NATIONAL QUALITY FORUM
Driving measurable health improvements together

Meeting Summary

Patient Safety Standing Committee – Measure Evaluation Web Meeting

The National Quality Forum (NQF) convened the Patient Safety Standing Committee for a web meeting on June 24 and 25, 2021, to evaluate six measures.

Welcome, Introductions, and Review of Meeting Objectives

NQF welcomed the Standing Committee and participants to the web meeting. NQF staff reviewed the meeting objectives. The Standing Committee members each introduced themselves and disclosed any conflicts of interest. Twenty Standing Committee members were present at the start of the meeting. Five total Standing Committee members identified conflicts with three measures: three Standing Committee members for NQF #0500 and one each for NQF #3501e and NQF #3612, which led to their recusal from those respective measures. Two Standing Committee members were recused from NQF #0500 due to their involvement with the measure developer as a consultant on the measure. One Standing Committee member was recused from NQF #0500 because he was employed by the same organization that developed the measure. One Standing Committee member was recused from NQF #3501e due to his direct collaboration with the measure developer on the measure. One Standing Committee member was recused from NQF #3612 because he was employed by the same organization that developed the measure.

Some Standing Committee members were unable to attend the entire meeting. The vote totals reflect members present and eligible to vote. Quorum was met and maintained for the entirety of all the meetings.

Topic Area Introduction and Overview of Evaluation Process

NQF staff provided an overview of the topic area and the current NQF portfolio of endorsed measures. There are currently 60 measures in the Patient Safety portfolio. Additionally, NQF reviewed the Consensus Development Process (CDP) and the measure evaluation criteria.

Measure Evaluation

During the meeting, the Patient Safety Standing Committee evaluated six measures, including four maintenance and two new measures for endorsement consideration. The summary of the Standing Committee's deliberations below will also be provided in the draft technical report. NQF will post the draft technical report on August 11, 2021, for public comment on the NQF website. The draft technical report will be posted for 30 calendar days.

Rating Scale: H – High; M – Medium; L – Low; I – Insufficient; NA – Not Applicable

#0500 Severe Sepsis and Septic Shock: Management Bundle (Henry Ford Hospital)

Measure Steward/Developer Representatives at the Meeting

Measure Stewards:

- Sean Townsend

- Emmanuel Rivers

CMS:

- Katrina Hoadley

Mathematica:

- Madeline Pearse
- Ana Talamas
- Robert Dickerson
- Raga Ayyagari
- Sharon Zhao

Standing Committee Votes

- **Evidence**: H-3; M-9; L-4; I-1 (Pass – 12/17)
- **Performance Gap**: H-6; M-9; L-2; I-0 (Pass – 15/17)
- **Composite - Quality Construct and Rationale**: H-6; M-9; L-2; I-0 (Pass – 15/17)
- **Reliability**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
  - The NQF Scientific Methods Panel's rating for Reliability: High, H-5; M-1; L-0; I-2 (Pass – 6/8)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
  - The NQF Scientific Methods Panel's rating for Validity: Moderate, H-3; M-2; L-1; I-2 (Pass – 5/8)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Composite Quality Construct**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
  - The NQF Scientific Methods Panel's rating for Composite Quality Construct: Moderate, H-2; M-3; L-0; I-1 (Pass – 5/6)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-3; M-13; L-1; I-0 (Pass – 16/17)
- **Use**: Pass-17; No Pass-0 (Pass – 17/17)
- **Usability**: H-10; M-5; L-2; I-0 (Pass – 15/17)

Standing Committee Recommendation for Endorsement: Yes-14; No-3 (Pass – 14/17)

The Standing Committee recommended this composite measure for continued endorsement.

