NATIONAL QUALITY FORUM

Memo

June 4, 2018

To: Consensus Standards Approval Committee (CSAC)

From: Surgery Project Team

Re: Surgery, Fall 2017 Cycle

CSAC Action Required

The CSAC will review recommendations from the Surgery Standing Committee at its June 4-5, 2018 meeting and vote on whether to uphold the recommendations from the Committee.

This memo includes a summary of the project, recommended measures, and identified themes and responses to the public and member comments. The following documents accompany this memo:

1. **Surgery, fall 2017 cycle draft report.** The draft report has been updated to reflect the changes made following the Standing Committee's discussion of public and member comments. The complete draft report and supplemental materials are available on the project webpage.
2. Comment Table. Staff has identified themes within the comments received. This table lists three comments received during the post-meeting comment period and the NQF and Standing Committee responses.

Background

This report reflects the review of measures in the surgery project. The measures in NQF's surgery endorsement project focus on key surgical care processes across an array of procedure types that include outcomes for general and subspecialty surgical procedures, including cardiac, orthopedic, ophthalmological, and vascular surgeries and procedures, and all phases of perioperative care. In this project, measures focused on lung resection and lobectomy for lung cancer and hospital visits following general surgery procedures.

The 24-member Surgery Standing Committee reviewed three measures—one maintenance measure and two new measures—and all were recommended for endorsement.

Draft Report

The Surgery fall 2017 cycle draft report presents the results of the evaluation of three measures considered for endorsement under the Consensus Development Process (CDP). All three measures are recommended for endorsement.

The measures were evaluated against the 2017 version of the measure evaluation criteria.

	Maintenance	New	Total
Measures under consideration	2	1	3
Measures recommended for endorsement	2	1	3
Measures recommended for inactive endorsement with reserve status	0	0	0
Measures approved for trial use	0	0	0
Measures not recommended for endorsement or trial use	0	0	0
Measures withdrawn from consideration	0	0	0
Reasons for not recommending	Importance – 0 Scientific Acceptability - 0 Overall – 0 Competing Measure – 0	Importance – 0 Scientific Acceptability – 0 Overall – 0 Competing Measure – 0	

CSAC Action Required

Pursuant to the CDP, the CSAC is asked to consider endorsement of three candidate consensus measures.

Measures Recommended for Endorsement

- 1790 Risk-Adjusted Morbidity and Mortality for Lung Cancer Resection for Lung Cancer (Society of Thoracic Surgeons)

  Overall Suitability for Endorsement: Yes-16; No-0

- 3294 STS Lobectomy for Lung Cancer Composite Score (Society for Thoracic Surgeons)

  Overall Suitability for Endorsement: Yes-16; No-0

- 3357 Facility Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers (Centers for Medicare & Medicaid Services/Yale CORE)

  Overall Suitability for Endorsement: Yes-16; No-0

Comments and Their Disposition

NQF received three comments from three organizations (including two member organizations) and individuals pertaining to the draft report and to the measures under consideration.

A table of comments submitted during the comment period, with responses to each comment and the actions taken by the Standing Committee and measure developers, is posted to the Surgery project webpage.

3357 Facility Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers

Generally, the three comments received addressed measure specifications, measure validity, sociodemographic risk adjustment, and whether the measure provides information about meaningful differences in performance and is actionable in order to improve the quality of surgical care. Commenters questioned the inclusion of skin procedures in the measure specifications because general surgeons do not routinely perform these procedures.

Commenters also questioned whether skin procedures were included to boost low case volume. Two comments noted that the measure should be risk adjusted for social risk factors, such as socioeconomic status (SES), for providers who serve low SES clients. One commenter pointed out that unplanned visits are not always in the surgeon's control and could be related to issues of access to care. Submitted comments also noted that the measure is "topped out" and does not provide enough variation in performance to discern quality of care.

Measure Steward/Developer Response

Measure Outcome - The commenter stated that the "readmission measure is not a good proxy for driving improvement in surgical care." We would like to clarify that the measure under review by the NQF (NQF ID 3357) is not a readmission measure. The measure's outcome is any unplanned hospital visit, defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission, occurring within 7 days of a general surgery procedure performed at an ambulatory surgical center (ASC). The outcome of hospital visits is the focus of this measure because it is a broad, patient-centered outcome that captures the full range of hospital visits resulting from adverse events or poor care coordination following an ASC general surgery procedure. This measure's goal is to assess and illuminate variation in risk-adjusted hospital visits following surgery, for quality improvement and public reporting.

Alignment with Registry-Based Measures - The commenter "believes that surgical measurement should be built on four key principles: 1) setting the standards, 2) building the right infrastructure, 3) using the right data, and 4) verifying with outside experts," and advocates that ASC measures "include standards-based facility-level verification programs, patient reported experience (PRE) and outcome (PRO) measures, and traditional quality measures including registry and claims-based measures."

This 7-day hospital visit measure fits within the above framework. It increases surgeon and ASC accountability for and awareness of patient outcomes during the post-surgical period, and provides facilities with patient-level data on outcomes to inform quality improvement that ASCs currently lack. The availability of linkable Medicare claims data to calculate the risk-adjusted measure and link patient outcomes, procedure, and risk factor data across settings makes this a low-burden measure that can provide valuable

information and complement registry-based, surgery-society developed measures such as PROs.

We would like to further clarify that our measure development process is aligned with the four key principles that the commenter identified as foundational for surgical quality measurement. We built the measure using standards for quality measure development set forth by CMS in its Measure Management System Blueprint. The measure was developed with input from general surgery consultants, a national Technical Expert Panel, and the public.

Usability and Actionable Information - Two commenters expressed concern that the measure would be of limited utility, and that the information it provides is not actionable.

For ASCs, we believe measuring and publicly reporting claims-based, risk-adjusted measure scores will encourage ASCs to engage in quality improvement and lead to better patient care over time. Further, CMS plans to implement the measure to optimize its usability. Prior to public reporting, CMS anticipates providing claims-detail and

facility-specific reports, as the Agency does for other outcome measures. These reports will allow ASCs to see patient 7-day outcomes that are currently not visible to them. This information will help ASCs understand their performance, inform quality improvement efforts, and improve the care they provide to patients.

Variation in Performance Scores and Identification of Outliers - All three commenters expressed their view that there is not enough variation in the measure results to show meaningful differences between facilities. One commenter expressed concern that the measure identifies relatively few outliers as better or worse than expected.

We appreciate the commenters' concerns. The measure score results, however, do present a clinically meaningful range in risk-adjusted outcome rates. As presented in the public comment technical report using Medicare FFS CY 2015 data, we found that the facility measures scores ranged from 0.94% to 4.55%, with a median risk-standardized hospital visit rate of 2.19% (the 25th and 75th percentiles were 2.03% and 2.46%, respectively). The variation in these rates provides a quality signal, and reporting facility-level measure scores will improve transparency and promote quality improvement.

