

NATIONAL QUALITY FORUM

CONFERENCE CALL OF THE RESOURCE USE STEERING COMMITTEE

July 15, 2011

Committee Members Participating: Tom Rosenthal, MD (*Co-Chair*, UCLA School of Medicine); Jephtha Curtis, MD (Yale University School of Medicine); Kurtis Elward, MD, MPH (Family Medicine of Albermarle); Lisa Grabert, MPH (American Hospital Association); Jack Needleman, PhD (University of California, Los Angeles School of Public Health); David Penson, MD, MPH (Vanderbilt University Medical Center); Steve Phillips, MPA (Ortho-McNeill-Janssen Pharmaceutical, Inc.); David Redfearn, PhD (WellPoint); Bill Rich, MD (Northern Virginia Ophthalmology Associates); Joe Stephansky, PhD (Michigan Health and Hospital Association).

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

Others Present: Cheri Zielinski, Ingenix; Thomas Lynn, Ingenix; Dan Dunn, Ingenix

MEETING PROCESS

Ms. Wilbon welcomed the Steering Committee and thanked them for their participation. The purpose of this conference call was to discuss the Feasibility criterion for NQF#1599 ETG Based Non-Condition Specific resource use measure and NQF# 1595 ETG Based Diabetes, as well as overall measure reviews for NQF# 1591 ETG Based Congestive Heart Failure (CHF) resource use measure and NQF# 1594 ETG Based Coronary Artery Disease (CAD) resource use measure; all of the measures were submitted by Ingenix.

The measure developers were available on the call to respond to questions from the Committee as needed. A NQF Member and public comment period occurred at the end of the call; no comments were made at that time. General project information can be found by clicking on the [Resource Use project page](#).

ABMS-REF MEASURE UPDATE

At this time the American Board of Medical Specialties Research & Education Foundation has decided to withdraw all of their measures from this project. The measure developers feel that the measures need more testing before undergoing the rigorous evaluation process. Since a number of the measures have undergone the evaluation process thus far, NQF will provide feedback to the development team so they can continue to refine and test the measures. For the remainder of the measures, mainly the pulmonary measures, the TAP and Steering Committee will still be asked to provide feedback.

NATIONAL QUALITY FORUM

MEASURE EVALUATION SUMMARY

The following summaries include a preliminary review of four measures submitted by Ingenix. The measure developers were present to give a brief overview of each measure.

<p>1599 ETG Based Non-Condition Specific resource use measure</p> <p>Description: The measure focuses on resources used to diagnose, manage and treat a population of patients (non-condition specific) during a defined 12-month period of time. The population included in the measurement can be described generally. Examples include a population of individuals enrolled with a health plan, individuals assigned to a patient-centered medical home or accountable care organization (ACO), or a panel of individuals managed by a primary care physician (PCP). A number of resource use measures are defined for this measure set, 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 member per month and comparisons with internal and external benchmarks are made using risk adjustment to support valid comparisons. Risk adjustment is based on the measure of risk assigned to each individual using the Episode Risk Group (ERG) methodology.</p> <p>Resource Use Type: Per capita (population- or patient-based)</p> <p>Data Type: Administrative claims, Other</p> <p>Resource Use Service Category: 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</p> <p>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</p> <p>Level of Analysis: Clinician : Group/Practice, Clinician : Individual, Clinician : Team, Facility, Health Plan, Integrated Delivery System, Population : County or City, Population : National, Population : Regional</p> <p>Measure Developer: Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02451</p>
<p>Steering Committee Recommendation for Endorsement:</p> <p>Rationale: <i>Pending Committee's official vote.</i></p>
<p>Conditions/Questions for Developer:</p> <ul style="list-style-type: none"> • What do the three tabs on the Ingenix pricing table mean? They are labeled ETG, ETG/ERG and ETG/PEG. • In order to use the software, does a user need technical support from Ingenix? <p>Developer Response:</p> <ul style="list-style-type: none"> • Some of the measures that were submitted for NQF endorsement are related to ETGs only, for example diabetes is an ETG specific relationship. The non-condition specific, population based measure is composed of both ETG and ERG, which is the episode treatment grouper and the episode risk-assessment grouper, respectively. The ETG and PEG methods pertain to the hip and knee replacement measures. The products contain all three as a suite, even if you do not want/need all of the measures; they all come in one package. • A user does not necessarily need the technical support of Ingenix, but it is included in the cost. <p>Steering Committee Follow-up: N/A</p> <p>If applicable, Questions to the Steering Committee: N/A</p>
<p>1. Importance to Measure and Report <i>**This criterion was discussed during the June 28-29 in person Steering Committee meeting. To access the summary and transcripts, please click here.</i></p>
<p>2. Scientific Acceptability of Measure Properties: <i>**This criterion was discussed during the June 28-29 in person Steering Committee meeting. To access the summary and transcripts, please click here.</i></p>
<p>3. Usability: <i>**This criterion was discussed during the June 28-29 in person Steering Committee meeting. To access the summary and transcripts, please click here.</i></p>
<p>4. Feasibility:</p> <p>4a. Data elements routinely generated during care process: Discussion: The Steering Committee believes that this sub criterion has been met; all of the data elements are generated during the care process.</p> <p>4b. Data elements available electronically: Discussion: The Steering Committee believes that this sub criterion has been met; all of the data is available electronically.</p> <p>4c. Susceptibility to inaccuracies/ unintended consequences identified:</p>

