TO: Consensus Standards Approval Committee (CSAC)

FR: Ashlie Wilbon, Quintin Dukes

RE: Cost and Resource Use Phase 3 Pulmonary Condition-Specific Measures: Member Voting Results

DA: November 12, 2014

The CSAC will review recommendations from the *Cost and Resource Use Phase 3 Pulmonary* project during its November 12 meeting.

This memo includes a summary of the project, recommended measures, public and member comment themes and their responses.

This project followed the National Quality Forum's (NQF) version 1.9 of the Consensus Development Process (CDP). Member voting on these recommended measures ended on October 21.

Accompanying this memo are the following documents:

1. Cost and Resource Use Phase 3 Pulmonary Draft Report. The draft report has been updated to reflect the changes made following Standing Committee discussion of public and member comments. The complete draft report and supplemental materials are available on the project page.
2. Comment Table. Staff has identified themes among the comments received. This table lists 18 comments received and the NQF/Standing Committee responses.

CSAC ACTION REQUIRED
Pursuant to the CDP, the CSAC may consider approval of 3 candidate consensus standards.

Cost and Resource Use Phase 3 Pulmonary Measures Recommended for Endorsement:

- 1560: Relative Resource Use for People with Asthma (NCQA)
- 1561: Relative Resource Use for People with COPD (NCQA)
- 2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care pneumonia (CMS/Yale)

Based on the recent Board of Directors decision to conditionally endorse two hospital-level, risk standardized payment episode of care cost measures for AMI and HF from phase two, it is recommended that measure #2579 be approved with the same conditions under which the cardiovascular cost measures were endorsed:

- **One-year look-back assessment of unintended consequences:** NQF staff will work with Cost and Resource Use Standing Committee and CMS to determine a plan for assessing potential unintended consequences of this measure in use. The evaluation of unintended consequences will be initiated in approximately one year and possible changes to the measures based on these data will be discussed at that time.

- **Consideration for the SDS trial period:** The Cost and Resource Use Standing Committee will consider whether the measure should be included in the NQF trial period for sociodemographic status adjustment.
- **Attribution:** NQF will consider opportunities to address the attribution issue.

BACKGROUND

The Cost and Resource Use is a three phase project that seeks to identify and endorse performance measures relative to total cost, efficiency, and risk-adjusted Relative Resource Use (RRU). In the first phase, a non-condition specific measure of total cost using a per-hospitalization episode approach for the Medicare population was endorsed. The second phase focuses on risk-standardized payment and relative resource use for cardiovascular condition-specific measures; subsequently, the third phase has similar performance measures on risk-standardized payment and relative resource use for pulmonary conditions.

The Cost and Resource Use Standing Committee reviewed measures: #1560: Relative Resource Use for People with Asthma (NCQA), #1561: Relative Resource Use for People with COPD (NCQA), and #2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care for pneumonia. All three performance measures were recommended for endorsement and progressed through the NQF Member and Public comment period. Pursuant to CDP process guidance, all three measures were posted for NQF member vote.

DRAFT REPORT

The Cost and Resource Use Phase 3 Pulmonary Draft report presents the results of the evaluation of three measures considered under the CDP. All three measures are recommended for endorsement as voluntary consensus standards suitable for accountability and quality improvement. The measures were evaluated against the NQF Resource Use Measure Evaluation Criteria.

	MAINTENANCE	NEW	TOTAL
Measures considered	2	1	3
Recommended	2	1	3
Not recommended	0	0	0
Reasons not Recommended	Importance- N/A Scientific Acceptability- N/A Overall- N/A Competing Measure- N/A	Importance- N/A Scientific Acceptability- N/A Overall- N/A Competing Measure- N/A	

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 June 12th through June 30th for the three measures under review. A total of ten pre-evaluation comments were received, pertaining to the two NCQA relative resource use maintenance measures for asthma and COPD and the newly proposed CMS and Yale risk-standardized payment pneumonia measure. These pre-evaluation comments were provided to the Committee prior to their initial deliberations held during the workgroups calls.

Post-evaluation comments

The draft report went out for public and member comment August 14th to September 12th. During this commenting period, NQF received 18 comments from 7 member organizations.

A table of comments submitted during the comment periods, with the responses to each comment and the actions taken by the Standing Committee and measure developers, is posted to the Cost and Resource Use project page under the Materials section.

Measure-specific Comments Themes and Committee Responses

Comments about specific measure specifications and/or rationale were forwarded to the developers, who were invited to respond.

At its review of all comments, the Standing Committee had the benefit of developer responses. Committee members focused their discussion on measures or topic areas with the most significant and recurring issues.

1560: Relative Resource Use for People with Asthma & 1561: Relative Resource Use for People with COPD

Theme 1: Reliability and Validity

Commenters raised concern about the validity and reliability of both measures. In particular, they noted that neither measure adequately measures the total cost of pulmonary conditions like asthma and COPD and questioned stability of the measure with lower sample sizes. The incidence of severe asthma and COPD cases is rare and treatment for patients consumes few resources. Further, health plans will have difficulty evaluating the quality and efficiency of care for asthma and COPD. Relative resource use cost measures do not adequately assess efficiency and total costs for specific conditions like asthma and COPD due to the low incidence of severe cases. Commenters proposed that the measure specification exclusions should not include all high cost diagnoses.

