NATIONAL QUALITY FORUM

IN PERSON MEETING OF THE RESOURCE USE PULMONARY TECHNICAL ADVISORY PANEL (TAP)

July 19, 2011

Committee Members Participating: Kurtis Elward, MD, MPH (*Co-Chair*), Family Medicine of Albermarle; Janet Maurer, MD, MBA (*Co-Chair),* American College of Chest Physicians; Gerene Bauldoff, PhD, RN, The Ohio State University, School of Nursing; Kathryn Blake, PharmD, Nemours Children's Clinic; Dale Bratzler, DO, MPH, Oklahoma Foundation for Medical Quality; Zab Mosenifar, MD, Cedars Sinai Medical Center; Linus Santo Tomas, MD, MS, Pulmonary & Critical Care, Medical College of Wisconsin; Michael Schatz, MD, MS, Kaiser Permanente; Richard Stanford, PharmD, GlaxoSmithKline.

NQF Staff Participating: Lauralei Dorian, Project Manager; Taroon Amin, MPH, MA, Senior Director; Ashlie Wilbon, MPH, BSN, Senior Project Manager; Sarah Fanta, Project Analyst; Heidi Bossley, MSN, MBA, Vice President, Performance Measures; Carlos Alzola, NQF Statistical Consultant; Sally Turbyville, MA, MS, NQF Project Consultant.

Measure developers present in person or via the telephone: Benjamin Hamlin, MPH, NCQA; Thomas Lynn, MD, Ingenix; Cheri Zielinsky, Ingenix.

Others present: Sheila Heitzig, Director of Practice and Policy, American Academy of Allergy, Asthma and Immunology.

MEETING PROCESS

Ms. Dorian welcomed the Pulmonary Technical Advisory Panel (TAP) and thanked them for their participation. The purpose of the one-day meeting was to evaluate using the Resource Use Consensus Development Process (CDP) five Pulmonary measures: two submitted by NCQA and three by Ingenix.

The measure developers were available in person and via telephone to respond to questions from the Panel as needed. A NQF Member and public comment period occurred at the end of the call; no comments were made at that time. The audio recordings and general project information can be found by clicking on the Resource Use project page.

MEASURE EVALUATION SUMMARY

The following summary includes a preliminary review of the measures submitted by NCQA and Ingenix:

1560 Relative Resource Use for People with Asthma (NCQA)
Description: This measure addresses the resource use of members identified as having asthma. Both encounter and pharmacy data are used to identify members for inclusion in the eligible population, and the results are adjusted to account for age, gender, and HCC-RRU risk classifications that predict cost variability (Refer to Attachment S8_Clinical Logic for additional information). Resource Use Type: Per capita (population- or patient-based)

Data Type: Administrative claims; Electronic Clinical Data : Electronic Health Record; Electronic Clinical Data : Imaging/Diagnostic Study; Electronic Clinical Data : Laboratory; Electronic Clinical Data : Pharmacy Paper Records
Resource Use Service Categories: Inpatient services: Inpatient facility services, Inpatient services: Evaluation and management, Inpatient services: Procedures and surgeries, Inpatient services: Imaging and diagnostic, Inpatient services: Lab services, Inpatient services: Admissions/discharges, Ambulatory services: Outpatient facility services, Ambulatory services: Emergency Department, Ambulatory services: Pharmacy, Ambulatory services: Evaluation and management, Ambulatory services: Procedures and surgeries, Ambulatory services: Imaging and diagnostic, Ambulatory services: Lab services
Care Setting: Ambulatory Care : Ambulatory Surgery Center (ASC), Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Hospital/Acute Care Facility, Imaging Facility, Laboratory, Pharmacy, Post Acute/Long Term Care Facility: Nursing Home/Skilled Nursing Facility
Level of Analysis: Clinician : Group/Practice, Health Plan, Integrated Delivery System, Population : National, Population : Regional
Measure Developer: National Committee for Quality Assurance (NCQA), 1100 13th Street NW, STE 1000, Washington, District Of Columbia, 20005

Conditions/Questions for Developer:

- Could this measure be improved by including other diagnostic criteria to ensure all appropriate asthma patients are captures?
- How have you come up with the age strata in your risk-adjustment?
- Can secondary diagnosis is taken into account within the measurement year?
- Is cost during the measurement year is part of the risk-adjustment strategy?
- Are your measure results published publically?

Developer Response:

- Using asthma as a principal diagnosis will make it difficult to identify most patients, especially those who are acute and come into the ER and are diagnosed with bronchitis first, and then asthma.
- The age strata for risk-adjustment are designed around known utilization patterns and clinical treatment patterns.
- *All* costs for anyone with asthma are counted.
- The HCC uses any services during the year to appropriately categorize patients into those 13 risk cohorts by severity of comorbidity. They also look at ICD-9 and procedural codes to categorize them and then go back and look at the number of times those services were offered to that population. Therefore, if a patient has multiple co-morbidities, that factors into the risk-adjustment, and will put a patient into a much higher risk-adjustment category.
- Results are published through NCQA's Quality Compass module which contains the individual plan results by detailed service category along with a quality score.

