

# NATIONAL QUALITY FORUM

## CONFERENCE CALL OF THE CARDIOVASCULAR/DIABETES TECHNICAL ADVISORY PANEL

June 9, 2011

*Committee Members Participating:* Jephtha Curtis, MD, Yale University School of Medicine (Co-Chair); Mary Ann Clark, MHA, Neocure Group; Michael O'Toole, MD, Midwest Heart Specialists; David Palestrant, MD, Cedars-Sinani Medical Center; Katherine Reeder, PhD, RN, University of Kansas School of Nursing.

*NQF Staff Participating:* Helen Burstin, MD, MPH, Senior Vice President; Sally Turbyville, MA, MS, Senior Director; Taroon Amin, MPH, Senior Director; Ashlie Wilbon, MPH, BSN, Project Manager; Sarah Fanta, Research Analyst.

*Others Present:* Cheri Zielinski, Ingenix; Thomas Lynn, Ingenix; Mohua Choudhury, National Committee for Quality Assurance (NCQA); Kevin Stroube, American Board for Medical Specialties Research and Education Foundation (ABMS-REF); Robin Wagner, ABMS-REF; Katie Harrell; Chad Heim, Health Partners; Sue Knudson, Health Partners.

### MEETING PROCESS

Ms. Wilbon welcomed the Steering Committee and thanked them for their participation. The purpose of this conference call was to discuss three cardiovascular measures submitted by NCQA, ABMS and Ingenix.

The measure developers and stewards 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](#).

### MEASURE EVALUATION SUMMARY

The following summary includes a preliminary review of the non-condition specific measure submitted by Health Partners. The measure developer gave an overview of the General methods approach and the measure submitted to the project.

1558 Relative Resource Use for People with Cardiovascular Conditions
--

<p><b>Description:</b> The risk-adjusted relative resource use by health plan members with specific cardiovascular conditions during the measurement year.</p>
--

<p><b>Resource Use Type:</b> Per capita (population- or patient-based)</p>
--

<p><b>Data Type:</b> Administrative claims, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Imaging/Diagnostic Study, Electronic Clinical Data : Laboratory, Electronic Clinical Data : Pharmacy, Paper Records</p>
---

<p><b>Resource Use Service Category:</b> Inpatient services: Inpatient facility services, Inpatient services: Evaluation and management Inpatient services: Procedures and surgeries, 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</p>
--

