TO: Consensus Standards Approval Committee (CSAC)

FR: NQF Staff

RE: Appeal of Measures for the Cost and Resource Use Standing Committee Ad Hoc Review of the Conceptual and Empirical Analysis of Sociodemographic Variables and Payment Outcomes

DA: August 2, 2016

In accordance with the NQF Consensus Development Process (CDP), the measures recommended for continuing endorsement by the Cost and Resource Use Standing Committee were released for a 30-day appeals period, which closed on April 5, 2016. NQF received two letters of appeal, one from the American Association (AMA), and one from four hospital associations, the American Hospital Association (AHA), the Federation of American Hospitals (FAH), the Association of American Medical Colleges (AAMC), and America's Essential Hospitals (AEH). The appellants are asking NQF to remove endorsement of three measures:

- #2431: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI) (CMS/Yale)
- #2436: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Heart Failure (HF) (CMS/Yale)
- #2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care pneumonia (CMS/Yale)

The following documents are appended to this memo:

1. Appendix A – Appeal Letter for from the AMA
2. Appendix B – Appeal Letter for from the AHA, FAH, AAMC, and AEH
3. Appendix C – NQF Response
4. Appendix D – CMS Response

CSAC ACTION REQUIRED

The CSAC will review the letter of appeal and this memo in consideration of the appeal. The CSAC will determine whether to uphold NQF endorsement of NQF #2431: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI) (CMS/Yale), NQF #2436: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Heart Failure (HF) (CMS/Yale), and #2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care pneumonia (CMS/Yale) .

SUMMARY OF ISSUES RAISED IN THE APPEAL

The National Quality Forum (NQF) has received appeals of its endorsement of the acute myocardial infarction (AMI) (NQF # 2431), heart failure (HF) (NQF #2436) and pneumonia (NQF #2579) 30-day episode-of-care payment measures. The Cost and Resource Use Standing Committee has deliberated on the scientific properties of these measures extensively and had made recommendations to CSAC and the Board prior to the start of the trial period, and, upon request from the Board, re-examined the measures using the sociodemographic (SDS) trial period guidance. The Cost and Resource Use Standing Committee

reviewed analyses from the developer and recommended the measures continue to be endorsed without the inclusion of SDS factors in their risk adjustment models. The decision was approved by the CSAC and ratified by the Executive Committee of the NQF Board of Directors. The appeals raise concerns regarding the application of the CDP on these measures, specifically:

- A flawed empirical analysis used to test whether cost and resource use measures should be SDS adjusted;
- Insufficient resolution of all of the conditions set by the NQF Board for endorsement in 2015.
- Implementation of the SDS trial period

The appellants raise a number of concerns related to the review of these measures for SDS adjustment. First, the appellants raise concerns about the testing of race as a possible factor for inclusion in the risk adjustment model of the measures. The appellants also note that the Cost and Resource Use Standing Committee urged the measure developer to explore in their conceptual model community and environmental factors, and to separate patient and community-level resources. The appellants feel that the developer did not sufficiently explore the variables included in the conceptual model. Additionally, the appellants raise concerns that the developers did not perform the analyses requested by the Standing Committee. In particular, the developer did not expand the analyses to the nine-digit zip code level and did not include Low Income Status along with the Medicaid enrollment/dual status variable

The appellants also raise concerns that the additional two conditions for endorsement have not been adequately met. First, the appellants raise concerns about the one-year look back assessment of unintended consequences of these measures in use. Secondly, the appellants raise concerns about the need to consider issues of attribution.

Finally the appellants raise concerns about the implementation of the trial period. Specifically the appellants have concerns about the guidance provided to Standing Committees on the selection and testing of SDS variables and CSAC approval of the revised measure evaluation criteria.