The developer started the discussion by providing an overview and description of the measure (which the developer termed *SEP-1*). This measure was last reviewed in 2017 by the Infectious Disease Standing Committee (which is no longer in existence). Since that time, the developer noted that the evidence for the elements supporting the measure had grown and that there has been no evidence of unintended consequences. The developer also described being approached by the Infectious Diseases Society of America (IDSA) in 2019 with concerns about the start time (i.e., time zero) for the measure as well as concerns that the measure may be promoting antibiotic overuse. The developer stated that both of these issues are under study. There were also concerns that septic shock patients received 30ml/kg of fluid administration without consideration for the presence of congestive heart failure or chronic renal insufficiency, in which fluid administration may lead to harm. The developer described that starting in

July, providers can still meet the measure by documenting that the 30 ml/kg fluid bolus may be detrimental to the patient. However, this information was not included in the submission that the Standing Committee reviewed.

The lead discussant continued by describing the evidence supporting the measure, specifically some concerns about the quality of the evidence for specific elements of SEP-1, such as the administration of high boluses of fluids, as well as rechecking the lactate after the first dose becomes elevated. In addition, there were concerns about differentiating between septic shock and severe sepsis, which have different evidence, particularly with respect to the elements of SEP-1. A co-discussant commented that the evidence supporting the giving of intravenous fluid resuscitation was inconsistent, while the evidence supporting lactate administration was more consistent. A Standing Committee member then clarified the definition of evidence criteria, specifically the association between a process and outcome. A Standing Committee member then redescribed the specific elements of evidence in the preliminary assessment, noting different levels of evidence for specific elements of SEP-1. A Standing Committee member asked about the "weight" of evidence, comparing the risk and benefits of the measure. Another Standing Committee member clarified that certain elements of the measure have clear evidence, such as the use of early antibiotics in the presence of severe infection, while others had less evidence. The developer further clarified that studies in the submission demonstrated that improved adherence to the guideline was associated with improved outcomes; however, this had not yet been discussed. Another Standing Committee member stated that antibiotic stewardship is important, but they also noted that half of the people with COVID-19 also had a co-infection and that 7–12 percent of patients with viral infections have a co-infection. The Standing Committee member stated that these data suggested that liberal antibiotic use in the critically ill, even of viral etiologies, may be appropriate. In addition, this suggested that early de-escalation of antibiotics rather than the avoidance of early antibiotics may be a better strategy, which supports the measure.

Next, the lead discussant summarized the public comments as well as letters of both support and non-support that were received from specialty societies, including the American Medical Association (AMA) and others. The discussant stated that commenters in support of the measure noted its importance, as sepsis can lead to severe morbidity and possibly death. Commenters that did not support the measure were concerned with the lack of alignment with current evidence and the potential for negative unintended consequences, such as incentivizing antibiotic overuse. One Standing Committee member stated that differences in mortality were likely less than improvements in morbidity, which were potentially greater. The developer then described that there were no studies that had quantified harm related to the measure. However, there were studies showing a single-center study that demonstrated increased use of antibiotics in urinary tract infections. Another Standing Committee member described a patient who had died due to a delay in antibiotics. Therefore, early interventions are vital; while antibiotic stewardship is also important, antibiotics should not have been restricted in this situation. Based on this discussion, the Standing Committee passed the measure on evidence.

The lead discussant then described the performance gap for the measure. A Standing Committee member asked about disparities in the measure; in response, the developer clarified that disparities are present at the patient and facility levels. No other concerns were raised about gap or composite construct, and the measure passed both criteria.

Moving to scientific acceptability, the lead discussant stated that the Scientific Methods Panel (SMP) passed the measure on reliability, validity, and the composite construction. A Standing Committee member expressed concern that a claims-based identification of sepsis may not be reliable. The Standing Committee considered the reliability results, which they agreed showed sufficient reliability.

Therefore, the Standing Committee voted to accept the SMP's high rating for reliability. The Standing Committee also accepted the SMP's moderate rating for validity and the composite construction without further discussion.

The lead discussant described the feasibility of the measure that used a combination of electronic claims as well as a chart review. The Standing Committee did not comment on this matter and voted "moderate" for feasibility. There were also no concerns discussed about use or usability; therefore, the Standing Committee gave a passing rating for use and a high rating for usability. Following this review, the Standing Committee passed the measure on its overall suitability for endorsement. The lead discussant reiterated that the developer should ensure they make the previous fixes to the measure as described and address concerns from specialty societies and other groups on an ongoing basis. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing. One Standing Committee member commented that they would like to see an outcome measure at some point within the composite.

