NATIONAL QUALITY FORUM

Memo

TO: Cost and Resource Use Standing Committee

FR: NQF Staff

RE: Post-Comment Call to Discuss Public and Member Comments

DA: June 1, 2017

Purpose of the Call

The Cost and Resource Use Standing Committee will meet via conference call on Tuesday, June 6, 2017 from 2:00 – 4:30 pm ET. The purpose of this call is to:

- Review and discuss comments received during the post-evaluation public and member comment period.
- Provide input on proposed responses to the post-evaluation comments.
- Determine whether reconsideration of any measures or other courses of action are warranted.

Due to time constraints, during this call we will review comments by exception, in the case the Committee disagrees with the proposed responses.

Standing Committee Actions

1. Review this briefing memo and Draft Report.
2. Review and consider the full text of all comments received and the proposed responses to the post-evaluation comments (see Comment Table).
3. Be prepared to provide feedback and input on proposed post-evaluation comment responses.

Conference Call Information

Please use the following information to access the conference call line and webinar:

Speaker dial-in #: 1-844-479-5351
Web Link: http://nqf.commpartners.com/se/Rd/Rg.aspx?302156
Registration Link: http://nqf.commpartners.com/se/Rd/Rg.aspx?302156

Background

Healthcare spending in the United States has increased exponentially, yet the United States ranks below other developed countries for quality of care and health outcomes. The goal of this project is to improve efficiency through measurement in order to reduce the rate of cost growth and improve the quality of care provided.

In this project, the 19-member Cost and Resource Use Standing Committee met during a one-day in-person meeting to evaluate a total of three maintenance measures. The Committee recommended all three measures for re-endorsement. The Standing Committee will reconvene on June 6 to discuss the comments received from the public regarding these three recommended maintenance measures.

Comments Received

NQF solicits comments on measures undergoing review in various ways and at various times throughout the evaluation process. First, NQF solicits comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). Second, NQF solicits member and public comments prior to the evaluation of the measures via an online tool located on the project webpage. Third, NQF opens a 30-day comment period to both members and the public after measures have been evaluated by the full committee and once a report of the proceedings has been drafted.

Pre-evaluation comments

The pre-evaluation comment period was open from February 20, 2017 to March 6, 2017 for the three maintenance measures under review. A total of 33 pre-evaluation comments were received on the three maintenance measures submitted for re-endorsement. Comments included questions about measure specifications, risk adjustment methods, and interpretation of submitted performance data, as well as comments from healthcare organizations and practitioners expressing support for the re-endorsement of NQF #1598 and NQF #1604. NQF staff provided all submitted comments to the Committee prior to its initial deliberations during the in-person meeting.

Post-evaluation comments

The Draft Report went out for Public and Member comment from April 20, 2017 to May 19, 2017. During this commenting period, NQF received 21 comments from nine member organizations:

Consumers – 0	Professional – 4
Purchasers – 1	Health Plans – 0
Providers – 2	QMRI – 1
Supplier and Industry – 0	Public & Community Health - 1

In order to facilitate discussion, the majority of the post-evaluation comments have been categorized into major topic areas or themes. Where possible, NQF staff has proposed draft responses for the Committee to consider. Although all comments and proposed responses are subject to discussion, we will not necessarily discuss each comment and response on the post-comment call. Instead, we will spend the majority of the time considering the major topics and/or those measures with the most significant issues that arose from the comments. Note that the organization of the comments into major topic areas is not an attempt to limit Committee discussion.

We have included all of the comments that we received (both pre- and post-evaluation) in the Comment Table. This comment table contains the commenter's name, comment, associated measure, topic (if applicable), and—for the post-evaluation comments—draft responses for the Committee's consideration. Please refer to this comment table to view and consider the individual comments received and the proposed responses to each.

A complete table of comments submitted pre- and post-evaluation, along with the responses to each comment and the measure recommendations made by the Standing Committee, is posted to the Cost and Resource Use project page on the NQF website, along with the measure submission forms.

Comments and their Disposition

Five major themes were identified in the post-evaluation comments, as follows:

1. Concerns about reliability and validity
2. Adjusting for social risk factors
3. Concerns about populations included in the measures
4. Support for measures
5. Updates to the Cost and Resource Use Measure Evaluation Criteria

In addition, there were a number of measure-specific comments, which are discussed in detail, measure by measure, below. In summary, the issues raised were:

- 1598 comments included:
  - Concerns around the measure's testing and usability in states outside of those two the measure was tested in, specifically unintended consequences, standardized prices, a risk adjustment approach, and acceptable sample sizes;
  - Concerns with the lack of adjustment for social risk factors (addressed in Theme 2).
  - Inclusion of non-generalist OB/GYNs (addressed in Theme 3);
  - Support for the measure (addressed in Theme 4)
- 1604 comments included:
  - Concerns around the measure's testing and usability in states outside of those two the measure was tested in, specifically unintended consequences, standardized prices, a risk adjustment approach, and acceptable sample sizes;
  - Concerns with the lack of adjustment social risk factors (addressed in Theme 2).
  - Inclusion of non-generalist OB/GYNs (addressed in Theme 3);
  - Support for the measure (addressed in Theme 4)

- 2158 comments included:
  - Concerns that the measure is only validated and endorsed at the facility level, and not physician level
  - Concerns with the measure's testing for reliability and validity (addressed in Theme 1).
  - Concerns with the lack of adjustment for social risk factors (addressed in Theme 2)
  - Concerns that the majority of variation in the measure is driven by post-acute spending. Commenters noted this measure is used in the Hospital Value-Based Purchasing Program and that there is a potential for negative unintended consequences from its use.
  - Support for the measure (addressed in Theme 4)

Theme 1 – Concerns about reliability and validity

A number of commenters questioned the reliability and validity of the measures. Comments on #2158*: Medicare Spending Per Beneficiary (MSPB) – Hospital* raised concerns about the weak association with measures of readmissions as well as the attribution model of the measure. Commenters noted that post-acute spending drives most of the variation in the measure and that hospitals may have limited ability to influence their results. Comments on #1598: *Total Resource Use Population-based PMPM Index* and #1604: *Total Cost of Care Population-based PMPM Index* raised concerns that testing occurred only in two states. Additionally commenters requested additional details on standardized prices, risk adjustment approaches, and acceptable sample sizes to ensure the measures enable accurate comparisons.

