Endorsing Cost and Resource Use Measures

TECHNICAL REPORT

January 31, 2014

This report is funded by the Department of Health and Human Services under contract HHSM-500-2012-00009I Task Order 1

NATIONAL QUALITY FORUM

Contents

Cost and Resource Use 2012

TECHNICAL REPORT

Introduction

Per capita healthcare spending in the United States is unmatched by any country in the world.[1] This high rate of spending, however, has not resulted in better health for Americans. In fact, higher spending has not decreased mortality, increased patient satisfaction, or led to improvements in access or higher quality of care.[2,3] This phenomenon of high spending with variable outcomes points to a system laden with waste. The contributing factors to this concerning trend are as complex as the healthcare system itself, with physician practice patterns, regional market influences, and access to care as major drivers. Meanwhile, the United States' healthcare spending continues to increase at a rate of 7% per year and is largely focused on treating acute and chronic illness rather than preventive care.[4] By improving efficiency, there is potential to reduce the rate of cost growth and improve the quality of care provided simultaneously. Evidence shows that not all care leads to better outcomes; thus, some portion of these current costs may be unnecessary. To identify and provide incentives for providers to deliver high quality, lower-cost care requires quality and resource use measures.

The National Quality Strategy's (NQS) three aims—better care, affordable care, and healthy people, healthy communities—have intensified the need to identify measures that address cost and align them with the relevant quality measures already in the marketplace. The NQS specifically identifies affordability as a target area for improvement, with goals of:

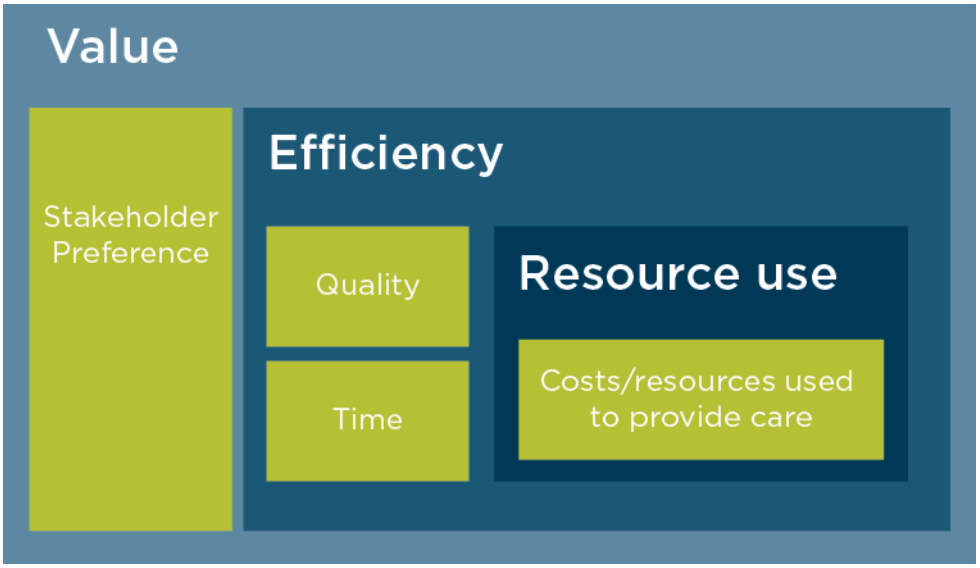
1) Ensuring affordable and accessible high quality health care for people, families, employers, and governments.
2) Supporting and enabling communities to ensure accessible, high quality care while reducing waste and fraud.

As ongoing health reform efforts focus on expanding coverage, increasing access to care, and reducing costs, it is important to understand how resources are currently being used in the system in the context of quality generally; however, the relationship to health outcomes is preferable. Aligning resource use (or cost) and quality measures will enable the system to better evaluate efficiency of care. Several provisions in the Affordable Care Act (ACA), slated to be implemented over in 2015, require using resource use data to further support efforts to move toward a value-based purchasing (VBP) payment model. Resource use data will be included on the physician compare website, as well as a physician value modifier that will be used to adjust fee-for-service (FFS) payments by combining physician performance on quality and resource use.

In January 2010, NQF released the *Measurement Framework: Evaluating Efficiency Across Patient-Focused Episodes of Care,* which addressed cost and resource use as one of the three overarching domains for assessing efficiency. This framework recommended that measures of resource use and cost should incorporate approaches to measure actual prices paid to providers, standardized prices, in

addition to overall utilization. Further, inappropriate care, including failing to provide an evidence-based intervention to an eligible patient or administering an intervention that is unwarranted, cannot be efficient.

NQF's work on endorsing cost and resource use measures has built on the concept within the Efficiency Framework report that measures of cost and quality must be aligned in order to truly understand efficiency and value (see the figure below). NQF has defined efficiency broadly as the resource use (or cost) associated with a specific level of performance with respect to the other five Institute of Medicine (IOM) aims of quality: safety, timeliness, effectiveness, equity, and patient-centeredness.



To expand the NQF portfolio of endorsed cost and resource use measures that in turn could be used as building blocks toward understanding efficiency and value, NQF embarked in 2010 on its first effort to evaluate and endorse cost and resource use measures. Laying the foundation for NQF's current work, the 2010 effort defined cost and resource use measures as broadly applicable and comparable measures of health services counts (in terms of units or dollars) that are applied to a population or event (broadly defined to include diagnoses, procedures, or encounters). A resource use measure counts the frequency of defined health system resources; some may further apply a dollar amount (e.g., allowable charges, paid amounts, or standardized prices) to each unit of resource use. This learning was captured in the 2012 report, *National Voluntary Consensus Standards for Cost and Resource Use* (and in its associated technical report). The report yielded the first eight endorsed cost and resource use measures in the NQF portfolio, and the NQF Resource Use Measure Evaluation Criteria. The work in this first consensus development project on cost and resource use measures serves as the foundation for this project.

Measure Evaluation

Consensus Development Process

Steering Committee Evaluation

On May 8-9, 2013 the Cost and Resource Use Steering Committee evaluated two new measures against NQF's Resource Use Measure Evaluation Criteria:

- 2158: Medicare Spending Per Beneficiary measure (CMS)
- 2165: Total Per Capita Cost Measure for Medicare Beneficiaries measure (CMS)

To facilitate the evaluation, each of the committee members completed preliminary evaluation of the measures prior to consideration by the entire Steering Committee at the in-person meeting. The Committee's discussion and ratings of the criteria are summarized in the evaluation tables beginning on page 11.

COST AND RESOURCE USE SUMMARY

	MAINTENANCE	NEW	TOTAL
Measures under consideration	0	2	2
Measures withdrawn from consideration	0	0	0
Measures Recommended	0	1	1
Not recommended	0	1	1

Steering Committee members discussed both measures on two occasions, at the May 8 and 9, 2013 in-person meeting and during a conference call on August 28, 2013 following the NQF member and public commenting period. The Cost and Resource Use Steering Committee evaluated the measures against NQF's resource use measure evaluation criteria and voted on both measures during the in-person meeting.

The Overarching Issues section outlines the concerns raised by the Committee with respect to each measure. At the in-person meeting, the Steering Committee voted as follows for the two measures:

Table 1. Voting Results from May 8 and 9, 2013 in-person meeting

Measure Number	Importance to Measure and Report (High, moderate, low, insufficient)	Scientific Acceptability of Measure Properties (high, moderate, low, insufficient)	Usability (high, moderate, low, insufficient)	Feasibility (high, moderate, low, insufficient)	Meet Criteria for Endorsement (YES, NO)
2158: Medicare Spending Per Beneficiary	H-9; M-15; L-1; I-0	*Reliability:* H-10; M-14; L-1; I-0 *Validity:* H-0; M-13; L-11; I-1	H-23; M-1; L-0; I-0	H-6; M-15; L-3; I-0	Y-15; N-10
2165: Total Per Capita Cost	H-11; M-10; L-4; I-0	*Reliability:* H-5; M-18; L-1; I-0; *Validity:* H-0; M-13; L-12; I-0	H-19; M-5; L-1; I-0	H-4; M-14; L-7; I-0	Y-11; N-14

Due to a lack of consensus among the Committee on various criteria, the Steering Committee expressed discomfort with being required to make a final vote on an endorsement recommendation for both measures without the benefit of reviewing the comments received during the upcoming NQF member and public commenting period. As such, the Steering Committee and NQF agreed to solicit comments on both measures and then allow the Steering Committee the opportunity to re-vote on the measures after consideration of all comments received.

On August 28, 2013, the Steering Committee met via conference call to review and discuss the submitted comments received during the public and member comment period. The issues raised during the comment period and the Steering Committee's discussions regarding both measures are captured in the measure evaluation summaries. The Steering Committee subsequently re-voted on both measures on the final endorsement recommendation only. Following the re-vote, measure #2158 Medicare Spending Per Beneficiary was recommended by the Committee for NQF endorsement. Measure #2165 Total Per Capita Cost was not recommended for endorsement.

Table 2. Updated Steering Committee Voting Results (Following August 28 call)

Measure Number	Meet Criteria for Endorsement (YES–NO)
2158: Medicare Spending Per Beneficiary	Y-17; N-8
2165: Total Per Capita Cost	Y-12; N-13

NQF Member Voting

Measure #2158: Medicare Spending Per Beneficiary was voted on by the NQF membership September 9 through 23, 2013. Membership voting results did not indicate clear consensus, with 43% of NQF member councils approving the measure. Representatives of 42 member organizations voted, with no votes received from the Public/Community Health Agency council. The Consumer, Purchaser, and Health Plan councils voted in favor of the measure; the Health Professional, Provider, and Supplier/Industry councils voted against the measure; and the QMRI council was split.

Table 3. NQF Member Voting Results for Measure #2158: Medicare Spending Per Beneficiary (CMS)

Member Council	Percent Approval	Total Votes
Consumer	100%	4
Health Plan	100%	4
Health Professional	13%	15
Provider Organizations	29%	7
Public/Community Health Agency	N/A	0
Purchaser	100%	5
QMRI	50%	4
Supplier/Industry	0%	3
All Councils	**43%**	**42**

Consensus Standards Approval Committee (CSAC)

At its October 8 conference call, the CSAC reviewed the recommendations from the Cost and Resource Use project, including the Steering Committee deliberations, public and member comments, and member voting results. Due to the lack of consensus noted among the councils represented in the voting results, the CSAC requested input from the NQF member councils to gain a better understanding

of the perspective of the NQF membership and determine whether consensus among the councils could be reached before making an endorsement recommendation.

