NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 2726

Corresponding Measures:

De.2. Measure Title: Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections

Co.1.1. Measure Steward: American Society of Anesthesiologists

De.3. Brief Description of Measure: Percentage of patients, regardless of age, who undergo central venous catheter (CVC) insertion for whom CVC was inserted with all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques followed

1b.1. Developer Rationale: Hospital-acquired bloodstream infections are a common complication that leads to increased costs and mortality. It is estimated that approximately 51% of hospital-acquired bloodstream infections occur in an intensive care unit (ICU), with the presence of a central venous catheter being the largest risk factor for the development of a bloodstream infection in the hospital. Catheter-related bloodstream infections (CRBSIs) commonly occur when the catheter becomes contaminated by microbes on the skin during insertion. The use of maximal sterile barriers, including sterile gloves, long-sleeved sterile gown, mask, cap, and full-sized sterile drape, during insertion of the catheter has been shown to cost effectively reduce CRBSI rates compared to the use of less stringent precautions.

A 2002 survey found that only 28% of ICUs have written policies requiring all five components of maximum sterile-barrier (MSB) precautions during central venous catheter insertion and 20% of ICUs had no written policies addressing sterile barrier precautions. Similarly, a 2005 survey of physicians showed that only 28% used all components of MSB precautions. This survey showed that low usage was driven by the belief that MSB precautions would not have a significant impact on infection rates. One study demonstrated that a one-day course on proper sterile techniques during catheter insertion for physicians resulted in increased use of MSB precautions and reduced CRBSI rates.

S.4. Numerator Statement: Patients for whom central venous catheter (CVC) was inserted with all elements of maximal sterile barrier technique*, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques** followed

Definitions:

*Maximal sterile barrier technique includes ALL of the following elements:

- cap
- mask
- sterile gown

- sterile gloves
- sterile full body drape

** Sterile ultrasound techniques require sterile gel and sterile probe covers

S.6. Denominator Statement: All patients, regardless of age, who undergo CVC insertion

S.8. Denominator Exclusions: None

The measure includes a denominator exception as indicated by reporting 6030F-1P for the numerator: Documentation of medical reason(s) for not following all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques during CVC insertion (including increased risk of harm to patient if adherence to aseptic technique would cause delay in CVC insertion)

De.1. Measure Type: Process

S.17. Data Source: Registry Data

S.20. Level of Analysis: Clinician : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Dec 10, 2015 **Most Recent Endorsement Date:** Dec 10, 2015

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2015

- Various recommendations statements from the CDC's Guidelines for the Prevention of Intravascular Catheter-Related Infections, (2011) are provided that support the measure.
    - The guidelines provides recommendation statements related to Hand Hygiene and Aseptic Technique, Maximal Sterile Barrier (MSB) Precautions, and Skin Preparation.

- The body of evidence examines the relationship between MSB precautions and catheter-related bloodstream infection (CRBSI) rates.
    - 23 studies support the body of evidence. Quality of evidence is high and consistently supports reducing the risk of bloodstream infections.
- Other studies provided by the developer (3 randomized controlled trials and 1 observational study) also supported the evidence base used in the Guidelines.
- The developer also provided an FDA recommendation that policies and standards be reviewed to include the use of sterile ultrasound gel.
- The Committee agreed that there is a very strong connection with outcomes.
    - AHRQ has reported a precipitous drop in CLABSI central line infections since this measure has been in use.

Changes to evidence from last review

☒ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☐ **The developer provided updated evidence for this measure:**

Questions for the Committee:

- The developer attests the underlying evidence for the measure has not changed since the last NQF endorsement review. Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion and vote on Evidence?

Guidance from the Evidence Algorithm

Process measure based on systematic review (Box 3) → QQC presented (Box 4) → Quantity: moderate/high; Quality: high; Consistency: high →High (Box 5a) → High

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- Developer provides MIPS performance data from 2016 and 2017 as well as 2018 Preliminary NACOR performance data for MIPS.

2016: Average Performance Rate - 93.3%, Standard Deviation - 15.6%

Decile 3: 91.85-96.14

Decile 4: 96.15-98.88

Decile 5: 98.89-99.99

Deciles 6-10: 100

2017: Average Performance Rate - 94.2%, Standard Deviation - 15.7%

Decile 3: 95.67-99.08

Decile 4: 99.09-99.99

Deciles 5-10: 100

2018: Average Performance Rate - 97.08%, Standard Deviation - 15.75%

- Results indicate a narrow gap with limited room from improvement.
- During the previous evaluation, the Committee discussed that there was a significant gap in reporting that indicated a potential gap in performance (though those who do report are successful).

Disparities

- The developer provides the age and gender breakdown of the reporting population but not performance scores by these patient characteristics.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- Does the standard deviation (approximately 15.7%) indicate there is continued room for improvement?

Preliminary rating for opportunity for improvement: ☐ High ☒ Moderate ☐ Low ☐ Insufficient

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence

Comments:

**Evidence has not changed since last endorsement

**no updated evidence but still strong

**nothing that changes the factors around this measure

**acceptable

1b. Performance Gap

Comments:

**2018: Average Performance Rate - 97.08%

**it would appear that there is still room for imporvement although measure does seem to be topping out

**Of course there is no pediatric inclusion in this measure, although the standards for CVC insertion is not different than in adults

**appears "topped out" however measure only applies to MDA/CRNAs who place central lines + chose to report to registry

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

Review A

Evaluation of Reliability and Validity:

- There have been no changes in reliably testing since 2015 evaluation. Mean reliability scores were consistently >0.9.
- During this maintenance review, the developer conducted empirical score-level validity testing by comparing average PQRS performance rates to nationally-reported central-line associated bloodstream infection (CLABSI) standardized infection ratios (SIR) for the same time period. Results showed as central line insertion practices improved, national CLABSI rates decreased. The developer states that the data sample did not allow for a correlation analysis (statistical test) to be performed.
- The developer also conducted face validity testing. 17 of 19 TEP members agreed that the scores from the measure as specified would provide an accurate reflection of quality and 2 disagreed. Mean Rating: 4.16 out of 5.
- Regarding the ability to meaningful performance differences, average performance is high (93.3%) and the distribution shows narrow differences between deciles. The developer states that the standard deviation (15.6%) supports the ability to identify meaningful differences in care.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Does the data indicating improvements in measures scores with decreases in CLABSI SIR for the same years, absent statistical correlation, plus supportive face validity testing prove the measure's validity?
- Is there concern about the relatively high measure performance rates across deciles and ability to detect meaningful differences in performance?

Preliminary rating for reliability: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Evaluation A: Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 2726

Measure Title: Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☒ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☒ **Clinician: Group/Practice** ☒ **Clinician: Individual** ☐ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

   None.