Developer Response:

See responses in "Measure Specific Comments" Section of Memo

Proposed Committee Responses:

The Committee has reviewed your comment and appreciates your input. The Committee recognizes the need to ensure NQF-endorsed cost and resource use measures are reliable and valid.

The Committee had in-depth conversations on the attribution of #2158. The Committee recognizes that hospitals may not have complete control over the spending captured by the measure. However, the Committee believes that there are actions hospitals can take to improve their performance on this measure. Additionally, the Committee noted the need for attribution models that support care coordination and team-based care as the system aims to transition from fee-for-service to population-based payment.

The Committee noted that #1598 and #1604 have been widely implemented and users have supported the usefulness of the information generated by the measures.

Theme 2 – Adjusting for Social Risk Factors

Three comments, two on measure *#2158: Medicare Spending Per Beneficiary (MSPB) – Hospital* and one on *#1598: Total Resource Use Population-based PMPM Index*, expressed concern regarding potentially insufficient adjustments made for social risk factors. Commenters were concerned the developers did not provide an adequate conceptual basis and justification for the risk factors included in the testing, and did not include several factors commonly available in the literature. The comments submitted to NQF urged the Committee to take a more in-depth look at the need for SDS adjustment, given the potentially negative impact these measures could have on providers. Commenters encouraged additional testing of SDS factors.

Developer Response:

See responses in "Measure Specific Comments" Section of Memo

Proposed Committee Response:

The Committee has reviewed your comment and appreciates your input. Consideration of social risk factors in risk adjustment models is a critical issue in measurement science. The Committee was charged with evaluating the measure specifications and testing submitted on the measure as developed by the measure developer. The Committee recognizes that there continues to be limitations in the available data elements to capture unmeasured clinical and social risk. Given the constraints on the current data elements available, the Committee relied on the methods used by the measure developers to test the conceptual and empirical relationship between social risk factors and readmissions.

While the Committee generally accepted the findings of the analyses conducted by the developer, the Committee agrees that more work is needed to identify more robust data elements and methods to isolate and account for unmeasured clinical and social risk for patients. The Committee recognized the impact that social risk can have on cost and resource use measures and encourages measure developers to test the impact of additional social risk variables. The Committee also encouraged exploration of the impact of community-level variables. However, the Committee generally agreed that

the risk adjustment method used in these measures met the NQF criteria given the data available to the developer, and the measure testing results presented.

Committee Action Item:

The Committee should review and discuss the comments and the developers' responses. Does the Committee agree with the proposed responses?

Theme 3- Concerns about populations included in the measures

Several commenters raised concerns about populations included in the measures, noting that spending can vary significantly for certain provider types and patient groups. One commenter asked for clarification on how all three measures address cancer patients. The commenter noted that there can be significant variation in treatment needs, comorbidities, and patient preferences that can influence cost and resource use.

One commenter raised expressed concern with the inclusion of all obstetrician-gynecologists and pharmacy resources in measures #1598 and #1604. They noted that non-generalist obstetrician-gynecologists provide specialty care and suggested only including generalists in these two measures. The commenter also noted that providers do not control insurer formularies and that information on the cost of pharmaceuticals is not available.

Developer Responses:

Response 1: HealthPartners

HealthPartners' agrees that cost measures are not a marker of quality, and should not be used to draw conclusions on quality.

As a provider of cancer care, (American Society of Clinical Oncology-certified practices HealthPartners Regions Hospital Cancer Care Center and Park Nicollet Frauenshuh Cancer Center), HealthPartners' understands the complexity of cancer diagnosis, the importance of early detection and the variation in treatments as we care for cancer patients in our medical groups.

We are sensitive to the complexity of including cancer patients in our full population measure and through our testing believe the measures sufficiently adjust for cancer patients in the population through the clinical risk adjustment process and application of the measure criteria. John's Hopkins ACGs accurately and reliably adjusts for the clinical risk of a population including the risk of cancer patients and this is a primary reason why HealthPartners recommends the use of a commercially available clinical risk adjuster rather than less effective open source adjusters. The measure criteria of a minimum of 9 month of enrollment ensures there is enough patient history to accurately assess risk and cost. In addition, the costs are truncated at $125,000 so no one patient can overly impact the performance of the measures. The peer group is also of vital importance when performing cost and resource use evaluations.

Response 2: Acumen/CMS

Thank you for your comment. We recognize that cancer patients often have complex comorbidities and require more intensive treatment. The MSPB-Hospital measure accounts for comorbidities through risk adjustment. Specifically, the MSPB-Hospital risk adjustment methodology adjusts the MSPB-Hospital measure for age, severity of illness, and enrollment status indicators. The methodology includes 12 age-categorical

variables; 79 hierarchical condition category (HCC) variables derived from the beneficiary's claims during the period 90 days prior to the start of the episode to measure severity of illness; as well as the MS-DRG of the index hospitalization. The risk adjustment methodology also includes the HCC interaction variables, status indicator variables for whether the beneficiary qualifies for Medicare through Disability or End-Stage Renal Disease (ESRD), and whether a beneficiary resides in a long-term care facility. The MS-DRG is included in the risk adjustment model to better account for the differences in cost of care that stem from different reasons for hospitalization, including cancer. This allows the MSPB-Hospital measure to compare cost of care across all conditions, rather than focusing on a specific disease. As such, the risk factor of the MS-DRG of the index hospitalization should account for the more intensive treatment that cancer patients may require.

