NATIONAL QUALITY FORUM

Memo

TO: Consensus Standards Approval Committee (CSAC)

FR: Melinda Murphy, Andrew Lyzenga, and Nadine Allen

RE: Appeal on Surgery Measure

DA: November 9, 2015

In accordance with the NQF Consensus Development Process (CDP), the measures recommended by the NQF Surgery Standing Committee for the Surgery 2014 project were released for a 30-day appeals period, which closed on October 6, 2015. NQF received one letter of appeal on behalf of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES). The appeal is pertinent to #2687: Hospital Visits after Hospital Outpatient Surgery (CMS/Yale-CORE). Measure #2687 is a measure of all-cause, unplanned hospital visits within 7 days of a same-day surgery performed in hospital outpatient departments; the level of measurement is hospitals.

The following documents are appended to this memo:

1. Appendix A - Appeal Letter: 2687: Hospital Visits after Hospital Outpatient Surgery
2. Appendix B - Response from measure developer (CMS/Yale-CORE)
3. Appendix C - Measure evaluation summary table

CSAC ACTION REQUIRED
The CSAC will review the letter of appeal, the response submitted by the developer, and this memo in consideration of the appeal. The CSAC will determine whether to uphold the endorsement decision or uphold the appeal for the measure.

Summary of Issues Raised in the Appeal:
The issues raised in the appeal include:

- The appellant questions the use of an all-cause approach to measuring readmissions, arguing that the measure does not adequately distinguish between admissions related to the surgical procedure and admissions that reflect underlying conditions or comorbidities unrelated to the procedure. The appellant suggests that "[hospital visits] for any reason would be unfairly pinned to the surgeon and the procedure, and would diminish the accuracy of the intended metric."
- The appellant also objects to the measure's approach to adjustment for socio-demographic (SDS) factors, claiming that "the measure, as written, attempts to address disparities among different races and different socio-economic status" but only addresses one race (African-American), ignoring potentially important factors such as language, culture, and education. The appellant also notes that the measure addresses socio-economic status using the proxy of dual-eligibility, which the appellant finds to be an unreliable indicator of potential disparities, particularly considering the variation in eligibility standards across states.

NATIONAL QUALITY FORUM

Memo

Summary of the Developer Response:
The developer defends the all-cause approach to measuring patient outcomes, noting that doing so encourages facilities to take action to minimize risks for a broad range of common problems that may be related or unrelated to a recent outpatient surgery. The developer also notes that the measure is risk adjusted for age, 24 comorbidities, and procedural complexity, and is reported as a ratio of predicted to expected number of visits so that facilities treating patients who are generally at higher risk are not disadvantaged in the measure.

In response to the appellant's concerns about adjustment for SDS factors, the developer points out that when this project was initiated, NQF had not yet started its SDS trial period, and therefore adjustment for SDS factors was prohibited per NQF's pre-trial policy. However, the developer notes that two SDS indicators (Medicaid dual-eligibility and race) that were readily available in CMS claims data were analyzed for their effect on measure scores; this analysis showed that these factors have little impact.

Summary of the Evaluation:
During its evaluation of this measure, the Surgery Standing Committee addressed each of the issues raised by the appellant. The Committee discussed the all-cause approach to readmissions measurement and the developer's risk adjustment approach, noting that other all-cause readmissions measures have appeared to drive improvement in readmission rates, and generally agreeing with the developer that the measure is likely to yield important information that will help facilities improve patient care.

The Committee also addressed the issue of SDS adjustment, suggesting that assessment of SDS factors' effect on the performance of measured entities will be important in affirming this measure's validity. Accordingly, the Committee requested that, as part of this measure's initial annual update, the developer provide updated information about the impact of SDS on the measure, including any changes that have been made to the measure to increase its sensitivity to SDS factors.

The Surgery Standing Committee approved the measure by a 18-1 vote, and both the CSAC and the NQF Board voted unanimously to approve the measure.

Additional details of the measure evaluation are included in Appendix C.

NATIONAL QUALITY FORUM

Memo

Appendix A — Appeal Letter: Measure 2687: Hospital Visits after Hospital Outpatient Surgery

The Quality, Safety, and Outcomes Committee of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) is pleased to be a partner organization with the National Quality Forum in creating, reviewing, and implementing safety measures focused on improving the experience and outcomes for all patients. We appreciate the opportunity to review the most recent phase 2 surgery recommendations, set for implementation in the latter part of this year. After a thorough evaluation of each of the proposed quality metrics, we would like to communicate significant concerns with one particular measure, #2687, Hospital Visits after Hospital Outpatient Surgery.