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**
6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   - The developer tested reliability based on the beta-binomial model using a signal-to-noise ratio.

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - There have been no changes in reliably testing since 2015 evaluation.
   - Mean reliability scores were consistently >0.9 from 2012-2014 for providers and practices using National Anesthesia Clinical Outcomes Registry (NACOR) data as well as for providers (n>2) in 2012-2013 using Medicare SAF 5% File data.

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☒ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☐ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**
    - Reliability testing has not been updated since the previous evaluation.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.
    - No exclusions. During 2015 review, a Committee member raised the concern that this measure should not apply to premature infants, who are likely to have adverse effects from one of the skin preparation solutions included in the specifications.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.
    - 2018 MIPS data indicated average performance of 93.3% with a standard deviation of 15.6%.
    - Decile 3 shows performance of 91.85-96.14%, while Deciles 6 and above are all at 100%.
    - Average performance is high and the distribution shows narrow differences between deciles. The developer states that the SD supports the ability to identify meaningful differences in care.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.
    N/A

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.
    - Developer states there is no missing data since NACOR requires all data elements in order to report.

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

    16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

    ☐ Yes ☐ No ☒ Not applicable

    16c. **Social risk adjustment:**

    16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☐ Not applicable

    16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

    16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

    16d. **Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5. Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

N/A

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☐ Data element ☐ Both**

18. **Method of establishing validity of the measure score:**

☒ Face validity

☒ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

19. **Assess the method(s) for establishing validity**

Submission document: Testing attachment, section 2b2.2

Empirical Validity

- Empirical validity testing includes data from the 2016 PQRS Experience Report and the CDC NHSN.
- To test validity the developer compared average PQRS performance rates to nationally-reported central-line associated bloodstream infection (CLABSI) standardized infection ratios (SIR) for the same time period. The developer states that the data sample is too small to conduct a statistical estimate of correlation.

Face Validity:

- Assessed by TEP of 19 physicians. Measure previously passed with a Moderate rating based on face validity.

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

Empirical Validity:
MIPS Avg. Performance Rate, NHSN National CLABSI SIR
2013: 82.8%, 0.53
2014: 84.5%, 0.49
2015: 88.3%, 0.60*
2016: 91.2%, 0.56*
**In 2015, the CDC changed their surveillance protocol, resulting in an increase in the reported SIR. The subsequent decline in infections from 2015 to 2016, combined with the downward trend in infections prior to 2015, support a likelihood that the increase from 2014-2015 is solely due to the surveillance protocol change.*

- Additionally, the CDC says, "dynamics in the microbiology of CLABSIs, along with lagging declines on wards where central lines are on average in place longer than ICU central lines, suggest prevention successes due primarily to improved central line insertion practices".
- The available data support that improvements in sterile technique for central line insertion may play a role in national reductions in CLABSIs.
- The analysis provided supports the measure's validity, but a statistical correlation test was not conducted.

Face Validity:

- 17 of 19 TEP members agreed that the scores from the measure as specified would provide an accurate reflection of quality and 2 disagreed. Mean Rating: 4.16 out of 5.

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?** *NOTE that data element validation from the literature is acceptable.*

Submission document*: Testing attachment, section 2b1.*

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

- Data indicates improvements in measures scores with decreases in CLABSI SIR over the same years.
    - Is the Committee concerned about the lack of a correlation coefficient or comfortable with the measure's validity and the developer's response that the sample size does not allow for the performance of a statistical correlation test?

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**
    - Is there concern about the relatively high measure performance rates across deciles and ability to detect meaningful differences in performance?
    - Is data indicating improvements in measures scores with decreases in CLABSI SIRs for the same years, absent statistical correlation, and positive face validity testing convincing of validity?

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability – Specifications

Comments:

**There have been no changes in reliably testing since 2015 evaluation. Mean reliability scores were consistently >0.9.

**No concerns

**I think the expected practice around CVC insertion can be fully implemented. My question is can you get reliable compliance data from a registry. I don't think so

**registry data used are dated

2a2. Reliability – Testing

Comments:

**No

**No concerns

**Yes, I am unsure that registry data correctly captures compliance with insertion practices with sufficient granularity. If this is self-reported compliance, it becomes useless

**would like to see updated reliability testing using the registry itself

2b1. Validity –Testing

Comments:

**No

**None

**see comments as under 6.2a2

**valid only for MDAs/CRNAs who place central lines and choose to report to the registry; would like to know how representative these folks are of ALL MDAs/CRNAs who place central lines and also what % of all central lines are NOT placed by MDA/CRNA?

2b4-7. Threats to Validity
2b4. Meaningful Differences

Comments:

**No

**we should discuss "topping out" is this really a meanigful measure still

**probably not

**way it is constructed they claim no missing data; self selection bias risk here

2b2-3. Other Threats to Validity
2b2. Exclusions
2b3. Risk Adjustment

Comments:

**Appropriate

**No concerns

**I don't think risk adjustment is needed. The problem with this measure is that it is ONLY a process measure, not tied to clinical outcomes (i.e. Catheter associated BSI)

**acceptable

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The measure uses registry data. Data are generated during the provision of care. Some data elements are in defined fields in electronic sources.
- The measure uses CPT codes to capture the process of maximum sterile barrier technique. Components of maximal barrier precautions are regularly included in EHRs and noted that the measure could potentially become an eCQM in the future.
- The developer noted limited proprietary coding is included in the specifications for convenience and users should obtain all necessary licenses from the owners of these code sets.

Questions for the Committee:

- Are the required data elements routinely generated and available?
- Is measure implementable without undue burden?
- Is the proprietary nature of the specifications a concern?

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility

Comments:

**Moderate

**No concerns

**all the elements of sterility during CVC insertion are NOT routinely collected. A separate process to capture all the elements will be needed. Also is compliance all or none - if one element is missing, does give you a zero for the denominator?? Many concerns with the reliability in capturing the data elements.

**how much training/education/coaching is really required?

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

Accountability program details

- The measure is used in the Merit-based Incentive Payment System (MIPS) and for external benchmarking in the Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (NACOR).

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- AQI NACOR provides performance results through a dashboard as well as provides assistance to users on reporting and interpreting data. Aggregate group-level reports are also available for users.
- Feedback is obtained through various avenues. Based on the request for clarification regarding who is expected to report the measure, implicit instructions were included in the specifications that ECs who perform CVC insertion should report. Account managers have successfully worked with users to address difficulties in capturing surgical billing information.