ADDITIONAL CONSEUS BUILDING

In June 2016, NQF convened the appellants, CMS, Yale/CORE, the CSAC co-chairs, and one of the chairs of the Cost and Resource Use Standing Committee. The goal of this call was to foster a dialogue between the affected parties and to lay out potential options as the appeal is considered. During the call the appellants asked for clarification for the conceptual basis for the expected effect of adjustment. Yale/CORE agreed to provide a clearer conceptual analysis and to perform additional empirical analyses to examine the impact of SES factors at the nine-digit zip code level to address the concerns raised by the appellants.

NEW EMPIRICAL ANALYSES

CMS/Yale CORE has submitted new analyses using nine-digit ZIP code data included in the CMS/Yale CORE response memo attached to this memo (Appendix C).

STANDING COMMITTEE REVIEW

In light of the new information provided by the developer and outstanding questions of attribution and unintended consequences the Cost and Resource Use Standing Committee met via webinar on July 28, 2016 to provide additional input to CSAC as they consider these appeals.

During the meeting, the Committee reviewed the new analyses provided by CMS/Yale CORE using 9-digit ZIP code data. The developer stated that using this data they found a slightly lower 30-day total payment for AMI, heart failure, and pneumonia for low SES patients. The developer suggested that this could be due to underutilization of services. The Committee noted that the difference was statistically significant but did not substantially impact hospital distribution.

Additionally, the Committee reviewed the conceptual model and noted challenges related to establishing a concrete link between socioeconomic status and risk-adjusted episode spending as well as limited SES factors available in current data sets. The Committee also noted that some costs related to care needed to support lower socioeconomic status patients may not be reimbursed and therefore would not show up in the results of these measures. Finally, the Committee reiterated the need to examine the impact of community factors in the future.

Given the limited time for the meeting, the Committee is providing additional feedback via survey. That feedback will be summarized and provided to CSAC as an addendum to this memo.

NEXT STEPS

CSAC will review the Standing Committee's input and consider the appeal during their August 9 meeting.

Appendix A

April 5, 2016

The American Medical Association (AMA) is writing to appeal the endorsement of the following three measures:

- #2431: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI) (CMS/Yale)
- #2436: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Heart Failure (HF) (CMS/Yale)
- #2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care for pneumonia (CMS/Yale)

The AMA examined the 2016 review of these cost and resource use measures and we believe that the National Quality Forum's (NQF) Consensus Development Process (CDP) was not followed. Specifically, we were unable to identify if and when the Consensus Standards Approval Committee (CSAC) approved the updates to the measure evaluation criteria and guidance prior to implementation in CDP projects. We believe that the revised criteria and guidance provided around risk adjustment and the inclusion of sociodemographic variables does not reflect the original intent of the expert panel and these differences impact whether the analyses completed by developers should be considered responsive to what was put forward by that panel. Our second concern relates to the omission of two of the three conditions placed on the three measures at the time of endorsement in February 2015 in this 2016 review.

As a result of these deviations from the CDP, we believe that the integrity of the NQF process has been compromised with respect to these measures. As such, we would ask that NQF:

- Remove endorsement on these measures until such time that all of the conditions can be adequately met and NQF should work with the developer to identify when the measures can be reevaluated;
- Work with developers to ensure that the measures within the SDS Trial Period are consistent with the recommendations from the Expert Panel on Risk Adjustment and Socioeconomic Status; and
- Reconsider the use of "Endorsement with Conditions" on any measures moving forward.

Our specific concerns and rationales for these requests are outlined below.

<u>Omission of two of the three conditions for endorsement during this current review</u>

Until recently, NQF maintained a simple structure regarding the types of endorsement available. Specifically, endorsement was limited to "NQF-endorsed" and for a few years, "time-limited endorsement" was also an available alternative. Previously, endorsement with any caveats or limitations was not considered in the CDP outside of those measures that were time-limited since all measures must meet the minimum set of measure endorsement criteria. These three cost and resource use measures are some of the first measures to our knowledge that have been endorsed with conditions by the NQF Board of Directors (BOD) Executive Committee. This new type of endorsement could be considered a deviation of the CDP as neither the measure

endorsement criteria, guidance to Standing Committees and the Consensus Standards Approval Committee (CSAC) and other documents provide information on what endorsement with conditions means, how measures can achieve this type of endorsement, and what the NQF's processes are to ensure that these conditions are met and reviewed in a timely manner.

