April 5, 2016

Helen Darling, MA
Interim President and CEO
National Quality Forum
1030 15th St., Suite 800
Washington, DC 20005

RE: Appeal of NQF #2431, 2436 and 2579 from the Cost and Resource Use Measure Endorsement Project

Dear Ms. Darling:

The undersigned associations representing the nation's hospitals and health care systems write to appeal the National Quality Forum's (NQF) endorsement of the acute myocardial infarction (NQF # 2431), heart failure (NQF #2436) and pneumonia (NQF #2579) 30-day episode-of-care payment measures. These three measures are among the first measures to be reviewed under the NQF's "Trial Period" for sociodemographic status (SDS) adjustment, which permits the consideration and endorsement of measures that use SDS adjustment.

We appreciate that NQF initiated the SDS Trial Period, as we have long urged NQF, Medicare and other stakeholders to ensure outcome measures are appropriately adjusted for factors beyond the control of providers, including SDS. In addition, hospitals continue to believe that well-designed measures of cost and resource use are important tools for facilitating improvements in the value of care – that is, delivering the same or better outcomes at lower cost.

However, we have several concerns regarding the application of the consensus development process (CDP) on these measures including:

- **Inaccurate representation of the recommendations of NQF's Expert Panel on Risk Adjustment and SDS in the measure evaluation criteria;**

- **A flawed empirical analysis used to test whether cost and resource use measures should be SDS adjusted;**

- **Insufficient criteria and materials provided by NQF staff to the Standing Committee and measure developers on what should be provided for SDS variable selection and testing to guide the evaluation; and**

- **Insufficient resolution of all of the conditions set by the NQF Board for endorsement in 2015.**

For these reasons, we recommend that NQF:

1. Remove endorsement on these measures at this time, and work with the developer to address the ongoing concerns around the scientific acceptability of the measures, including additional analyses on SDS adjustment prior to reconsideration;

2. Ensure NQF's criteria and processes for the SDS Trial Period are clear, consistent with the original intent of the expert panel, transparent to all stakeholders, and approved by the CSAC prior to further implementation of the SDS criteria in NQF projects; and

3. Reexamine the use of "endorsement with conditions" on any measures moving forward, including further discussion with the NQF membership and public.

We provide additional detail on our concerns and recommendations below.

SDS TRIAL PERIOD IMPLEMENTATION CONCERNS

We do not believe that the intent of the SDS Expert Panel's recommendations is accurately represented in the measure evaluation criteria and associated SDS Trial Period guidance. As a result, the evaluation by the Cost and Resource Use Standing Committee conflicted with the original intent of the trial period. More importantly, one of the few criteria clearly articulated in both the Expert Panel report and in the evaluation criteria is the expectation that there would be a conceptual basis for believing that the SDS factor(s) being tested represents a legitimate reason for variation in the results of what is being be measured (in this case, cost per episode). We do not believe that the developer provided adequate justification of the conceptual relationships each of its chosen variables had with the three measures. As a result, the empirical model used to evaluate whether these 3 measures should be SDS-adjusted is neither robust nor well-specified enough to warrant the conclusions drawn by the measure developers.

Our specific concerns are as follows:

1. **The inclusion of race in the analysis of these cost and resource use measures is not justified by the material presented to the Standing Committee and is inconsistent with the original SDS Adjustment Expert Panel recommendations**. In the analysis submitted by the measure developers, individuals' race was coded as either "Black" or "Not Black". By aggregating majority-Whites and groups who, like Black Americans, suffer disproportionately from inequities in health care (e.g. Latinos, Native Americans, etc.), differences between the "Black" and "Not Black" groups will necessarily be attenuated, masking important disparities evident in the literature. Racial groups should not be "collapsed" unless there is a valid conceptual reason to do so. We believe the measure developer failed to adequately articulate a conceptual basis for the use of race as a variable, and further, it did not adequately

explain why it was appropriate to collapse the groups. This falls short of the recommendation of the expert panel that developers articulate a clear conceptual link between adjustment variables and outcomes.

Moreover, the SDS Expert Panel expressed significant concerns about the general conceptual basis for using race as a proxy for SDS. Indeed, the panel's final report suggests that race and ethnicity can be "…confounded by [SDS]. That is, income, education, and related factors (including language and insurance) represent key contributors to racial and ethnic disparities in healthcare." Since the developer's analysis included only one other SDS adjustment variable – dual eligibility – the relationship between race and the outcomes of interest are likely to remain confounded, further masking any conceptual relationship.

2. The Cost and Resource Use Standing Committee urged the measure developer to explicitly include in their conceptual model community and environmental factors, and to separate patient- from community-level resources. However, the empirical model used to test for SDS-adjustment only contains patient-level factors (race and dual-eligibility) and ignores completely the influence of community-defined SDS variables on the outcomes of interest. This is a significant flaw as multilevel analyses show distinct and direct effects of both individual- and community-level drivers on health and health care outcomes.

3. Throughout this review, the Standing Committee explicitly requested that additional variables be included in the analyses such as the expansion of the zip codes from 5-digits to 9-digits and the addition of Low Income Status along with the Medicaid Enrollment/Dual Status. In addition, four variables were initially identified in the conceptual model. Yet, one could argue that only one variable was adequately addressed in the empirical analyses and the others were addressed through the use of a proxy.

