with a two-

Topic	Commenter	Comment
		dimensional echocardiogram with doppler flow study to screen for cardiotoxicity. The oncologist is attributed the beneficiary's costs for a twelve-month period, despite no other management of the patient. • A newly diagnosed cancer patient requests a second opinion from an oncologist other than their primary clinician. The oncologist conducts an evaluation and management service which happens to take place within +/- 3 days of other designated primary care services. The oncologist performing the second opinion confirms the primary clinician's treatment plan and the patient returns to their primary clinician for continued management and treatment. The consulting oncologist is attributed the beneficiary's costs for a twelve-month period, despite never having managed the beneficiary's care. • A nurse practitioner is employed by a cancer practice to assist in management of cancer patients receiving chemotherapy and/or radiation therapy. The nurse practitioner does not qualify for an exemption from the measure given that a physician's NPI, rather than theirs, is used to bill for the chemotherapy services. • An oncologist whose TIN includes in-office chemotherapy services is attributed a patient who receives chemotherapy services within 90 days after an E&M primary care service, but outside of +/- 3 days of the E&M primary care service. An equivalent hospital-based oncologist is not attributed a similar case, as the chemotherapy services are billed by the hospital TIN. • An oncologist whose TIN includes in-office chemotherapy services qualifies for an exemption from the measure due to their NPI-TIN being used to bill for chemotherapy services. An equivalent hospital-based oncologist does not qualify for an exemption, as the chemotherapy services are billed by the hospital TIN. In each of these examples, an oncologist will not know if they qualify for the TPCC measure, as the exemption is applied retrospectively based on a measurement of candidate events for which the oncologist bills for chemotherapy or radiation therapy services. We feel it is inappropriate for a clinician to be included in a measure for which they are unaware of which beneficiaries they may be attributed, or whether they will receive an exemption. We have previously recommended that all medical and radiation oncologists be excluded from the TPCC measure. The analysis of the extent to which nurse practitioners and physician assistants have this measure attributed to them found that 5.5% of practices (just over 4,000 of the 74,191 TINs) were ones in which the specialty is not considered to provide primary

Topic	Commenter	Comment
		care. We believe that these findings are the result of the decision to make exclusions at the specialty level and not at the service level. While the measure excludes certain specialties, the results as outlined in Table 7 (Frequency of Most Common HCFA Specialties in TINs Attributed TPCC Measure via Nurse Practitioners and Physician Assistants Alone) of the testing form confirm that there is potential for the measure to attribute patients to clinicians who do not provide primary care services. These results from both analyses lead to questions on the validity of the attribution methodology as it creates a fairness issue by sometimes including certain specialties regarded as not providing primary care, but it also holds primary care clinicians responsible for the costs of non-primary-care services that they do not provide and cannot control. ASCO requests that these gaps in attribution be addressed prior to endorsement of this measure. We appreciate the Committee's consideration of our comments.
3575: Total Per Capita Cost (TPCC)	Submitted by American Academy of Neurology (AAN).	The American Academy of Neurology (AAN) appreciates the opportunity to comment on this measure and hopes the Cost and Efficiency Technical Advisory Panel takes these comments into consideration when deliberating on the measure. The AAN echoes the American Medical Association's concerns related to the measure score reliability, empirical validity and attribution methodology for this measure. While The Centers for Medicare & Medicaid Services (CMS) developed the Total Per Capita Cost (TPCC) measure for use in the Merit-based Incentive Payment System (MIPS) to meaningfully and reliably distinguish individual and groups by measuring costs associated with the care provided to patients, the measure testing results are unclear and fail to demonstrate reliable and valid results that support moving forward with measure implementation at this time. Our top concerns based on the measure testing results include: • An inadequate moderate reliability threshold • A lack of correlation to quality measures used in MIPS; cost of care assessments should be rooted within the context of quality measure assessment, which this measure falls to do • An unreliable attribution methodology that has several potential unintended consequences including, that multiple clinicians can be attributed to the measure unrelated to practicing team-based care. • Neurologists and neurology advanced practice providers often consult with primary care and other specialists as they care for patients with neurological conditions, many of them chronic. Without clearer attribution methodology, the measure as written could give the perverse incentive to any of the clinicians involved to schedule follow up visits within three months so the