We appreciate your comment that cost measures should capture and categorize costs throughout the cycle of care. Hospitals that have an MSPB-Hospital measure receive a Hospital-Specific Report that provides a cost breakdown by claim type for the hospital's MSPB-Hospital episodes for three categories: 3 days prior to index admission, during-index admission, and 30 days after hospital discharge. This breakdown is provided for informational purposes to allow hospitals to evaluate its episode spending before, during, or after the index hospital admission.

We also wanted to acknowledge and address your comment that cost measures are not necessarily good markers of quality. For this reason, we note in our public documentation that the MSPB-Hospital measure alone is not intended to necessarily reflect the quality of care provided by hospitals. Accordingly, a lower MSPB-Hospital measure score across performance periods (i.e., lower Medicare spending per beneficiary) in isolation, should not be interpreted as better care. The MSPB Measure is most meaningful when presented in the context of other quality measures, which are part of the Hospital Value-Based Purchasing (VBP) Program. As part of the Hospital VBP Program, the MSPB-Hospital measure is combined with current quality of care measures to facilitate profiling hospital value (payments and quality).

Proposed Committee Response:

The Committee will discuss and draft responses to these comments during the June 6 post-comment call.

Committee Action Item:

The Committee should review the comments and the responses in detail. After reviewing the responses from Acumen and HealthPartners, the Committee should discuss a response.

Theme 4 – Support for Measures

Seven of the comments received were in support of the measures' continued endorsement.

- Measure #1598: *Total Resource Use Population-based PMPM Index* received two supportive comments.
- Measure #1604: *Total Cost of Care Population-based PMPM Index* received three supportive comments.
- Measure #2158: *Medicare Spending Per Beneficiary (MSPB) – Hospital* received one supportive comment.

In addition, one general comment was received, noting the gap in measures in this area and supporting the continued endorsement of these three measures.

Proposed Committee Response:

Thank you for your comment.

Theme 5 – Updates to the Cost and Resource Use Measure Evaluation Criteria

Overall, commenters were supportive of the revisions to the Cost and Resource Use Measure Evaluation Criteria. However, commenters asked for additional clarifications on what information should be provided by developers to address the performance gap subcriterion.

NQF Response:

NQF thanks the commenters for their support for the revisions to the Cost and Resource Use Measure Evaluation Criteria. The performance gap subcriterion is meant to address the question of whether there is actually a cost and resource use problem that is addressed by a particular measure. Because the measurement enterprise is resource intensive, NQF's position is to endorse measures that address areas of known gaps in performance (i.e., those for which there is actually opportunity for improvement). Opportunity for improvement can be demonstrated by data that indicate overall poor performance (in the activity or outcome targeted by the measure), substantial variation in performance across providers, or variation in performance for certain subpopulations (i.e., disparities in care). The proposed update removes subcriterion 1c to streamline the criteria, harmonize with the quality measure evaluation criteria, and prevent redundancies with the reliability and validity subcriterion.

Measure Specific Comments

1598: Total Resource Use Population-based PMPM Index

This measure received a total of five comments. Two supported continued endorsement, and one, regarding the inclusion of non-generalist OB/GYNs, is addressed above in Theme 3. The remaining two comments, both submitted by the AMA, focus on concerns around the measure's testing and usability in states outside of those two the measure was tested in. The commenters requested more detail on the measure's current performance and implementation experiences, given that it has been in use for five years. Specific concerns raised include unintended consequences, standardized prices, a risk adjustment approach, and acceptable sample sizes. Additional comments focusing on concerns with SDS are addressed in Theme 2, above.

Developer Responses:

Response 1: HealthPartners

HealthPartners thanks the American Medical Association (AMA) for sharing its comments.

To address the AMA's first comment regarding standardized pricing, the Total Resource Use measure uses the Total Care Relative Resource Values (TCRRVs). 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. TCRRVs are based on industry standard weighting systems (RVU, DRG, APC). 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.

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.

Details regarding standardized prices can be found under section S.9.6. "Costing Method" within the measure submission form. The detailed development of the TCRRV methodology is described in a technical white paper publicly available on HealthPartners' Total Cost of Care website.

https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/cntrb_039627.pdf

The full TCRRV tables are available via our website and licensed, free of charge at: www.healthpartners.com/tcoc. Below is a sample TCRRV table:

https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/entry_188112.pdf

In regards to the comment shared about the risk adjustment approach, HealthPartners' Total Cost of Care and Resource Use measures are specified for use of the Johns Hopkins' Adjusted Clinical Groups (ACG System). However, we recognize the practicality of communities and users who have financially invested in different risk adjustment groupers. Users opting to use different risk adjuster for their analysis should test for both reliability and validity of the measures. Additionally, for comparability of results across different users, each user must use the same risk adjustment tool.

The Society of Actuaries Accuracy of Claims-Based Risk Scoring Models (2016) findings that suggest other commercially available risk groupers perform similarly.

https://www.soa.org/Files/Research/research-2016-accuracy-claims-based-risk-scoring-models.pdf

To clarify the testing was performed on provider groups with a minimum of 600 members for both Total Cost of Care and Total Resource Use measures. While HealthPartners performed the testing at the provider group level the unit of analysis can be applied to a variety of units, such as the health plan, employer group, provider group, clinic, physician or geographical area. The measures' constructs remain constant and are not dependent on level of analysis.

The HealthPartners measures of Total Cost of Care and Total Resource Use are not measures of value, they are measures that represent the affordability arm of the Triple Aim. HealthPartners is focused on the Triple Aim, quality, experience and affordability of health care to provide value for our patients. We believe it is essential when measuring affordability to complement it with quality and experience measurement, which allows members and patients to make their own value determination. The majority of the contracted provider groups in our health plan network are focusing on the same work, each having their own process improvement plan. HealthPartners' health plan implemented a Triple Aim risk sharing program over 5 years ago with Total Cost of Care representing the affordability component of the Triple Aim. To date, tens of millions of dollars have been paid out by the health plan and self-insured employer groups. Essentia Health, CentraCare Health and Fairview Health Services are provider groups outside the HealthPartners family of providers who have shared their success

stories through letters of support in the usability section of the National Quality Forum measure submission process.