> **Developer Response (#1560):** The RAS measure is limited to capturing the resources used by health plan members with persistent asthma. Members are identified as having persistent asthma through claims using a NQF endorsed, validated algorithm. NCQA's Relative Resource Use measures do not measure cost. The current risk adjustment approach provides a satisfactory O/E variance at the conservative min sample size of 200 eligible members and is very specific with regard to assigning health plan members to risk cohorts based on data available in administrative claims. The purpose of the measure is not to map resources to severity, rather to compare health plans' resource use managing their members with persistent asthma with other plans in their peer group.
>
> **Developer Response (#1561):** NCQA's Relative Resource Use measures do not measure cost. The current risk adjustment approach provides a satisfactory O/E variance at the conservative min sample size of 200 eligible members and is very specific with regard to assigning health plan members to risk cohorts based on data available in administrative claims. The purpose of the

measure is not to map resources to severity or other co-morbidities, rather to compare health plans' resource use managing their members with COPD with other plans in their peer group.

Committee Response: The Committee has weighed each of these concerns in their deliberations to evaluate this and other relative resource use measures. Based on the NQF criteria for reliability and validity, the Committee agreed the measures have met these criteria and affirmed the developer's responses in relation to the reliability and validity of measures #1560 and #1561.

Theme 2: Usability

While some commenters were not in support of using both measures for public reporting and a decision-making tool for consumers, others indicated strong support of these measures for use by health plans. Those expressing concerns with the usability of the measure noted that the limited usability of this measure would negatively impact both health plans and consumers. Those in support of this measure and its usability noted that this measure facilitates a collaborative network between health plans and providers in order to improve measure results.

Committee Response: The Committee also weighed these benefits and challenges with the measures' usability when evaluating these measures. Given that the intent of these measures as specified is to measure the cost of care from the health plan perspective to care for asthmatics and those with COPD, and the current widespread use of these measures by health plans, the Committee ultimately recommended the measures.

Theme 3: Risk Adjustment

During the evaluation of these measures by the Committee, some committee members raised concern with the risk adjustment approach and requested additional clarification from the developers on their approach to risk adjusting and testing the risk model. The r^2 values for both measures were 0.48. This led to questions of the developers to further describe how the value was attained and what it represents. In particular, there were concerns that the current risk adjustment model is unable to discriminate within a specified health condition (i.e., asthma or COPD), as opposed to discriminating across them; by testing the model on a heterogeneous population (including members with asthma, COPD and cardiovascular conditions) it becomes difficult to discern what is causing the variation. This also impacts the coefficients used in the model and raised questions on how those coefficients may have been assigned.

Developer Response: The NCQA developers provided a response to these concerns: NCQA Response

Committee Response: The Committee generally accepted the developer's rationale for pooling data for health plan members across all five chronic conditions to estimate the regression for total annual spending. The group generally agreed that since the measure seeks to profile the total cost of all medical services for health plan members that this approach was reasonable. Some members of the Committee urged the developer to reconsider this design approach in updates to the measure.

2579: Hospital-level, Risk-standardized Payment Associated with a 30-day Episode of Care for Pneumonia

Theme 1: Appropriateness of the Attribution Approach

Commenters raised concern about the attribution approach for hospitalized patients with pneumonia, suggesting that the approach s is inappropriate and only reflects an episode-of-care and not the care of multiple providers across the health care delivery system. Commenters stated that measures should assess processes and outcomes over which the measured entity (e.g., hospital, physician group) can exercise a reasonable level of control, and that these measures may be more appropriate for an organization accepting bundled payments on behalf of all measured entities.

Committee Response: The Committee acknowledges and shares this concern; however, they also stated that hospitals are increasingly responsible for care delivered up to 30 days after discharge. Consequently, hospitals are in the unique position of being able to push coordination of care, and this measure may serve as an impetus for this to occur.

Theme 2: Risk Adjustment for Sociodemographic Status

A few commenters noted that it would be appropriate to stratify the claims to calculate this measure by sociodemographic status (SDS). The purpose for integrating SDS (i.e. low income, poor housing, no/low access to social service, and no/low access to community supports) is to document their negative impact on patient outcomes. These commenters expressed concern about penalizing providers for poor patient outcomes that are exacerbated by non-clinical factors.

Committee Response: The Committee recognizes the importance of adequately adjusting for sociodemographic status in the appropriate applications. While NQF continues to work on their implementation of the guidance from the SDS Expert Panel, measures currently under review have been recommended with additional guidance to stratify for SDS, as appropriate.

Theme 3: Validity of Exclusions

A commenter raised a concern about the specification of the measure and proposes the inclusion of the ICD-9 code 507.0 in the denominator for aspiration pneumonia. This code is used for 15% of Medicare patients discharged with pneumonia and this will address the poor risk adjustment for high cost patients that are hospitalized for pneumonia.

