NATIONAL QUALITY FORUM

Memo

TO: NQF Members

FR: NQF Staff

RE: Voting Draft Report: *National Voluntary Consensus Standards for Neurology, Phase I*

DA: September 11, 2012

Background

Neurological conditions and injuries affect millions of Americans each year, taking a tremendous toll on patients, families, and caregivers, and costing billions of dollars in treatment, rehabilitation, and lost or reduced earnings. Specifically:

- Strokes were the fourth leading cause of death in the United States in 2009, as well as a leading cause of disability.[1]
- Each year, approximately 795,000 people suffer a stroke.[2]
- Health care costs for stroke-related morbidity reached $73.7 billion in 2010.[3]

NQF has endorsed a number of consensus standards to evaluate the quality of care for neurological conditions over the past decade. As quality measurement has matured, better data systems have become available, electronic health records are closer to widespread adoption, and the demand for meaningful performance measures has prompted development of more sophisticated measures of healthcare processes and outcomes for neurological conditions. An evaluation of the NQF-endorsed® neurology measures and consideration of new measures will ensure the currency of NQF's portfolio of voluntary consensus standards.

In Phase 1 of this project, NQF sought performance measures that could be used for accountability and public reporting on stroke and transient ischemic events for adults and children in all settings of care. The project reviewed measures on a variety of topics in this area, including treatments, diagnostic studies, interventions, and procedures.

---

[1] Centers for Disease Control. Available at http://www.cdc.gov/nchs/fastats/stroke.htm Last accessed February 2012.
American Stroke Association. Available at http://www.strokeassociation.org/STROKEORG/AboutStroke/About-Stroke_UCM_308529_SubHomePage.jsp Last accessed February 2012

[2] The Internet Stroke Center. Available at http://www.strokecenter.org/patients/about-stroke/stroke-statistics/ Last accessed February 2012

[3] American Stroke Association. Available at http://www.strokeassociation.org/STROKEORG/AboutStroke/Impact-of-Stroke_UCM_310728_Article.jsp Last accessed February 2012

A 23-member steering Committee reviewed 29 measures, and recommended 14 of these measures for endorsement. Public and member commenting took place from July 13-August 13, 2012.

Comments and Revised Voting Report

NQF received 53 comments from two members of the public and ten NQF members. The comments, with their final responses, are posted in the project page.

Revisions to the draft report and the accompanying measure specifications are identified as redlined changes. (NOTE: Typographical errors and grammatical changes have not been red-lined to assist in reading.)

Comments and their Disposition

The Steering Committee reviewed the comments and focused its discussion on those specific measures and topic areas with the most significant and recurring issues that arose from the comments. Comments about specific measure specifications and rationale also were forwarded to the measure developers, who were invited to respond.

Major Themes

Three major themes were identified in the comments, as follows:

1. Feasibility
2. Harmonization of stroke rehabilitation measures
3. Inclusion of a stroke severity indicator in risk-adjustment models for stroke mortality and readmission measures

Theme 1: Feasibility

Description: We received 13 comments regarding the feasibility of several measures. Most of these comments noted that the measures "may require burdensome electronic health record data extraction or medical chart review."

Specifically, 12 of these comments addressed the following measures:

- 0240: Stroke and Stroke Rehabilitation: Deep Vein Thrombosis (DVT) Prophylaxis for Ischemic Stroke or Intracranial Hemorrhage (AMA-PCPI)
- 0241: Stroke and Stroke (AMA-PCPI)
- Rehabilitation: Anticoagulant Therapy Prescribed for Atrial Fibrillation at Discharge (AMA-PCPI)
- 0243: Stroke and Stroke Rehabilitation: Screening for Dysphagia (AMA-PCPI)
- 0244: Stroke and Stroke Rehabilitation: Rehabilitation Services Ordered (AMA-PCPI)
- 0325: Stroke and Stroke Rehabilitation: Discharged on Antithrombotic Therapy (AMA-PCPI)
- 0434: STK-01: Venous Thromboembolism (VTE) Prophylaxis (TJC)
- 0435: STK 02: Discharged on Antithrombotic Therapy (TJC)
- 0436: STK-03: Anticoagulation Therapy for Atrial Fibrillation/Flutter (TJC)

- 0437: STK 04: Thrombolytic Therapy (TJC)
- 0438: STK 05: Antithrombotic Therapy By End of Hospital Day Two (TJC)
- 0439: STK-06: Discharged on Statin Medication (TJC)
- 0441: STK-10: Assessed for Rehabilitation (TJC)

Also addressed in these comments was a concern that the measures will be difficult to implement from administrative claims because there are limitations in identifying relevant physician behavior in hospital claims. However, the AMA-PCPI measures that were addressed in these comments are clinician-level measures that use CPT-II codes to record the measure focus, and the TJC measures that were addressed are facility-level measures that are not specified for administrative claims.

