NATIONAL QUALITY FORUM

CONFERENCE CALL OF THE RESOURCE USE STEERING COMMITTEE

June 22, 2011

Committee Members Participating: Bruce Steinwald, MBA (Co-Chair), Independent Consultant; Tom Rosenthal, MD (Co-Chair), UCLA School of Medicine; Paul Barnett, PhD, VA Palo Alto Health Care System; Jeptha Curtis, MD, Yale University School of Medicine; Mary Kay O'Neill, MD, MBA, CIGNA HealthCare; Steve Phillips, MPA, Ortho-McNeill-Janssen Pharmaceutical, Inc.; Barbara Rudolph, PhD, MSSW, The Leapfrog Group; Joseph Stephansky, PhD, Michigan Health and Hospital Association; Dolores Yanagihara, MPH, Integrated Healthcare Association.

NQF Staff Participating:; Taroon Amin, MPH, MA, Senior Director; Ashlie Wilbon, MPH, BSN, Senior Project Manager; Lauralei Dorian, Project Manager; Sarah Fanta, Project Analyst; Carlos Alzola, NQF Statistical Consultant; Sally Turbyville, MA, MS, NQF Project Consultant.

Others Present: Cheri Zielinski, Ingenix; Thomas Lynn, Ingenix; Daniel 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 Scientific Acceptability criterion for the non-condition specific measures 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. 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 non-condition specific measure submitted by Ingenix. The measure developer gave an overview of the General methods approach and the measure submitted to the project.

1599 ETG Based Non-Condition Specific resource use measure
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. **Resource Use Type:** Per capita (population- or patient-based) **Data Type:** Administrative claims **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 : 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:

- How does the risk score correlate with the actual expenditures?
- What is the distinction between ETGs and ERGs?
- Can this measure be applied to the Medicare population?
- Have there been any changes in the underlying risk model used in the ETGs since what has been published on the Ingenix web site a year ago?
- How are the carve outs, pharmacy and mental health data handled? How was this data validated?

Developer Response:

- Ingenix provides options for expenditure thresholds for a patient's annual member costs: $25,000, $100,000, and $250,000. Ingenix explained that these thresholds would vary depending on the application.
- ETGs are episode-based measures. For example, an episode of diabetes, congestive heart failure or COPD--the severity models are built separately for each of the conditions which allows for risk adjustment for each separate condition-based episode. The results are then tagged for each episode for a member not only by condition, but also by the level of severity. There are hundreds of ETGs that map into the ERGs. Ingenix maps to the ERG designed for the population-based risk adjustment; they weight each of the ERG markers to the final ERG score. The ERGs looks at age, in which case they may be applied to the Medicare population, however not all of the ETGs take age into account in the risk adjustment model. During the developer testing they didn't find that age had much explanatory power so they are not included in all of the ERGs. The ERG will point to a different weight depending on the age of the individual. However, since this measure has only been tested in a commercial database, per NQF policy, it can only be endorsed for use in commercial populations.
- The ETG models and the risk models related to the ETGs have not been updated or recalibrated within the last year; therefore the information on the Ingenix website is still applicable.
- Ingenix works with a population that has pharmacy and medical data. Mental health is excluded because the claims are not often available in addition to lack of coding for mental health services. Pharmacy data hasn't been an issue because it's up to the user whether they want to include and compare populations who have pharmacy data. The methodology can be adjusted, you are able to a have a mixed population of both medical and pharmacy benefits, and the user is able to isolate the medical resource use data if they choose to.

Importance to Measure and Report – This criterion was discussed during the June 6 conference call. To access the webinar presentation, please click here.
1a.High Impact: *Pending Committee's Final Evaluation*
Discussion: The Steering Committee has deemed this measure to be a high cost/high impact area, this sub criteria has been met.
1b. Resource use/cost problems: *Pending Committee's Final Evaluation*
Discussion: The Steering Committee believes this measure has met the sub criteria.
1c. Purpose clearly described: *Pending Committee's Final Evaluation*
Discussion: The Steering Committee believes the measure has met this sub criterion, as the measure's purpose is clearly described.
1d. Resource use service categories consistent and representative: *Pending Committee's Final Evaluation*
Discussion: The Steering Committee believes the measure has met this sub criterion, the resource use service categories are representative of the measure.

