# National Quality Forum

# Memo

**Date:** May 19, 2015

**To**: Cost and Resource Use Standing Committee &

CMS/Yale Measure Development Team

**From**: NQF Cost and Resource Use Project Team

**Re**: Review of SDS Conceptual Analysis for CMS/Yale Cost Measures

The Cost & Resource Use Standing Committee will meet via webinar on Thursday, May 21.

**The purpose of the meeting is to**:

- Provide an overview of the process and plan for reviewing the (3) CMS/Yale cost measures for cardiovascular and pneumonia conditions under the new guidance for sociodemographic status (SDS) risk adjustment.
- Review and discuss the conceptual analysis of the selected SDS risk adjustment factors for the (3) cost measures
- Determine whether further empirical analysis of the impact of SDS factors in the risk model is warranted for the measures.
- Discuss and provide guidance on next steps for empirical analysis (if warranted) of the impact of the SDS factors in the risk model.

---

**Standing Committee Action**:

1. Review the Yale submission of the conceptual model and memo discussing the conceptual analysis of SDS risk factors and hospital-level cost measurement.
2. Review this memo; prepare to provide input and discuss the Committee discussion questions on page 4.

---

**Conference Call & Webinar Information:** *Thursday, May 21, 2pm-4pm ET*

- Conference call dial in: (888) 802-6696
- Web Link: http://nqf.commpartners.com/se/Rd/Mt.aspx?262682

---

## Agenda

2:00pm   Welcome & Roll Call

2:05pm   Background
  - How did we get here?
  - Goals and purpose of this call

2:15pm   Review of Conceptual Analysis
  - Developer overview and summary of submission
  - Committee Discussion

3:45pm  Public and Member Comment

3:55pm  Next Steps

4:00pm  Adjourn

## Background

The NQF Board of Directors Executive Committee ratified the CSAC's recommendation to endorse the following cost 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)

Only 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 this measure in use. The evaluation of unintended consequences will be initiated in approximately one year and possible changes to the measures based on this 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 sociodemographic status adjustment.

## Overview of the Sociodemographic Status (SDS) Adjustment Trial Period

The trial period approved by the NQF Board of Directors is designated as a 2-year period of time during which SDS factors should be considered as potential factors in the risk-adjustment model if there is a conceptual reason for doing so. If there is a conceptual relationship between potential SDS risk factors and the outcome of interest, the developer should conduct empirical analyses to determine whether such factors improve the risk-adjustment model. Based on that analysis, measure developers may submit measures with SDS factors included in the risk model. The trial period begins April 2015.

Prior to this decision, NQF criteria and policy prevented the inclusion of SDS factors in the risk model and only allowed for the inclusion of a patient's clinical factors present at the start of care. Rather than including SDS factors related to the outcome in the risk-adjustment model, NQF criteria required that measures enable the stratification of these variables.

## Reviewing the Cost Measures during the SDS Trial Period

In collaboration with the CMS/Yale measure development team, NQF agreed to divide the assessment of the impact of SDS variables on the risk model and performance scores for the cost measures into two stages (and webinars):

- **Stage 1/Webinar #1 (May 21, 2-4pm ET): Conceptual Analysis**
    - Review of conceptual analysis of selected variables
    - Determine whether further empirical analysis is warranted
    - Identify the variables to be pursued in empirical analysis

- Provide input on the plan or approach to empirical analysis of the selected variables.
- **Stage 2/Webinar #2 (October 27, 3-5pm ET): Empirical Analysis**
    - Review empirical analysis of the impact of SDS risk factors in the risk model
    - Determine endorsement status:
        - Recommend [continued] endorsement of the measure.
        - Recommend to de-endorse the measure.

## Conceptual Analyses Review

A conceptual relationship refers to a logical theory or rationale that explains the association between an SDS factor(s) and the outcome of interest. The conceptual basis may be informed by prior research and/or healthcare experience related to the outcome of interest, but does not require a direct causal relationship (i.e., it could be a direct cause, an indirect cause, or serve as a surrogate for a cause for which data are lacking).

An assessment of a conceptual relationship between an SDS factor and an outcome of interest includes a consideration of whether the effect of the SDS is primarily mediated by the quality of care delivered (i.e., does the SDS factor lead to the delivery of inferior care processes, which in turn affect the outcome?).

Some potential questions that can be considered to describe the conceptual relationship between an outcome measure and possible SDS risk factors include:

- Does prior research indicate a relationship between SDS and the outcome?
- Is there a logical relationship or theory about the relationship between SDS and the outcome?
- Is there a significant passage of time between the healthcare unit intervention and measured outcome during which other factors may have an effect?
- Do patient actions or decisions influence the outcome or process and are the decisions affected by SDS (e.g., ability to purchase medications)?
- Does the patient community have an influence (e.g., distance to pharmacies, groceries, healthcare services)?

- Risk factors should not be confounded with the effect of the healthcare unit
    - Risk factors should be present at the start of care
    - Risk factors should not be an indicator or characteristic of the care provided (e.g., treatments, interventions, expertise of staff)
- Data for risk factors should be captured reliably and feasibly

*Variables under Consideration (based on Yale Submission)*
- Patient zip code (proxy for educational attainment or income)
- Medicaid status (proxy for low income and insurance coverage)
- Black or white race

**Committee Discussion:**
- Has the developer adequately demonstrated that there is (or is not) a conceptual relationship between the risk factors and the payment/resource utilization/cost for each measure or condition (e.g., pneumonia, AMI, HF)? (i.e., Does the Committee believe there is a conceptual relationship?)
- How well do these variables proxy for the intended SDS factors and align with the conceptual model?
- If there is a conceptual relationship, are the data available, feasible and accessible (for this population) in order for these factors to be used in empirical testing of risk-adjustment?
- Based on the conceptual analysis provided by the developers, does the Committee believe that further empirical analysis is warranted?
  - If so, which factors does the Committee recommend the developers pursue in the empirical analysis?

## Preparing for Empirical Analysis

If the Committee believes a conceptual relationship exists between the sociodemographic factor(s) and the outcome (i.e, resource utilization or cost), it should be tested empirically to confirm that relationship. NQF does not recommend any particular analytic approach with which to assess empirical associations between sociodemographic factors and outcomes, nor any specific cutoff or threshold value to use for declaring the presence of an association.

Current NQF guidance for the submission of empirical analysis of SDS factors in the risk model requires the submission of:
- Analyses and interpretation resulting in decision to include or not include SDS factors in section.
- Compare performance scores and risk model performance with and without SDS factors in the risk adjustment model (including method and results).
- An interpretation of their results in terms of the differences in performance scores for the same entities.
- If the developer has decided to SDS adjust they will need to submit, updated reliability and validity testing and specifications for a stratified version of the measure using these factors.

**Committee Discussion:**
- If the developer has a plan for the empirical analysis for the Committee to consider, what recommendations or input does the Committee have on the proposed approach?
- If a plan has not been submitted, what considerations or recommendations would the Committee like to provide to the developers as they develop their approach?

**Appendix A: Sociodemographic Factors – PROs and CONs**

Table 6  (page 44), excerpted from the NQF Technical Report: Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors.

**Table 6. Sociodemographic Factors – PROs and CONs**

| Factors/Concepts (specific variables) | PROs | CONs | Caveats |
|---|---|---|---|
| **Factors that should be considered, depending on:  data availability and the specific outcome or process** | | | |
| **Income** | • Allows for use of various ranges | • Hard to collect privately (e.g., in clinician office)<br>• Not easily collected with a single question<br>• May not be an acceptable question to all patients<br>• Meaning is not geographically consistent due to difference | • For national performance measures, need to consider standardization to account for area wage and cost of living differences |
| Income in relation to federal poverty level | • Definition is standard<br>• Being used under ACA<br>• Researchers | • Doesn't include receipt of other benefits (e.g., food stamps)<br>• Doesn't account for cost of living or community | . |
| Household income | • May be more meaningful | • Requires assessment of household size | . |
| Medicaid status as proxy | • Relatively easy to collect in claims data | • Eligibility not consistent across states | • Potentially becomes more useful as more States expand Medicaid to 138% federal poverty level |

| Factors/Concepts (specific variables) | PROs | CONs | Caveats |
|---|---|---|---|
| Social Security Supplemental Income (SSI) | | • Correlated with Medicaid status, but not consistently across states | • In many states, receipt of SSI automatically makes one eligible for Medicaid |
| **Education** | • Perceived to be valid (i.e., less misreporting than for income)<br>• Definitions fairly consistent across various subgroups (e.g., answers from immigrants comparable to those from others)<br>• Fairly stable across time, | • Not widely collected by healthcare units<br>• If collected (e.g., in EHR text fields) may not be easily retrievable | |
| **Homelessness** | • Strongly associated with health outcomes<br>• Measures something "beyond" income | • Multiple other definitions<br>• Data often not collected<br>• Status can change | • Prevalence tends to cluster among safety net healthcare units |
| Housing instability | • May be better indicator than homelessness which can change | • More difficult to define than homelessness | |
| **English Proficiency** | • Standard definition exists<br>• Tied to need for translation services/other resource needs and therefore should be collected<br>• Increasingly being collected (required by | | |
| **Insurance Status** | • Readily available<br>• Some indication of access and resources<br>• Benefit coverage strongly related to | • Wide variability in insurance coverage<br>• Data for underinsurance | |
| Medicaid status | • Readily available<br>• Some indication of limited income and | • Not consistent across states | |

| No insurance | • Readily available<br>• Standard meaning | | • Difficult to capture information about these patients (particularly if |
| --- | --- | --- | --- |
| **Community/ Neighborhood-level data used as proxy for individual data or as contextual variable** | • Many variables available from Census data<br>   • Income<br>   • Education<br>   • Immigration status<br>   • Language<br>   • Unemployment<br>   • Home ownership<br>   • Single parents<br>   • Others | • Census data do not include all potentially important variables<br>• Residential heterogeneity will affect whether it is a good proxy for data about individuals.<br>• Heterogeneity may differ based on levels of socioeconomic segregation and potentially population density.<br>• Requires geocoding for Census Tract and smaller areas. | |
| Contextual - Proportion vacant housing | • Seen as indicator for other related issues such as poverty, crime, lack of resources | | |
| Contextual- Crime rate | • May be an indicator for other related issues such as poverty, lack of resources | | |
| **Other factors that could be considered** | | | |
| **Factors/Concepts** (specific variables) | **PROs** | **CONs** | **Caveats** |
| **Social Support** | • Some brief items have been used in previous research<br>• Captures something that other variables do not | • Multidimensional construct that typically requires multiple questions<br>• Lack of agreement about how to | |

| | | | |
|---|---|---|---|
| Living alone | • Available in OASIS data for home health | • Directionality may not be consistent. In some situations such as frailty or impairment, it could be a risk factor. In other situations, it might be an indicator of ability to live alone due to good health and function. | |
| Marital status | • Often collected | | |
| **Occupation** | • May capture other concepts (e.g., environmental exposures) | • Multiple definitions<br>• Potentially large data collection burden due to the complexity of the concept<br>• Marginal value (i.e., over and above that contributed through use of other variables) may be limited<br>• Unclear how to handle certain population subgroups (e.g., retirees, students, homemakers) | |
| **Employment Status** | • Often collected | • Employment status does not reflect income or availability of insurance<br>• Simple yes/no does not reflect desire/happiness with situation (e.g., retirees may be happy to be unemployed)<br>• Subject to change requiring continuous updating | |

| | | | |
|---|---|---|---|
| **Literacy** | • This concept may also be able to partially capture health literacy | • No standardized definitions<br>• May be easy to game | If the correlation with education is high, then education could be used. |
| **Health Literacy** | • Potentially more relevant to healthcare<br>• Three-item and single-item validated questions | • Not consistently collected/ available | |
| **Local/state funding for safety net providers** (e.g., tax base) | • Affect resources available to safety net providers beyond insurance | • Data not easily collected/ available | • Not a patient characteristic<br>• Risk for unintended consequences (setting a lower standard for poorly supported institutions might send the wrong messages to tax payers) |
| **Race/ Ethnicity** | • Correlated with SES and may be more available than other variables | • May be more correlated with bias | • Should not generally be used as proxy for SES |

# MEMORANDUM

**TO:** NQF Cost and Resource Use Project Team
**FROM:** Nancy Kim and Susannah Bernheim, Yale New Haven Health Services Corporation - Center for Outcome Research and Evaluation (CORE)
**THROUGH:** Lein Han and Vinitha Meyyur, CMS
**DATE:** Thursday, May 7, 2015
**SUBJECT**: Proposed SDS Trial Period Evaluation Process for Cost and Resource Use Measures

---

In early 2015, the NQF Cost and Resource Use Standing Committee decided the payment measures will enter the SDS Trial Period. On February 12, 2015, NQF proposed a process for the SDS Trial Period Evaluation of the Cost and Resource Use Measures to which Yale proposed a response on February 22, 2015. This response was in agreement with a two staged process that will include 2 webinars: Webinar #1 conceptual analysis and determination of SDS variables to be used in empiric analysis and Webinar #2 empiric analysis.

This memo and the attached conceptual model is being submitted to NQF in preparation for Webinar #1 to address the question of what variables we plan to include in our future analyses and to present an initial conceptual framework for the causal pathways by which SDS may influence episode payments. This memo and the conceptual model relate to all three of the following 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)

In general, as evidenced by the national debate around SDS adjustment, the question of the relationship between SDS and episode-of-care payment is complex and not well understood. Given the timeframe of this request we have provided a very brief overview of the findings of our literature search, a discussion about our approach to risk-adjustment for SDS, its implication for the payment measures, and the variables currently available to us in claims data.

## Literature Review

We performed 3 separate focused literature reviews of the recent scientific literature to examine what is known regarding the relationship between socioeconomic status (SES) and/or sociodemographic (SDS) factors and costs of care for AMI, HF, or pneumonia patients.

Of the small number of relevant studies that examined this question for each condition, few were focused on the older, Medicare insured population that comprise the cohorts for our measures. Nevertheless, when we synthesized the existing literature, we found that most studies used the following markers of disadvantage to examine this relationship: income, insurance status, education, or race. Many studies suggested that at a patient-level low income, worse insurance coverage, lower education, and non-white race were associated with higher costs (measured in a number of different ways). However the studies and findings were quite heterogeneous.

Interestingly one paper (Cram, et al.) identified in the HF literature search suggested that patients with more generous insurance were more readily admitted to specialty cardiology hospitals. This article concluded that specialty hospitals may contribute to differential healthcare along socioeconomic lines, but hints that more generous insurance may also result in higher costs.

Two other studies (Barnato, et al.; Jha, et al.) suggested that the mechanism for existing racial disparities in cost may be mediated by a hospital effect. In other words, disadvantaged patients went to hospitals utilizing evidence-based medical treatments at lower rates than other hospitals or that ranked among the worst hospitals in terms of quality and cost (low quality, high cost). These studies suggest that risk adjustment for race could potentially mask hospital cost signals that the measures aim to illuminate.

Based on this literature review and data availability, we limited the variables we would consider for risk adjustment in the episode of care payment measures to (1) educational attainment or income based on census data from a patient's zip code, (2) Medicaid status as a proxy for both low income and insurance coverage for some post-discharge services, and (3) black or white race. Further details of our literature search for each condition can be found in our appendix. This literature review also helped to inform the development of our conceptual model.

**Conceptual Approach to Risk-adjustment**

Below we present a conceptual model for the pathways by which SDS may influence the costs of an episode of care and a number of key questions and starting assumptions about risk-adjustment in general.

**Question 1. Why risk adjust hospital episode payment measures?**

1. The overarching goal of risk-adjustment is to adjust for differences in patient case-mix among hospitals in so far as these patient characteristics at the time of admission influence the episode of care payment, but are outside of the hospital's control. Factors that are within the control of the hospital may or may not be appropriate for adjustment depending upon the goal of the measure and the underlying mechanism.

**Question 2a: Which factors are reasonably outside a hospital's control and influence the payment outcome, and therefore appropriate to consider for risk adjustment for the purpose of identifying differences in hospital payment?**

2a. Although hospitals can help mitigate community health needs, most patient characteristics present before or at the time of admission are beyond a hospital's control. Therefore, there should

be adjustment for these factors that are present on admission if they have a conceptual relationship with the payment outcome. These include but are not necessarily limited to: age, gender, patient comorbidities, and prior procedures.

**Question 2b: How does SDS relate to these variables?**

2b. Relationship of SDS to health at admission. A lifetime of SDS disadvantage often leads to worse general health status and therefore patients who have lower income/education/literacy, tenuous housing, and non-white race may present for their hospitalization with a greater severity of underlying illness. These factors may also contribute to worse health status at admission due to patients not having the resources to respond to early symptoms and presenting later in their disease process, which may then contribute to greater health care needs and costs. (Although in some cases greater disease severity and comorbidity burden may limit interventions and therefore decrease costs.)

To the extent that SDS exerts its effect on health status at admission through higher comorbidity burden, we account for this with the current risk-adjustment strategy in all 3 payment measures.

**Question 3a: Which factors are fully within the control of the hospital and influence the payment outcome?**

3a. The clinical care delivered throughout the hospitalization and the discharge process is within a hospital's control. Clinical care decisions that affect payment include, but are not necessarily limited to: (1) decisions about care management, procedures, and medications, and (2) tailoring care that is appropriate for a patient's needs including decisions about care transitions and follow-up care (e.g. discharge disposition, instructions, support for follow-up appointments).

**Question 3b: How does SDS interact with these variables?**

3b. Relationship of SDS to care within a hospital. SDS factors can contribute to the payment outcome if patients do not receive equivalent care within a facility due to SDS. For example, patients of low income or minority race may experience differential, lower quality, or discriminatory care within a given facility.[1] This may lead to lower use of procedures and lower cost care (or vice versa). Additionally, patients with SDS risk factors may require differentiated care (e.g. provision of information at a lower reading level). Such differentiated care may be more or less costly for a hospital to provide but necessary for high quality care.

Determination of whether risk-adjustment for SDS is appropriate in such situations is complex as these pathways may lead to lower or higher cost care. Also the appropriateness of adjustment depends, in part, on the extent to which patients of low SDS consistently require care that is more (or less) costly for the given condition (rather than differences that are the result of differential or discriminatory care).

**Question 4: Which factors are partially within a hospital's control and influence the payment outcome?**

4a. Many aspects of clinical care that contribute to episode of care payments occur post-discharge and are partially within a hospital's control. These include but are not necessarily limited to: (1)

---

[1] Trivedi AN, Nsa W, Hausmann LRM, et al. Quality and Equity of Care in U.S. Hospitals. New England Journal of Medicine. 2014;371(24):2298-2308.

selection of post-discharge care that is in a location and price point that meets patient needs, (2) aiding patients in complying with care plans (filling medications before discharge and sending reminders before appointments), (3) communication with patients (responsiveness to patient needs, receiving and reading discharge summary), and (4) mitigation of environmental factors affecting health (providing a safe place to recover, transportation, access to social services, social support).