Each council gathered input during the month of October, via conference call or email, and a representative from each council was invited to attend the November 6 CSAC meeting. At their in-person meeting on November 6, CMS, Acumen, and the council chairs provided input to the CSAC on the measure.

Representatives from both CMS and the measure developer, Acumen, were present at the meeting to clarify several committee and CSAC concerns as well as address questions from the CSAC and Council representatives. Using this presentation, Acumen was able to clarify concerns regarding the use of measure #2158 Medicare Spending Per Beneficiary in conjunction with quality measures, and provide a response to concerns that costs captured are largely driven by post-acute services. Acumen was also able to clarify the risk adjustment model, and the exclusions of deaths and transfers. Representatives from all councils provided input on the measure, stating the importance of the measure but also raising methodological concerns discussed in the overarching issues section.

After considering the input from the NQF member councils, in addition to the Steering Committee recommendation, public and member comment, and NQF member voting, the CSAC decided to proceed with a vote on the measure. The CSAC voted to endorse measure #2158 Medicare Spending Per Beneficiary by a vote of 10-yes; 3-no.

NQF Board of Directors

On December 6, the Board discussed the CSAC recommendations and process to resolve the lack of consensus noted among the councils voting on measure #2158 Medicare Spending Per Beneficiary. The Board discussed many of the concerns previously raised throughout the endorsement process, including attribution, risk adjustment and SES, and appropriate use of the measure. Ultimately, the Board voted to ratify endorsement of measure #2158, noting the value of the process used by CSAC to gain additional input and build consensus.

The robust discussions surrounding the measure led to the recommendation that further work be considered in this area, specifically focused on the appropriateness of including dual eligible populations in risk adjustment models. An expert panel focused on risk adjustment and socioeconomic status will be convened by NQF in early 2014.

Appeals Period

Measure #2158 Medicare Spending Per Beneficiary was posted for a 30-day appeals period. No appeals were received during that time.

Overarching Issues

During the Steering Committee's discussion of the measures, several overarching issues emerged that were factored into the Committee's ratings and recommendations for multiple measures:

Risk adjustment

Socioeconomic Status

The NQF guidance supporting the scientific acceptability criteria for risk adjustment (2b.4) indicates that risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status (SES), or gender. NQF recommends that measures be stratified by race and socioeconomic status rather than adjusting away differences that may be due to disparities in the quality of care provided. Given this guidance, the Committee discussed the appropriateness of including markers of socioeconomic status in the risk model at length in regards to both measures, each with a different approach to accounting for these differences.

During the evaluation of Total Per Capita Cost Measure for Medicare Beneficiaries measure (NQF #2165), the developers described that gender and dual eligibility status are both included in the version of the hierarchical condition category (HCC) risk adjustment model used in this measure. In the importance section of the measure submission, the developers indicated that there is data demonstrating disparities by population group, specifically dual eligible, noting that Medicare spending on dual eligible beneficiaries was almost two times higher than spending on non-dual eligible beneficiaries in 2008[5]. Therefore, including a marker of dual eligibility in the measure's risk adjustment model has the potential of masking this difference.

Given that this modified HCC risk adjustment model was originally developed for Medicare Advantage plans, these demographic factors have been used historically by CMS actuaries to determine payments to these plans. The original intent of the model is to avoid risk selection of patients based on gender and dual eligibility status. The Committee was concerned that given the intended use of the original model it may not be suited for performance measurement where these factors are preferably excluded from risk adjustment to determine whether disparities in care exist. Additionally, concern was raised over whether or not the risk adjustment approach could appropriately adjust for rare conditions. The Committee noted that the HCC model does not include as many diagnostic categories as many commercially available risk adjustment models and therefore may not be as accurate in assigning the appropriate risk categories for rare conditions. However, given the broad use of HCCs across Medicare programs, the Committee agreed that this approach was sufficient for this application.

In the evaluation of the Medicare Spending Per Beneficiary measure (NQF #2158), the measure developers did not include adjustments for dual eligibility. While this measure also uses a version of the HCC risk adjustment model, the developers explained that they tested an exclusion of dual eligible beneficiaries, as well as an inclusion of dual eligible beneficiaries as a risk adjuster, but these adjustments did not result in major differences in the measure performance. Thus, the decision was made for the measure to include dually eligible beneficiaries in the measure population, but not to include a dual eligibility risk adjuster in the developers' version of the HCC risk adjustment model.

The Committee ultimately agreed that more guidance in this area is needed, particularly using markers of SES variables in outcome and resource use measures.

Look back period

In the evaluation of the Medicare Spending Per Beneficiary measure (NQF #2158), some Committee members were concerned with the application of the HCC risk adjustment to capture and identify pre-existing conditions. While the measure identifies these pre-existing conditions by looking back to conditions present in the 90 days prior to admission, the HCC model is designed for a full 12-month look back period. In response to this concern, the developers described their testing of various look back periods and concluded that the 90-day look back period offered marginally superior performance to the 12-month look back period. This is possibly because the conditions that occurred closer to the hospitalization were more relevant to predicting patient severity. The Committee was ultimately satisfied with the response from the developer and the level of testing conducted to justify the 90-day look back period used in the measure.

Exclusion of Deaths

In both measures, patients who died were excluded from the measurement period. This decision was made based on testing showing that this subset of patients has a bimodal distribution of costs caused by a significant number of patients who are high cost and a significant number of patients who are low cost. This distortion in the distribution of data may limit the validity of the model to predict costs for this subset of patients, compared to predicting costs of patients within a normal distribution. The high cost group likely represented those beneficiaries that received high intensity end-of-life care and died toward the end of the measurement period. Conversely, those who died earlier in the episode would show as low cost as they likely used resources for a shorter period of time. The Committee generally disagreed with this exclusion arguing that end-of-life care is a high cost area for the Medicare program and is important for measurement and improvement. Further, excluding patients who die during the measurement period may create unintended negative consequences. For example, hospitals that provide intense care keeping patients alive will appear more costly than hospitals providing equally intense care but ultimately resulting in a patient death because the costs associated with the death have been excluded. Further, in the evaluation of Medicare Spending Per Beneficiary measure (NQF #2158), the Committee questioned the developers on the appropriateness of including hospice costs when deaths are excluded from the measure. This seems counterintuitive as patients entering hospice are expected to die, and thus costs associated with hospice care may be excluded from the measure.

Attribution

In cost measurement, attribution is the step in specifying measures that identifies the responsible entity(s) for the performance results. While similar to the level of analysis, attribution specifically determines what proportion of the costs or resources are assigned to a single provider, divided amongst a group of providers, or some combination thereof. The level of analysis often aligns with the attribution approach, however, often focuses on the lowest level at which the costs can be rolled up and reported (e.g., physician, physician group, state, national). This is often dependent on measure testing to determine stability (or reliability) of the measure results with certain sample sizes.

While NQF seeks to endorse standardized performance measures intended for both accountability and performance improvement that can be used for national comparisons, users of cost and resource use measures often prefer flexibility in the attribution approach to accommodate specific applications, the

unique attributes of their healthcare system or market, and allow the opportunity to consider input from the attributable entities. Further, with no accepted gold standard for attribution or uniform guidance on best practices for attribution, it becomes difficult to determine how best to integrate it into the measure submission and evaluation process while trying to meet various needs. In response to the need for flexibility, under the direction of NQF's first Resource Use Steering Committee in 2010, the attribution approach was allowed to be submitted as measure specifications or as optional guidelines for users to consider when implementing the measure.

In response to some confusion resulting from attempting to decipher specifications versus guidelines in the submission, the attribution approach was included on the submission form as guidelines only in the current resource use project. In both of the measure submissions to this project, however, the attribution approach was included in the submission as specifications key to the implementation of the measure, and they were evaluated as such by the Committee. In an effort to continually evaluate and improve the NQF evaluation process for cost and resource use measures, the current Resource Use Committee was asked to reconsider the implications of requesting the attribution approach as guidelines or specifications. Specifically, they were asked to provide input on the potential for variation in measure results and impact on comparability of resource use measures that are implemented with various attribution approaches. Additionally, recognizing that quality measures also use attribution to determine the responsible entity for performance results that are included in the specifications, the Committee was asked to consider whether resource use measures were sufficiently unique such that the attribution approach would not need to be specified and could continue to be submitted as guidelines enabling user flexibility.

In general, the Committee agreed that the attribution approach should be specified for resource use measures (not allowing guidelines) since allowing flexibility would result in different measure results and has implications for comparability. For example, hospital A implements a measure and chooses to attribute all costs for their patients' episodes to the primary care provider (PCP); using the same measure, hospital B chooses to divide the costs of their patients' episodes by attributing costs among all providers that touched the patient. Comparing PCPs at both hospitals in this scenario would be unfair as the rules for how costs are assigned are different and disproportionate. Further, some argued that the attribution approach must be described clearly in order to understand the context in which providers are measured and the results are computed, emphasizing the need for the application of the measures to align with the intended use. On the other hand, other members pointed out that given the lack of a gold standard for an attribution method, flexibility in the attribution approach would allow for innovation. The Committee ultimately agreed that resource use measures should move to a more standardized approach of requiring the attribution approach to be submitted as a specification to the measure.

Risk Adjustment

In response to requests from various stakeholders and NQF members, NQF was asked to consider the implications of endorsing a single cost/resource use measure that has been tested with multiple risk adjustors. This would enable the measure to be used interchangeably with different risk adjustors based on user need. The need for flexibility in risk adjustors reflects the healthcare market in which different

regions and healthcare systems have invested in a single risk adjustor that may not be one that was used in an endorsed measure of interest. In order to use the endorsed measure (including the risk adjustor), a potential user must weigh the benefits of the measure against the additional investment in another (proprietary) risk adjustor. This introduces a major barrier to the market and the uptake of endorsed measures as organizations often have limited resources; transitioning to another tool is financially inefficient as it introduces new license fees and opportunity costs.

While allowing flexibility would potentially enable markets and users to continue with the risk adjustment model that they have purchased and already have in place, to facilitate national comparisons a single tool must be used. Even with acceptable testing of the measure with each risk adjustor individually, comparability among the measures using the different adjustors is limited. Further, in order to ensure that measure users were not inadvertently comparing measure results from the measure using a different risk adjustor, each of the measure-risk adjustor combinations would need to be endorsed separately and have different endorsement numbers.