# NATIONAL QUALITY FORUM

<p>Ambulatory services: Lab services  <b>Care Setting:</b> Administrative claims, Cardiovascular : Cardiovascular, Electronic Clinical Data : Electronic Health Record, Electronic Clinical Data : Imaging/Diagnostic Study, Population Health : Population Health  <b>Level of Analysis: Administrative claims :</b> Administrative claims, Cardiovascular, Electronic Clinical Data : Electronic Health Record            Electronic Clinical Data : Imaging/Diagnostic Study, Population Health  <b>Measure Developer:</b> National Committee for Quality Assurance (NCQA), 1100 13th Street NW, STE 1000, Washington, District Of Columbia, 20005</p>
<p><b>Steering Committee Recommendation for Endorsement:</b>  <b>Rationale:</b> <i>Pending Committee's official vote.</i></p>
<p><b>Conditions/Questions for Developer:</b></p> <ul style="list-style-type: none"> <li>• Are Coronary Artery Disease equivalents included, such as ischemic heart disease?</li> <li>• How does the stratification discern between high and low risk patients?</li> <li>• What is the time frame for exclusions?</li> <li>• How would a provider know which behaviors to curb based on the report?</li> </ul> <p><b>Developer Response:</b></p> <ul style="list-style-type: none"> <li>• This is based on the HEDIS measure, covers acute and sub-acute, ischemic heart disease, cardiovascular unspecified, angina, atherosclerosis of extremity, etc. CAD related codes diverged into family history etc. Did not try to account for something other than what CAD is described as in the code set. Will look into including code sets that are non CAD specific, non-traditional patients.</li> <li>• In terms of stratification for the risk adjustment it is dependent on the number of comorbidities. Section 10.1 risk adjustment methods, identifies based on qualifying and HCC rankings.</li> <li>• The time frames align with the eligible population period; patients are looked at a year prior to the measurement year and are looked at the year prior to and during the eligibility period.</li> <li>• The reports are divided up by resource categories; user would need to look into measure specifications which are fairly broad.</li> </ul> <p><b>Steering Committee Follow-up:</b> N/A</p> <ul style="list-style-type: none"> <li>• Has this type of risk adjustment model been validated in the past? HCC are well validated. RTI evaluated this in April 2011 and it continues to be a valid stratification method. Follow up needed on who has validated methodology.</li> <li>• Time period for eligibility for risk adjustment/ exclusions.</li> </ul> <p><b>If applicable, Questions to the Steering Committee:</b> N/A</p>
<p><b>1. Importance to Measure and Report</b>  <b>1a.High Impact:</b> <i>Pending Committee's Final Evaluation</i>  <b>Discussion:</b> The TAP agreed that this subcriterion has been met.  <b>1b. Resource use/cost problems:</b> <i>Pending Committee's Final Evaluation</i>  <b>Discussion:</b> The TAP agreed this subcriterion has been met and is supported by the evidence.  <b>1c. Purpose clearly described:</b> <i>Pending Committee's Final Evaluation</i>  <b>Discussion:</b> Inclusion criteria for this measure is very broad – PCI and CABG but not other codes associated with chronic conditions. In the future, if someone is being rated by an institution they wouldn't know where to begin based on how broad the category is. The costs of carotid disease are included in the category, it does capture costs but there is the issue of which costs are incorporated and which costs are not. Given the broad category, the calculation of costs is difficult for a user to understand. This measure covers all costs across all procedures and excludes those who were screened and had plaque in their carotid paired equally as those with PCIs and that early detection may become a preponderance of those grouped in cardiovascular disease.  <b>1d. Resource use service categories consistent and representative:</b> <i>Pending Committee's Final Evaluation</i>  <b>Discussion:</b> The TAP agreed that this subcriterion has been met.</p>