Topic	Commenter	Comment
		patient and his or her costs are not attributed to that clinician. With these concerns in mind, the AAN does not support the measure based on the testing results provided and these gaps should be addressed before endorsement and implementation
3575: Total Per Capita Cost (TPCC)	Submitted by American Society of Retina Specialists.	The American Society of Retina Specialists appreciates the opportunity to provide comments on the Standing Committee related to its consideration of the Total Per Capital Cost (TPCC) measure. ASRS is the largest retinal organization in the world, representing over 3,000 members in all 50 US states, the District of Columbia, Puerto Rico, and 63 countries. The Society serves as a national advocate and primary source of clinical and scientific information and education for its members. ASRS opposes the inclusion of TPCC in the Merit-Based Incentive Payment System (MIPS) and urges the Standing Committee not to recommend it for endorsement because it potentially holds physicians responsible for the cost of care that they did not provide. While we continue to oppose the measure concept overall due to its attribution of costs, we appreciate that CMS has taken steps to target this measure more specifically to primary care physicians by excluding most specialists and surgeons, including all ophthalmologists. If this measure is not removed from the MIPS program, however, at a very minimum it must retain the specialty exclusions. This measure is not appropriate for measuring or influencing individual physician performance and thus should not be included in the MIPS program. TPCC seeks to assign the total costs of all care for individual Medicare beneficiaries to primary care physicians tasked with coordinating and overseeing the total healthcare of the patient. While CMS listened to stakeholder feedback that the previous attribution methodology for the measure potentially held physicians responsible for care that happened before the primary care physician saw the patient or well after the patient had left that particular physician's care, its revisions to the methodology to shorten the time frame for attribution have created an additional problem of potentially double-counting costs by assigning them to more than one physician or group. This risks confusion over who is responsible for the costs and may inappropriately label a physician as high or low cost. Furthermore, the measure updates fail to address the underlying issue that the attributed physician is still being held responsible for the cost of care that he or she neither provides nor has any ability to influence. Under the TPCC methodology, a primary care physician will be attributed the cost of care provided by a retina specialist, such as macular degeneration treatment or surgical repair of a torn or detached retina, even though he or she has no influence over the cost or quality of that treatment. By including

Topic	Commenter	Comment
		all costs of care for a particular beneficiary, the measure loses its overall usefulness to the attributed physician since he or she is limited in the ability to modify or influence the behavior of other physicians caring for the patient. The physician will not ultimately take action that lowers the cost of care for the beneficiary and the measure score will not accurately assign physicians as high or low cost. Despite ASRS' overall opposition to the measure concept, we applaud CMS for listening to feedback and excluding specialists and surgeons, including all ophthalmologists, who had patients inappropriately attributed to them by application of the TPCC measure under its prior methodology. Previously, beneficiaries who did not see a primary care physician sometime during the performance year were attributed to whichever physician or group billed the plurality of evaluation and management (E/M) services during the year, which could be a retina specialist. Retina specialists provide care only for diseases of the retina and macula and do not provide overall or systemic healthcare for the patient. While retina specialists treat diabetic eye disease, such as diabetic retinopathy, they do not manage the patient's overall diabetes care. Although any physician is limited in his or her ability to influence their TPCC score, retina specialists are especially disadvantaged since the care they provide is so specialized. The new attribution methodology appropriately excludes them and all other ophthalmologists from the measure and must be retained if this flawed measure is included in the MIPS program. Thank you for the opportunity to provide comments on the TPCC measure. ASRS continues to oppose the inclusion of this measure in the MIPS program because it holds physicians responsible for the cost of care they did not provide, thereby limiting their ability to influence their performance on the measure. While this measure should be removed from MIPS entirely, retention of the specialty exclusions is necessary to ensure specialists and surgeons, such as retina specialists, do not have patients and costs inappropriately attributed to them under the measure. We urge the Standing Committee not to recommend this measure for endorsement. For additional information, please contact Allison Madson, director of health policy, at allison.madson@asrs.org or (312) 578-8760.
3575: Total Per Capita Cost (TPCC)	Submitted by Infectious Diseases Society of America.	IDSA appreciates the opportunity to provide comments to the NQF Cost and Efficiency Standing Committee. IDSA agrees with the findings of the American Medical Association's (AMA) more detailed analysis of the MIPS Medicare Spending per Beneficiary (MSPB) and Total per Capita Cost (TPCC) measures. We continue to have concerns about the ability of these measures to accurately and reliably distinguish performance among clinicians,