#3621 Composite Weighted Average for Computerized Tomography (CT) Exam Types: Overall Percent of CT Exams for Which Dose Length Product Is at or Below the Size-Specific Diagnostic Reference Level (for CT Abdomen-Pelvis With Contrast/Single Phase Scan, CT Chest Without Contrast/Single (American College of Radiology [ACR])

Measure Steward/Developer Representatives at the Meeting

- Karen Campos (presenter)
- Mahadevappa Mahesh (presenter)
- Judy Burleson
- Mythreyi Chatfield

Standing Committee Votes

- **Evidence**: H-0; M-15; L-3; I-1 (Pass – 15/19)
- **Performance Gap**: H-0; M-18; L-0; I-0 (Pass – 18/18)
- **Composite - Quality Construct and Rationale**: H-2; M-14; L-1; I-1 (Pass – 16/18)
- **Reliability**:
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Reliability: High, H-5; M-2; L-0; I-1 (Pass – 7/8)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**:
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's Consensus Not Reached (CNR) rating for Validity: H-0; M-4; L-2; I-2 (Consensus Not Reached – 4/8)
    - The Standing Committee voted on validity: H-0; M-12; L-3; I-2 (Pass – 12/17)
- Composite Quality Construct:
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Validity: Moderate, H-2; M-3; L-0; I-1 (Pass – 5/6)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-4; M-14; L-0; I-0 (Pass – 18/18)

- **Use**: Pass-18; No Pass-0 (Pass – 18/18)
- **Usability**: H-4; M-14; L-0; I-0 (Pass – 18/18)

Standing Committee Recommendation for Endorsement: Yes-16; No-2 (Pass – 16/18)

The Standing Committee recommended this composite measure for initial endorsement.

The developer described that the intent of this measure is to optimize the way that computerized tomography (CT) scans are performed, adjusting for diagnostic-specific radiation levels and by the dose length. The goal is to safely reduce radiation exposure by holding entities accountable for proper radiation dosing. The developer then described how the measure received a "consensus not reached" verdict for validity from the SMP and that face validity was conducted only because criterion validity testing could not be conducted due to data limitations. Following this description, the developer described increases in the number of CTs in the United States (U.S.) and gave further information as to why this measure is important for quality control in radiology.

The lead discussant began the evaluation by describing the evidence supporting the measure, specifically the presence of variation in the diagnostic reference levels that are used in CT scans. This is a particular concern for patients who receive multiple CT scans over time, as they are at higher risk for radiation exposure and potential adverse events from high levels of radiation (i.e., development of cancer). The lead discussant noted that the preliminary rating was moderate for evidence. NQF staff clarified that this is an intermediate outcome measure, which should be evaluated similarly to a process measure, according to NQF's evaluation criteria. The lead discussant then asked whether there was any linkage to actual outcomes. The developer clarified that if there is no adjustment of the dosing, patients could possibly be exposed to excessive radiation; however, the developer did not specifically describe any link to other outcomes. A Standing Committee member clarified that the intent of the measure is to limit the amount of radiation delivered to patients in order to limit the risk of cancer. The developer clarified that the information linking radiation to cancer was primarily drawn from radiation exposure in World War II from Nagasaki, Japan.

A Standing Committee member suggested limiting this measure to certain patients, particularly those who are at higher risk for high radiation exposure. Another Standing Committee member asked how the recommended amount of radiation was determined. The developer replied that it is based on experience as well as studies that balance the amount of radiation and the quality of the image. There was also a question about trauma exclusions, or potentially acute stroke, in which dosing may be difficult to optimize and attempts to optimize dosing (i.e., by requiring a patient's weight to be estimated by an emergency physician) may interfere with patient care. The developer replied that these were rare cases and would not have a large impact. The developer also clarified that the measure only included CT head, chest, and abdomen and may not include other protocols such as perfusion studies. The lead discussant proceeded to describe a public comment that stated the importance of exposure to ionizing radiation; however, evidence is unclear as to whether this affected specific protocols within facilities. The Standing Committee agreed that this is an important measure and passed the measure on evidence.