**Question 4b: How does SDS interact with these variables?**

4b. <u>Relationship of SDS to post-discharge period.</u>  Patients of lower SDS may have greater post-discharge care needs. However they may also live in areas of limited availability of certain services. Hospitals have some but not complete influence on the cost of post-discharge care. Some discharge strategies may mitigate some of the negative effects of low SDS during a patient's recovery; by improving recovery these strategies can limit the costs associated with worsening clinical status (such as readmission) however some strategies (e.g. providing transportation) may increase post-discharge costs). In other instances a hospital may make appropriate care decisions and provide tailored care and education but a lower-income patient may not follow the care plan because limited resources create competing priorities for that patient.

As with the factors fully within a hospital's control, determination of the appropriateness of risk-adjustment for SDS will depend in part on the extent to which the above pathways are contributing to the outcome.

**Question 5: Given these pathways how will empiric analyses guide risk-adjustment decisions?**

5. Each of the pathways described above has different implications for the treatment of SDS variables within the risk model. For example, to the extent SDS influences risk by affecting admission health status, this may already be accounted for in the clinical risk-adjustment of the measures and further adjustment may not be warranted. Yet if patients of low SDS for a given condition consistently require differential care within the hospital to achieve good outcomes and this results in higher (or lower) payments across the episode of care, risk adjustment may be warranted if it does not simultaneously mask the potential effects of discriminatory care. If a predominant pathway is apparent from empiric analyses this will guide risk adjustment decisions. All of these pathways may be implicated, but the relative influence of each has been little studied to date.

**Question 6: What are the implications of SDS risk-adjustment on the episode of care payment measures?**

6. There is little available scientific evidence to suggest whether SDS adjustment will increase or decrease the total risk-standardized payment. Although low SDS may result in a "sicker" patient on admission, making the patient potentially more expensive to care for; that patient may not have access to healthcare providers or testing in the post-discharge setting, which may lower payments for that patient.

# Appendix

## Overview of Literature Reviews

<u>**AMI Literature Review**</u>

**Research Question:** What is the relationship between socioeconomic status (SES) and/or sociodemographic factors, and costs for acute myocardial infarction (AMI)?

**Methodology:** The review was an iterative process allowing for continuous improvement (that is, refinement of inclusion criteria; and, use of exclusion criteria [for example, exclusion of international studies, exclusion of studies with low reliability and/or validity, etc.]). Our final search string (below) provided 287 publications which were assessed for empiric and conceptual evidence related to SDS and cost/payment.

**Search String:** ((((Myocardial infarction) OR "AMI")) AND (((((((socioeconomic factors [MeSH Terms]) OR race [MeSH Terms]) OR socioeconomic status) OR sociodemographic) OR healthcare disparities [MeSH Terms]) OR health status disparities [MeSH Terms]) OR minority health [MeSH Terms])) AND economics [MeSH Terms]

**Summary of Results:** 287 publications were retrieved. Of these studies, 207 (72%) focused on myocardial infarction, 73 (25%) on socioeconomic status and/or sociodemographic factors, and 69 (24%) on cost and/or payment. After reviewing the titles and abstracts, 267 studies were excluded. Of the excluded studies, 232 (87.9%) were excluded due to lack of empiric and/or conceptual evidence related to the research question and 35 (13.1%) were excluded because the abstracts and/or full-text were not available. The remaining 22 articles were reviewed in full. These articles represented a broad focus on the clinical condition (AMI) and outcomes of interest (cost/payment).

**Reference List:**

1. Shen JJ, Wan TT, Perlin JB. An exploration of the complex relationship of socioecologic factors in the treatment and outcomes of acute myocardial infarction in disadvantaged populations. *Health services research.* Aug 2001;36(4):711-732.
2. Jha AK, Orav EJ, Epstein AM. Low-quality, high-cost hospitals, mainly in South, care for sharply higher shares of elderly black, Hispanic, and medicaid patients. *Health affairs (Project Hope).* Oct 2011;30(10):1904-1911.
3. Georgiades A, Janszky I, Blom M, Laszlo KD, Ahnve S. Financial strain predicts recurrent events among women with coronary artery disease. *International journal of cardiology.* Jun 26 2009;135(2):175-183.
4. Hasan O, Orav EJ, Hicks LS. Insurance status and hospital care for myocardial infarction, stroke, and pneumonia. *Journal of hospital medicine.* Oct 2010;5(8):452-459.
5. Shaw LJ, Merz CN, Bittner V, et al. Importance of socioeconomic status as a predictor of cardiovascular outcome and costs of care in women with suspected myocardial ischemia. Results from the National Institutes of Health, National Heart, Lung and Blood Institute-sponsored Women's Ischemia Syndrome Evaluation (WISE). *Journal of women's health (2002).* Sep 2008;17(7):1081-1092.
6. Barnato AE, Lucas FL, Staiger D, Wennberg DE, Chandra A. Hospital-level racial disparities in

acute myocardial infarction treatment and outcomes. *Medical care.* Apr 2005;43(4):308-319.

7.  Alter DA, Naylor CD, Austin PC, Chan BT, Tu JV. Geography and service supply do not explain socioeconomic gradients in angiography use after acute myocardial infarction. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* Feb 4 2003;168(3):261-264.

## HF Literature Review

**Research Question:** What is the relationship between socioeconomic status (SES) and/or sociodemographic factors, and costs for heart failure (HF)?

**Methodology:** The review was an iterative process allowing for continuous improvement (that is, refinement of inclusion criteria; and, use of exclusion criteria [for example, exclusion of international studies, exclusion of studies with low reliability and/or validity, etc.]). Our final search string (below) provided 280 publications which were assessed for empiric and conceptual evidence related to SDS and cost/payment.

**Search String:** ((heart failure) AND (((((((socioeconomic factors [MeSH Terms]) OR race [MeSH Terms]) OR socioeconomic status) OR sociodemographic) OR healthcare disparities [MeSH Terms]) OR health status disparities [MeSH Terms]) OR minority health [MeSH Terms])) AND economics [MeSH Terms]

**Summary of Results:** 280 publications were retrieved. After reviewing the titles and abstracts, 262 studies were excluded. Of the excluded studies, 237 (90.5%) were excluded due to lack of empiric and/or conceptual evidence related to the research question and 25 (9.5%) were excluded because the abstracts and/or full-text were not available. The remaining 18 articles were reviewed in full. These articles represented a broad focus on the clinical condition (HF) and outcomes of interest (cost/payment).

**Reference List:**

1.  Cram P, Pham HH, Bayman L, Vaughan-Sarrazin MS. Insurance status of patients admitted to specialty cardiac and competing general hospitals: are accusations of cherry picking justified? *Medical care.* May 2008;46(5):467-475.
2.  Heisler M, Choi H, Rosen AB, et al. Hospitalizations and deaths among adults with cardiovascular disease who underuse medications because of cost: a longitudinal analysis. *Medical care.* Feb 2010;48(2):87-94.
3.  Nelson EC, McHorney CA, Manning WG, Jr., et al. A longitudinal study of hospitalization rates for patients with chronic disease: results from the Medical Outcomes Study. *Health services research.* Feb 1998;32(6):759-774.
4.  Piamjariyakul U, Yadrich DM, Russell C, et al. Patients' annual income adequacy, insurance premiums and out-of-pocket expenses related to heart failure care. *Heart & lung : the journal of critical care.* Sep-Oct 2014;43(5):469-475.
5.  Unroe KT, Greiner MA, Johnson KS, Curtis LH, Setoguchi S. Racial differences in hospice use and patterns of care after enrollment in hospice among Medicare beneficiaries with heart failure. *American heart journal.* Jun 2012;163(6):987-993.e983.

**Pneumonia Literature Review**

**Research Question:** What is the relationship between socioeconomic status (SES) and/or sociodemographic factors, and costs for pneumonia?

**Methodology:** The review was an iterative process allowing for continuous improvement (that is, refinement of inclusion criteria; and, use of exclusion criteria [for example, exclusion of international studies, exclusion of studies with low reliability and/or validity, etc.]). Our final search string (below) provided 175 publications which were assessed for empiric and conceptual evidence related to SDS and cost/payment.

**Search String:** ((heart failure) AND (((((((socioeconomic factors [MeSH Terms]) OR race [MeSH Terms]) OR socioeconomic status) OR sociodemographic) OR healthcare disparities [MeSH Terms]) OR health status disparities [MeSH Terms]) OR minority health [MeSH Terms])) AND economics [MeSH Terms]

**Summary of Results:** 175 publications were retrieved. After reviewing the titles and abstracts, 173 studies were excluded. Of the excluded studies, 162 (93.6%) were excluded due to lack of empiric and/or conceptual evidence related to the research question (for example, many studies focused on infectious diseases [AIDS, influenza, etc.]) and 11 (6.4%) were excluded because the abstracts and/or full-text were not available. The remaining 2 articles were reviewed in full.

**Reference List:**

1. Hasan O, Orav EJ, Hicks LS. Insurance status and hospital care for myocardial infarction, stroke, and pneumonia. *Journal of hospital medicine.* Oct 2010;5(8):452-459.
2. Carrie AG, Kozyrskyj AL. Disease, temporal and sociodemographic influences on initial treatment of community-acquired pneumonia in Manitoba, Canada. *International journal of antimicrobial agents.* Aug 2006;28(2):95-100.

SDS variables under consideration for risk adjustment like **Income, Education, Medicaid Status, and Race** may act through several pathways to influence episode of care payments

| Lifetime health effects of low SDS | Differential services or discrimination | Differential services or discrimination | Fewer resources |
|---|---|---|---|

## Pre-Admission

**Health status at admission**

—Comorbidities

—Prior procedures

—Time to arrival

## Hospitalization

**Clinical care**

—Procedures

—Utilization of services

—Medical management

**Additional services**

—Translation

—Rehabilitation

—Nutrition

## Care Transition

**Discharge planning**

—Communication with post-acute providers

—Schedule follow-up appointments

—Easy to understand discharge instructions

**Mitigation of patient needs**

—Access to necessary medications

—Transportation to follow-up appointments

—A place to recover post-discharge

—Connections to community resources

## Post-Discharge Environment

**Community factors**

—Lack of community services

—Lack of social supports/caregivers

**Patient Behavior**

—May not use services provided

—May not fully adhere to care plan

Hospital influence on episode of care payment:

(1) Direct costs associated with care decisions and (2) Costs associated with the outcome of care

# Memo

**Date:**   October 22, 2015

**To**:      Cost and Resource Use Standing Committee &

CMS/Yale Measure Development Team

**From**:   NQF Cost and Resource Use Project Team

**Re**:      Review of SDS Empirical Analysis for CMS/Yale Cost Measures

The Cost & Resource Use Standing Committee will meet via webinar on Tuesday, October 27.

**The purpose of the meeting is to**:
- Provide an overview of the process and plan for reviewing the (3) CMS/Yale cost measures for cardiovascular and pneumonia conditions under the new guidance for sociodemographic status (SDS) risk adjustment.
- Review and discuss the empirical analysis of the selected SDS risk adjustment factors for the (3) cost measures
- Prepare the Committee to make final recommendations on the validity and endorsement status for the measures under review.

---

**Standing Committee Action**:
1. Review the Yale memo discussing responses to prior recommendations and empirical analysis of SDS risk factors.
2. Review this memo; prepare to provide input and discuss the Committee discussion questions on page 6.
3. Submit vote on validity criterion and endorsement recommendation.

---

## Agenda
3:00pm   Welcome & Roll Call
3:05pm   Background
     o   How did we get here?
     o   Goals and purpose of this call
3:15pm   Review of Empirical Analysis
     o   Developer overview and summary of memo and findings
     o   Committee Discussion
4:45pm   Public and Member Comment
4:55pm   Next Steps
5:00pm   Adjourn

## Overview:  Reviewing the Cost Measures during the SDS Trial Period
The SDS trial period approved by the NQF Board of Directors is designated as a 2-year period of time during which SDS factors should be considered for potential inclusion in the risk-adjustment approaches if there is a conceptual reason for doing so.  If there is a conceptual relationship between potential SDS risk factors and the outcome of interest, the developer

should conduct empirical analyses to determine whether such factors should be included in the risk-adjustment approach.

Following the NQF Board of Directors Executive Committee decision to endorse the cost measures with the condition that they be considered under the trial period guidance, NQF, in collaboration with the CMS/Yale measure development team, agreed to divide the assessment of the impact of SDS variables into two stages (and webinars).

### *Measures under Review*

- #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)

### Stage 1/Webinar #1 (May 21, 2-4pm ET): Conceptual Analysis

- Review conceptual analysis of relationships between SDS factors and the outcomes of interest
- Determine whether further empirical analysis is warranted
- Identify the variables to be pursued in empirical analysis
- Provide input on the plan or approach to empirical analysis of the selected variables

### *Stage 2/Webinar #2 (October 27, 3-5pm ET): Empirical Analysis*

- Review empirical analysis of the impact of SDS risk factors in the risk model and measure scores:
    - Vote on Validity Criterion
- Make a recommendation on endorsement status:
    - Recommend [continued] endorsement of the measure OR
    - Recommend to de-endorse the measure

## *Conceptual Analyses Review*

For the May webinar, the CMS/Yale Core development team submitted a memo summarizing their literature review and conceptual model diagram illustrating the potential relationships between various factors during the episode of care captured by the measures (i.e., hospital admission through 30 days post-discharge). Based on this conceptual analysis, they identified three variables that have been identified in the literature to have a conceptual relationship to utilization and payment.

### *Proposed Variables:*
- Patient zip code (proxy for educational attainment or income)
- Medicaid status (proxy for low income and insurance coverage)
- Black or white race

*Summary: Committee Discussion and Recommendations on Conceptual Analyses (Webinar 1)*

1. **Broaden the conceptual model**. The Committee was concerned that the conceptual model seemed too medical-oriented and should be broadened to account for more public health variables. For example, the model did not address community, environmental, or patient factors (e.g., social supports, lack of money to buy medication, no refrigerator). The conceptual model should reflect resources available for care within individual hospitals. While these should not be included in the risk-adjustment approach, because differential resources can impact quality of care, they should be noted in the conceptual model.

2. **Additional literature review.** 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. Members also suggested that the developers do a broader search of literature to include readmissions and impact of SDS on health.

3. **Conceptual Relationships.** Based on the research performed by the developers, the Committee agreed there is a conceptual relationship between the selected variables and cost/payment outcomes.

In response to the Committee's recommendations (1 &2 above) in May, the developers performed additional literature review and revised their conceptual model (Appendix 1 of Yale Memo). Text from this memo summarizing the key findings from this additional research has been excerpted below:

- They reviewed 14 additional articles that examined within and between hospital differences in outcomes related to SDS variables:
  - "Taken together these papers do not present a conclusive or consistent picture about the role of within hospital differences in treatment of patients based on SDS nor the subsequent impact on outcomes or cost. However they provide some evidence that in certain settings differential care by race could contribute to differences in costs and outcome." (page 3)
  - "Taken together, the body of literature reveals an inconsistent and complex association of low SDS and health outcomes. Most studies used race as their independent variable with less attention to income or other measures of poverty (e.g. Medicaid status). The literature demonstrates both within and between hospital differences in outcomes among racial/ethnic groups that can be partially explained by the use of lower quality hospitals by minorities." (page 3)

---

**Standing Committee Discussion**:
1. Does the committee believe the developer has adequately supplemented their conceptual analysis based on previous Committee recommendations?
2. Does the conceptual model adequately reflect the impact of SDS factors in the episode of care that is captured in these measures?

---

## Empirical Analyses Review

NQF guidance for the evaluation of SDS factors states that if the Committee believes a conceptual relationship exists between the sociodemographic factor(s) and the outcome (i.e, resource utilization or cost), developers should conduct empirical analyses to confirm that relationship.

During the May webinar, the Committee determined there is conceptual relationship between the proposed variables and the three cost outcomes. Their discussion yielded the following recommendations regarding the examination and consideration of these variables in empirical analyses:

- **Race:** The Committee recommended that the Yale team review the data and consider including other race variables beyond black.
- **Income and educational attainment**: The Committee was not in favor of the developers beginning empirical analysis using 5-digit zip code data. The Committee would prefer for the developers to use their resources analyzing the 9-digit zip code data once it is available to them.
- **Medicaid/dual eligibility status**: The Committee was in support of empirical analysis on this (Medicaid status) variable, but only in combination with the Low Income Subsidy (LIS) data as proxy for insurance status and income.

Current NQF guidance for the submission of empirical analysis of SDS factors in the risk model requires the submission of:

- Analyses and interpretation resulting in decision to include or not include SDS factors in section.
- Performance scores and risk model performance of the model with and without SDS factors included (including method and results).
- An interpretation of their results in terms of the differences in performance scores for the same entities.
- Submission of updated reliability and validity testing and specifications for a stratified version of the measure using these factors, if SDS factors are included in the risk-adjustment approach.

The importance of the SDS variables in the risk adjustment model should be evaluated by the size of the SDS coefficients in the risk adjustment model, the p-values associated with the SDS coefficients, and the impact of adjusting for the SDS variables on the measure results.  Reasons for including the SDS variables in the risk-adjustment approach include (1) demonstration of the contribution of the SDS factor(s) to unique variation in the outcome that is not due to between-unit effects, (2) adjustment leads to substantial differences between measure scores (although this doesn't have to result in change in rankings), or (3) if needed for face validity of the approach.

### *Summary: Key Findings Discussed in the Yale Empirical Analysis Memo*

**Variables Used in the Empirical Analysis:**

1. Race: Categorized as Black and Non-Black
2. Medicaid enrollment/Dual Status (as proxy for low income): Categorized as Medicaid and Non-Medicaid.

**Definitions (Table 1):**

- Identified conceptual relationship: The Committee's determination of whether the variables had a conceptual relationship to the cost outcomes.
- Variation in prevalence of SDS factors across entities: If the prevalence of a particular factor does not vary across the measured entities, then adjustment likely is not necessary. But if it does vary substantially, then there is reason to believe that one should potentially control for it in the risk-adjustment approach.
- Bi-variate relationship between SDS factors and outcome: If the SDS factor is associated with the outcome, then it is a potential confounder and may be a candidate for risk adjustment.
- Significant to the multivariable model: Analysis demonstrating that the SDS factor is statistically associated with the outcome of interest, after controlling for other (clinical) factors
- Good model fit: The risk-adjustment model adequately "reflects" the data.
- Impact on the risk model and measure scores with the inclusion of the SDS Factors: How the measure scores and risk model fit is impacted when you add in the SDS factors.
    o Improvement in model fit with the addition of variables (based on Quasi R2)
    o Change in distribution of scores:
    o Change in mean Risk Standardized Payment (RSP):
    o Impact on Risk Standardized Payment:
- Decision to include variables in the model: The developer's decision on whether or not to include the variables in the risk model.