To assist consumers with interpreting the measure, we provide a descriptive category of facility quality (better than, worse than, or no different than the national rate); to provide ASCs and other users with richer insight into performance, we provide the estimated 7-day hospital visit rate and the 95% interval estimate (uncertainty estimate) around that rate. As the commenters pointed out, the descriptive approach categorizes relatively few facilities as outliers. The approach to categorizing facility outliers is very conservative by design. It uses 95% confidence interval (uncertainty) estimates to identify outliers.

Measuring quality of care associated with general surgery procedures performed at ASCs would bring awareness to ASCs and provide valuable data to patients. As intended,

we expect this measure will promote patient-centered improvement in care provided at ASCs, because measurement coupled with transparency will make visible, for the first time, the rate of hospital visits after general surgery ASC procedures to both patients and ASCs.

Emergency Department (ED) Visits and Observation Stays - One commenter was concerned that the planned admissions algorithm was not applied to ED visits or hospital observation stays, and that counting all ED visits as unplanned did not account for lack of "patient access to care in the desirable setting." As the commenter noted, we have previously stated that while we understand that the ED and hospital observation setting may be used for planned care at times, the measure is structured to count these events because these settings are not usually a desirable setting for planned care from the patient's point of view.

Also, ASCs are expected to limit their cases to those that can be safely performed outside the hospital setting. Higher rates of serious but potentially preventable complications that result in ED visits and observation stays may be a sign of poorer quality ASC care. ASCs can reduce the likelihood of these serious complications by emphasizing patient safety and reducing rates of complications or harm events, and by discharge planning that anticipates the need for potential office-based follow up care.

Measure Cohort - One commenter noted that this measure "is meant to capture all other routinely performed outpatient surgical procedures," unlike the other measures in CMS's ASCQR program that focus on colonoscopy, orthopedic, or urology procedures.

For this measure, we clarify that we targeted procedures that fall within the scope of general surgery, including those performed by other surgical specialists. We combined these procedures because they share the risk of post-surgery hospital visits within 7 days, they share common reasons for return to the hospital, and the risk of hospital visits following these procedures can be mitigated through similar strategies.

Additionally, three commenters expressed concerns that the measure includes many skin and plastic surgery procedures. We understand the commenters' concerns that over half the procedures in the cohort are skin procedures and concern that these procedures were included to increase sample size. The commenters also stated that many of the included procedures are performed by surgeons other than general surgeons.

As noted above, we included procedures within the scope of practice of general surgery even though these procedures are often performed by other subspecialists because they share common features that allow us to combine them for assessing quality.

However, the commenters are correct that in addition to general surgeons, other types of surgeons and non-surgical specialists perform skin and other procedures included in the cohort. As we clarified above, we included these procedures in a single measure because the procedures share (1) a risk of post-surgery hospital visits within 7 days and

(2) relatively similar reasons for return to the hospital. Members of our TEP also felt the care practices that would best lower the risk of hospital visits were similar across these procedures. Procedural volume was not a criterion for inclusion of procedures in the cohort.

Further, we identified and refined the group of procedures to include in the cohort through multi-stakeholder review and input from a national TEP, general surgery consultants, and the public. For example, in light of comments receive on the measure during measure development public comment, we re-reviewed all of the individual CPT codes within CCS categories and removed 15 individual procedures (CPT® codes) from the measure that were outside the scope of general surgery practice.

Adjustment for Social Risk Factors -Two commenters expressed concern that the measure is not adjusted for social risk factors.

As previously acknowledged in the conceptual model we presented in the NQF application, we agree that patients' socioeconomic status (SES) affects health and health outcomes in important ways.

Our intent when developing and testing the measure was to be responsive to the NQF guidelines for measure developers on SES. We therefore examined three patient-level indicators of social risk that are reliably available for all Medicare beneficiaries: 1) Medicare-Medicaid dual eligibility, 2) race, and 3) the AHRQ SES Index. The variables used are aligned with those the National Academy of Medicine committee identified as available for use in outcome measures. We examined whether these factors were associated with increased risk in hospital visits after adjusting for other risk factors and evaluated the impact of social risk factors on ASC-level measure scores.

While including each of these risk factors in our models indicated a statistically significant association after controlling for other risk-adjusters, results showed that the effect of social risk factors on hospital visit rates in the fully adjusted model was significant but small. Additionally, inclusion of these variables did not change ASCs' risk-standardized hospital visit ratios (RSHVRs) or their performance on the measures.

Correlation coefficients between RSHVRs with and without adjustment for these factors were near 1 (0.998, 1.000, and 0.999 for dual-eligible, African-American, and low SES patients, respectively) and mean differences in RSHVRs were near zero (0.0000, -0.0001, and -0.0002 for dual-eligible, African-American, and low SES patients, respectively).

In addition, we examined the relationship between the proportion of low SES patients and the facility-level score, focusing on facilities with the highest proportion of dual eligible patients (fourth quartile). This analysis did not show a clear relationship between the proportion of low SES patients and the facility-level score (Pearson correlation coefficient = -0.17).

Based on these findings, and a consideration of how social risk factors affect patients in the ambulatory setting and the importance of efforts to address all patients' needs, CMS decided to not adjust the models for these social risk factors; not adjusting is not likely

to lead to unintended consequences, or burden providers that serve low SES patients, and adjusting may mask quality differences. However, once the measure is implemented, CMS will monitor this measure, as with others, for unintended consequences related to disparities.

Finally, we acknowledge the importance of optimizing measures to incentivize high quality care for all while ensuring providers caring for low SES patients are not disadvantaged on the measures. CORE, with CMS, is exploring alternative modeling approaches that better illuminate how ASCs and their patients contribute to SES-related risks, and will continue to explore incorporating social risk factors into quality measures.

Attribution of Outcomes - We appreciate the commenter's continued review of the top reasons for any hospital visit within 7 days of general surgery procedures.

As previously clarified, the diagnoses referred to by the commenter (Table 4 of measure documentation) can occur during an admission, ED visit, or observation stay. If they occur during an admission, then all but one type (acquired absence of breast and nipple) are identified as planned admissions and are not counted in the measure outcome. If these diagnoses (including cancer diagnoses) occur as part of an ED visit or observation stay, they are included in the measure outcome because ED visits and observation stays are not routinely used for planned care. We understand that the ED and hospital observation setting may be used for planned care at times, but the measure is structured to count these events because these settings are not usually a desirable setting for planned care from the patient's point of view.

CORE is committed to evaluating whether refining the CMS planned admission algorithm will better capture planned admissions for the diagnoses flagged by comments. As such, during measure reevaluation, CORE will consider updating the planned admission algorithm to include acquired absence of breast and nipple so that admissions with this diagnosis would not be counted in the measure outcome.

Measure Title - We appreciate the commenter's suggestion to rename the measure. We have already renamed the measure to address this concern. The measure was previously named "Hospital Visits After General Surgery Ambulatory Surgical Center Procedures," and we revised its title to be "Facility-Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers." Our intent was to emphasize the scope of the procedures included in the measure cohort rather than the types of specialists performing them. The scope of the measure was defined by the scope of practice of general surgeons. Thus, we have chosen not to include "skin procedures" or "plastic repair" in the title given that many types of procedures in the measure are performed by both general surgeons and other specialists, and including one specific procedure type in the title would make the scope of the measure less clear. We believe the measure title accurately reflects what it assesses. CMS welcomes continued suggestions on the best name for the measure.