Next, the lead discussant described the gap information the developer provided, which revealed that a gap exists for this measure. Moving to a vote, the Standing Committee gave a moderate rating for the performance gap criterion. There was no discussion on the composite construct, which the Standing Committee also deemed as moderate.

The lead discussant noted that the SMP passed the measure on reliability. The Standing Committee did not express any concerns regarding the SMP's high reliability rating for the measure and voted to accept

it. However, the SMP did not reach consensus on validity. The lead discussant noted that the SMP questioned the level of analysis (i.e., clinician group versus facility), specifically whether face validity was conducted at the clinician group or facility level of analysis, or both levels. The developer clarified within their submission that face validity was conducted at both levels of analysis. A Standing Committee member asked whether the measure would exclude certain types of patients, such as pregnant patients, to which the developer replied that this is a very small population. Based upon this discussion, the Standing Committee voted to pass the measure on validity with a moderate rating. There were no concerns or discussion on the composite construct, and the Standing Committee voted to accept the SMP's rating of moderate for the quality construct.

The lead discussant described the feasibility of the measure, which can be extracted from electronic platforms. In addition, the American College of Radiology (ACR) charged fees to submit the data for the Merit-Based Incentive Payment System (MIPS). Based on this information, the Standing Committee gave the measure a moderate rating for feasibility. There were no concerns about use and usability, both of which received a pass and moderate rating, respectively. The Standing Committee ultimately recommended the measure for endorsement. The Standing Committee also observed that there are several related measures to this metric, but it did not consider these measures to be competing.

#0679 Percent of High-Risk Residents With Pressure Ulcers (Long Stay) (Centers for Medicare & Medicaid Services)

Measure Steward/Developer Representatives at the Meeting

Acumen:

- Cheng Lin
- Sri Nagavarapu
- Stephen McKean
- Becky Clearwater
- Aathira Santhosh
- Howard He
- Layla Taha

CMS:

- Rebekah Natanov
- Alan Levitt
- Mary Pratt
- Shequila Purnell-Saunder

Standing Committee Votes

- **Evidence**: Pass-17; No Pass-0 (Pass – 17/17)
- **Performance Gap**: H-10; M-7; L-0; I-0 (Pass – 17/17)
- **Reliability**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
  - The NQF Scientific Methods Panel's rating for Reliability: Moderate, H-0; M-6; L-2; I-0 (Pass – 6/8)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.

    - The NQF Scientific Methods Panel's rating for Validity: Moderate, H-2; M-4; L-2; I-0 (Pass – 6/8)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-13; M-5; L-0; I-0 (Pass – 18/18)
- **Use**: Pass-18; No Pass-0 (Pass – 18/18)
- **Usability**: H-4; M-12; L-2; I-0 (Pass – 16/18)

Standing Committee Recommendation for Endorsement: Yes-17; No-1 (Pass – 17/18)

The Standing Committee recommended this outcome measure for continued endorsement.

The developer provided an overview of the measure, emphasizing pressure ulcers as important and unwanted adverse events, particularly in nursing homes. The appearance of pressure ulcers can be reduced with appropriate staffing and staff composition. This measure has a long history, dating back to 2002, and is publicly reported on Nursing Home Compare. In addition, clarification was given on the staging of pressure ulcers, which described that the measure captures any ulcers that are staged two or greater, and unstageable ulcers.

The lead discussant described the evidence supporting the outcome measure, which is associated with one or more healthcare processes within nursing homes, which can reduce the incidence of pressure ulcers. Based on this information, the Standing Committee passed the measure on evidence.

The lead discussant then noted the measure identifies a gap in care and that there are disparities that can be identified with this measure, particularly by age, race, and socioeconomic status (SES). A Standing Committee member asked why non-Medicaid patients were at higher risk of pressure ulcers. The developer explained that research shows the older population may have lower function than others, which puts them at increased risk. In addition, these patients can have a longer healthcare stay and may be sicker. This Standing Committee member also requested that improved stratification be done in future submissions, to which the developer agreed. Based upon this discussion, the Standing Committee voted "high" on performance gap.