**Table 1: Assessment of Empirical Analyses Results and Validity of the Risk Adjustment Approach (Applies to Black Race + Medicaid Variables)**

| Assessment of Validity of the Risk Model | AMI | HF | PN |
|---|---|---|---|
| *Identified conceptual relationship* | Yes | Yes | Yes |
| *Variation in prevalence of SDS factors across entities* | Yes (Tables 2 &3) | Yes (Tables 11 &12) | Yes (Table 20 &21) |
| *Bi-variate relationship* | Yes (Table 7) | Yes (Table 16) | Yes (Table 25) |
| *Significant to the model* | Yes (Table 8) | Yes (Table 17) | Yes (Table 26) |

| Assessment of Validity of the Risk Model | AMI | HF | PN |
|---|---|---|---|
| *Good model fit* | Yes (Table 9) | Yes (Table 18) | Yes (Table 27) |
| *Impact on the risk model and measure scores with the inclusion of the SDS Factors:* | | | |
| *Improvement in model fit with the addition of variables (based on Quasi $R^2$)* | Slight (Table 9) | Slight (Table 18) | None (Table 27) |
| *Change in distribution of scores* | Slight (Table 6) | Slight (Table 15) | Slight (Table 24) |
| *Change in Mean Risk Standardized Payment (RSP)* | None (Table 4) | $1 (Table 3) | $1 (Table 22) |
| *Impact on RSP* | Black: Lower payment: beta: -0.058) | Black: Lower payment: beta: -0.030) | Black: Higher payment: $287 |
| | Medicaid: Lower payment (beta: -0.017) (Appendix 3) | Medicaid: Higher payment (beta: 0.012) (Appendix 4) | Medicaid: Higher payment: $496 (Appendix 5) |
| *Decision to include variables in the model* | No | No | No |

---

**Standing Committee Discussion**:
3. Does the committee believe the developer has adequately demonstrated validity of their risk adjustment approach?
4. Has the developer adequately supported their decision to not include the SDS variables in the risk model?

**Standing Committee Action**:

1. Submit your votes on validity and endorsement with rationale. *(Please only submit your vote ON or AFTER the October 27 webinar)*

   a. Validity: High, Moderate, Low or insufficient

      i. Consider the other validity sub-criteria when submitting your votes (See *Appendix A* for Resource Use Measure Evaluation Criteria for Validity and *Appendix B* for prior voting results on validity).

   b. Recommendation for endorsement:

      i. Recommend for [continued] endorsement (without the inclusion of SDS) factors in the risk model

      ii. Do not recommend for [continued] endorsement

## Appendix A: *Resource Use Measure Evaluation Criteria (Validity Only)*
**2b. Validity**

2b1. The measure specifications are consistent with the measure intent described under criterion 1c and captures the most inclusive target population.

2b2. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the cost of care or resources provided.

2b3. Exclusions are supported by the clinical evidence.

AND/OR

There is a rationale or analysis demonstrating that the measure results are sufficiently distorted due to the magnitude and/or frequency of the non-clinical exclusions;

AND

− Measure specifications for scoring include computing exclusions so that the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);

AND

− If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b4. For resource use measures and other measures when indicated:

− an evidence-based risk-adjustment strategy (e.g., risk models, risk-stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not factors related to disparities in care or the quality of care) and are present at start of care and has demonstrated adequate discrimination and calibration OR

− rationale/data support no risk-adjustment/-stratification.

2b5. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful13 differences in performance.

2b6. If multiple data sources/methods are specified, there is demonstration that they produce comparable results.

2c. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender) OR rationale/data justifies why stratification is not necessary or not feasible.

## Appendix B: *Summary of Committee Deliberations of Scientific Acceptability (Initial Endorsement)*

_____

### 2579 Hospital-level, risk-standardized payment associated with a 30-day episode of care for pneumonia

**Description**: This measure estimates hospital-level, risk-standardized payment for a pneumonia episode of care starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of pneumonia.

**Resource Use Measure Type:** Per episode

**Level of Analysis:** Facility

**Costing Method:** Standardized pricing

**Target population:** Senior Care

**Data Source**: Administrative claims

**Measure Steward**: Centers for Medicare & Medicaid Services

**STANDING COMMITTEE MEETING [06/25/2014]**

**2. Scientific Acceptability of Measure Properties: The measure meets the Scientific Acceptability criteria**

*(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)*

2a. Reliability: **H-10**; **M-11**; **L-1**; **I-0** 2b. Validity: **H-3**; **M-18**; **L-1**; **I-0**

Rationale:

• The Committee stated that the measure specifications were precise and that the measure was well-constructed. This measure captures risk-standardized payments for a thirty-day episode of care for Medicare patients diagnosed admitted to the hospital with a diagnosis of pneumonia through administrative claims data.

• The developer provided reliability testing at the level of the performance measure score; testing was performed by calculating the Intraclass Correlation Coefficient (ICC) score by calculating the risk standardized payment using a split-sample of the combined 2008-2009 data from hospitals. The ICC score was 0.825, indicating significant agreement between the two samples, which the Committee found sufficient.

• The Committee questioned the validity of specifying the measure for a thirty-day episode triggered by admission for pneumonia, as the treatment of pneumonia may require care coordination post-discharge that may extend past thirty days. The Committee stated that this could affect payments captured during the post-discharge period, artificially inflating or deflating the costs for some patients simply because of the construct of the measure.

• The Committee raised concerns regarding the attribution approach and the implications for attribution of costs if a patient were transferred to another hospital. The developer clarified that only 0.4 percent of cohorts are transferred for pneumonia, which represents a small number of beneficiaries.

In the case of transfer patients, costs for the patient will be attributed to the initial admitting hospital, as hospitals are increasingly responsible for care delivered up to 30 days after discharge. The Committee found this approach to attribution to be acceptable.

• The Committee stated concern that the low r-squared value (.07) for the risk model may indicate that case mix is not being appropriately adjusted for through the risk model. The developer clarified that at lower patient volumes, there is less certainty when estimating cost. The measure uses a continuous outcome which results in a more accurate estimate than would result from a binary outcome. Additionally, the measure uses hierarchical risk modeling that adjusts hospitals with low patient volume towards the mean. The Committee found this explanation to be sufficient.

• The Committee questioned whether adjustments for sociodemographic status (SDS) factors should be incorporated into the risk adjustment model. NQF clarified that it is in the early stages of reviewing our policy on risk adjusting for SDS factors. The recommendations for modifying NQF's current policy on adjusting for SDS factors have not yet been finalized. As such, we ask that Committees continue to evaluate measures according to our current guidelines, that SDS factors are not included in the risk adjustment model, but are used to stratify the measure. If in the future the recommendations for adjusting for SDS factors become NQF policy, measures that may be improved from incorporating these adjustments will be updated and reviewed by the Committee through one of NQF's measure maintenance processes.

_____

**2431 Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for Acute Myocardial Infarction (AMI)\***

**Submission | Specifications**

**Description**: This measure estimates hospital-level, risk-standardized payment for an AMI episode-of-care starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of AMI.

**Resource Use Measure Type:** Per episode
**Level of Analysis:** Facility
**Costing Method:** Standardized pricing
**Target Population:** Senior Care
**Data Source:** Administrative Claims
**Measure Steward:** Centers for Medicare and Medicaid Services
**STANDING COMMITTEE MEETING [March 4-5, 2014]**

_____

**2. Scientific Acceptability of Measure Properties**
*(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)*
2a. Reliability: **H-3; M-16; L-2; I-0** 2b. Validity: **H-0; M-9; L-7; I-4**
Rationale:

• The Committee raised concern about the ability to assess performance of low volume hospitals given the hierarchical modeling approach and the potential implications it could pose for the reliability and validity of the measure. The developers responded that at lower patient volumes, the less certainty you have about your estimates for cost. This measure uses a continuous outcome so the estimate is more accurate than a binary outcome.

Additionally, this measure uses hierarchical risk modeling that adjusts hospitals with low patient volume towards the mean. Furthermore, reporting is only done for hospitals that have 25 or more cases.

• The Committee further questioned the decision to attribute the entire cost of an episode to the initial hospital in the case of a transfer to another facility. The developers responded that the decision was made not to exclude these cases because transfers account for approximately 8 percent of AMI episodes. This represented too many cases to exclude. Furthermore, the initial hospital begins the episode of care and can have a great influence over the coordination of care.

• The Committee raised concerns about whether the supplied reliability testing was done with the amount of data required by the specification of the measure. The measure is specified for a 12-month period and the testing used combined 2008 and 2009 data. The developers responded that the measure will eventually be implemented with three years of data but when the testing was performed, only two years of data was available. The decision to include three years of data was made to include as many hospitals in the measurement as possible. Many hospitals do not have 25 AMI cases in a year and would therefore not meet the threshold for reporting.

• In addition to the risk adjustment provided in the overarching issues section, the Committee was concerned that the developer did not do empiric measure-level validity testing for the measure as specified. The developers acknowledged that they relied on prior research on risk adjustment testing for mortality measures and also relied on face validity testing with their technical expert panel.

_____

## 2436 Hospital-level, risk-standardized payment associated with a 30-day episode-of-care for heart failure (HF)

**Description**: This measure estimates hospital-level, risk-standardized payment for a HF episode of care starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of HF.

**Resource Use Measure Type:** Per episode
**Level of Analysis:** Facility
**Costing Method:** Standardized pricing
**Target Population:** Senior Care
**Data Source:** Administrative Claims
**Measure Steward:** Centers for Medicare and Medicaid Services
**STANDING COMMITTEE MEETING [March 4-5, 2014]**
_____

### 2. Scientific Acceptability of Measure Properties
*(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)*
2a. Reliability: **H-7; M-11; L-2; I-1** 2b. Validity: **H-0; M-9; L-6; I-5**
Rationale:
• The Committee questioned the description of a "typical heart failure" patient considering that many patients have chronic heart failure and a hospitalization occurs for an acute incidence of the disease.

The developer responded that they meant non-LVAD, non-transplant, non-major surgical procedure heart failure patients. These conditions dramatically change the payment outcome. They are sicker patients and were excluded from the measure.

• The Committee also questioned the methodology for choosing the index admission for patients who might have multiple hospitalizations in the same year for heart failure. The developer responded that the hospitalization is randomly selected and any re-hospitalization within 30 days of that index admission would be considered a re-admission and counted in the total hospitalization cost.

• The Committee expressed concern that attributing costs to hospitals was inappropriate for heart failure patients and that the real accountability should be with the ambulatory providers. Furthermore, the 30-day time period for costs does not align with the typical disease progression for a heart failure patient. A longer period, perhaps 12 months, would be more appropriate for the chronic nature of this disease.

• The developer defended the attribution to the hospital by stating that heart failure is a leading cause of hospitalization for the elderly and it represented a high-leverage opportunity to measure and evaluate spending. Additionally, the 30-day time period was short enough that the associated spending would be attributable to the hospital admission.

• In addition to the risk adjustment discussion provided in the overarching issues section, the Committee was concerned that the developer did not do empiric measure-level validity testing for the measure as specified. The developers acknowledged that they relied on prior research on risk adjustment testing for mortality measures and also relied on face validity testing with their technical expert panel.

_____

# MEMORANDUM

**TO:** NQF Cost and Resource Use Standing Committee
**FROM:** Nancy Kim and Susannah Bernheim, Yale New Haven Health Services Corporation - Center for Outcomes Research and Evaluation (CORE)
**THROUGH:** Lein Han, CMS
**DATE:** Monday, October 5, 2015
**SUBJECT**: Empiric analyses for payment measures in the sociodemographic status (SDS) trial period

The National Quality Forum (NQF) endorsed three payment measures, developed by the Center for Outcomes Research and Evaluation (CORE) under contract with the Centers for Medicare & Medicaid (CMS), in 2014 with the consideration that additional testing and analyses, focusing on sociodemographic status (SDS) risk factors, is performed and considered under an ad hoc review process by the NQF Cost and Resource Use Standing Committee (hereinafter "Standing Committee"). Consequently, NQF staff proposed, evaluated, and finalized a two-staged process that includes the following: Webinar #1 conceptual analysis and determination of SDS variables to be used in empiric analysis and Webinar #2 empiric analysis.

Webinar #1 took place on May, 21, 2015. During this meeting Yale presented an initial conceptual framework for the causal pathways by which SDS may influence episode payments and presented the variables that we planned to use in the empiric analysis. In response to Webinar #1, the Standing Committee recommended that Yale broaden the conceptual model and literature review to determine if there were other variables that would merit consideration in risk-adjustment. They further suggested that Yale consider a number of variables including race and Medicaid status. With regard to race, the Standing Committee specifically felt that further literature review was necessary "to determine the within and between effects of race on hospital performance."[i]

In this memo, we provide:

1) A revised conceptual model;

2) A summary of the expanded literature review with a particular focus on within and between effects of race on health outcomes; and

3) Empiric analyses that examine the effect of adjustment for race (black/non-black) and Medicaid status on hospital-level risk-standardized payments for the three payment measure conditions (acute myocardial infarction [AMI], heart failure [HF], and pneumonia [PN]).

---

[i] "Meeting Summary: Cost and Resource Use Standing Committee Webinar: Discussion of Conceptual Relationships between SDS Variables and Payment Outcomes: May 21, 2015 (2-4pm ERT)" via correspondence with Ashlie Wilbon, NQF, June 2015.

This memo and conceptual model pertain to the following three measures:

- 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)

- NQF #2579: Hospital-level, risk-standardized payment associated with a 30-day episode of care pneumonia (CMS/Yale)

## SECTION 1. UPDATED CONCEPTUAL APPROACH TO RISK-ADJUSTMENT

For Webinar #1, we presented a conceptual framework with which we approached risk-adjustment. The conceptual model was not intended to be an exhaustive overview of the many ways that SDS factors can affect population health but rather a focused view of the hospital experience guided by a narrow question of whether or not to risk adjust our payment measures. The Standing Committee provided feedback on the conceptual model and felt that the model should be broadened to account for more public health variables. Specifically, they felt the model should: 1) more explicitly include community, environmental, or patient factors, 2) differentiate lack of patient resources from lack of community resources, 3) reflect resources available for care within individual hospitals, and 4) change "patient behavior" title as it seemed to blame patients for poor outcomes.

In response to the above recommendations, we have made a number of modifications to highlight the context of the patient's community and larger environment both before and after admission to the hospital (Appendix 1). We have changed the titles within pre-admission and post-discharge to capture the many patient and community factors that reflect differential SDS and can impact episode of care payments. We have changed the title "patient behavior" to patient factors in the post-discharge setting. Additionally, given the chance to revisit the model, we chose to reorient the model to emphasize the potential pathways by which low SDS may be exerting influence on care provided by hospitals that may be captured in the episode of care payments.

## SECTION 2. UPDATED LITERATURE REVIEW

For Webinar #1, we performed three focused literature reviews to examine the relationship between socioeconomic status (SES) and/or sociodemographic factors and costs associated with AMI, HF, or pneumonia care. In the course of the discussion of the conceptual model and the potential pathways by which low sociodemographic status (SDS) may exert influence on health outcomes or payments, the Standing Committee asked for an expanded literature review to determine: 1) whether other SDS factors should be considered for risk-adjustment and 2) the within and between effects of race on hospital performance.

In response to NQF Webinar #1, we further examined the medical literature that assessed the association of SDS and health outcomes beyond cost and payment, with a particular focus on the

hospital's role in healthcare outcomes. We reviewed a number of studies shared by Standing Committee member, Dr. Andrew Ryan as well as additional relevant articles.

In total, we evaluated 14 articles (Appendix 2). Based on the Standing Committee's recommendations, we organized the articles conceptually by their focus on the following categories: 1) "within hospital" papers; i.e. those examining differences in quality or outcome between populations of different SDS cared for in the same institution, and 2) "between hospital" differences between populations of different SDS, i.e. papers examining whether minorities or patients of low socioeconomic status are cared for at lower quality hospitals based on outcomes.

Among the four articles that examined only within hospital differences in outcomes, all used race/ethnicity as their independent variable.[1-5] These articles had mixed findings and most focused on cardiovascular diseases and procedures. Gaskin et al. found that whites and minorities received the same quality of care as measured by AHRQ inpatient quality (mortality after certain procedures or conditions) and safety (complications and adverse events following surgeries and procedures) indicators.[1] In contrast, Schulman et al found that race and gender, independent of clinical factors, influenced physician management of chest pain.[2] Similarly Chen et al. and Epstein et al. found that blacks received fewer invasive procedures for AMI that whites.[3-5] Further, Chen et al. exposed the complexity of this relationship demonstrating that despite differences in the receipt of intermediate outcomes such as invasive procedures, mortality between blacks and whites was no different.[4,5] Taken together these papers do not present a conclusive or consistent picture about the role of within hospital differences in treatment of patients based on SDS nor the subsequent impact on outcomes or cost. However they provide some evidence that in certain settings differential care by race could contribute to differences in costs and outcome.

To gain further insights into the interplay between the hospital and SDS variables, we reviewed nine articles that examined between hospital differences in outcomes.[6-14] Among these, several also examined within hospital differences. Two focused specifically on the use of lower quality hospitals by minorities.[9,10] Eight analyzed the effect of race while only one, Reames et al.,[6] used Census derived SES data as their independent variable. In general, these articles consistently found both within and between hospital differences across a breadth of outcomes including inpatient perioperative complications, cardiovascular procedures, readmission rates, and mortality rates. Helland et al. found that black race was associated with reduced spending for heart conditions, but the race effect went away after considering blacks from out of state, suggesting that local care was responsible for disparities.[7] The two articles that examined the use of lower quality hospitals by minorities offered different conclusions. Gaskin et al. examined black, Hispanic, and Asians,[10] while Dimick et al. looked only at blacks versus whites.[9] Gaskin et al. found that minorities do not necessarily use lower quality hospitals and this type of characterization will differ based on the choice of quality indicator.[10] Dimick et al. found a strong relationship between racial segregation and use of low-quality hospitals.[9]

**Summary:** Taken together, the body of literature reveals an inconsistent and complex association of low SDS and health outcomes. Most studies used race as their independent variable with less attention to income or other measures of poverty (e.g. Medicaid status). The literature demonstrates both within and between hospital differences in outcomes among racial/ethnic groups that can be partially explained by the use of lower quality hospitals by minorities.

**SECTION 3. EMPIRIC ANALYSES**

To investigate the impact of adding SDS variables to the episode of care payment models for the three conditions, we performed a number of empiric analyses. We used the same strategy for each condition. To facilitate review of our results, we briefly outline our step-wise approach after we selected the SDS variables.

We began by examining the prevalence of SDS variables in each cohort (AMI, HF, pneumonia) and the distribution of SDS variables across hospitals. We then analyzed the bivariate relationship between SDS variables and total payment at the patient level as well the impact on other covariates when SDS variables were added. We then performed a number of model diagnostics to assess model performance when SDS variables were added and further considered the model performance among subgroups (black/non-black and Medicaid/non-Medicaid). We then evaluated the correlation of SDS variables and finally investigated the hospital risk-standardized payment when SDS variables added

### 3.1 Variable Selection

We selected SDS variables based on our literature review and their availability in national data sources. Few SDS variables are available nationally at a patient-level and can be linked to Medicare data. We settled on race and a variable that would proxy for low-income as the only two feasible patient-level SDS variables to examine directly. This was discussed with the NQF Cost and Resource use Committee during Webinar #1. The Committee was in favor of this approach and did not recommend the use of the 5-digit zip code data as a proxy for low SDS.