NATIONAL QUALITY FORUM

Discussion: Mental health is not available and pharmacy data rarely is, when pharmacy data is included it is stratified. Ingenix does not have a formal audit system to ensure that all of the numbers are included & correct. In general, when dealing with any measure that uses administrative data there are various inaccuracies, pertaining particularly to coding inaccuracies and variation. Ingenix provides guidelines how to use small volumes/ sample sizes, however there is not content available to demonstrate this approach. This measure appears less prone to "gaming", as there is not much a user can do to manipulate the start or end of an episode.

4d. Data collection strategy can be implemented:

Discussion: Annually, for physician's the cost of the small package (less than 800 physicians) is \$70,000, the medium package is \$90,000 and the large is \$110,000 (over 2000 physicians in the group). For health plans the package comes in small, medium and large; the small package is \$90,115 (less than 400,000 covered lives) and the large package is \$135,000 (over a million covered lives), all prices are annual rates. The Steering Committee came to the conclusion that this is no more expensive than another product of its caliber (for example, ACGs used by HealthPartners). The cost is not only for the licensure of the proprietary software, the cost includes all of the measures, over 558 ETGs. The Steering Committee acknowledges that the system may be used without Ingenix's technical support, it is complicated, but if the documentation is thoroughly read, it is doable.

1595 ETG Based Diabetes resource use measure

Description: The measure focuses on resources used to deliver episodes of care for patients with Diabetes. Diabetes 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 diabetes. A number of resource use measures are defined for diabetes 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 Diabetes episodes and will cover both measures at the Diabetes base and severity level and also a Diabetes composite measure where Diabetes episode results are combined across Diabetes severity levels. At the most detailed level, the measure is defined as the base condition of Diabetes and an assigned level of severity (e.g., resources per episode for Diabetes, 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 Diabetes is derived by combining Diabetes episode results across Diabetes severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of Diabetes episodes by severity level when supporting a Diabetes composite comparison). The focus of this measure is on Diabetes. However, Diabetes episode results could also be included in an "endocrinology", "chronic care", or other clinical composite for a physician, combining episodes in clinical areas similar to Diabetes. 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 Category: 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

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 : states

Measure Developer: Ingenix, 950 Winter Street, Waltham, Massachusetts, 02154

Steering Committee Recommendation for Endorsement:

Rationale: *Pending Committee's official vote.*

Conditions/Questions for Developer: N/A

Developer Response: N/A

Steering Committee Follow-up: N/A

If applicable, Questions to the Steering Committee: N/A

3. Importance to Measure and Report

****This criterion was discussed during the June 28-29 in person Steering Committee meeting. To access the summary and transcripts, please click [here](#).**

4. Scientific Acceptability of Measure Properties:

****This criterion was discussed during the June 28-29 in person Steering Committee meeting. To access the summary and transcripts, please click [here](#).**

3. Usability:

NATIONAL QUALITY FORUM

****This criterion was discussed during the June 28-29 in person Steering Committee meeting. To access the summary and transcripts, please click [here](#).**

4. Feasibility:

4a. Data elements routinely generated during care process:

Discussion: The Steering Committee believes that this sub criterion has been met; all of the data elements are generated during the care process.

4b. Data elements available electronically:

Discussion: The Steering Committee believes that this sub criterion has been met; all of the data is available electronically.

4c. Susceptibility to inaccuracies/ unintended consequences identified:

Discussion: Mental health is not available and pharmacy data rarely is, when pharmacy data is included it is stratified. Ingenix does not have a formal audit system to ensure that all of the numbers are included & correct. In general, when dealing with any measure that uses administrative data there are various inaccuracies, pertaining particularly to coding inaccuracies and variation. It's not clear if the published data defines or uses small sample sizes. Ingenix provides guidelines how to use small volumes/ sample sizes, however there is not content available to demonstrate this approach. This measure appears less prone to "gaming", as there is not much a user can do to manipulate the start or end of an episode.