The lead discussant noted that the SMP passed the measure on reliability. The lead discussant then asked a question about the reliability of monitoring pressure ulcers and how this may differ between facilities. The developer replied that this measure is looking at stage 2–4 pressure ulcers, and starting at stage 2, the wound easier to detect clinically. The lead discussant then described the validity testing that was done with the RAND Corporation, comparing the rating with a gold standard; however, the data were old. The developer did also describe a follow-up study showing similar data, and the Minimum Data Set (MDS) form has not changed. Therefore, although the data are old, the results should still be relevant. The Standing Committee voted to accept the SMP's vote on reliability.

Next, the lead discussant described the validity testing for the measure, which consisted of comparing the measure to other measures of nursing home quality, including the facility Star Ratings for Medicare. The SMP passed the measure on validity with a moderate rating. A Standing Committee member commented about potentially stratifying the measure by the degree of the ulcer (i.e., stage) in the future. This is because higher-level ulcers may provide interesting insight into the level of nursing quality. Another Standing Committee member noted that this could reduce the reliability of the measure due to the inter-rater variation between nurses in stage two and three ulcers. A question was raised regarding whether risk adjustment should be discussed as part of the validity, which was confirmed by NQF. The Standing Committee discussed whether risk adjustment should be implemented, particularly because certain patients may be at higher risk for pressure ulcer development. The Standing Committee also discussed that the measure could be stratified by risk, such as by the elderly or frail or

by patients with paraplegia, all of whom are at higher risk for pressure ulcers. There was an additional concern that risk stratification could create additional burden and may be less practical. The developer noted the presence of a tradeoff between simplicity and risk-adjusted or stratification. Another concern was raised regarding the exclusion of stage one ulcers, as this is a stage at which an intervention can be implemented to reduce the progression to later stages. The developer mentioned that the MDS is being respecified and that this is under consideration for the future. However, this has not been changed to date because the measure is working well in practice. Based on this information, the Standing Committee accepted the SMP's moderate rating for validity.

The lead discussant then described the feasibility of the measure, which is based on the MDS. The Standing Committee did not have any concerns related to feasibility and passed the measure with a high rating for this criterion. For use and usability, the lead discussant noted that the measure is broadly used in public programs and for accountability, and its rates have improved over time. Therefore, the Standing Committee gave the measure a passing rating for use and a high rating for usability. Lastly, the Standing Committee voted to recommend the measure for endorsement and ultimately recommended the measure for endorsement. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing.

#0674 Percent of Residents Experiencing One or More Falls With Major Injury (Long Stay) (Centers for Medicare & Medicaid Services)

Measure Steward/Developer Representatives at the Meeting

Acumen:

- Cheng Lin
- Sri Nagavarapu
- Stephen McKean
- Becky Clearwater
- Aathira Santhosh
- Howard He
- Layla Taha

CMS:

- Rebekah Natanov
- Alan Levitt
- Mary Pratt
- Shequila Purnell-Saunder

Standing Committee Votes

- **Evidence**: Pass-18; No Pass-0 (Pass – 18/18)
- **Performance Gap**: H-1; M-17; L-0; I-0 (Pass – 18/18)
- **Reliability**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
  - The NQF Scientific Methods Panel's rating for Reliability: Moderate, H-0; M-6; L-2; I-0 (Pass – 6/8)
  - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**:
  - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.

    - The NQF Scientific Methods Panel's rating for Validity: Moderate, H-1; M-6; L-1; I-0 (Pass – 7/8)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-7; M-12; L-0; I-0 (Pass – 19/19)
- **Use**: Pass-18; No Pass-0 (Pass – 18/18)
- **Usability**: H-5; M-13; L-0; I-0 (Pass – 18/18)

Standing Committee Recommendation for Endorsement: Yes-19; No-0 (Pass – 19/19)

The Standing Committee recommended this outcome measure for continued endorsement.