We are concerned that due to the lack of clear processes and procedures, there is great potential for NQF to inadvertently omit or inadequately address these conditions around endorsement. We believe that these three measures serve as a good example of this concern.

These three measures according to the final Technical Report for the Cost and Resource Use project released in February 2015 were endorsed with the following conditions:

- One-year look-back assessment of unintended consequences: 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.
- Consideration for the SDS trial period: 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 what was included in this review of these measures, only the second condition has been addressed. Assessment of any unintended consequences of these measures was not included in the review and 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 review no longer list this assessment as one of the conditions. We would also note that the third condition around attribution would not be considered actionable and responsive to the concerns raised by the NQF membership and public and it appears to have been removed completely from the list of conditions by NQF staff in a memo sent to the Cost and Resource Use Standing Committee and measure developer on May 19, 2015.

These omissions demonstrate that the conditions placed by the NQF BOD Executive Committee in February 2015 have not been adequately addressed nor has the CDP been followed. These omissions also are examples of our concerns about the lack of transparency via the NQF web site. 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 the three measures were included in the SDS Trial Period. This new type of endorsement and the underlying conditions are not sufficiently clear to the NQF membership and public and could have unintended consequences for those seeking to implement NQF-endorsed measures who remain uninformed and unaware of serious concerns around these measures.

Lack of oversight and approval of current measure evaluation criteria by the CSAC

NQF released updated measure evaluation criteria that went into effect in April 2015. Several modifications were made to the criteria including updates to the Scientific Acceptability

subcriterion and specifically to the language around risk adjustment. These modifications included additional guidance to measure developers and Standing Committees on what must be provided and evaluated during the SDS Trial Period and were based on the recommendations made by the Expert Panel on Risk Adjustment and Sociodemographic Status.

While informational items on the SDS Trial Period were provided to the CSAC in April and August of that same year, we were unable to find documentation of any CSAC approvals of these changes and the associated guidance on the NQF web site, which is contrary to the process followed when other modifications were made to the criteria. This lack of oversight and approval by the CSAC is troubling given the degree of interest and support by the NQF membership on the inclusion of these variables in risk adjustment models, the support of the membership of the recommendations of the Expert Panel, and the desire of many stakeholders to sufficiently address this ongoing measure methodology concern.

In addition, we do not believe that the intent of the Expert Panel's recommendations is adequately represented in the SDS Trial Period guidance. In the final report, the Expert Panel stated that of race/ethnicity should not be considered as acceptable proxies for socioeconomic (SES) because SES often confounds their effects. We would expect that this intent and explicit statements around what should be considered acceptable variables or proxies would be included in the SDS Trial Period guidance, but it is not. In the case before us, the developer of these three measures included race as one of the SDS risk variables. This inclusion is not consistent with the original Expert Panel recommendations and raises significant concerns that the important SDS variables for risk adjustment of these measures were not sufficiently identified and tested.

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.

We are therefore concerned that the empirical analyses provided by the developer were not fully responsive to the Committee's requests and that the use of proxies should not be considered adequate based on the conceptual model provided. For example, as stated by the Committee, 5-digit zip codes do not provide sufficient information around SDS factors. While we understand that access to the 9-digit zip code data is not yet available to the developer, the absence of data should not justify the use of proxies or inadequate data. Other measures such as eMeasures for which validity is directly impacted by availability of the data, have led committees to not recommend endorsement; yet, despite there having been a similar concern raised with these three measures, endorsement continues to be recommended. Also, we do not believe that the developer adequately demonstrated that Medicaid Enrollment/Dual Status could be considered a valid proxy for the variables identified in the conceptual analysis.

