

August 21, 2012

Suzanne C. Theberge, MPH  
Project Manager, Performance Measures  
National Quality Forum  
1030 15th Street, NW, Suite 800  
Washington, DC 20005

Dear Ms. Theberge,

During the public comment period for the CMS 30-day stroke measure, a number of commenters stated concerns that an administrative claims model that does not directly risk adjust for stroke severity cannot adequately profile hospital quality based on 30-day mortality rates. Specifically, concerns were expressed that hospitals could be misclassified due to the lack of including stroke severity as a risk-adjustment variable, or that the measure will discourage hospitals from accepting severe stroke patients. These commenters referenced a recent paper in JAMA (Fonarow et al).

We are using this letter to respond to the JAMA study below. To place this paper in context we remind the Steering Committee that this analysis is one of three separate analyses designed to examine the importance of medical record data for the stroke mortality measure. The first two -- our own chart validation and a recent paper within the VA system (Keyhani et al) -- both showed a high degree of correlation between hospital-level results on the claims model and a measure that included stroke severity. The latest contribution, the Fonarow paper, is discussed below.

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 help 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 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).

The high proportion of patients without an NIHSS also highlights the lack of a feasible alternative medical record model for profiling hospitals. The finding that among hospitals enrolled in a registry the NIHSS is only available on 45% of patients suggests that there are many hurdles to producing a registry or medical record-abstracted measure that adequately captures a hospital's ischemic stroke patients.

We have brought forth to NQF a feasible measure of 30-day stroke mortality. The measure has a good performance (c-statistic 0.74) and displays good variation. The measure results strongly correlate with those of a medical-record based model. The measure includes a risk-adjustment variable for ED transfer patients so it will not discourage hospitals from accepting this high-risk group of patients. Also, in response to concerns that hospitals that care for the most complex patients would not be fairly characterized by the measure, during the measure development, we examined hospitals' results by various hospital characteristics including teaching status and stroke center certification. We find no evidence that teaching hospitals or stroke centers perform systematically worse on the measure.

We appreciate the concerns raised in public comment and have responded to them individually as well, but we wanted to provide the Steering Committee with a cohesive discussion of why we do not believe the recent paper by Fonarow should change the Steering Committee's assessment of our measure.

Sincerely,

Susannah Bernheim, MD, MHS  
Director, Quality Measurement Programs

cc: CMS

### **References**

Fonarow GC, Pan W, Saver JL, et al. Comparison of 30-Day Mortality Models for Profiling Hospital Performance in Acute Ischemic Stroke With vs Without Adjustment for Stroke Severity. *JAMA*. 2012. 308(3): 257-64.

Keyhani S, Cheng E, Arling G, et al. Does the Inclusion of Stroke Severity in a 30-Day Mortality Model Change Standardized Mortality Rates at Veterans Affairs Hospitals? *Circ Cardiovasc Qual Outcomes*. 2012. 5(4):508-13.