NATIONAL QUALITY FORUM

CONFERENCE CALL OF THE RESOURCE USE BONE/JOINT TECHNICAL ADVISORY PANEL (TAP)

August 5, 2011

Committee Members Participating: James Weinstein, DO, MS (*Chair*), Dartmouth Institute for Health Policy; Mary Kay O'Neill, MD, MBA, CIGNA HealthCare; John Ratliff, MD, FACS, Thomas Jefferson University; Catherine Roberts, MD, Mayo Clinic; Craig Rubin, MD, University of Texas Southwestern Medical School; Patricia Sinnott, PT, PhD, MPH, VA health Economics Resource Center, Menlo Park, CA

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 via telephone: Cheri Zielinski, Ingenix; Tom Lynn, Ingenix.

MEETING PROCESS

Ms. Fanta welcomed the Bone/Joint Technical Advisory Panel (TAP) and thanked them for their participation in their final conference call. The purpose of the call was to discuss the Scientific Acceptability criteria for updated Ingenix measure# **1603:** ETG Based Hip Fracture Resource Use Measure. The TAP were reminded that for this updated version of this measure, Ingenix had separated out hip fractures of what had originally been a combined hip fracture / pelvic fracture measure. Ingenix had done so at the request of the TAP during their in person meeting on July 7, 2011.

Ingenix measure developers were available on the call to respond to questions from the TAP 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 measure submitted by Ingenix.

1603 ETG PEG Based Hip Fracture resource use measure (Ingenix)
Description: The measure focuses on resources used to deliver episodes of care for patients with Hip Fracture. Hip Fracture 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 Hip Fracture. A number of resource use measures are defined for Hip Fracture 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 Hip Fracture episodes and will cover both measures at the Hip Fracture base and severity level and also a Hip Fracture composite measure where Hip Fracture episode results are combined across Hip Fracture severity levels. At the most detailed level, the measure is defined as the base condition of Hip Fracture and an assigned level of severity (e.g., resources per episode for Hip Fracture, 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 Hip Fracture is derived by combining Hip Fracture episode results across Hip Fracture severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of Hip Fracture episodes by severity level when supporting a Hip Fracture composite comparison).The focus of this measure is on Hip Fracture. However, Hip Fracture episode results could also be included in an "orthopedics", "acute care", or other clinical composite for a physician, combining episodes in clinical areas similar to Hip Fracture. 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; Admissions/discharged
Ambulatory services: Outpatient facility services; Emergency Department; Pharmacy; Evaluation and management; Procedures and surgeries; Imaging and diagnostic; 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:

- Why are different age groups are assigned the same risk coefficients, when they will have extremely different risk factors?
- How does the episode grouper work in terms of low and high outliers? Are you able to provide information on exactly how many episodes have been excluded?
- Why do you cut the bottom costs from being included in the measure?

Developer Response:

- This represents a limit of the data set. Due to the minimal number of people over 65 in commercial programs, we didn't have the numbers to further stratify.
- We exclude cases that are low in cost. We have the data to talk about the number of cases that are excluded by varying a low outlier, yes.
- The hypothesis that that these low cost episodes – ones under 2.5 percent – are either mistakes or miscodes. They are probably incomplete episodes, so we don't count them.

1. Importance to Measure and Report
1a.High Impact: H-2; M-1; L-2; I- 0; N-0
TAP Discussion: There was general agreement that hip fracture is a major cause of morbidity, mortality and high resource use. The TAP did, however, question the importance of measuring hip fractures in a predominately under 65 group of patients. Ingenix acknowledged that this was a significant limitation of using administrative data.
1b. Resource use/cost problems: H-2; M-2; L-1; I-0; N-0
TAP Discussion: No issues were identified.
1c. Purpose clearly described: H-1 ; M-4; L- 0; I-0; N-0
TAP Discussion: No issues were identified.
1d. Resource use service categories consistent and representative: H-2; M-2; L-1; I-0; N-0
TAP Discussion: The TAP were concerned that resource use service categories omit nursing homes and inpatient or outpatient rehab services.