While recognizing the challenge with leveraging data sources, the AMA had the expectation that measure developers would be required to obtain new data sources to account for SDS variables. Therefore, we are disappointed that this did not occur and calls into question the effectiveness of the SDS Trial Period. We stand ready to work with the NQF and other relevant

health care stakeholder groups to improve the current quality measure endorsement processes, specifically the SDS Trial Period. Please feel free to contact Koryn Rubin, Assistant Director of Federal Affairs, at koryn.rubin@ama-assn.org or (202) 789-7408 for more information.

Thank you for considering our appeal and concerns.

Appendix B

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

Appendix C

TO: Koryn Rubin, American Medical Association, Nancy Foster, American Hospital Association, Jayne Hart Chambers, Federation of American Hospitals, Ivy Baer, Association of American Medical Colleges, Beth Feldpush, America's Essential Hospitals

FR: Helen Burstin, Marcia Wilson, Elisa Munthali, National Quality Forum

RE: Appeal of NQF #2431: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI) (CMS/Yale); #2436: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Heart Failure (HF) (CMS/Yale); #2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care for pneumonia (CMS/Yale)

DA: May 10, 2016

The National Quality Forum (NQF) has received two appeals of its endorsement of the acute myocardial infarction (AMI) (NQF # 2431), heart failure (HF) (NQF #2436) and pneumonia (NQF #2579) 30-day episode-of-care payment measures. NQF takes the concerns of the appellants of these three cost and resource use measures seriously. The Cost and Resource Use Standing Committee has deliberated on the scientific properties of these measures extensively and had made recommendations to CSAC and the Board prior to the start of the trial period, and, upon request from the Board, re-examined the measures using the SDS trial period guidance. NQF recognizes it has an important role important role to build consensus and work though challenging scientific and policy issues such as these. To begin to work through these challenges, NQF will convene the appellants and CMS/Yale on May 10, 2016 to discuss the issues raised. In preparation for that meeting, NQF has outlined its reponses to the concerns raised by the appeallents.

SDS Trial Period Concerns

- The appellants raise concerns about the testing of race as a possible factor for inclusion in the risk adjustment model of the measures. In particular the appellants raise two concerns about the developer's use of race:
    - The appellants believe the developer did not provide an adequate conceptual basis for the use of race as a variable and did not explain why it was appropriate to aggregate individuals into "black or non-black;"
    - The developer tested only one other SDS adjustment variable (dual eligibility). The appellants note the SDS Expert Panel stated that race should not be used as a proxy for SES; rather race is confounded by SES. The appellants believe the developers did not test enough variables to unmask any conceptual relationship and that the relationship between race and the measures' outcomes are likely to remain confounded.
- NQF Response:
    - Guidance was provided to the measure developers and the Standing Committee based on the recommendations of the SDS expert panel that race should not be used as proxy for SDS and should not be used in adjustment unless there is a clear conceptual rationale.

- During its May 21, 2015 webinar to review the developer's conceptual analysis, the Cost and Resource Use Standing Committee raised concerns about the inclusion of race as a variable. The Committee believed that further literature review was needed to determine the within and between effects of race on hospital performance. Some members strongly suggested that between and within hospital differences should be a lens through which this information should be analyzed.
- In a memo dated October 5, 2015, the developer summarizes the results of their expanded literature search. The developer found that most studies uses race and their independent variable with less attention to income or other measures of poverty. The developer concluded that the literature demonstrates that both within and between hospital differences in outcomes among racial/ethnic groups can be partially explained by the use of lower quality hospitals by minorities.
- During the May webinar, the Standing Committee raised similar concerns to the appellants about the aggregation of racial categories. However, in the October 5 memo, the developer confirms that while they considered creating categorizations of black/white/other or black/white/other/Hispanic, data from CMS suggests that black and white race are the only categories with both high sensitivity and specificity in the Beneficiary Race Code variable.
- Race was not included as a variable in the final risk adjustment model; rather it was only explored by the developer.
- NQF agrees with the appellants that race should not be used as proxy for SES. This guidance was explicitly stated in the SDS Expert Panel's final report. The Disparities Standing Committee is currently examining this issue and is in the process of providing additional guidance to measure developers and NQF Standing Committees about the use of race as a variable in risk adjustment models.
- Additionally, the measure developer could clarify their rationale for testing race as a possible variable for inclusion during the May 10, 2016 meeting.