This measure quantifies the frequency with which patients discharged after an outpatient procedure return to the Emergency Department (ED) within a 7 day period. As written, this is an all-cause measure, with no discernment of relationship to the procedure or of patient demographics. There are significant problems with using this as a measure of quality for the surgeon or the institution. First, ED visits for any reason would be unfairly pinned to the surgeon and the procedure, and would diminish the accuracy of the intended metric. Additionally, the measure also would include all patients who are admitted to the hospital after a planned outpatient procedure. This is alarming as a potential metric for quality--especially in the elderly population. Many older patients undergoing outpatient procedures are ultimately admitted to the hospital for conditions that are not preventable and have very little to do with the surgical procedure. An example would be patients who are admitted after a planned outpatient surgery for observation for respiratory or cardiac monitoring secondary to baseline comorbidities. The measure proposed, as written, would therefore also penalize surgeons and hospitals for a problem that is related, in part, to the expected higher rate of certain comorbidities in elderly patients.

Second, and much more concerning, are the methods used to address the potential disparities of outcomes among populations. The measure, as written, attempts to address disparities among different races and different socio-economic status. The racial analysis only addresses one race, with no explanation of why that race was singled out, ignoring many races and cultures whose language barriers would make ED returns more likely as a result of potential communication difficulties. It also cannot be assumed that these disparities are accounted for by the evaluation of socioeconomic status, the other analysis performed in the measure. That potential disparity is assessed by equating low socioeconomic status to dual-eligibility. This comparison has inherent potential inaccuracies. While there are minimum national standards for Medicaid eligibility, states set their own eligibility standards, making a national comparison prone to errors. Additionally, these eligibility standards cannot be relied upon to fully address the potential disparities in language, culture, and education. Our fear is that the ultimate result of this is not just a flawed method of determining quality, but that the perception of these analysis flaws by surgeons and institutions would result in a potential barrier to access for these patients.

We appreciate the consideration of these concerns. SAGES is committed to partner with the NQF and recognizes the need for metrics to assess outcomes of outpatient procedures, but strives to support measures that are true and accurate representations of the quality of the care given to our patients.

NATIONAL QUALITY FORUM

Memo

Jonathan Dort, MD, FACS
Vice Chairman of Education
Director, Surgery Residency Program
Department of Surgery
Inova Fairfax Medical Campus

Member, Quality, Outcomes, and Safety Committee Society, American Gastrointestinal and Endoscopic Surgeons (SAGES)

NATIONAL QUALITY FORUM

Memo

Appendix B — Response from measure developer (CMS/Yale-CORE)

TO: National Quality Forum (NQF) Surgery Standing Committee

FROM: Mayur Desai, PhD, MPH, Yale New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (YNHHSC/CORE)

THROUGH: The Centers for Medicare & Medicaid Services (CMS)
Vinitha Meyyur, PhD

DATE: Wednesday, October 28, 2015

SUBJECT: Response to SAGES letter appealing approval of NQF 2687: Hospital Visits after Hospital Outpatient Surgery; Submitted October 21, 2015

Commenter: Jonathan Dort, MD, FACS, Society of American Gastrointestinal and Endoscopic Surgeons (SAGES)

Comment: The Quality, Safety, and Outcomes Committee of the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) is pleased to be a partner organization with the National Quality Forum in creating, reviewing, and implementing safety measures focused on improving the experience and outcomes for all patients. We appreciate the opportunity to review the most recent phase 2 surgery recommendations, set for implementation in the latter part of this year. After a thorough evaluation of each of the proposed quality metrics, we would like to communicate significant concerns with one particular measure, #2687, Hospital Visits after Hospital Outpatient Surgery.

This measure quantifies the frequency with which patients discharged after an outpatient procedure return to the Emergency Department (ED) within a 7 day period. As written, this is an all-cause measure, with no discernment of relationship to the procedure or of patient demographics. There are significant problems with using this as a measure of quality for the surgeon or the institution. First, ED visits for any reason would be unfairly pinned to the surgeon and the procedure, and would diminish the accuracy of the intended metric. Additionally, the measure also would include all patients who are admitted to the hospital after a planned outpatient procedure. This is alarming as a potential metric for quality--especially in the elderly population. Many older patients undergoing outpatient procedures are ultimately admitted to the hospital for conditions that are not preventable and have very little to do with the surgical procedure. An example would be patients who are admitted after a planned outpatient surgery for observation for respiratory or cardiac monitoring secondary to baseline comorbidities. The measure proposed, as written, would therefore also penalize surgeons and hospitals for a problem that is related, in part, to the expected higher rate of certain comorbidities in elderly patients.