Action Taken: While SC members recognize that these measures may require a fair amount of data abstraction, they rated the measures moderate to high on NQF's feasibility criterion, and did not wish to revisit their recommendation.

A final comment related to measure feasibility concerned measure #1952 (Time to Intravenous Thrombolytic Therapy). Specifically, the comment addressed the difficulty in implementing this measure from administrative claims alone. This measure was mistakenly specified for administrative claims when originally submitted; however, the developer agreed that the measure cannot be captured by claims data only and has revised the submission to specify the measure for electronic registry data only.

Action Taken: No action by the Committee was required.

Theme 2- Harmonization

Description: We received three comments suggesting that the numerators, denominator exclusions, and timeframe for measures #0244 and #0441 be harmonized.

These measures were identified by NQF staff as related measures because both address rehabilitation services for stroke patients. The measures differ in the following ways:

Number, Title, and Developer	0244 Stroke and Stroke Rehabilitation: Rehabilitation Services Ordered (AMA-PCPI)	0441 STK-10: Assessed for Rehabilitation (The Joint Commission)
Measure focus	Rehab services ordered OR documentation that no rehab needed	Assessed for or received rehab services
Patient population	Patients 18+, dx=ischemic stroke or intracranial hemorrhage	Patients 18+, dx=ischemic stroke or hemorrhagic stroke

Number, Title, and Developer	0244 Stroke and Stroke Rehabilitation: Rehabilitation Services Ordered (AMA-PCPI)	0441 STK-10: Assessed for Rehabilitation (The Joint Commission)
Denominator exclusions	None	Length of Stay > 120 days, comfort measures only documented, enrolled in clinical trials related to stroke, admitted for elective carotid intervention, discharged to another hospital, left against medical advice, expired, discharged to home for hospice care, discharged to a health care facility for hospice care
Timeframe	At/prior to discharge	Hospital admission to discharge
Level of analysis	Clinician	Facility
Data source	Administrative claims, electronic clinical data, EHR, registry	Electronic clinical data, EHR, paper medical records

In their initial discussion of these measures during the in-person meeting, the Committee did not identify any harmonization issues to be addressed by the developers.

Developer response (AMA-PCPI): The AMA PCPI uses measure exclusions and measure exceptions, where appropriate. However, as this measure numerator is constructed to capture patients for whom rehabilitation services were ordered and patients for whom the physician has documented that no rehabilitation services were indicated, there is no need for exclusions or exceptions. All patients that receive the appropriate care are captured in the numerator of the measure. Therefore, if rehabilitation services were not ordered and there is no reason documented for the services not being indicated, the physician will not meet the measure.

Developer response (TJC): Thank you for your comment and support of the Joint Commission's performance measure. The stroke core measures were developed in collaboration with the American Heart Association/American Stroke Association in 2003. Since that time, extensive efforts have been made to harmonize the measures, including STK-10: Assessed for Rehabilitation, with the American Medical Association Physician Consortium for Performance Improvement Stroke and Stroke Rehabilitation measures, The American Heart Association/American Stroke Association Get With the Guidelines Patient Management Tool, and the Centers for Disease Control and Prevention Paul Coverdell National Acute Stroke Registry. The Joint Commission will continue to work towards further harmonization with these organizations, and appreciates the feedback for future measure development.

Action Taken: Committee members agreed that measures should be harmonized to the extent possible, but recognized that measures specified for different levels of analysis (e.g., clinician vs. facility) may require different specifications. The Committee also recommended continued and aggressive efforts at harmonization and requested a progress update at the annual review.