1. Importance to Measure and Report
1a.High Impact: H-9; M-0; L-0; N-0
TAP Discussion: Agreement that asthma is an important area of healthcare to measure due to its high cost and the potential for improvements in care.
1b. Resource use/cost problems: H-7; M-2; L-0 ; I-0; N-0
TAP Discussion: The TAP agreed that asthma represents a resource use problem and noted that there is a well-documented opportunity for improvement.
1c. Purpose clearly described: H-9 ; M-0 ; L- X; I- X; N-0
Discussion: Purpose and objective are clear.
1d. Resource use service categories consistent and representative: H-9 ; M-0; L-0 ; I-0; N-0
TAP Discussion: No issues were raised.

2. Scientific Acceptability of Measure Properties:
2a. Reliability:
2a1.Measure well defined and precisely specified: H-9 ; M-0 ; L-0 ; I-0; N-0
TAP Discussion: No issues raised.
2a2. The results are repeatable: H-8 ; M-1 ; L-0 ; I-0; N-0
TAP Discussion: There was general agreement from the committee that that the balance is in favor of following a methodology of including *all* costs, to avoid having to consider what costs should or should not be associated with asthma. The developer reaffirmed that the measures are valid for any health plan; they are population-based measures and have been tested and used in physician groups. However, a population of at least 400 members is needed for the methodology to be valid, so it consequently tends to be larger physician groups that can use the measures.
Overall Reliability: H-8; M-1; L-0 ; I-0; N-0
2b. Validity:
2b1. Evidence is consistent with intent: H-6; M-3; L-0 ; I-0; N-0
TAP Discussion: There was a good overall evidence of face validity, but also a general desire to see more specific discussion around

the face validity of the HCC's.
2b2.Score/Analysis: H-6 ; M-3 ; L-0 ; I-0; N-0
TAP Discussion: The face validity of HCC's was found to be clear, but the categorizations based on age were unclear. The developers use the HCC approach for categorizing the population and clarified some of the methodology behind that approach.
2b3. Exclusions: H-6 ; M-3 ; L-0 ; I-0; N-0
TAP Discussion: There was an in-depth discussion regarding measure exclusions. The measure developer explained that cardiovascular conditions are not specifically excluded, but are used in the risk adjustment model. Patients with COPD are excluded. Exclusions affect the denominator population over either year within the two-year criteria, which is similar to the HEDIS asthma measure. There was agreement that the exclusion of COPD (which resulted in 38% of the initial population being eliminated) seems appropriate, particularly in light of the age range increasing to 64. The TAP did express concern that excluding acute respiratory failure could exclude the poorly managed patients. However, NCQA noted that acute respiratory failure were only accounted for 3% of the population, so it and don't meet their 5% threshold of concern.
2b4. Risk Adjustment: H-7; M-2; L-0 ; I-0; N-0
TAP Discussion: The risk-adjustment strategy seems appropriate. Several strategies are tested by NCQA, and the same methodology is used for all of their measures, which is to stratify by age and gender and use HCC's to categorize the population.
2b5. Identification of statistically significant/meaningful differences: H-8; M-1; L-0 ; I-0; N-0
TAP Discussion: Agreement that the distribution of the scores' detail score was appropriate. There was concern regarding whether the measure score could differentiate between statistically significant versus clinically significant variations.
2b6. Multiple data sources:
TAP Discussion: N/A (using all administrative data)
Overall Validity: H-5 ; M-4 ; L-0 ; I-0; N-0
2c. Stratification for disparities: H-5 ; M-3; L-0 ; I-1; N-0
TAP Discussion: Stratification is needed although data isn't available.

Usability:
3a. Measure performance results are publicly reported: H-8 ; M-1 ; L-0 ; I-0; N-0
TAP Discussion: The TAP was satisfied that NCQA publically reports measure results and provides support to enable understanding of those results. Purchasers are using this information, along with NCQA quality measures, to improve value for their employees. Asthma is a bit more difficult because there is only one NCQA quality measure to associate with this cost measure, however NCQA believes there are more quality measures in the pipeline.
3b. Measure results are meaningful/useful for public reporting and quality improvement: H-6 ; M-3 ; L-0 ; I-0; N-0
TAP Discussion: The measure is straightforward and easy to interpret. NCQA uses standardized pricing tables, which are reviewed annually. Health plans are the main users for this data. However, purchasers and the large employers will also drive a need for this info. The TAP wondered how smaller businesses would implement this measure, and NCQA explained that they provide help through their annual conferences, webinar services and a dedicated webpage.
3c. Data and results can be decomposed for transparency and understanding: H-8; M-1 ;L-0 ; I-0; N-0
TAP Discussion: Methodology was generally very transparent and appropriate.
3d. Harmonized or justification for differences:
TAP Discussion: N/A

4. Feasibility:
4a. Data elements routinely generated during care process: H-9 ; M-0 ; L-0; I-0; N-0
TAP Discussion: Data is a byproduct of care.
4b. Data elements available electronically: H-9 ; M-0 ; L-0 ; I-0; N-0
TAP Discussion: Data available electronically.
4c. Susceptibility to inaccuracies/ unintended consequences identified: H-7 ; M-2; L-0; I-0; N-0
TAP Discussion: Agreement that NCQA did a good job recognizing where the challenges with data inaccuracies are and have addressed these challenges.
4d. Data collection strategy can be implemented: H-8 ; M-1; L-0 ; I-0; N-0
TAP Discussion: All the data has to go through a certified auditor before it's reported to NCQA. As part of their annual analysis, NCQA reviews outliers, but currently the outliers are less than half a percent for measures.