The developer provided an overview of the measure, noting that it is an outcome measure and that there are interventions that can be implemented to reduce falls with injury in nursing homes. In addition, this measure is captured in the MDS and is publicly reported in Nursing Home Compare. The developer also noted that this measure is both reliable and valid.

The lead discussant presented the evidence for the measure. There were some updates to the evidence since it was last endorsed in 2015; these updates demonstrate structural interventions that can be implemented to reduce falls with injury in long-term care facilities. There were also facility processes, such as using less restraints, that can be performed to reduce falls. Based on this information, the Standing Committee passed the measure on evidence.

The lead discussant did mention that a quality gap in care still exists and that only 19 percent of nursing homes have a rate of zero. There were also differences in the measure rate according to age, race, and SES. However, one Standing Committee member noted that the race disparities were somewhat counterintuitive, as the rates for minorities were lower than would be expected. The developer thought it could be due to staffing levels and that there may be an interaction with other effects that they could look into in the future. Based on this information, the Standing Committee voted "moderate" on performance gap.

Next, the lead discussant described the reliability data, which caused some concerns to emerge, specifically regarding the low reliability results. According to the SMP, this was due to a fifth of the sample having a rate of zero. A Standing Committee member mentioned that it was not necessarily believable that any facility would have a zero rating for this measure. One Standing Committee member commented that this measure is not only looking at falls, but also falls that result in a reportable injury, which may explain the zero-event rate for some facilities. Moving to vote, the Standing Committee accepted the SMP's rating for reliability.

Then, the lead discussant described the validity data, which correlated the measures with other measures of quality. The lead discussant noted a validity concern with respect to reporting bias, as falls are self-reported by the facility. The Standing Committee considered evidence from the literature, which found that the MDS only identified 57 percent of falls in claims and that White patients had 60 percent of falls reported compared with 46 percent of non-White patients. A Standing Committee member recommended that consideration be given to assess underreporting or consider validating with claims data. The developer mentioned that they are planning to conduct quarterly monitoring to assess this matter in the future, linking MDS information to Medicare claims to assess the degree of underreporting. It was also mentioned that this would be difficult in the Medicaid population, as well as Medicare Advantage claims, which are not consistently reported. Based on this discussion, the Standing Committee accepted the SMP's moderate rating for validity.

Lastly, the lead discussant described the feasibility of the measure; this measure is captured in the MDS. The Standing Committee did not have any concerns and gave a moderate rating for feasibility. In addition, use and usability information were described to the Standing Committee. This measure is publicly reported, and improvements have occurred over time with regard to the measure. Ultimately, the Standing Committee gave use a passing rating and usability a moderate rating. The Standing Committee voted to recommend the measure for endorsement and ultimately recommended the measure for endorsement. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing.

#3501e Hospital Harm – Opioid-Related Adverse Events (Centers for Medicare & Medicaid Services/IMPAQ International, LLC)

This is an electronic clinical quality measure (eCQM).

Measure Steward/Developer Representatives at the Meeting

IMPAQ International

- Kendall Hall
- Patrick Romano
- Bo Feng
- Mia Nievera
- Michelle Lefebvre
- Chana West
- Katie Magoulick
- Anna Michie
- Hannah Klein

Standing Committee Votes

- **Evidence**: Pass-10; No Pass-6 (Pass – 10/16)
- **Performance Gap**: H-0; M-7; L-5; I-4 (Consensus Not Reached – 7/16)
- **Reliability**:
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Reliability: Moderate, H-2; M-5; L-0; I-1 (Pass – 7/8)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Validity**:
    - This measure is deemed as complex and was evaluated by the NQF Scientific Methods Panel.
    - The NQF Scientific Methods Panel's rating for Validity: Moderate, H-1; M-6; L-1; I-0 (Pass – 7/8)
    - The Standing Committee accepted the NQF Scientific Methods Panel's rating.
- **Feasibility**: H-7; M-11; L-0; I-0 (Pass – 18/18)
- **Use**: Pass-17; No Pass-1 (Pass – 17/18)
- **Usability**: H-1; M-11; L-2; I-4 (Pass – 12/18)

Standing Committee Recommendation for Endorsement: Yes-X; No-X

The Standing Committee did not vote on the recommendation for endorsement for this outcome measure because they did not reach consensus on performance gap—a must-pass criterion. The

Standing Committee will re-vote on this criterion and vote on the overall suitability for endorsement during the post-comment web meeting on October 13, 2021.