Additional Feedback:

N/A

***Questions for the Committee**:*

- How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

RATIONALE:

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- Performance has improved over time from 93.9% in 2016 to 97.06% in 2018. User uptake has increase based on use in MIPS. Approximately 50,000 providers reported the measure in 2017.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- None.

Potential harms

- The time required to complete the documentation is trivial and supported by existing aesthesia records. The time required to comply is trivial compared to preventing a CVC-associated infection.

Additional Feedback:

- N/A

***Questions for the Committee**:*

- Can the performance results continue to be used to further the goal of high-quality, efficient healthcare?

Preliminary rating for Usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency

Comments:

**The measure is used in the Merit-based Incentive Payment System (MIPS) and for external benchmarking in the Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (NACOR).

**None

**Not sure

**acceptable

4b1. Usability – Improvement

Comments:

**N/A

**None

**None

**no real down side; if I know I am compliant I will choose to report and look good

Criterion 5: Related and Competing Measures

Related or competing measures

- Related measure: 0139: National Healthcare Safety Network (NHSN) Central line-associated Bloodstream Infection (CLABSI) Outcome Measure.
    - Differences include measure type (process versus outcome) and different levels of analysis (2726 is specified at the clinician level, while 0139 is specified at the facility level).

Harmonization

- The Committee previously discussed that both process and outcome measures exist around this issue, and the developer explained that the measures are complimentary and serve different purposes.
- The Committee will discuss potential harmonization if necessary.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing

Comments:

**Related measure: 0139

**related measure -- the standard is the outcome from NHSN but there seems to still be a role for the process measure

**not that I am aware of

**understand the complementarity w/ the NHSN outcome measure

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 6/5/2019

- **No NQF Members have submitted support/non-support choices as of this date.**

Developer Submission

Brief Measure Information

NQF #: 2726

Corresponding Measures:

De.2. Measure Title: Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections

Co.1.1. Measure Steward: American Society of Anesthesiologists

De.3. Brief Description of Measure: Percentage of patients, regardless of age, who undergo central venous catheter (CVC) insertion for whom CVC was inserted with all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques followed

1b.1. Developer Rationale: Hospital-acquired bloodstream infections are a common complication that leads to increased costs and mortality. It is estimated that approximately 51% of hospital-acquired bloodstream infections occur in an intensive care unit (ICU), with the presence of a central venous catheter being the largest risk factor for the development of a bloodstream infection in the hospital. Catheter-related bloodstream infections (CRBSIs) commonly occur when the catheter becomes contaminated by microbes on the skin during insertion. The use of maximal sterile barriers, including sterile gloves, long-sleeved sterile gown, mask, cap, and full-sized sterile drape, during insertion of the catheter has been shown to cost effectively reduce CRBSI rates compared to the use of less stringent precautions.

A 2002 survey found that only 28% of ICUs have written policies requiring all five components of maximum sterile-barrier (MSB) precautions during central venous catheter insertion and 20% of ICUs had no written policies addressing sterile barrier precautions. Similarly, a 2005 survey of physicians showed that only 28% used all components of MSB precautions. This survey showed that low usage was driven by the belief that MSB precautions would not have a significant impact on infection rates. One study demonstrated that a one-day course on proper sterile techniques during catheter insertion for physicians resulted in increased use of MSB precautions and reduced CRBSI rates.

S.4. Numerator Statement: Patients for whom central venous catheter (CVC) was inserted with all elements of maximal sterile barrier technique*, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques** followed

Definitions:

*Maximal sterile barrier technique includes ALL of the following elements:

- cap
- mask
- sterile gown
- sterile gloves
- sterile full body drape

** Sterile ultrasound techniques require sterile gel and sterile probe covers

S.6. Denominator Statement: All patients, regardless of age, who undergo CVC insertion

S.8. Denominator Exclusions: None

The measure includes a denominator exception as indicated by reporting 6030F-1P for the numerator: Documentation of medical reason(s) for not following all elements of maximal sterile barrier technique, hand hygiene,

skin preparation and, if ultrasound is used, sterile ultrasound techniques during CVC insertion (including increased risk of harm to patient if adherence to aseptic technique would cause delay in CVC insertion)

De.1. Measure Type: Process

S.17. Data Source: Registry Data

S.20. Level of Analysis: Clinician : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Dec 10, 2015 **Most Recent Endorsement Date:** Dec 10, 2015

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? NA

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

2726_evidence_attachment.docx

1a.1 <u>For Maintenance of Endorsement:</u> Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

No

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 2726

Measure Title: Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: Click here to enter a date

Instructions

- ***Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.***
- ***Complete EITHER 1a.2, 1a.3 or 1a.4 as applicable for the type of measure and evidence.***
- ***For composite performance measures:***

- *A separate evidence form is required for each component measure unless several components were studied together.*
- *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*

- **All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.**
- **If you are unable to check a box, please highlight or shade the box for your response.**
- **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.**
- **Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.**
- **Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.**
- **Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.**
- **Efficiency: [6] evidence not required for the resource use component.**
- **For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.**
- **Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.**

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: CVC insertion with all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques followed

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

By following maximal sterile barrier (MSB) techniques during central venous catheter (CVC) insertion, providers reduce the risk of the patient acquiring a catheter-related bloodstream infection. This is a process-health outcome relationship. CPT procedure codes are used to identify patients who are included in the measure's denominator. When reporting via claims, CPT Category II codes are used to report the numerator. When reporting via registry, listed numerator options are used to report the numerator.

Anesthesia provider follows maximal sterile barrier (MSB) techniques during central venous catheter (CVC) insertion

↓

Reduces the risk of the patient acquiring a catheter-related bloodstream infection

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

N/A

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

☒Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	• **Guidelines for the Prevention of Intravascular Catheter-Related Infections** • **O'Grady NP, Alexander M, Burns LA, et al.** • **2011** • **O'Grady NP, Alexander M, Burns LA, et al. (2011). Guidelines for the Prevention of Intravascular Catheter-Related Infections, 2011. *Centers for Disease Control and Prevention*. Washington, DC pp. 28-30.** • **https://www.cdc.gov/infectioncontrol/guidelines/pdf/bsi/bsi-guidelines-H.pdf** • **FDA Safety Communication: UPDATE on Bacteria Found in Other-Sonic Generic Ultrasound Transmission Gel Poses Risk of Infection. U.S. Food and Drug Administration, US Dept of Health and Human Services. June 8, 2012.**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	Hand Hygiene and Aseptic Technique (p.12) 1. Perform hand hygiene procedures, either by washing hands with conventional soap and water or with alcohol-based hand rubs (ABHR). Hand hygiene should be performed before and after palpating catheter insertion sites as well as before and after inserting, replacing, accessing, repairing, or dressing an intravascular catheter. Palpation of the insertion site should not be performed after the application of antiseptic, unless aseptic technique is maintained. Category IB