### 3.2 Description of Variables

The two SDS variables we chose to use for race and low-income in the empiric analyses are found in CMS administrative claims data.

For the race variable, we used Beneficiary Race Code. This has been used for CORE disparities analyses in the past and is present in the data used for 2015 production of the payment measures. We considered creating categorizations of black/white/other or black/white/other/Hispanic, but data from CMS suggests that black and white are the only categories with both high sensitivity and specificity in the Beneficiary Race Code variable (Table 1). Therefore, we created an indicator variable for black/non-black for use in our empiric analyses.

**Table 1. Medicare Race and Ethnicity Data Validation Results[α]**

| Racial/Ethnic Classification | Accuracy Measures for CMS EDB | | | | |
|---|---|---|---|---|---|
| | Sensitivity | Specificity | PPV | NPV | Kappa |
| White | 99.3 | 92.9 | 98.7 | 96.1 | 0.93 |
| Black | 98.2 | 99.6 | 96.8 | 99.8 | 0.97 |
| Hispanic | 28.6 | 99.9 | 96.7 | 94.2 | 0.42 |

---

[α] Source: Validating Medicare's Race and Ethnicity Data. Kimberly Proctor and Carla Hodge. CMS, Office of Minority Health. (Using 2010 and 2000 Census data and 2011-2009 American Community Survey)

| | Accuracy Measures for CMS EDB | | | | |
|---|---|---|---|---|---|
| Asian/Pacific Islander | 57.4 | 99.8 | 91.4 | 98.6 | 0.70 |
| American Indian/Alaska Native | 54.3 | 99.9 | 71.9 | 99.7 | 0.62 |
| Other | 15.7 | 98.3 | 0.6 | 99.9 | 0.01 |

As a proxy for low-income, we chose Medicaid enrollment. After discussion with Committee members who have used a different approach to define low-income status, we performed analyses among potential low-income proxy variables and chose to use Dual Status Code. Specifically, where Dual Status Code equaled 01, 02, 03, 04, 05, 06, or 08 (indicating full or partial Medicaid benefits) we categorized patients as "Medicaid". All other patients were categorized as "Non-Medicaid." The Dual Status Code variable has been used for CORE payment measure disparities analyses in the past. Ultimately, we chose to use the Dual Status Code variable as a marker of poverty because we felt it best reflected those with the lowest income. Additionally, CCW and ResDAC technical guidelines suggest that this variable best captures beneficiaries that are dual eligible (i.e. also receiving Medicaid benefits).

CORE currently has access to the race variable for the full current public reporting periods (July 2011 to June 2014), however we only have access to the Medicaid variable for July 2011 to December 2013. To maintain consistency in these empiric analyses, we chose to use the two and a half year span from July 2011 to December 2013 for all analyses.

## 3.3 Methods

We began our analyses on the effect of the addition of SDS variables to the payment models at the patient level. This is aligned with our general approach to considering clinical risk-adjustment variables for inclusion during measure development. We examined changes in predicted total episode standardized payments with the addition of the SDS variables at the patient level and computed summary statistics to assess model performance: predictive ratios by deciles and top 1% of predicted payment and quasi-$R^2$.[v] We compared residuals for subgroups of patients (i.e. black/non-black; Medicaid/non-Medicaid) to discern whether the addition of SDS variables produced a better model fit for those subgroups. We also investigated whether the SDS variables were strongly correlated (i.e. collinear) with one another or with any other risk-adjustment variables.

We then conducted analyses to examine whether hospital-level RSPs were affected by the inclusion of SDS variables in the payment model. We compared the distributions of RSPs calculated with and without the addition of each SDS variable. We used a Spearman rank correlation coefficient to examine whether the ranking of hospitals' RSP estimates shifted with the addition of the SDS variables. We also examined the percent change in same-hospital RSPs with the addition of the SDS variables.

---

[v] A predictive ratio is an estimator's ratio of predicted outcome to observed outcome. A predictive ratio of 1.0 indicates an accurate prediction. A ratio greater than 1.0 indicates overprediction, and a ratio less than 1.0 indicates underprediction. The quasi-$R^2$ calculated is the $R^2$ from a regression of observed outcome on the predicted outcome. (Reference: Jones AM. Models for Health Care. Health, Econometrics and Data Group (HEDG) Working Papers. 2010.)

### 3.4 Summary of Results

To facilitate discussion of the results, we are providing a summary of key findings. We begin with the results of the hospital-level RSP analyses since the question posed by NQF is whether the measure outcome, or RSP, is affected by the addition of SDS variables. We then provide patient-level results. Additional, detailed results are available on request.

**Overall we found that there is minimal association between race and/or Medicaid status and the episode of care payment.** Specifically, the addition of the race and/or the Medicaid variable to the payment model had little to no effect on hospital RSPs (Tables 4; 13; 22) for all three payment measure conditions. The Spearman rank correlation coefficients showed very little difference between the ranking of RSPs produced by original model and those produced by the model that included the SDS variables. The overall change in RSPs was minimal. The average percent change in same-hospital RSPs for all three of the payment measures was effectively 0 after adding the SDS variable(s) (Tables 5-6; 14-15; 23-24). Thus, the impact of these SDS variables was small to negligible on hospital profiling.

At the patient level, for most variable/condition combinations, we saw little change in predicted total episode payments when the SDS variables were added to the model. However, the relationship between AMI payment and race was slightly more substantial. Though the coefficients for both race and Medicaid were statistically significant for all models, the large sample size should be taken into account when considering significance. Moreover, for some variable/condition combinations, the relationship between the SDS variables and payment is in the opposite direction than what has been the expressed concern of stakeholders interested in adding such adjustment to the models. For example, the relationship between AMI payment and race indicates that an AMI episode of care is significantly less expensive for black patients than for non-black patients.

We also found that the impact of the SDS variables was small to negligible on model performance. There was no appreciable difference in the coefficients of the other risk-adjustment variables after adding race and/or Medicaid or race and Medicaid and an interaction between race and Medicaid to the model. Similarly, the results of the collinearity analyses did not suggest that there is a significant relationship between the race and Medicaid variables or between these variables and any other risk-adjustment variables in the model.

### 3.5 Conclusion

The addition of race and/or Medicaid status had a negligible impact on hospital RSPs, and are therefore unlikely to affect hospital profiling. Additionally, the direction of the association of race and/or Medicaid status (whether their addition increased or decreased predicted total episode payment) was not consistent across the three conditions (AMI, HF, pneumonia).

**3.5 Condition-Specific Results: AMI Payment**

**3.5.1 Cohort**

The AMI payment cohort, based on July 2011-December 2013 data, included 379,923 index admissions (Table 2). Among black patients, 41.86% were also receiving Medicaid benefits and among Medicaid patients, 17.31% were black. 3.33% of the total index admissions were both black and Medicaid patients. This indicates that there is not substantial overlap in the populations captured by these variables.

**Table 2. Prevalence of SDS factors in AMI payment measure cohort**

| Variable category | Number of index admissions (% of total index admissions) |
|---|---|
| Black | 30,249 (7.96%) |
| Non-Black | 349,674 (92.04%) |
| Medicaid | 73,151 (19.25%) |
| Non-Medicaid | 306,772 (80.75%) |
| Black and Medicaid | 12,663 (3.33%) |

The prevalence of SDS factors in the AMI payment cohort varies substantially across hospitals (Table 3). For hospitals whose AMI payment measure results are publicly reported (i.e. those with at least twenty-five cases during the measurement period) the percent of index admissions that were black ranged from 0% to 100%, with a mean of 8.6% and a median of 3.4% (IQR 0% to 10.3%). For these hospitals, the percent of index admissions that were receiving Medicaid benefits ranged from 0% to 91.9%, with a mean of 21.5 % and a median of 17.7% (IQR 11.8% to 26.8%).

**Table 3. Distribution of percent Black and Medicaid index admissions in AMI payment measure cohort across hospitals**

| Distribution | % Black index admissions | | % Medicaid index admissions | |
|---|---|---|---|---|
| | All hospitals (N=4,287) | Hospitals with at least 25 cases (N=2,287) | All hospitals (N=4,287) | Hospitals with at least 25 cases (N=2,287) |
| Minimum | 0 | 0 | 0 | 0 |
| 10th Percentile | 0 | 0 | 0 | 8.1 |
| 25th Percentile | 0 | 0.71 | 10.8 | 11.8 |
| Median | 0.67 | 3.4 | 20.0 | 17.7 |
| Mean | 7.5 | 8.6 | 25.6 | 21.5 |
| 75th Percentile | 7.7 | 10.3 | 34.7 | 26.8 |
| 90th Percentile | 23.0 | 22.9 | 53.8 | 40.0 |
| 99th Percentile | 77.8 | 67.3 | 100 | 75.9 |
| Maximum | 100 | 100 | 100 | 91.9 |

### 3.5.2 Hospital-Level Results

For AMI, the addition of the race variable had virtually no impact on the overall distribution of hospital-level RSPs (Table 4). Moreover, a Spearman rank correlation coefficient[†] (rho) of 0.997 (p<=0.0001) indicated that the ranking of hospital RSPs were virtually unchanged after the addition of the race variable. The findings were similar for both the addition of the Medicaid variable alone as well as the addition of the race and Medicaid variable together. The Spearman rank correlation coefficients were 0.999 (p<=0.0001) and 0.996 (p<=0.0001), respectively.

Roughly 97% of hospitals showed less than a 1% change in RSP when the race variable was added to the risk-adjustment model (Table 5). The maximum change for any hospital was 5% (Table 6). The finding was similar for the addition of the Medicaid variable alone, where nearly all hospitals showed less than a 1% change in RSP and the maximum change was 1% as well as for the addition of the race and Medicaid variable together, where roughly 97% of hospitals showed less than a 1% change in RSP and the maximum change was 5%.

**Summary:** These analyses indicate that the addition of the race and/or the Medicaid variable to the risk-adjustment model had little to no effect on overall and same-hospital RSPs.

**Table 4. AMI RSPs calculated with the current model vs. the current model with the addition of SDS variables**

| Variables included in model | Number of Hospitals | Minimum | Median | Mean | Standard Deviation | Maximum |
|---|---|---|---|---|---|---|
| Current[*] | 4,287 | $13,760 | $21,635 | $21,806 | $1,350 | $29,594 |
| Current + Black | 4,287 | $13,688 | $21,642 | $21,807 | $1,359 | $29,547 |
| Current + Medicaid | 4,287 | $13,712 | $21,634 | $21,806 | $1,349 | $29,537 |
| Current + Black + Medicaid | 4,287 | $13,655 | $21,641 | $21,806 | $1,358 | $29,507 |

**Table 5. Percent change in AMI RSPs calculated with the addition of SDS variables to the current model**

| Variables included in model | % Change in RSP | Number of Hospitals | % of Hospitals |
|---|---|---|---|
| Current[*] + Black | - 1[+] | 0 | 0 |
| | - 0-1 | 3,326 | 77.6 |
| | 0-1 | 829 | 19.3 |
| | 1[+] | 132 | 3.1 |
| Current + Medicaid | - 1[+] | 0 | 0 |
| | - 0-1 | 2,441 | 56.9 |
| | 0-1 | 1,843 | 43.0 |
| | 1[+] | 3 | 0.1 |
| Current + Black + Medicaid | - 1[+] | 0 | 0 |
| | - 0-1 | 3,043 | 71.0 |
| | 0-1 | 1,099 | 25.6 |
| | 1[+] | 145 | 3.4 |

---

[†] A Spearman correlation coefficient, rho, equal to 1 indicates that the ranking of hospital RSPs were virtually unchanged after the addition of the race variable

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

**Table 6. Distribution of percent change in AMI RSPs with addition of SDS variables to the current model**

| Distribution | Current[*] + Black (% RSP Change) | Current + Medicaid (% RSP Change) | Current + Black + Medicaid (% RSP Change) |
|---|---|---|---|
| Minimum | -0.53 | -0.38 | -0.76 |
| 10th Percentile | -0.31 | -0.18 | -0.36 |
| 25th Percentile | -0.19 | -0.087 | -0.2 |
| Median | -0.064 | -0.013 | -0.054 |
| Mean | 0.00084 | 0.00013 | 0.00093 |
| 75th Percentile | -0.0079 | 0.054 | 0.021 |
| 90th Percentile | 0.34 | 0.17 | 0.37 |
| 99th Percentile | 1.91 | 0.58 | 1.88 |
| Maximum | 5.06 | 1.11 | 5.00 |

### 3.5.3 Patient-Level Results

While our primary goal was to assess the effect of the addition of SDS variables on hospital RSPs, we began our analyses, as we do when building any outcome measure, at the patient level. The results of these analyses may help to inform the hospital-level results seen above.

We began by assessing the individual relationship between the payment outcome and the SDS variables. A bivariate regression (i.e. a regression of total episode payment on the individual SDS variables) indicated little to no effect on total episode payment (Table 7).

**Table 7. Bivariate relationship between total AMI payment and SDS variables[ᵗ]**

| Variables included in model | Estimate | Payment Ratio[ʳ] | p-value |
|---|---|---|---|
| Black | 0.0085 | 1.01 | 0.0261 |
| Medicaid | 0.0048 | 1.00 | 0.0657 |

The bivariate regression between total payment and the race variable estimated that payments were, on average, roughly 1% higher for black index admissions than for non-black index admissions (payment ratio of 1.01). The bivariate regression between total payment and the Medicaid variable estimated that payments were, on average, roughly the same as for non-Medicaid index admissions (payment ratio of 1.00).

We then added the SDS variables to the current risk-adjustment model, both individually and jointly (Table 8). With the sole addition of race to the patient-level model, black index admissions were estimated to be roughly 6% less expensive than non-black index admissions (payment ratio of 0.94). With the sole addition of Medicaid status to the patient-level model, Medicaid index admissions were estimated to be 2% less expensive than non-Medicaid index admissions (payment ratio of 0.98). With the addition of both race and Medicaid status to the model, the same individual results hold. The interpretation of these results is that, holding all other risk-

---

[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[ᵗ] Used log link and inverse Gaussian distribution as with current AMI payment measure patient-level model
[ʳ] Payment ratio is equal to exponentiated estimate

adjustment variables constant, both black and Medicaid patients are slightly less expensive for an AMI episode of care than non-black and/or non-Medicaid patients.

**Table 8. Relationship between total AMI payment and SDS variables[ŧ]**

| Variables included in model | Variable | Estimate | Payment Ratio[r] | p-value |
|---|---|---|---|---|
| Current[*] + Black | Black | -0.062 | 0.94 | <0.0001 |
| Current + Medicaid | Medicaid | -0.023 | 0.98 | <0.0001 |
| Current + Black + Medicaid | Black | -0.058 | 0.94 | <0.0001 |
| | Medicaid | -0.017 | 0.98 | <0.0001 |

### 3.5.4 Model Diagnostics

**Model Performance**

The addition of the SDS variables to the patient-level model did not significantly improve model performance compared the current model as demonstrated by the comparison of predictive ratios and quasi-$R^2$s ([Table 9](#)).

**Table 9. AMI payment model performance**

| Diagnostic[γ] | Current[*] | Current + Black | Current + Medicaid | Current + Black + Medicaid |
|---|---|---|---|---|
| Predictive Ratio, 1st Decile (lowest) | 0.9540 | 0.9546 | 0.9535 | 0.9537 |
| Predictive Ratio, 2nd Decile | 1.0021 | 1.0028 | 1.0031 | 1.0025 |
| Predictive Ratio, 3rd Decile | 1.0098 | 1.0113 | 1.0114 | 1.0128 |
| Predictive Ratio, 4th Decile | 1.0290 | 1.0221 | 1.0206 | 1.0207 |
| Predictive Ratio, 5th Decile | 1.0492 | 1.0506 | 1.0544 | 1.0525 |
| Predictive Ratio, 6th Decile | 1.0521 | 1.0540 | 1.0510 | 1.0555 |
| Predictive Ratio, 7th Decile | 1.0302 | 1.0274 | 1.0331 | 1.0258 |
| Predictive Ratio, 8th Decile | 0.9907 | 0.9962 | 0.9903 | 0.9951 |
| Predictive Ratio, 9th Decile | 0.9590 | 0.9553 | 0.9594 | 0.9584 |
| Predictive Ratio, 10th Decile (highest) | 0.9356 | 0.9363 | 0.9343 | 0.9340 |
| Quasi-$R^2$ | 0.0717 | 0.0722 | 0.0719 | 0.0724 |

---

[ŧ] Used log link and inverse Gaussian distribution as with current AMI payment measure patient-level model

[r] Payment ratio is equal to exponentiated estimate

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

[γ] A predictive ratio is an estimator's ratio of predicted outcome to observed outcome. A predictive ratio of 1.0 indicates an accurate prediction. A ratio greater than 1.0 indicates overprediction, and a ratio less than 1.0 indicates underprediction. The quasi-$R^2$ calculated is the $R^2$ from a regression of observed outcome on the predicted outcome. (Reference: Jones AM. Models for Health Care. Health, Econometrics and Data Group (HEDG) Working Papers. 2010.)

**Residual Analysis**

To assess whether the addition of the SDS variables produced a better model fit (i.e. improved the predicted total payment) for sub-groups of patients, we compared residuals[Φ] from the current model to the residuals after the SDS variables were added (Table 10). Model fit was improved, on average, for all sub-groups (though the addition of a variable is likely to produce this result for any model). The prediction for black patients improved more than the prediction for non-black patients. The same is true for Medicaid and non-Medicaid patients, though the magnitude of improvement is decreased.

**Table 10. Average residual for current model vs. current model with the addition of SDS variables for sub-groups of patients**

| Variable category | Current[*] ($)[ⁿ] | Current + Black + Medicaid ($)[ⁿ] |
|---|---|---|
| Black | -1,233 | -8 |
| Non-Black | 192 | 87 |
| Medicaid | -488 | -116 |
| Non-Medicaid | 214 | 126 |

**Collinearity Analyses**

With respect to assessing collinearity among the variables, the addition of the SDS variables did not increase the standard errors of the current risk-adjustment variables or appreciably alter their estimated coefficients (Appendix 3; full results available upon request). However, formal testing for collinearity among the risk-adjustments variables after the addition of the SDS variables was also performed (analyses available upon request). No indications of collinearity were found.

---

[Φ] The residual is the difference between the observed total payment and the predicted total payment
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[ⁿ] Values are in standardized dollars

### 3.6 Condition-Specific Results: HF Payment

### 3.6.1 Cohort

The HF payment cohort, based on July 2011-December 2013 data, included 736,511 index admissions (Table 11). Among black patients, 45.8% were also receiving Medicaid benefits and among Medicaid patients, 21.81% were black. 5.24% of the total index admissions were both black and Medicaid patients. This indicates that there is not substantial overlap in the populations captured by these variables.