The exclusions of admissions for this measure does not provide clinical significance; currently, same day discharges or one day length of stays are not included within the analysis, when these time points could be indicative of highly efficient care.

> **Developer Responses:** We appreciate the commenter's concern. To clarify, the pneumonia cohort specifications were closely aligned with the 30-day pneumonia mortality measure which is NQF endorsed and publicly reported. Initially, aspiration pneumonias were considered a potential complication of care and therefore, they were not included in the pneumonia cohort. However, given the prevalence of this code, we plan to reevaluate including aspiration pneumonia in this measure in the future. We appreciate the thoughtful comment and agreement with our decision to exclude same- or next-day discharges from the measure. Given the shift over time to rapid treatment and timely care, we do plan to reevaluate these exclusion criteria in the future.

Committee Responses: Based on the NQF criteria for validity, the Committee has agreed that this measure has met the criteria for validity and has recommended it for endorsement. A few committee members support the inclusion of ICD-9 code 507.0 within this measure, which will assist with documenting the presence of aspiration pneumonia among admission.

NQF MEMBER VOTING RESULTS

Representatives from 18 member organizations submitted votes for the three measures recommended for endorsement by the Cost and Resource Use Standing Committee. The two NCQA measures were approved with 80 percent approval of member councils. The third measure receiving votes (#2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care for pneumonia) received approval from 50 percent of member councils and therefore consensus among the membership was not reached for this measure. No votes were received from Public/Community Health or the Supplier/Industry Councils. Results for each measure are provided below. (Links are provided to the full measure summary evaluation tables.)

NQF Member Council	Voting Organizations	Eligible to Vote	Rate
Consumer	2	28	7%
Health Plan	4	15	27%
Health Professional	4	87	5%
Provider Organizations	2	134	1%
Public/Community Health Agency	0	33	0%
Purchaser	4	24	17%
QMRI	2	69	3%
Supplier/Industry	0	29	0%
All Councils	**18**	**419**	**7%**

#1560 RELATIVE RESOURSE USE FOR PEOPLE WITH ASTHMA

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	2	0	0	2	100%
Health Plan	4	0	0	4	100%
Health Professional	0	1	3	4	0%
Provider Organizations	0	0	2	2	
Public/Community Health Agency	0	0	0	0	
Purchaser	4	0	0	4	100%
QMRI	1	0	1	2	100%
Supplier/Industry	0	0	0	0	
All Organizations	**11**	**1**	**6**	**18**	**92%**
Percentage of councils approving (>60%)					**80%**
Average council percentage approval					80%

*equation: Yes/ (Total - Abstain)

Voting Comments:

- American College of Nurse-Midwives: It is hard to understand how these measures made the top priority list when CHILDBIRTH is perhaps the most wasteful, resource intensive episode of care. Beginning of life care should be the priority for this project and measure.

#1561 RELATIVE RESOURCE USE FOR PEOPLE WITH COPD

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	2	0	0	2	100%
Health Plan	4	0	0	4	100%
Health Professional	0	1	3	4	0%

Provider Organizations	0	0	2	2	
Public/Community Health Agency	0	0	0	0	
Purchaser	4	0	0	4	100%
QMRI	1	0	1	2	100%
Supplier/Industry	0	0	0	0	
All Organizations	**11**	**1**	**6**	**18**	**92%**
Percentage of councils approving (>60%)					**80%**
Average council percentage approval					80%

*equation: Yes/ (Total - Abstain)

#2579 HOSPITAL-LEVEL RISK-STANDARDIZED PAYMENT ASSOCIATED WITH A 30-DAY EPISODE OF CARE FOR PNEUMONIA

Member Council	Yes	No	Abstain	Total Votes	% Approval*
Consumer	2	0	0	2	100%
Health Plan	4	0	0	4	100%
Health Professional	0	2	2	4	0%
Provider Organizations	0	2	0	2	0%
Public/Community Health Agency	0	0	0	0	
Purchaser	4	0	0	4	100%
QMRI	1	1	0	2	50%
Supplier/Industry	0	0	0	0	
All Organizations	**11**	**5**	**2**	**18**	**69%**
Percentage of councils approving (>60%)					**50%**
Average council percentage approval					58%

*equation: Yes/ (Total - Abstain)

Voting Comments:

- AAMC: The Association of American Medical Colleges (AAMC) thanks the NQF for the opportunity to vote on the Cost and Resource Use, Phase 3 measures. The AAMC has specific concerns with the hospital-level risk-standardized payment associated with a 30-day episode-of-care for pneumonia measure due to the lack of sociodemographic status (SDS) factors in the risk-adjustment methodology.

  During the Steering Committees discussion of this measure, members voiced concerns that SDS factors should be included in the measure methodology to account for patient characteristics that are outside of the control of the hospital. The measure currently does not adjust for non-clinical factors, such as income, housing, access to social services, and community supports even though there is robust evidence that such factors affect health outcomes, including resource

use.

The NQF Board of Directors recently approved the implementation of a trial period to adjust performance measures for SDS factors; however the trial period has not yet started. The AAMC asks that the Committee postpone further action on this measure until the SDS trial period has concluded and the NQF has had sufficient time to adopt a new policy relating to an SDS adjustments. In the meantime, the AAMC recommends that NQF consider the inclusion of the pneumonia 30 day episode of care measure as part of the SDS trial period so that a proper evaluation can be made.