Endorsing a single measure with multiple risk adjustors or separate measures (each with different adjustors) presents challenges in applying some of NQF's evaluation criteria and guiding principles for endorsement and national comparisons:

- NQF endorses national standards for performance measures that are intended for both accountability and performance improvement.
- In order to be useful to make conclusions about performance, especially relative performance, all entities need to be measured exactly the same way.
- NQF seeks to endorse the best from among competing measures whenever possible in order to minimize the confusion created when accountable entities are scored and ranked differently based on differences in measure specifications.
- The measure evaluation criteria must be applied to every measure submitted for endorsement. The following criteria are specifically challenging as it relates to multiple risk adjustors:
    - Scientific Acceptability Criteria 2b4 Risk Adjustment/Stratification for Outcome or Resource Use Measures. For resource use measures, an evidence-based risk adjustment strategy (e.g., risk model) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care or the quality of care) and are present at the start of care.
    - Scientific Acceptability Criteria 2b6 Comparability of Multiple Data Sources. If multiple data sources/methods (e.g., risk adjustment approaches) are specified, there is a demonstration that they produce comparable results.
    - Scientific Acceptability Criteria 2b2. Validity testing. NQF criteria allow testing of either the data elements or the measure score.

To facilitate discussion of this issue and fully solicit input from multiple stakeholders, the Resource Use Steering Committee was asked to provide input for consideration of NQF policy on this issue. In addition to the Committee, other stakeholders were invited to participate in this discussion including measure developers, statisticians, purchasers, and other measure users impacted by this issue. Specifically, a member of the Society of Actuaries (SOA) was invited to present to the Committee to discuss the

implications of a 2007 SOA report, *A Comparative Analysis of Claims-Based Tools for Health Risk Assessment*. In this report, it was summarized that several risk adjustors evaluated in this study (e.g., ACGs, ETGs, DxCGs) had comparable performance[6].

Specifically related to the challenges this issue presents to current NQF policy and guidelines, the Committee was asked to consider and provide a rationale for the following questions and potential options for future measure submissions:

- Should NQF consider changes to current policy on this issue and adopt one of the following:

    1. Endorse one measure, with one measure number, including multiple risk adjustment models for the user to pick from.
        - Developers would be required to demonstrate comparability of the results with each adjuster.
    2. Endorse multiple measures with different numbers, each measure with a different risk adjustment model.
        - Assume performance scores are not comparable
        - Each additional measure must be evaluated against criteria
        - All specifications, other than the risk model variables and coefficients should be identical
        - Requires justification of endorsement of multiple competing measures
    3. Endorse one measure, with one measure number, with multiple risk adjustment models
        - Developer would not be required to demonstrate comparability.

In response to the various concerns raised by this issue, the Committee also considered the state of resource use measure development and commercially available risk adjustment methodologies determining that:

1. In order to be useful in making valid conclusions about performance, particularly relative performance, all entities should be measured in the same manner.
2. If a measure is submitted using multiple risk adjustment models, the developers must submit empirical analyses to demonstrate comparability of measure results. These analyses should compare the same patients, distribution of diseases, distribution of cost, as well as measure and compare different risk adjusters. The results should analyze both the differences in relative ranking of providers and the differences in the performance measure results.

More broadly, the Committee raised concern that there may be flaws of risk adjustment systems using claims or administrative data for resource use measurement. Patients who receive higher intensity treatment and thus generate more claims may potentially be assigned to a higher risk category or severity level, resulting in a higher expected cost. The Committee recommended monitoring of unintended negative consequences of this phenomenon.

Measure Specific Issues

During the Steering Committee's discussion of the measures, several issues specific to individual measures emerged.

NQF #2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries

The Committee noted several issues in their discussion of this measure, all linked to validity. The measure's attribution approach and the exclusion of pharmacy costs and Medicare Advantage patients were of specific concern.

Attribution

The attribution methodology chosen for NQF #2165 was strongly questioned by many members of the Steering Committee. While this attribution approach has been used in other CMS programs, including the Physician Group Practice (PGP) demonstration and more recently in the group practice reporting option (GPRO) of the Physician Quality Reporting System (PQRS), the Committee expressed concern about a number of factors. First, the measure includes a two-step attribution rule in which the first step attributes beneficiaries to a medical group with affiliated primary care physicians (PCPs) whose services account for the largest amount of Medicare allowable charges within the measurement period. If beneficiaries are not assigned in the first step, they are assigned to any medical group in which they have seen at least one physician in the group, regardless of specialty, who has provided primary care services. Attribution to the medical group is based on which medical group provided the largest amount of Medicare allowable charges during the measurement period (inclusive of the charges by specialist physicians, nurse practitioners, physician assistants, and clinical nurse specialists). The attribution methodology assigns all healthcare services and associated costs for the beneficiaries to the medical group identified in either step one or step two of the attribution rule.

Members of the Committee expressed strong reservation that visits with non-physician primary care providers, specifically nurse practitioners and physician assistants, are not included as eligible visits for attribution to a provider group in the first stage. The developers described that this approach was guided by statute and is based on CMS' goal to provide feedback on resource use to physicians, but agreed that the inclusion of these providers should be considered in future iterations of this measure given their growing role in primary care. The Committee strongly encourages CMS to include non-physician providers in the first stage of the attribution approach.

Some members of the Steering Committee were also concerned that the attribution method limited the utility of the measure to improve cost performance for two primary reasons. First, even with the reports provided for this measure, there is limited information on the within group variation of costs. Feedback given to provider groups is rolled up to the group level and thus individual provider cost variation may be masked. Second, primary care providers may have limited ability to influence the cost of specialists, inpatient care, and post-acute care and may be ultimately held responsible for these costs. In markets with integrated care delivery networks, this may be expected of primary care; however, the current fragmented state of care delivery does not support this attribution approach. On the other hand, several other committee members stated that this level of accountability for providers is the entire rationale for the measure and should help push providers to be better organized to reduce costs.

Due to the many concerns expressed about this attribution approach, many members ultimately agreed that the approach significantly impacted the validity of the measure and that the accountability for costs should be explored differently to allow for shared accountability across providers.

Outpatient Pharmacy costs

Given that not all Medicare beneficiaries have Medicare Part D coverage, the measure does not include outpatient pharmacy costs. The Committee recognized the limitation in the availability of these data; however, they encouraged the measure developers to consider additional strategies to include these costs in the future. They argued that since more than half of Medicare beneficiaries have Part D coverage, it is important to understand the cost drivers and variation in drug utilization and costs. Members suggested that the developer consider aggregating and reporting the measure by those who have Medicare Part D coverage and those who do not.

Exclusion of Medicare Advantage

Committee members were concerned about the exclusion of Medicare Advantage patients from the measure. Medicare Advantage Plans, also known as Medicare Part C, are health plans offered by private companies approved by Medicare. Members of the Committee argued that measuring cost for beneficiaries in the Medicare Advantage plans is equally as important as measuring cost for the Medicare fee-for-service population. The measure developers argued that it is often difficult to obtain utilization data for beneficiaries enrolled in Medicare Advantage Plans; as such, these beneficiaries were not included in the measure. There was concern that there could be gaming with large, sophisticated practices encouraging higher cost fee-for-service beneficiaries to switch to Medicare Advantage plans, enabling the practice to continue seeing these patients without inclusion of their costs in the measure.

The Committee requested that the developers reconsider this exclusion. Many members agreed that this issue would not ultimately influence the final endorsement recommendation but does challenge the validity of the measure as constructed.

Harmonization Discussion

The Steering Committee considered potential harmonization issues between the NQF #2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries and the previously endorsed NQF #1598 Total Resource Use Population-based PMPM Index developed by HealthPartners. A summary comparison of these measures is captured in Appendix D. The goal of this harmonization effort is to reduce measurement burden for providers and implementers, while improving interpretability for patients and facilitating alignment of measurement across public and private sector. In its preliminary recommendation at the in-person meeting, the Steering Committee did not recommend endorsement of Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries (#2165), but agreed to discuss areas of potential harmonization between the two measures.

The Steering Committee reviewed areas of conceptual (intent of the measure) and technical (how the intent of the measure is operationalized) similarities and differences between the two per capita measures. Prior to the Committee's discussion of potential harmonization areas between the two measures, NQF staff facilitated early discussions with the developers of each measure to identify

possible areas of alignment. The developers were asked to submit a joint letter to the Committee outlining areas of potential alignment and key differences. Upon review, the Committee considered how the two measures would provide consistent measure results by improved interpretability across levels of analysis and data sources.

Similar resource use measures are defined as the following:

- same measure types (e.g., per episode, per capita),
- measure the same costs/resources (e.g., actual cost vs. standard prices, resource service categories), and
- address the same population (e.g., diabetic patients).

NQF #2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries and NQF #1598 Total Resource Use Population-based PMPM Index are both per capita, non condition-specific measures that capture standard prices; however, NQF #2165 addresses the Medicare population, and NQF #1598 addresses the commercially insured population. The measures are both risk adjusted; however, NQF #1598 Total Resource Use Population-based PMPM Index uses a commercial risk adjustment methodology developed and calibrated specifically for the commercially insured population (Johns Hopkins University's Adjusted Clinical Groups (ACG) Case Mix System) whereas NQF #2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries uses the CMS Hierarchical Condition Category (CMS-HCC) risk adjustment methodology designed for Medicare FFS beneficiaries.

While evaluating these similar resource use measures for potential harmonization, the Steering Committee was asked to assess both the value and the burden of recommending that the measures not be harmonized and remain distinct. Specifically, the Committee reviewed whether the differences in the technical specifications were necessary, whether they affected interpretability across the measures, and the effect on the burden of data collection. The Committee considered whether the measures had both sufficiently different populations and standardized costing approaches to justify the burden of having two similar measures. The Steering Committee was not asked to review the measures for harmonization of risk adjustment models, risk stratification approaches, and statistical methods for estimating measure results, as this is not recommended under NQF guidance.

The Committee reviewed the key differences between the measures and agreed that there was little room for increased alignment between the measures given the unique characteristics of the two target populations and measure intent. The Committee stated that the differences in the data sources resulting from the differences in the target populations for the two measures drive the differences in the technical specifications for the measures, including risk adjustment methodologies. Given the different patient populations, the Committee discussed the challenges to align the risk adjustment methods and the payment standardization methodologies. Some members of the Committee suggested that the developers consider potential harmonization of their attribution approach. They discussed that providers could better interpret how their patients are assigned to them if the attribution approach is similar for their Medicare and commercial patients. The Committee also discussed differences of pharmacy data; the HealthPartners measure includes pharmacy data when available and the CMS

measure does not. Members of the Committee recommended that CMS consider experience from commercial payers in handling missing pharmacy data. For example, it may be possible to calculate the measure for people who have a pharmacy benefit and those without to create a blended per-member per-month measure result.