**Table 11. Prevalence of SDS factors in HF payment measure cohort**

| Variable category | Number of index admissions (% of total index admissions) |
|---|---|
| Black | 84,311 (11.45%) |
| Non-Black | 652,200 (88.55%) |
| Medicaid | 176,849 (24.01%) |
| Non-Medicaid | 559,662 (75.99%) |
| Black and Medicaid | 38,579 (5.24%) |

The prevalence of SDS factors in the HF payment cohort varies substantially across hospitals (Table 12). For hospitals whose HF payment measure results are publicly reported (i.e. those with at least twenty-five cases during the measurement period) the percent of index admissions that were black ranged from 0% to 98.7%, with a mean of 10.2% and a median of 3.3% (IQR 0% to 12.9%). For these hospitals, the percent of index admissions that were receiving Medicaid benefits ranged from 0% to 95.2%, with a mean of 27.2% and a median of 22.9% (IQR 15.7% to 34.7%).

**Table 12. Distribution of percent Black and Medicaid index admissions in HF payment measure cohort across hospitals**

| Distribution | % Black index admissions | | % Medicaid index admissions | |
|---|---|---|---|---|
| | All hospitals (N=4,629) | Hospitals with at least 25 cases (N=3,536) | All hospitals (N=4,629) | Hospitals with at least 25 cases (N=3,536) |
| Minimum | 0 | 0 | 0 | 0 |
| 10th Percentile | 0 | 0 | 9.1 | 10.8 |
| 25th Percentile | 0 | 0 | 15.0 | 15.7 |
| Median | 2.0 | 3.3 | 23.2 | 22.9 |
| Mean | 9.5 | 10.2 | 27.5 | 27.2 |
| 75th Percentile | 11.3 | 12.9 | 36.4 | 34.7 |
| 90th Percentile | 29.4 | 29.9 | 51.4 | 49.3 |
| 99th Percentile | 80.2 | 76.8 | 90.0 | 81.1 |
| Maximum | 100 | 98.7 | 100 | 95.2 |

### 3.6.2 Hospital-Level Results

For HF, the addition of the race variable had virtually no impact on the overall distribution of hospital-level RSPs (Table 13). Moreover, a Spearman rank correlation coefficient[†] (rho) of 0.999 (p<=0.0001) indicated that the ranking of hospital RSPs were virtually unchanged after the addition of the race variable. The findings were similar for both the addition of the Medicaid variable alone as well as the addition of the race and Medicaid variable together. The Spearman rank correlation coefficients were 0.999 (p<=0.0001) and 0.996 (p<=0.0001), respectively.

Roughly 97% of hospitals showed less than a 1% change in RSP when the race variable was added to the risk-adjustment model (Table 14). The maximum change for any hospital was 2.6% (Table 15). The finding was similar for the addition of the Medicaid variable alone, where all hospitals showed less than a 1% change in RSP and the maximum change was 0.3% as well as for the addition of the race and Medicaid variable together, where roughly 97% of hospitals showed less than a 1% change in RSP and the maximum change was approximately 2.7%.

**Summary:** These analyses indicate that the addition of the race and/or the Medicaid variable to the risk-adjustment model had little to no effect on overall and same-hospital RSPs.

**Table 13. HF RSPs calculated with the current model vs. the current model with the addition of SDS variables**

| Variables included in model | Number of Hospitals | Minimum | Median | Mean | Standard Deviation | Maximum |
|---|---|---|---|---|---|---|
| Current[*] | 4,629 | $11,762 | $15,188 | $15,346 | $1,376 | $22,181 |
| Current + Black | 4,629 | $11,709 | $15,193 | $15,347 | $1,383 | $22,178 |
| Current + Medicaid | 4,629 | $11,722 | $15,193 | $15,347 | $1,379 | $22,230 |
| Current + Black + Medicaid | 4,629 | $11,727 | $15,198 | $15,347 | $1,388 | $22,240 |

**Table 14. Percent change in HF RSPs calculated with the addition of SDS variables to the current model**

| Variables included in model | % Change in RSP | Number of Hospitals | % of Hospitals |
|---|---|---|---|
| Current[*] + Black | - 1[+] | 0 | 0 |
| | - 0-1 | 3,319 | 71.7 |
| | 0-1 | 1,189 | 25.7 |
| | 1[+] | 121 | 2.6 |
| Current + Medicaid | - 1[+] | 0 | 0 |
| | - 0-1 | 2,050 | 44.3 |
| | 0-1 | 2,579 | 55.7 |
| | 1[+] | 0 | 0 |
| Current + Black + Medicaid | - 1[+] | 8 | 0.2 |
| | - 0-1 | 2,928 | 63.3 |
| | 0-1 | 1,564 | 33.8 |
| | 1[+] | 129 | 2.8 |

---

[†] A Spearman correlation coefficient, rho, equal to 1 indicates that the ranking of hospital RSPs were virtually unchanged after the addition of the race variable
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)

**Table 15. Distribution of percent change in HF RSPs with addition of SDS variables to the current model**

| Distribution | Current* + Black (% RSP Change) | Current + Medicaid (% RSP Change) | Current + Black + Medicaid (% RSP Change) |
|---|---|---|---|
| Minimum | -0.45 | -0.7 | -1.22 |
| 10th Percentile | -0.24 | -0.16 | -0.3 |
| 25th Percentile | -0.19 | -0.062 | -0.18 |
| Median | -0.094 | 0.014 | -0.059 |
| Mean | 0.00056 | 0.000087 | 0.00067 |
| 75th Percentile | 0.026 | 0.089 | 0.083 |
| 90th Percentile | 0.36 | 0.15 | 0.37 |
| 99th Percentile | 1.47 | 0.23 | 1.47 |
| Maximum | 2.59 | 0.29 | 2.74 |

### 3.5.3 Patient-Level Results

While our primary goal was to assess the effect of the addition of SDS variables on hospital RSPs, we began our analyses, as we do when building any outcome measure, at the patient level. The results of these analyses may help to inform the hospital-level results seen above.

We began by assessing the individual relationship between the payment outcome and the SDS variables. A bivariate regression (i.e. a regression of total episode payment on the individual SDS variables) indicated little to no effect on total episode payment (Table 16).

**Table 16. Bivariate relationship between total HF payment and SDS variables[‡]**

| Variables included in model | Estimate | Payment Ratio[r] | p-value |
|---|---|---|---|
| Black | 0.0129 | 1.01 | <0.0001 |
| Medicaid | 0.0576 | 1.06 | <0.0001 |

The bivariate regression between total payment and the race variable estimated that payments were, on average, roughly 1% higher for black index admissions than for non-black index admissions (payment ratio of 1.01). The bivariate regression between total payment and the Medicaid variable estimated that payments were, on average, roughly 6% for Medicaid index admissions as for non-Medicaid index admissions (payment ratio of 1.06).

We then added the SDS variables to the current risk-adjustment model, both individually and jointly (Table 17). With the sole addition of race to the patient-level model, black index admissions were estimated to be roughly 3% less expensive than non-black index admissions (payment ratio of 0.97). With the sole addition of Medicaid status to the patient-level model, Medicaid index admissions were estimated to be 1% more expensive than non-Medicaid index admissions (payment ratio of 1.01). With the addition of both race and Medicaid status to the model, the same individual results hold. The interpretation of these results is that, holding all other risk-

---

[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[‡] Used log link and Gamma distribution as with original HF payment measure patient-level model
[r] Payment ratio is equal to exponentiated estimate

adjustment variables constant, both black patients are slightly less expensive for an HF episode of care than non-black patients and Medicaid patients are slightly more expensive for an HF episode of care than non-Medicaid patients.

**Table 17. Relationship between total HF payment and SDS variables[t]**

| Variables included in model | Variable | Estimate | Payment Ratio[r] | p-value |
|---|---|---|---|---|
| Current[*] + Black | Black | -0.028 | 0.97 | <0.0001 |
| Current + Medicaid | Medicaid | 0.0080 | 1.01 | <0.0001 |
| Current + Black + Medicaid | Black | -0.030 | 0.97 | <0.0001 |
| | Medicaid | 0.012 | 1.01 | <0.0001 |

### 3.5.4 Model Diagnostics

#### Model Performance

The addition of the SDS variables to the patient-level model did not significantly improve model performance compared the current model as demonstrated by the comparison of predictive ratios and quasi-$R^2$s (Table 18).

**Table 18. HF payment model performance**

| Diagnostic[y] | Current[*] | Current + Black | Current + Medicaid | Current + Black + Medicaid |
|---|---|---|---|---|
| Predictive Ratio, 1st Decile (lowest) | 1.0273 | 1.0307 | 1.0271 | 1.0299 |
| Predictive Ratio, 2nd Decile | 1.0184 | 1.0182 | 1.0184 | 1.0175 |
| Predictive Ratio, 3rd Decile | 1.0062 | 1.0020 | 1.0061 | 1.0017 |
| Predictive Ratio, 4th Decile | 0.9903 | 0.9895 | 0.9903 | 0.9939 |
| Predictive Ratio, 5th Decile | 0.9918 | 0.9939 | 0.9920 | 0.9908 |
| Predictive Ratio, 6th Decile | 0.9868 | 0.9858 | 0.9873 | 0.9852 |
| Predictive Ratio, 7th Decile | 0.9834 | 0.9826 | 0.9848 | 0.9859 |
| Predictive Ratio, 8th Decile | 0.9840 | 0.9838 | 0.9813 | 0.9811 |
| Predictive Ratio, 9th Decile | 0.9871 | 0.9886 | 0.9879 | 0.9891 |
| Predictive Ratio, 10th Decile (highest) | 1.0300 | 1.0304 | 1.0301 | 1.0303 |
| Quasi-$R^2$ | 0.0399 | 0.0400 | 0.0399 | 0.0401 |

---

[t] Used log link and Gamma distribution as with original HF payment measure patient-level model
[r] Payment ratio is equal to exponentiated estimate
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[y] A predictive ratio is an estimator's ratio of predicted outcome to observed outcome. A predictive ratio of 1.0 indicates an accurate prediction. A ratio greater than 1.0 indicates overprediction, and a ratio less than 1.0 indicates underprediction. The quasi-$R^2$ calculated is the $R^2$ from a regression of observed outcome on the predicted outcome. (Reference: Jones AM. Models for Health Care. Health, Econometrics and Data Group (HEDG) Working Papers. 2010.)

**Residual Analysis**

To assess whether the addition of the SDS variables produced a better model fit (i.e. improved the predicted total payment) for sub-groups of patients, we compared residuals[Φ] from the current model to the residuals after the SDS variables were added (Table 19). Model fit was improved, on average, for all sub-groups (though the addition of a variable is likely to produce this result for any model). The prediction for black patients improved more than the prediction for non-black patients. The same is true for Medicaid and non-Medicaid patients, though the magnitude of improvement is decreased.

**Table 19. Average residual for current model vs. current model with the addition of SDS variables for sub-groups of patients**

| Variable category | Current[*] ($)[ʌ] | Current + Black + Medicaid ($)[ʌ] |
|---|---|---|
| Black | -330 | 33 |
| Non-Black | 38 | -9 |
| Medicaid | 85 | -4 |
| Non-Medicaid | -33 | -4 |

**Collinearity Analyses**

With respect to assessing collinearity among the variables, the addition of the SDS variables did not increase the standard errors of the current risk-adjustment variables or appreciably alter their estimated coefficients (Appendix 4; full results available upon request). However, formal testing for collinearity among the risk-adjustments variables after the addition of the SDS variables was also performed (analyses available upon request). No indications of collinearity were found.

---

[Φ] The residual is the difference between the observed total payment and the predicted total payment
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[ʌ] Values are in standardized dollars

Confidential-Do Not Distribute

### 3.7 Condition-Specific Results: PN Payment

### 3.7.1 Cohort

The PN payment cohort, based on July 2011-December 2013 data, included 740,244 index admissions (Table 20). Among black patients, 51.66% were also receiving Medicaid benefits and among Medicaid patients, 13.61% were black. 3.63% of the total index admissions were both black and Medicaid patients. This indicates that there is not substantial overlap in the populations captured by these variables.

**Table 20. Prevalence of SDS factors in PN payment measure cohort**

| Variable category | Number of index admissions (% of total index admissions) |
|---|---|
| Black | 51,966 (7.02%) |
| Non-Black | 688,278 (92.98%) |
| Medicaid | 197,142 (26.63%) |
| Non-Medicaid | 543,102 (73.37%) |
| Black and Medicaid | 26,844 (3.63%) |

The prevalence of SDS factors in the PN payment cohort varies substantially across hospitals (Table 21). For hospitals whose PN payment measure results are publicly reported (i.e. those with at least twenty-five cases during the measurement period) the percent of index admissions that were black ranged from 0% to 100%, with a mean of 6.9% and a median of 2.0% (IQR 0% to 8.0%). For these hospitals, the percent of index admissions that were receiving Medicaid benefits ranged from 0% to 100%, with a mean of 29.3 % and a median of 25.9% (IQR 18.4% to 36.7%).

**Table 21. Distribution of percent Black and Medicaid index admissions in PN payment measure cohort across hospitals**

| Distribution | % Black index admissions | | % Medicaid index admissions | |
|---|---|---|---|---|
| | All hospitals (N=4,670) | Hospitals with at least 25 cases (N=4,094) | All hospitals (N=4,670) | Hospitals with at least 25 cases (N=4,094) |
| Minimum | 0 | 0 | 0 | 0 |
| 10th Percentile | 0 | 0 | 11.1 | 12.2 |
| 25th Percentile | 0 | 0 | 17.8 | 18.4 |
| Median | 1.5 | 2.0 | 25.9 | 25.9 |
| Mean | 6.9 | 6.9 | 29.4 | 29.3 |
| 75th Percentile | 7.6 | 8.0 | 37.5 | 36.7 |
| 90th Percentile | 20.1 | 19.6 | 52.6 | 51.1 |
| 99th Percentile | 71.4 | 64.7 | 87.3 | 83.5 |
| Maximum | 100 | 100 | 100 | 100 |

### 3.5.2 Hospital-Level Results

For PN, the addition of the race variable had virtually no impact on the overall distribution of hospital-level RSPs (Table 22). Moreover, a Spearman rank correlation coefficient[†] (rho) of 0.999 (p<=0.0001) indicated that the ranking of hospital RSPs were virtually unchanged after the addition of the race variable. The findings were similar for both the addition of the Medicaid variable alone as well as the addition of the race and Medicaid variable together. The Spearman rank correlation coefficients were 0.999 (p<=0.0001) and 0.999 (p<=0.0001), respectively.

Roughly 99.9% of hospitals showed less than a 1% change in RSP when the race variable was added to the risk-adjustment model (Table 23). The maximum change for any hospital was 0.2% (Table 24). The finding was similar for the addition of the Medicaid variable alone, where roughly 96% of all hospitals showed less than a 1% change in RSP and the maximum change was approximately 1% as well as for the addition of the race and Medicaid variable together, where roughly 97% of hospitals showed less than a 1% change in RSP and the maximum change was approximately 1%.

**Summary:** These analyses indicate that the addition of the race and/or the Medicaid variable to the risk-adjustment model had little to no effect on overall and same-hospital RSPs.

**Table 22. PN RSPs calculated with the current model vs. the current model with the addition of SDS variables**

| Variables included in model | Number of Hospitals | Minimum | Median | Mean | Standard Deviation | Maximum |
|---|---|---|---|---|---|---|
| Current[*] | 4,670 | $9,171 | $14,249 | $14,299 | $1,396 | $23,661 |
| Current + Black | 4,670 | $9,182 | $14,245 | $14,299 | $1,392 | $23,647 |
| Current + Medicaid | 4,670 | $8,985 | $14,247 | $14,300 | $1,398 | $23,769 |
| Current + Black + Medicaid | 4,670 | $8,882 | $14,244 | $14,300 | $1,396 | $23,761 |

**Table 23. Percent change in PN RSPs calculated with the addition of SDS variables to the current model**

| Variables included in model | % Change in RSP | Number of Hospitals | % of Hospitals |
|---|---|---|---|
| Current[*] + Black | - 1[+] | 2 | 0.04 |
| | - 0-1 | 1,211 | 25.9 |
| | 0-1 | 3,457 | 74.0 |
| | 1[+] | 0 | 0 |
| Current + Medicaid | - 1[+] | 153 | 3.3 |
| | - 0-1 | 1,777 | 38.1 |
| | 0-1 | 2,740 | 58.7 |
| | 1[+] | 0 | 0 |
| Current + Black + Medicaid | - 1[+] | 175 | 3.8 |
| | - 0-1 | 1,714 | 36.7 |
| | 0-1 | 2,781 | 59.6 |
| | 1[+] | 0 | 0 |

---

[†] A Spearman correlation coefficient, rho, equal to 1 indicates that the ranking of hospital RSPs were virtually unchanged after the addition of the race variable

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

**Table 24. Distribution of percent change in PN RSPs with addition of SDS variables to the current model**

| Distribution | Current[*] + Black (% RSP Change) | Current + Medicaid (% RSP Change) | Current + Black + Medicaid (% RSP Change) |
|---|---|---|---|
| Minimum | -1.09 | -2.49 | -2.42 |
| 10th Percentile | -0.14 | -0.58 | -0.6 |
| 25th Percentile | -0.004 | -0.22 | -0.23 |
| Median | 0.048 | 0.088 | 0.094 |
| Mean | 0.0031 | 0.0059 | 0.0073 |
| 75th Percentile | 0.075 | 0.32 | 0.33 |
| 90th Percentile | 0.089 | 0.48 | 0.49 |
| 99th Percentile | 0.11 | 0.7 | 0.71 |
| Maximum | 0.19 | 0.95 | 0.98 |

### 3.5.3 Patient-Level Results

While our primary goal was to assess the effect of the addition of SDS variables on hospital RSPs, we began our analyses, as we do when building any outcome measure, at the patient level. The results of these analyses may help to inform the hospital-level results seen above.

We began by assessing the individual relationship between the payment outcome and the SDS variables. A bivariate regression (i.e. a regression of total episode payment on the individual SDS variables) indicated little to no effect on total episode payment (Table 25).

**Table 25. Bivariate relationship between total PN payment and SDS variables[ǂ]**

| Variables included in model | Estimate | Payment Ratio | p-value |
|---|---|---|---|
| Black | $1,708 | N/A | <0.0001 |
| Medicaid | $1,600 | N/A | <0.0001 |

The bivariate regression between total payment and the race variable estimated that payments were, on average, roughly $1,708 higher for black index admissions than for non-black index admissions. The bivariate regression between total payment and the Medicaid variable estimated that payments were, on average, roughly $1,600 higher for Medicaid index admissions as for non-Medicaid index admissions.

We then added the SDS variables to the current risk-adjustment model, both individually and jointly (Table 26). With the sole addition of race to the patient-level model, black index admissions were estimated to be roughly $391 more expensive than non-black index admissions. With the sole addition of Medicaid status to the patient-level model, Medicaid index admissions were estimated to be $516 more expensive than non-Medicaid index admissions. With the addition of both race and Medicaid status to the model, black index admissions were estimated to be roughly $287 more expensive than non-black index admission and Medicaid index admissions were

---

[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[ǂ] Used identity link and Gamma distribution as with original PN payment measure patient-level model. Estimates can be directly interpreted as standardized dollars.

estimated to be $496 more expensive than non-Medicaid index admissions. The interpretation of these results is that, holding all other risk-adjustment variables constant, both black and Medicaid patients are slightly more expensive for a PN episode of care than non-black and/or non-Medicaid patients.