Measure Evaluation Summary Tables

LEGEND: Y = Yes; N = No; H = High; M = Moderate; L = Low; I = Insufficient

1560 RELATIVE RESOURCE USE FOR PEOPLE WITH ASTHMA
Submission | Specifications
Description: The risk-adjusted relative resource use by health plan members with asthma during the measurement year. **Resource Use Measure Type:** Per capita (population- or patient-based) **Level of Analysis:** Health Plan **Costing Method:** Standardized pricing **Target population**: Populations at Risk **Data Source**: Administrative claims **Measure Steward**: National Committee for Quality Assurance
STANDING COMMITTEE MEETING [06/25/2014] **1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>** *(1a. High Priority, 1b. Opportunity for Improvement, 1c. Measure Intent)* 1a. High Priority: **H-19**; **M-3**; **L-0**; **I-0; IE-0**; 1b. Opportunity for Improvement: **H-13**; **M-9**; **L-0**; **I-0;** 1c. Measure Intent: **H-10**; **M-9**; **L-2**; **I-0** 1. Overall Importance: **H-16; M-5; L-1; I-0** <u>Rationale</u>: • The Committee stated that asthma is a prevalent and costly condition, affecting more than 23 million Americans and accounting for over $20 billion spent annually on health care in the United States. • The developer provided data on performance trends for Commercial and Medicaid plans demonstrating significant variation in health plan resource use from an overall perspective and with respect to specific service areas and regions, which the Committee agreed indicated a substantial opportunity for improvement. • The Committee questioned whether trend data for health plans was available to enable health plans to understand which areas to investigate for potential cost reductions. The developer stated that there is not trend data available, as that would require actual prices and patient populations to be standardized year to year. • The Committee and the Technical Expert Panel's opinion that asthma is a condition for which disparities impact outcomes was substantiated by the evidence submitted by the developer demonstrating disparities; these studies indicated that race/ethnicity, socioeconomic class, and health insurance status impacted utilization for asthma related services. • The Committee found that the measure intent was clear; that is, to reduce variation in risk adjusted resource use among patients with asthma. However, the Committee stated that, given that the measure captures all costs for asthma patients in a given year, and the proportion of patient costs associated with asthma is unclear, the measure may have been better specified using an episode-based approach. The developer stated that this approach was selected because parsing out which episode costs should or should not be attributed to the condition was subject to much debate and little consensus among the developer-convened experts during the development process.

1560 RELATIVE RESOURCE USE FOR PEOPLE WITH ASTHMA

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-8**; **M-13**; **L-0**; **I-0** 2b. Validity: **H-5**; **M-17**; **L-0**; **I-0**
<u>Rationale</u>:

- The Committee stated that overall the measure is well defined with clear inclusion and exclusion criteria.
- The Committee reiterated the TEP concern that asthma is over diagnosed and questioned whether the specifications should allow for patients with any diagnosis of asthma or the proxy of filling a prescription for a medication used to treat asthma to move a patient into the denominator. The Committee questioned if the measure specifications should include objective verification that the patient has asthma in order to be counted in the denominator. The measure developer clarified that even though the measurement period is one year, the measure uses a two-year look back period to determine whether a patient should be included in the denominator. A patient must have a diagnosis of asthma each consecutive year and/or meet the criteria for the denominator over both years to be counted in the denominator; the developer believes this will reduce the probability of false positives being included in the measure population.
- The reliability testing provided by the developer was conducted at the data element level and at the performance measure score level. Testing results indicated that at the data element level, the mean percentage of dollars with acceptable coding across plans was 92.8%. At the performance measure score level, the developer submitted testing assessing whether plan rankings by quartile were stable year to year; the data presented indicated that plan performance compared to other health plans remained generally stable over time.
- The developer presented information describing the process for and results of assessing face validity for the measure, as well as empirical evidence of validity obtained from a study demonstrating that for a given health plan and clinical category, measures of relative resource utilization were generally similar across different types of service, with only some modest variations. The Committee found the information presented related to validity of the measure to be sufficient.
- Some Committee members expressed concern that the r-squared value for the risk adjustment model was .48, which was considered to be somewhat high. The Committee acknowledged that this issue of high r-squared values has occurred before, when a variety of diseases are included in the risk adjustment model; however, this does not allow for discrimination within a condition. For instance, a model which includes both young people who are not very sick with few comorbidities (e.g. asthma) versus older people who are quite sick and with a lot of comorbidities (e.g. COPD) may seem to offer great characteristics (e.g. high r-squared or c indices), but this means little. What you want is discriminate within asthma or within COPD, not across them. Further, they were concerned about the lack of clarity regarding what variance the r-square was capturing and on which risk adjustment model cost was being regressed As such, the Committee members requested clarity from the developers on both points and will consider any additional information during their call to review comments received during the NQF Member and Public comment period.
- The developer stated that the Hierarchical Clinical Conditions for Relative Resource Use (HCC-RRU) model was developed based on components of the CMS Hierarchical Clinical Conditions (CMS-HCC) risk adjustment methodology and accounts for age, gender, and HCC-RRU risk classifications that predict cost