1561 Relative Resource Use for People with COPD (NCQA)

Description: This measure addresses the resource use of members identified with COPD.
Clinical diagnosis of COPD during the measurement year is used to identify members for inclusion in the eligible population and the results are adjusted to account for age, gender, and HCC-RRU risk classifications that predict cost variability (Refer to Attachment S8_Clinical Logic for additional information).

Resource Use Type: Per capita (population- or patient-based)
Data Type: Administrative claims, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Pharmacy, Paper Records
Resource Use Service Categories: Inpatient services: Inpatient facility services, Inpatient services: Evaluation and management, Inpatient services: Procedures and surgeries, Inpatient services: Imaging and diagnostic, Inpatient services: Lab services, Inpatient services: Admissions/discharges, Ambulatory services: Outpatient facility services, Ambulatory services: Emergency Department, Ambulatory services: Pharmacy, Ambulatory services: Evaluation and management, Ambulatory services: Procedures and surgeries, Ambulatory services: Imaging and diagnostic, Ambulatory services: Lab services
Care Setting: Ambulatory Care : Ambulatory Surgery Center (ASC), Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Hospital/Acute Care Facility, Imaging Facility, Laboratory, Pharmacy, Post Acute/Long Term Care Facility: Nursing Home/Skilled Nursing Facility
Level of Analysis: Clinician : Group/Practice, Health Plan, Integrated Delivery System, Population: Community, Population: National, Population : Regional
Measure Developer: National Committee for Quality Assurance (NCQA), 1100 13th street NW, STE 1000, Washington, District Of Columbia, 20005

Conditions/Questions for Developer:

- If the goal is to eventually marry these measures with quality measures and stratification is different, how will that be plausible?
- What is the upper age limit to be included in this measure?
- How do you ensure similar populations are compared?

Developer Response:

- The resource use strata are different than they are for clinical quality strata, which are not risk-adjusted. As the quality measures further increase and perhaps in the future become risk-adjusted, there will be more room for comparability.
- There is no upper age limit to this measure, but the dataset age limitation is 85 years.
- By risk adjusting to the specified level using the HCC's and the 13 different cohorts, NCQA end up comparing relatively similar plan populations. The quality index for this measure is use of diagnostic spirometry and exacerbations measures. There is no attribution of specific procedures to COPD yet.

1. Importance to Measure and Report

1a.High Impact: H-9; M-0; L-0; N-0
TAP Discussion: Agreement that this is an important area of measurement. COPD is the fourth leading cause of death in the US, and the only one in the top four increasing in percentage.
1b. Resource use/cost problems: H-9; M-0; L-0 ; I-0; N-0
TAP Discussion: While variation in resource use was identified in other parts of the submission, the information submitted in the form for this item only discussed the variations in clinical care provided.
1c. Purpose clearly described: H-8; M-1; L- 0; I- 0; N-0
TAP Discussion: The TAP was concerned that the measure submission applied only to newly diagnosed patients. The developer clarified that it is supposed to apply to anyone with a diagnosis with COPD. Otherwise, the purpose of the measure is to evaluate the total cost of care for COPD patients within a 1 year timeframe was clear.
1d. Resource use service categories consistent and representative: H-9 ; M-0; L-0 ; I-0; N-0
TAP Discussion: No issues raised.

2. Scientific Acceptability of Measure Properties:

2a. Reliability:
2a1.Measure well defined and precisely specified: H-9 ; M-0 ; L-0 ; I-0; N-0
TAP Discussion: The specifications provided are clear and precise. The developer provided clarification on age stratification for resource use categories indicating that they are based on utilization patterns in the data-set, not clinical factors.
2a2. The results are repeatable: H-8 ; M-1; L-0 ; I-0; N-0
TAP Discussion: A similar methodology was used for this measure as for NCQA measure #1560, the primary difference being in the selection of the population. The TAP was concerned about the multiple populations being studied including commercial, Medicare, and Medicaid, due to the age range (unlike Measure 1560, where the age range cut off at 64). There was also concern that NCQA did not distinguish the fee-for-service versus the general eligible population in Medicare. It remains difficult to separate these populations because right now they are aggregated.
Overall Reliability: H-7 ; M-2 ; L-0 ; I-0; N-0
2b. Validity:
2b1. Evidence is consistent with intent: H-8; M-1; L-0 ; I-0; N-0
TAP Discussion: The measure is clearly delineated; however, one of the challenges will be the fact that COPD has multiple co-morbidities, particularly when compared to asthma. It will therefore be difficult to know if you are measuring exactly COPD.