4d. Data collection strategy can be implemented:

Discussion: Annually, for physician's the cost of the small package (less than 800 physicians) is \$70,000, the medium package is \$90,000 and the large is \$110,000 (over 2000 physicians in the group). For health plans the package comes in small, medium and large; the small package is \$90,115 (less than 400,000 covered lives) and the large package is \$135,000 (over a million covered lives), all prices are annual rates. The Steering Committee came to the conclusion that this is no more expensive than another product of its caliber (for example, ACGs used by HealthPartners). The cost is not only for the licensure of the proprietary software, the cost includes all of the measure, over 558 ETGs. The Steering Committee acknowledges that the system may be used without Ingenix's technical support, it is complicated, but if the documentation is thoroughly read, it is doable.

1591 ETG Based Congestive Heart Failure (CHF) resource use measure

Description: The measure focuses on resources used to deliver episodes of care for patients with Congestive Heart Failure (CHF). CHF 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 CHF. A number of resource use measures are defined for CHF 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 CHF episodes and will cover both measures at the CHF base and severity level and also a CHF composite measure where CHF episode results are combined across CHF severity levels. At the most detailed level, the measure is defined as the base condition of CHF and an assigned level of severity (e.g., resources per episode for CHF, 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 CHF is derived by combining CHF episode results across CHF severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of CHF episodes by severity level when supporting a CHF composite comparison). The focus of this measure is on CHF. However, CHF episode results could also be included in a "cardiology", "chronic care", or other clinical composite for a physician, combining episodes in clinical areas similar to CHF. 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 Category: 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

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 : states

Measure Developer: Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02451

Steering Committee Recommendation for Endorsement:

Rationale: Pending Committee's official vote.

Conditions/Questions for Developer: N/A

Developer Response: N/A

NATIONAL QUALITY FORUM

<p>Steering Committee Follow-up: N/A If applicable, Questions to the Steering Committee: N/A</p>
<p>5. Importance to Measure and Report 1a. High Impact: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this is a high impact, high cost area; this sub criterion has been met. 1b. Resource use/cost problems: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this subcriterion has been met. 1c. Purpose clearly described: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this subcriterion has been met. 1d. Resource use service categories consistent and representative: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this subcriterion has been met.</p>
<p>6. Scientific Acceptability of Measure Properties: 2a. Reliability: <i>Pending Committee's Final Evaluation</i> 2a1. Well defined/precise specifications: Discussion: The Steering Committee discussion focused on how clearly specified these codes are, and how well they capture systolic heart failure. This is a measure of systolic heart failure, there is a paired measure of diastolic heart failure that they did not submit; because the Steering Committee couldn't take into account the existence of the diastolic measure there was concern as to the completeness and accuracy with which this measure would capture systolic heart failure. The diagnosis codes specified are limited to the 428s that used the word "systolic", they do not use some of the 404s and 402s that the other measures have used to capture the larger heart failure population. 2a2. Reliability testing: Discussion: The measure specifications have been in use for a long time; Ingenix has demonstrated that if you do the same measure in the same population, at the same time, then the result will be the same. 2b. Validity: <i>Pending Committee's Final Evaluation</i> 2b1. Specifications consistent with resource use/cost problem: Discussion: The Steering Committee believes this subcriterion has been met, the specifications are consistent with the resource use and cost problem. 2b2. Validity testing: Discussion: Although there are carve outs for mental health & pharmacy data, comparisons within the health plan are the same or likely to be the same. It's comparing across health plans or across physician groups where this issue may become relevant. 2b3. Exclusions: Discussion: There are no exclusions within this measures, the Steering Committee believes this subcriterion has been met. 2b4. Risk adjustment: Discussion: There were concerns that the measure was adjusting for comorbidities identified during the measurement episode as opposed to comorbidities identified prior to. There was also concern that the risk adjustment is "over -adjusting", or possibly "adjusting away" significant differences. 2b5. Identification of statistically significant/meaningful differences: Discussion: The Steering Committee believes that this measure did not identify statistically significant or meaningful differences across groups. There was a general concern that something may be classified as statistically significant, when it really is clinically significant. 2b6. Multiple data sources: Discussion: N/A 2c. Stratification for disparities: Discussion: N/A</p>
<p>3. Usability: 3a. Measure performance results are publicly reported: Discussion: The Steering Committee was concerned with the availability of this data to the public and requested clarification from NQF on what is required for "public reporting". The NQF Consensus Standards Approval Committee (CSAC) and the Board of Directors (BOD) continue to discuss this issue. 3b. Measure results are meaningful/useful for public reporting and quality improvement: Discussion: The Steering Committee agreed that more information would be needed to explain the results of this measure to the public and to be used for internal quality improvement. 3c. Data and results can be decomposed for transparency and understanding: Discussion: The Steering Committee agreed there are challenges for the use of this measure, which include its complexity and lack of clarity in the specifications. The Steering Committee agreed it is difficult to assess the extent of which the measure can be decomposed as it is currently specified. 3d. Harmonized or justification for differences: Discussion: N/A</p>