   While we recognize the challenge of leveraging various data sources, the absence of data is not sufficient to justify the use of proxies or inadequate data. Indeed, prior NQF committees have recommended against the endorsement of several measures (e.g., some eMeasures) for which a lack of available data directly impacts measure validity. Yet, despite similar concerns with these three measures, endorsement was recommended. Moreover, the developer has not adequately demonstrated that Medicaid Enrollment/Dual Status could be considered a valid proxy for the variables identified in the conceptual analysis.

We are therefore concerned that the conceptual model was insufficient, and the empirical analyses provided by the developer were not fully responsive to the Committee's requests. **Given the mismatch between the conceptual model and its empirical operationalization, and the flawed application of 'race' by the measure developers, the NQF should remove endorsement from these measures, and work with the developer to identify when the measures can be reevaluated.** The reevaluation should

address the ongoing concerns around the scientific acceptability of the measures, and likely would include additional analyses on SDS adjustment.

We also are concerned that NQF provided insufficient criteria and materials on the selection and testing of SDS variables to guide the Standing Committee's evaluation. NQF released updated measure evaluation criteria that went into effect in April 2015. The updated criteria included modifications to the Scientific Acceptability subcriterion, and specifically to the language around risk adjustment and consideration of SDS variables. These modifications included additional guidance to measure developers and Standing Committees on what must be provided and evaluated during the SDS Trial Period. As discussed above the evaluation criteria used by the Standing Committee do not accurately represent the recommendations of the Expert Panel.

While informational items on the SDS Trial Period were provided to the CSAC in April and August of 2015, we are unable to find documentation of any CSAC approvals of these changes in the measure evaluation criteria nor in the associated guidance on the NQF web site. The lack of explicit approval is contrary to the process followed when other modifications were made to the criteria. The lack of oversight and approval by the CSAC is troubling since NQF members and users of measures rely on CSAC for a thorough and complete review of measures, including risk adjustment models. It also is problematic given the degree of interest and support by the NQF membership for recommendations of the Expert Panel, and the desire of many stakeholders to sufficiently address this ongoing measure methodology concern. **Therefore, we ask that the criteria and guidance on the SDS Trial Period be revised to address the current inaccuracies and to further clarify what is expected of measure developers. We also urge that the criteria be reviewed and approved by the CSAC prior to further implementation.**

INSUFFICIENT RESOLUTION OF THE CONDITIONS OF ENDORSEMENT

When initially endorsed, these three measures were endorsed with conditions by the NQF Board of Directors (BOD) Executive Committee to specifically address the concerns of NQF members. The conditions placed on the measures according to the February 2015 final Technical Report for the Cost and Resource Use were:

- <u>One-year look-back assessment of unintended consequences:</u> NQF staff will work with the Cost and Resource Use Standing Committee and CMS to determine a plan for assessing potential unintended consequences of these measures in use. The evaluation of unintended consequences will begin in approximately 1 year, and possible changes to the measures based on these data will be discussed at that time.

- <u>Consideration for the SDS trial period:</u> The Cost and Resource Use Standing Committee will consider whether the measure should be included in the NQF trial

period for consideration of sociodemographic status adjustment.

- Attribution: NQF will consider opportunities to address the attribution issue.

Based on information provided by NQF during the review of these measures and posted to the NQF web site, it appears that only the second condition has been addressed. No assessment of the unintended consequences of these measures was included in the review. To our knowledge, a plan to assess the potential unintended consequences has not been released and could not be found on the NQF web site. In fact, materials to the CSAC and BOD Executive Committee on this ad hoc review no longer list this assessment as one of the conditions. Furthermore, the third condition (attribution) raised during the previous review of these measures does not seem to have been addressed, and without explanation appears to have been removed from the list of conditions. An NQF staff memorandum to the Cost and Resource Use Standing Committee and measure developers dated May 19, 2015 makes no mention of the attribution condition.

These omissions demonstrate that the conditions placed by the NQF BOD Executive Committee in February 2015 have not been adequately addressed and the CDP has not been followed.

We also have concerns about the lack of information provided via the NQF web site to identify which measures carry what endorsement. Measures that are endorsed with conditions do not carry this label on the NQF measure search engine (QPS) nor are the conditions included in any materials or measure information with the exception of the final technical report. QPS also does not indicate that any of these three measures were included in the SDS Trial Period.

It was our understanding that a permanent endorsement category was not being created at the time of the Cost and Resource Use measures' endorsement; yet, other measures have since been endorsed with conditions. If it is NQF's intent to expand the endorsement categories, then member input should be solicited and the CDP should be revised to clearly articulate what constitutes a condition, how and when the condition could be used, how these conditions will be displayed and communicated to members and the public, and what the NQF's processes are to ensure that these conditions are met and reviewed in a timely manner.

For all the reasons listed above, we appeal of the endorsement of the Myocardial Infarction (#2431), Health Failure (#2436) and Pneumonia (#2579) Cost Resource Use Measures, and urge NQF to develop and publish a transparent plan addressing the concerns listed above prior to further Committee review.

Thank you for your consideration of these important issues. If you have further questions, please contact Nancy Foster at nfoster@aha.org, Jayne Hart Chambers at jchambers@fah.org, Ivy Baer at ibaer@aamc.org, and Beth Feldpush at bfeldpush@essentialhospitals.org.

Sincerely,

American Hospital Association
Federation of American Hospitals
Association of American Medical Colleges
America's Essential Hospitals

cc: Helen Burstin, MD, MPH
Marcia Wilson, PhD, MBA