Scientific Acceptability of Measure Properties:
2a. Reliability:
2a1.Measure well defined and precisely specified:
Discussion: This measure appears to be well defined and specified. This methodology is used in a number of organizations and appears to work well. This sub criterion has been met.
2a2. The results are repeatable:
Discussion: The TAP agreed that this sub criterion has been met; the results have shown to be repeatable.
Overall Reliability:
Discussion: The reliability of the data appears to be quite strong. The Ingenix team has a robust system where they double code the data – the steps that lead to the production of the data has a 99.9% match between the two approaches, which makes it highly reliable.

The TAP felt that tables that depict the reliability are not clearly labeled or defined. The TAP agreed that this sub criterion has been met; the results have shown to be repeatable.

2b. Validity:

2b1. Evidence is consistent with intent:

Discussion: The specifications appear to be consistent with the intent; the TAP agrees this sub criterion has been met.

2b2.Score/Analysis:

Discussion: This measure has been tested for face validity but has not been tested against other databases.

2b3. Exclusions:

Discussion: Ingenix noted that there are no exclusions based on cost or other criteria; however the committee noted that since some plans have pharmacy data and others do not, comparability would be difficult.

2b4. Risk Adjustment

Discussion: When looking at the ETG codes, Ingenix assigns a severity score. The methodology takes into account not only the ETG severity codes but also the number of comorbidities. A retrospective model contains the observed episodes that may occur during that year, but a user will not be able to observe any markers or costs for people who did not undergo services. The ERG risk level determines the individual's ERG risk score which drives the risk adjustment. The risk adjustment was adequately described, no follow up with Ingenix is necessary.

2b5. Identification of statistically significant/meaningful differences:

Discussion: There is a way to stratify those with or without pharmacy data. The TAP expressed concern that valid comparisons cannot be made across organizations with different levels of data completeness and consistency.

2b6. Multiple data sources:

Discussion: N/A (using all administrative data)

Overall Validity:

Discussion: In the submission, Ingenix states that they apply the methodology to data from several different organizations, but this is not detailed in any of the results. They tested the face validity by intentionally modifying the data and converting resource pre and post modifications, however there is not any description of the results within the submission. The tables that depict the validity are not clearly labeled or defined.

2c. Stratification for disparities:

Discussion: Ingenix does not stratify by race and ethnicity. This may be handled in the future, but at the present time this information is not available.

Usability:

3a. Measure performance results are publicly reported:

Discussion: Ingenix conducted a survey of their customers, some users are publicly reporting the data and others are sharing information with physicians for incentive based programs. Some users have decided to put the information on a website that goes to their providers, which allows them to access their risk scores and score card. Providers are then able to drill down on the scorecard to the claim base level, the patient level and then the overall claims level.

***The remainder of Usability criterion will be discussed during an upcoming conference call.*

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

Discussion:

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

Discussion:

3d. Harmonized or justification for differences:

Discussion: N/A

***The Feasibility criterion will be discussing during an upcoming conference call.*

4. Feasibility:

4a. Data elements routinely generated during care process:

Discussion:

4b. Data elements available electronically:

Discussion:

4c. Susceptibility to inaccuracies/ unintended consequences identified:

Discussion:

4d. Data collection strategy can be implemented:

Discussion:

PUBLIC COMMENT

Kay Jewell from Centers for Consumers of Healthcare commented that the issue of public reporting is extremely relevant if the measurers will be used by public payers. She stated the "public" should be broadly defined as "anyone who would be influenced or affected by the measure". Public reporting would affect payers, as this may influence incentives or value based purchasing and directly affect how healthcare providers are reimbursed. It may also be relevant for the consumer to know what may be influencing the provider that they have selected.

The Steering Committee thanked Ms. Jewell for this comment.

NEXT STEPS

Ms. Wilbon indicated that project staff will continue with preparations for the in person Steering Committee meeting on June 29-30 in Washington, D.C. To access the dial-in information and agenda for that meeting, please click here.