In addition, HealthPartners compares its own family of provider groups against the health plan network wide benchmark to identify total cost and resource use improvement opportunities. While working towards better affordability of care and reducing costs, we are still able to maintain the highest quality of care. HealthPartners' family of providers were recently reported as three of the highest performers statewide in most quality measures by Minnesota Community Measurement.

In response to the question raised about comparisons across different medical specialties, to clarify, the measure is specified as a full population measure, including all care, from all provider specialties. The unit of analysis, or attribution, is a measurement guideline for both Total Cost of Care and Total Resource Use measures. The risk adjustment accounts for variation in age, gender and the clinical risk of patients treated by various specialists (e.g. typically pediatric patients receive lower risk scores).

Minnesota Community Measurement (MNCM) is a community collaborative organization that measures Total Cost of Care and Total Resource Use annually, according to the NQF-endorsed specifications, for all provider groups in the state of Minnesota and bordering communities to drive improvement by showcasing variation through transparency. The Network for Regional Healthcare Improvement (NRHI) represents more than 30 Regional Health Improvement Collaboratives (RHICs) across the United States and three state- affiliated partners. With funding from the Robert Wood Johnson Foundation, five pioneering regional health improvement collaboratives (RHICs) are now joined by six additional regions to standardize how they report cost information. NRHI is driving a national effort to make care affordable by using the NQF-endorsed specifications to make cost and resource information consistent and transparent. Both MNCM (third year) and NRHI (first year) results were included in the usability portion of our submission documents for the committee's review.

You can't improve what you don't measure. The uptake of these measures across the country and provider engagement are the first steps to reaching our nation's goal of providing affordable care for our patients.

To address the AMA's last concern about testing sociodemographic (SDS) factors, the Total Cost of Care and Total Resource Use measures are risk adjusted for age, gender, and clinical risk profile based on diagnosis. The measures are also specified for the commercially insured population. Income and education status were explored as potential socioeconomic (SES) variables for additional adjustment due to their conceptual alignment, along with their likely data availability. Income has been viewed as a main contributor to healthcare access and affordability along with education influencing a patient's approach to the healthcare system1,2,3. Income is a continuous and granular variable. Education status is a categorical variable and difficult to create an average or median. Because income and education have been found to be correlated and because income was a more continuous and granular variable HealthPartners focused the analysis on income.

Testing was done on a data element reasonably available to HealthPartners or other users, which would not include the majority of factors listed by the AMA. HealthPartners used two separate data sources to evaluate income. The first was U.S. Census Tracts. The second was a more robust commercially licensed data source that HealthPartners has access to for other business purposes, which provided us with household level income.

To ensure the study population included lower income ranges, HealthPartners Medicaid population was included along with the full commercial book of business for testing. The Medicaid population has a different reimbursement rate (typically significantly lower) than the commercial reimbursement rate, which would result in a lower total cost of care. The Medicaid population was included to prove that the product delineation between Medicaid and Commercial sufficiently controls for the variation in cost and resource use, therefore, adding income in the model resulted in no additional explanatory power.

As stated in the SES testing analysis, after risk adjusting for age, gender, and clinical risk, and limiting by commercial product, income did not significantly impact a patient's total cost or resource use. There was less than a 1% change in performance for all provider groups when income was introduced into the model for both measures when using Census Tract data and less than a 0.5% change when using the commercially licensed data source with more granular income data.

Citations:

1. Alter D, et. al. Lesson From Canada's Universal Care: Socially Disadvantaged Patients Use More Health Services, Still Have Poorer Health.. Health Affairs doi: 10.1377/hlthaff.2009.0669Health Aff February 2011 vol. 30 no. 2 274-283 http://content.healthaffairs.org/content/30/2/274.abstract?sid=94d288f0- 331d-469e-8c11-023b272bed92

2. Lemstra, M, et.al. High health care utilization and costs associated with lower socio-economic status: results from a linked data set. Can J Public Health. 2009 May-Jun;100(3):180-3. https://www.jstor.org/stable/41995241?seq=1#page_scan_tab_contents 4.

3. United States Department of Labor, Bureau of Labor Statistics 2015. Earnings and Unemployment by Educational Attainment Status. Last Modified March 15, 2016. http://www.bls.gov/emp/ep_chart_001.htm

Response 2: HealthPartners

HealthPartners thanks the American College of Obstetricians and Gynecologists (the College) for sharing its comment. The intent of the Total Cost of Care measure is to measure a provider's risk adjusted cost effectiveness at managing the population they care for. Similarly, the Total Resource Use measure is a risk adjusted measure of the frequency and intensity of services utilized to manage a provider's patients. While all costs and resources associated with treating patients are included for evaluation, implementation of the measures and how results are used and reported are decisions that need to be considered and defined by the users.

Measure reporting guidelines (guidelines are not a part of the specifications), which include attribution methodology, have been shared to assist with implementation of the measures and appropriate comparisons across specified reporting entities. For comparability purposes, the attribution method used in Total Cost of Care and Total Resource Use measurement must be consistently applied across the population measured. In addition, a peer group or benchmark must be defined. Users must determine both the method of attribution and the peer group to be used with their own market and specific business needs in mind.

HealthPartners' attribution process, which has been vetted and accepted locally for use in our market, includes the following specialties: family medicine, internal medicine, pediatrics, geriatrics, obstetrics and gynecology (OB/GYN). We agree with the College and recognize that subspecialties and specific areas of care may not reflect primary care

services. HealthPartners' measurement approach excludes specific OB/GYN specialties from attribution when measuring primary care providers (e.g. Gynecological Surgery). The measures are however versatile and could evaluate subspecialties if the peer group was limited to the subspecialty being evaluated. It is up to the user to ensure the intended use of the measure aligns with the providers being measured and is reflected in the peer group.