NQF #2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)

Exclusion of Transfers

In this measure all beneficiaries who are transferred are excluded from the measure; the Committee discussed the appropriateness of this exclusion at length. The developers explained that during their public comment on the measure, community hospitals argued that they should not be responsible for patients whom they stabilize and transfer to another facility. Facilities that receive transfers argued that they should not be responsible for care that was provided prior to the patients entering their facility. To account for both perspectives, the developer chose to exclude all transfers from the measure. The Committee noted that hospitals are increasingly responsible for care delivered up to 30 days after discharge; thus they agreed that hospitals should be responsible for the utilization and associated costs for patients that they transfer to other facilities. The developer acknowledged that it was challenging to address the various perspectives on attribution of transfers but agreed to reconsider the specification based on the Committee's feedback.

Recommendations for Future Measure Development

The Committee identified opportunities for linking cost and resource use measures with quality measures in order to better understand efficiency, as well as areas for future cost and resource use measure development.

In order to understand efficiency, cost and resource use measures should be linked with:

- appropriateness/overuse measures
- outcome measures
- process measures
- clinical data and patient reported outcomes

Other gaps noted included:

- Measures capturing variations in cost and outcomes for potentially high cost patients (e.g., cardiovascular or diabetes patients)
- Episode-based cost and resource use measures for high impact conditions and procedures
- Measures capturing actual prices paid to providers by health plans

Measure Evaluation Summary

Endorsed measure

Measure not endorsed

Endorsed Measure

#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)

Steward: Centers for Medicare and Medicaid Services

Description: The MSPB Measure assesses the cost of services performed by hospitals and other healthcare providers during an MSPB hospitalization episode, which comprises the period immediately prior to, during, and following a patient's hospital stay. Beneficiary populations eligible for the MSPB calculation include Medicare beneficiaries enrolled in Medicare Parts A and B who were discharged from short-term acute hospitals during the period of performance.

Resource Use Measure Type: Per episode

Data Source: Administrative claims

Level of Analysis: Facility

Costing Method: Standardized pricing

Target Population: Senior Care

Resource Use Service Categories: Inpatient services: Inpatient facility services; Inpatient services: Evaluation and management; Inpatient services: Procedures and surgeries; Inpatient services: Imaging and diagnostic; Inpatient services: Lab services; Inpatient services: Admissions/discharges; Ambulatory services: Outpatient facility services; Ambulatory services: Emergency Department; Ambulatory services: Evaluation and management; Ambulatory services: Procedures and surgeries; Ambulatory services: Imaging and diagnostic; Ambulatory services: Lab services; Durable Medical Equipment (DME)

STEERING COMMITTEE MEETING [May 8-9, 2013]

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

(1a. High Impact: 1b. Performance Gap, 1c. Measure Intent)

1a. Impact: **H-23**; **M-2**; **L-0**; **I-0** 1b. Performance Gap: **H-12**; **M-12**; **L-1**; **I-0** 1c. Measure Intent: **Y-6**; **N-16**; **I-3**; **L-0**

1. Overall: **H-9**; **M-15**; **L-1**; **I-0**

<u>Rationale:</u> While evaluating the measure's importance to measure and report, the Committee agreed that the subcriteria were met and provided the following rationale:

- General agreement that healthcare cost is a high impact area of healthcare.
- Affordability of healthcare has been identified as an area of focus as part of the Triple Aim and under the National Quality Strategy.
- Inpatient costs are a major driver of total costs; capturing this may incentivize hospitals to examine causes of these expenditures.
    - Readmissions and Skilled Nursing Facility costs will be significant drivers of cost captured through this measure; these are high impact areas where Medicare spends the most money with respect to hospitalizations.
- Though the developers stated that a benefit of the measure would be to improve care coordination, the Steering Committee did not agree that the evidence submitted substantiated this claim.
- Though the measure was described as a cost measure, the Steering Committee clarified that this is a Medicare expenditure measure, which can be used as a proxy for cost.

#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)

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-10**; **M-14**; **L-1**; **I-0** 2b. Validity: **H-0**; **M-13**; **L-11**; **I-1**

<u>Rationale:</u> While evaluating the measure's scientific acceptability, the Committee agreed that the subcriteria were met and identified 5 major issues:

1) Reliability concerns relating to 30% of hospitals moving quintiles as demonstrated in the test, re-test results
2) Validity concerns relating to the exclusion of deaths and transfers
3) Concern regarding the construct validity testing results which demonstrated low correlation with measures of readmissions in heart attack, heart failure, and pneumonia
4) Concern regarding the look back period for the HCC risk adjustment model
5) Concern regarding the appropriateness of not incorporating the dual eligible population into the risk adjustment model

1) Reliability concerns relating to 30% of hospitals moving quintiles as demonstrated in the test, re-test results
   - Many Committee members expressed concern that the test, re-test results demonstrated that approximately 30% of hospitals in the lowest-spending quintile in one sample were not in the lowest-spending quintile in the next sample; similarly, approximately 30% of hospitals in the highest-spending quintile in one sample were not in the highest-spending quintile in the next sample.
   - Committee members questioned whether this level of reliability would be sufficient, particularly with respect to establishment of cutoff thresholds when the measure is reported.
   - The developer stated that Spearman rank correlation for a hospital across samples is 0.835, demonstrating a linear relationship between the rank of the hospitals in the test and re-test samples. This indicates that using a different random group of patients does not result in significant variation of the hospital's relative performance.

2) Validity concerns relating to the exclusion of deaths and transfers
   - Several Committee members expressed concern that the exclusion of deaths and transfer patients from the measure is unnecessary.
     - Exclusion of deaths removes from the measure calculation some of the patients who use the highest resources and thus are the most expensive. Additionally, the Committee questioned the rationale for inclusion of hospice costs when deaths are excluded.
     - Exclusion of transfer patients accounts for approximately 5% of patients, and the rationale for excluding them is unclear. The Committee members stated that, given that the measure holds the hospital accountable for patients 30 days after discharge, it isn't clear why transfers are excluded. Additionally, the Committee members stated that exclusion of transfers may result in gaming of the measure, as hospitals may simply transfer high cost patients.
   - The developer stated that deaths were excluded because of the bimodal distribution of costs, with an average cost for patients who die 40% higher than those patients who do not die. However, many episodes cost far under what was predicted, potentially because the patient died early in the episode and thus did not utilize resources.
   - The developer stated that transfer patients were excluded because of difficulties with attributing the patients to a hospital.
   - Several Committee members stated concern that the rationale provided by the developer for excluding deaths and transfers was insufficient and suggested the measure developer consider updating the measure to address this concern.

#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)

3) Concern regarding the construct validity testing results which demonstrated low correlation with measures of readmissions in heart attack, heart failure, and pneumonia
   - Several Committee members stated that high correlation between the MSPB measure and a measure capturing readmissions is expected because of the high cost of readmissions for these diseases.
   - The Committee stated concern that the testing results for the measure demonstrated weak correlations with the readmissions measures (0.08, 0.07, and 0.06 for heart attack, heart failure, and pneumonia readmission rates respectively), particularly because the developers used this to demonstrate validity of the measure.
   - The developer speculated that the weak correlation resulted from the fact that the MSPB measure assesses the cost to Medicare of all services performed by hospitals and other healthcare providers during an MSPB episode; as a result, a hospital's MSPB measure value is driven by both acute and post-acute spending.
   - Several Committee members stated that the rationale provided by the developer on why spending and readmissions should be correlated needs to be substantiated by further testing, as these results provided demonstrated weak validity of the measure.
   - The developer also submitted validity testing results for the 30-day MSPB post-discharge window, demonstrating a positive correlation (0.13) between MSPB measure values and the percent of beneficiaries with multiple episodes. This analysis was intended to demonstrate that the measure is sensitive to the length of the 30-day post discharge window. The analysis aimed to analyze whether hospitals whose beneficiaries incurred multiple 30-day episodes performed better on the measure by virtue of the beneficiaries' care having been split into more episodes that were less expensive individually. The analysis, however, found that high cost hospitals are more likely to have beneficiaries with multiple episodes. This indicates that the 30-day window is not strongly affecting the measure.
     - Additionally, the developer further explored the validity of the 30-day MSPB post-discharge window by testing rank correlation against a 90-day window. The developer found a positive rank correlation (0.897), suggesting that hospitals with high MSPB measures using the 30-day window also had high MSPB measures using the 90-day window.

4) Concern regarding the look back period for the HCC risk adjustment model
   - Several Committee members stated the concern that the HCC risk adjustment model only captures health status variables derived from claims during the 90 days prior to the start of an episode. Committee members stated that accuracy of the HCC model drops off dramatically with less than 7 months of data; 12 months of data is the gold standard.
   - The developer stated that testing was done to evaluate the health status variables in the risk adjustment model by using one year of data prior to the start of an episode rather than 90 days. The developer found that 6% of episodes are dropped, and the R-squared value actually decreases from 0.4621 (90 days data) to 0.4601 (one year data). Summarized, the developer found that capturing 90 days of data rather than one year of data resulted in no significant trade-off between the number of episodes included and the model fit.

5) Concern regarding the appropriateness of not incorporating the dual eligible population into the risk adjustment model
   - Steering Committee members voiced opinions on both sides of this issue, with some stating that dual eligible patients should be included in the risk adjustment model and others stating that they should not.
     - Those in favor of including dual eligible patients in the model stated concern that a potentially significant unintended consequence of not including dual eligible patients in the risk adjustment model would be the refusal of hospitals to accept dual eligible patients, as they are known to be higher cost than traditional Medicare patients.
     - Those opposed to including dual eligible patients in the model stated concern that adjusting for dual eligible status would mask any disparities in the cost of care for these patients.

#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)

- The developer stated that although dual eligible patients are included in the measure population, a dual eligible risk adjuster is not currently included in the risk adjustment model.
- A commenter from the public stated that dual eligible patients share characteristics beyond socioeconomic status, such as multiple chronic conditions, complex societal issues, and disparities in healthcare literacy. They are a population with chronic, complex disease that needs to be accounted for, particularly as relates to the impact on Safety Net hospitals within the context of this measure.