The developer provided an overview of the measure, noting that it is a facility-level measure that assessed the percentage of patients who have received opioids as well as naloxone. The Patient Safety Standing Committee last reviewed this measure in spring 2019. Since that time, the developer made changes to the measure based on the feedback received from the Standing Committee during the spring 2019 evaluation. The specific changes include the following: (1) the denominator has been changed to those receiving at least one opioid during the hospitalization; (2) any naloxone administration needs to be preceded by an opioid with a time parameter; (3) measure value sets have been updated to include all opioids; and (4) it will now be determined whether there is enough variation across sites.

Next, the lead discussant described the evidence for the measure. Several studies have demonstrated how naloxone administration is used to identify adverse drug events in the hospital and that there are healthcare actions that can be used to reduce opioid-related adverse events. One Standing Committee member asked whether naloxone administration is an appropriate outcome. Another Standing Committee member asked whether naloxone administration is an actual adverse event; they also commented that this may overly capture some appropriate medical care. In particular, naloxone can be used for many other reasons beyond opioid overdose only, such as diagnostic reasons to assess whether opioid overdose is the correct diagnosis. The developer replied that nurse reviewers assessed why patients received the medication as well as the response, which was performed in most of the cases for respiratory depression related to opioids (98 percent of the time), and that it was given for opioid reversal and did result in improvement in the patient's level of consciousness (76 percent of the time). Moving to a vote, the Standing Committee passed the measure on evidence.

Following this vote, the lead discussant described the performance gap for the measure, which was tested in six hospitals with measure rates ranging from 0.11 to 0.45 percent. The Standing Committee discussed whether a performance gap was truly present because the absolute rate was low. Some Standing Committee members noted that a gap exists; there were four-fold differences across the six sites tested. A discussion was also held as to whether the number of events, which was low, truly showed differences across sites. Moving to a vote, the Standing Committee did not reach consensus on performance gap.

The lead discussant then described the reliability results, which were a comparison of electronically extracted versus manually extracted data. The SMP reviewed the reliability testing and gave the measure a moderate, passing rating for reliability. The Standing Committee did not have any major concerns related to reliability and voted to accept the SMP's rating.

Next, the lead discussant described the validity testing for the measure, which demonstrated excellent accuracy in detecting whether naloxone was given after an opioid administration. This assessment used several elements in the charts, including nursing notes and other documentation. The SMP reviewed the validity testing and passed the measure with a moderate rating for validity. A Standing Committee member asked a clarifying question, specifically whether the Standing Committee was evaluating the clinical validity of this measure, which was confirmed by the NQF staff. Discussion was held with regard to the exclusion of patients who were in the operating room and how this was identified. In two of the 23 measure testing sites, an issue occurred with detecting whether the patient was in the operating room. However, there were other proxy ways to measure this detection, such as the location of the administering provider. Based upon this discussion, the Standing Committee voted to uphold the SMP's assessment of validity.

Lastly, the lead discussant described the feasibility of the measure. The Standing Committee commented that there may be some feasibility challenges with anesthesiologists documenting naloxone use on paper charts. Therefore, the Standing Committee voted to pass the measure with a moderate rating for feasibility. Regarding use and usability, the lead discussant mentioned that the developer envisioned using this measure in public programs in the future, as this was a novel measure. There was some discussion about unintended consequences that should be evaluated with this measure, given the prior discussion after this is implemented. It was also mentioned that naloxone could be used as a trigger tool in hospitals to identify problems and target quality improvement efforts. Based on this discussion, the Standing Committee gave the measure a passing rating for use and a moderate rating for usability. Because consensus was not reached on performance gap, no vote was taken on overall suitability for endorsement nor was a related and competing measure(s) discussion held. If the measure passes on performance gap and is recommended for endorsement during the October 2021 post-comment call, the Standing Committee will then proceed with a related and competing measure(s) discussion.