Low ASC Case Volumes - The commenter expressed concern that this measure would provide insufficient information for low-volume facilities. We understand the

commenter's concern that the measure would not provide sufficient information about the quality of care in individual low-volume facilities. For this measure, as is done for other risk-adjusted outcome measures, CMS will set minimum-volume requirements for reporting and only report scores for ASCs that have an adequate number of cases to generate reliable estimates. For example, for publicly reported outcome measures such as CMS's hospital readmission measures, CMS implemented minimum volume requirements for reporting.

The commenters also expressed concern about CMS implementing this measure using an inadequate amount of data. The commenter referenced CMS's use of only 1 year of claims data for implementation of another risk-adjusted outcome measure (ASC-12: Facility Seven-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy) in the ASCQR program, instead of 3 years of claims data. For this general surgery measure, CMS will consider using multiple years of claims data for public reporting. We calculated measure score (test, retest) reliability for a 2-year reporting period and found that the agreement (intraclass coefficient) between the two RSHVR values for each ASC was 0.526, indicating moderate measure score reliability. NQF committees consider their evaluation criteria to be rigorous, which state that moderate or high reliability is typically required for endorsement.

Prior to measure implementation, CMS will evaluate the amount of data required for reliable measure score calculation, and determine the number of years of data to use, weighing the tradeoffs between having an adequate number of cases for the greatest number of facilities and ensuring data used are timely.

Committee Member Response

Committee members stated that they are satisfied with the developer's response to the public comments, and agreed that the inclusion of skin procedures is not detrimental to the measure. The Committee stated that skin procedures should be included because they are an increasing part of care provided at ambulatory surgical centers (ASC). They also noted that the measure fills a gap in surgical care and provides valuable performance data to ASCs that they would not otherwise receive. The Committee also noted that the ability of the risk model to discriminate differences in patient characteristics could be improved (c-statistic = 0.69) and that they would like to see a more robust c-statistic and additional data on the upward mobility of patients when the measure is resubmitted for endorsement consideration.

NQF Response

The preliminary analysis for this measure included a summary from the MAP 2018 Considerations for Implementing Measures in Federal Programs: Hospitals Final Report. NQF provided the Standing Committee with the 10 comments submitted to the MAP Hospital Workgroup for review prior to the Post-Comment Call held on May 3, 2018.

Member Expression of Support

Throughout the 16-week continuous public commenting period, NQF members had the opportunity to express their support ('support' or 'do not support') for each measure submitted

for endorsement consideration to inform the Committee's recommendations. Two NQF members were not supportive of 3357 *Facility Level 7-day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers*. Appendix B details the expression of support.

Removal of NQF Endorsement

One measure previously endorsed by NQF was not re-submitted by the measure steward/developer, and endorsement has been removed.

Measure	Measure Description	Reason for Removal of Endorsement
0534 Hospital Specific Risk-Adjusted Measure of Mortality or One or More Major Complications Within 30 Days of a Lower Extremity Bypass (LEB)	Hospital specific risk-adjusted measure of mortality or one or more of the following major complications (cardiac arrest, myocardial infarction, CVA/stroke, on ventilator >48 hours, acute renal failure (requiring dialysis), bleeding/transfusions, graft/prosthesis/flap failure, septic shock, sepsis, and organ space surgical site infection), within 30 days of a lower extremity bypass (LEB) in patients age 16 and older	Developer does not have the resources and personnel to maintain the measure.

Appendix A: CSAC Checklist

The table below lists the key considerations to inform the CSAC's review of the measures submitted for endorsement consideration.

Key Consideration	Yes/No	Notes
Were there any process concerns raised during the CDP project? If so, briefly explain.	No	
Did the Standing Committee receive requests for reconsideration? If so, briefly explain.	No	
Did the Standing Committee overturn any of the Scientific Methods Panel's ratings of Scientific Acceptability? If so, state the measure and why the measure was overturned.	No	
If a recommended measure is a related and/or competing measure, was a rationale provided for the Standing Committee's recommendation? If not, briefly explain.	No	
Were any measurement gap areas addressed? If so, identify the areas.	No	
Are there additional concerns that require CSAC discussion? If so, briefly explain.	No	

Appendix B: NQF Member Expression of Support Results

Two NQF members provided their expressions of support for one measure. Results for the measure are provided below.

3357 Facility-Level 7-Day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers (Centers for Medicare & Medicaid Services/Yale CORE)

Member Council	Support	Do Not Support	Total
Health Professional	0	1	1
Provider Organization	0	1	1

Appendix C: Details of Measure Evaluation

Rating Scale: H=High; **M**=Moderate; **L**=Low; **I**=Insufficient; **NA**=Not Applicable; **Y**=Yes; **N**=No

Measures Recommended

3357 Facility Level 7-day Hospital Visits after General Surgery Procedures Performed at Ambulatory Surgical Centers

Submission

Description: Facility-level risk-standardized rate of acute, unplanned hospital visits within 7 days of a general surgery procedure performed at an ambulatory surgical center (ASC) among Medicare Fee-For-Service (FFS) patients aged 65 years and older. An unplanned hospital visit is defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission

Numerator Statement: The outcome being measured is acute, unplanned hospital visits (ED visit, observation stay, or unplanned inpatient admission) occurring within 7 days of a general surgery procedure performed at an ASC.

Denominator Statement: Target Population

Included patients:

The target population for this measure is Medicare FFS patients aged 65 years and older, who are undergoing outpatient general surgery procedures in ASCs that are within the scope of general surgery training. Specifically, the cohort of procedures includes the following types of surgeries: abdominal, alimentary tract, breast, skin/soft tissue, wound, and varicose vein.

The Medicare FFS population was chosen because of the availability of a national dataset (Medicare claims) that could be used to develop, test, and publicly report the measure. We limit the measure to patients who have been enrolled in Medicare FFS Parts A and B for the 12 months prior to the date of surgery to ensure that we have adequate data for identifying comorbidities for risk adjustment.

Included procedures:

The target group of procedures is surgical procedures that (1) are routinely performed at ASCs, (2) involve risk of post-surgery hospital visits, and (3) are within the scope of general surgery training. The scope of general surgery overlaps with that of other specialties (for example, vascular surgery and, plastic surgery). For this measure, we targeted surgeries that general surgeons are trained to perform with the understanding that other subspecialists may also be performing many of these surgeries at ASCs. Since the type of surgeon performing a particular procedure may vary across ASCs in ways that affect quality, the measure is neutral to surgeons' specialty training.