Theme 3: Severity of stroke in risk-adjustment models

Description: We received 18 comments expressing the concern that an indicator of stroke severity (particularly, the value of the NIH Stroke Scale) is not included in the risk-adjustment models for stroke mortality and readmissions. These comments on stroke severity pertain to the following measures:

- 0467: Acute Stroke Mortality Rate (IQI 17) (AHRQ)
- 2026: Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following an acute ischemic stroke hospitalization (CMS/Yale)
- 2027: Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following an acute ischemic stroke hospitalization (CMS/Yale)

Most of these comments specifically cited a recent article by Fonarow[4] and colleagues. The conclusion of this article states:

> *Adding stroke severity assessed with the NIHSS score to a hospital 30-day mortality model based on claims data for Medicare beneficiaries with acute ischemic stroke is associated with substantial improvement in model discrimination and changes in mortality performance ranking for a considerable proportion of hospitals. These findings suggest that it may be critical to collect and include stroke severity for optimal hospital risk adjustment of 30-day mortality for Medicare beneficiaries with acute ischemic stroke.*

Responses from AHRQ: ***(****NOTE: The following text is taken from the responses to comments #2673 and #2707):*

AHRQ acknowledges that optimal risk-adjustment would include clinical markers of stoke severity, such as the NIH Stroke Scale, which may vary across hospitals in association with socioeconomic factors (Kleindorfer D, et al. Stroke 2012;43:2055-9). However, the recent paper by Fonarow et al. is likely to exaggerate the magnitude of this problem, for reasons described fully below. AHRQ will continue to work with the "Get With The Guidelines" team, the VA, and other interested entities that have linked clinical and administrative data to test and improve risk-adjustment modeling. The only currently available data set that links clinical and administrative data on a large population of stroke patients in the US is potentially useful , but it has historically suffered from two disadvantages: (1) it includes only Medicare fee-for-service beneficiaries aged 65 year s or older, and thus underrepresents younger and healthier patients, and patients from states with high managed care penetration; and (2) up until 4/1/12 it has it only had 9 diagnosis fields, as opposed to the 25 or more diagnosis fields available in most data sets used to estimate this AHRQ measure. It will be some time before a year of data will be available with expanded diagnosis codes. AHRQ will continue to collaborate with other interested parties to improve the data and take advantage of recently improved data that are available for testing and validation of risk-adjustment models. Analysis of a relatively small linked data set from the Veterans Health Administration suggests that the NIH Stroke Scale may not have as much impact on risk-standardized mortality rates in the VA setting as among

[4] Fonarow, et al. (July 18, 2012). Comparison of 30-day mortality models for profiling hospital performance in acute ischemic stroke with vs without adjustment for stroke severity. *JAMA, 308(3)*, 257-264.

hospitals participating in "Get With The Guidelines" (Keyhani S, et al. Circ Cardiovasc Qual Outcomes 2012; 5:508-13).

AHRQ acknowledges that optimal risk-adjustment would include clinical markers of stoke severity, such as the NIH Stroke Scale. AHRQ has carefully reviewed Fonarow's findings and held two meetings with his team. However, the applicability of their findings to the AHRQ measure is uncertain, because the risk-adjustment model that Fonarow et al. estimated using Medicare administrative data is markedly inferior to AHRQ's model using all-payer administrative data. Specifically, Fonarow et al. initially (J Am Heart Assoc 2012; 1:42-50) reported c statistics for ischemic stroke mortality of 0.71 (95% CI, 0.70-0.72), 0.82 (0.81-0.83), and 0.84 (0.84-0.85) using demographic and comorbidity information from administrative data, the NIH Stroke Scale alone, and both data sources, respectively. In a more recent paper (JAMA 2012; 308:257-64), the same authors built a more robust risk-adjustment model with 87 covariates derived from longitudinal claims data, and reported c statistics of 0.772 (0.769-0.776) and 0.864 (0.861-0.867) for models without and with the NIH Stroke Scale, respectively. By comparison, the AHRQ model, fully stratified for ischemic stroke, has a c statistic of 0.866, which is similar to that of Fonarow et al's combined model and much higher than their model based only on administrative data. The superiority of AHRQ's risk-adjustment model is not due to combining ischemic and hemorrhagic stroke, and it is also not due to adjustment for procedures performed after admission. Re-estimating the AHRQ ischemic stroke model without procedure-related APR DRGs, the c statistic dropped slightly from 0.866 to 0.858, and the weighted hospital-level correlation of adjusted rates between models with and without procedure-related APR DRGs was 0.977. The superiority of AHRQ's risk-adjustment model appears to be attributable to: (1) more complete data, with 25 or more available diagnosis fields instead of 9; (2) inclusion of a wider age spectrum, with adjustment for age; and (3) adjustment for markers of stroke severity that are present on admission and codable in ICD-9-CM, such as coma, other alteration of consciousness, convulsions, and hemiplegia. For example, among patients with ischemic stroke (APR DRG 045), we are able to stratify patients into four risk of mortality categories, with the following numbers of patients and death rates: Minor (referent) 112,533 0.0038 (0.38%); Moderate (OR=2.92) 160,536 0.0282 (2.82%); Major (OR=10.99) 53,457 0.0883 (8.83%); and Extreme (OR=98.15) 23,077 0.3916 (39.2%).