<p><b>2. Scientific Acceptability of Measure Properties:</b>  <b>2a. Reliability:</b> <i>Pending Committee's Final Evaluation</i>  <b>2a1. Well defined/precise specifications:</b>  <b>Discussion:</b> The specifications don't consider the cost, rather use what RVUs would be, i.e., the actual resource use versus the cost. The Steering Committee believed this to be a relevant way to approach the measure, each grouping and person is stratified according to risk. It is unclear which risk adjustment is used for each patient, public algorithm in terms of how this is interpreted. This measure is calculated by using databases from insurers, up to age 75 and only reports on organizations with more than 400 people in the measure; restrict use for larger groups.  <b>2a2. Reliability testing:</b>  <b>Discussion:</b> The reliability testing uses data from 15 months, the results are consistent with other models.  <b>2b. Validity:</b> <i>Pending Committee's Final Evaluation</i>  <b>2b1. Specifications consistent with resource use/cost problem:</b>  <b>Discussion:</b> Discussion similar to 2a1. It is unclear which risk adjustment is used for which patient.</p>

# NATIONAL QUALITY FORUM

<p><b>2b2. Validity testing:</b>  <i>Discussion:</i> NCQA publicly reported the results annually and continues to publicly report. The costs are standardized and are good measures of the resources being used. There is a track record of data being clean, includes resource use not what was actually charged.</p> <p><b>2b3. Exclusions:</b>  <i>Discussion:</i> the measure is unclear regarding the time period for exclusions.</p> <p><b>2b4. Risk adjustment:</b>  <i>Discussion:</i> It is difficult to discern what is included in risk adjustment criteria, once an institution is using this data the severity of illness and compare one to another is based largely on stratification. Unclear how stratification is working and if the groups produced is legitimate?</p> <p><b>2b5. Identification of statistically significant/meaningful differences:</b>  <i>Discussion:</i> The Steering Committee has agreed this subcriterion has been met.</p> <p><b>2b6. Multiple data sources:</b>  <i>Discussion:</i> N/A</p> <p><b>2c. Stratification for disparities:</b>  <i>Discussion:</i> N/A</p>
<p><b>3. Usability:</b></p> <p><b>3a. Measure performance results are publicly reported:</b>  <i>Discussion:</i> This measure has been utilized for a short amount of time (since 2007), difficult to assess if the manner in which they are reporting is useful.</p> <p><b>3b. Measure results are meaningful/useful for public reporting and quality improvement:</b>  <i>Discussion:</i> There is no data on how consumers are utilizing the data and making changes based on this measure. It is unclear what would or would not affect the score and change practices in the long run. The measure would rate fairly low for this subcriterion, it may not be extremely useful for public reporting as it's not easily interpreted. <b>Other steering committee members argue that there are many types of audiences and that, in fact, consumer communication experts can translate the results, but in order to do this, require measures are being used.</b></p> <p><b>3c. Data and results can be decomposed for transparency and understanding:</b>  <i>Discussion:</i> The measure is very broad; moving forward with changes is difficult as its unclear how providers can change behavior.</p> <p><b>3d. Harmonized or justification for differences:</b>  <i>Discussion:</i> N/A</p>
<p><b>4. Feasibility:</b></p> <p><b>4a. Data elements routinely generated during care process:</b>  <i>Discussion:</i> This subcriterion has been met.</p> <p><b>4b. Data elements available electronically:</b>  <i>Discussion:</i> All data is available electronically.</p> <p><b>4c. Susceptibility to inaccuracies/ unintended consequences identified:</b>  <i>Discussion:</i> This subcriterion has been met.</p> <p><b>4d. Data collection strategy can be implemented:</b>  <i>Discussion:</i> This subcriterion has been met.</p>