**Table 26. Relationship between total PN payment and SDS variables[t]**

| Variables included in model | Variable | Estimate | Payment Ratio[r] | p-value |
|---|---|---|---|---|
| Current[*] + Black | Black | $391 | N/A | <0.0001 |
| Current + Medicaid | Medicaid | $516 | N/A | <0.0001 |
| Current + Black + Medicaid | Black | $287 | N/A | <0.0001 |
|  | Medicaid | $496 | N/A | <0.0001 |

### 3.7.4 Model Diagnostics

**Model Performance**

The addition of the SDS variables to the patient-level model did not significantly improve model performance compared the current model as demonstrated by the comparison of predictive ratios and quasi-$R^2$s (Table 27).

**Table 27. HF payment model performance**

| Diagnostic[γ] | Current[*] | Current + Black | Current + Medicaid | Current + Black + Medicaid |
|---|---|---|---|---|
| Predictive Ratio, 1st Decile (lowest) | 1.0567 | 1.0565 | 1.0593 | 1.0567 |
| Predictive Ratio, 2nd Decile | 1.0026 | 1.0021 | 0.9978 | 1.0026 |
| Predictive Ratio, 3rd Decile | 0.9832 | 0.9827 | 0.9843 | 0.9832 |
| Predictive Ratio, 4th Decile | 0.9729 | 0.9740 | 0.9747 | 0.9729 |
| Predictive Ratio, 5th Decile | 0.9698 | 0.9706 | 0.9704 | 0.9698 |
| Predictive Ratio, 6th Decile | 0.9728 | 0.9726 | 0.9731 | 0.9728 |
| Predictive Ratio, 7th Decile | 0.9818 | 0.9824 | 0.9823 | 0.9818 |
| Predictive Ratio, 8th Decile | 0.9943 | 0.9935 | 0.9925 | 0.9943 |
| Predictive Ratio, 9th Decile | 1.0173 | 1.0180 | 1.0181 | 1.0173 |
| Predictive Ratio, 10th Decile (highest) | 1.0580 | 1.0568 | 1.0571 | 1.0580 |
| Quasi-$R^2$ | 0.0893 | 0.0895 | 0.0898 | 0.0893 |

---

[t] Used identity link and Gamma distribution as with original PN payment measure patient-level model.
[r] Payment ratio is equal to exponentiated estimate
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[γ] A predictive ratio is an estimator's ratio of predicted outcome to observed outcome. A predictive ratio of 1.0 indicates an accurate prediction. A ratio greater than 1.0 indicates overprediction, and a ratio less than 1.0 indicates underprediction. The quasi-$R^2$ calculated is the $R^2$ from a regression of observed outcome on the predicted outcome. (Reference: Jones AM. Models for Health Care. Health, Econometrics and Data Group (HEDG) Working Papers. 2010.)

**Residual Analysis**

To assess whether the addition of the SDS variables produced a better model fit (i.e. improved the predicted total payment) for sub-groups of patients, we compared residuals$^{\Phi}$ from the current model to the residuals after the SDS variables were added (Table 28). Model fit was improved, on average, for all sub-groups (though the addition of a variable is likely to produce this result for any model). The prediction for black patients improved more than the prediction for non-black patients. The same is true for Medicaid and non-Medicaid patients.

**Table 28. Average residual for current model vs. current model with the addition of SDS variables for sub-groups of patients**

| Variable category | Current[*] ($)[ʌ] | Current + Black + Medicaid ($)[ʌ] |
|---|---|---|
| Black | 533 | 187 |
| Non-Black | -72 | -46 |
| Medicaid | 289 | -51 |
| Non-Medicaid | -145 | -21 |

**Collinearity Analyses**

With respect to assessing collinearity among the variables, the addition of the SDS variables did not increase the standard errors of the current risk-adjustment variables or appreciably alter their estimated coefficients (Appendix 5; full results available upon request). However, formal testing for collinearity among the risk-adjustments variables after the addition of the SDS variables was also performed (analyses available upon request). No indications of collinearity were found.
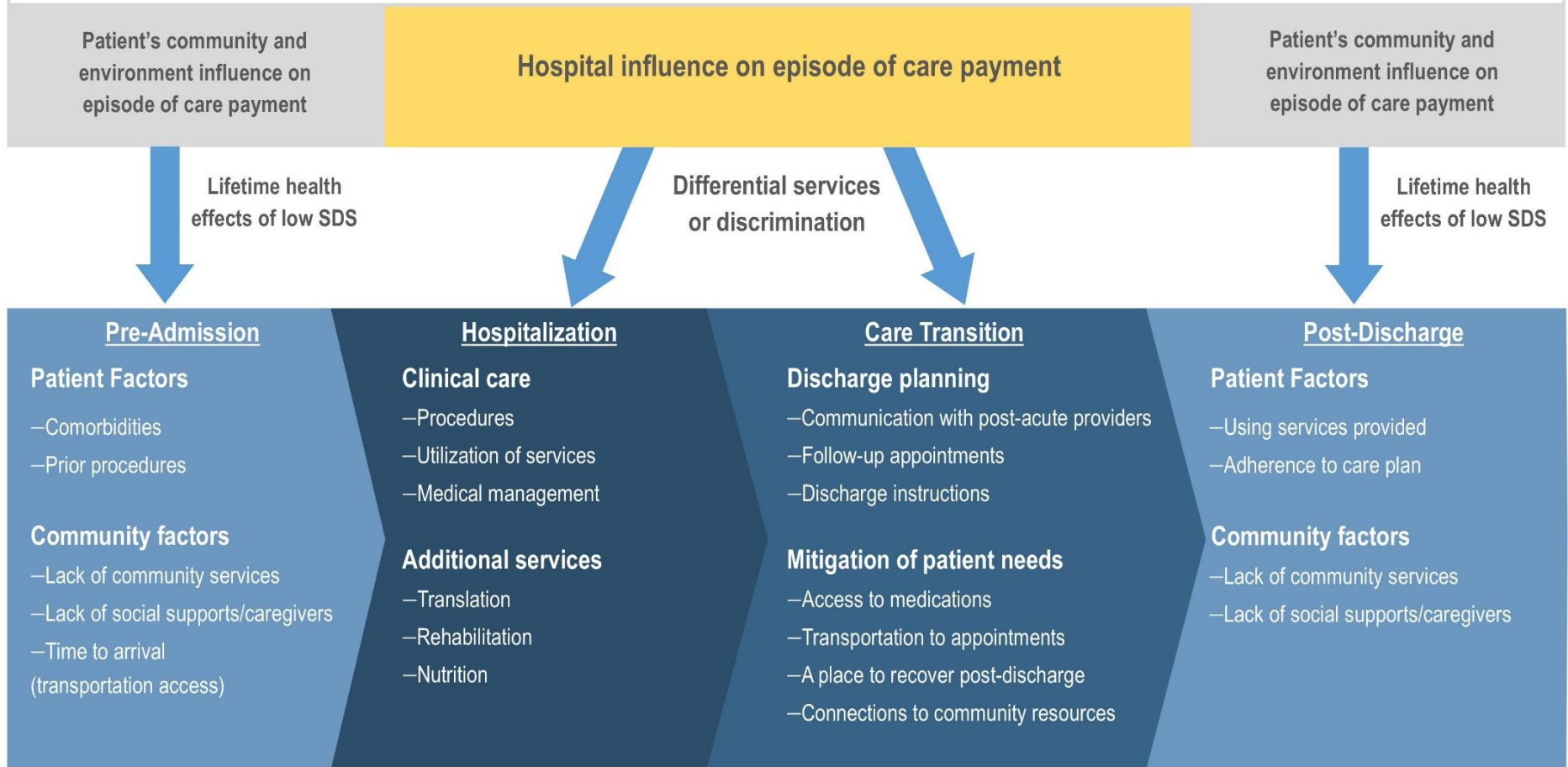
---

[Φ] The residual is the difference between the observed total payment and the predicted total payment
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)
[ʌ] Values are in standardized dollars

**APPENDIX 1.**
**UPDATED CONCEPTUAL MODEL**

SDS variables under consideration for risk adjustment, **Medicaid Status and Race,** may act through several pathways to influence episode of care payments

Patient's community and environment influence on episode of care payment

**Hospital influence on episode of care payment**

Patient's community and environment influence on episode of care payment

Lifetime health effects of low SDS

**Differential services or discrimination**

Lifetime health effects of low SDS

**Pre-Admission**

**Patient Factors**

—Comorbidities

—Prior procedures

**Community factors**

—Lack of community services

—Lack of social supports/caregivers

—Time to arrival (transportation access)

**Hospitalization**

**Clinical care**

—Procedures

—Utilization of services

—Medical management

**Additional services**

—Translation

—Rehabilitation

—Nutrition

**Care Transition**

**Discharge planning**

—Communication with post-acute providers

—Follow-up appointments

—Discharge instructions

**Mitigation of patient needs**

—Access to medications

—Transportation to appointments

—A place to recover post-discharge

—Connections to community resources

**Post-Discharge**

**Patient Factors**

—Using services provided

—Adherence to care plan

**Community factors**

—Lack of community services

—Lack of social supports/caregivers

# APPENDIX 2

## ARTICLES INCLUDED IN UPDATED LITERATURE REVIEW

1.    Gaskin DJ, Spencer CS, Richard P, Anderson GF, Powe NR, Laveist TA. Do hospitals provide lower-quality care to minorities than to whites? *Health Aff (Millwood).* 2008;27(2):518-527.
2.    Schulman KA, Berlin JA, Harless W, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med.* 1999;340(8):618-626.
3.    Epstein AM, Weissman JS, Schneider EC, Gatsonis C, Leape LL, Piana RN. Race and gender disparities in rates of cardiac revascularization: do they reflect appropriate use of procedures or problems in quality of care? *Med Care.* 2003;41(11):1240-1255.
4.    Stolberg SG. Blacks Found on Short End Of Heart Attack Procedure. http://www.nytimes.com/2001/05/10/us/blacks-found-on-short-end-of-heart-attack-procedure.html.
5.    Chen J, Rathore SS, Radford MJ, Wang Y, Krumholz HM. Racial differences in the use of cardiac catheterization after acute myocardial infarction. *N Engl J Med.* 2001;344(19):1443-1449.
6.    Reames BN, Birkmeyer NJ, Dimick JB, Ghaferi AA. Socioeconomic disparities in mortality after cancer surgery: failure to rescue. *JAMA Surg.* 2014;149(5):475-481.
7.    Helland E, Klick, J, Sridhar, A. The Role of Selection Effects in Estimated Racial Healthcare Disparities: Evidence from Travelers. Not published.
8.    Barnato AE, Lucas FL, Staiger D, Wennberg DE, Chandra A. Hospital-level racial disparities in acute myocardial infarction treatment and outcomes. *Med Care.* 2005;43(4):308-319.
9.    Dimick J, Ruhter J, Sarrazin MV, Birkmeyer JD. Black patients more likely than whites to undergo surgery at low-quality hospitals in segregated regions. *Health Aff (Millwood).* 2013;32(6):1046-1053.
10.   Gaskin DJ, Spencer CS, Richard P, Anderson G, Powe NR, LaVeist TA. Do minority patients use lower quality hospitals? *Inquiry.* 2011;48(3):209-220.
11.   Hasnain-Wynia R, Kang R, Landrum MB, Vogeli C, Baker DW, Weissman JS. Racial and ethnic disparities within and between hospitals for inpatient quality of care: an examination of patient-level Hospital Quality Alliance measures. *J Health Care Poor Underserved.* 2010;21(2):629-648.
12.   He D, Mellor JM, Jankowitz E. Racial and ethnic disparities in the surgical treatment of acute myocardial infarction: the role of hospital and physician effects. *Med Care Res Rev.* 2013;70(3):287-309.
13.   Joynt KE, Orav EJ, Jha AK. Thirty-day readmission rates for Medicare beneficiaries by race and site of care. *JAMA.* 2011;305(7):675-681.
14.   Skinner J, Chandra A, Staiger D, Lee J, McClellan M. Mortality after acute myocardial infarction in hospitals that disproportionately treat black patients. *Circulation.* 2005;112(17):2634-2641.

## AMI PAYMENT PATIENT-LEVEL SDS ANALYSES

**Patient-level risk-adjustment model results for the current model and with the addition of the SDS variables‡**

| Variable | Current* Model | | Current Model + Race + Medicaid | |
|---|---|---|---|---|
| | Estimate | p-value | Estimate | p-value |
| Intercept | 9.78 | - | 9.78 | |
| Age (65 – 74) | 0.184 | <.0001 | 0.187 | <.0001 |
| Age (75 – 84) | 0.172 | <.0001 | 0.173 | <.0001 |
| Age (>=85) (reference group) | 0.000 | - | 0.000 | - |
| History of Percutaneous Transluminal Coronary Angioplasty (PTCA) (ICD, 9 codes V45.82, 00.66, 36.06, 36.07) | -0.057 | <.0001 | -0.058 | <.0001 |
| History of Coronary Artery Bypass Graft (CABG) (ICD, 9 codes V45.81, 36.10, 36.16) | -0.187 | <.0001 | -0.189 | <.0001 |
| Metastatic cancer, acute leukemia and other severe cancers (CC 7, 8) | -0.092 | <.0001 | -0.092 | <.0001 |
| Diabetes mellitus (DM) or DM complications (CC 15, 19, 119, 120) | 0.079 | <.0001 | 0.082 | <.0001 |
| Protein, calorie malnutrition (CC 21) | 0.205 | <.0001 | 0.207 | <.0001 |
| Other significant endocrine and metabolic disorders (CC 22) | 0.062 | <.0001 | 0.063 | <.0001 |
| Other endocrine/metabolic/ nutritional disorders (CC 24) | -0.017 | <.0001 | -0.019 | <.0001 |
| Other gastrointestinal disorders (CC 36) | -0.029 | <.0001 | -0.029 | <.0001 |
| Osteoporosis and other bone/cartilage disorders (CC 41) | -0.045 | <.0001 | -0.046 | <.0001 |
| Iron deficiency or other unspecified anemias and blood disease (CC 47) | 0.199 | <.0001 | 0.201 | <.0001 |
| Delirium and encephalopathy (CC 48) | -0.027 | <.0001 | -0.025 | <.0001 |
| Dementia (CC 49) | -0.074 | <.0001 | -0.070 | <.0001 |
| Drug/alcohol psychosis (CC 51) | 0.0096 | 0.5220 | 0.0076 | 0.6100 |
| Drug/alcohol abuse/dependence (CC 52, 53) | 0.014 | <.0001 | 0.015 | <.0001 |
| Severe mental illness (CC 54, 55) | 0.030 | <.0001 | 0.033 | <.0001 |
| Reactive and unspecified psychosis (CC 56) | 0.0029 | 0.6271 | 0.0030 | 0.6139 |
| Depression/anxiety (CC 58, 59) | -0.025 | <.0001 | -0.027 | <.0001 |
| Congestive heart failure (CC 80) | -0.052 | <.0001 | -0.049 | <.0001 |
| Angina pectoris/old myocardial Infarction (CC 83) | -0.035 | <.0001 | -0.036 | <.0001 |
| Heart infection/inflammation, except rheumatic (CC 85) | 0.210 | <.0001 | 0.209 | <.0001 |
| Valvular or rheumatic heart disease (CC 86) | 0.068 | <.0001 | 0.067 | <.0001 |
| Congenital cardiac/circulatory defect (CC 87, 88) | 0.113 | <.0001 | 0.112 | <.0001 |
| Hypertension and hypertension complications (CC 89, 91) | -0.028 | <.0001 | -0.025 | <.0001 |
| Precerebral arterial occlusion and transient cerebral ischemia (CC 97) | 0.014 | <.0001 | 0.012 | 0.0002 |

‡Used log/inverse Gaussian as with original AMI Payment patient-level model
*Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)

| Variable | Current[*] Model | | Current Model + Race + Medicaid | |
|---|---|---|---|---|
| | Estimate | p-value | Estimate | p-value |
| Vascular disease and complications (CC 104, 105) | -0.0004 | 0.8872 | 0.0010 | 0.6963 |
| Other lung disorders (CC 115) | 0.057 | <.0001 | 0.057 | <.0001 |
| Legally blind (CC 116) | -0.035 | 0.0005 | -0.031 | 0.0021 |
| Dialysis status (CC 130) | 0.113 | <.0001 | 0.122 | <.0001 |
| Internal injuries (CC 160) | 0.148 | <.0001 | 0.148 | <.0001 |
| Black | - | - | -0.058 | <.0001 |
| Medicaid | - | - | -0.017 | <.0001 |

# HF PAYMENT PATIENT-LEVEL SDS ANALYSES

**Patient-level risk-adjustment model results for the current model and with the addition of the SDS variables[‡]**

| Variable | Current[*] Model | | Current Model + Race + Medicaid | |
|---|---|---|---|---|
| | Estimate | p-value | Estimate | p-value |
| Intercept | 9.50 | <.0001 | 9.50 | <.0001 |
| Age (65 – 74) | 0.054 | <.0001 | 0.056 | <.0001 |
| Age (75 – 84) | 0.044 | <.0001 | 0.045 | <.0001 |
| Age (>=85) (reference group) | 0.000 | | 0.000 | |
| History of infection (CC 1, 3-5) | 0.065 | <.0001 | 0.066 | <.0001 |
| Other infectious diseases (CC 6) | 0.024 | <.0001 | 0.023 | <.0001 |
| Protein-calorie malnutrition (CC 21) | 0.156 | <.0001 | 0.156 | <.0001 |
| Other significant endocrine and metabolic disorders (CC 22) | 0.074 | <.0001 | 0.075 | <.0001 |
| Other endocrine/metabolic/ nutritional disorders (CC 24) | 0.0002 | 0.9221 | -0.0002 | 0.9235 |
| Other gastrointestinal disorders (CC 36) | 0.0003 | 0.8565 | -0.0001 | 0.9550 |
| Bone/joint/muscle infections/necrosis (CC 37) | 0.046 | <.0001 | 0.046 | <.0001 |
| Other musculoskeletal and connective tissue disorders (CC 43) | 0.0064 | 0.0002 | 0.0066 | 0.0001 |
| Delirium and encephalopathy (CC 48) | 0.024 | <.0001 | 0.024 | <.0001 |
| Dementia and senility (CC 49, 50) | 0.050 | <.0001 | 0.049 | <.0001 |
| Schizophrenia/major depressive/ bipolar disorders (CC 54-55) | 0.062 | <.0001 | 0.060 | <.0001 |
| Other psychiatric disorders (CC 60) | 0.0060 | 0.0011 | 0.0045 | 0.0144 |
| Respiratory arrest/ cardiorespiratory failure/respirator dependence (CC 77-79) | 0.025 | <.0001 | 0.025 | <.0001 |
| Angina pectoris/old myocardial infarction (CC 83) | -0.015 | <.0001 | -0.016 | <.0001 |
| Heart infection/inflammation, except rheumatic (CC 85) | 0.092 | <.0001 | 0.093 | <.0001 |
| Major congenital cardiac/circulatory defect (CC 87) | 0.042 | <.0001 | 0.042 | <.0001 |
| Hypertension (CC 91) | -0.055 | <.0001 | -0.054 | <.0001 |
| Arrhythmias (CC 92, 93) | -0.025 | <.0001 | -0.025 | <.0001 |
| Cerebrovascular disease (CC 97-99) | 0.018 | <.0001 | 0.017 | <.0001 |
| Vascular or circulatory disease (CC 104-106) | 0.017 | <.0001 | 0.018 | <.0001 |
| History of pneumonia (CC 111-113) | 0.112 | <.0001 | 0.112 | <.0001 |
| Other ear, nose, throat, and mouth disorders (CC 127) | -0.019 | <.0001 | -0.019 | <.0001 |
| Dialysis status (CC 130) | 0.155 | <.0001 | 0.157 | <.0001 |