Additional Issues:

- *MS-DRG Regression.* Using Medicare Severity - Diagnostic Related Groups (MS-DRG) variables in the regression has the potential to mask variation attributable to quality of care, as patients can be bumped into a higher DRG through comorbidities or complications. The Committee questioned whether the standardized core DRG had been subtracted before the regression to see how much variance in the rest of the payments are explained by the other health status variables included in the risk adjustment model. The developer is willing to do this analysis.
- *Fiscal year payment rates.* Measure uses payment rates at the time of the claim (for the relevant fiscal year); potential for bias exists if admission rates vary significantly between fiscal years between hospitals.
- *Pre- and post-hospitalization services.* Several Committee members stated the concern that the major sources of variation between hospitals after risk adjustment are the pre-hospitalization and the post-hospitalization care; the Committee questioned whether the measure allowed for understanding of which sets of post-acute services result in higher cost when the measure is calculated.

3. Feasibility: H-23; **M-1**; **L-0**; **I-0**

(3a. Byproduct of Care Processes; and 3b. Electronic sources; and 3c. Data Collection Strategy)

Rationale: While evaluating the measure's feasibility, the Committee agreed that the subcriteria were met and provided the following rationale:

- Data for the measure is being collected and is available.
- Data is generated electronically.
- The Committee generally agreed that the measure is feasible to implement.

4. Usability: H-6; M-15; L-3; I-0

(4a. Accountability/transparency (used in accountability w/in 3 yr, public reporting w/in 6 yr, or if new - credible plan); and 4b. Improvement – progress demonstrated (if new - credible rationale); and 4c. Unintended Consequences - benefits outweigh evidence of unintended negative consequences (to patients/populations); and 4d. Measure Deconstruction – can be deconstructed to facilitate transparency and understanding)

Rationale: While evaluating the measure's usability and use, the Committee agreed that the subcriteria were met and provided the following rationale:

- The Committee largely agreed that the measure will be most useful when paired with quality outcome measures.
- The measure is in use for accountability purposes.
- The Committee expressed concern that many hospitals may not have the analytic capacity to understand the data and understand the impact of care outside of the hospitalization on the measure result.
    - The Committee recommended that the reports from CMS provide hospitals with analysis to allow hospitals to identify cost drivers outside of the hospitalization.
    - The Committee recommended that CMS provide the hospitals with information on which post-acute care providers are using the most resources, so that hospitals can partner with providers who are utilizing fewer resources and providing quality care.
- The developer stated that hospitals are provided with several different files to understand costs, including hospital-specific reports on its performance on the MSPB measure and patient-level data. The reports also provide comparison of a hospital's performance compared to other hospitals in the same state or across the nation, and provide a breakdown of spending by claim type.

#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)
• From a consumer's perspective, the small variation in performance will make it difficult for the consumer to distinguish the best performers. The data is presented in a way that may be challenging for a consumer to deconstruct. • The developer stated that downloadable files are available online which will provide more detail on the measure results for consumers. Unintended Consequences: While evaluating the measure's usability, the Committee identified the following potential unintended consequences: • Consumers may choose the most expensive hospital, believing that increased cost corresponds to higher quality healthcare. • Hospitals may transfer patients based on expected high expenditures post-discharge, resulting in the patient being excluded from the measure.
5. Related and Competing Measures • No related or competing measures noted.
Steering Committee Recommendation for Endorsement: Y-15; **N-10** Rationale: • The measure focus is high impact; healthcare costs in the United States are very high. • The measure allows hospitals to begin looking at and understanding cost; when paired with quality outcome measures, this will help hospitals gain an understanding of the value and efficiency of healthcare services provided.
Public and Member Comment [July 9 – August 7, 2013] **Table of Comments and Responses** This measure received comments from twenty organizations/ individuals. Three of these comments were supportive, noting that this measure is an important first step "towards an optimal measure of hospital resource use." One commenter noted that the measure "does have methodologic concerns but its intent is clear and necessary." The remaining comments addressed several themes listed below, along with a description of the comments received. **Exclusion of Deaths** • Two commenters questioned the exclusion of deaths, noting that they "believe the measure would be a stronger measure of costs if patient deaths were included." • One commenter supported the exclusion of deaths and also called for the exclusion of hospice payments in order to "maintain the internal consistency of the measure." **Steering Committee Response:** The Resource Use Steering Committee generally agreed with the Commenter that the inclusion of episodes where the patient dies would create a stronger measure. End of life care is a high-cost area for the Medicare program and is important for measurement and improvement. The developer acknowledged that the exclusion was finalized through notice and comment rulemaking based on the fact that these are incomplete episodes where significant data could be missing, but that CMS agreed it would consider including episodes in which the beneficiary dies in future updates to the MSPB measure. **Exclusions** • One commenter expressed concern that exclusion of transfer patients from other acute care facilities may affect a larger portion of PPS-exempt Cancer Center patient admissions when compared to PPS Hospital admissions.

- One commenter expressed concern that exclusion of transfer patients could remove more seriously ill patients, which represent significant opportunities for reduced spending.
- One commenter stated that inclusion of Medicare Part D data would result in a stronger measure.

Steering Committee Response:
The Committee agreed that additional analysis would need to be conducted to determine the transferability of validity results to a cancer patient population. As specified, the measure currently excludes cancer hospitals.

Furthermore, additional analysis on risk adjustment approach specific to PPS-Exempt Cancer Centers' patient population and the 90-day look back period would need to be conducted before the measure was specified for a cancer patient population.

The Resource Use Steering Committee generally agreed with the commenter that facilities being held responsible for the utilization and associated costs for patients that they transfer to other facilities would foster better collaboration resulting in more efficient and effective care. This collaboration fits with the philosophy of holding a facility responsible for care delivered up to 30 days post discharge.

The Committee agreed that inclusion of Part D data would create a stronger measure. They recognized the limitation in the availability of these data; however, they encouraged the measure developers to consider additional strategies to include these costs in the future.

Attribution

- Commenters cautioned that this measure is only suitable for reporting at the facility level and should not be analyzed or reported at the individual clinician level. Commenters stated concern that the measure has not been tested or specified for this analysis at the individual clinician level.
- Commenters also agreed with the Steering Committee recommendation that this measure be reported with quality measures, in order to provide meaningful information about the efficiency of health care delivery.

Steering Committee Response:
The Steering Committee unanimously agreed that cost and resource use measures must be paired with quality measures in order to understand and make decisions about care. The Committee agreed with the commenter that measures of efficiency, and ultimately value are critical tools needed to improve the efficiency of US health care system, specifically encouraging shared accountability and team-based care.

Measures endorsed by NQF are only endorsed for use at the specified level of analysis that the measure developer has provided testing for; in this case, that would be at the facility level. The Steering Committee has only recommended this measure for endorsement for analysis at the facility level.

Risk Adjustment

- Several commenters stated concern that the risk adjustment methodology was not valid for the following reasons:
    - Lack of a socio-economic status (SES) adjustment (i.e. dual-eligible status).
        - Comment #3249: CMS's analysis demonstrates that dual-eligible patients have $859 more spending per episode than other patients. The agency finds that including patient dual-eligible status as a risk adjuster marginally improved the fit (R-squared value) of the risk adjustment model. But, the same analysis also demonstrates that about 12% of hospitals would have their MSPB measure values change by more than 1 percentage point if dual-eligible status were included in the risk adjustment model. About 10.8% of hospital scores would decrease by between 1 and 3

percentage points. Nevertheless, CMS chose to not include a dual-eligible adjustment in the measure.

- Testing results demonstrating clustering of large, urban, teaching hospitals that treat a large proportion of low income patients with higher MSPB index rates than their community hospital counterparts, possibly due to the risk adjustment not accounting for the ranges of patient complexity that exist between and within MS-DRGs or that case mix is driving he differences in measure score.
  - Comment #3258: The actual results of the MSPB suggest that the case mix adjustment isn't working properly. In Minnesota, for example, hospitals in urban areas have similar scores, clustering around .93. In Greater Minnesota, however, the scores are almost all less than .88. Because there are large differences in the types of conditions treated by urban and rural hospitals, it raises a concern that the case mix is driving the differences vs. actual differences in adjusted resource use.
  - Comment #3261: We thank NQF for the opportunity to comment on the cost and resource use measure, #2158 Medicare Spending Per Beneficiary (MSPB). We have several areas of concern with this measure that should be addressed prior to endorsement. In the MSPB results that have been published by the measure developer, there is a noticeable clustering of large, urban, teaching hospitals that treat a large proportion of low-income patients with higher MSPB index rates than their community hospital counterparts. We believe this is due to insufficient severity adjustment in the measure that does not account for the ranges of patient complexity that exist between and within MS-DRGs. Since large, urban, teaching hospitals, with a large share of low income patients have the capability to treat more complex patients, whereas community hospitals often do not, they have a higher proportion of complex cases that require more hospital resources and are also more likely to have home care or skilled nursing care following the inpatient admission. This mix of more complex patients could be a contributing factor to the clustering being seen in the results. We would expect a more normal distribution of MSPB results across all hospitals if the measure were appropriately severity/risk adjusted and adjusted for outliers.
- Risk stratification using MDC criteria alone is inadequate and will introduce significant variability in the MSPB rating based upon patient-specific and diagnosis-specific factors that are not adequately encompassed in the MDC classification.
  - Comment #3271: NASS is concerned that the risk stratification for MSPB does not have adequate granularity to differentiate significant cost drivers. Specifically, the proposal to stratify cases by major diagnostic category (MDC).There is significant evidence that MDC classification does not accurately encompass the factors that contribute to cost of care, and there are significant inaccuracies in the administrative data that contributes to the MDC Classification. Risk stratification using MDC criteria alone is inadequate and will introduce significant variability in the MSPB rating based upon patient-specific and diagnosis-specific factors that are not adequately encompassed in the MDC classification. A more comprehensive classification that includes an algorithm that includes CPT codes and procedure specific information would be more useful than a stratification based upon MDC alone.
- Concern that the 90-day look-back period to capture a patient's comorbidities in order to determine the HCC score is insufficient.

#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)

Steering Committee Response:
The Committee reviewed the concern around the clustering of large, urban, teaching hospitals that treat a large proportion of low-income patients raised by the commenters. They voiced concerns on both sides of the issue of including SES adjustment with some members agreeing with the Commenter that disadvantaged patients with multiple complex conditions will require more resources to treat, while other members argued that including SES variables in the risk adjustment model will mask disparities in cost performance among different groups of patients. The Committee acknowledged that hospitals are legitimately held accountable for taking appropriate care of patients within the case mixes and making sure that these patients receive appropriate post-acute care, however, the availability of these support services will vary from community to community.

The Committee recommended that additional work be considered in this area, specifically the appropriateness of including dual-eligibility in risk adjustment models for resource use measurement.
The Committee considered the major diagnostic category (MDC) risk stratification criteria, specifically applying the risk adjustment within MDC to be generally appropriate for this application. For the purposes of performance measurement, factors that are included in the risk adjustment model should be present at the start of care – thus including procedure codes that occur during the measurement period would not be appropriate.