The measures are population-based, patient-centered and cross all categories of health care services, including pharmacy. Pharmacy contributes an estimated 20% of the total costs and resources and are driving steep trends. Therefore it is imperative to include pharmacy costs when measuring total cost of care and resource use. At HealthPartners we have and are enhancing our approaches to make the costs of drugs available to providers and consumers so the decision-making process can be fully informed. We would encourage others to do the same. While a provider may not have control over formulary drug lists, providers do have an opportunity to help educate patients on alternative drug options when there is clinical equivalence yet a large cost difference. Alternative therapies, generics or less expensive brand drugs may be options for patients that providers can help coordinate, leading to lower overall costs, lower out of pocket costs for patients, increased patient experience and most importantly, a better chance consumers can afford to fill and take their medicine as prescribed. Providers also have the ability to manage potential overuse of medications which not only reduces costs, but also improves quality of care for patients.

Proposed Committee Response:

The Committee has reviewed the comments and appreciates the additional insights on the measure. After reviewing the comments and responses from the developer the Committee believes this measure is appropriately specified and tested and continues to meet the criteria for NQF endorsement.

Action Item:

The Committee should review the comments and the developer's responses, and discuss on the post-comment call.

1604: Total Cost of Care Population-based PMPM Index

This measure received a total of five comments. Three supported continued endorsement, and one, regarding the inclusion of OB/GYNs, is addressed above in Theme 3. The remaining comment focuses on concerns around the measure's testing and whether the measure's specifications are precise enough to ensure consistent implementation. Specific concerns raised include unintended consequences, standardized prices, a risk adjustment approach, and acceptable sample sizes.

Developer Responses:

Response 1: HealthPartners

HealthPartners thanks the American Medical Association (AMA) for sharing its comments.

To address the AMA's first comment regarding standardized pricing, the Total Resource Use measure uses the Total Care Relative Resource Values (TCRRVs). 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. TCRRVs are based on industry standard weighting systems (RVU, DRG, APC). 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.

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.

Details regarding standardized prices can be found under section S.9.6. "Costing Method" within the measure submission form. The detailed development of the TCRRV methodology is described in a technical white paper publicly available on HealthPartners' Total Cost of Care website.

https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/cntrb_039627.pdf

The full TCRRV tables are available via our website and licensed, free of charge at: www.healthpartners.com/tcoc. Below is a sample TCRRV table:

https://www.healthpartners.com/ucm/groups/public/@hp/@public/documents/documents/entry_188112.pdf

In regards to the comment shared about the risk adjustment approach, HealthPartners' Total Cost of Care and Resource Use measures are specified for use of the Johns Hopkins' Adjusted Clinical Groups (ACG System). However, we recognize the practicality of communities and users who have financially invested in different risk adjustment groupers. Users opting to use different risk adjuster for their analysis should test for both reliability and validity of the measures. Additionally, for comparability of results across different users, each user must use the same risk adjustment tool.

The Society of Actuaries Accuracy of Claims-Based Risk Scoring Models (2016) findings that suggest other commercially available risk groupers perform similarly. https://www.soa.org/Files/Research/research-2016-accuracy-claims-based-risk-scoring-models.pdf

To clarify the testing was performed on provider groups with a minimum of 600 members for both Total Cost of Care and Total Resource Use measures. While HealthPartners performed the testing at the provider group level the unit of analysis can be applied to a variety of units, such as the health plan, employer group, provider group, clinic, physician or geographical area. The measures' constructs remain constant and are not dependent on level of analysis.

The HealthPartners measures of Total Cost of Care and Total Resource Use are not measures of value, they are measures that represent the affordability arm of the Triple Aim. HealthPartners is focused on the Triple Aim, quality, experience and affordability of health care to provide value for our patients. We believe it is essential when measuring affordability to complement it with quality and experience measurement, which allows members and patients to make their own value determination. The majority of the contracted provider groups in our health plan network are focusing on the same work, each having their own process improvement plan. HealthPartners' health plan implemented a Triple Aim risk sharing program over 5 years ago with Total Cost of Care representing the affordability component of the Triple Aim. To date, tens of millions of dollars have been paid out by the health plan and self-insured employer groups. Essentia Health, CentraCare Health and Fairview Health Services are provider

groups outside the HealthPartners family of providers who have shared their success stories through letters of support in the usability section of the National Quality Forum measure submission process.

In addition, HealthPartners compares its own family of provider groups against the health plan network wide benchmark to identify total cost and resource use improvement opportunities. While working towards better affordability of care and reducing costs, we are still able to maintain the highest quality of care. HealthPartners' family of providers were recently reported as three of the highest performers statewide in most quality measures by Minnesota Community Measurement.

In response to the question raised about comparisons across different medical specialties, to clarify, the measure is specified as a full population measure, including all care, from all provider specialties. The unit of analysis, or attribution, is a measurement guideline for both Total Cost of Care and Total Resource Use measures. The risk adjustment accounts for variation in age, gender and the clinical risk of patients treated by various specialists (e.g. typically pediatric patients receive lower risk scores).

Minnesota Community Measurement (MNCM) is a community collaborative organization that measures Total Cost of Care and Total Resource Use annually, according to the NQF-endorsed specifications, for all provider groups in the state of Minnesota and bordering communities to drive improvement by showcasing variation through transparency. The Network for Regional Healthcare Improvement (NRHI) represents more than 30 Regional Health Improvement Collaboratives (RHICs) across the United States and three state- affiliated partners. With funding from the Robert Wood Johnson Foundation, five pioneering regional health improvement collaboratives (RHICs) are now joined by six additional regions to standardize how they report cost information. NRHI is driving a national effort to make care affordable by using the NQF-endorsed specifications to make cost and resource information consistent and transparent. Both MNCM (third year) and NRHI (first year) results were included in the usability portion of our submission documents for the committee's review.