To identify eligible ASC general surgery procedures, we first identified a list of procedures from Medicare's 2014 and 2015 ASC lists of covered procedures, which include procedures for which ASCs can be reimbursed under the ASC payment system. This lists of surgeries is publicly available at: https://www.cms.gov/medicare/medicare-fee-for-service-payment/ascpayment/11_addenda_updates.html (download January 2014 and January 2015

ASC Approved HCPCS Code and Payment Rates, Addendum AA). Surgeries on the ASC list of covered procedures do not involve or require: major or prolonged invasion of body cavities, extensive blood loss, major blood vessels, or care that is either emergent or life-threatening. The ASC list is annually reviewed and updated by Medicare, and includes a transparent public comment submission and review process for addition and/or removal of procedure codes. Using an existing, defined list of surgeries, rather than defining surgeries de novo, is useful for long-term measure maintenance. Procedures listed in Medicare's list of covered ASC procedures are defined using Healthcare Common Procedure Coding System (HCPCS) and Common Procedural Terminology (CPT®) codes.

Ambulatory procedures include a heterogeneous mix of non-surgical procedures, minor surgeries, and more substantive surgeries. The measure is not intended to include very low-risk (minor) surgeries or non-surgical procedures, which typically have a high volume and a very low outcome rate. Therefore, to focus the measure only on the subset of surgeries on Medicare's list of covered ASC procedures that impose a meaningful risk of post-procedure hospital visits, the measure includes only "major" and "minor" procedures, as indicated by the Medicare Physician Fee Schedule global surgery indicator (GSI) values of 090 and 010, respectively. The GSI code reflects the number of post-operative days that are included in a given procedure's global surgical payment and identifies surgical procedures of greater complexity and follow-up care. This list of GSI values is publicly available for calendar year (CY) 2014 at: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Federal-Regulation-Notices-Items/CMS-1600-FC.html and for CY 2015 at: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Federal-Regulation-Notices-Items/CMS-1612-FC.html (download PFS Addenda, Addendum B).

Finally, to identify the subset of general surgery ASC procedures, we reviewed with consultants and Technical Expert Panel (TEP) members the Clinical Classifications Software (CCS) categories of procedures developed by the Agency for Healthcare Research and Quality (AHRQ). We identified and included CCS categories within the scope of general surgery, and only included individual procedures within the CCS categories at the procedure (CPT® code) level if they were within the scope of general surgery practice. We did not include in the measure gastrointestinal endoscopy, endocrine, or vascular procedures, other than varicose vein procedures, because reasons for hospital visits are typically related to patients' underlying comorbidities.

See the attached Data Dictionary, sheet S.9 "Codes Used to Define Cohort" for a complete list of all CPT procedure codes included in the measure cohort.

Exclusions: The measure excludes surgeries for patients without 7 or more days of continuous enrollment in Medicare FFS Parts A and B after the surgery. The measure excludes these patients to ensure all patients have full data available for outcome assessment.

Adjustment/Stratification: Not Applicable

Level of Analysis: Facility

Setting of Care: Outpatient services

Type of Measure: Outcome

Data Source: Claims, Enrollment Data

Measure Steward: Centers for Medicare & Medicaid Services (CMS)

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Y-17**; **N-0**; 1b. Performance Gap: **H-9**; **M-8**; **L-0**; **I-0;**

Rationale:

- The evidence base for the measure includes several studies regarding the factors that can predict unplanned readmissions, complications, and mortality following surgery. The Committee agreed that the evidence base supports that patient selection and preparation, post-operative care, and post-discharge planning can affect the rate of adverse events and unplanned admissions following outpatient surgery.
- Data submitted by the developer suggest variation in patient outcomes associated with ambulatory surgical center surgeries, with performance scores ranging from 0.42 to 2.13. Data also showed that among African Americans (3.1 percent), and patients with low socioeconomic status (2.2 percent) or dual eligible patients (3.7 percent), that dual eligible patients had higher readmission rates. A Committee member questioned whether this measure presents a disadvantage to ambulatory surgical centers that serve a large number of dual eligible patients.
- The developer noted that ambulatory surgical centers that serve a greater number of dual eligible patients are not doing any worse that centers with lower number of dual eligible patients. Another Committee member noted that ambulatory surgical centers' perception of risk in taking on dual eligible patients could drive discriminatory behavior.
- Ultimately, Committee members agreed that the measure met both the evidence and performance gap subcriteria.

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-3**; **M-15**; **L-0**; **I-0** 2b. Validity: **M-16**; **L-1**; **I-0**

Rationale:

- This measure calculates the ratio of the predicted to the expected number of post-surgical unplanned hospital visits among ambulatory surgical center (ASC) patients.
- The developer tested the reliability of the measure by calculating the intra-class correlation coefficient (ICC) of the measure score with the Medicare fee for service (FFS) calendar year 2012-2015 dataset. Reliability testing yielded an ICC of 0.530.
- The Committee agreed that the ICC score indicates moderate measure score reliability. (Results at the unit level of the ASC were higher.)
- A Committee member asked whether adding in isolated ICD-9 and ICD-10 codes, such as obesity or neurocognitive problems, would be helpful in improving the reliability of the measure.
- The developer explained they used a grouper approach for grouping ICD-9 and ICD-10 codes, which pulls in diagnoses for diseases. The developer stated that for some variables for obesity, instead of using the grouper, the most extreme cases of obesity were identified using codes. An expert panel reviewed a list of candidate variables and agreed on the conceptual risk variables they would like to test. The developer noted that obesity and neurodegenerative codes were initially included in the measure but

were not statistically significant. The developer explained that the expert panel was able to iterate on the inclusion of both variables. Ultimately, obesity codes were put back in the measure but neurodegenerative codes were not.
- The developer noted they would continue to collect information regarding clinical risk factors that could be included in the model.
- Ultimately, the Committee agreed the measure met the reliability and validity criterion.

---

3. Feasibility: H-14; M-2; 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 Committee noted that the measure is claim based and there are no registry fees associated with this measure. The Committee agreed the measure is feasible to collect.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-17; No Pass-0** 4b. Usability: **H-2; M-3; L-0; I-11**

Rationale:

- The Committee stated that although the measure is not currently in use, the developer provides a path for the use of this measure in the Medicare Ambulatory Surgical Center Quality (ASCQ) reporting program. The Committee also noted that the measure was supported for pre-rulemaking by the Measure Applications Partnership (MAP) in 2017.
- The Committee noted that the developer's technical advisory panel asked that additional procedure specific information be included in the measure and continued inclusion of opioid use as a risk in the model. One Committee member suggested that the developer clarify the variable for opioid use since the ICD-9 and ICD-10 codes definitions specify opioid abuse. The Committee member noted the importance of accurately assessing the causes of readmissions. The developer agreed to clarify the variable for opioid dependency or abuse.
- The Committee agreed that this measure could be an important feedback mechanism for ASCs to identify root causes for readmissions and ultimately improve outcomes.
- The Committee questioned whether the Committee should be looking for demonstrated usability or potential usability of the measure. NQF clarified that the measure was rated insufficient on usability since it was a new measure and no information was provided on the planned use of this measure. The developer will need to provide usability information upon maintenance of endorsement if the measure is recommended for endorsement. There is a stronger emphasis on use and usability for maintenance measures.
- Ultimately, the Committee agreed that the measure met this criterion.