The Committee also expressed concern over the 90-day look back period but ultimately agreed that the performance of the models did have a slightly improved model fit over the models with a year of look-back.

Final Steering Committee Recommendation for Endorsement (9/5/2013): Y-17; N-8

Consensus Standards Approval Committee (CSAC) Vote (11/6/2013): Y–10; N–3

NQF Board of Directors: Ratified endorsement of the measure on December 6, 2013

Measure Not Endorsed

#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries

Steward: Centers for Medicare & Medicaid Services

Description: The Payment-Standardized Total Per Capita Cost Measure for Medicare FFS Beneficiaries assesses the per capita (per beneficiary) cost of health care services for Medicare FFS beneficiaries enrolled in Parts A and B and attributed to medical group practices. The measure includes all Medicare Part A and Part B costs during a calendar year and is payment-standardized and risk-adjusted (using patient demographics and medical conditions) to account for any potential differences in costs among providers that result from circumstances beyond the physician's control. Under CMS' attribution rule, beneficiaries are attributed on the basis of the plurality of primary care services, to those medical group practices with the greatest potential to influence the quality and cost of care delivered to Medicare FFS beneficiaries.

Resource Use Measure Type: Per capita (population- or patient-based)

Data Source: Administrative claims

Level of Analysis: Clinician : Group/Practice

Costing Method: Standardized pricing

#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries

Target Population: Senior Care

Resource Use Service Categories: Inpatient services: Inpatient facility services; Inpatient services: Evaluation and management; Inpatient services: Procedures and surgeries; Inpatient services: Imaging and diagnostic; Inpatient services: Lab services; Inpatient services: Admissions/discharges; Inpatient services: Labor (hours, FTE, etc.); Ambulatory services: Outpatient facility services; Ambulatory services: Emergency Department; Ambulatory services: Evaluation and management; Ambulatory services: Procedures and surgeries; Ambulatory services: Imaging and diagnostic; Ambulatory services: Lab services; Durable Medical Equipment (DME); Other services not listed

STEERING COMMITTEE MEETING [May 8-9, 2013]

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

(1a. High Impact: 1b. Performance Gap, 1c. Measure Intent)

1a. Impact: **H-20**; **M-2**; **L-2**; **I-0**; 1b. Performance Gap: **H-11**; **M-10**; **L-4**; **I-0**; 1c. Measure Intent: **H-8**; **M-13**; **L-4**; **I-0**; 1. Overall: **H-11**; **M-10**; **L-4**; **I-0**

<u>Rationale:</u> While evaluating the measure's importance to measure and report, the Committee agreed that the subcriteria were met and provided the following rationale:

- There was general agreement that this represents a high impact area of healthcare.
- The Committee was concerned, however, that the results of the measure may not be actionable because of the attribution method.
- The measure does not present a consistent breakdown of disparities (race, dual eligible status, etc.).
- The inclusions of pharmacy costs would present a more accurate picture of costs.
- The measure applies to 7,000 groups across the country and covers 75% of physicians. This represents a majority of physicians but a minority of groups. This measure would therefore benefit large groups with a value modifier.
    - After the Steering Committee meeting, the developer clarified that the measure covers 45% of physicians, not 75% as stated during the in-person meeting.

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-5**; **M-18**; **L-1**; **I-0**; 2b. Validity: **H-0**; **M-13**; **L-12**; **I-0**

Rationale: While evaluating the measure's scientific acceptability, the Committee agreed that the subcriteria were met and identified 3 major issues:

1) Attribution Method
2) Risk Adjustment Model
3) Exclusions

1) Attribution Method
    - The Committee was concerned about the general construction of the attribution approach.
    - Stage 1 of the attribution model assigns patients to physician groups by looking at number of visits with a primary care physician. The first stage of the attribution model does not consider the number of visits with Physician Assistants (PA) or Nurse Practitioners (NP). The lack of consideration of PA and NP visits was questioned, as PAs and NPs increasingly deliver more primary care.
    - The developer responded that this measure was designed according to requirements in statue to

capture per capita costs for services delivered by physicians; thus physicians serve as the entry point to the attribution model.

2) Risk Adjustment Model

- The Committee expressed concern about the inclusion of dual-eligible status and gender in the risk adjustment model. The developer responded that the model was originally developed for the Medicare Advantage program and not necessarily for this measure.
- Questions arose over whether the inclusion of SES and demographic factors could obscure the identification of disparities in care.
- The Committee found the Hierarchical Condition Category (HCC) risk adjustment methodology with demographic factor adjustments to be weak in this application.

3) Exclusions

- The Committee questioned the exclusion of deaths, part-year beneficiaries, and Medicare Advantage beneficiaries; these areas represent significant opportunities for improvement in reducing spending.
- The Committee was concerned that excluding patients with Medicare Advantage presented a significant opportunity for "gaming" of the measure. High cost patients could be shifted to Medicare Advantage to prevent costs from being captured and attributed to the practice.
- The developer stated that some physician stakeholders did not agree that the inclusion of part-year beneficiaries was a fair representation of the cost of care for their patients.

Other issues

- The developer calculated a reliability score by measuring the between medical group variance compared to within medical group variance. The Committee expressed concern regarding these reliability testing results which showed that for medical group practices with at least 25 EPs and 20 attributed beneficiaries, the average reliability was 0.95, and 99 percent of groups had a reliability exceeding 0.50, and 96 percent of groups had a reliability exceeding 0.70.
- The Committee was concerned that the use of Tax ID numbers (TIN) may not be an accurate way to identify physician groups. Several small groups may bill under the same TIN giving the appearance of a larger group for the purposes of this measure. The developer agreed that this may be a legitimate concern; however, because the TIN is the unit of payment, it is still a legitimate method to aggregate costs.

3. Feasibility: H-19; **M-5**; **L-1**; **I-0**

(3a. Byproduct of Care Processes; and 3b. Electronic sources; and 3c. Data Collection Strategy)

Rationale: While evaluating the measure's feasibility, the Committee agreed that the subcriteria were met and provided the following rationale:

- The data for the measure is being collected and is a byproduct of the care process.
- Data is generated electronically
- Providers are not able to implement this measure without CMS. Commitment must be made from those with the data to make it publicly available.

4. Usability: H-4; M-14; L-7; I-0

(4a. Accountability/transparency (used in accountability w/in 3 yr, public reporting w/in 6 yr, or if new - credible plan); and 4b. Improvement – progress demonstrated (if new - credible rationale); and 4c. Unintended Consequences - benefits outweigh evidence of unintended negative consequences (to patients/populations); and 4d. Measure Deconstruction – can be deconstructed to facilitate transparency and understanding)

#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries

Rationale: While evaluating the measure's usability, the Committee agreed that the subcriteria were met and provided the following rationale:

- The Committee largely agreed that the measure will be most useful when paired with quality outcome measures.
- This measure can drive change by placing primary care physicians as the responsible entity.
- Groups are in the best position to impact coordination of care and affect the access of care by the individual.
- Significant variation within groups can be masked by group-level reporting; physician-level reporting would eliminate that masking, but presents its own challenges.
- Consumers/Purchasers would find physician-level reporting to be the most actionable.

5. Related and Competing Measures

Potential harmonization issues relating to #1598 Total Resource Use Population-based PMPM Index (HealthPartners) were discussed by the Committee:

- The Committee reviewed areas of conceptual and technical similarities and differences between the two measures, noting that both measures are per capita, non-condition specific and capture standardized prices; however, the measures address different but overlapping target populations. NQF#1598 addresses the commercially insured population, and NQF#2165 addresses the Medicare population.
- The Committee considered whether the differences in target population and the differences in approach to standardization of prices were sufficient to justify recommending that the two measures not be harmonized and remain distinct. As part of this discussion, the Committee considered the potential value and burden for this; specifically, whether the differences in the technical specifications are necessary, affect interpretability across the measures, or affect data collection burden.
- The Committee reviewed the key differences between the measures and agreed that there was little room for increased alignment between the measures given the unique characteristics of the two target populations and measure intent. The Committee stated that the differences in the data sources resulting from the differences in the target populations for the two measures drive the differences in the technical specifications for the measures.
- Some members of the Committee suggested that the developers consider potential harmonization of their attribution approach. They discussed that providers could better interpret how their patients are assigned to them if the attribution approach is similar for their Medicare and commercial patients.
- The Committee also discussed differences of pharmacy data; the HealthPartners measure does include pharmacy data when available and the CMS measure does not. Members of the Committee recommended that CMS consider experience from commercial payers in handling missing pharmacy data.

Steering Committee Recommendation for Endorsement: Y-11; N-14

Rationale:

- The Committee was concerned about the construction of the measure and the ability of the attribution approach to capture costs appropriately and assign them to appropriate providers.
- The exclusion of Medicare Advantage, part-year beneficiaries, Part D, and deaths limit the utility of the measure to address high-cost, high-priority areas of healthcare.
- Reporting at the group level may not provide actionable information and mask significant intra-group variation.
- The Committee did not reach consensus on this measure. The Committee considered this vote "preliminary" and will likely reconsider after the developer's responses and public and member comments have been reviewed and discussed.

#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries

Public and Member Comment [July 9 – August 7, 2013]

Table of Comments and Responses

This measure received comments from eighteen organizations/ individuals. Several commenters shared support for the concept and intent of the measure, urging CMS to make revisions to the attribution approach, risk adjustment algorithm, reliability and validity of the measure and to bring the measure back to NQF for endorsement, as this is an area where measures are needed and would provide insight into the costs of healthcare to Medicare. In addition to the support for the concept and intent of the measure, one commenter also acknowledged "provider concerns over the attribution of total cost of care to primary care physicians who are not part of an organized health system. But purchasers have come to expect care to be coordinated among providers and see the need to incentivize such coordination. Moreover, measures such as this one will help primary care physicians to understand the cost implications of their referral recommendations."

Steering Committee Response:

The Steering Committee raised several concerns with the construct of the measure, which the developer has been working to analyze and address during the past few months. Responses to the committee's concerns and additional analysis performed by the developer were shared with the committee on their August 28th call. The Steering Committee had the opportunity to review all comments and the developer's analysis and re-affirmed their decision to not recommend the measure for endorsement.

NQF stated that they will work with the developers to determine when the Total Per Capita Cost measure can next be reviewed for endorsement. This would happen when the next Cost and Resource Use project is scheduled.