1560 RELATIVE RESOURCE USE FOR PEOPLE WITH ASTHMA
variability. The developer stated that r-squared testing was done by comparing four different risk adjustment approaches including the HCC-RRU model; from this analysis, the developer determined that their risk adjustment model yielded similar observed to expected results to the other models across health plans. The developer posited that the r-squared values would be expected to be slightly higher, as the HCC categories are based on a Medicare population, and this measure population includes a broader age range. • The Committee questioned whether adjustments for sociodemographic status (SDS) factors should be incorporated into the risk adjustment model. NQF clarified that it is in the early stages of reviewing our policy on risk adjusting for SDS factors. The recommendations for modifying NQF's current policy on adjusting for SDS factors have not yet been finalized. As such, we ask that Committees continue to evaluate measures according to our current guidelines, that SDS factors are not included in the risk adjustment model, but are used to stratify the measure. If in the future the recommendations for adjusting for SDS factors become NQF policy, measures that may be improved from incorporating these adjustments will be updated and reviewed by the Committee through one of NQF's measure maintenance processes.
3. Feasibility: H-20; M-1; L-1; I-0 *(3a. Byproduct of Care Processes; and 3b. Electronic sources; and 3c. Data Collection Strategy)* Rationale: • The Committee stated that because the measure is already in use and is calculated using claims data at the health plan level as part of collecting HEDIS data, the measure is very feasible to implement.
4. Usability: H-7; M-13; L-2; 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: • The Committee agreed that at a high level this is a useful measure for health plans to look at their own data and see where they can make improvements. • The Committee questioned how consumers and patients would use the measure as the data is reported at the health plan level, which may not be granular enough for these stakeholders in particular. The developer acknowledged that this measure is less usable for consumers and patients. • The Committee questioned whether this measure would be actionable by health plans because the trend data is not available. The developer did not see this as a weakness of the measure, as the measure does allow for comparisons between health plans.
5. Related and Competing Measures • No related or competing measures noted.
Standing Committee Recommendation for Endorsement: **Y-22**; **N-0**
6. Public and Member Comment: August 14, 2014 – September 12, 2014 Comments received: • Comments addressed the inability of both measures to adequately assess efficiency and total costs for specific conditions like asthma and COPD, as well as the stability of the measures at smaller sample sizes. .

1560 RELATIVE RESOURCE USE FOR PEOPLE WITH ASTHMA
• Five comments were received regarding usability of this measure. Some commenters were in support of these measures being used by health plans; others were concerned with the limited actionability for other stakeholders and were not in support of measures used for public reporting and a decision-making tool for consumers. • During their deliberations, the Committee raised concern regarding the risk adjustment model for this measure and reviewed the developer's response to these issues during their comment call. The concern was that the r^2 values for both measures were 0.48 and the current risk model adjustment model is unable to discriminate within a specified health condition (i.e., asthma or COPD), as opposed to discriminating across them; by testing the model on a heterogeneous population (including members with asthma, COPD and cardiovascular conditions) it becomes difficult to discern what is causing the variation. Developer response: • NCQA addressed the concerns of commenters and identified that these measures are not intended to measure cost or severity of asthma or COPD; therefore, these measures assess a health plan's resource use of a member with asthma or COPD and compare the resource use of the health plan's peer group. • NCQA submitted a response to the Committees concerns used a single regression model to define risk strata and relationship between HCCs and cost; asthma cases were assessed due to low severity, COPD cases have broader mix of severity, and for both conditions costs rise substantially with patient severity. Committee responses: • The Committee generally accepted the developer's rationale for pooling data for health plan members across all five chronic conditions to estimate the regression for total annual spending. The group generally agreed that since the measure seeks to profile the total cost of all medical services for health plan members that this approach was reasonable. Some members of the Committee urged the developer to reconsider this design approach in updates to the measure. • The Committee also weighed these benefits and challenges with the measures' usability when evaluating these measures. Given that the intent of these measures as specified is to measure the cost of care from the health plan perspective to care for asthmatics and those with COPD, and the current widespread use of these measures by health plans, the Committee ultimately recommended the measures.
7. NQF Member Voting: October 2014
8. Consensus Standards Approval Committee (CSAC) Vote: November 2014
9. Board of Directors Vote: November 2014
10. Appeals: December 2014

1561 RELATIVE RESOURCE USE FOR PEOPLE WITH COPD
Submission | Specifications
Description: The risk-adjusted relative resource use by health plan members with COPD during the measurement year. **Resource Use Measure Type:** Per capita (population- or patient-based) **Level of Analysis:** Health Plan, Integrated Delivery System, Population : National, Population : Regional **Costing Method:** Standardized pricing **Target population**: Populations at Risk **Data Source**: Administrative claims **Measure Steward**: National Committee for Quality Assurance

1561 RELATIVE RESOURCE USE FOR PEOPLE WITH COPD

STANDING COMMITTEE MEETING [06/25/2014]