You can't improve what you don't measure. The uptake of these measures across the country and provider engagement are the first steps to reaching our nation's goal of providing affordable care for our patients.

To address the AMA's last concern about testing sociodemographic (SDS) factors, the Total Cost of Care and Total Resource Use measures are risk adjusted for age, gender, and clinical risk profile based on diagnosis. The measures are also specified for the commercially insured population. Income and education status were explored as potential socioeconomic (SES) variables for additional adjustment due to their conceptual alignment, along with their likely data availability. Income has been viewed as a main contributor to healthcare access and affordability along with education influencing a patient's approach to the healthcare system1,2,3. Income is a continuous and granular variable. Education status is a categorical variable and difficult to create an average or median. Because income and education have been found to be correlated and because income was a more continuous and granular variable HealthPartners focused the analysis on income.

Testing was done on a data element reasonably available to HealthPartners or other users, which would not include the majority of factors listed by the AMA. HealthPartners used two separate data sources to evaluate income. The first was U.S. Census Tracts. The second was a more robust commercially licensed data source that HealthPartners has access to for other business purposes, which provided us with household level income.

To ensure the study population included lower income ranges, HealthPartners Medicaid population was included along with the full commercial book of business for testing. The Medicaid population has a different reimbursement rate (typically significantly lower) than the commercial reimbursement rate, which would result in a lower total cost of care. The Medicaid population was included to prove that the product delineation between Medicaid and Commercial sufficiently controls for the variation in cost and resource use, therefore, adding income in the model resulted in no additional explanatory power.

As stated in the SES testing analysis, after risk adjusting for age, gender, and clinical risk, and limiting by commercial product, income did not significantly impact a patient's total cost or resource use. There was less than a 1% change in performance for all provider groups when income was introduced into the model for both measures when using Census Tract data and less than a 0.5% change when using the commercially licensed data source with more granular income data.

Citations:

1. Alter D, et. al. Lesson From Canada's Universal Care: Socially Disadvantaged Patients Use More Health Services, Still Have Poorer Health.. Health Affairs doi: 10.1377/hlthaff.2009.0669Health Aff February 2011 vol. 30 no. 2 274-283 http://content.healthaffairs.org/content/30/2/274.abstract?sid=94d288f0- 331d-469e-8c11-023b272bed92

2. Lemstra, M, et.al. High health care utilization and costs associated with lower socio-economic status: results from a linked data set. Can J Public Health. 2009 May-Jun;100(3):180-3. https://www.jstor.org/stable/41995241?seq=1#page_scan_tab_contents 4.

3. United States Department of Labor, Bureau of Labor Statistics 2015. Earnings and Unemployment by Educational Attainment Status. Last Modified March 15, 2016. http://www.bls.gov/emp/ep_chart_001.htm

Response 2: HealthPartners

HealthPartners thanks the American College of Obstetricians and Gynecologists (the College) for sharing its comment. The intent of the Total Cost of Care measure is to measure a provider's risk adjusted cost effectiveness at managing the population they care for. Similarly, the Total Resource Use measure is a risk adjusted measure of the frequency and intensity of services utilized to manage a provider's patients. While all costs and resources associated with treating patients are included for evaluation, implementation of the measures and how results are used and reported are decisions that need to be considered and defined by the users.

Measure reporting guidelines (guidelines are not a part of the specifications), which include attribution methodology, have been shared to assist with implementation of the measures and appropriate comparisons across specified reporting entities. For comparability purposes, the attribution method used in Total Cost of Care and Total Resource Use measurement must be consistently applied across the population measured. In addition, a peer group or benchmark must be defined. Users must determine both the method of attribution and the peer group to be used with their own market and specific business needs in mind.

HealthPartners' attribution process, which has been vetted and accepted locally for use in our market, includes the following specialties: family medicine, internal medicine, pediatrics, geriatrics, obstetrics and gynecology (OB/GYN). We agree with the College and recognize that subspecialties and specific areas of care may not reflect primary care

services. HealthPartners' measurement approach excludes specific OB/GYN specialties from attribution when measuring primary care providers (e.g. Gynecological Surgery). The measures are however versatile and could evaluate subspecialties if the peer group was limited to the subspecialty being evaluated. It is up to the user to ensure the intended use of the measure aligns with the providers being measured and is reflected in the peer group.

The measures are population-based, patient-centered and cross all categories of health care services, including pharmacy. Pharmacy contributes an estimated 20% of the total costs and resources and are driving steep trends. Therefore it is imperative to include pharmacy costs when measuring total cost of care and resource use. At HealthPartners we have and are enhancing our approaches to make the costs of drugs available to providers and consumers so the decision-making process can be fully informed. We would encourage others to do the same. While a provider may not have control over formulary drug lists, providers do have an opportunity to help educate patients on alternative drug options when there is clinical equivalence yet a large cost difference. Alternative therapies, generics or less expensive brand drugs may be options for patients that providers can help coordinate, leading to lower overall costs, lower out of pocket costs for patients, increased patient experience and most importantly, a better chance consumers can afford to fill and take their medicine as prescribed. Providers also have the ability to manage potential overuse of medications which not only reduces costs, but also improves quality of care for patients.

Proposed Committee Response:

The Committee has reviewed the comments and appreciates the additional insights on the measure. After reviewing the comments and responses from the developer the Committee believes this measure is appropriately specified and tested and continues to meet the criteria for NQF endorsement.

Action Item:

The Committee should review the comments and the developer's responses, and discuss on the post-comment call.

2158: Medicare Spending Per Beneficiary (MSPB) – Hospital

This measure received a total of five comments, including one supporting its endorsement. Two commenters noted that this measure is only validated and endorsed at the facility level. Commenters expressed concerns that this measure should be validated at the clinician level before adoption for the Merit-Based Incentive Payment System (MIPS) and urged the developer to expedite clinician level testing. Three comments address concerns with the measure's testing for reliability and validity. Commenters questioned the weak association between this measure and measures of readmission. Commenters also raised concerns that the majority of variation in the measure is drive by post-acute spending. Commenters noted this measure is used in the Hospital Value-Based Purchasing Program and that there is a potential for negative unintended consequences from its use.