The Steering Committee unanimously agreed that cost and resource use measures must be paired with quality measures in order to understand and make decisions about care. The Committee agreed with the commenters that measures of efficiency, and ultimately value are critical tools needed to improve the efficiency of US health care system, specifically encouraging shared accountability and team-based care.

The Steering Committee acknowledged the consumer perspective that care should be coordinated among providers; however, the Steering Committee was split over the idea that it may be inappropriate to hold primary care providers accountable for the cost of care provided to patients by other specialists, through inpatient care or through post-acute care, as primary care providers have limited ability to control these costs. In the current state of care delivery, health care is accessed in many ways. Many patients select their own primary and specialty care physicians, making decisions to see providers on their own, without coordination with their PCP or PCP group. Several Committee members stated that this may be appropriate in markets with integrated care delivery networks or where patients identify with a PCP or PCP group voluntarily or by assignment; however, in the current fragmented state of care delivery this attribution approach is not preferred. Several other committee members stated that this level of accountability for providers is the entire rationale for the measure and should help push providers to be better organized to reduce costs.

The remaining comments addressed several themes listed below, along with a description of the comments received.

Attribution

- Several commenters agreed with the Committee that primary care physicians or specialists who may be attributed patients because they provided primary care services to that patient have limited ability to control the cost of care provided to patients by other specialists, through inpatient care or through post-acute care. The majority of commenters agreed that it may be inappropriate to hold these providers accountable for these costs of care. The commenters also agreed that this may be appropriate in markets with integrated care delivery networks; however, in the current fragmented state of care delivery this attribution approach is not supported.

- Additionally, several commenters shared the Steering Committee concerns that patients and their associated costs may potentially be attributed to specialists who provide primary care services that are Medicare allowable charges and questioned the appropriateness of this.
- Several commenters shared the Steering Committee concern that visits with non-physician providers (PAs and NPs) are not taken into account in the attribution model until the second stage, as non-physician providers are increasingly delivering more primary care.
- Given the various concerns about the attribution approach, several commenters called into question the reliability and validity of the measure, noting the Steering Committee's split vote as to whether the measure was in fact valid.

Steering Committee Response:
The Steering Committee acknowledged many of the same concerns with the attribution approach. The Steering Committee stated concern that patients and their associated costs may potentially be attributed to specialists who provide primary care services that are Medicare allowable charges. This is particularly significant in the case of patients who receive long-term care for chronic conditions, who may receive many primary care services from specialists treating them for their chronic conditions, who are then attributed to a medical group practice based on the plurality of Medicare allowable charges. The Committee noted the distinction that specialists can provide primary care services through visits other than primary care visits.

The Committee was ultimately split on the concern that physicians have little ability to control the cost of care provided to patients by other specialists, through either inpatient care or post-acute care. Several Steering Committee members raised concern that it may be inappropriate to hold these providers accountable for these costs of care. Further, several Committee members stated that this may be appropriate in markets with integrated care delivery networks; however, in the current fragmented state of care delivery this attribution approach is not preferred. On the other hand, several other committee members stated that this level of accountability for providers is the entire rationale for the measure and should help push providers to be better organized to reduce costs.

The Steering Committee agreed with commenters that there are issues with both the first and second stage of the attribution approach. In the first stage, visits with non-physician providers (PAs and NPs) are not taken into account in the attribution model until the second stage, as non-physician providers are increasingly delivering more primary care. The Committee strongly encouraged CMS to include non-physician providers in the first stage of the attribution approach. Further, primary care services as defined by this measure may not always represent actual primary care visits by primary care providers. The Committee encouraged CMS to update this attribution approach.

Exclusions

- One commenter expressed concern that the exclusions of death and Medicare Advantage beneficiaries impact the usability of the measure.
- One commenter expressed concern that Medicare Part D (prescription medications) was excluded from the measure.

Steering Committee Response:
The committee was split on the reliability and validity of this measure but ultimately agreed that a number of issues, including the exclusions of death needed to be addressed before recommending this measure for endorsement. Additionally, Medicare Part D payment is an important area for measurement and improvement. CMS should consider approaches to including this data for beneficiaries with Part D coverage.

#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries

Reliability

- One commenter requested that the measure developer not publically report results for any provider group with reliability scores less than 0.70.
- One commenter stated that the measure is only reliable for groups of 25 or more eligible professionals; however, nearly half of all Medicare physicians practice in groups of fewer than 10 eligible professionals. As the measure will be used as part of CMS' value-based modifier calculation, the commenter questioned how this will impact smaller physician groups and solo practitioners.

Steering Committee Response:
While NQF does not require a specific cut-off for reliability testing, the Committee did encourage CMS to report information on provider groups that have adequate reliability in performance score and sample size. This measure should only be used for 25 or more eligible professions since this is the scope of measure testing.

Risk Adjustment

- One commenter expressed concern that the risk adjustment model might not adequately capture the differences in patient population for different specialties, particularly those who treat patients with uncommon and very severe diseases.
    - Comment #3253: We appreciate the opportunity to comment on the measure #2165 (Payment-Standardized Total Per Capita cost Measure for Medicare Fee-for-Service Beneficiaries). We have some concerns about this measure and its potential for use as a component of the value-based modifier. From the measure description and information provided, it is unclear how this measure would be applied. We are concerned about the broad nature of this measure and the fact that it looks across different specialties rather than within each specialty. We understand that the risk adjustment takes into account complexity of disease; however, we are concerned that the risk adjustment model might not adequately capture the differences in the patient population for different specialties. We are concerned that certain specialties, particularly cognitive specialists like rheumatology caring for patients with uncommon and very severe diseases, as a whole might fare worse than others if this measure is applied across specialties. In addition, we reviewed the risk adjustment model and do not believe it adequately captures the scope and complexity of conditions that rheumatologists care for. The exclusion of consideration of specific patient populations in the risk adjustment model would put providers or centers who treat a large number of these patients at a disadvantage. We would urge any assessment of providers for efficiency to look within a specialty rather than across specialties and that the risk adjustment model be thoroughly reviewed through specialty societies.
- Several commenters stated that the HCC model, which was developed for the Medicare Advantage program, does not adequately account for risk for purposes of analyzing physician group resource use, as it was designed to risk adjust large patient populations for insurance rate determination.

Steering Committee Response:
The Committee generally agreed that while this HCC-risk adjustment model was developed for Medicare Advantage it was appropriate but weak in this application. The HCC model does not include as many diagnostic categories as many commercially available risk adjustment models and therefore may not be as accurate in assigning the appropriate risk categories for rare conditions. However, given the broad use of HCCs across Medicare programs the Committee agreed that this approach was sufficient for this application.

Final Steering Committee Recommendation for Endorsement (9/5/2013): Y-12; N-13

The measure was not recommended for endorsement by the Steering Committee and thus was not put out for NQF member vote or reviewed by the CSAC or NQF Board of Directors.

Notes

[1] Catlin A, Cowan C, Hartman M, et al. National health spending in 2006: a year of change for prescription drugs. *Health Aff*. 2008; 27(1):14–29.

[2] Banks J, Marmot M, Oldfield Z, et al. Disease and disadvantage in the United States and in England. *JAMA*. 2006;295(17):2037–2045.

[3] Hoyert DL, Matthews TJ, Menacker F, et al., Annual summary of vital statistics: 2004. *Pediatrics*. 2006;117(1):168–183.

[4] Weiss JE, Mushinski M. International mortality rates and life expectancy: selected countries. *Stat Bull Metrop Insur Co*. 1999;80(1):13–21.

[5] NQF. #2165 Payment-Standardized Total Per Capita Cost for Medicare Beneficiaries. Baltimore, MD:CMS;2013. Available at http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=73024. Last accessed June 2013.

[6] Winkelman R, Mehmud S. *A Comparative Analysis of Claims-Based Tools for Health Risk Assessment*. Schaumburg, IL: Society of Actuaries; 2007. Available at http://www.soa.org/Files/Research/Projects/risk-assessmentc.pdf. Last accessed June 2013.

Appendix A: Measure Specifications

	#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)
Steward	Centers for Medicare & Medicaid Services (CMS)
Description	The MSPB Measure assesses the cost of services performed by hospitals and other healthcare providers during an MSPB hospitalization episode, which comprises the period immediately prior to, during, and following a patient's hospital stay. Beneficiary populations eligible for the MSPB calculation include Medicare beneficiaries enrolled in Medicare Parts A and B who were discharged from short-term acute hospitals during the period of performance.
Resource Use Measure Type	Per Episode
Data Source	Administrative Claims
Level of Analysis	Facility
Construction Logic Description	The MSPB Measure assesses the cost to Medicare of services performed by hospitals and other healthcare providers during an MSPB episode. An MSPB episode is risk adjusted and includes Medicare payments for services provided to a beneficiary with start date falling between 3 days prior to an IPPS hospital admission (index admission) through 30 days post-hospital discharge.
Clinical Framework Description	Objective: The MSPB 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. Clinical Topic Area: Inpatient Admissions, all conditions Accounting for Comorbidities: Application of a variant of the CMS-HCC risk adjustment model. The model includes a select number of interaction terms between comorbidities. Measure of Episode Severity: Risk Adjustment model includes indicators for the MS-DRG of the index admission. Concurrency of Clinical Events. The MSPB Episode spans the period 3 days prior to the index hospital admission through 30 days post-discharge. All events that occur during this time period are included in the MSPB episode.
Costing Method	Standardized Pricing
Tested Population	Medicare
Resource Use Service Categories	Inpatient services: Inpatient facility services; Inpatient services: Evaluation and management; Inpatient services: Procedures and surgeries; Inpatient services: Imaging and diagnostic; Inpatient services: Lab services; Inpatient services: Admissions/discharges; Ambulatory services: Outpatient facility services; Ambulatory services: Emergency Department; Ambulatory services: Evaluation and management; Ambulatory services: Procedures and surgeries; Ambulatory services: Imaging and diagnostic; Ambulatory services: Lab services; Durable Medical Equipment (DME)

	#2158 Payment-Standardized Medicare Spending Per Beneficiary (MSPB)
Attribution Approach	The MSPB episode is attributed to the hospital on the trigger inpatient claim for the index hospital admission that begins an MSPB episode. Specifically, for any period of performance selected, the first set of hospitalizations that can be included in the MSPB Measure are those that begin on the fourth day of the period of performance. This permits sufficient data for the 3-day pre-hospitalization period. Hospitalizations eligible to start an MSPB episode also must end in a discharge 30 days prior to the end of the period of performance to permit the collection of claim information during the post-discharge period. For instance, for the current MSPB figures available on Hospital Compare, the period of performance is May 1, 2011 to December 31, 2011. In this case, hospitalizations that start on May 4 and have a discharge date before December 1 are eligible to be included as index admissions. As discussed in S.9.1., however, 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 are not considered index admissions for the purposes of the MSPB Measure. In other words, these cases will not generate new MSPB 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.
Risk Adjustment	Statistical risk model
Stratification	The risk-adjustment model is stratified by major diagnostic category (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/