	2. Maintain aseptic technique for the insertion and care of intravascular catheters. Category IB 3. Wear clean gloves, rather than sterile gloves, for the insertion of peripheral intravascular catheters, if the access site is not touched after the application of skin antiseptics. Category IC 4. Sterile gloves should be worn for the insertion of arterial, central, and midline catheters. Category IA 5. Use new sterile gloves before handling the new catheter when guidewire exchanges are performed. Category II 6. Wear either clean or sterile gloves when changing the dressing on intravascular catheters. Category IC Maximal Sterile Barrier Precautions (p.12) 1. Use maximal sterile barrier precautions, including the use of a cap, mask, sterile gown, sterile gloves, and a sterile full body drape, for the insertion of CVCs, PICCs, or guidewire exchange. Category IB 2. Use a sterile sleeve to protect pulmonary artery catheters during insertion. Category IB Skin Preparation (p.13) 1. Prepare clean skin with an antiseptic (70% alcohol, tincture of iodine, or alcoholic chlorhexidine gluconate solution) before peripheral venous catheter insertion. Category IB 2. Prepare clean skin with a >0.5% chlorhexidine preparation with alcohol before central venous catheter and peripheral arterial catheter insertion and during dressing changes. If there is a contraindication to chlorhexidine, tincture of iodine, an iodophor, or 70% alcohol can be used as alternatives. Category IA 3. No comparison has been made between using chlorhexidine preparations with alcohol and povidone-iodine in alcohol to prepare clean skin. Unresolved issue. 4. No recommendation can be made for the safety or efficacy of chlorhexidine in infants aged <2 months. Unresolved issue 5. Antiseptics should be allowed to dry according to the manufacturer's recommendation prior to placing the catheter. Category IB Sterile Ultrasound The Food and Drug Administration recommends that policies and clinical practice standards be reviewed to ensure the use of sterile ultrasound gel. Once a container of sterile or non-sterile ultrasound gel is opened, it is no longer sterile and contamination during ongoing use is possible.

Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The evidence supporting the recommendations related to skin preparation has been assigned a grade of IA. The evidence supporting the recommendations related to maximal sterile barrier technique, hand hygiene has been assigned a grade of IB. The recommendation related to sterile ultrasound technique was not assigned a grade.
Provide all other grades and definitions from the evidence grading system	Category IA – Strongly recommended for implementation and strongly supported by well-designed experimental, clinical, or epidemiologic studies. Category IB – Strongly recommended for implementation and supported by some experimental, clinical, or epidemiologic studies and a strong theoretical rationale; or an accepted practice (e.g., aseptic technique) supported by limited evidence. Category IC – Required by state or federal regulations, rules, or standards. Category II – Suggested for implementation and supported by suggestive clinical or epidemiologic studies or a theoretical rationale. Unresolved Issue – Represents an unresolved issue for which evidence is insufficient or no consensus regarding efficacy exists.
Grade assigned to the **recommendation** with definition of the grade	The grade assigned to the strength of recommendation is included in the grade assigned to the evidence above.
Provide all other grades and definitions from the recommendation grading system	See above.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	14 citations provided in support of the recommendations cited above were pre-post observational studies, 6 were prospective randomized controlled trials, 1 was a retrospective cohort study, and 2 were reviews (1 about Swanz-Ganz catheters, the other HICPAC guidelines for hand hygiene in health care settings). The quality of evidence across studies is high. Of the 13 recommendations provided in the CDC guidelines, 8 were graded as levels IA or IB (the two highest levels of evidence). Almost half of the studies cited by the CDC were randomized controlled trials. Of the 12 studies that were not reviews, 7 had very large sample sizes (>400, highest >4,000), 4 had relatively large sample sizes (>100, <400), and the smallest sample was 82 patients (Carrer 2005). The study designs minimize bias, and the large study samples increase statistical power.

Estimates of benefit and consistency across studies	The direction of the effect shows uniform support for MSB techniques reducing the risk for bloodstream infections. Studies that examine the same outcomes show similar results. The evidence consistently shows that when sterile techniques are used, the rate of bloodstream infections is reduced by approximately 25-35 percent (highest was 66 percent, Coopersmith et al 2002). The evidence consistently shows that when sterile techniques are not used, the risk of bloodstream infections and/or bacterial contamination is approximately 2-3 times higher (highest was 6.3, Raad et al 1994). The evidence consistently shows that after provider-focused interventions to increase compliance with sterile techniques, compliance approximately doubled (sometimes tripled, as in Sherertz et al 2000).
What harms were identified?	None of the studies reported adverse events or any harms associated with using MSB techniques. Although it may require additional man hours to properly train physicians in MSB techniques, this minor negative effect is heavily outweighed by improvements to patient safety and economic cost.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Gerolemou L, Fidellaga A, Rose K, Cooper S, Venturanza M, Aqeel A, Han Q, Jones J, Shapiro J, Khouli H: Simulation-based training for nurses in sterile techniques during central vein catheterization. Am J Crit Care 2014; 23(1):40-8 • A pre-post observational study surrounding an intervention of simulation-based nurse training in sterile techniques during CVC insertion. After intervention, there was an 85 percent reduction in catheter-related infections in the critical care unit. This supports the conclusions of the body of evidence. Khouli H, Jahnes K, Shapiro J, Rose K, Mathew J, Gohil A, Han Q, Sotelo A, Jones J, Aqeel A, Eden E, Fried E: Performance of medical residents in sterile techniques during central vein catheterization: randomized trial of efficacy of simulation-based training. Chest 2011; 139(1):80-7 • A pre-post observational study surrounding an intervention of simulation-based training in sterile techniques during CVC insertion. After intervention, there was a 70 percent reduction in the incidence of CRBSI in the medical ICU. This supports the conclusions of the body of evidence.