Specifications should be explored on how to develop disease severity; however, this is difficult to do with administrative datasets.
2b2.Score/Analysis: H-6; M-3; L-0 ; I-0; N-0
TAP Discussion: Overall the validity treating was appropriate. Outliers are identified by tagging O/E ratios below .3 or above 3 as outliers.
2b3. Exclusions: H-4 ; M-5; L-0 ; I-0; N-0
TAP Discussion: The exclusions are well stated and similar to the asthma measure.
2b4. Risk Adjustment: H-6 ; M-3; L-0 ; I-0; N-0
TAP Discussion: Cardiovascular disease may be biggest driver around severity of disease. The risk adjustment approach appears reasonable for the data available. The intent is to compare across populations.
2b5. Identification of statistically significant/meaningful differences: H-5 ; M-4 ; L-0 ; I-0; N-0
TAP Discussion: NCQA was complimented for doing a good job of presenting their data in a transparent manner.
2b6. Multiple data sources:
TAP Discussion: N/A (using all administrative data)
Overall Validity: H-4 ; M-5 ; L-0 ; I-0; N-0
2c. Stratification for disparities: H-5; M-4; L-0 ; I-0; N-0
TAP Discussion: Examining differences in racial disparities for this data set is not yet possible, but there is stratification by gender. Race is not a required field for most provider systems and is usually unavailable except in the Medicare population.

Usability:
3a. Measure performance results are publicly reported: H-9 ; M-0 ; L-0 ; I-0; N-0
TAP Discussion: NCQA does extensive audits of their material on a regular basis.
3b. Measure results are meaningful/useful for public reporting and quality improvement: H-5 ; M-4 ; L-0 ; I-0; N-0
TAP Discussion: Results are usable and understandable.
3c. Data and results can be decomposed for transparency and understanding: H-6; M-3 ;L-0 ; I-0; N-0
TAP Discussion: NCQA does extensive audits of their material on a regular basis, and the measure can be deconstructed to facilitate transparency.
3d. Harmonized or justification for differences:
TAP Discussion: N/A

4. Feasibility:
4a. Data elements routinely generated during care process: H-9 ; M-0 ; L-0; I-0; N-0
TAP Discussion: Data is a byproduct of care.
4b. Data elements available electronically: H-9 ; M-0 ; L-0 ; I-0; N-0
TAP Discussion: Data available electronically.
4c. Susceptibility to inaccuracies/ unintended consequences identified: H-6 ; M-3; L-0; I-0; N-0
TAP Discussion: Feasibility rationale is similar to NCQA Asthma measure already discussed.
4d. Data collection strategy can be implemented: H-8 ; M-1; L-0 ; I-0; N-0
TAP Discussion: No new issues raised.

1605 ETG Based Asthma Resource Use Measure(Ingenix)

Description: The measure focuses on resources used to deliver episodes of care for patients with Asthma.
Asthma episodes are defined using the Episode Treatment Groups (ETG) methodology and describe the unique presence of the condition for a patient and the services involved in diagnosing, managing and treating asthma. A number of resource use measures are defined for asthma episodes, including overall cost of care, cost of care by type of service, and the utilization of specific types of services. Each resource use measure is expressed as a cost or a utilization count per episode and comparisons with internal and external benchmarks are made using risk adjustment to support valid comparisons.
As requested by NQF, the focus of this submission is for Asthma episodes and will cover both measures at the Asthma base and severity level and also an Asthma composite measure where Asthma episode results are combined across Asthma severity levels. At the most detailed level, the measure is defined as the base condition of Asthma and an assigned level of severity (e.g., resources per episode for Asthma, severity level 1 episodes). Composite measures can then be created using these measurement units to meet a specific need. For example, a composite measure for Asthma is derived by combining Asthma episode results across Asthma severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of Asthma episodes by severity level when supporting an Asthma composite comparison). The focus of this measure is on Asthma. However, Asthma episode results could also be included in a "pulmonologist", "chronic care", or other clinical composite for a physician, combining episodes in clinical areas similar to Asthma. Further, an "overall" composite for a physician can be created, again by aggregating episode results across appropriate conditions and severity levels and applying proper risk adjustment when making comparisons.

Resource Use Type: Per episode
Data Type: Administrative claims, Other
Resource Use Service Categories:
Inpatient services: Inpatient facility services, Inpatient services: Admissions/discharges, Ambulatory services: Outpatient facility services, Ambulatory services: Emergency Department, Ambulatory services: Pharmacy, Ambulatory services: Evaluation and management, Ambulatory services: Procedures and surgeries, Ambulatory services: Imaging and diagnostic, Ambulatory services: Lab services
Care Setting: Ambulatory Care : Ambulatory Surgery Center (ASC), Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Emergency Medical Services/Ambulance, Home Health, Hospice, Hospital/Acute Care Facility, Imaging Facility, Laboratory, Post Acute/Long Term Care Facility: Nursing Home/Skilled Nursing Facility
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team, Facility, Health Plan, Integrated Delivery System, Population: Community, Population: County or City, Population : National, Population : Regional, Population: State
Measure Developer: Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02451

Conditions/Questions for Developer:

- Can you give us more information on how repeatability and "consistency" were determined? The results don't appear consistent.
- Are patients with COPD excluded?
- How are results reported and interpreted?
- How would a smaller health plan implement this measure? It seems it might be too complex and burdensome.