Appendix B: Project Steering Committee and NQF Staff

STEERING COMMITTEE

Eugene Nelson, DSc, MPH (Co-Chair)
Dartmouth Institute for Health Policy and Clinical Practice
Norwich, VT

David Penson, MD, MPH (Co-Chair)
Vanderbilt University
Nashville, TN

Lawrence Becker
Xerox Corporation
Rochester, NY

Mary Ann Clark, MHA
Intralign
Washington, DC

Cheryl Damberg, PhD
RAND Corporation
Santa Monica, CA

Jennifer Eames-Huff, MPH
Pacific Business Group on Health
San Francisco, CA

Nancy Garrett, PhD
Hennepin County Medical Center
St. Paul, MN

Andrea Gelzer, MD, MS, FACP
AmeriHealth Mercy Family of Companies
Philadelphia, PA

David Gifford, MD, MPH
American Health Care Association
Washington, DC

Stanley Hochberg, MD
Boston Medical Center
Boston, MA

Lisa Latts, MD, MSPH, MBA, FACP
LML Health Solutions, LLC
Denver, CO

Matthew McHugh, PhD, JD, MPH, RN, CRNP, FAAN
University of Pennsylvania
Philadelphia, PA

Martin Marciniak, MPP, PhD
GlaxoSmithKline
Research Triangle Park, NC

James Naessens, ScD, MPH
Mayo Clinic
Rochester, MN

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

Carolyn Pare
Minnesota Health Action Group
Bloomington, MN

David Redfearn, PhD
Wellpoint
Las Vegas, NV

Andrew Ryan, PhD
Weill Cornell Medical College
New York, NY

Joseph Stephansky, PhD
Michigan Health & Hospital Association
East Tawas, MI

Thomas Tsang, MD, FACP
Merck
Boston, MA

Lina Walker, PhD
AARP – Public Policy Institute
Washington, DC

William Weintraub, MD, FACC
Christiana Care Health System
Newark, DE

Daniel Wolfson, MHSA
ABIM Foundation
Philadelphia, PA

Herbert Wong, PhD
Agency for Healthcare Research and Quality
Rockville, MD

Dolores Yanagihara, MPH
Integrated Healthcare Association
Oakland, CA

NQF STAFF

Helen Burstin, MD, MPH
Senior Vice President
Performance Measurement

Taroon Amin, MA, MPH
Senior Director
Performance Measurement

Ashlie Wilbon, RN, MPH
Managing Director
Performance Measurement

Lindsey Tighe, MS
Senior Project Manager
Performance Measurement

Evan M. Williamson, MPH, MS
Project Manager
Performance Measurement

Appendix C: Measures Endorsed in Cost and Resource Use Since April 2012

NQF Number	Title	Steward
1557	Relative Resource Use for People with Diabetes	National Committee for Quality Assurance (NCQA)
1558	Relative Resource Use for People with Cardiovascular Conditions	National Committee for Quality Assurance (NCQA)
1560	Relative Resource Use (RRU) for People with Asthma	National Committee for Quality Assurance (NCQA)
1561	Relative Resource Use for People with Chronic Obstructive Pulmonary Disease (COPD)	National Committee for Quality Assurance (NCQA)
1598	Total Resource Use Population-based PMPM Index	HealthPartners
1604	Total Cost of Care Population-based PMPM Index	HealthPartners
1609	ETG based Hip/Knee Replacement Cost of Care	Ingenix/OptumInsight
1611	ETG based Pneumonia Cost of Care	Ingenix/OptumInsight

Appendix D: Related and Competing Measures

Comparison of NQF #1598 and NQF #2165

	#1598 Total Resource Use Population-based PMPM Index	#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries
Steward	HealthPartners	Centers for Medicare & Medicaid Services (CMS)
Description	The Resource Use Index (RUI) is a risk adjusted measure of the frequency and intensity of services utilized to manage a provider group's patients. Resource use includes all resources associated with treating members including professional, facility inpatient and outpatient, pharmacy, lab, radiology, ancillary and behavioral health services.	The Payment-Standardized Total Per Capita Cost Measure for Medicare FFS Beneficiaries assesses the per capita (per beneficiary) cost of health care services for Medicare FFS beneficiaries enrolled in Parts A and B and attributed to medical group practices. The measure includes all Medicare Part A and Part B costs during a calendar year and is payment-standardized and risk-adjusted (using patient demographics and medical conditions) to account for any potential differences in costs among providers that result from circumstances beyond the physician's control. Under CMS' attribution rule, beneficiaries are attributed on the basis of the plurality of primary care services, to those medical group practices with the greatest potential to influence the quality and cost of care delivered to Medicare FFS beneficiaries.
Resource Use Measure Type	Per capita (population- or patient-based)	Per capita (population- or patient-based)
Data Source	Administrative claims, Other: Users administrative claims data base, Risk-adjustment Tool, Johns Hopkins ACG System Version 9.0, Standardized costing code table, Total Care Relative Resource Values (TCRRV) specification provided	Administrative Claims
Level of Analysis	Clinician: Group/Practice; Population: Community	Clinician : Group/Practice

	#1598 Total Resource Use Population-based PMPM Index	#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries
Construction Logic Description	The measure examines total resource use of a commercial population between for a given measurement year (e.g. January 1 and December 31), for all members eligible for the measure	The Payment-Standardized Total Per Capita Cost Measure for Medicare FFS Beneficiaries is formed by first attributing beneficiaries to medical group practices. Then, unadjusted per capita costs are calculated as the sum of all Medicare Part A and Part B costs for all beneficiaries attributed to a medical group practice, divided by the number of attributed beneficiaries. All unadjusted costs are then payment-standardized and risk adjusted to accommodate differences in costs between peers that result from circumstances beyond physicians' control. Risk-adjusted costs are computed as the ratio of a medical group practice's payment-standardized (but not risk-adjusted) per capita costs to its expected per capita costs, as determined by the risk adjustment algorithm. Finally, to express the risk-adjusted cost in dollars and for ease of interpretation, the ratio is multiplied by the mean cost of all beneficiaries attributed to all practices.
Clinical Framework Description	Not applicable. This is a population-based measure that applies to all service categories, care settings and conditions.	This is an annual payment-standardized per capita cost measure for medical group practices that applies to all clinical topic areas. Comorbidities and clinical hierarchies are accounted for during the risk-adjustment process.

	#1598 Total Resource Use Population-based PMPM Index	#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries
Costing Method	**Description:** The Total Care Relative Resource Values (TCRRVs) are a grand linear scale of relative values designed to evaluate resource use across all types of medical services, procedures and places of service. The values are independent of price and can be used to evaluate providers, hospitals, physicians and health plans against their peers on their efficiency of resource use in treating like conditions. **General Overview of Application:** The TCRRVs are applied at the procedure level for each component of care with the exception of inpatient, which is applied at the full admission level. There is a TCRRV lookup table for each component of care where each claim's procedure is matched with the corresponding value. The TCRRV weights that are applied to the claim is tested for accuracy and a total TCRRV is calculated. The final step is to calibrate the total TCRRVs to the paid ratio between components of care using the paid adjustment factor. www.healthpartners.com/files/56500.pdf OR www.healthpartners.com/tcoc.	Standardized Pricing
Tested Population	Commercial	Medicare; Medicaid

	#1598 Total Resource Use Population-based PMPM Index	#2165 Payment-Standardized Total Per Capita Cost Measure for Medicare Fee-for-Service (FFS) Beneficiaries
Resource Use Service Categories	Inpatient services: Inpatient facility services; Inpatient services: Evaluation and management; Inpatient services: Procedures and surgeries; Inpatient services: Imaging and diagnostic; Inpatient services: Lab services; Inpatient services: Admissions/discharges; Inpatient services: Labor (hours, FTE, etc.); Ambulatory services: Outpatient facility services; Ambulatory services: Emergency Department; Ambulatory services: Pharmacy; Ambulatory services: Evaluation and management; Ambulatory services: Procedures and surgeries; Ambulatory services: Imaging and diagnostic; Ambulatory services: Lab services; Ambulatory services: Labor (hours, FTE, etc.); Durable Medical Equipment (DME)	Inpatient services: Inpatient facility services; Inpatient services: Evaluation and management; Inpatient services: Procedures and surgeries; Inpatient services: Imaging and diagnostic; Inpatient services: Lab services; Inpatient services: Admissions/discharges; Inpatient services: Labor (hours, FTE, etc.); Ambulatory services: Outpatient facility services; Ambulatory services: Emergency Department; Ambulatory services: Evaluation and management; Ambulatory services: Procedures and surgeries; Ambulatory services: Imaging and diagnostic; Ambulatory services: Lab services; Durable Medical Equipment (DME); Other services not listed; Hospice; Home health; skilled nursing facility; Anesthesia; Ambulance services; Chemotherapy; Drugs administered in an ambulatory setting or used with DME (covered by Medicare Part B); Orthotics, chiropractic, enteral and parenteral nutrition; some vision services; some hearing and speech services; immunizations
Attribution Approach	**Guidelines:** To determine which members to include in the Total Resource Use measure, there are several options available depending upon your business purpose and unit of measure. If the unit of measure is an entire health plan or employer group, all members will be included in the Total Resource Use measure. If the unit of measure is a provider and members are required to select a primary care provider, we recommend using the member selected provider. When the member is not required to select a primary care provider, we recommend the use of an attribution algorithm to identify the member's primary care provider. The measure was tested using this methodology.	Beneficiaries are attributed to medical group practices that provided the plurality of primary care services (PCS). Only beneficiaries that received PCS from at least one physician during the measurement period are eligible for assignment.
Risk Adjustment	For Total Resource Use measurement, risk adjustment is performed using Adjusted Clinical Groups (ACG) developed by Johns Hopkins University.	Statistical risk model
Stratification	This is a population-based measure that is fully inclusive.	This measure uses risk-adjusted costs for comparison purposes and further stratification is not done.

National Quality Forum
1030 15th St NW, Suite 800
Washington, DC 20005
http://www.qualityforum.org

Contract HHSM-500-2012-00009I
Task Order 1

ISBN 978-1-933875-58-3