Commenters were also concerned about the SDS adjustment, as addressed above in Theme 2.

Developer Responses:

Response 1: Acumen/CMS

The measure developer appreciates the AMA's feedback on the MSPB-Hospital measure construction and the testing of sociodemographic (SDS) factors in the measure's risk adjustment model. The developer believes that the MSPB-Hospital measure does meet the scientific acceptability criteria of validity, and the NQF committee agreed that the measure met the Scientific Acceptability criterion. The NQF committee had 4 members vote that the measure demonstrated high validity, 9 members vote that the measure demonstrated medium validity, and 5 members vote that the measure had low reliability.

The MSPB-Hospital measure aims to improve care coordination in the period between 3 days prior to an acute inpatient hospital admission through the period 30 days after discharge. The MSPB-Hospital measure recognizes lower costs associated with a reduction in unnecessary services, preventable complications, readmissions, and shifting post-acute care from more expensive to less expensive services when appropriate. The MSPB-Hospital measure creates parallel incentives for hospitals and post-acute care providers. The developer would also like to clarify that 84% of the variance in episode cost is accounted for by post-acute care costs, rather than 84% of total episode costs being attributed to the hospital during the 30 day post-discharge period. This finding is consistent with expectations. The risk adjustment model predicts a certain level of post-discharge spending based upon the beneficiary's prior health history and MS-DRG. Specifically, the MSPB-Hospital risk adjustment methodology adjusts the MSPB-Hospital measure for age, severity of illness, and enrollment status indicators.

Variance in provider scores based on post-discharge spending emphasizes the importance of care transitions and care coordination in improving patient care. Hospitals receive a Hospital-Specific Report (HSR) that provides information on the hospital's performance on the MSPB-Hospital measure, as well as three supplementary hospital-specific data files (an index admission file, a beneficiary risk score file, and an MSPB-Hospital episode file) related to the hospital's MSPB-Hospital measure. Together, these files provide an overview of how the hospital performed on the MSPB-Hospital measure and identify other providers involved in care for their beneficiaries, which facilitates better coordination of care with those providers. No evidence of unintended consequences to individuals or populations, such as changes in referral patterns, have been identified during testing and since implementation.

The developer would also like to note that they submitted an updated measure testing form to the NQF on March 31st, 2017 that contained an appendix with additional analyses responding to NQF feedback and further description of the original submission. That appendix notes that analyses comparing the MSPB-Hospital measure with the condition-specific readmission measures were excluded in the 2016 submission because the condition-specific readmission measures examine hospital performance on a specific set of conditions, while the MSPB-Hospital measure is intended to capture hospital performance across all acute conditions. Consequently, comparisons could be misleading. Since MSPB-Hospital is an all cost measure that includes all conditions, the developer believes that it is more appropriate to look at the correlation between MSPB-Hospital and another broad-based all cost measure (i.e., the risk-adjusted, standardized total Medicare spending at the Hospital Referral Regions (HRR) level). The developer agrees that the MSPB-Hospital measure is most meaningful when presented in the context of other quality measures, which are part of the Hospital Value-Based Purchasing (VBP) Program. As part of the Hospital VBP Program, the MSPB-Hospital

measure is combined with current quality of care measures to facilitate profiling hospital value (payments and quality).

The developer believes that the MSPB-Hospital measure submission did meet the requirements of the NQF's SDS trial period and the NQF committee confirmed this by passing the MSPB-Hospital measure on the Scientific Acceptability criterion. The developer noted in the original submission that the inclusion of SDS factors (i.e., family income-to-poverty ratio and race) had a minimal impact on hospital's measure scores. The developer recognizes the commenter's concerns that additional factors could be included in the SDS measure testing. The developer selected family income-to-poverty ratio to strike a balance between the individual and community factors related to SES and listed by the commenter, as individual family members may pool financial resources to provide care for older relatives. The developer also conducted additional analyses based on feedback from the NQF committee to examine the impact of including a dual eligibility flag in the risk adjustment model, which are included in the appendix of the measure testing form that was submitted to the NQF on March 31st, 2017. These analyses showed that including a dual eligibility flag had a low impact on MSPB-Hospital measure scores and that hospitals on the tails of score distributions were not disproportionately affected. A recent ASPE report showed some differences in measure performance between hospitals with a high amount of Disproportionate Share Hospital payments and a low amount.* The analysis in the appendix's Supplementary Table 7 suggests that these differences may be driven by hospitals with a very high concentration of dual eligible beneficiaries (above 60%), and that measure scores are high for both duals and non-duals in these hospitals. This suggests that these hospitals are relatively higher-cost hospitals for all types of patients.

The developer also appreciates the commenter's feedback on the separate clinician-level measure (MSPB-TIN). The developer would like to clarify that while there are similarities in measure logic and measure construction between MSPB-TIN and the facility-level MSPB-Hospital measure currently under consideration for NQF re-endorsement, MSPB-TIN differs in attribution methodology. The MSPB-TIN measure is still under reevaluation. To ensure the reliability and validity of the measures being implemented, CMS reevaluates the measures annually and plans NQF submission of the measures by taking into account program needs and measure implementation timelines to meet the statutory requirements.

*Office of the Assistant Secretary for Planning and Evaluation (ASPE). "Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs." December, 2016. Available at https://aspe.hhs.gov/system/files/pdf/253971/ASPESESRTCfull.pdf.

Response 2: Acumen/CMS

The measure developer appreciates FAH's feedback on the MSPB-Hospital measure construction and the testing of sociodemographic (SDS) factors in the measure's risk adjustment model. The developer believes that the MSPB-Hospital measure does meet the scientific acceptability criteria of validity, and the NQF committee agreed that the measure met the Scientific Acceptability criterion. The NQF committee had 4 members vote that the measure demonstrated high validity, 9 members vote that the measure demonstrated medium validity, and 5 members vote that the measure had low reliability.