Developer Response:

- Repeatability was demonstrated by programming the measure in SAS code and the Ingenix software and comparing results. Because there are differences in what geographies these health plans are pulling from, variation is expected. But while differences across HCO's are expected, whether the differences are too high or low is difficult to know.
- Patients are excluded from the asthma episode if they have more costs attributable to COPD than asthma.
- The main measurement is the O/E ratio metric - the numerator of which is the cost of all the episodes of asthma, and the denominator which is the expected costs.
- The burden depends on the plan's familiarity with ETGs and similar products, and for those who are just starting out, there is unlimited training involved (i.e. help desk support, etc.). There is another option where Ingenix takes the data and runs it themselves - or uses their PCQ Connect product that prepared the data into report-ready formats.

1.Importance to Measure and Report
1a.High Impact: H-9; M-0; L-0; N-0
TAP Discussion: Agreement that asthma is a very important health care area to measure.
1b. Resource use/cost problems: H-8; M-1; L-0 ; I-0; N-0
TAP Discussion: Measure demonstrates cost problems and opportunity for improvement.
1c. Purpose clearly described: H-7; M-2; L- X; I- X; N-0
TAP Discussion: Purpose and objective are clear.
1d. Resource use service categories consistent and representative: H-7 ; M-2; L-0 ; I-0; N-0
TAP Discussion: No issues raised.

2. Scientific Acceptability of Measure Properties:
2a. Reliability:
2a1.Measure well defined and precisely specified: H-2 ; M-6 ; L-1 ; I-0; N-0
TAP Discussion: This measure is one that's part of a suite of episodes around diseases and conditions included in Ingenix's episode treatment grouper. This product identifies claims that should be part of an episode of asthma and divides them into year-long segments, looking at asthma as a chronic disease. The episodes are severity adjusted using clinical markers called condition status factors. Anchor episodes, or face-to-face encounters, are merged together into one episode (i.e. "asthma").
2a2. The results are repeatable: H-3 ; M-5 ; L-1 ; I-0; N-0
TAP Discussion: The TAP didn't understand why Ingenix used three different population samples, rather than taking a portion of the larger population and testing it multiple times. They would like better communication on the approach as well as more detailed depiction of the data. Repeatability was generally determined to be demonstrated adequately, but for the above reasons, some did question the reliability of the measure score.
Overall Reliability: H-0 ; M-8 ; L-1 ; I-0; N-0
2b. Validity:
2b1. Evidence is consistent with intent: H-2 ; M-5; L-1 ; I-1; N-0
TAP Discussion: It was unclear to the panel whether Ingenix is actually measuring asthma costs as intended. The determination of what is an asthma cost and what is not isn't transparent. They also agreed that any results are going to be questioned when over 50% of the costs (the pharmacy costs) are not represented. There were suggestions to stratify those with and without pharmacy costs.

2b2.Score/Analysis: H-1 ; M-4 ; L-2 ; I-2; N-0
TAP Discussion: Face validity was determined to be appropriate. The TAP continued to express concern about the exclusion of pharmacy costs, which were agreed to be a significant component of asthma care. Pharmacy data is not a requirement to get into the episode (for all ETGs).
2b3. Exclusions: H-1 ; M-7 ; L-1 ; I-0; N-0
TAP Discussion: The TAP was concerned about the lack of transparency regarding which costs were excluded, and why. Confusion existed around what the grouper identified as outliers or exclusions. Windsorizing very high cost people, the top 2%, effectively excludes those kinds of patients that would be important to know about. Addition information such as sensitivity analyses would have helped explain the impact of these high cost cases.
2b4. Risk Adjustment: H-1 ; M-4 ; L-2 ; I-2; N-0
TAP Discussion: The TAP expressed the same concerns regarding the risk-adjustment methodology as they had for previous Ingenix measures. The TAP was apprehensive that because the measure doesn't require use of standardized costs, the playing field is not level and it can't be implemented consistently across organizations if one is using standard and another actual pricing. To examine how refined the risk-adjustment is, R-squares for different severity levels and how they predict resource utilization should be provided.
2b5. Identification of statistically significant/meaningful differences: H-0 ; M-8 ; L-0 ; I-1; N-0
TAP Discussion: The panel felt confident in Ingenix's methodology after it was explained.
2b6. Multiple data sources:
TAP Discussion: N/A (using all administrative data)
Overall Validity: H-0 ; M-6 ; L-1 ; I-2; N-0
2c. Stratification for disparities: H-2 ; M-6; L-0 ; I-1; N-0
TAP Discussion: Gender and age can be stratified, but race data is not available.