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

(1a. High Priority, 1b. Opportunity for Improvement, 1c. Measure Intent)

1a. High Priority: **H-19**; **M-2**; **L-1**; **I-0; IE-0**; 1b. Opportunity for Improvement: **H-13**; **M-8**; **L-1**; **I-0;** 1c. Measure Intent: **H-11**; **M-9**; **L-2**; **I-0** 1. Overall Importance: **H-18; M-3; L-1; I-0**

<u>Rationale:</u>

- The Committee stated that COPD is a prevalent and costly condition. COPD affects more than 12 million people who have been diagnosed with COPD and another 12 million who are not aware they have the disease; it is the fourth leading cause of death in the United States. COPD also accounts for over $18 billion spent annually on health care in the United States.
- The developer provided data on performance trends for Commercial and Medicaid plans demonstrating significant variation in health plan resource use from an overall perspective and with respect to specific service areas and regions, which the Committee agreed indicated a substantial opportunity for improvement.
- The Committee questioned whether trend data for health plans was available to enable health plans to understand which areas to investigate for potential cost reductions. The developer stated that there is not trend data available, as that would require actual prices and patient populations to be standardized year to year.
- The Committee found that the measure intent was clear; that is, to reduce variation in risk adjusted resource use among patients with COPD. However, the Committee stated that, given that the measure captures all costs for COPD patients in a given year, and the proportion of patient costs associated with COPD is unclear, the measure may have been better specified using an episode-based approach. The developer stated that this approach was selected because parsing out which episode costs should or should not be attributed to the condition was subject to much debate and little consensus.

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-6**; **M-14**; **L-2**; **I-0** 2b. Validity: **H-4**; **M-17**; **L-1**; **I-0**

<u>Rationale</u>:

- The Committee stated that overall the measure is well defined with clear inclusion and exclusion criteria.
- The reliability testing provided by the developer was conducted at the data element level and at the performance measure score level. Testing results indicated that at the data element level, the mean percentage of dollars with acceptable coding across plans was 92.8%. At the performance measure score level, the developer submitted testing assessing whether plan rankings by quartile were stable year to year; the data presented indicated that plan performance compared to other health plans remained generally stable over time.
- The developer presented information describing the process for and results of assessing face validity for the measure, as well as empirical evidence of validity obtained from a study demonstrating that for a given health plan and clinical category, measures of relative resource utilization were generally similar across different types of service, with only some modest variations. The Committee found the information presented related to validity of the measure to be sufficient.

1561 RELATIVE RESOURCE USE FOR PEOPLE WITH COPD

- Some Committee members expressed concern that the r-squared value for the risk adjustment model was .48, which was considered to be somewhat high. The Committee acknowledged that this issue of high r-squared values has occurred before, when a variety of diseases are included in the risk adjustment model; however, this does not allow for discrimination within a condition. For instance, a model which includes both young people who are not very sick with few comorbidities (e.g. asthma) versus older people who are quite sick and with a lot of comorbidities (e.g. COPD) may seem to offer great characteristics (e.g. high r-squared or c indices), but this means little. What you want is discriminate within asthma or within COPD, not across them. Further, they were concerned about the lack of clarity regarding what variance the r-square was capturing and on which risk adjustment model cost was being regressed As such, the Committee members requested clarity from the developers on both points and will consider any additional information during their call to review comments received during the NQF Member and Public comment period.
- The developer stated that the Hierarchical Clinical Conditions for Relative Resource Use (HCC-RRU) model was developed based on components of the CMS Hierarchical Clinical Conditions (CMS-HCC) risk adjustment methodology and accounts for age, gender, and HCC-RRU risk classifications that predict cost variability. The developer stated that r-squared testing was done by comparing four different risk adjustment approaches including the HCC-RRU model; from this analysis, the developer determined that their risk adjustment model yielded similar observed to expected results to the other models across health plans. The developer posited that the r-squared values would be expected to be slightly higher, as the HCC categories are based on a Medicare population, and this measure population includes a broader age range.
- The Committee questioned whether adjustments for sociodemographic status (SDS) factors should be incorporated into the risk adjustment model. NQF clarified that it is in the early stages of reviewing our policy on risk adjusting for SDS factors. The recommendations for modifying NQF's current policy on adjusting for SDS factors have not yet been finalized. As such, we ask that Committees continue to evaluate measures according to our current guidelines, that SDS factors are not included in the risk adjustment model, but are used to stratify the measure. If in the future the recommendations for adjusting for SDS factors become NQF policy, measures that may be improved from incorporating these adjustments will be updated and reviewed by the Committee through one of NQF's measure maintenance processes.

3. Feasibility: H-18; M-4; L-0; I-0

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

Rationale:

- The Committee stated that because the measure is already in use and is calculated using claims data at the health plan level as part of collecting HEDIS data, the measure is very feasible to implement.