---

5. Related and Competing Measures

- This measure is related to #2687 *Hospital Visits after Hospital Outpatient Surgery* and #2539 *Facility 7-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy*.
- The developer clarified that #2687 and #2539 include exclusions for same day procedure claims (i.e., hospital or emergency department visits that occur the same day) but this measure (#3357) does not. In the rationale for #2687 and #2539, the developer notes that the exclusion is included since with claims it is difficult to tell which came first, the hospital visit after surgery or a patient's visit to the outpatient department.
- The developer noted that the other measures exclude emergency or hospital visits on the same day because in the outpatient setting, the patient can go to the emergency department at the same hospital and then have their procedure or vice versa. The developer stated that this was not an issue in the ASC setting because it is possible to use claims to determine if a patient underwent surgery at an ASC then went to the emergency department at another facility; therefore, this exclusion was not included.
- The Committee raised no concerns regarding the exclusions for this measure.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-0**

6. Public and Member Comment

NQF received three public and member comments concerning the measure's specifications, measure validity, the lack of sociodemographic risk adjustment, and concerns about whether meaningful differences in performance are actionable enough to improve the quality of surgical care. Commenters questioned the inclusion of skin procedures in the measure specifications because general surgeons do not routinely perform these procedures. Commenters also questioned whether skin procedures were included to boost low case volume. Two comments noted that the measure should be risk adjusted for social risk factors, such as socioeconomic status (SES), for providers who serve low SES clients. One commenter pointed out that unplanned visits are not always in the surgeon's control and could be related to issues of access to care. Submitted comments also noted that the measure is "topped out" and does not provide enough variation in performance to discern quality of care.

The developer provided the following response:

> **Measure Outcome** - The commenter stated that the "readmission measure is not a good proxy for driving improvement in surgical care." We would like to clarify that the measure under review by the NQF (NQF ID 3357) is not a readmission measure. The measure's outcome is any unplanned hospital visit, defined as an emergency department (ED) visit, observation stay, or unplanned inpatient admission, occurring within 7 days of a general surgery procedure performed at an ambulatory surgical center (ASC). The outcome of hospital visits is the focus of this measure because it is a broad, patient-centered outcome that captures the full range of hospital visits resulting from adverse events or poor care coordination following an ASC general surgery procedure. This measure's goal is to assess and illuminate variation in risk-adjusted hospital visits following surgery, for quality improvement and public reporting.

Alignment with Registry-Based Measures - The commenter "believes that surgical measurement should be built on four key principles: 1) setting the standards, 2) building the right infrastructure, 3) using the right data, and 4) verifying with outside experts," and advocates that ASC measures "include standards-based facility-level verification programs, patient reported experience (PRE) and outcome (PRO) measures, and traditional quality measures including registry and claims-based measures."

This 7-day hospital visit measure fits within the above framework. It increases surgeon and ASC accountability for and awareness of patient outcomes during the post-surgical period, and provides facilities with patient-level data on outcomes to inform quality improvement that ASCs currently lack. The availability of linkable Medicare claims data to calculate the risk-adjusted measure and link patient outcomes, procedure, and risk factor data across settings makes this a low-burden measure that can provide valuable information and complement registry-based, surgery-society developed measures such as PROs.

We would like to further clarify that our measure development process is aligned with the four key principles that the commenter identified as foundational for surgical quality measurement. We built the measure using standards for quality measure development set forth by CMS in its Measure Management System Blueprint. The measure was developed with input from general surgery consultants, a national Technical Expert Panel, and the public.

Usability and Actionable Information - Two commenters expressed concern that the measure would be of limited utility, and that the information it provides is not actionable.

For ASCs, we believe measuring and publicly reporting claims-based, risk-adjusted measure scores will encourage ASCs to engage in quality improvement and lead to better patient care over time. Further, CMS plans to implement the measure to optimize its usability. Prior to public reporting, CMS anticipates providing claims-detail and facility-specific reports, as the Agency does for other outcome measures. These reports will allow ASCs to see patient 7-day outcomes that are currently not visible to them. This information will help ASCs understand their performance, inform quality improvement efforts, and improve the care they provide to patients.

Variation in Performance Scores and Identification of Outliers - All three commenters expressed their view that there is not enough variation in the measure results to show meaningful differences between facilities. One commenter expressed concern that the measure identifies relatively few outliers as better or worse than expected.

We appreciate the commenters' concerns. The measure score results, however, do present a clinically meaningful range in risk-adjusted outcome rates. As presented in the public comment technical report using Medicare FFS CY 2015 data, we found that the facility measures scores ranged from 0.94% to 4.55%, with a median risk-standardized hospital visit rate of 2.19% (the 25th and 75th percentiles were 2.03% and 2.46%,

respectively). The variation in these rates provides a quality signal, and reporting facility-level measure scores will improve transparency and promote quality improvement.

To assist consumers with interpreting the measure, we provide a descriptive category of facility quality (better than, worse than, or no different than the national rate); to provide ASCs and other users with richer insight into performance, we provide the estimated 7-day hospital visit rate and the 95% interval estimate (uncertainty estimate) around that rate. As the commenters pointed out, the descriptive approach categorizes relatively few facilities as outliers. The approach to categorizing facility outliers is very conservative by design. It uses 95% confidence interval (uncertainty) estimates to identify outliers.

Measuring quality of care associated with general surgery procedures performed at ASCs would bring awareness to ASCs and provide valuable data to patients. As intended, we expect this measure will promote patient-centered improvement in care provided at ASCs, because measurement coupled with transparency will make visible, for the first time, the rate of hospital visits after general surgery ASC procedures to both patients and ASCs.

Emergency Department (ED) Visits and Observation Stays - One commenter was concerned that the planned admissions algorithm was not applied to ED visits or hospital observation stays, and that counting all ED visits as unplanned did not account for lack of "patient access to care in the desirable setting." As the commenter noted, we have previously stated that while we understand that the ED and hospital observation setting may be used for planned care at times, the measure is structured to count these events because these settings are not usually a desirable setting for planned care from the patient's point of view.

Also, ASCs are expected to limit their cases to those that can be safely performed outside the hospital setting. Higher rates of serious but potentially preventable complications that result in ED visits and observation stays may be a sign of poorer quality ASC care. ASCs can reduce the likelihood of these serious complications by emphasizing patient safety and reducing rates of complications or harm events, and by discharge planning that anticipates the need for potential office-based follow up care.

Measure Cohort - One commenter noted that this measure "is meant to capture all other routinely performed outpatient surgical procedures," unlike the other measures in CMS's ASCQR program that focus on colonoscopy, orthopedic, or urology procedures. For this measure, we clarify that we targeted procedures that fall within the scope of general surgery, including those performed by other surgical specialists. We combined these procedures because they share the risk of post-surgery hospital visits within 7 days, they share common reasons for return to the hospital, and the risk of hospital visits following these procedures can be mitigated through similar strategies.