- The appellants note that the Cost and Resource Use Standing Committee urged the measure developer to explore in their conceptual model community and environmental factors, and to separate patient- from community-level resources.
- NQF Response:
  - During its October 27, 2015 webinar to review the developer's empirical analysis, the Committee had extensive discussion about the inclusion of community-level factors into the risk-adjustment model given the inclusion of a 30-day post discharge period in the episode. The Committee acknowledged that for some of the post-hospitalization services, the community context is a critical variable and that these factors may or may not be fully captured by the patient-level SDS adjustment.
  - The developers expressed interest in potentially considering these factors in the model, but sought Committee input and recommendations on how to approach this.
  - This issue could be further explored between the appellants and the developer during the May 10, 2016 meeting.

- The developer did not sufficiently explore the variables included in the conceptual model. Additionally, the appellants raise concerns that the developers did not perform the analyses requested by the Standing Committee. In particular, the developer did not expand the analyses to the nine-digit zip code level and did not include Low Income Status along with the Medicaid enrollment/dual status variable.
- NQF Response:
    - The developer expanded the conceptual model in response to the Cost and Resource Use Standing Committee's concerns. The CMS/Yale team revised the model to broaden the scope of community-level factors included in the model. In doing so, they updated the pre-admission and post discharge phases of the conceptual model to capture the many patient and community factors that reflect differential impact of SDS on episode of care payments. The developer also revised the model to reflect "patient factors" rather than "patient behaviors." Patient factors included variables such as using services provided and adherence to care plan. Community factors included variables such as lack of community services and lack of social supports/caregiver. Finally, the model also was reoriented to capture the potential pathways by which low SDS may impact the care provided to patients. Details of the final memo can be found in the developer's October 5 memo.
    - The Cost and Resource Use Standing Committee noted significant gaps in the literature specific to the impact of SDS on cost, utilization, or payment outcomes. Specifically, the Committee questioned whether the use of standardized payments based on diagnosis-related groups may mitigate the relationship between SDS and costs.
    - In the October 5 memo, the developer clarified they chose to use the Dual Status variable because it best reflected those with the lowest income.
    - The developer and the appellants could discuss the use of additional patient and community level variables at the May 10 meeting as well as the possibility of exploring these variables at the nine-digit zip code level.
- The appellants raise concerns about the implementation of the trial period. Specifically the appellants have concerns about:
    - The guidance provided to Standing Committees on the selection and testing of SDS variables.
    - Consensus Standards Approval Committee (CSAC) approval of the revised measure evaluation criteria.
- NQF Response:
    - NQF recognizes that the SDS trial period marks a significant change the Consensus Development Process. NQF staff have worked to provide guidance to measure developers, Standing Committees, and the public to educate them on the input of the SDS expert panel and on how measures should be reviewed during the trial period. Web meetings have been held with measure developers and Standing Committees are briefed on the changes during their orientation and Question and Answer calls. NQF will

work to improve the clarity and breadth of the educational materials and opportunities provided to developers, Standing Committees, and the public.

- However, NQF maintains a non-prescriptive approach to the selection and testing of variables included in risk adjustment models. NQF does not require that certain variables be tested and does not set requirements around the inclusion of any specific variables. Similarly NQF does not set certain "cut-points" for the statistical testing of a risk adjustment model. The evaluation of the model is the left to the Standing Committee reviewing the measure. This approach applies to both clinical and SDS variables.
- The Disparities Standing Committee is charged with evaluating the trial period. Results to date were presented to the Disparities Standing Committee during their April 26, 2016 webinar. The Committee is currently drafting additional guidance based on the findings and challenges of the trial period to date. This guidance will be provided to the Standing Committees, developers, and public by early summer 2016.
- Updates to the measure evaluation criteria were made as part of the CSAC's approval of the SDS Expert Panel's recommendations during its July 9-10, 2014 meeting. Specifically, the Expert Panel's Recommendation 4 revised the criteria. These recommendations passed with the consensus of the CSAC.