	Timsit JF, Mimoz O, Mourvillier B, Souweine B, Garrouste-Orgeas M, Alfandari S, Plantefeve G, Bronchard R, Troche G, Gauzit R, Antona M, Canet E, Bohe J, Lepape A, Vesin A, Arrault X, Schwebel C, Adrie C, Zahar JR, Ruckly S, Tournegros C, Lucet JC: Randomized controlled trial of chlorhexidine dressing and highly adhesive dressing for preventing catheter-related infections in critically ill adults. Am J Respir Crit Care Med 2012; 186(12):1272-8 • A randomized controlled trial with chlorhexidine dressings acting as the intervention. With chlorhexidine dressings, the major catheter-related infection rate was 67 percent lower than with standard dressings. This supports the conclusions of the body of evidence. Wu PP, Liu CE, Chang CY, Huang HC, Syu SS, Wang CH, Huang YC: Decreasing catheter-related bloodstream infections in the intensive care unit: interventions in a medical center in central Taiwan. J Microbiol Immunol Infect 2012; 45(5):370-6 A pre-post observational study surrounding an intervention standardizing the process of CVC insertion using MSB precautions. After intervention, catheter-related infection rates decreased to zero, which was sustained for 6 months after the intervention. This supports the conclusions of the body of evidence. Centers for Disease Control and Prevention. 2017 Recommendations on use of chlorhexidine-impregnated dressings for prevention of intravascular catheter-related infections: An update to the 2011 guidelines for the prevention of intravascular catheter-related infections from the Centers for Disease Control and Prevention. Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Diseases, Division of Healthcare Quality and Promotion. https://www.cdc.gov/infectioncontrol/guidelines/pdf/bsi/c-i-dressings-H.pdf. Accessed April 2, 2019. In 2017, the CDC performed a focused update of their 2011 guideline that targeted evolving evidence related to chlorhexidine-impregnated dressings. This focused update did not result in any changes to the recommendations cited in support of this measure.

1a.4 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

Hospital-acquired bloodstream infections are a common complication that leads to increased costs and mortality. It is estimated that approximately 51% of hospital-acquired bloodstream infections occur in an intensive care unit (ICU), with the presence of a central venous catheter being the largest risk factor for the development of a bloodstream infection in the hospital. Catheter-related bloodstream infections (CRBSIs) commonly occur when the catheter becomes contaminated by microbes on the skin during insertion. The use of maximal sterile barriers, including sterile gloves, long-sleeved sterile gown, mask, cap, and full-sized sterile drape, during insertion of the catheter has been shown to cost effectively reduce CRBSI rates compared to the use of less stringent precautions.

A 2002 survey found that only 28% of ICUs have written policies requiring all five components of maximum sterile-barrier (MSB) precautions during central venous catheter insertion and 20% of ICUs had no written policies addressing sterile barrier precautions. Similarly, a 2005 survey of physicians showed that only 28% used all components of MSB precautions. This survey showed that low usage was driven by the belief that MSB precautions would not have a significant impact on infection rates. One study demonstrated that a one-day course on proper sterile techniques during catheter insertion for physicians resulted in increased use of MSB precautions and reduced CRBSI rates.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

2016: MIPS Historical Benchmarks using 2016 Performance Data

Average Performance Rate- 93.3%

Standard Deviation- 15.6%

Performance Deciles:

3: 91.85-96.14

4: 96.15-98.88

5: 98.89-99.99

6: --

7: --

8: --

9: --

10: 100.00

2017: MIPS Historical Benchmarks using 2017 Performance Data

Measured Entities: 49, 715

Average Performance Rate- 94.2%

Standard Deviation- 15.7%

Performance Deciles:

3: 95.67-99.08

4: 99.09-99.99

5: --

6: --

7: --

8: --

9: --

10: 100.00

2018:Preliminary AQI NACOR Performance Data for MIPS

# of Measured Entities: 8,465

Average Performance Rate: 97.08%

Standard Deviation: 15.75%

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

Not Applicable

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

Calendar Year 2016 AQI NACOR Cases by Age and Gender

Percentage of Total Cases by Patient Age:

<15 years: 6.3% (42.8% female, 57.2% male)

15-24 years: 6.4% (66.8% female, 33.2% male)

25-44 years: 21.4% (73.4% female, 26.6% male)

45-64 years: 33.8% (55.4% female, 44.6% male)

65-74 years: 19.2% (53.4% female, 46.6% male)

75+ years: 12.9% (54.2% female, 45.8% male)

Calendar Year 2017 AQI NACOR Cases by Age and Gender

Percentage of Total Cases by Patient Age:

<15 years: 5.6% (43.0% female, 57.0% male)

15-24 years: 5.9% (67.8% female, 32.2% male)

25-44 years: 21.2% (73.3% female, 26.7% male)

45-64 years: 33.8% (55.0% female, 45.0% male)

65-74 years: 20.1% (53.2% female, 46.8% male)

75+ years: 12.3% (53.8% female, 46.2% male)

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

Disparities data provided in 1b.4

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Cancer, Cancer : Bladder, Cancer : Breast, Cancer : Colorectal, Cancer : Gynecologic, Cancer : Hematologic, Cancer : Liver, Cancer : Lung, Esophageal, Cancer : Prostate, Cardiovascular, Cardiovascular : Congestive Heart Failure, Gastrointestinal (GI), Musculoskeletal : Falls and Traumatic Injury, Musculoskeletal : Joint Surgery, Neurology : Brain Injury, Respiratory, Surgery, Surgery : Cardiac Surgery, Surgery : General Surgery, Surgery : Perioperative and Anesthesia, Surgery : Thoracic Surgery, Surgery : Vascular Surgery

De.6. Non-Condition Specific*(check all the areas that apply):*

Safety, Safety : Complications, Safety : Healthcare Associated Infections

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Children, Elderly, Populations at Risk, Populations at Risk : Individuals with multiple chronic conditions, Populations at Risk : Veterans, Women

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2019_Measure_076_MIPSCQM.pdf

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

No data dictionary **Attachment:**

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

No

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

There have been no changes to the measure specifications since the previous measure update.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Patients for whom central venous catheter (CVC) was inserted with all elements of maximal sterile barrier technique*, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques** followed

Definitions:

*Maximal sterile barrier technique includes ALL of the following elements:

- cap
- mask
- sterile gown
- sterile gloves
- sterile full body drape

** Sterile ultrasound techniques require sterile gel and sterile probe covers

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Performance Met: CPT® II Code: 6030F- All elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques followed

Denominator Exception: CPT® II Code: 6030F-1P- Documentation of medical reason(s) for not following all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound

techniques during CVC insertion (including increased risk of harm to patient if adherence to aseptic technique would cause delay in CVC insertion).

Performance Not Met: CPT® II Code: 6030F-8P- All elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques not followed, reason not otherwise specified.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

All patients, regardless of age, who undergo CVC insertion

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Patient procedure during the performance period (CPT): 36555, 36556, 36557, 36558, 36560, 36561, 36563, 36565, 36566, 36568, 36569, 36570, 36571, 36572, 36573, 36578, 36580, 36581, 36582, 36583, 36584, 36585, 93503

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

None

The measure includes a denominator exception as indicated by reporting 6030F-1P for the numerator: Documentation of medical reason(s) for not following all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques during CVC insertion (including increased risk of harm to patient if adherence to aseptic technique would cause delay in CVC insertion)

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b.)*

NA

The measure includes denominator exception as indicated by reporting 6030F-1P for the numerator: Documentation of medical reason(s) for not following all elements of maximal sterile barrier technique, hand hygiene, skin preparation and, if ultrasound is used, sterile ultrasound techniques during CVC insertion (including increased risk of harm to patient if adherence to aseptic technique would cause delay in CVC insertion)

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b.)*

The measure is not stratified.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

1. Start with Denominator

2. Check Procedure Performed:

a. If Procedure as Listed in the Denominator equals No, do not include in Eligible Population. Stop

Processing.

b. If Procedure as Listed in the Denominator equals Yes, include in the Eligible Population.