Response: The Centers for Medicare & Medicaid Services (CMS) measures all-cause hospital visits to encourage facilities to minimize all types of risks that may lead to the need for a hospital visit after a same-day outpatient surgery. We agree that in an elderly population exacerbation of underlying

conditions may be related or unrelated to the procedure. Measuring all-cause patient outcomes encourages facilities to minimize the risk of a broad range of outcomes, including the risk of dehydration, pain, urinary retention, and arrhythmia. These are common problems that may be related or unrelated to a recent outpatient surgery. We have structured the measure so that facilities who most effectively minimize patient risk will perform better on the measure. As you point out, some providers care for patients who are inherently at higher risk; however, the measure is risk adjusted for baseline comorbidities so facilities treating patients who are more likely to experience visits because they have a generally higher risk are not disadvantaged in the measure.

The measure score is the ratio of the predicted to expected number of post-surgical hospital visits among the facility's patients. For all facilities, the expected number of hospital visits, given the facility's case mix and surgical procedure mix, is greater than zero since, as you point out, some patients will have visits for reasons unrelated to the surgical procedure. The numerator of the risk-standardized ratio is the number of hospital visits predicted for the facility's patients accounting for its observed rate, the number of surgeries performed at the facility, the case mix, and the surgical procedure mix. A ratio of less than one indicates the facility's patients were estimated as having fewer post-surgical visits than expected compared to facilities with similar surgical procedures and patients, and a ratio of greater than one indicates the facility's patients were estimated as having more visits than expected.

For fairness, the statistical model used to calculate the measure score adjusts for age, 24 comorbidity variables, and surgical complexity that vary across patient populations, are unrelated to quality, and influence the outcome in order to help ensure differences in the measure score do not reflect differences in case mix and surgical procedure mix across facilities. The measure risk adjusts for surgical procedural complexity using two variables. First, it adjusts for surgical procedural complexity using the Work RVU of the procedure. Second, it classifies each surgery into an anatomical body system group using the Agency for Healthcare Research and Quality (AHRQ) Clinical Classification System (CCS). The measure uses the body system variable, in addition to the Work RVU of the surgery, to account for organ-specific differences in risk and complications which are not adequately captured by the Work RVU alone. This approach to risk adjustment for surgical procedural complexity is similar to that described in the literature and used for risk adjustment in the American College of Surgeons' National Surgical Quality Improvement Program (NSQIP).

Comment: Second, and much more concerning, are the methods used to address the potential disparities of outcomes among populations. The measure, as written, attempts to address disparities among different races and different socio-economic status. The racial analysis only addresses one race, with no explanation of why that race was singled out, ignoring many races and cultures whose language barriers would make ED returns more likely as a result of potential communication difficulties. It also cannot be assumed that these disparities are accounted for by the evaluation of socioeconomic status, the other analysis performed in the measure. That potential disparity is assessed by equating low socioeconomic status to dual-eligibility. This comparison has inherent potential inaccuracies. While there are minimum national standards for Medicaid eligibility, states set their own eligibility standards, making a national comparison prone to errors. Additionally, these eligibility standards cannot be relied upon to fully address the potential disparities in language, culture, and education. Our fear is that the

ultimate result of this is not just a flawed method of determining quality, but that the perception of these analysis flaws by surgeons and institutions would result in a potential barrier to access for these patients.

Response: We appreciate your concern about the potential effects of sociodemographic status (SDS) on the measure score.

Please note that the measure was submitted to NQF before NQF's SDS trial period went into effect on April 15, 2015. Consistent with the pre-trial NQF guidance on SDS, we evaluated the potential effects of risk adjusting for two SDS indicators – Medicaid-dual eligibility and race. These variable are readily available in the CMS claims data. In addition, use of Medicaid eligibility status as a proxy for SDS is consistent with prior research as well as NQF recommendations (https://www.qualityforum.org/projects/Patient_Outcome_Measures_Phases1-2.aspx). Our results show that adjusting for these factors at the patient level does little to change the measure scores; unadjusted and adjusted hospital outpatient department (HOPD) riskstandardized hospital visit (RSHV) ratios are highly correlated (Pearson correlation 0.990 and 0.998 for adjustment for Medicaid-dual eligibility and race, respectively). This suggests that including a patient-level risk adjuster for SDS will make little difference in the measure results after accounting for other factors already adjusted for in the model, such as age, comorbidities, and the complexity of the surgery.