Insufficient Resolution of the Conditions of Endorsement

- The appellants raise concerns that the three conditions for endorsement have not been adequately met. First, the appellants raise concerns about the one-year look back assessment of unintended consequences of these measures in use.
- NQF Response:
  - There is general agreement that these measures need to be monitored as they are endorsed and implemented into federal quality initiative programs. These measures have been recently adopted for the Hospital Inpatient Quality Reporting program for FY 2016 (AMI) and FY 2017 (HF and pneumonia). NQF will need implementation data from CMS as experience with the measures has been demonstrated. The May 10 meeting will allow the appellants and CMS the chance to opportunities to develop a path forward on the look back period issue
- Secondly, the appellants raise concerns about the need to consider issues of attribution.
- NQF Response:
  - With funding from HHS, NQF has launched a project on attribution. The expert panel guiding this work includes representation from both hosptials and the American Medical Association to ensure attribution issues such as the ones illustrated by these measures are addressed. As part of this project, NQF will commission an environmental scan identify different attribution models and examine their strengths and weaknesses. The environmental scan will be used as a foundation for establishing a set of principles and recommendations for applying the models within a complex healthcare delivery system. Throughout this project, NQF will solicit input from NQF's multi-stakeholder audience,

including NQF membership and public stakeholders at key points throughout the project.

Next Steps

The appellants have raised a number of important concerns around the continuing endorsement of these measures. NQF recognizes that the concerns raised around the analytic approach used by the developer will require additional work with the appellants, the developer/CMS, and the leadership of the Standing Committee/CSAC to discuss an agreeable path forward. However, there may be legitimate challenges to leveraging existing data to examine the variables requested by the appellants, but additional work can be undertaken to identify a path forward that explores the issue further. NQF will continue to convene the Disparities Standing Committee to evaluate the trial period and will work with that group to address the concerns and challenges that arise. In the short term, the Disparities Standing Committee will develop additional guidance to the Standing Committees, developers, and the public about the use and testing of SDS variables. Additionally, NQF staff will work to continue to educate Standing Committees, developers, and the public about the SDS trial period and how measures should be evaluated during the trial.

NQF has an important leadership role in building consensus on these issues and will work with the affected parties to determine an agreeable path forward that respects the limitations that may exist in the field, in terms of data and variables, but is also responsive to concerns raised by the appellants.

DEPARTMENT OF HEALTH & HUMAN SERVICES
Centers for Medicare & Medicaid Services
7500 Security Boulevard, Mail Stop S2-26-12
Baltimore, Maryland 21244-1850

CENTERS FOR MEDICARE & MEDICAID SERVICES
CENTER FOR MEDICAID & CHIP SERVICES

Memorandum

DATE: **Monday, May 16, 2016**

TO: **The National Quality Forum (NQF)**

FROM: **Lein Han, PhD, Contracting Officer Representative**
Division of Quality Measurement (DQM)
The Centers for Medicare & Medicaid Services (CMS)

Kate Goodrich, MD, MHS, Director
Center for Clinical Standards and Quality
The Centers for Medicare & Medicaid Services (CMS)

SUBJECT: **CMS Response to Appeal of Acute Myocardial Infarction (NQF # 2431), Heart Failure (NQF #2436) and Pneumonia (NQF #2579) 30-Day Episode-Of-Care Payment Measures**