Additionally, three commenters expressed concerns that the measure includes many skin and plastic surgery procedures. We understand the commenters' concerns that over half the procedures in the cohort are skin procedures and concern that these

procedures were included to increase sample size. The commenters also stated that many of the included procedures are performed by surgeons other than general surgeons.

As noted above, we included procedures within the scope of practice of general surgery even though these procedures are often performed by other subspecialists because they share common features that allow us to combine them for assessing quality. However, the commenters are correct that in addition to general surgeons, other types of surgeons and non-surgical specialists perform skin and other procedures included in the cohort. As we clarified above, we included these procedures in a single measure because the procedures share (1) a risk of post-surgery hospital visits within 7 days and (2) relatively similar reasons for return to the hospital. Members of our TEP also felt the care practices that would best lower the risk of hospital visits were similar across these procedures. Procedural volume was not a criterion for inclusion of procedures in the cohort.

Further, we identified and refined the group of procedures to include in the cohort through multi-stakeholder review and input from a national TEP, general surgery consultants, and the public. For example, in light of comments receive on the measure during measure development public comment, we re-reviewed all of the individual CPT codes within CCS categories and removed 15 individual procedures (CPT® codes) from the measure that were outside the scope of general surgery practice.

Adjustment for Social Risk Factors -Two commenters expressed concern that the measure is not adjusted for social risk factors.

As previously acknowledged in the conceptual model we presented in the NQF application, we agree that patients' socioeconomic status (SES) affects health and health outcomes in important ways.

Our intent when developing and testing the measure was to be responsive to the NQF guidelines for measure developers on SES. We therefore examined three patient-level indicators of social risk that are reliably available for all Medicare beneficiaries: 1) Medicare-Medicaid dual eligibility, 2) race, and 3) the AHRQ SES Index. The variables used are aligned with those the National Academy of Medicine committee identified as available for use in outcome measures. We examined whether these factors were associated with increased risk in hospital visits after adjusting for other risk factors and evaluated the impact of social risk factors on ASC-level measure scores.

While including each of these risk factors in our models indicated a statistically significant association after controlling for other risk-adjusters, results showed that the effect of social risk factors on hospital visit rates in the fully adjusted model was significant but small. Additionally, inclusion of these variables did not change ASCs' risk-standardized hospital visit ratios (RSHVRs) or their performance on the measures. Correlation coefficients between RSHVRs with and without adjustment for these factors were near 1 (0.998, 1.000, and 0.999 for dual-eligible, African-American, and low SES

patients, respectively) and mean differences in RSHVRs were near zero (0.0000, -0.0001, and -0.0002 for dual-eligible, African-American, and low SES patients, respectively).

In addition, we examined the relationship between the proportion of low SES patients and the facility-level score, focusing on facilities with the highest proportion of dual eligible patients (fourth quartile). This analysis did not show a clear relationship between the proportion of low SES patients and the facility-level score (Pearson correlation coefficient = -0.17).

Based on these findings, and a consideration of how social risk factors affect patients in the ambulatory setting and the importance of efforts to address all patients' needs, CMS decided to not adjust the models for these social risk factors; not adjusting is not likely to lead to unintended consequences, or burden providers that serve low SES patients, and adjusting may mask quality differences. However, once the measure is implemented, CMS will monitor this measure, as with others, for unintended consequences related to disparities.

Finally, we acknowledge the importance of optimizing measures to incentivize high quality care for all while ensuring providers caring for low SES patients are not disadvantaged on the measures. CORE, with CMS, is exploring alternative modeling approaches that better illuminate how ASCs and their patients contribute to SES-related risks, and will continue to explore incorporating social risk factors into quality measures.

Attribution of Outcomes - We appreciate the commenter's continued review of the top reasons for any hospital visit within 7 days of general surgery procedures.

As previously clarified, the diagnoses referred to by the commenter (Table 4 of measure documentation) can occur during an admission, ED visit, or observation stay. If they occur during an admission, then all but one type (acquired absence of breast and nipple) are identified as planned admissions and are not counted in the measure outcome. If these diagnoses (including cancer diagnoses) occur as part of an ED visit or observation stay, they are included in the measure outcome because ED visits and observation stays are not routinely used for planned care. We understand that the ED and hospital observation setting may be used for planned care at times, but the measure is structured to count these events because these settings are not usually a desirable setting for planned care from the patient's point of view.

CORE is committed to evaluating whether refining the CMS planned admission algorithm will better capture planned admissions for the diagnoses flagged by comments. As such, during measure reevaluation, CORE will consider updating the planned admission algorithm to include acquired absence of breast and nipple so that admissions with this diagnosis would not be counted in the measure outcome.

Measure Title - We appreciate the commenter's suggestion to rename the measure. We have already renamed the measure to address this concern. The measure was previously named "Hospital Visits After General Surgery Ambulatory Surgical Center Procedures," and we revised its title to be "Facility-Level 7-Day Hospital Visits after

General Surgery Procedures Performed at Ambulatory Surgical Centers." Our intent was to emphasize the scope of the procedures included in the measure cohort rather than the types of specialists performing them. The scope of the measure was defined by the scope of practice of general surgeons. Thus, we have chosen not to include "skin procedures" or "plastic repair" in the title given that many types of procedures in the measure are performed by both general surgeons and other specialists, and including one specific procedure type in the title would make the scope of the measure less clear. We believe the measure title accurately reflects what it assesses. CMS welcomes continued suggestions on the best name for the measure.

Low ASC Case Volumes - The commenter expressed concern that this measure would provide insufficient information for low-volume facilities. We understand the commenter's concern that the measure would not provide sufficient information about the quality of care in individual low-volume facilities. For this measure, as is done for other risk-adjusted outcome measures, CMS will set minimum-volume requirements for reporting and only report scores for ASCs that have an adequate number of cases to generate reliable estimates. For example, for publicly reported outcome measures such as CMS's hospital readmission measures, CMS implemented minimum volume requirements for reporting.

The commenters also expressed concern about CMS implementing this measure using an inadequate amount of data. The commenter referenced CMS's use of only 1 year of claims data for implementation of another risk-adjusted outcome measure (ASC-12: Facility Seven-Day Risk-Standardized Hospital Visit Rate after Outpatient Colonoscopy) in the ASCQR program, instead of 3 years of claims data. For this general surgery measure, CMS will consider using multiple years of claims data for public reporting. We calculated measure score (test, retest) reliability for a 2-year reporting period and found that the agreement (intraclass coefficient) between the two RSHVR values for each ASC was 0.526, indicating moderate measure score reliability. NQF committees consider their evaluation criteria to be rigorous, which state that moderate or high reliability is typically required for endorsement.

Prior to measure implementation, CMS will evaluate the amount of data required for reliable measure score calculation, and determine the number of years of data to use, weighing the tradeoffs between having an adequate number of cases for the greatest number of facilities and ensuring data used are timely.