1. Scientific Acceptability of Measure Properties:
2a. Reliability:
2a1.Measure well defined and precisely specified: H-1 ; M-2 ; L-2; I-0; N-0
TAP Discussion: The TAP was concerned that the measure didn't capture certain co-morbid conditions such as dementia which are critical to understanding resource use for this clinical condition. There was substantial unease that the data does not examine the Medicare population, where the majority of hip-fractures occur.
2a2. The results are repeatable: H-1 ; M-2; L-2; I-0; N-0
TAP Discussion: The panel questioned whether one could infer grouper reliability from the tables submitted by Ingenix. Ingenix explained that the tables illustrate expected variability in results and point to a relatively consistent cost across health care organizations.
Overall Reliability: H-1 ; M-0; L-4; I-0; N-0
2b. Validity:
2b1. Evidence is consistent with intent: H-0; M-0; L-5; I-0; N-0

TAP Discussion: The TAP reiterated their concern that the measure hasn't captured the patient population most likely to be affected by hip fractures. Therefore, the measure may have limited applicability, due to the limitations of using only commercial data. The panel also felt that hip fractures in younger populations versus older populations represent two very different clinical situations.
2b2.Score/Analysis: H-0; M-1 ; L-4; I-0; N-0
TAP Discussion: The TAP was uncomfortable with the fact that all age groups were assigned the same risk coefficients. Ingenix explained that this also represents a limitation of the data set, where they did not have the numbers over 65 to further stratify. Members of the panel believed that certain clinically relevant co-morbities and complications such as dementia and post-op delirium should be reported on in a hip-fracture measure.
2b3. Exclusions: H-0; M-1; L-4; I-0; N-0
TAP Discussion: The TAP felt that the reasoning behind the exclusion criteria was unclear and not based on clinical evidence.
2b4. Risk Adjustment: H-0; M-0; L-4 ; I-1; N-0
TAP Discussion: The developer described how the measure contains low dollar exclusions. The assumption is that these claims represent incomplete episodes.
2b5. Identification of statistically significant/meaningful differences: H-0; M-0 ; L-4 ; I-1; N-0
TAP Discussion: There was a discussion regarding the relative cost of care ratio and a question about what numbers represent statistically significant differences. Ingenix explained that the numbers would depend on the confidence interval, the underlying variance of episode cost and the number of total cases.
2b6. Multiple data sources:
TAP Discussion: N/A (using all administrative data)
Overall Validity: H-0; M-1 ; L-3; I-0; N-0
2c. Stratification for disparities: H-0; M-1; L-1 ; I-3; N-0
TAP Discussion: Racial disparities were addressed in the submission, but the data limits a further examination into these disparities.

Usability:
3a. Measure performance results are publicly reported: H-0 ; M-2 ; L-3; I-0; N-0
TAP Discussion: No issues were raised.
3b. Measure results are meaningful/useful for public reporting and quality improvement: H-0; M-1; L-4; I-0; N-0
TAP Discussion: The TAP acknowledged the impressive amount of work Ingenix put into this measure, but again articulated concern that the measure would have limited meaningful use as it is not capturing the appropriate population. The panel was uneasy with the grouping of two clinically different age cohorts together into one measure; they felt that the clinical situation, treatment path and mortality for a younger population with hip fractures versus an older population were different enough to warrant two separate measures.
3c. Data and results can be decomposed for transparency and understanding: H-0 ; M-2 ;L-3; I-0; N-0
TAP Discussion: The TAP agreed data could be decomposed.
3d. Harmonized or justification for differences:
TAP Discussion: N/A

4. Feasibility: Feasibility criteria were not discussed on the call. The Bone/Joint TAP had previously discussed feasibility for other Ingenix measures during their in-person meeting on July 17, 2011. Their comments remain applicable to this measure.
4a. Data elements routinely generated during care process: H-3 ; M-1 ; L-1; I-0; N-0
4b. Data elements available electronically: H-4 ; M-0 ; L-1; I-0; N-0
4c. Susceptibility to inaccuracies/ unintended consequences identified: H-1 ; M-1 ; L-3; I-0; N-0
4d. Data collection strategy can be implemented: H-0 ; M-2 ; L-2; I-1; N-0

PUBLIC COMMENT

There were no public comments.

NEXT STEPS

Ms. Fanta thanked the TAP for their dedication in reviewing Bone/Joint measures. Dr. Weinstein was also thanked for his leadership of the panel. The Steering Committee will meet in person to review this and six other measures on August 30-31st.