Background

On February 18, 2016, the National Quality Forum's (NQF) Board of Directors ratified NQF #2431: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI), NQF #2436: Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for heart failure (HF), and Hospital-level, risk-standardized payment associated with a 30-day episode of care for pneumonia (PN) for continued endorsement, followed by a 30-day appeals period. We received two letters of appeal on the February 18, 2016 endorsement decision. Several stakeholders, including the American Hospital Association (AHA), the Federation of American Hospitals (FAH), the Association of American Medical Colleges (AAMC), the America's Essential Hospitals (AEH), and the American Medical Association (AMA), offered comments addressing the following: use of race variable, consideration of community and environmental factors, and use of additional patient-level variables. We appreciate their interest and thoughtful comments made on the measures. Although some comments will not be addressed in this memo, we have discussed with NQF and the Yale Center for Outcomes Research and Evaluation (CORE). This memo is organized to summarize and respond to the appellant's comments on each issue identified above.

I. Use of Race Variable

Comment: Stakeholders expressed concern on use of the race variable, commenting on the quality of race/ethnicity data and noting that race/ethnicity should not be used as a proxy for socioeconomic status (SES).

Response: In regards to the issue of using race as a proxy for SES, we agree with the appellants that race generally should not serve as a proxy for SES. We feel it is useful to examine race not as a proxy for SES but as an important comparator. Although the NQF Expert Panel on Risk Adjustment for Sociodemographic Factors did not provide clear guidance regarding inclusion of race, the panel did broaden the term from SES to SDS to account for consideration of racial disparities, and we feel it is useful to understand the pattern of racial disparities along with SES disparities in these payment measures. Moreover, the Cost and Resource Use Standing Committee did agree with CORE's analytic plan to examine race. We believe it is helpful to show analyses with race, not because it should be incorporated in risk adjustment models, but as a point of comparison with other SES variables. The conceptual rationale for not adjusting for SES has important parallels with race in that both SES and race are associated with access to high quality care and can lead to differential care within hospitals. These comparisons can be helpful in understanding causal pathways and for making decisions about incorporation of SES in risk adjustment models.

We share concerns regarding the quality of national race/ethnicity data. However, CMS data are not yet specific or sensitive enough to determine race/ethnicity at a more granular level. To be specific, CMS research has shown that "black" and "white" are the only categories of CMS' beneficiary race code variable with high sensitivity and specificity. In the future, when other race/ethnicity categories are more reliable or when other race/ethnicity variables are reliably available, we would certainly support their inclusion in SDS evaluation, but only as a comparator with other SES variables.

II. Consideration of Community and Environmental Factors

Comment: Stakeholders expressed interest in incorporating community-level factors in analyses and risk models.

Response: We appreciate the stakeholder's consideration of community-level factors. We believe the use of ZIP code-linked variables – e.g., the Agency for Healthcare Research and Quality (AHRQ) SES Index that is derived from the American Community Survey (ACS) census block group level data and linked to a patient's ZIP code – can capture community factors and are tested in models at the patient-level as a proxy for patient SES. Additionally, conducting analyses using patient-level variables was consistent with the guidance from NQF: "If a conceptual relationship exists between a *patient-level* sociodemographic factor and outcome, it should be tested empirically."

In terms of using community-level factors that are not at the patient level within the risk adjustment model, we see a few challenges. First there, there is insufficient evidence on which community factors influence health care utilization and episode payment and what would be appropriate to incorporate in risk models. There is also a need to carefully consider the policy implications of incorporating community factors into episode payment models since many potential variables are related to availability of services (such as nursing homes or primary care) which may be driving utilization patterns that the measures are meant to illuminate. So although we are open to considering new approaches to modelling and potential incorporation of community variables, we felt this was not the charge of the NQF guidance, and we do not feel the evidence is sufficient to do so at this time.

III. Use of Additional Patient-Level Variables

Comment: Stakeholders expressed concern with performing analyses using only dual-eligible status and expressed interest in the use of 9-digit zip code data in analyses.