In addition, to explore whether there might be differences in HOPD RSHV ratios by the proportion of lower SDS patients hospitals care for, we examined the distribution of measure scores by quartiles of both percentage of dual-eligible patients and percentage of African American patients. Although the results show a trend toward higher measure scores in the highest quartile of lower SDS patients, they also show that some hospitals with relatively high proportions of lower SDS patients can and do perform well on the measure.

We cannot tell from these analyses what is causing the observed differences across quartiles of proportion of lower SDS patients. One of the potential causes is differences related to quality. For example, some hospitals may be better able than other hospitals to meet the needs of patients with language barriers or more limited resources. Given these findings, on balance we do not recommend adjusting the measure for SDS at this time. Doing so will not appreciably change the measure scores and might contribute to masking disparities in care. Indeed, we have designed the measure to incentivize facilities and providers to improve systems of care that will lead to better outcomes for all patients.

CMS is participating fully in the NQF trial and is actively working to further consider issues related to adjusting for SDS. In addition, CMS notes that the Office of the Assistant Secretary for Planning and Evaluation (ASPE) is conducting research on the issue of risk adjustment for socioeconomic status as directed by the IMPACT Act and will issue a report to Congress by October 2016. CMS will closely examine the recommendations issued by ASPE and consider how they apply to this and other CMS quality measures.

NATIONAL QUALITY FORUM

Memo

Appendix C — Measure 2687 Evaluation Summary Table

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed
Submission | Specifications
Description: Facility-level, post-surgical risk-standardized hospital visit ratio (RSHVR) of the predicted to expected number of all-cause, unplanned hospital visits within 7 days of a same-day surgery at a hospital outpatient department (HOPD) among Medicare fee-for-service (FFS) patients aged 65 years and older. **Numerator Statement**: The outcome is all-cause, unplanned hospital visits, defined as 1) an inpatient admission directly after the surgery or 2) an unplanned hospital visit (emergency department [ED] visit, observation stay, or unplanned inpatient admission) occurring after discharge and within 7 days of the surgical procedure. **Denominator Statement**: Outpatient same-day surgeries performed at HOPDs for Medicare FFS patients aged 65 years and older with the exception of eye surgeries and same day surgeries performed concurrently with high-risk procedures. **Exclusions**: The measure excludes surgeries for patients without continuous enrollment in Medicare FFS Parts A and B in the 1 month after the surgery. The measure excludes these patients to ensure all patients have full data available for outcome assessment. The exclusion prevents unfair distortion of performance results. The measure excludes surgeries for patients without continuous enrollment in Medicare FFS Parts A and B in the 1 month after the surgery. **Adjustment/Stratification**: No risk adjustment or risk stratification **Level of Analysis:** Facility **Setting of Care:** Other **Type of Measure**: Outcome **Data Source**: Administrative claims **Measure Steward**: The Centers for Medicare & Medicaid Services (CMS)
STANDING COMMITTEE MEETING [03/19/2015-03/20/2015] **1. Importance to Measure and Report: The measure meets the Importance criteria** *(1a. Evidence: 1b. Performance Gap)*

NATIONAL QUALITY FORUM

Memo

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed

1a. Evidence: **Y-15**; **N-3**; 1b. Performance Gap: **H-5**; **M-13**; **L-2**; **I-0**

Rationale:

- The developers provided a rationale for the measure, specifically that there are interventions and strategies that may reduce unplanned hospital visits after same-day surgery, including appropriate patient selection, patient education, and nausea and pain management. The developer clarified the difference between an unplanned and planned visit and noted that they recommend reporting the measure as a ratio rather than a rate.
- The Committee concluded there is minimal evidence that ties specific processes to the outcome but that the rationale is sufficient to support the measure.
- The developer assessed provider-level variation in performance scores using data from a 20 percent sample of 2010 Medicare fee-for-service claims that represented 4,234 HOPDs and 212,104 surgeries. The measure developers found that the high performing HOPD's (at or below the 5th percentile) had at least 24 percent fewer than expected surgical hospital visits and those in the 95th percentile had at least 34 percent more hospital visits than what they were expecting given the case and surgical procedure mix.
- Some Committee members had concerns about being able to determine if there is a performance gap given a small sample size; however, the Committee generally agreed that the evidence is sufficient.