[‡] Used log link and Gamma distribution as with original HF payment measure patient-level model
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)

| Variable | Current[*] Model | | Current Model + Race + Medicaid | |
|---|---|---|---|---|
| | Estimate | p-value | Estimate | p-value |
| Renal failure (CC 131) | 0.030 | <.0001 | 0.031 | <.0001 |
| Decubitus ulcer of skin (CC 148) | 0.039 | <.0001 | 0.039 | <.0001 |
| Chronic ulcer of skin, except decubitus (CC 149) | 0.067 | <.0001 | 0.067 | <.0001 |
| Cellulitis, local skin infection (CC 152) | 0.011 | <.0001 | 0.010 | <.0001 |
| Hip fracture/dislocation (CC 158) | 0.033 | <.0001 | 0.032 | <.0001 |
| Internal injuries (CC 160) | 0.084 | <.0001 | 0.085 | <.0001 |
| Black | - | - | -0.030 | <.0001 |
| Medicaid | - | - | 0.012 | <.0001 |

# APPENDIX 5.

## PN PAYMENT PATIENT-LEVEL SDS ANALYSES

**Patient-level risk-adjustment model results for the current model and with the addition of the SDS variables[‡]**

| Variable | Current [*] Model | | Current Model + Race + Medicaid | |
|---|---|---|---|---|
| | Estimate ($) | p-value | Estimate ($) | p-value |
| Intercept | 10,141 | <.0001 | 10,095 | <.0001 |
| Age (65 – 74) | -1,164 | <.0001 | -1,208 | <.0001 |
| Age (75 – 84) | -683 | <.0001 | -694 | <.0001 |
| Age (>=85; reference group) | 0 | -- | 0 | -- |
| History of infection (CC 1, 3-5) | 2,246 | <.0001 | 2,252 | <.0001 |
| Other infectious diseases (CC 6) | 434 | <.0001 | 398 | <.0001 |
| Metastatic cancer and acute leukemia (CC 7) | 1,405 | <.0001 | 1,422 | <.0001 |
| Lung, upper digestive tract, and other severe cancers (CC 8) | 813 | <.0001 | 830 | <.0001 |
| Lymphatic, head and neck, brain, and other major cancers (CC 9) | 872 | <.0001 | 909 | <.0001 |
| Diabetes mellitus (DM) or DM complications (CC 15-19, 119, 120) | 472 | <.0001 | 427 | <.0001 |
| Protein-calorie malnutrition (CC 21) | 3,593 | <.0001 | 3,575 | <.0001 |
| Other significant endocrine and metabolic disorders (CC 22) | 1,237 | <.0001 | 1,246 | <.0001 |
| Other endocrine/metabolic/ nutritional disorders (CC 24) | -105 | <.0001 | -75 | 0.0012 |
| Other gastrointestinal disorders (CC 36) | -134 | <.0001 | -142 | <.0001 |
| Bone/joint/muscle infections/necrosis (CC 37) | 890 | <.0001 | 887 | <.0001 |
| Osteoporosis and other bone/cartilage disorders (CC 41) | -187 | <.0001 | -185 | <.0001 |
| Severe hematological disorders (CC 44) | 1,106 | <.0001 | 1,140 | <.0001 |
| Iron deficiency or other unspecified anemias and blood disease (CC 47) | 1,356 | <.0001 | 1,336 | <.0001 |
| Delirium and encephalopathy (CC 48) | 486 | <.0001 | 491 | <.0001 |
| Dementia and senility (CC 49-50) | 1,253 | <.0001 | 1,183 | <.0001 |
| Drug/alcohol dependence/psychosis (CC 51-52) | -87 | 0.2352 | -89 | 0.2228 |
| Drug/alcohol abuse, without dependence (CC 53) | -8 | 0.7688 | -34 | 0.2349 |
| Major psychiatric disorders (CC 54-56) | 1,052 | <.0001 | 985 | <.0001 |
| Plegia, paralysis, spinal cord disorder and amputation (CC 67-69, 100-101, 177-178) | 1,358 | <.0001 | 1,296 | <.0001 |
| Muscular dystrophy and/or polyneuropathy (CC 70-71) | 602 | <.0001 | 622 | <.0001 |
| Multiple sclerosis and Parkinson's (CC 72-73) | 1,237 | <.0001 | 1,238 | <.0001 |
| Coma, brain compression/anoxic damage | 1,289 | <.0001 | 1,287 | <.0001 |

---

[‡] Used identity link and Gamma distribution as with original PN payment measure patient-level model
[*] Current indicates inclusion of all current risk-adjustment variables (age, comorbidities)

| Variable | Current* Model | | Current Model + Race + Medicaid | |
|---|---|---|---|---|
| | Estimate ($) | p-value | Estimate ($) | p-value |
| (CC 75) | | | | |
| Respirator dependence/respiratory arrest/cardiorespiratory failure (CC 77-79) | 949 | <.0001 | 946 | <.0001 |
| Congestive heart failure (CC 80) | 644 | <.0001 | 609 | <.0001 |
| Angina pectoris/old myocardial infarction (CC 83) | -299 | <.0001 | -291 | <.0001 |
| Heart infection/inflammation, except rheumatic (CC 85) | 1,255 | <.0001 | 1,249 | <.0001 |
| Valvular and rheumatic heart disease (CC 86) | 492 | <.0001 | 519 | <.0001 |
| Hypertension (CC 91) | -57 | 0.0196 | -81 | 0.0010 |
| Arrhythmias (CC 92-93) | 201 | <.0001 | 229 | <.0001 |
| Stroke (CC 95-96) | 416 | <.0001 | 400 | <.0001 |
| Vascular or circulatory disease (CC 104-106) | 156 | <.0001 | 155 | <.0001 |
| Chronic Obstructive Pulmonary Disease (COPD) (CC 108) | 813 | <.0001 | 799 | <.0001 |
| Fibrosis of lung or other chronic lung disorder (CC 109) | 406 | <.0001 | 430 | <.0001 |
| Asthma (CC 110) | -694 | <.0001 | -708 | <.0001 |
| Aspiration and specified bacterial pneumonias (CC 111) | 461 | <.0001 | 442 | <.0001 |
| Pleural effusion/pneumothorax (CC 114) | 548 | <.0001 | 567 | <.0001 |
| Other ear, nose, throat, and mouth disorders (CC 127) | -484 | <.0001 | -472 | <.0001 |
| Dialysis status (CC 130) | 3,239 | <.0001 | 3,161 | <.0001 |
| Renal failure (CC 131) | 473 | <.0001 | 462 | <.0001 |
| Decubitus ulcer or chronic skin ulcer (CC 148-149) | 1,159 | <.0001 | 1,142 | <.0001 |
| Head injury (CC 154-156) | 327 | <.0001 | 343 | <.0001 |
| Vertebral fractures (CC 157) | 1,017 | <.0001 | 1,044 | <.0001 |
| Hip fracture/dislocation (CC 158) | 584 | <.0001 | 593 | <.0001 |
| Major fracture, except of skull, vertebrae, or hip (CC 159) | 770 | <.0001 | 780 | <.0001 |
| Internal injuries (CC 160) | 1,512 | <.0001 | 1,530 | <.0001 |
| Major symptoms, abnormalities (CC 166) | 689 | <.0001 | 669 | <.0001 |
| Black | - | - | 287 | <.0001 |
| Medicaid | - | - | 496 | <.0001 |

# Cost and Resource Use Standing Committee Ad Hoc Review: Conceptual & Empirical Analysis of SDS Variables and Payment Outcomes

The Cost and Resource Use Standing Committee met over two webinars in May and October 2015 to evaluate three risk-standardized payment measures under the NQF Sociodemographic Status (SDS) Adjustment Trial Period guidance. This report summarizes the Committee's review and recommendations.

*Measures under Review*
- #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)

## Overview of the Sociodemographic Status (SDS) Adjustment Trial Period

The NQF Board of Directors approved SDS trial period is designated as a 2-year period of time during which SDS factors should be considered as potential factors in the risk-adjustment approach of measures submitted to NQF if there is a conceptual reason for doing so. If there is a conceptual relationship between potential SDS risk factors and the outcome of interest, the developer should conduct empirical analyses to determine whether such factors improve the risk-adjustment model and/or result in meaningful difference in performance rates. Based on those analyses, measure developers decide whether to include SDS in their risk-adjustment approach. The trial period began January 2015.

Prior to this SDS trial period, NQF criteria and policy prohibited the inclusion of SDS factors in the risk adjustment approach and only allowed for the inclusion of a patient's clinical factors present at the start of care. Rather than including SDS factors related to the outcome in statistical risk models, NQF guidance indicated that measure results should be stratified by these variables.

## Reviewing the Cost Measures during the SDS Trial Period

The evaluation of the aforementioned measures began and ended prior to the inception of the SDS trial period, and therefore the Committee did not consider SDS factors as part of the risk-adjustment approach during their initial evaluation. When the NQF Board of Directors (BoD) Executive Committee ratified the Consensus Standards Approval Committee's (CSAC) approval to endorse the measures, it did so with conditions in recognition of the potential impact of SDS on cost and payment outcomes and the impending start of the SDS trial period. The conditions for endorsement included:

- Consideration of the measures to enter the SDS trial period; and
- A one-year look-back assessment of unintended consequences.

Following the NQF Board of Directors Executive Committee decision to endorse the cost measures with the condition that they be considered under the trial period guidance, NQF, in collaboration with the CMS/Yale measure development team, agreed to divide the assessment of the impact of SDS variables into two stages (and webinars).

### *Stage 1/Webinar #1 (May 21, 2015): Conceptual Analyses Review*

- o Review the conceptual analysis of the relationship between SDS factors and the hospital-level, risk-standardized payment associated with a 30-day episode-of-care for the three conditions (pneumonia, HF, and AMI)
- o Determine whether further empirical analysis is warranted
- o Identify the variables to be pursued in empirical analysis
- o Provide input on the plan or approach to empirical analysis of the selected variables

### *Stage 2/Webinar #2 (October 27, 2015): Empirical Analyses Review*

- o Review empirical analysis of the impact of SDS risk factors in the risk model and measure score
- o Make a recommendation on endorsement status

## Webinar 1: Conceptual Analyses Review

A conceptual relationship refers to a logical theory or rationale that explains the association between a SDS factor(s) and the outcome of interest. The conceptual basis may be informed by prior research and/or healthcare experience related to the outcome of interest, but does not require a direct causal relationship (i.e., it could be a direct cause, an indirect cause, or serve as a surrogate for a cause for which data are lacking). An assessment of a conceptual relationship between a SDS factor and an outcome of interest includes a consideration of whether the effect of the SDS is primarily mediated by the quality of care delivered (i.e., does the SDS factor lead to the delivery of inferior care processes, which in turn affects the outcome?).

The CMS/Yale Core development team submitted a memo and conceptual model diagram illustrating the potential relationships of various factors during the episode of care captured by the measures (i.e., hospital admission through 30 days post-discharge). Of the factors identified in their conceptual analysis, they selected three variables that have also been noted in the literature to have a conceptual relationship to utilization and payment and that can be represented by data to which they have access. They also identified the relevant data that are currently available to them for potential empirical testing.

| SDS factors with conceptual relationship to utilization and payment selected by Yale Core | Variables and data source |
| --- | --- |
| Educational Attainment | Educational attainment obtained from Census data linked to patient's 5-digit ZIP Code |
| Income | Income level obtained from Census data linked to patient's 5-digit ZIP Code<br><br>Medicaid (Dual Eligibility) Status as a proxy for low income, obtained from Medicare enrollment data |
| Insurance coverage | Medicaid (Dual Eligibility) Status as a proxy, obtained from Medicare enrollment data |
| Race | Operationalized as black or white race, obtained from Medicare enrollment data |

In their overview of the conceptual model, the CMS/Yale team noted the following regarding the appropriateness of adjusting on these variables:

- The association of low socioeconomic status and hospital cost is uncertain and exerts itself at multiple points in episode of care. The impact of SDS may be intrinsic to the patient or extrinsic and it is unclear whether hospitals should be held responsible and whether these factors should be included in the adjustment.
- During hospitalization, the hospital has control of a patient's care and therefore any differences in care influenced by SDS should not be adjusted for. Once a patient is discharged, the hospital only has partial control over the patient's care and environmental, community, and patient factors play a larger role.
- The risk standardized payments captured by the measures are based on DRGs (which do not account for length of stay, translational services, or the cost of care coordination). The risk-standardized payments captured by the measures are only linked to procedures, complications of care, and, sometimes, comorbidities.

### Committee Discussion
The Committee discussed the conceptual model as well as the literature review summary submitted by the developer. The Committee expressed concerns about some elements of the conceptual model and offered suggestions on how to make it and the literature review broader and more comprehensive.

1. **Broaden the conceptual model**. The Committee was concerned that the conceptual model seemed too medically-oriented and should be broadened to account for more public health variables.  For example, the model did not address community, environmental, or patient factors (e.g., social supports, lack of money to buy medication, no refrigerator). The conceptual model should reflect resources available for care within individual hospitals. While these should not be included in the risk-adjustment approach, because differential resources can impact quality of care, they should be noted in the conceptual model.

2. **Additional literature review.** 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. Members also suggested that the developers do a broader search of literature to include readmissions and impact of SDS on health.
3. **Conceptual Relationships.** Based on the research performed by the developers, the Committee agreed there is a conceptual relationship between the selected variables and cost/payment outcomes.

NQF guidance for the evaluation of SDS factors states that if the Committee believes a conceptual relationship exists between the sociodemographic factor(s) and the outcome (i.e, resource utilization or cost), developers should conduct empirical analyses to confirm that relationship. The Committee determined there is conceptual relationship between the proposed variables and the three payment outcomes. Their discussion yielded the following recommendations regarding the examination and consideration of these variables in empirical analyses:

- **Race:** The Committee recommended that the CMS/Yale team review the data and consider including other race variables beyond black.
- **Income and educational attainment**: The Committee was not in favor of the developers beginning empirical analysis using data linked on the basis of 5-digit ZIP Code. The Committee preferred the developers to use their resources analyzing the 9-digit ZIP Code data once it is available to them.
- **Medicaid/dual eligibility status**: The Committee was in support of empirical analysis on this (Medicaid status) variable, but only in combination with the Low Income Subsidy (LIS) data as proxy for insurance status and income.

## Webinar 2: Empirical Analyses Review

### *Follow up on Conceptual Analysis*
In response to the Committee's recommendations in May, the developer submitted a second memo that included a summary of their review of 14 additional articles and a revised conceptual model. These additional articles examined within and between hospital differences in outcomes related to SDS variables; the key findings from this review have been excerpted below:

- "Taken together these papers do not present a conclusive or consistent picture about the role of within hospital differences in treatment of patients based on SDS nor the subsequent impact on outcomes or cost. However they provide some evidence that in certain settings differential care by race could contribute to differences in costs and outcome."
- "Taken together, the body of literature reveals an inconsistent and complex association of low SDS and health outcomes. Most studies used race as their independent variable with less attention to income or other measures of poverty (e.g. Medicaid status). The literature demonstrates both within and between hospital differences in outcomes among racial/ethnic groups that can be partially explained by the use of lower quality hospitals by minorities."

The CMS/Yale Team also revised the conceptual model to broaden the scope of community-level factors that are considered. In doing so, they updated the pre-admission and post-discharge phases of the 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". And finally, the model also was reoriented to capture the potential pathways by which low SDS may impact the care provided to patients.

Upon review of these modifications to the conceptual model and the literature review, the Committee and developers noted the significant gaps in the literature specific to the impact of SDS on cost, utilization or payment outcomes. One Committee member raised the issue of whether there is a relationship between the quality of the hospitals that low SDS patients are likely to be treated in and the resources available to those facilities, which may be hidden by the use of standardized payments based on diagnosis- related groups (DRGs). In particular, the Committee member questioned whether any of the literature identified by the developers addressed this issue, to which the developer confirmed their literature search did not find anything specific to this issue.

Another concern raised by one of the Committee members was in reference to summary of the literature review (page 3, Yale Memo), which indicated that the body of literature identified by CMS/Yale suggests an inconsistent and complex association of low SDS and health outcomes. The Committee member cautioned that this language could be taken to suggest that race or ethnicity could be used to proxy for high or low SES since their literature review predominantly focused on race/ethnicity as the dependent variable.

## Review of Empirical Analyses

The importance of the SDS variables in the risk adjustment model should be evaluated by the size of the SDS coefficients in the risk adjustment model, the p-values associated with the SDS coefficients, and the impact of adjusting for the SDS variables on the measure results. Reasons for including the SDS variables in the risk-adjustment approach include (1) demonstration of the contribution of the SDS factor(s) to unique variation in the outcome that is not due to between-unit effects, (2) adjustment leads to substantial differences between measure scores (although this doesn't have to result in change in rankings), or (3) if needed for face validity of the approach.

### Variables Used in the Empirical Analyses
1. Race: Categorized as Black and Non-Black
2. Medicaid enrollment/Dual Status (as proxy for low income): Categorized as Medicaid and Non-Medicaid.

The CMS/Yale Team explained that while the Committee recommended the use of the LIS variable in conjunction with the Medicated variable as a proxy for income, when they performed their analysis of the LIS data they chose not to use it as the patients captured with their current method to identify patients based on dual status alone sufficiently overlapped with those captured with the dual plus LIS variables.

*Methods*

For each measure/condition, the Yale team conducted the following analyses:

1. Determined the prevalence of the SDS variables for each condition across the measured entities.
2. Determined whether there is a bi-variate association between each of the SDS variables and the outcome.. [Bi-variate relationship]
3. Determined whether the inclusion of variables in the risk model improved the risk model's ability to account for variation in the data.
4. Determined whether the risk-standardized payment changed with the inclusion of the SDS variables (i.e., how much did the payment increase or decrease for the hospitals in the sample with the inclusion of the variables in the model?)
5. Determined whether and how much the ranking of hospitals shifted with the addition of the SDS variables.

*Results*

These measures estimate hospital-level, risk-standardized episode-of-care payment starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of AMI, HF or Pneumonia.