NATIONAL QUALITY FORUM

<p>4. Feasibility:</p> <p>4a. Data elements routinely generated during care process: <i>Discussion:</i> The Steering Committee believes that this sub criterion has been met; all of the data elements are generated during the care process.</p> <p>4b. Data elements available electronically: <i>Discussion:</i> The Steering Committee believes that this sub criterion has been met; all of the data is available electronically.</p> <p>4c. Susceptibility to inaccuracies/ unintended consequences identified: <i>Discussion:</i> Mental health is not available and pharmacy data rarely is, when pharmacy data is included it is stratified. Ingenix does not have a formal audit system to ensure that all of the numbers are included & correct. In general, when dealing with any measure that uses administrative data there are various inaccuracies, pertaining particularly to coding inaccuracies and variation. It's not clear if the published data defines or uses small sample sizes. Ingenix provides guidelines how to use small volumes/ sample sizes, however there is not content available to demonstrate this approach. This measure appears less prone to "gaming", as there is not much a user can do to manipulate the start or end of an episode.</p> <p>4d. Data collection strategy can be implemented: <i>Discussion:</i> Annually, for physician's the cost of the small package (less than 800 physicians) is \$70,000, the medium package is \$90,000 and the large is \$110,000 (over 2000 physicians in the group). For health plans the package comes in small, medium and large; the small package is \$90,115 (less than 400,000 covered lives) and the large package is \$135,000 (over a million covered lives), all prices are annual rates. The Steering Committee came to the conclusion that this is no more expensive than another product of its caliber (for example, ACGs used by HealthPartners). The cost is not only for the licensure of the proprietary software, the cost includes all of the measure, over 558 ETGs. The Steering Committee acknowledges that the system may be used without Ingenix's technical support, it is complicated, but if the documentation is thoroughly read, it is doable.</p>
--

<p>1594 ETG Based Coronary Artery Disease (CAD) resource use measure</p> <p>Description: The measure focuses on resources used to deliver episodes of care for patients with CAD. CAD 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 CAD. A number of resource use measures are defined for CAD 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 CAD episodes and will cover both measures at the CAD base and severity level and also a CAD composite measure where CAD episode results are combined across CAD severity levels. At the most detailed level, the measure is defined as the base condition of CAD and an assigned level of severity (e.g., resources per episode for CAD, 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 CAD is derived by combining CAD episode results across CAD severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of CAD episodes by severity level when supporting a CAD composite comparison). The focus of this measure is on CAD. However, CAD episode results could also be included in a "cardiology", "chronic care", or other clinical composite for a physician, combining episodes in clinical areas similar to CAD. 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.</p> <p>Resource Use Type: Per episode</p> <p>Data Type: Administrative claims, other</p> <p>Resource Use Service Category: 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</p> <p>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</p> <p>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 : states</p> <p>Measure Developer: Ingenix, 950 Winter Street, Waltham, Massachusetts, 02154</p>
<p>Steering Committee Recommendation for Endorsement: <i>Rationale: Pending Committee's official vote.</i></p>
<p>Conditions/Questions for Developer: N/A Developer Response: N/A Steering Committee Follow-up: N/A</p>