Usability:
3a. Measure performance results are publicly reported: H-2 ; M-4; L-2; I-1; N-0
TAP Discussion: This product is generally used with a suite of ETG's, usually in combination with the pneumonia and COPD measures. There was uncertainty about the measure's usefulness on its own. Since Ingenix can't ascertain if this measure is being used individually the concern from the panel is how the individual measure could be used.
3b. Measure results are meaningful/useful for public reporting and quality improvement: H-0 ; M-6; L-2; I-1; N-0
TAP Discussion: The TAP was concerned about the possibility of misinterpretation of results because of the transparency and usability of the results of this measure.
3c. Data and results can be decomposed for transparency and understanding: H-3; M-5;L-1; I-0; N-0
TAP Discussion: The TAP reiterated their concern of the transparency of the score. Ingenix clarified that there are ways to drill into different aspects of care to see how they might be driving the score.
3d. Harmonized or justification for differences:
TAP Discussion: N/A

4. Feasibility:
4a. Data elements routinely generated during care process: H-7 ; M-2; L-0; I-0; N-0
TAP Discussion: Data is a byproduct of care.
4b. Data elements available electronically: H-7 ; M-2; L-0 ; I-0; N-0
TAP Discussion: Data available electronically.
4c. Susceptibility to inaccuracies/ unintended consequences identified: H-1 ; M-8; L-0; I-0; N-0
TAP Discussion: The TAP was generally comfortable with the error checks built into the product.
4d. Data collection strategy can be implemented: H-4 ; M-4; L-0 ; I-1; N-0
TAP Discussion: The TAP was concerned the burden this measure would place on a programmer to implement, particularly at smaller health plans.

1608 ETG Based Chronic Obstructive Pulmonary Disease Resource Use Measure (COPD) (Ingenix)

Description: The measure focuses on resources used to deliver episodes of care for patients with COPD.
COPD episodes are defined using the Episode Treatment Groups (ETG) methodology and describe the unique presence of the condition for a patient and the services involved in diagnosing, managing and treating COPD. A number of resource use measures are defined for COPD episodes, including overall cost of care, cost of care by type of service, and the utilization of specific types of services. Each resource use measure is expressed as a cost or a utilization count per episode and comparisons with internal and external benchmarks are made using risk adjustment to support valid comparisons.
As requested by NQF, the focus of this submission is for COPD episodes and will cover both measures at the COPD base and severity level and also a COPD composite measure where COPD episode results are combined across COPD severity levels. At the most detailed level, the measure is defined as the base condition of COPD and an assigned level of severity (e.g., resources per episode for

COPD, severity level 1 episodes). Composite measures can then be created using these measurement units to meet a specific need. For example, a composite measure for COPD is derived by combining COPD episode results across COPD severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of COPD episodes by severity level when supporting a COPD composite comparison). The focus of this measure is on COPD. However, COPD episode results could also be included in a "pulmonary" "chronic care", or other clinical composite for a physician, combining episodes in clinical areas similar to COPD. Further, an "overall" composite for a physician can be created, again by aggregating episode results across appropriate conditions and severity levels and applying proper risk adjustment when making comparisons.
Resource Use Type: Per episode
Data Type: Administrative claims, Other
Resource Use Service Categories:
Inpatient services: Inpatient facility services, Inpatient services: Admissions/discharges, Ambulatory services: Outpatient facility services, Ambulatory services: Emergency Department, Ambulatory services: Pharmacy, Ambulatory services: Evaluation and management, Ambulatory services: Procedures and surgeries, Ambulatory services: Imaging and diagnostic, Ambulatory services: Lab services
Care Setting: Ambulatory Care : Ambulatory Surgery Center (ASC), Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Emergency Medical Services/Ambulance, Home Health, Hospice, Hospital/Acute Care Facility, Imaging Facility, Laboratory, Post Acute/Long Term Care Facility: Nursing Home/Skilled Nursing Facility
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team, Facility, Health Plan, Integrated Delivery System, Population: Community, Population: County or City, Population : National, Population : Regional, Population: State
Measure Developer: Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02451

Conditions/Questions for Developer:

- What was the clinical logic of using 180 days, particularly since your Asthma measure had used 365 days, and both are similar chronic conditions?

Developer Response:

- We will have to examine that further.

1. Importance to Measure and Report
1a.High Impact: H-7; M-0; L-0; N-0
TAP Discussion: The TAP felt Ingenix did well with articulating the high impact of COPD.
1b. Resource use/cost problems: H-7; M-0; L-0 ; I-0; N-0
TAP Discussion: COPD represents a resource use issue that can be addressed.
1c. Purpose clearly described: H-7 ; M-0 ; L- X; I- X; N-0
TAP Discussion: Purpose and objective are clear.
1d. Resource use service categories consistent and representative: H-6 ; M-1; L-0 ; I-0; N-0
TAP Discussion: No issues raised.

2. Scientific Acceptability of Measure Properties:
2a. Reliability:
2a1.Measure well defined and precisely specified: H-4 ; M-3; L-0 ; I-0; N-0
TAP Discussion: Questions were raised around the clinical logic around the timeframes chosen.
2a2. The results are repeatable: H-5 ; M-2 ; L-0 ; I-0; N-0
TAP Discussion: Agreement from the TAP that reliability for this measure is similar to the previously discussed Ingenix asthma measure.
Overall Reliability: H-4; M-3 ; L-0 ; I-0; N-0
2b. Validity:
2b1. Evidence is consistent with intent: H-2 ; M-5; L-0 ; I-0; N-0
TAP Discussion: No issues raised.
2b2.Score/Analysis: H-0 ; M-7 ; L-0 ; I-0; N-0
TAP Discussion: The TAP remained concerned about Ingenix's testing method for customization, the inability to compare actual versus standardized prices, and the high level of pharmacy exclusions.
2b3. Exclusions: H-1 ; M-6; L-0 ; I-0; N-0
TAP Discussion: There are no clinical exclusions, only administrative ones. The TAP felt it was unclear how tie-breaking logic works and noted that it was not specified in the submission how COPD and asthma ETG's interact.
2b4. Risk Adjustment: H-0 ; M-4 ; L-3 ; I-0; N-0
TAP Discussion: While Ingenix had a nice description of how they developed their risk-adjustment approach, the panel would have liked to see more modeling presented in the submission.
2b5. Identification of statistically significant/meaningful differences: H-0 ; M-7; L-0 ; I-0; N-0
TAP Discussion: The TAP questioned whether the practical significance of the measure is justified by the relative cost ratio.
2b6. Multiple data sources:

TAP Discussion: N/A (using all administrative data)
Overall Validity: H-0 ; M-7 ; L-0 ; I-0; N-0
2c. Stratification for disparities: H-0 ; M-5; L-0 ; I-0; N-0
TAP Discussion: Only gender and age are stratified for. Race data is not available.

Usability:
3a. Measure performance results are publicly reported: H-0 ; M-7 ; L-0 ; I-0; N-0
TAP Discussion: The TAP expressed doubts regarding whether the measure could be implemented in a user-friendly manner.
3b. Measure results are meaningful/useful for public reporting and quality improvement: H-0 ; M-7; L-0 ; I-0; N-0
TAP Discussion: The panel agreed that measure provides useful information for individual health plans. However, they expressed concern about how useful it would be to compare across health plans, due to the fact that standardized pricing is not required.
3c. Data and results can be decomposed for transparency and understanding: H-3; M-4;L-0 ; I-0; N-0
TAP Discussion: It was agreed that previous discussions regarding Ingenix transparency would also apply to this measure.
3d. Harmonized or justification for differences:
TAP Discussion: N/A

4. Feasibility:
4a. Data elements routinely generated during care process: H-5 ; M-2 ; L-0; I-0; N-0
TAP Discussion: Data is a byproduct of care.
4b. Data elements available electronically: H-7 ; M-0 ; L-0 ; I-0; N-0
TAP Discussion: Data available electronically.
4c. Susceptibility to inaccuracies/ unintended consequences identified: H-3 ; M-4; L-0; I-0; N-0
TAP Discussion: The TAP were comfortable that Ingenix can accurately identify inaccuracies and errors.
4d. Data collection strategy can be implemented: H-6 ; M-2 ; L-0 ; I-0; N-0
TAP Discussion: No new issues raised.

1611 ETG Based Pneumonia Resource Use Measure (Ingenix)

Description: The measure focuses on resources used to deliver episodes of care for patients with pneumonia.
Pneumonia episodes are defined using the Episode Treatment Groups (ETG) methodology and describe the unique presence of the condition for a patient and the services involved in diagnosing, managing and treating pneumonia. A number of resource use measures are defined for pneumonia episodes, including overall cost of care, cost of care by type of service, and the utilization of specific types of services. Each resource use measure is expressed as a cost or a utilization count per episode and comparisons with internal and external benchmarks are made using risk adjustment to support valid comparisons. As requested by NQF, the focus of this submission is for pneumonia episodes and will cover both measures at the pneumonia base and severity level and also a pneumonia composite measure where pneumonia episode results are combined across pneumonia severity levels. At the most detailed level, the measure is defined as the base condition of pneumonia and an assigned level of severity (e.g., resources per episode for pneumonia, severity level 1 episodes). Composite measures can then be created using these measurement units to meet a specific need. For example, a composite measure for pneumonia is derived by combining pneumonia episode results across pneumonia severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of pneumonia episodes by severity level when supporting a pneumonia composite comparison). The focus of this measure is on pneumonia. However, pneumonia episode results could also be included in a "pulmonary" or other clinical composite for a physician, combining episodes in clinical areas similar to pneumonia. Further, an "overall" composite for
a physician can be created, again by aggregating episode results across appropriate conditions and severity levels and applying proper risk adjustment when making comparisons.
Resource Use Type: Per episode
Data Type: Administrative claims, Other
Resource Use Service Categories:
Inpatient services: Inpatient facility services, Inpatient services: Admissions/discharges, Ambulatory services: Outpatient facility services, Ambulatory services: Emergency Department, Ambulatory services: Pharmacy, Ambulatory services: Evaluation and management, Ambulatory services: Procedures and surgeries, Ambulatory services: Imaging and diagnostic, Ambulatory services: Lab services
Care Setting: Ambulatory Care : Ambulatory Surgery Center (ASC), Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Emergency Medical Services/Ambulance, Home Health, Hospice, Hospital/Acute Care Facility, Imaging Facility, Laboratory, Post Acute/Long Term Care Facility: Nursing Home/Skilled Nursing Facility, Post Acute/Long Term Care Facility: Rehabilitation
Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team, Facility, Health Plan, Integrated Delivery System, Population: Community, Population: County or City, Population : National, Population : Regional, Population: State
Measure Developer: Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02451

Conditions/Questions for Developer:

- Would it be possible to break down the measure by bacterial versus non-bacterial to try to separate out pneumonia types?