Responses from CMS/YALE: ***(****NOTE: The following text is taken from a detailed letter from Susannah Bernheim, MD, MHS of Yale, in response to the recent JAMA paper by Fonarow, et al.; the full letter was made available to the Committee and is posted on the NQF project page):*

A few important considerations limit the interpretability of the Fonarow paper with reference to our measure.

1. The first concern is the high percent of patients missing National Institutes of Health Stroke Scale (NIHSS) -- over half of the patients in the study do not have a measured NIHSS. The authors provide little information on the potential bias that could be introduced by the missing stroke scales –such as how the degree of missing NIHSS scores relates to median NIHSS for a hospital. If the hospitals with low percentage of completed NIHSS scores also have particularly high NIHSS median scores, this may account for the handful of hospitals whose profile changes with the addition of the score in the model.
2. Secondly, the measure described within the JAMA paper, though described as being modeled after our measure, differs in important respects from ours: 1) the cohort includes hemorrhagic patients, 2) the risk-adjustment includes different variables and is much less parsimonious (including 87 variables in total), and most importantly, 3) the measure does not risk adjust for transfers from Emergency Departments (ED).

The inclusion of a risk-adjustment variable indicating that a patient transferred into their index admission from an outside ED is important in our measure for two reasons. First, it will helps to account for the increased severity of cases at hospitals which frequently accept such patients (because the ED transfer patients are often more acutely ill). Second, it removes incentives to turn away transferred patients because the severity of such patients is accounted for within the model.

3. Finally, there are a number of issues about the modeling strategy used and the comparisons provided within the JAMA paper that affect our interpretation of the results of the paper.
    a. Unlike our medical record validation which directly compares hospital rates estimated by measures developed using two independent data sources (clinical versus administrative data), the JAMA paper compares a primary administrative model with a second model that includes one additional clinical predictor. Therefore, the administrative primary model is nested within this bigger model. As a mathematical certainty, adding an additional covariate will reduce overall variance at the hospital level and have the effect of pulling some outliers in as seen in the reclassification analysis.
    b. Furthermore, the re-classification analysis provided in the paper is based solely on the hospital random intercept, rather than comparing hospital results based on risk-standardized rates. This approach is, in essence, comparing hospitals' performance on one standard patient (in this case a patient with no comorbid disease). We find interpretation of such results is uncertain because they comparison of intercepts does not capture the full case mix of the hospitals as a risk-standardized rate would.
    c. Finally, and perhaps most importantly, the article does not allow evaluation of the degree of differences between the two models. The reported results in addition to reclassification refer to changes in ranking (based on hospital random intercepts and a standard patient) rather than actual rate estimates. All estimates have a degree of uncertainty. A small perturbation in the estimates may change ranking without meaningfully changing hospital estimates. The paper does not provide information about how similar the new estimates are to the original estimates, or whether the new estimates fall within the uncertainty of the original estimates. Nor does it present the correlation between the original model results and new results for hospitals.

In summary, although the stated goal of the Fonarow paper is to as assess the additional value of inclusion of stroke severity in our 30-day mortality measure, we find that the model used differs substantially from the measure we have put forward at NQF. The paper does not address critical questions about the impact of missing NIHSS on the majority of patients, nor do the final analyses fundamentally answer the question of whether hospital profiles differ meaningfully with inclusion of the severity score (for all the specific reasons described above).