3. Denominator Population:

a. Denominator Population is all Eligible Procedures in the Denominator.

4. Start Numerator

5. Check All Elements of Maximal Sterile Barrier Technique Followed:

a. If All Elements of Maximal Sterile Barrier Technique Followed equals Yes, include in Data Completeness Met and Performance Met.

b. If All Elements of Maximal Sterile Barrier Technique Followed equals No, proceed to check Documentation of Medical Reasons for All Elements of Maximal Sterile Barrier Technique Not Followed.

6. Check Documentation of Medical Reasons for All Elements of Maximal Sterile Barrier Technique Not Followed:

a. If Documentation of Medical Reasons for All Elements of Maximal Sterile Barrier Technique Not Followed equals Yes, include in Data Completeness Met and Denominator Exception.

b. If Documentation of Medical Reasons for All Elements of Maximal Sterile Barrier Technique Not Followed equals No, proceed to check All Elements of Maximal Sterile Barrier Technique Not Followed, Reason Not Otherwise Specified.

7. Check All Elements of Maximal Sterile Barrier Technique Not Followed, Reason Not Otherwise Specified:

a. If All Elements of Maximal Sterile Barrier Technique Not Followed, Reason Not Otherwise Specified equals Yes, include in the Data Completeness Met and Performance Not Met.

b. If All Elements of Maximal Sterile Barrier Technique Not Followed, Reason Not Otherwise Specified equals No, proceed to check Data Completeness Not Met.

8. Check Data Completeness Not Met:

a. If Data Completeness Not Met, the Quality Data Code or equivalent was not submitted.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

The measure is not based on a sample.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

The measure is not based on a survey.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Registry Data

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

Measure data was collected from the Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (NACOR).

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Clinician : Group/Practice, Clinician : Individual

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Inpatient/Hospital

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

The measure is not a composite performance measure

2. Validity – See attached Measure Testing Submission Form

2726_testing_attachment_2019.04.09_addition.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

No

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 2726

Measure Title: Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections

Date of Submission: 1/7/2019

Type of Measure:

☐ Outcome (*including PRO-PM*)	☐ Composite – ***STOP – use composite testing form***
☐Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.
- **For outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b5** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*including instructions, minimum font size 11 pt, do not change margins*)

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding how the testing results for this measure meet NQF's evaluation criteria for testing. ***Contact NQF staff if more codes are needed.***

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **instrument-based measures** (including PRO-PMs) **and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b3. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to

conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

15. With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐claims	☐claims
☒registry	☒registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

Reliability testing was performed using data from the Medicare Limited Data Set Carrier SAF – 5% File and data from the Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (NACOR). Validity testing was performed on data from the Centers for Medicare & Medicaid Services (CMS) Physician Quality Reporting System and the Centers for Disease Control and Prevention (CDC)'s National Healthcare Safety Network (NHSN).

1.3. What are the dates of the data used in testing? Medicare Limited Data Set Carrier SAF-5% File (2012-2013); AQI NACOR (2012-2014); CMS PQRS (2013-2016); CDC NHSN (2013-2016)

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	**Measure Tested at Level of:**
☒individual clinician	☒individual clinician
☒group/practice	☒group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency

☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (2012-2014)

All cases in NACOR Public Use File (PUF) Q4 2014 where known providers are MD/DO or CRNAs[1]				
	2012	**2013**	**2014**	**Total**
Number of Practices	70	89	112	127
Number of Providers	1,886	2,547	3,149	4,775
Number of Cases	10,595	11,859	12,357	34,811

Medicare Limited Data Set Carrier SAF – 5% File

Medicare Limited Data Set Carrier SAF – 5% File		
	2012	**2013**
Cases	3,476	3,669

CMS PQRS Data (2016)

2016 CMS PQRS Experience Report	
	2016
Individuals	35, 405
Groups	468

CDC NHSN Data

CDC NHSN Healthcare-Associated Infection Data for Acute Care Hospitals	

[1] Totals represent unique practices, facilities and providers over the three year period and not simply the sum of individual years.

	Hospitals
2016	2,345
2015	2,328
2014	2,442
2013	2,389

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Patient Age (AQI NACOR: Medicare PQRS #76: Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections)			
Age Group	**2012 (%)**	**2013 (%)**	**2014 (%)**
< 1	0 (0)	0 (0)	0 (0)
1 – 18	5 (0.05)	9 (0.08)	9 (0.07)
19 – 49	425 (4.01)	605 (5.10)	582 (4.71)
50 – 64	1,063 (10.03)	1,489 (12.56)	1,543 (12.49)
65 – 79	4,893 (46.18)	6,960 (58.69)	7,471 (60.46)
80+	1,774 (16.74)	2,538 (21.40)	2,695 (21.81)
Not Reported	2,435 (22.98)	258 (2.18)	57 (0.46)

Patient Sex (AQI NACOR: Prevention of Central Venous Catheter (CVC) - Related Bloodstream Infections)			
Sex	**2012 (%)**	**2013 (%)**	**2014 (%)**
Female	3,465 (32.70)	4,852 (40.91)	5,146 (41.64)
Male	4,624 (43.64)	6,618 (55.81)	7,103 (57.48)
Not Reported	2,506 (23.65)	389 (3.28)	108 (0.87)

ASA Physical Status (AQI NACOR: Prevention of Central Venous Catheter (CVC) - Related Bloodstream Infections)			
ASA Physical Status	**2012 (%)**	**2013 (%)**	**2014 (%)**
I – II	4,733 (44.67)	5,292 (44.62)	4,888 (39.56)

III	1,832 (17.29)	1,796 (15.15)	2,053 (16.61)
IV	3,222 (30.41)	4,021 (33.91)	4,665 (37.75)
V	63 (0.60)	89 (0.75)	124 (1.00)
Not Reported	745 (7.03)	661 (5.57)	627 (5.07)

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

Data used to for reliability testing include the Medicare Limited Data Set Carrier SAF – 5% File and data from AQI NACOR. Data for validity testing includes data from the 2016 PQRS Experience Report and the CDC NHSN.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

Patient-level sociodemographic data variables were not available or analyzed in the data sample used.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Reliability, measured from 0 to 1, can be characterized as the proportion of observed physician variation that can be explained by "true" differences in the quality measure. A reliability of 0 means that all of the variance in the outcomes is the result of measurement error, and a reliability of 1 means that all of the variance in outcomes is from true differences in performance.