1561 RELATIVE RESOURCE USE FOR PEOPLE WITH COPD

4. Usability: H-8; M-13; L-1; 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:

- The Committee agreed that at a high level this is a useful measure for health plans to look at their own data and see where they can make improvements.
- The Committee questioned how consumers and patients would use the measure as the data is reported at the health plan level, which may not be granular enough for these stakeholders in particular. The developer acknowledged that this measure is less usable for consumers and patients.
- The Committee questioned whether this measure would be actionable by health plans because the trend data is not available. The developer did not see this as a weakness of the measure, as the measure does allow for comparisons between health plans.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: Y-20; N-2

6. Public and Member Comment: August 14, 2014 – September 12, 2014

Comments received:

- Comments addressed the inability of both measures to adequately assess efficiency and total costs for specific conditions like asthma and COPD, as well as the stability of the measures at smaller sample sizes. .
- Five comments were received regarding usability of this measure. Some commenters were in support of these measures being used by health plans; others were concerned with the limited actionability for other stakeholders and were not in support of measures used for public reporting and a decision-making tool for consumers.
- During their deliberations, the Committee raised concern regarding the risk adjustment model for this measure and reviewed the developer's response to these issues during their comment call. The concern was that the r^2 values for both measures were 0.48 and the current risk model adjustment model is unable to discriminate within a specified health condition (i.e., asthma or COPD), as opposed to discriminating across them; by testing the model on a heterogeneous population (including members with asthma, COPD and cardiovascular conditions) it becomes difficult to discern what is causing the variation.

Developer response:

- NCQA addressed the concerns of commenters and identified that these measures are not intended to measure cost or severity of asthma or COPD; therefore, these measures assess a health plan's resource use of a member with asthma or COPD and compare the resource use of the health plan's peer group.
- NCQA submitted a response to the Committees concerns used a single regression model to define risk strata and relationship between HCCs and cost; asthma cases were assessed due to low severity, COPD cases have broader mix of severity, and for both conditions costs rise substantially with patient severity.

Committee responses:

- The Committee generally accepted the developer's rationale for pooling data for health plan members across all five chronic conditions to estimate the regression for total annual spending. The group generally agreed that since the measure seeks to profile the total cost of all medical services for health plan members that this approach was reasonable. Some members of the Committee urged the developer to reconsider this design approach in updates to the measure.

1561 RELATIVE RESOURCE USE FOR PEOPLE WITH COPD
• The Committee also weighed these benefits and challenges with the measures' usability when evaluating these measures. Given that the intent of these measures as specified is to measure the cost of care from the health plan perspective to care for asthmatics and those with COPD, and the current widespread use of these measures by health plans, the Committee ultimately recommended the measures.
7. NQF Member Voting: October 2014
8. Consensus Standards Approval Committee (CSAC) Vote: November 2014
9. Board of Directors Vote: November 2014
10. Appeals: December 2014

2579 HOSPITAL-LEVEL, RISK-STANDARDIZED PAYMENT ASSOCIATED WITH A 30-DAY EPISODE OF CARE FOR PNEUMONIA
Submission | Specifications
Description: This measure estimates hospital-level, risk-standardized payment for a pneumonia episode of care starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of pneumonia. **Resource Use Measure Type:** Per episode **Level of Analysis:** Facility **Costing Method:** Standardized pricing **Target population:** Senior Care **Data Source**: Administrative claims **Measure Steward**: Centers for Medicare & Medicaid Services
STANDING COMMITTEE MEETING [06/25/2014] **1. Importance to Measure and Report: <u>The measure meets the Importance criteria</u>** 1a. High Priority: **H-17**; **M-5**; **L-0**; **I-0; IE-0**; 1b. Opportunity for Improvement: **H-19**; **M-2**; **L-1**; **I-0;** 1c. Measure Intent: **H-18**; **M-4**; **L-0**; **I-0 1. Overall Importance: H-18; M-4; L-0; I-0** <u>Rationale</u>: • The Committee stated that the measure is high priority given that pneumonia is one of the leading causes of hospitalization for Medicare patients sixty-five years of age and older, with Medicare paying roughly ten billion dollars in aggregate costs for hospitalized beneficiaries with pneumonia. • The developer presented evidence indicating that there is a threefold variation in cost for the medical treatment of pneumonia patients, which the Committee agreed signified that there is a substantial opportunity for improving the overall costs for pneumonia patients. • The Committee stated that by using this measure in conjunction with a measure capturing the quality of care for pneumonia patients, there is an opportunity to begin to understand the value of the care provided by the hospitals and other providers in treating this condition.
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)*

2579 HOSPITAL-LEVEL, RISK-STANDARDIZED PAYMENT ASSOCIATED WITH A 30-DAY EPISODE OF CARE FOR PNEUMONIA

2a. Reliability: **H-10**; **M-11**; **L-1**; **I-0** 2b. Validity: **H-3**; **M-18**; **L-1**; **I-0**

Rationale:

- The Committee stated that the measure specifications were precise and that the measure was well-constructed. This measure captures risk-standardized payments for a thirty-day episode of care for Medicare patients diagnosed admitted to the hospital with a diagnosis of pneumonia through administrative claims data.
- The developer provided reliability testing at the level of the performance measure score; testing was performed by calculating the Intraclass Correlation Coefficient (ICC) score by calculating the risk standardized payment using a split-sample of the combined 2008-2009 data from hospitals. The ICC score was 0.825, indicating significant agreement between the two samples, which the Committee found sufficient.
- The Committee questioned the validity of specifying the measure for a thirty-day episode triggered by admission for pneumonia, as the treatment of pneumonia may require care coordination post-discharge that may extend past thirty days. The Committee stated that this could affect payments captured during the post-discharge period, artificially inflating or deflating the costs for some patients simply because of the construct of the measure.
- The Committee raised concerns regarding the attribution approach and the implications for attribution of costs if a patient were transferred to another hospital. The developer clarified that only 0.4 percent of cohorts are transferred for pneumonia, which represents a small number of beneficiaries. In the case of transfer patients, costs for the patient will be attributed to the initial admitting hospital, as hospitals are increasingly responsible for care delivered up to 30 days after discharge. The Committee found this approach to attribution to be acceptable.
- The Committee stated concern that the low r-squared value (.07) for the risk model may indicate that case mix is not being appropriately adjusted for through the risk model. The developer clarified that at lower patient volumes, there is less certainty when estimating cost. The measure uses a continuous outcome which results in a more accurate estimate than would result from a binary outcome. Additionally, the measure uses hierarchical risk modeling that adjusts hospitals with low patient volume towards the mean. The Committee found this explanation to be sufficient.
- The Committee questioned whether adjustments for sociodemographic status (SDS) factors should be incorporated into the risk adjustment model. NQF clarified that it is in the early stages of reviewing our policy on risk adjusting for SDS factors. The recommendations for modifying NQF's current policy on adjusting for SDS factors have not yet been finalized. As such, we ask that Committees continue to evaluate measures according to our current guidelines, that SDS factors are not included in the risk adjustment model, but are used to stratify the measure. If in the future the recommendations for adjusting for SDS factors become NQF policy, measures that may be improved from incorporating these adjustments will be updated and reviewed by the Committee through one of NQF's measure maintenance processes.

2579 HOSPITAL-LEVEL, RISK-STANDARDIZED PAYMENT ASSOCIATED WITH A 30-DAY EPISODE OF CARE FOR PNEUMONIA

3. Feasibility: H-20; M-2; L-0; I-0

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

Rationale:

- The Committee stated that this measure is feasible to implement because the measure is specified using administrative claims data which is created as a byproduct of care delivery and available electronically.

4. Usability: H-10; M-11; L-1; 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:

- The Committee found the measure to be useful for providers, giving them access to detailed data of cost for hospital care for pneumonia.
- The Committee questioned the availability of information on costs for providers other than the hospital to which the patient has been attributed, stating that for this measure to be most useful there needs to be documentation of the reimbursement amounts for each provider treating the patient.

5. Related and Competing Measures

- No related or competing measures noted.

Standing Committee Recommendation for Endorsement: **Y-21**; **N-1**

6. Public and Member Comment: August 14, 2014 - September 12, 2014

Comments received:

- One measure-specific comment was received regarding the appropriateness of the attribution approach for measure #2579. The commenter suggested that the current attribution approach is inappropriate and only reflects an episode-of-care attributed to a hospital as the responsible entity and does not account for the care of multiple providers across the health care delivery system. The commenter suggested this approach would be more appropriate for an integrated health system or an organization accepting bundled payments.
- Two comments regarding risk adjustment for sociodemographic status for this measure. Some commenters believed that it would be appropriate to stratify claims by sociodemographic factors and document non-clinical elements that negatively impact patient outcomes when calculating risk adjusted costs.
- One measure-specific comment was received regarding validity of exclusions for measure this measure. A commenter proposed the inclusion of ICD-9 code 507.0 in the denominator for aspiration pneumonia, which was estimated to account for 15% of Medicare patients discharged with pneumonia.

Developer response:

- Yale addressed the concern of integrating the ICD-9 code 507.0 in the denominator for aspiration pneumonia and based on the prevalence of the code, developers will plan to reevaluate including aspiration pneumonia in future versions of the measure.

Committee responses:

- The Committee acknowledged and many shared the concerns with the attribution approach used in this measure; however, they also stated that hospitals are increasingly responsible for care delivered up to 30

2579 HOSPITAL-LEVEL, RISK-STANDARDIZED PAYMENT ASSOCIATED WITH A 30-DAY EPISODE OF CARE FOR PNEUMONIA
days after discharge. Consequently, hospitals are in the unique position of being able to push coordination of care, and this measure may serve as an impetus for this to occur. • The Committee recognizes the importance of adequately adjusting for sociodemographic status in the appropriate applications. While NQF continues to work on their implementation of the guidance from the SDS Expert Panel, measures currently under review have been recommended with additional guidance to stratify for SDS, as appropriate. • Based on the NQF criteria for validity, the Committee has agreed that this measure has met the criteria for validity and has recommended it for endorsement. A few committee members support the inclusion of ICD-9 code 507.0 within this measure, which will assist with documenting the presence of pneumonia aspiration among admission.
7. NQF Member Voting: October 2014
8. Consensus Standards Approval Committee (CSAC) Vote: November 2014
9. Board of Directors Vote: November 2014
10. Appeals: December 2014