Topic	Commenter	Comment
		the ongoing failure of these cost measures to link to relevant quality measures under MIPS, and the ongoing failure of these measures to produce meaningful and comprehendible information that clinicians can use to enhance patient care and value. We are also concerned that ID physicians may be held accountable simultaneously for both cost measures under MIPS. While recent revisions to these measures were intended to avoid this situation, many members of our specialty work in both inpatient and outpatient settings. As a result, they may be captured under the MSPB measure under the medical E/M attribution rule, but also under the TPCC measure since the ID specialty is not specifically excluded from this measure.
3575: Total Per Capita Cost (TPCC)	Submitted by American College of Physicians.	The ACP appreciates the opportunity to comment in advance of the NQF Cost and Efficiency Standing Committee's review of several measures submitted for endorsement consideration during the Spring 2020 cycle. The Total per Capita Cost measure represents an important move towards cost assessment in pay-for-performance programs. However, the methods that policymakers and measure developers apply to assessing costs is critical to the success of this initiative. In this regard, several inherent limitations to the measure exist. The Centers for Medicare & Medicaid Services (CMS) should consider addressing the concerns listed below in the interest of enhancing the validity of the measure. The Performance Measure Committee (PMC) of the ACP prefers that all cost measures be attributed to the level of the group/practice or higher for the following reasons: • If health plan administrators and government payers intend to create individual cost profiles to generate incentives to decrease health care costs, it is important that these profiles provide insights into which care management interventions are most effective in reducing costs year-over-year, even if what is measured does not encompass the totality of the cost to Medicare for the items and services provided to a patient during an episode of care. Measuring what is actionable could build trust with clinicians, feed a cycle of participation, and discourage dysfunctional behaviors such as avoiding attribution. Stratifying and comparing results based on costs related to 1) services that are under the direct control of the individual clinician, 2) indirect costs, and 3) services under the control of the facility could help to mitigate this concern by identifying behaviors that correspond with opportunities for improvement. • While improvements have been made to the attribution model, revisions do not address the possibility of multiple clinicians being held accountable for the total costs associated with a single episode. CMS attributes each

Topic	Commenter	Comment
		beneficiary to a single Taxpayer Identification Number-National Provider Identifier (TIN-NPI) if the beneficiary received more primary care services from primary care clinicians in that TIN-NPI than any other TIN-NPI or CMS Certification Number (CCN). If two TIN-NPIs tie for the largest share of a beneficiary's primary care services, CMS attributes the beneficiary to the TIN-NPI that provided primary care services most recently. According to this model, multiple clinicians could be accountable for the annualized costs of care for beneficiaries attributed to the TIN-NPI. While it is reasonable to apply this model to health plans, it is unclear how this approach will provide meaningful information to individual clinicians that will appropriately inform quality improvements. While we generally support the attribution model at the facility, system, and health plan levels, we caution CMS that attributing patient costs to individual clinicians can be technically challenging. Healthcare costs are influenced not only by the actions of one clinician but often by the actions of multiple clinicians as well as a patient's social, economic, and environmental factors. It is difficult to determine the relative influence that an individual clinician has on a patient's expenses. Understanding who is responsible is essential to driving improvements in care as well as for securing long-term buy-in from clinicians and facilitating the ability of value-based purchasing programs to influence clinician behavior. The current model does not speak to the care coordination system that most clinicians would likely endorse. For example, Accountable Care Organizations that build on the value-based purchasing framework to enhance care coordination and promote responsibility for clinical and efficiency outcomes. Additional areas of concern are as follows: • The implications of the risk-adjustment model as currently specified are unclear. The model estimates expected episode costs in recognition of the different levels of care beneficiaries may require due to comorbidities, disability, age and other risk factors. This model is not sufficient to control for all significant social determinants of health (SDOH) that may influence the clinical health status of patients as well as the outcome of acute admissions. The Centers for Medicare & Medicaid Services (CMS) should consider revising the risk-adjustment model to include SDOH that are most likely to influence the clinical health status of the denominator population under consideration. Aligning the model for risk-adjustment with more robust methods for statistical