We estimated the reliability based on the beta-binomial model[2], using the following formula: reliability = signal/(signal + noise). This reliability estimate assumes the beta distribution for the "true" physician scores.[1] Although less common than normal hierarchical linear models, the beta-binomial model is a more natural fit for estimating the reliability of simple pass/fail rate measures (e.g., HEDIS measures). There are also computational advantages to using the beta-binomial model. The beta distribution is usually defined by two parameters, alpha and beta. The mean and variance of the distribution are:

$$\mu = \frac{\alpha}{(\alpha+\beta)}$$

$$\sigma^2_{provider-to-provider} = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

The usual binomial variance for the error where p is the provider specific probability of passing the indicator is:

$$\sigma^2_{error} = \frac{\hat{p}(1-\hat{p})}{n}$$

Where $\hat{p}$ is the observed pass rate for the provider. Then, the standard formula for reliability is applied. Reliability estimates based on the beta-binomial distribution were developed using SAS 9.3 (Carey, NC).

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

Reliability was tested on providers and practices where at least 2 cases were reported.

Anesthesia Quality Institute – National Anesthesia Clinical Outcomes Registry			
	2012	**2013**	**2014**
Number of Providers	2,417	2,991	3,812
Reliability (Mean of Providers)	0.969	0.968	0.967

Medicare SAF 5% File		
	2012	**2013**
Sample Size	1,309	1,384

[2] Adams, J. (2009). The reliability of provider profiling: a tutorial. Santa Monica, CA: RAND Health.

Average Provider Reliability (n>=2)	0.946	0.952

Anesthesia Quality Institute – National Anesthesia Clinical Outcomes Registry			
	2012	**2013**	**2014**
Number of Practices	65	88	108
Reliability (Mean of Practices)	0.971	0.983	0.964

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Reliability scores vary from 0.0 to 1.0, with a score of zero indicating that all variation is attributable to measurement error (noise, or variation across patients within providers) where as a reliability of 1.0 implies that all variation is caused by real difference in performance across accountable entities.

Reliability for the process Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections measure is consistently greater than 0.9, and thus can be considered to be very good. This reflects the inclusion of that measure in public reporting programs, the number of years that the measure has been reported and the number of cases available to test and analyze. In the three years of NACOR data analyzed, reliability has remained stable and consistent.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒Performance measure score

☒**Empirical validity testing**

☒**Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

To assess the relationship between performance on this process measure related to use of sterile technique for central venous catheter insertion and the incidence of subsequent infection, we compared the average performance rates on this measure in the PQRS program to nationally-reported central-line associated bloodstream infection (CLABSI) standardized infection ratios (SIR) for the same time period.

Face validity of the measure score as an indicator of quality was systematically assessed as follows. After the measure was fully specified, a group of experts was assembled to rate face validity. The experts included 19 physicians.

We provided the detailed measure specifications to the experts and asked them to rate their agreement with the following statement: The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good from poor quality.

The rating scale had five levels (1-5) with the following narrative anchors:

1 = Disagree; 3 = Moderate Agreement; 5 = Agree

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

	MIPS 076 Avg. Performance Rate	**NHSN National CLABSI SIR**
2013	82.8%	0.53
2014	84.5%	0.49
2015	88.3%	0.60*
2016	91.2%	0.56*

*In 2015, the CDC changed their surveillance protocol, resulting in an increase in the reported SIR. The subsequent decline in infections from 2015 to 2016, combined with the downward trend in infections prior to 2015, support a likelihood that the increase from 2014-2015 is solely due to the surveillance protocol change. More information is available at: https://www.cdc.gov/hai/data/archive/data-summary-assessing-progress.html

To assess the face validity of this measure, it was examined by a group of experts. Out of the 19 participants, 17 agreed that the scores from the measure as specified would provide an accurate reflection of quality and 2 disagreed.

Mean rating = (4.16 out of 5)

Rating Scale	**Number Who Selected this Rating**
1 – Strongly Disagree	1
2 – Disagree	1
3 – Neither	0
4 – Agree	9
5 – Strongly Agree	8

Total	19

2b1.4. What is your interpretation of the results in terms of demonstrating validity? (i*.e., what do the results mean and what are the norms for the test conducted?*)

As performance on this measure has improved over time from 82.8% in 2013 to 91.2% in 2016, rates of central line infections in hospitals have continued to decrease. While this data sample is too small to conduct a statistical estimate of correlation, the available data support that improvements in sterile technique for central line insertion may play a role in national reductions in CLABSIs. In their analysis of progress related to healthcare-associated infections (available at: https://www.cdc.gov/hai/data/archive/data-summary-assessing-progress.html) , the CDC states "dynamics in the microbiology of CLABSIs, along with lagging declines on wards where central lines are on average in place longer than ICU central lines, suggest prevention successes due primarily to improved central line insertion practices". Because the CDC changed their CLABSI surveillance protocol between 2014 and 2015, we are unable to assess longitudinal correlation between measure performance and national CLABSI rates from 2013-2016, the years for which we have available data. Instead, we would need to split the data into two sets, each including a single surveillance methodology: 2013-2014 and 2015-2016. Because each data set would only include 2 years, this is not sufficient data to establish a statistical trend or correlation over time. In their analysis of longitudinal CLABSI trends, the CDC concluded that the increase from 2014-2015 is most likely solely due to the surveillance protocol change, as supported by the subsequent decrease in CLABSI rates from 2015-2016.

The results of the assessment of face validity indicate that an independent group of experts (different from those who advised on measure development) had high levels of agreement with the statement: "The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality."

2b2. EXCLUSIONS ANALYSIS

NA ☒no exclusions — *skip to section 2b3*

2b2.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

2b2.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased*

data collection and analysis. *Note*: ***If patient preference is an exclusion****, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b4.

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

Also discuss any "ordering" of risk factor inclusion; for example, are social risk factors added after all clinical factors?