Changes to measure specifications (#2026 and #2027)

As part of their measure harmonization efforts after the in-person meeting, CMS/Yale made one material change to measure #2026 and two material changes to measure #2027, as follows:

- Measure #2026: Hospital 30-day, all-cause, risk-standardized mortality rate (RSMR) following an acute ischemic stroke hospitalization
    - This measure now includes all-payer patients ages 18 and over (rather than Medicare FFS patient ages 65+ only)
- Measure #2027: Hospital 30-day, all-cause, risk-standardized readmission rate (RSRR) following an acute ischemic stroke hospitalization

- This measure now includes all-payer patients ages 18 and over (rather than Medicare FFS patient ages 65+ only)
- This measure now incorporates an algorithm for identifying and excluding planned readmissions from the measure
  - Originally, the measure excluded readmissions that were planned for procedures that are related to follow-up care after an ischemic stroke (e.g., carotid endarterectomy). The revised algorithm identifies commonly planned readmissions for all types of patients, not just those that are planned as follow-up post-stroke (e.g., maintenance chemotherapy, rehabilitation).
  - The new planned readmission algorithm harmonizes the stroke readmission measures with other CMS/Yale readmission measures.

The developer provided detailed reports describing the effects of these changes on the measures; these reports are available on the NQF project page. Because the revisions reflect material changes to these measures, an additional 15-day public and member comment period will open on September 12. ***Depending on the final recommendations from the Committee after this second comment period, Member voting on these two measures will be held.***

Action Taken: In their discussion of the comments regarding the inclusion of stroke severity in the risk-adjustment model for the three measures, the Committee considered the following:

- The need for adjustment for stroke severity
- The success (or not) in adjustment for severity using only administrative data
- The potential timing and feasibility of collecting the NIH stroke scale value
- The findings from the Fonarow paper that inclusion of the NIH stroke score resulted in changes in hospital rankings
- The trade-offs between a possibly imperfect measure against having no measure of readmissions at all
- The high discriminatory power of the AHRQ risk-adjustment model even though only administrative data are used
- The potential discriminatory ability of the risk-adjustment model for the CMS/Yale mortality measure if the NIH stroke scale also had also been included in the model
- The face validity of the risk-adjustment model for the CMS/Yale readmission measure, given that some covariates seem to be paradoxically protective against readmission
- The concern that the CMS/Yale mortality and readmission measures unfairly categorize tertiary care facilities that accept many transfer patients (e.g., stroke centers/safety net hospitals)

Because of the concern regarding inclusion of stroke severity in the risk-adjustment models for all three measures, as well as the material changes made to measures #2026 and #2027, the Committee agreed to re-vote on all three measures. Supplemental

materials addressing several of the above issues were provided to the Committee[5]. To inform their decisions, the Committee considered all comments and developer responses, supplemental materials, and the revised specifications for measures #2026 and #2027. Upon re-vote, the Committee:

- Recommended measure #0467 (yes-14, no-8)
- Could not reach consensus on measure #2026 (yes-11, no-11)
- Could not reach consensus on measure #2027 (yes-10, no-12)

Additional Areas for Measure Development

Several comments included suggestions for additional measure development, as follows:

New gap areas

- An outcome measure that is a combined endpoint of death and severe disability (i.e. Rankin Score 4-6), for a patient centered approach that would incorporate a patient's values on quality of life.
- A measure to document patient and family training and education in acute and post acute settings to reduce disability, burden of care, and primary and secondary prevention.

Suggested edit

- Change "Measures of **post hospital care** (prescriptions use at timed intervals after stroke, whether health problems are controlled over time, etc.)" to "Measures of **post-acute care and rehabilitation care** (prescriptions use at timed intervals after stroke, whether health problems are controlled over time, etc.)" (see page 13 of the project report).

Action Taken: After review by the Committee, the report was updated to include these suggestions.

NQF Member Voting

Information for electronic voting has been sent to NQF Member organization primary contacts. Accompanying comments must be submitted via the online voting tool.

Please note that voting concludes on September 26, 2012 at 6:00 pm ET – no exceptions

[5] See http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=71817 and http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=71815 .