NATIONAL QUALITY FORUM

TO: Cardiovascular Endorsement Maintenance Steering Committee

FR: Reva Winkler, MD, MPH; Kathryn Streeter, MS

SU: NQF policy on "reserve" measures

DA: May 31, 2011

At the recent meetings, the Steering Committee recommended that three measures be retired because current performance is very high and there seems to be little opportunity for improved performance. These measures have been successful in driving improvement in performance, but concerns have been raised about possible decline in performance if measurement is discontinued. The NQF Board of Directors has recently approved a policy for a special designation -- "reserve" measures -- that meet the following criteria:

> A subset of measures evaluated during Endorsement Maintenance that have consistently demonstrated highest levels of performance may maintain endorsement with a special designation in the portfolio if they still address critical aspects of quality that require reinforcement and completely meet all other endorsement criteria. This category should be used by exception only for those measures that receive high ratings on evidence, validity, and reliability.

"Reserve" status would apply only to highly credible, reliable, and valid measures that have high levels of performance due to quality improvement actions (often facilitated or motivated through public reporting and pay-for-performance programs) rather than problems with the measure specifications. The key issue is the opportunity cost associated with measuring processes at high levels of performance—rather than focusing on areas where there is really a gap in care. NQF does not want to move into reserve status measures that are really not needed because they are too far from the desired outcome.

CONSIDERATIONS FOR RESERVE ENDORSEMENT STATUS

The data provided in measure submissions are frequently limited. In determining whether there is further opportunity for improvement, the Steering Committee should review data on representation, variation, and disparities:

- What is the representativeness of the data, i.e., is it national data from a majority of hospitals or is the data from a single state or payer group?
- What is the range in performance, particularly in the lowest decile or quartile?
- What is the performance among possible disparities population(s)?
- Is the measure performance data indicating high levels of performance consistent with other evidence (epidemiologic or research)?

- What is the size of the population at risk, effectiveness of an intervention, and consequences of a quality problem (e.g., even modest variation would be sufficient justification for some highly effective, potentially life-saving treatments)?

Other considerations include:

- Is this a measure with strong, direct evidence of a link to a desired health outcome? Generally measures more distal to the desired outcome with only indirect evidence would not qualify, e.g., assessment of blood pressure (BP) measurement rather than the BP value.
- Measures with a focus more distal to a desired outcome are not needed if there is a measure with a focus more proximal to the desired outcome (e.g., venous thromboembolism (VTE) prophylaxis ordered versus VTE prophylaxis administered).
- Is the measure needed if outcomes (i.e., mortality, readmission) of care are being measured?

MEASURES FOR POSSIBLE RESERVE ENDORSEMENT STATUS

The Steering Committee has recommended three measures that may be candidates for reserve endorsement status. The Committee voted that the following measures did not meet the *Importance to Measure and Report* criteria due to high performance and lack of opportunity for improvement:

- 0160 Beta blocker prescribed at discharge [for AMI] (CMS)
- 0142 Aspirin prescribed at discharge for AMI (CMS)
- 0135 Heart failure: Evaluation of left ventricular systolic dysfunction (CMS)

DATA ON OPPORTUNITY FOR IMPROVEMENT

160 Beta blocker prescribed at discharge

National performance rates: 1Q10: 98.2% 4Q09 98.3% 3Q09 98.2% 2Q09 98.1%

Representative: 1Q10: 105,436 acute myocardial infarction (AMI) patients, 3111 hospitals

Range/variation: additional data on percentile distribution is attached

Disparities: see Centers for Medicare & Medicaid Services (CMS) disparities spreadsheets

142 Aspirin prescribed at discharge for AMI

National performance rates: 1Q10: 98.5% 4Q09 98.5% 3Q09 98.4% 2Q09 98.3%

Representative: 1Q10: 107,852 AMI patients, 3096 hospitals

Range/variation: additional data on percentile distribution is attached

Disparities: see CMS disparities spreadsheets

135 Evaluation of left ventricular systolic dysfunction

National performance rates: 2Q09: 97.2%, 3Q09: 97.3%, 4Q09: 97.6%, 1Q10: 97.8%

Representative: 1Q10: 212,985 HF patients, 4,087 hospitals

Range/variation: additional data on percentile distribution is attached

Disparities: see CMS disparities spreadsheets

COMPLETE MEASURE EVALUATION

ACTION ITEM: Since measures under consideration for reserve endorsement must meet all criteria, the Committee must rate the measures on all four endorsement criteria. Measures 160 and 142 were only evaluated on the Importance criteria. The measure submission forms are provided again for your reference. The preliminary ratings from the work group members are provided below:

160 Beta blocker prescribed at discharge (CMS)

SCIENTIFIC ACCEPTABILTY		
2a Specifications	C=4	2a. The measure is precisely specified. 2b. Testing demonstrates reliability. 2c. The TAP has accepted the face validity of the measure. 2d. Exclusions are consistent with current ACC/AHA performance measures. 2e. Risk adjustment is not necessary. 2f. Meaningful differences in rates are reported. 2g. Paper record abstraction and extraction of the data from an electronic health record (EHR) have not been compared. 2g. Not addressed. 2h. In addition to the measure steward's recommendation to control for the simultaneous effect of other potential factors, this cardiac measure should be stratified by race and ethnicity, since the performance data suggest potential disparities. 2h. Preliminary analyses suggest that disparities are present, but definitive analyses have not been performed.
2b Reliability	C=4	
2c Validity	C=4	
2d Exclusions	C=4	
2e Risk adjustment	C=1, NA=3	
2f Meaningful differences	C=4	
2g Comparability	C=1 M=1 N=2	
2h Disparities	C=2 P=2	
USABILITY		
3a Understandable	C=4	3a. The measure is currently in use. 3b. This measure's specifications are not harmonized with NQF #0613 measure specifications, as the latter's measure population uses the outpatient setting and includes patients diagnosed with MI at any time in the past. 3c. No other NQF measure addresses this target population. 3c. NQF #0071 is a more appropriate measure (long-term adherence versus in-hospital treatment) and is a better reflection of appropriate clinical care.
3b Harmonization	C=1 P=1 N=1 NA=1	
3c Added value	C=2 N=1 NA=1	
FEASIBILITY		
4a Data a by-product of care	C=4	4a. The data are generated during routine clinical care. 4b. The data must be abstracted from paper records. 4c. Exclusions do not require additional data. 4d. Efforts are underway to minimize errors of inclusion/exclusion. 4d. Monitoring of the use of "other reason" is important to guarantee that this category is used consistently and appropriately. 4e. The strategy is already implemented and modifications have eased the data collection burden.
4b Electronic	C=2 P=1 N=1	
4c Exclusions	C=4	
4d Inaccuracies	C=2 P=2	
4e Implementation	C=3	

0142 Aspirin prescribed at discharge (CMS)

SCIENTIFIC ACCEPTABILTY		
2a Specifications	C=3	2f. National performance rate has stayed between 98.3% and 98.5% from 2Q09-1Q10. No data offered on different regions, providers. It is possible that this is such a widely used, accepted metric that the differences are not great but that this measure pushes performance. 2h. Univariate analyses suggest potential disparities, rates range from 96.5% for Hispanic/Latinos, to 97.4% for African Americans, to 98.5 for White/Caucasians to 98.6% for native Americans. To date, stratification analysis has not been performed. Further analyses need to control for other potential confounding factors. The listed performance rates of 98.5% call into question the purpose of this measure in 2011. Do the trivial differences in rates justify the expense of data collection for this measure? Also given the need to report the same thing for other patient subsets, should this measure be subsumed under other reported measures?
2b Reliability	C=3	
2c Validity	C=3	
2d Exclusions	C=1 P=2	
2e Risk adjustment	C=1, NA=2	
2f Meaningful differences	C=1 M=2	
2g Comparability	M=2 NA=1	
2h Disparities	N=2 NA=1	
USABILITY		
3a Understandable	C=1 P=1 N=1	3a. Hospital Inpatient Quality Reporting Program—measures can be used by individual hospitals for internal quality improvement. Results not available. Used successfully in hospital inpatient quality reporting programs. 3b. Not harmonized with NQF #0631, which evaluates primarily patients in the outpatient setting. 3c. No NQF endorsed measures with same topic and target population. The reported minimal differences in rates do not seem to allow for meaningful public reporting. Likely that the measure has accomplished its goal and question its ongoing use. Disagree that there are no other similar measures given the PCI, ischemic vascular disease measures which have large overlap.
3b Harmonization	N=3	
3c Added value	C=1 P=1 M=1	
FEASIBILITY		
4a Data a by-product of care	C=3	4b All the data elements are not presently available in an electronic health record, but retooling work with the Dept of Health and Human Services (HHS) is expected to be completed in 2011. 4c/4d. There are important exclusions that are common: allergy, bleeding diathesis, concomitant therapy with other anti-thrombotics/ anticoagulants that do require additional data sources and are not always easy to retrieve/document. This can affect the accuracy of the measure and lead to errors. . Exclusions are varied and in the past “false exclusions” were relatively common. Changes in the measure such that patients prescribed the medication stayed in the measure attenuated this problem. Data elements in the measure are closely tracked to see if problems arise. 4e. The frequency of questions submitted by abstractors pertaining to aspirin prescription and No Aspirin at discharge amounted to only 3.3% during close tracking of the data elements.
4b Electronic	C=1 P=1 M=1	
4c Exclusions	C=2 P=2	
4d Inaccuracies	C=1 P=2	
4e Implementation	C=2 P=1	

Measure 135 was evaluated by the Committee at the April 8, 2011 meeting:

0135 Evaluation of Left ventricular systolic function (LVSF)

Percentage of heart failure patients with documentation in the hospital record that left ventricular systolic (LVS) function was evaluated before arrival, during hospitalization, or is planned for after discharge. (CMS)

Committee evaluation:
IMPORTANCE: Yes - 15, No – 3
SCIENTIFIC ACCEPTABILTY: C=7, P=6, M=5, N=0
USABILTY: C= 5, P=10 , M=4, N=0
FEASIBILTY: C= 5, P=8 , M=6 , N=0
Meets criteria: Yes - 5 No-13

Discussion: Current performance very high. Concern with misinterpretation of measure so that testing is done at every hospitalization which is not required by the measure. Unintended consequence may be to encourage overuse.

ACTION ITEM: After consideration of the criteria for inactive endorsement, the Steering Committee will vote on a final recommendation for all three measures:

- Continue endorsement
- Reserve endorsement
- Remove endorsement