<p><b>1572 Episode of care for management of chronic coronary artery disease</b></p> <p><b>Description:</b> Resource use and costs associated with management of chronic coronary artery disease (CAD) care over a one-year period. Patients are identified with a diagnosis of CAD in the year prior to the measurement year and the resource use and costs associated with CAD during the measurement year are assessed.</p> <p><b>Resource Use Type:</b> Per episode</p> <p><b>Data Type:</b> Administrative claims, other</p> <p><b>Resource Use Service Category:</b> 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, Durable Medical Equipment (DME)</p> <p><b>Care Setting:</b> Ambulatory Care : Ambulatory Surgery Center (ASC), Ambulatory Care : Clinic/Urgent Care, Ambulatory Care : Clinician Office, Hospital/Acute Care Facility, Imaging Facility, Laboratory, Pharmacy</p> <p><b>Level of Analysis:</b> Clinician, Individual</p> <p><b>Measure Developer:</b> American Board of Medical Specialties Research and Education Foundation, 222 N. LaSalle St., Suite 1500,</p>
--

# NATIONAL QUALITY FORUM

Chicago, Illinois, 60601
<b>Steering Committee Recommendation for Endorsement:</b> <b>Rationale:</b> <i>Pending Committee's official vote.</i>
<b>Conditions/Questions for Developer:</b> <ul style="list-style-type: none"> <li>• For inclusion criteria should 414.10 (aneurysm) be included?</li> <li>• Are patients with angina included or excluded?</li> </ul> <b>Developer Response:</b> <ul style="list-style-type: none"> <li>• Inclusion with PCI or revascularization prior AMI, acute coronary, standard exclusions with vasculitis.</li> <li>• Patients with angina are not exclusively included or excluded – will follow up with the TAP members.</li> </ul> <b>Steering Committee Follow-up:</b> N/A <ul style="list-style-type: none"> <li>• Provide evidence/rationale why angina is not included. For the most common and specific diagnosis of coronary artery disease, angina is usually a coincidence. Angina without CAD diagnosis would not be necessary. This would only include a very small population and would not make a large difference. There are also codes, bypass surgery for something other than CAD is far and few between. Computer algorithm that runs all day and night and automatically gives them a 4.4 code.</li> <li>• More rationale for inclusion/exclusion criteria.</li> </ul>
<b>If applicable, Questions to the Steering Committee:</b> N/A
<b>3. Importance to Measure and Report</b> <b>1a. High Impact:</b> <i>Pending Committee's Final Evaluation</i> <b>Discussion:</b> This measure demonstrates high impact. <b>1b. Resource use/cost problems:</b> <i>Pending Committee's Final Evaluation</i> <b>Discussion:</b> This subcriterion has been met, high cost, high utilization and high morbidity condition. <b>1c. Purpose clearly described:</b> <i>Pending Committee's Final Evaluation</i> <b>Discussion:</b> This subcriterion has been met. <b>1d. Resource use service categories consistent and representative:</b> <i>Pending Committee's Final Evaluation</i> <b>Discussion:</b> This subcriterion has been met.
<b>4. Scientific Acceptability of Measure Properties:</b> <b>2a. Reliability:</b> <i>Pending Committee's Final Evaluation</i> <b>2a1. Well defined/precise specifications:</b> <b>Discussion:</b> To qualify for inclusion had to have atherosclerosis in any diagnosis field and then the exclusions were previous bypass in year prior, PCI or AMI, acute coronary syndrome and standard exclusions common to other – cancer, ESRD, HIV, pregnancy, etc. Angina is not excluded or included, everyone is being captured but trying to look at patients with chronic artery disease, only using atherosclerosis code – should angina also be included? The measure may be too narrowly focused; other codes capture chronic ischemic heart disease. Patients that died are not included, the developers attempted to create a homogenous population to assess resource use. If someone died during measurement yet, and went through a long hospitalization period, but not included in resource use it may be unreasonable. If someone has an AMI the costs are only counted if the individual lives – however, this is not expected to be a high count. This measure is not capturing mortality, rather, capturing hospitalization – mortality is more of a quality issue. <b>2a2. Reliability testing:</b> <b>Discussion:</b> This measure uses market scan commercial payer data sets to test the measure, there are reports showing the results of the testing. <b>2b. Validity:</b> <i>Pending Committee's Final Evaluation</i> <b>2b1. Specifications consistent with resource use/cost problem:</b> <b>Discussion:</b> The specifications may be too narrowly focused and exclude certain patients. The measure developers should calculate the score and compare it to differences with different groups in order to gauge its significance. <b>2b2. Validity testing:</b> <b>Discussion:</b> Validity testing uses commercial payer data sets to test the measure; this is adequate for this measure. <b>2b3. Exclusions:</b> <b>Discussion:</b> Measure developer must provide more rationale for exclusion criteria. <b>2b4. Risk adjustment:</b> <b>Discussion:</b> HCC risk adjustment, several were looked at and it seems to be working well for this measure. Sampling to test was around the risk adjustment development. <b>2b5. Identification of statistically significant/meaningful differences:</b> <b>Discussion:</b> This is a measure for individual physicians to the calculation of the score was similar to another measure – so they were looking at attributing to individual physicians for management of claims 70% or more. Otherwise, it's possible to attribute to others for the other 30% so the score at the individual physician level and is one cost score. The sample report used several different types of service categories (this is associated with the codes). <b>2b6. Multiple data sources</b>

# NATIONAL QUALITY FORUM

<p><i>Discussion:</i> N/A</p> <p><b>2c. Stratification for disparities:</b></p> <p><i>Discussion:</i> N/A</p>
<p><b>3. Usability:</b></p> <p><b>3a. Measure performance results are publicly reported:</b></p> <p><i>Discussion:</i> These measures have not yet been publicly reported; therefore it is difficult to say whether or not they are adequately reported in their current state.</p> <p><b>3b. Measure results are meaningful/useful for public reporting and quality improvement:</b></p> <p><i>Discussion:</i> There is no data on how consumers are utilizing the data and making changes based on this measure. It is unclear what would or would not affect the score and change practices.</p> <p><b>3c. Data and results can be decomposed for transparency and understanding:</b></p> <p><i>Discussion:</i> This criterion has been met.</p> <p><b>3d. Harmonized or justification for differences:</b></p> <p><i>Discussion:</i> N/A</p>
<p><b>4. Feasibility:</b></p> <p><b>4a. Data elements routinely generated during care process:</b></p> <p><i>Discussion:</i> This subcriterion has been met.</p> <p><b>4b. Data elements available electronically:</b></p> <p><i>Discussion:</i> All data is available electronically.</p> <p><b>4c. Susceptibility to inaccuracies/ unintended consequences identified:</b></p> <p><i>Discussion:</i> This subcriterion has been met.</p> <p><b>4d. Data collection strategy can be implemented:</b></p> <p><i>Discussion:</i> This subcriterion has been met.</p>