The MSPB-Hospital measure aims to improve care coordination in the period between 3 days prior to an acute inpatient hospital admission through the period 30 days after discharge. The MSPB-Hospital measure recognizes lower costs associated with a

reduction in unnecessary services, preventable complications, readmissions, and shifting post-acute care from more expensive to less expensive services when appropriate. The MSPB-Hospital measure creates parallel incentives for hospitals and post-acute care providers. The developer would also like to clarify that 84% of the variance in episode cost is accounted for by post-acute care costs, rather than 84% of total episode costs being attributed to the hospital during the 30 day post-discharge period. This finding is consistent with expectations. The risk adjustment model predicts a certain level of post-discharge spending based upon the beneficiary's prior health history and MS-DRG. Specifically, the MSPB-Hospital risk adjustment methodology adjusts the MSPB-Hospital measure for age, severity of illness, and enrollment status indicators.

Variance in provider scores based on post-discharge spending emphasizes the importance of care transitions and care coordination in improving patient care. Hospitals receive a Hospital-Specific Report (HSR) that provides information on the hospital's performance on the MSPB-Hospital measure, as well as three supplementary hospital-specific data files (an index admission file, a beneficiary risk score file, and an MSPB-Hospital episode file) related to the hospital's MSPB-Hospital measure. Together, these files provide an overview of how the hospital performed on the MSPB-Hospital measure and identify other providers involved in care for their beneficiaries, which facilitates better coordination of care with those providers. No evidence of unintended consequences to individuals or populations, such as changes in referral patterns, have been identified during testing and since implementation.

The developer would also like to note that they submitted an updated measure testing form to the NQF on March 31st, 2017 that contained an appendix with additional analyses responding to NQF feedback and further description of the original submission. That appendix notes that analyses comparing the MSPB-Hospital measure with the condition-specific readmission measures were excluded in the 2016 submission because the condition-specific readmission measures examine hospital performance on a specific set of conditions, while the MSPB-Hospital measure is intended to capture hospital performance across all acute conditions. Consequently, comparisons could be misleading. Since MSPB-Hospital is an all cost measure that includes all conditions, the developer believes that it is more appropriate to look at the correlation between MSPB-Hospital and another broad-based all cost measure (i.e., the risk-adjusted, standardized total Medicare spending at the Hospital Referral Regions (HRR) level). The developer agrees that the MSPB-Hospital measure is most meaningful when presented in the context of other quality measures, which are part of the Hospital Value-Based Purchasing (VBP) Program. As part of the Hospital VBP Program, the MSPB-Hospital measure is combined with current quality of care measures to facilitate profiling hospital value (payments and quality).

The developer believes that the MSPB-Hospital measure submission did meet the requirements of the NQF's SDS trial period and the NQF committee confirmed this by passing the MSPB-Hospital measure on the Scientific Acceptability criterion. The developer noted in the original submission that the inclusion of SDS factors (i.e., family income-to-poverty ratio and race) had a minimal impact on hospital's measure scores. The developer recognizes the commenter's concerns that additional factors could be included in the SDS measure testing. The developer selected family income-to-poverty ratio to strike a balance between the individual and community factors related to SES and listed by the commenter, as individual family members may pool financial resources to provide care for older relatives. The developer also conducted additional analyses based on feedback from the NQF committee to examine the impact of including a dual eligibility flag in the risk adjustment model, which are included in the appendix of the measure testing form that was submitted to the NQF on March 31st, 2017. These

analyses showed that including a dual eligibility flag had a low impact on MSPB-Hospital measure scores and that hospitals on the tails of score distributions were not disproportionately affected. A recent ASPE report showed some differences in measure performance between hospitals with a high amount of Disproportionate Share Hospital payments and a low amount.* The analysis in the appendix's Supplementary Table 7 suggests that these differences may be driven by hospitals with a very high concentration of dual eligible beneficiaries (above 60%), and that measure scores are high for both duals and non-duals in these hospitals. This suggests that these hospitals are relatively higher-cost hospitals for all types of patients.

*Office of the Assistant Secretary for Planning and Evaluation (ASPE). "Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs." December, 2016. Available at https://aspe.hhs.gov/system/files/pdf/253971/ASPESESRTCfull.pdf.

Response 3: Acumen/CMS

The MSPB-Hospital measure developer appreciates the commenter's feedback on the separate clinician-level measure (MSPB-TIN) used in the Merit-Based Incentive Payment System (MIPS). The developer would like to clarify that while MSPB-TIN and the facility-level MSPB-Hospital measure currently under consideration for NQF re-endorsement are alike, MSPB-TIN differs in attribution methodology. The MSPB-TIN measure is still under reevaluation. To ensure the reliability and validity of the measures being implemented, CMS reevaluates the measures annually and plans NQF submission of the measures by taking into account program needs and measure implementation timelines to meet the statutory requirements. MSPB-TIN was finalized for inclusion in the MIPS Cost Category as part of the Quality Payment Program Final Rule (https://www.federalregister.gov/documents/2016/11/04/2016-25240/medicare-program-merit-based-incentive-payment-system-mips-and-alternative-payment-model-apm). The first performance period for the Cost Category is calendar year 2017, and the category is weighted at zero percent for the associated payment year, meaning that it will not impact payments under the program in its first year.

Proposed Committee Responses:

The Committee agrees that the measure is only validated and recommended for use at the facility level, and needs further testing before it can be considered for endorsement at the physician level.

The Committee has reviewed the comments and appreciates the additional insights on the measure. After reviewing the comments and responses from the developer, the Committee believes this measure is appropriately specified and tested and continues to meet the criteria for NQF endorsement.

Action Item:

Does the Committee agree with the proposed response regarding the comments on facility level endorsement?

The Committee should review the comments submitted on testing and SDS, and the developer's responses, and discuss.