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-2**; **M-15**; **L-0**; **I-0**; 2b. Validity: **H-3**; **M-16**; **L-0**; **I-0**

Rationale:

- The data used in testing the reliability of the performance measure score were derived from 2009-2011 Medicare fee-for-service (FFS) claims. These data included a 20 percent sample of same-day surgery claims from Part B (physician) claims, which were then matched to the corresponding hospital claims. The developer conducted a "test-retest" approach by randomly selecting half of the patients from each HOPD into two datasets. They then calculated the risk-standardized hospital visit ratios for each HOPD in each of the datasets, then compared the agreement between the scores for the HOPDs using the Intraclass Correlation Coefficient (ICC) The ICC value was 0.50 (95 percent CI: 0.48-0.53), indicating "moderate" agreement according to the categorization by Landis and Koch.
- Face validity of the performance measure score was assessed by a Technical Expert Panel comprised of 15 patient representatives, expert clinicians, methodologist, researchers, and providers. Of the 13 experts who responded, 92.3 percent either strongly or moderately agreed that this measure can accurately distinguish better and worse quality facilities.
- The Committee generally found the reliability and validity information submitted by the developers to be sufficient.

NATIONAL QUALITY FORUM

Memo

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed

3. Feasibility: H-16; M-3; L-0; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c.Susceptibility to inaccuracies/ unintended consequences identified 3d. Data collection strategy can be implemented)

Rationale:

- The data source for this measure is Medicare administrative claims and enrollment data, and therefore all data elements are in defined fields.
- The Committee was satisfied with the feasibility of this measure.

4. Usability and Use: H-6; M-11; L-1; I-0

(Meaningful, understandable, and useful to the intended audiences for 4a. Public Reporting/Accountability and 4b. Quality Improvement)

Rationale:

- The Committee was generally satisfied with the use and usability of this measure and would like the comments that have been made to be addressed at the next cycle for the measure.

5. Related and Competing Measures

- This measure is related to 2539 Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy, Rate of risk-standardized, all-cause, unplanned hospital visits within 7 days of an outpatient colonoscopy among Medicare fee-for-service (FFS) patients aged 65 years and older.
- The Committee recommended that the need for two similar measures, as well as harmonization and unintended consequences should be assessed during annual updates once the two new measures have been in use for some time so that any potentially needed adjustments could be considered for each measure independently.

Standing Committee Recommendation for Endorsement: **Y-18**; **N-1**

6. Public and Member Comment: April 17, 2015 – May 18, 2015 (Additional 15-day Public and Member Comment: May 22, 2015 to June 5, 2015)

Comments received:

- One commenter expressed uncertainty about the feasibility of this measure, citing that a free-standing surgical center would have no mechanism to recall patients. Additionally, hospitals and ambulatory surgical centers that have urgent care facilities would be penalized for providing patient access, per the current measure language.
- Another commenter noted that CMS Planned Readmission Algorithm 3.0 was used to identify those procedures or conditions that typically result in planned admissions. The commenter noted that this algorithm has been tested for the inpatient care and has not been tested for the ambulatory care setting. The commenter further noted that outpatient surgery procedures that

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed

are planned admissions are different and unique to this setting; and questioned that by using this inpatient algorithm, that there has been a compromise in developing a comprehensive list of planned admissions for procedures performed in ambulatory surgery centers.

- Lastly, two commenters noted that NQF is currently holding a trial period under which measures may be risk-adjusted for patients' socioeconomic status and other demographic factors (SDS). The commenters suggested that SDS adjustment for measure #2687 (Hospital Visits After Outpatient Surgery) may be appropriate, and questioned why this had not been discussed or considered by the Standing Committee. Commenters also observed that a measure (#2539: Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy) similar to measure #2687 (Hospital Visits after Hospital Outpatient Surgery) was recently endorsed by NQF's Readmissions Standing Committee, and questioned why the Surgery Standing Committee had not addressed harmonization of these two measures

NQF response:

- NQF appreciates your comment and the opportunity to provide clarification. Previous NQF policy prohibited the inclusion of sociodemographic status (SDS) factors in risk-adjustment approaches out of concern that doing so might conceal inequalities in care and result in lower standards of provider performance for certain subpopulations. However, in 2014, NQF convened a multi-stakeholder panel of experts in healthcare performance measurement and disparities to consider if, when, and how performance measures should be adjusted for SDS. After its deliberations, the Expert Panel recommended that NQF should allow inclusion of SDS factors in the risk-adjustment approach for performance measures when conceptual reasons and empirical evidence demonstrate it is appropriate. The NQF Board of Directors reviewed the Expert Panel's recommendations and decided to temporarily change NQF's policy and evaluate its impact during the course of a two-year trial period. This trial period went into effect on April 15, 2015, meaning that projects with measure submission deadlines before that date fell under NQF's previous policy/guidance on SDS adjustment, while projects with measure submission deadlines after that date are subject to the trial policy on SDS adjustment. The2015 Surgery project's measure submission deadline was January 14, 2015, prior to the start of NQF's SDS trial period. Therefore, both the developer and the Surgery Standing Committee conformed to the previous policy regarding inclusion of SDS factors in the risk-adjustment approach.

Developer response:

- Thank you for raising these two potential concerns; we would like to clarify, however, that the measure as designed does not assess either ambulatory surgery centers or free standing urgent care facilities. The measure includes outpatient same-day surgeries performed at hospital outpatient departments only; it does not include procedures performed at ambulatory surgery centers. Likewise, the measure does not affect urgent care facilities. They are not measured, and visits to urgent care facilities are not counted in the measure outcome, which only includes hospital emergency department visits, observation stays, or unplanned inpatient admissions.
- We appreciate the question and the opportunity to clarify why it makes sense to use an algorithm developed for hospital readmission measures in this measure, which as you note focuses on same-day surgery rather than admitted patients. The CMS Planned Readmission Algorithm was developed to identify all admissions (rather than readmissions per se) that are planned. That is, it uses condition and procedure codes to distinguish between admissions to address acute illness and injury from admissions of stable patients that are for planned procedures (such as for chemotherapy or a hip replacement). We use the algorithm in this

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed

measure because our goal here is the same as it was for the hospital readmission measures – we do not want to include in our measure outcome admissions that are planned, since they are not a signal of care quality. We did review the algorithm carefully to make sure the way we identify the planned admissions makes sense in the context of this surgery measure, and shared the details of the algorithm with our technical expert panel, the public, and NQF reviewers. If you have specific suggestions for ways the algorithm should be adapted for this particular measure, we are happy to consider them.

- We appreciate your concern about the potential effects of SDS on the measure score. We wanted to address your comments on both the process of review and the substance of our conclusions in the NQF application based on the SDS analysis we conducted for the application. Regarding the process, the surgery measure is not technically in NQF's SDS pilot. "This trial period went into effect on April 15, 2015. This means that projects with measure submission deadlines before that date fell under NQF's previous policy/guidance on SDS adjustment, while projects with measure submission deadlines after that date are subject to the trial policy on SDS adjustment. Since the 2015 Surgery project's measure submission deadline was January 14, 2015, both the developer and the Surgery Standing Committee conformed to the [pre-trial] policy regarding inclusion of SDS factors in the risk-adjustment approach (email from Andrew Lyzenga at NQF, June 15, 2015)."

  Regarding the substance of your concern, consistent with the pre-trial NQF guidance on SDS, we evaluated the potential effects of risk adjusting for two SDS indicators – Medicaid-dual eligibility and race. These variable are readily available in the CMS claims data. In addition, use of Medicaid eligibility status as a proxy for SDS is consistent with prior research as well as NQF recommendations (http://www.qualityforum.org/projects/Patient_Outcome_Measures_Phases1-2.aspx). Our results show that adjusting for these factors at the patient level does little to change the measure scores; unadjusted and adjusted HOPD risk-standardized hospital visit (RSHV) ratios are highly correlated (Pearson correlation 0.990 and 0.998 for adjustment for Medicaid-dual eligibility and race, respectively). This suggests that including a patient-level risk adjuster for SDS will make little difference in the measure results after accounting for other factors already adjusted for in the model, such as age, comorbidities, and the complexity of the surgery.