<p><b>1591 ETG based Congestive Heart Failure (CHF) resource use measure</b></p> <p><b>Description:</b> 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.</p> <p><b>Resource Use Type:</b> Per episode</p> <p><b>Data Type:</b> Administrative claims, other</p> <p><b>Resource Use Service Category:</b> 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><b>Care Setting:</b> 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><b>Level of Analysis:</b> Clinician : Group/Practice, Clinician : Individual, Clinician : Team, Facility, Health Plan, Integrated Delivery System</p> <p><b>Population :</b> Community, Population : County or City, Population : National, Population : Regional, Population : states</p> <p><b>Measure Developer:</b> Ingenix, 950 Winter Street, suite 3800, Waltham, Massachusetts, 02451</p>
<p><b>Steering Committee Recommendation for Endorsement:</b></p> <p><i>Rationale: Pending Committee's official vote.</i></p>
<p><b>Conditions/Questions for Developer:</b></p>

# NATIONAL QUALITY FORUM

- What does a complete episode group consist of? 12 month period?
- How does it handle diastolic vs. systolic?

**Developer Response:**

- Claims and dollars associated with treating a disease within a 12 month period.
- Did not single out type of CHF. There is a separate episode so the claims are not included for cardiac myopathy and then if there is a mix of diastolic and systolic – goes to systolic coding and the severity model.

**Steering Committee Follow-up:** N/A

- The mix between systolic and diastolic should be included in the documentation.
- Diagnostic codes for CHF should be included – is it CHF, systolic, doesn't need to be congested at the time just less than 40%.

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

**5. Importance to Measure and Report**

**1a. High Impact:** *Pending Committee's Final Evaluation*

*Discussion:* The Steering Committee believes this subcriterion has been met.

**1b. Resource use/cost problems:** *Pending Committee's Final Evaluation*

*Discussion:* The Steering Committee believes this subcriterion has been met.

**1c. Purpose clearly described:** *Pending Committee's Final Evaluation*

*Discussion:* The Steering Committee believes this subcriterion has been met.

**1d. Resource use service categories consistent and representative:** *Pending Committee's Final Evaluation*

*Discussion:* The Steering Committee believes this subcriterion has been met.

**6. Scientific Acceptability of Measure Properties:**

**2a. Reliability:** *Pending Committee's Final Evaluation*

**2a1. Well defined/precise specifications:**

*Discussion:* The specifications target commercial, non-elderly individuals between the years 2006-2010. It appears that comparisons were made to Ingenix's own internal benchmarking database; they need to show more external comparison. Large commercial insurers are included in the Ingenix database and are spread out geographically. Those identified as having CHF appear to be adequate with this measure.

**2a2. Reliability testing:**

*Discussion:* Strong reliability testing was conducted; this subcriterion has been adequately met.

**2b. Validity:** *Pending Committee's Final Evaluation*

**2b1. Specifications consistent with resource use/cost problem:**

*Discussion:* This is not just CHF, CHF, systolic specifically. If you have both diastolic and systolic you would be included in both but recorded as systolic. This is found in medical record data and claims data. The developer could possibly restrict the codes for those with systolic failure 42840 and 42841 are used.

**2b2. Validity testing:**

*Discussion:* The Steering Committee believes there is quite a bit of content validity within Ingenix's large database, as well as regression analysis. This is not of great deal of concern except in regards to disparities; there was a broad range of clinical context for measurement.

**2b3. Exclusions:** *This subcriterion will be discussed during an upcoming conference call.*

*Discussion:*

**2b4. Risk adjustment:** *This subcriterion will be discussed during an upcoming conference call.*

*Discussion:*

**2b5. Identification of statistically significant/meaningful differences:** N/A

**2b6. Multiple data sources:** N/A

**2c. Stratification for disparities:** N/A

**3. Usability:** *This criterion will be discussed during an upcoming conference call.*

**3a. Measure performance results are publicly reported:**

**3b. Measure results are meaningful/useful for public reporting and quality improvement:**

**3c. Data and results can be decomposed for transparency and understanding:**

**3d. Harmonized or justification for differences:**

**4. Feasibility:** *This criterion will be discussed during an upcoming conference call.*

**4a. Data elements routinely generated during care process:**

**4b. Data elements available electronically:**

**4c. Susceptibility to inaccuracies/ unintended consequences identified:**

**4d. Data collection strategy can be implemented:**

# NATIONAL QUALITY FORUM

## **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. An additional call will be scheduled to finish the discussion of 1591 ETG based Congestive Heart Failure (CHF) resource use measure submitted by Ingenix.