#3389 Concurrent Use of Opioids and Benzodiazepines (COB) (Pharmacy Quality Alliance)

Measure Steward/Developer Representatives at the Meeting

- Ben Shirley
- Lisa Hines
- Meghan Gabriel

Standing Committee Votes

- **Evidence**: H-6; M-12; L-0; I-0 (Pass – 18/18)
- **Performance Gap**: H-11; M-6; L-1; I-0 (Pass – 17/18)
- **Reliability**: H-4; M-14; L-0; I-0 (Pass – 18/18)
- **Validity**: H-3; M-14; L-1; I-0 (Pass – 17/18)
- **Feasibility**: H-6; M-12; L-0; I-0 (Pass – 18/18)
- **Use**: Pass-18; No Pass-0 (Pass – 18/18)
- **Usability:** H-11; M-7; L-0; I-0 (Pass – 18/18)

Standing Committee Recommendation for Endorsement: Yes-17; No-1 (Pass – 17/18)

The Standing Committee recommended this process measure for continued endorsement.

The developer presented an overview of the measure, noting that this is a maintenance measure developed for health plan accountability. The developer described the importance of this measure by highlighting the healthcare problems related to opioid overdose and the need for opioid-related measures. To address this matter, the Centers for Disease Control and Prevention (CDC) issued class A recommendation and the Food and Drug Administration (FDA) issued black box warning against the use of opioids with benzodiazepines, which can increase the risk of overdose. The developer mentioned that this measure had been used for public accountability and still has room for improvement.

The lead discussant started the review with presenting the evidence for the measure. Specifically, the CDC guidelines give a category A recommendation for this measure. The developer also provided additional studies that support the measure's continued measurement. It was also mentioned that the Medicare population was more adversely affected by opioid and benzodiazepine combination prescribing than other groups. Furthermore, it was mentioned that sickle cell, cancer, and hospice were not included in the denominator for the measure. The Standing Committee did not have any major concerns and voted to pass the measure on evidence.

Next, the lead discussant described the performance gap data for this measure. The Standing Committee agreed that a substantial gap remains and passed the measure on performance gap. Following this review, the lead discussant presented the data on reliability and validity. The Standing Committee did not raise any questions or concerns and voted to pass the measure with moderate ratings for both reliability and validity. In addition, the Standing Committee did not have any concerns with feasibility and voted to pass the measure on feasibility.

The lead discussant presented the data on use and usability, noting that this measure has seen improvements over time and that the developer noted its future use in accountability programs. There was no discussion from the Standing Committee on either use or usability. Therefore, the Standing Committee passed the measure on use and usability. Lastly, the Standing Committee voted to recommend the measure for endorsement and ultimately recommended the measure for endorsement. The Standing Committee observed that there are several related measures to this metric, but it did not consider these measures to be competing.

Public Comment

A public comment was submitted about the sepsis measure (NQF #0500), which expressed support. The commenter focused on the epidemiology of sepsis, noting that the burden of sepsis is increased among communities of color and that early interventions are important. With respect to the overuse of antibiotics, the commenter described the importance of antibiotic stewardship. A second commenter on the sepsis measure focused on sepsis care in rural hospitals, stating that the measure was very important, particularly in smaller hospitals, to protocolize sepsis care, which has been a benefit to patient care. A third commenter from the IDSA raised concerns about the evidence for the sepsis measure, unintended consequences regarding antibiotic overuse, and outsized burden on providers. A public comment on the CT measure (NQF #3621) did not express support for the measure due to validity concerns. The commenter expressed that the measure ignores whether variation is present in the choice of protocol, which is greater than the variation in the actual radiation within the protocol. The Standing Committee co-chairs thanked the commenters for their comments.

Next Steps

NQF will post the draft technical report on August 11, 2021, for public comment for 30 calendar days. The continuous public commenting period with member support will close on September 9, 2021. NQF will reconvene the Standing Committee for the post-comment web meeting on October 13, 2021.