Response: At the time of CORE's meeting with the NQF Cost and Resource Use Standing Committee, CORE identified all feasible variables for use in measures based on the Medicare administrative claims dataset. Among the identified variables, the Committee discouraged CORE from further examination of the AHRQ SES Index linked to a patient's 5-digit ZIP code. (CORE was not able to link the AHRQ SES Index at the 9-digit zip code level at the time of the Standing Committee's in-person meeting.) Secondly, CORE considered the Low-Income Subsidy (LIS) variable and the Supplemental Security Income (SSI) variable. LIS was not used because it has a slightly higher income threshold and does not capture many additional patients above dual eligible status. Patient-level SSI is unavailable for use by developers (only used by CMS to calculate disproportionate share hospital [DSH] status but not otherwise available).

We note that CORE has now completed analyses for the acute myocardial infarction, heart failure, and pneumonia payment measures using 9-digit ZIP code linked to the AHRQ SES Index (a composite of 7 SES variables including housing, income and education from the American Community Survey) at the census block group level. We also adjusted the AHRQ SES Index for cost of living. The results of these analyses are similar to the results of the analyses using the black/non-black and dual-eligible status indicator variables.

CORE Payment Measures: Using 9-digit ZIP Code

Table 1. Relationships between Total Payment and SES or Race Variables

Measure	Variable in the Model	Bivariate Model		Multivariate Model (Current* + SES/Race Variable)	
		Payment Ratio[r]/ Estimate	P-Value	Payment Ratio[r]/ Estimate	P-Value
AMI	Race	1.01	0.0261	0.94	<0.0001
	Dual Eligibility	1.00	0.0657	0.98	<0.0001
	Low SES census block group (AHRQ SES index, linked to 9-digit ZIP – Adjusted for Cost of Living)†	1.01	<0.0001	0.98	<0.0001
HF	Race	1.01	<0.0001	0.97	<0.0001
	Dual Eligibility	1.06	<0.0001	1.01	<0.0001
	Low SES census block group (AHRQ SES index, linked to 9-digit ZIP – Adjusted for Cost of Living)†	1.00	0.4171	0.98	<0.0001
PN	Race	$1,708	<0.0001	$391	<0.0001
	Dual Eligibility	$1,600	<0.0001	$516	<0.0001
	Low SES census block group (AHRQ SES index, linked to 9-digit ZIP – Adjusted for Cost of Living)†	$191	<0.0001	-$134	<0.0001

* Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)

† AHRQ SES index score is less than or equal to 42.7

[r] Payment ratio is equal to exponentiated estimate

Table 2. Distribution of Percent Change in RSPs using the Current Model with Each SES or Race Indicator Added (July 2011-December 2013)

Measure	Distribution	Current* + Race (% RSP Change)	Current* + Dual Eligibility (% RSP Change)	Current* + Low SES census block group (AHRQ SES index, linked to 9-digit ZIP – Adjusted for Cost of Living)† (%RSP Change)
AMI	Minimum	-0.53	-0.38	-0.28
	10th Percentile	**-0.31**	**-0.18**	**-0.15**
	25th Percentile	-0.19	-0.087	-0.071
	Median	-0.064	-0.013	-0.0014
	Mean	0.00084	0.00013	**0.000076**
	75th Percentile	-0.0079	0.054	0.051
	90th Percentile	0.34	0.17	**0.15**
	Maximum	5.06	1.11	0.65
HF	Minimum	-0.45	-0.7	-0.31
	10th Percentile	-0.24	-0.16	**-0.20**
	25th Percentile	-0.19	-0.062	-0.12
	Median	-0.094	0.014	-0.028
	Mean	0.00056	0.000087	**0.00015**
	75th Percentile	0.026	0.089	0.087
	90th Percentile	0.36	0.15	**0.25**
	Maximum	2.59	0.29	0.68
PN	Minimum	-1.09	-2.49	-0.11
	10th Percentile	-0.14	-0.58	**-0.076**
	25th Percentile	-0.004	-0.22	-0.057
	Median	0.048	0.088	-0.016
	Mean	0.0031	0.0059	**-0.00014**
	75th Percentile	0.075	0.32	0.039
	90th Percentile	0.089	0.48	**0.11**
	Maximum	0.19	0.95	0.31