**Level of Analysis:** Facility   **Costing Method:** Standardized pricing   **Target Population:** Senior Care   **Data Source:** Administrative Claims

| | #2431: AMI | #2436: HF | #2579: Pneumonia |
|---|---|---|---|
| **SUMMARY OF EMPIRICAL ANALYSIS & RESULTS:** | | | |
| *Was there sufficient variation of the SDS variables within the sample to warrant additional analysis?* | Yes | Yes | Yes |
| *Was there bivariate association between the SDS variables and payment outcomes?* | Yes | Yes | Yes |
| *Was there a statistically significant relationship between the variables and payment outcomes after accounting for other clinical variables?* | Yes | Yes | Yes |
| *Did the risk model's ability to account for variation in payment improve with the addition of the variables?* | Minimal | Minimal | No |
| *Was there a change in the hospital payment with the inclusion of the variables?* | Black: Slightly lower payment<br><br>Medicaid: Slightly lower payment | Black: Slightly lower payment<br><br>Medicaid: Slightly higher payment | Black: Higher payment +$287<br><br>Medicaid: Higher payment, +$496 |

| | #2431: AMI | #2436: HF | #2579: Pneumonia |
|---|---|---|---|
| *Did the ranking of hospital payments change with the inclusion of the variables in the risk model?* | No | No | No |
| **DEVELOPER DECISION TO INCLUDE FACTORS IN THE MODEL:** | Based on the results of the empirical analysis, the developers chose **NOT** to include the variables in the model. The developers cited the nominal impact of the SDS variables on the risk model performance and payment outcomes as their rationale not to change the measure. | | |
| **COMMITTEE RECOMMENDATIONS:** | Validity: **H-4; M-9; L-0; I-0** Recommendation for continued endorsement: **Yes- 13; No-0** | Validity: **H-4; M-7; L-1; I-0** Recommendation for continued endorsement: **Yes- 12; No-0** | Validity: **H-4; M-8; L-0; I-0** Recommendation for continued endorsement: **Yes- 12; No-0** |

*Committee Discussion*

Ultimately the Committee voted to continue endorsement of the measures without inclusion of SDS factors in the risk-adjustment approach. The empirical results do not suggest that accounting for black versus non-black race and Medicaid dual-eligibility status is needed when estimating facility-level episode-of-care payments for AMI, heart failure, or pneumonia.

The Committee discussed the minimal impact of the SDS variables on the payment outcomes. The episode of care captured by the measures extends 30 days after discharge for each of the conditions, capturing payments for any readmissions, admissions to skilled nursing facilities, and other post-acute activities within that timeframe. It is this post-acute timeframe that has been shown to have the most variation for these measures and is the time within the episode where patient-level SDS and community-level factors presumably exert the most influence on outcomes. Through dialogue with the developers, the Committee spent much of their discussion on these measures trying to understand why the results of the empirical analysis did not align with their expectations. In doing so, the Committee members and developers identified several important points for clarification in understanding the measures and what they capture, as well as some possible considerations for why the analysis yielded the results it did.  These points are discussed below.

### Measuring Hospital Payments

Diagnosis-related groups (DRGs) are a classification system used for grouping similar patients into groups based on several factors including diagnosis, procedures, age, sex, and comorbidities (e.g., AMI patients with X severity level are grouped into a DRG). These groups are then used to determine the amount Medicare pays a hospital based on estimates of the amount of resources that would be used to care for the patients within the group. Understanding the concept of DRGs is important to understanding what these payment measures are measuring, as they are based on standardized DRG hospital payments and therefore do not capture itemized costs for procedures or interventions during the hospital portion of the episode. In fact, these measures do not capture cost at all, as only payments are being measured. The Committee acknowledged that although there are differences between hospitals and in the resources available to them to manage patients and invest in quality improvement, the payments captured in the measure are standardized and are not necessarily a reflection of the resources that are available or utilized within a hospital to care for the patients.  Further, even though these measures may be in use for public reporting and potentially other programs, risk adjustment for the examined SDS factors would only potentially address the "mismeasurement" issue and would not address the problem of uneven distribution of resources among hospitals in order to reduce disparities. This, as the Committee pointed out, is a payment policy issue, not a measurement issue.

The Committee also discussed how the distribution of payments across the episode might relate to the impact of the SDS variables on the payment outcomes. For example, in the AMI population, the CMS/Yale developers indicated that approximately 70 percent of the payment is

allocated for inpatient hospital charges (i.e., DRG-based payment). The remaining 30 percent of the payment covers the post-acute care phase of the episode, during which SDS factors are presumably most likely to exert an impact. Given the small proportion of the payment that is potentially impacted by SDS in the episode, the Committee acknowledged that this might be one reason for why any empirical differences might have been minimal. The proportion of the payment allocated for inpatient and post-acute segments of the episode varies among the three measures. In response to this concern, the developers explained that they had not performed analyses of the post-acute portion of the payment to determine whether the SDS factors had a greater impact; they focused their analysis on the entire episode rather than segments of the episode. The Committee noted that conducting this analysis might shed additional light on whether adjustment for the SDS factors during the post-acute phase would be more impactful.

### Differences in the Empirical Results among the Measures

The CMS/Yale developers also offered some possible causes for the minimal impact on the empirical results. Initial analyses indicated variation in the number of procedures received by certain subpopulations with AMI, potentially reflecting poor quality care.  For example, these analyses suggested that both Blacks and Medicaid patients received fewer procedures than non-Blacks and non-Medicaid patients, which could result in lower DRG payments. While this might explain why the payments were lower in the AMI measure for these variables, it might not explain those differences for the other two measures. Given that the literature indicated inconsistent relationships in utilization within certain populations, the developer also suggested that perhaps that both high and low payments for patients within these groups  washes out any significant differences in the end. For the heart failure and pneumonia measures, the developer suggested that because these conditions are predominantly treated medically (with medications rather than surgery or major procedures) during the hospitalization, there is less variation in the DRG payments.

### Relationship to Clinical Risk-Adjustment

When discussing the clinical risk-adjustment model, the developers pointed out that due to the homogeneity of the samples used for analysis, the effect of the clinical risk adjustment is less impactful. This was a major topic of discussion among the Committee members during their initial review of the measures, as the risk model predicted very little variation (R-squared =0.07). The risk strategy was ultimately accepted by the Committee with further explanation from the developers on the composition of the sample and their approach. For this analysis of the impact of SDS, the SDS variables were added into a model that has already been extensively clinically adjusted for conditions present on admission. The CMS/Yale developers suggested that conditions present on admission might in fact be occurring due to differences in SDS that have impacted a patient's health. Thus the impact of SDS might indirectly be adjusted for in the clinical adjustment. This might also explain why the SDS variables had minimal impact.

### Community-Level Adjustment

The Committee had an 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. They acknowledged that for some of the post-hospitalization services, the community context is a critical variable. For example, if high quality nursing homecare is available, or primary care physician services are available in the community, it makes a difference in outcomes and these

factors may or may not be fully captured by the patient-level SDS adjustment.  In communities where a high proportion of the post-acute population is in need of these services, there may be less capacity to adequately care for them, suggesting that the neighborhood a hospital is in may also have an impact on payment outcomes. The developers expressed interest in potentially considering these factors in the model, but sought Committee input and recommendations on how to approach this.

## Next Steps

Based on the NQF ad hoc review process, the results of this review will be posted for public and NQF member commenting, followed by review by the Consensus Standards Approval Committee (CSAC), review by the Board of Directors, and an appeals period. The initial review by CSAC will take place during their in person meeting on November 17-18, 2015. Once comments from the commenting period have been compiled, they will be submitted to CSAC for a final review during their January 12, 2016 call. The results and learnings of this project will also be shared with the Disparities Committee for discussion during their January meeting and as needed with other Standing Committee that are considering measures for SDS adjustment.

TO:      Consensus Standards Approval Committee (CSAC)

FR:      Cost and Resource Use Standing Committee

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

DA:      January 6, 2016

The CSAC will review recommendations from the Cost and Resource Use Standing Committee Ad Hoc Review of the Conceptual and Empirical Analysis of socioeconomic status and demographic variables (SDS) Variables and Payment Outcomes project at its January 12 conference call.

This memo includes a summary of the project, the three recommended measures, and themes identified from and responses to the public and member comments.

Accompanying this memo are the following documents:

1. Cost and Resource Use Standing Committee Ad Hoc Review: Conceptual & Empirical Analysis of SDS Variables and Payment Outcomes Draft Report.
2. Comment table. Staff has identified themes within the comments received. This table lists 11 comments received.

**CSAC ACTION REQUIRED**
The CSAC will consider approval of the Standing Committee recommendation to continue 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 pneumonia (CMS/Yale)

**BACKGROUND**
In early 2015, NQF began a two year trial period during which sociodemographic status (SDS) factors can be considered in the risk-adjustment approach of measures submitted to NQF if there is a conceptual and empirical rationale for doing so. Prior to January 2015, NQF criteria and policy prohibited the inclusion of such factors in the risk adjustment approach and only allowed for inclusion of a patient's clinical factors present at the start of care.

Because the evaluation of the three measures listed above began and ended prior to the start of the SDS trial period, the Cost and Resource Use Standing Committee did not consider SDS factors as part of the risk-adjustment approach during their initial evaluation. When the NQF Board of Directors Executive Committee ratified the CSAC's approval to endorse the measures, it did so with the condition that these

measures enter the SDS trial period because of the questions raised throughout the project about the potential impact of SDS on payment outcomes and the impending start of the SDS trial period.

To meet this condition for endorsement, the Cost and Resource Use Standing Committee reviewed the conceptual and empirical relationship between sociodemographic factors and payment outcomes. The measure developers were asked to submit additional analysis in a two-phased approach:
- Webinar #1: Examine the conceptual relationship between SDS factors and the outcome
- Webinar #2: Examine the empirical relationship between SDS factors and the outcome

During the first webinar, the Standing Committee reviewed the conceptual analysis of selected SDS variables provided by the measure developer and determined that further empirical analysis was warranted. The Committee reviewed the proposed variables to be pursued in the empirical analysis by the measure developer and provided input on the approach to empirical analysis.

During the second webinar, the Standing Committee reviewed the empirical analysis of the impact of SDS variables in the risk model and the measure score. The Standing Committee evaluated the validity of the developer's decision to not include SDS adjustment in the risk adjustment model based on the empirical analysis provided. The Committee ultimately decided to recommend continued endorsement for the three measures without SDS adjustment.

**MEASURE REVIEW SUMMARY**
These measures estimate hospital-level, risk-standardized episode-of-care payment starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of AMI, HF or Pneumonia.

The developers explored the impact of race categorized as Black and Non-Black and Medicaid enrollment/Dual Status (as a proxy for low income) categorized as Medicaid and Non-Medicaid on the risk adjustment model as these variables were often cited in the literature for these outcomes and other similar outcomes. The developer and the Committee agreed that there was sufficient conceptual rationale for an exploration of these variables for consideration in the risk adjustment approach. Based on the results of the empirical analysis, the developers chose **NOT** to include the variables in the model. The developers cited the nominal impact of the SDS variables on the risk model performance and payment outcomes as their rationale not to change the measure.

Ultimately the Committee voted to continue endorsement of the measures without inclusion of SDS factors in the risk-adjustment approach. The empirical results do not suggest that accounting for Black versus non-Black and Medicaid dual-eligibility status is needed when estimating facility-level episode-of-care payments for AMI, HF, or pneumonia.

**Comment Themes**

The results of the Standing Committee's review were posted for public and NQF member commenting. Eleven comments were received from five organizations.

*Theme 1 – Guidelines for Variables Reviewed*

Historically, NQF has not been prescriptive in its approach to the variables included in risk adjustment models. Measure developers are responsible for the selection of the variables included in the model and for defending the selection of those variables to the Standing Committees. This approach applies to both the selection of clinical and sociodemographic factors. However, a number of commenters raised concerns with this approach and asked for NQF to establish guidelines for what SDS factors should be considered. Commenters recommend that providing developers this additional guidance would allow for a more consistent and thorough trial period.

Potential Response: Historically, NQF has not been prescriptive in its approach to the variables included in risk adjustment models. Measure developers are responsible for the selection of the variables included in the model and for defending the selection of those variables to the Standing Committees. This approach applies to both the selection of clinical and sociodemographic factors. The selection of SDS variables to include in risk adjustment models should be guided by the conceptual relationship between the SDS factor and the outcome, and the results of the empirical testing. The NQF Disparities Standing Committee will consider this issue during their January 20-21, 2016 meeting.

### Theme 2 – Concern about the Variables Selected
Commenters raised a number of concerns about the variables selected by the developer for inclusion in the risk adjustment model. First, commenters expressed concerns about the inclusion of race as a factor as well as the limited number of categories used to express this construct. Commenters also expressed their concern with the potential use of race as a proxy for sociodemographic status. Next, commenters expressed disappointment that the developers did not analyze the data by nine digit zip code. Some commenters suggested that five digit zip code be considered as an interim step until nine digit zip code information becomes available. Finally, commenters suggested that the Low Income Subsidy (LIS) should be used in combination with Medicaid status.

Potential Response: The CMS Yale Team justified the inclusion of race in the empirical analysis as it is often used as the SDS factor examined in these outcomes in the literature. The Standing Committee agreed that the use of this variable was sufficiently justified to allow for additional empirical examination. The Standing Committee was not in favor of the developers beginning empirical analysis using data linked on the basis of 5-digit ZIP Code. The Committee preferred the developers to use their resources analyzing the 9-digit ZIP Code data once it is available to them since the 5-digit ZIP code data is often too heterogeneous. The CMS/Yale Team explained that while the Committee recommended the use of the low income subsidy (LIS) variable in conjunction with the Medicaid variable as a proxy for income, when they performed their analysis of the LIS data they chose not to use it as the patients captured with their current method to identify patients based on dual status alone sufficiently overlapped with those captured with the dual plus LIS variables.

### Theme 3 – No Analysis of Post-Acute Care (PAC) Portion of the Payment
Commenters expressed concern that the developers did not perform analyses of the post-acute portion of the payment to assess the impact of SDS factors. Commenters noted that the post-acute expenses are a significant source of variation in these measures and are not mitigated by using a standardized DRG payment for the hospital expenditures. Thus, commenters noted that analysis of the post-acute portion of the payment may be more sensitive to SDS factors.

Potential Response:  The measure developers focused their analysis of the impact of SDS factors on the entire episode payment rather than segments of the episode.


**NEXT STEPS**
The results of the NQF ad hoc review process will considered by the Board of Directors and will be subject to an appeals period. The results of this review will also be shared with the Disparities Standing Committee during their January 2016 meeting.

TAB 2

TO:     Executive Committee

FR:     Helen Burstin, Chief Scientific Officer

        Marcia Wilson, Senior Vice President, Quality Measurement

RE:     Ratification 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:     February 2, 2016


**ACTION REQUIRED**

The Executive Committee is asked to ratify the Consensus Standards Approval Committee's (CSAC) recommendation to continue 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)


**BACKGROUND**

In early 2015, NQF began a two year trial period during which sociodemographic status (SDS) factors could be considered in the risk-adjustment approach of measures submitted to NQF if there is a conceptual and empirical rationale for doing so. Prior to this, NQF criteria and policy prohibited the inclusion of such factors in the risk adjustment approach and only allowed for inclusion of a patient's clinical factors present at the start of care.

Because the evaluation of the three measures listed above began and ended prior to the start of the SDS trial period, the Cost and Resource Use Standing Committee did not consider SDS factors as part of the risk-adjustment approach during their initial evaluation.  When the NQF Board of Directors Executive Committee ratified the CSAC's approval to endorse the measures, it did so with the condition that these measures enter the SDS trial period because of the questions raised throughout the project about the potential impact of SDS on payment outcomes and the impending start of the SDS trial period.

To meet this condition for endorsement, the Cost and Resource Use Standing Committee reviewed the conceptual and empirical relationship between sociodemographic factors and payment outcomes. The measure developers were asked to submit additional analysis in a two-phased approach:
- Webinar #1: Examine the conceptual relationship between SDS factors and the outcome
- Webinar #2: Examine the empirical relationship between SDS factors and the outcome

During the first webinar, the Standing Committee reviewed the conceptual analysis of selected SDS variables provided by the measure developer and determined that further empirical analysis was warranted. The Committee reviewed the proposed variables to be pursued in the empirical analysis by the measure developer and provided input on the approach to empirical analysis.

During the second webinar, the Standing Committee reviewed the empirical analysis of the impact of SDS variables in the risk model and the measure score. The Standing Committee evaluated the validity of the developer's decision to not include SDS adjustment in the risk adjustment model based on the empirical analysis provided. The Committee ultimately decided to recommend continued endorsement for the three measures without SDS adjustment.

**Measure Review Summary**
Standing Committee Evaluation: These measures estimate hospital-level, risk-standardized episode-of-care payment starting with inpatient admission to a short term acute-care facility and extending 30 days post-admission for Medicare fee-for-service (FFS) patients who are 65 years of age or older with a principal discharge diagnosis of AMI, HF or pneumonia.

The developers explored the impact of race categorized as Black and Non-Black and Medicaid enrollment/Dual Status (as a proxy for low income) categorized as Medicaid and Non-Medicaid on the risk adjustment model as these variables were often cited in the literature for these outcomes and other similar outcomes. The developer and the Committee generally agreed that there was sufficient conceptual rationale for an exploration of these variables for consideration in the risk adjustment approach. Based on the results of the empirical analysis, the developers chose **NOT** to include the variables in the model as the empirical results do not suggest that accounting for Black versus non-Black and Medicaid dual-eligibility status is needed when estimating facility-level episode-of-care payments for AMI, HF, or pneumonia. The developers cited the nominal impact of the SDS variables on the risk model performance and payment outcomes as their rationale not to change the measures. Ultimately the Committee voted to continue endorsement of the measures without inclusion of SDS factors in the risk-adjustment approach.

CSAC Review: CSAC recommended continuing endorsement of the three measures as recommended by the Cost and Resource Use Standing Committee. The CSAC noted the minimal impact that including the SDS variables had on the results of the measures. CSAC members raised concerns about the SDS variables selected by the developers and the limited data available for analysis. The CSAC encouraged the measure developers to continue to explore additional data sets and other SDS variables in future updates to the measure. CSAC recommended a progress report on the consideration of other SDS factors at the next annual update of the measure.

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 <u>conceptual</u> 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

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.

Omission of two of the three conditions for endorsement during this current review

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.

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.

- o 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.
  - o 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.
  - o 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.
  - o Race was not included as a variable in the final risk adjustment model; rather it was only explored by the developer.
  - o 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.
  - o 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:
  - o 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.
  - o The developers expressed interest in potentially considering these factors in the model, but sought Committee input and recommendations on how to approach this.
  - o 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:
  o 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.
  o 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.
  o 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.
  o 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:
  o The guidance provided to Standing Committees on the selection and testing of SDS variables.
  o Consensus Standards Approval Committee (CSAC) approval of the revised measure evaluation criteria.
- NQF Response:
  o 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.

- o 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.
- o 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.
- o 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:
  - o 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:
  - o 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.

## *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 |