Topic	Commenter	Comment
		analyses that consider all factors that are independently and significantly associated with outcomes and that vary across measurement participant (e.g., the Society for Thoracic Surgeons Adult Cardiac Surgery Risk Model) could enhance individual clinician acceptance of outcomes measures and helps to mitigate risk aversion. • While we note that the current use of this measure requires that clinicians and clinician groups meet a 35-episode case minimum which is referenced in a few sections of the submission form, we would recommend that this minimum requirement be included in the technical measure specifications - either in the denominator requirements or exclusions. This is particularly important given that the measure's reported reliability results rely on a minimum volume threshold of 35 episodes. • Additionally, CMS should consider establishing a premortem approach for evaluating the impact of performance measures to combat the unintended consequences of implementation and correctly identify reasons for future outcomes. • CMS should independently establish a robust minimum average reliability rating and evaluate all future cost measures based on that same standard, not pre-determine a set of measures the Agency wishes to use then selecting whatever low reliability standard allows them to adopt all of those measures without raising case minimums. CMS designed this measure to seemingly reward the creation of Patient-Centered Medical Homes; however, PCMH models have not been uniformly successful in achieving care quality improvements.
General comments on the draft report	Submitted by American Academy of Physical Medicine and Rehabilitation.	The American Academy of Physical Medicine and Rehabilitation (AAPM&R) appreciates the opportunity to comment on the Medicare Spending per Beneficiary (MSPB) and Total Per Capita Cost (TPCC) measures. AAPM&R has submitted comments expressing our concerns about these measures to CMS on multiple occasions in the past several years. AAPM&R asserts it is inappropriate to use broad measures such as MSPB and TPCC to evaluate the resource use of individual physicians. Many Medicare beneficiaries have multiple health problems, and in most cases those different health problems are treated by multiple physicians and non-physician providers. Moreover, under Medicare rules, beneficiaries have the freedom to see any physicians they wish to obtain treatment for their health problems. Even if each of the individual physicians who a patient sees are "efficient" in the services they deliver or order, the overall spending on the patient's care may be higher than for

Topic	Commenter	Comment
		other patients because of the number and types of physicians and non-physician providers the patient chooses to access. We therefore discourage NQF endorsement of the MSPB and TPCC measures. • AAPM&R is aware that the American Medical Association (AMA) has developed detailed comments regarding the MSPB and TPCC measures. AAPM&R supports the AMA's comments and detailed analysis. We support the AMA's assertion that cost must be assessed within the context of the quality of care provided. Further, we agree that cost measures against which no quality measure can be assessed should not achieve NQF endorsement.
General comments on the draft report	Submitted by American Society for Radiation Oncology.	The American Society for Radiation Oncology1 (ASTRO) appreciates the opportunity to provide comments on proposed revisions to the two Merit Based Incentive Payment System (MIPS) cost measures: Medicare Spending Per Beneficiary (MSPB) and Total Per Capita Cost Measures (TPCC). While ASTRO appreciates the intent of these measures, we are concerned that the proposed revisions to these measures fall short of their intended purpose. ASTRO shares the concerns expressed by the American Medical Association (AMA) regarding the measures and testing results provided, particularly for measure score reliability, empirical validity and the risk adjustment approach, which do not provide the information needed to ensure that the MSPB Clinician or TPCC Clinician produces the desired results. ASTRO urges the Committee to address gaps in testing prior to final endorsement of the measures. Additionally, ASTRO appreciates previous efforts to exclude specialty specific services, including radiation therapy, from the MSPB and TPCC methodologies. However, radiation therapy continues to be an included service associated with diseases of the ear, nose, mouth and throat in the post-trigger inpatient component of the MPSB. Radiation therapy is excluded for all other services, regardless of diagnosis. ASTRO urges the Agency to correct this oversight and exclude radiation therapy from the post-trigger inpatient component for services involving disease of the ear, nose, mouth and throat. ASTRO appreciates the opportunity to comment on the revisions to the MSPB and TPCC measures. If you have any questions, they can be directed to Anne Hubbard, Director of Health Policy at 703-839-7394.

National Quality Forum
1099 14th Street NW, Suite 500
Washington, DC 20005
http://www.qualityforum.org