Committee members were satisfied with the developer's response to the public comments, and agreed that the inclusion of skin procedures is not detrimental to the measure. The Committee stated that skin procedures should be included because they are an increasing part of care provided at ASCs. They also noted that the measure fills a gap in care and provides valuable performance data to ASCs that they would not otherwise receive. The Committee also noted that the ability of the risk model to discriminate differences in patient characteristics could be improved (c-statistic =0.69) and that they would like to see a more robust c-statistic and additional data on the

upward mobility of patients when the measure is resubmitted for endorsement consideration.

---

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

1790 Risk-Adjusted Morbidity and Mortality for Lung Resection for Lung Cancer

Submission

Description: Percentage of patients greater than or equal to 18 years of age undergoing elective lung resection (Open or VATS wedge resection, segmentectomy, lobectomy, bilobectomy, sleeve lobectomy, pneumonectomy) for lung cancer who developed any of the following postoperative complications: reintubation, need for tracheostomy, initial ventilator support > 48 hours, ARDS, pneumonia, pulmonary embolus, bronchopleural fistula, unexpected return to the operating room, myocardial infarction or operative mortality (death during the index hospitalization, regardless of timing, or within 30 days, regardless of location).

Numerator Statement: Number of patients greater than or equal to 18 years of age undergoing elective lung resection (Open or VATS wedge resection, segmentectomy, lobectomy, bilobectomy, sleeve lobectomy, pneumonectomy) for lung cancer who developed any of the following postoperative complications: reintubation, need for tracheostomy, initial ventilator support > 48 hours, ARDS, pneumonia, pulmonary embolus, bronchopleural fistula, unexpected return to the operating room, myocardial infarction or operative mortality (death during the index hospitalization, regardless of timing, or within 30 days, regardless of location).

Denominator Statement: Number of patients greater than or equal to 18 years of age undergoing elective lung resection (Open or VATS wedge resection, segmentectomy, lobectomy, bilobectomy, sleeve lobectomy, pneumonectomy) for lung cancer

Exclusions: Patients were excluded if they had an extrapleural pneumonectomy, completion pneumonectomy, carinal pneumonectomy, occult carcinoma or benign disease on final pathology, or an urgent, emergent, or palliative operation. Furthermore, patients with missing age, sex, discharge mortality status, and predicted forced expiratory volume in 1 second were also excluded.

Adjustment/Stratification: Statistical risk model

Level of Analysis: Facility

Setting of Care: Inpatient/Hospital

Type of Measure: Outcome

Data Source: Other, Registry Data

Measure Steward: The Society of Thoracic Surgeons

STANDING COMMITTEE MEETING [02/01/2018]

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Y-16**; **N-0**; 1b. Performance Gap: **H-5**; **M-11**; **L-0**; **I-0;**

Rationale:

- Updated evidence for this maintenance measure is based on a study showing that operative mortality and complication rates are low for lung cancer resection among surgeons participating in the Society of Thoracic (STS) General Thoracic Surgery Database (GTSD).
- Performance data showed variation in performance from 0.47 percent to 2.37 percent, among 217,844 patient records at 213 sites.
- The Committee agreed that there is still a gap in performance among the different practices.
- The Committee raised no concerns with the evidence and performance gap for this maintenance measure.

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-7**; **M-9**; **L-0**; **I-0** 2b. Validity: **M-8**; **L-8**; **I-0**

Rationale:

- Using Bayesian inference methods on data from 231 hospitals and 27,844 patient records, reliability was measured at 0.50 for 231 hospitals, rising to 0.53 for 216 hospitals performing at least 10 procedures, and to 0.84 in hospitals performing at least 200 procedures. Validity was assessed using percent agreement of data elements; the agreement rate was 96.78 percent for overall data accuracy, ranging from 94.3 to 99.0 percent.
- The Committee noted there is a change to the measure specifications – from monitoring "bleeding requiring reoperation" to "an unexpected return to the operating room"; the Committee agreed that this was an appropriate change. The developer clarified that this change included more events or indications for return to the operating room rather than just bleeding.
- In their discussion on validity, the Committee noted that this measure was similar to #3294 *STS Lobectomy for Lung Cancer Composite Score* in that validity was high and the percent of missing data had been decreasing but did not lead to a bias in calculation of the measure. The developer noted that they had addressed missing data by requiring programs to have a 95.0 percent completion rate for the outcome field "operative mortality" in 2016, and a completion rate of 98.0 percent by 2017.

3. Feasibility: H-6; M-9; 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 developer reports that data elements are generated and used by healthcare personnel during the provision of care. STS GTSD participants pay an annual participant fee of $550-$700 depending on whether the participant is an STS member.
- The Committee raised no concerns with the feasibility of the measure.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-16; No Pass-0** 4b. Usability: **H-7; M-9; L-0; I-0**

Rationale:

- The Committee questioned whether data were publically reported. The developer articulated plans public report results of the measure in 2019.
- The Committee questioned what percentage of total thoracic surgeries are captured in the STS database and how many thoracic surgeons participate in the registry. The developer noted that the database captures data on surgeries from a large number of primary thoracic surgeons; however, this information is not widely available for adult cardiac surgeons or general surgeons who may perform a small number of thoracic surgeries. The Committee agreed these surgeons could benefit from participating in the database. Another Committee member noted that the STS GTSD is now open to all surgeons and non-STS members, demonstrating opportunities to share data in the database with other surgeons.
- A Committee member shared data from a presentation Penetration, Completeness, and Representativeness of the STS GTSD for Lobectomy at the annual STS meeting. The Committee member reported that although STS lobectomy procedures are low volume, representing 25.0 percent of lobectomy procedures in the United States when compared to lobectomy procedures in the Centers for Medicare & Medicaid Services (CMS) database, STS participants outperformed CMS participants. Participants in the STS database had shorter length of stay and half the mortality rate compared to the participants reporting to the CMS database; both outcomes were reported to be highly statistically significant.

---

5. Related and Competing Measures

- Measure #1790 is related to #3294 *STS Lobectomy for Lung Cancer Composite Score*. Measure #1790 includes a broader range of lung resection procedures than the Lobectomy Composite, and therefore includes a larger number of cases and potentially provides performance data to more general thoracic surgeons.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-0**

6. Public and Member Comment

No public or member comments submitted.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals

3294 STS Lobectomy for Lung Cancer Composite Score

Submission

Description: The STS Lobectomy Composite Score comprises two domains:

1. Operative Mortality (death during the same hospitalization as surgery or within 30 days of the procedure)

2. Presence of at least one of these major complications: pneumonia, acute respiratory distress syndrome, bronchopleural fistula, pulmonary embolus, initial ventilator support greater than 48 hours, reintubation/respiratory failure, tracheostomy, myocardial infarction, or unexpected return to the operating room.