NATIONAL QUALITY FORUM

<p>If applicable, Questions to the Steering Committee: N/A</p>
<p>1. Importance to Measure and Report 1a. High Impact: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this is a high impact, high cost area; this sub criterion has been met. 1b. Resource use/cost problems: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this subcriterion has been met. 1c. Purpose clearly described: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this subcriterion has been met. 1d. Resource use service categories consistent and representative: <i>Pending Committee's Final Evaluation</i> Discussion: The Steering Committee believes this subcriterion has been met.</p>
<p>2. Scientific Acceptability of Measure Properties: 2a. Reliability: <i>Pending Committee's Final Evaluation</i> 2a1. Well defined/precise specifications: Discussion: The measure identified the primary incurring diagnosis codes as 410s through 414s. Within those strata there is a range of conditions – ranging from chronic, stable coronary artery disease to patients with cardiogenic shock complicated by a flail mitral posterior leaflet. Therefore, there is a huge spectrum of risk adverse outcomes within this population. Furthermore, this carries the risk of different resource use for each specific condition included in the measure. The measure was submitted for implementation across various levels of analysis, however for individual clinicians there is not a sample size guideline. 2a2. Reliability testing: Discussion: The measure is specified in a way that it has been used over a long period of time, Ingenix demonstrated that if the user does the same measure in the same population, as the same time, then the result will be the same. 2b. Validity: <i>Pending Committee's Final Evaluation</i> 2b1. Specifications consistent with resource use/cost problem: Discussion: The Steering Committee believes this subcriterion has been met. 2b2. Validity testing: Discussion: Although there are carve outs for mental health & pharmacy data, comparisons within the health plan are the same or likely to be the same. It's comparisons across different health plans or across physician groups where the issues become relevant. 2b3. Exclusions: Discussion: There are no exclusions within this measures, the Steering Committee believes this subcriterion has been met. 2b4. Risk adjustment: Discussion: The CV/DM TAP requested on July 14, that the developer demonstrate proof of the concept that this is accurately accounting for differences in the population – the risk adjustment method does not appear to be robust. 2b5. Identification of statistically significant/meaningful differences: Discussion: The Steering Committee believes that this measure did not identify statistically significant or meaningful differences across groups. There was general concern that something may be classified as statistically significant, when it really is not clinically significant. 2b6. Multiple data sources: Discussion: N/A 2c. Stratification for disparities: Discussion: N/A</p>
<p>3. Usability: 3a. Measure performance results are publicly reported: Discussion: The Steering Committee was concerned with the availability of this data to the public and requested clarification from NQF on what is required for "public reporting". The NQF Consensus Standards Approval Committee (CSAC) and the Board of Directors (BOD) continue to discuss this issue. 3b. Measure results are meaningful/useful for public reporting and quality improvement: Discussion: The Steering Committee agreed that more information would be needed to explain the results of this measure to the public and to be used for internal quality improvement. 3c. Data and results can be decomposed for transparency and understanding: Discussion: The Steering Committee agreed there are challenges for the use of this measure, which include its complexity and lack of clarity in the specifications. The Steering Committee agreed it is difficult to assess the extent of which the measure can be decomposed as it is currently specified. 3d. Harmonized or justification for differences: Discussion: N/A</p>
<p>4. Feasibility: 4a. Data elements routinely generated during care process: Discussion: The Steering Committee believes that this sub criterion has been met; all of the data elements are generated during the</p>

NATIONAL QUALITY FORUM

care process.

4b. Data elements available electronically:

Discussion: The Steering Committee believes that this sub criterion has been met; all of the data is available electronically.

4c. Susceptibility to inaccuracies/ unintended consequences identified:

Discussion: Mental health is not available and pharmacy data rarely is, when pharmacy data is included it is stratified. Ingenix does not have a formal audit system to ensure that all of the numbers are included & correct. In general, when dealing with any measure that uses administrative data there are various inaccuracies, pertaining particularly to coding inaccuracies and variation. It's not clear if the published data defines or uses small sample sizes. Ingenix provides guidelines how to use small volumes/ sample sizes, however there is not content available to demonstrate this approach. This measure appears less prone to "gaming", as there is not much a user can do to manipulate the start or end of an episode.

4d. Data collection strategy can be implemented:

Discussion: Annually, for physician's the cost of the small package (less than 800 physicians) is \$70,000, the medium package is \$90,000 and the large is \$110,000 (over 2000 physicians in the group). For health plans the package comes in small, medium and large; the small package is \$90,115 (less than 400,000 covered lives) and the large package is \$135,000 (over a million covered lives), all prices are annual rates. The Steering Committee came to the conclusion that this is no more expensive than another product of its caliber (for example, ACGs used by HealthPartners). The cost is not only for the licensure of the proprietary software, the cost includes all of the measure, over 558 ETGs. The Steering Committee acknowledges that the system may be used without Ingenix's technical support, it is complicated, but if the documentation is thoroughly read, it is doable.

PUBLIC COMMENT

There were no public comments.

NEXT STEPS

Ms. Wilbon indicated that project staff will continue with preparations for the next Steering Committee conference call based on the Committee's availability. The July 28 and September 20 conference calls will be canceled; the September 14 call will be used to discuss the remaining ABMS-REF measures.