  In addition, to explore whether there might be differences in HOPD RSHV ratios by the proportion of lower SDS patients hospitals care for, we examined the distribution of measure scores by quartiles of both percentage of dual-eligible patients and percentage of African American patients. Although the results show a trend toward higher measure scores in the highest quartile of lower SDS patients, they also show that some hospitals with relatively high proportions of lower SDS patients can and do perform well on the measure. We cannot tell from these analyses what is causing the observed differences across quartiles of proportion of lower SDS patients. One of the potential causes is differences related to quality. For example, some hospitals may be better able than other hospitals to meet the needs of patients with low literacy. Given these findings, on balance we do not recommend adjusting the measure for SDS at this time. Doing so will not appreciably change the measure scores and might contribute to

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed

masking disparities in care.

CMS is participating fully in the NQF trial and is actively working to further consider issues related to adjusting for SDS. In addition, CMS notes that the Office of the Assistant Secretary for Planning and Evaluation (ASPE) is conducting research on the issue of risk adjustment for socioeconomic status as directed by the IMPACT Act and will issue a report to Congress by October 2016. CMS will closely examine the recommendations issued by ASPE and consider how they apply to this and other CMS quality measures.

CMS did consider the effect of adjusting for SDS and reported the results in the NQF application. As discussed in the application and in response to the question above, we do not recommend adjusting for SDS at this time, so testing the reliability of the measure with SDS adjustment is not necessary at this time. As you note, reliability testing yielded an intraclass correlation coefficient (ICC) of 0.50, which according to conventional interpretation is "moderate." It should be noted that this ICC value is consistent with those of other CMS claims-based measures. In addition, measure testing was conducted using a 20% sample of Medicare Fee-for-Service data. We expect the reliability score will be higher in the national 100% sample where individual facility volumes would be higher yielding more reliable individual facility results. The 100% sample would be used for public reporting.

The present measure (NQF # 2687) is already fully harmonized with NQF # 2539 (Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy) on areas of the methodology that are analogous. Specifically, both measures use the same outcome. For both the outpatient surgery measure and the outpatient colonoscopy measure, the outcome is identically specified as all-cause, unplanned hospital visits, defined as 1) an inpatient admission directly after the procedure, or 2) an unplanned hospital visit (emergency department visit, observation stay, or unplanned inpatient admission) occurring after discharge and within 7 days of the outpatient procedure.

We believe that the measure will yield important information that will help facilities improve patient care. Measure testing demonstrated significant variation in risk-standardized performance across facilities, indicating opportunities for quality improvement. Facilities with a higher than expected number of outcomes will be able to review and improve their processes around preparing the patient for surgery, the surgery itself, and follow-up care. In addition, in implementing the measure, CMS would provide each facility with patient-level data so that facilities could examine the specific causes of higher than expected outcome.

Reliability testing yielded an intraclass correlation coefficient (ICC) of 0.50, which according to conventional interpretation is "moderate." It should be noted that this ICC value is consistent with those of other CMS claims-based measures. In addition, measure testing was conducted using a 20% sample of Medicare Fee-for-Service data. We expect the reliability score will be higher in the national 100% sample where individual facility volumes would be higher yielding more reliable individual facility results. The 100% sample would be used for public reporting.

Committee Response:

- The Committee appreciates the opportunity to provide clarification regarding the setting of interest. Given the care setting to which the measure applies, the Committee believes the

2687 Hospital Visits after Hospital Outpatient Surgery: Endorsed
expressed concerns are mitigated. • The Committee also appreciates the precision requested by the commenter as well as the clarity provided by the developer. During the in-person meeting the Committee agreed that the specifications of the measure were appropriate. • Finally, the Committee appreciates the position of NQF, the participation by CMS in the SDS trial as outlined in NQF policy, and CMS commitment regarding recommendations from ASPE research. During the in-person meeting the Committee agreed that the datasets, approach to testing and testing outcome was sufficient to move the measure forward. As part of the annual update to the measure, the Committee anticipates updated information about SDS impact including any changes to the measure to increase SDS sensitivity as well as any changes required to ensure its full alignment with 2539. With respect to harmonization, the Committee agreed that it was appropriate to assess the impact and implementation of the two new measures independently before further consideration about how additional alignment might occur.
7. Consensus Standards Approval Committee (CSAC) Vote (August 18, 2015): Y-13; N-0; A-0 • **Decision**: Approved for Endorsement
8. Board of Directors Vote (September 2, 2015): Y-6; N-0; Recusal-1 • **Decision:** Ratified for Endorsement
9. Appeals