The composite score is created by a weighted combination of the above two domains resulting in a single composite score. In addition to receiving a numeric score, participants are assigned to rating categories designated by the following:

1 star: lower-than expected performance

2 stars: as-expected-performance

3 start: higher-than-expected-performance

Numerator Statement: The STS Lobectomy Composite Score comprises two domains:

1. Operative Mortality (death during the same hospitalization as surgery or within 30 days of the procedure)

2. Presence of at least one of these major complications: pneumonia, acute respiratory distress syndrome, bronchopleural fistula, pulmonary embolus, initial ventilator support greater than 48 hours, reintubation/respiratory failure, tracheostomy, myocardial infarction, or unexpected return to the operating room.

The composite score is created by a weighted combination of the above two domains resulting in a single composite score. Operative mortality and major complications were weighted inversely by their respective standard deviations across participants. This procedure is equivalent to first rescaling mortality and complications by their respective standard deviations and then assigning equal weighting to the rescaled mortality rate and rescaled complication rate. This is the same methodology used for other STS composite measures.

In addition to receiving a numeric score, participants are assigned to rating categories designated by the following:

1 star: lower-than expected performance

2 stars: as-expected-performance

3 start: higher-than-expected-performance

Patient Population: The STS GTSD was queried for all patients treated with lobectomy for lung cancer between January 1, 2014, and December 31, 2016. We excluded patients with non-elective status, occult or stage 0 tumors, American Society of Anesthesiologists class VI, and with missing data for age, sex, or discharge mortality status.

Time Window: 01/01/2014 - 12/31/2016

Model variables: Variables in the model: age, sex, year of operation, body mass index, hypertension, steroid therapy, congestive heart failure, coronary artery disease, peripheral vascular disease, reoperation, preoperative chemotherapy within 6 months, cerebrovascular disease, diabetes mellitus, renal failure, dialysis, past smoker, current smoker, forced expiratory volume in 1 second percent of predicted, Zubrod score (linear plus quadratic), American Society of Anesthesiologists class (linear plus quadratic), and pathologic stage.

Denominator Statement: Number of patients greater than or equal to 18 years of age undergoing elective lobectomy for lung cancer

Exclusions: Patients were excluded with non-elective status, occult or stage 0 tumors, American Society of Anesthesiologists class VI, and with missing data for age, sex, or discharge mortality status.

Adjustment/Stratification: Statistical risk model

Level of Analysis: Facility

Setting of Care: Inpatient/Hospital

Type of Measure: Composite

Data Source: Other, Registry Data

Measure Steward: The Society of Thoracic Surgeons

STANDING COMMITTEE MEETING [02/06/2018]

1. Importance to Measure and Report: The measure meets the Importance criteria

(1a. Evidence, 1b. Performance Gap, 1c. Composite – Quality Construct and Rationale)

1a. Evidence: **Y-16**; **N-0**; 1b. Performance Gap: **H-3**; **M-12**; **L-0**; **I-0;** 1c. Composite – **H-9**; **M-7**; **L-0**; **I-0**

Rationale:

- The developer reported that data in the Society of Thoracic Surgeons (STS) General Thoracic Surgery Database (GTSD) show a reduction in perioperative morbidity and equivalent long-term survival when minimally invasive approaches for lobectomy are used. Specifically, STS data have shown that minimally invasive lung cancer resection has a 50.0 percent reduction in major complications compared with a thoracotomy approach, adjusted for age, sex, and comorbidities.
- The Committee stated that the evidence presented supports the measure.
- The Committee agreed there is a gap based on the performance data presented by the developer. Data collected during two separate timeframes during 2013-2016 indicated a performance rate of 95.0 percent to 98.0 percent, for approximately 200-300 participants and over 24,000 operations.
- In terms of quality construct, this measure is based on a combination of an operative mortality outcome and the risk-adjusted occurrence of any of nine major complications.

Participants are scored for each domain (mortality and complication), and an overall composite score which is created by a weighted combination of the two domains. Participants are also assigned a performance rating designated by one to three stars. The developer reported that since mortality rates for thoracic surgery have declined, it can be difficult to differentiate performance based on mortality alone because it does not take into account that not all operative survivors received equal quality care. The Committee agreed that a composite score from a weighted combination of mortality and operative complications provides a more comprehensive measure of overall surgical quality.

- A Committee member noted that operative mortality is weighted approximately four times that of a major complication in the composite, consistent with STS adult cardiac surgery quality measures. Committee members also noted that this is an improvement from its previous lung cancer resection model in which mortality and major morbidity were weighted equally.
- Overall, the Committee agreed that the quality construct and rationale for the composite are explicitly stated and logical; and the weighting and approach to the measure construction is described clearly and has been vetted by an expert panel.

---

2. Scientific Acceptability of Measure Properties: <u>The measure meets the Scientific Acceptability criteria</u>

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

2a. Reliability: **H-8**; **M-8**; **L-0**; **I-0** 2b. Validity: **M-16**; **L-0**; **I-0** 2c. Composite Construction: **H-9**; **M-6**; **L-1**; **I-0**

<u>Rationale</u>:

- The Committee noted that the measure is well and clearly specified, and that the measure can consistently be implemented. Committee members also noted that reliability of data elements was supported by external audit of the GTSD, demonstrating high agreement rates and validation of data accuracy. In addition, committee members noted that the NQF Scientific Methods Panel was satisfied with the reliability testing for the measure.
- Committee members noted that validity of the performance score was not tested. Confidence interval testing was performed and only percent agreement was assessed in the analysis. NQF staff clarified that while score-level validity testing is desired, data element testing is acceptable because this is a new measure. For future maintenance evaluations, score-level testing will be required. Overall, the Committee did not have any major concerns regarding the validity of the measure and noted that the NQF Scientific Methods Panel was satisfied with the validity analyses for the measure.
- The Committee did not express additional concerns with the construct of the composite measure and agreed the information provided was sufficient to satisfy the criterion for composite construct.

---

3. Feasibility: H-5; M-11; 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:

- Feasibility was addressed in terms of its similarity across STS measures; i.e., data for the measure is captured in a standardized way through the STS database of which most surgeons and programs in the United States are members. STS GTSD participants pay an annual participant fee of $550-$700 depending on whether the participant is an STS member.

---

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Pass-15; No Pass-1** 4b. Usability: **H-9; M-7; L-0; I-0**

Rationale:

- According to the developer, this measure is publicly reported through the STS Public Reporting Task force. The task force develops public report cards that are consumer centric. In terms of accountability, results are shared with participants in the STS GTSD for quality improvement purposes.
- Committee members noted that while performance variation is not wide, performance results for this measure are still considered useful for both accountability and performance improvement activities.

---

5. Related and Competing Measures

- This measure is related to 1790 *Risk-Adjusted Morbidity and Mortality for Lung Resection for Lung Cancer*. Measure #1790 includes a broader range of lung resection procedures than the Lobectomy Composite, and therefore includes a larger number of cases and potentially provides performance data for more general thoracic surgeons.

Standing Committee Recommendation for Endorsement: **Y-16**; **N-0**

6. Public and Member Comment

No public or member comments submitted.

7. Consensus Standards Approval Committee (CSAC) Vote: Y-X; N-X

8. Appeals