Developer Response:

- Yes, the measure rule is stratified so that they are looked at in different episodes.

1. Importance to Measure and Report

1a.High Impact: H-8; M-0; L-0; N-0

TAP Discussion: The TAP agreed that pneumonia is a high impact and high cost area.

1b. Resource use/cost problems: H-8; M-0; L-0 ; I-0; N-0

TAP Discussion: Ingenix have done a good job of presenting the data.

1c. Purpose clearly described: H-8 ; M-0 ; L- X; I- X; N-0

TAP Discussion: Purpose and objective are clear.

1d. Resource use service categories consistent and representative: H-7 ; M-1; L-0 ; I-0; N-0

TAP Discussion: No issues raised.

2. Scientific Acceptability of Measure Properties:

2a. Reliability:

2a1.Measure well defined and precisely specified: H-3 ; M-4; L-0 ; I-0; N-0

TAP Discussion: Members of the TAP felt uncomfortable with the lack of transparency in the risk adjustment specifications and felt that the severity weights, particularly for the elderly, were unclear. The panel also had a hard time identifying clean periods. There was a strong feeling that there should be some separation between community-acquired and healthcare-acquired pneumonia, as they represent very different clinical conditions.

2a2. The results are repeatable: H-6 ; M-1; L-0 ; I-0; N-0

TAP Discussion: Concerns were expressed regarding the fact that that there is no way to ascertain how Ingenix came up with the specific weights assigned to comorbidities.

Overall Reliability: H-3 ; M-3 ; L-0 ; I-1; N-0

2b. Validity:

2b1. Evidence is consistent with intent: H-4 ; M-3; L-0 ; I-0; N-0

TAP Discussion: The panel again asked for clarification regarding why the measure has different weighted scores for the elderly.

2b2.Score/Analysis: H-0 ; M-5 ; L-2; I-0; N-0

TAP Discussion: The TAP was concerned that they weren't provided with enough information to understand how Ingenix assigned risk scores. Questions regarding how diagnostic descriptions lead to increased utilization were raised. The TAP remained doubtful as to whether this measure should be counted as one distinct population.

2b3. Exclusions: H-2 ; M-4 ; L-1; I-0; N-0

TAP Discussion: The panel felt that more data around the impact of exclusions (e.g. sensitivity analysis) would be helpful. Ingenix confirmed that there are no clinical exclusions from the measure, only cost exclusions.

2b4. Risk Adjustment: H-1 ; M-3 ; L-2 ; I-1; N-0

TAP Discussion: As with other Ingenix measures, the TAP believed that the risk-adjustment methodology is not readily transparent. More information on how risk scores are assigned was requested.

2b5. Identification of statistically significant/meaningful differences: H-0 ; M-7 ; L-0 ; I-0; N-0

TAP Discussion: Data submitted does demonstrate variation in resource use. However, there was a general feeling that meaningfulness is questionable since types of pneumonia cannot be separated out.

2b6. Multiple data sources:

TAP Discussion: N/A (using all administrative data)

Overall Validity: H-0 ; M-7 ; L-0 ; I-0; N-0

2c. Stratification for disparities: H-2 ; M-5; L-0 ; I-0; N-0

TAP Discussion: Gender and age can be stratified, but race data is not available.

Usability:

3a. Measure performance results are publicly reported: H-0 ; M-6; L-1; I-0; N-0

TAP Discussion: The TAP agreed that despite the fact that multiple care organizations are currently using this measure, the inability to distinguishing between types of pneumonia severely limits the usability of the measure. They agreed that for individual organizations this limitation might be acceptable, but the measure wouldn't be useful in a comparative setting. NQF clarified that the measure has a specified level of analysis, and the ratings need to be reflective of that specification.

3b. Measure results are meaningful/useful for public reporting and quality improvement: H-1 ; M-5; L-1; I-0; N-0

TAP Discussion: No new issues raised.

3c. Data and results can be decomposed for transparency and understanding: H-1; M-5;L-1; I-0; N-0

TAP Discussion: The measure would be more transparent if more user-friendly detail were provided.

3d. Harmonized or justification for differences:

TAP Discussion: N/A

4. Feasibility:
4a. Data elements routinely generated during care process: H-7 ; M-0 ; L-0; I-0; N-0
TAP Discussion: Data is a byproduct of care.
4b. Data elements available electronically: H-7; M-0 ; L-0 ; I-0; N-0
TAP Discussion: Data available electronically.
4c. Susceptibility to inaccuracies/ unintended consequences identified: H-1 ; M-5; L-0; I-1; N-0
TAP Discussion: The TAP concluded there was a lack of information in the submission regarding data cleaning and missing data to sufficiently understand those areas.
4d. Data collection strategy can be implemented: H-5 ; M-2 ; L-0 ; I-0; N-0
TAP Discussion: No new issues raised.

PUBLIC COMMENT

There were no public comments.

NEXT STEPS

Ms. Dorian thanked the TAP for their dedication in reviewing Pulmonary measures. Doctors Elward and Maurer were also thanked for their leadership of the panel. The Steering Committee will meet in person to review these and two other measures on August 30-31st.