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

- ☐ **Published literature**
- ☐ **Internal data analysis**
- ☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

2b3.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b3.9. Results of Risk Stratification Analysis:

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i*.e., what do the results mean and what are the norms for the test conducted*)

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

To analyze the identification of statistically significant and meaningful differences in performance, we analyzed the distribution of performance scores into deciles and calculated the mean performance score and standard deviation. These measures provide an estimate of the variation in provider performance, and thus the ability to distinguish high performers from low performers.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

2018 MIPS Performance Benchmark data based on 2016 PQRS data for this measure was distributed as follows:

Average Performance: 93.3%
Standard Deviation: 15.6%

Decile 3: 91.85-96.14
Decile 4: 96.15-98.88
Decile 5: 98.05-99.99
Decile 6: 100
Decile 7: 100
Decile 8: 100
Decile 9: 100
Decile 10: 100

2b4.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

While overall performance is rather high for this measure, the moderate standard deviation of 15.6% indicates a moderate amount of variation in performance on the measure and ability to identify those differences in care.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without social risk factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b5.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing

data (or differences between responders and nonresponders) and how the specified handling missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

NACOR requires participants to submit all of the measure's data elements in order for a performance score to be calculated and reported. As a result, there is no missing data in this measure's dataset.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Because all of the measure's data elements are required to be reported to NACOR as a pre-requisite to measure calculation, missing data does not factor into a provider's performance rate.

2b6.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

NACOR prevents bias in measure performance due to systematic missing data by requiring participants to submit all included measure data elements as a pre-requisite for measure calculation and reporting. Registry staff proactively work with providers to address data completeness issues and ensure that providers can submit all needed measure data prior to measure calculation and submission.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score), Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

Some data elements are in defined fields in electronic sources

3b.2. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

Electronic documentation of compliance with the maximal sterile barrier technique is not universally available in EHRs. Instead, the use of CPT II Codes to capture the process of care is needed in the interim to collect this measure. For those providers who use EHRs, components of the maximal barrier precautions for CVC placement are regularly included in those electronic systems. Therefore, ASA believes that this measure will easily transition into electronic health record capture in the future.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Some anesthesia billing vendors who work with AQI NACOR members to report this measure reported some difficulty in obtaining the needed surgical data elements to report this measure. ASA and AQI staff conducted a range of educational activities as well as working one on one with practices to ensure they are able to appropriately collect and report the needed data elements. As a result of these activities, users no longer report difficulty in collecting and reporting this measure.

3c.2. Describe any fees, licensing, or other requirements to use any aspect of the measure as specified *(e.g., value/code set, risk model, programming code, algorithm)*.

Limited proprietary coding is contained in the measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The denominator of this measure does include the use of Current Procedural Terminology (CPT®), which is copyrighted by the American Medical Association (AMA). Use of CPT in Measure(s) is limited to Non-Commercial Use. Any commercial use of CPT beyond fair use requires a license from the AMA.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
	Public Reporting www.cms.gov/pqrs Physician Quality Reporting System Payment Program Merit-based Incentive Payment System https://qpp.cms.gov Merit-based Incentive Payment System https://qpp.cms.gov Regulatory and Accreditation Programs www.jointcommission.org Joint Commission Quality Improvement (external benchmarking to organizations) Anesthesia Quality Institute National Anesthesia Clinical Outcomes Registry (AQI NACOR) www.aqihq.org

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

The Merit-based Incentive Payment System (MIPS) is sponsored by the Centers for Medicare & Medicaid Services (CMS)

- The purpose of this program is to reward value and outcomes in healthcare. This program was designed to tie payments to quality and cost efficient care, drive improvement in care processes and health outcomes, increase the use of healthcare information, and reduce the cost of care.
- The MIPS program is a national program that applies to eligible clinicians across the entire United States who provide Medicare-covered services.

The National Anesthesia Clinical Outcomes Registry (NACOR) is sponsored by the Anesthesia Quality Institute (AQI)

- The purpose of NACOR is to be the primary source of information for quality improvement in the clinical practice of anesthesia. Through education and quality feedback, AQI helps to improve the quality care of patients, lower anesthesia mortality, and lower anesthesia incidents.
- AQI NACOR includes participants from throughout the United States and U.S. territories. In 2017, more than 25,000 anesthesia providers reported data to NACOR.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

The measure is currently in an accountability/payment program.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

The measure is currently in an accountability/payment program.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

AQI NACOR makes data available to those being measured through the use of a measure data dashboard that is continuously accessible and reflects real-time data. AQI also assigns each participating practice an account manager to help a practice correctly capture and report data, as well as to correctly understand and interpret their data reports and to troubleshoot any issues with the data. AQI also hosts monthly office hours with users where they address relevant educational topics and address questions from users. In addition, AQI makes available various educational resources and webinars on their websites to support users in appropriately collecting and reporting data, as well as reviewing and interpreting data reports.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

Measure data is available to users within 48 hours of submission to NACOR via an online dashboard. The dashboard includes information on performance and reporting rates, which can be viewed down to a individual provider/case level. Aggregate group-level reports are also available for users. Dedicated account managers then work with practices to trouble shoot any data reporting issues and to help them interpret their data reports.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

Stakeholder feedback, including feedback from those being measured, is obtained through a number of avenues including the submission of registry help-desk tickets, direct feedback to their account manager or other AQI staff via phone or email, and through participation in monthly office hours sessions.

4a2.2.2. Summarize the feedback obtained from those being measured.

In early implementation of this measure, some entities being measured expressed confusion on whether or not a provider who provides anesthesia for central line insertion is expected to report this measure. As a result ASA and AQI provided a number of educational offerings to clarify reporting expectations for this measure.

4a2.2.3. Summarize the feedback obtained from other users

Some anesthesia billing vendors who partner with members to report this measure to NACOR have expressed difficulty in capturing surgical billing information.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

To clarify questions on who is expected to report this measure, we have included explicit instructions in the specifications that it should be reported by ECs who perform CVC insertion. Additionally, ASA and AQI have conducted extensive user and vendor educational activities to clarify how the measure data should be collected and reported. Dedicated account managers have worked with individual NACOR members to address difficulties in reporting the required surgical billing data. As a result of these educational and outreach activities, reporting of the measure has increased and user feedback on the measure is positive.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included.)

If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

Performance has shown steady improvement over time, increasing from an average performance rate of 93.3% in 2016 to an average performance rate of 97.08% in 2018. As the measure has continued to be included in the MIPS program, user uptake of the measure has continued to increase with nearly 50,000 providers reporting the measure to CMS in 2017.

4b2. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

Documentation of compliance with maximum barrier precautions when placing a CVC has not led to any unintended consequences. Compliance with the recommended standards has a strong association with improved patient outcomes, and this fact is generally well-recognized in anesthesia practice. The increased time required to comply, and the increased use of resources such as drapes, gloves, and prep solution, are trivial compared to the benefits of preventing even a single CVC-associated infection. The time required to complete the documentation itself is trivial, is supported by existing anesthesia records (both paper and electronic) and will become more so as Anesthesia Information Management Systems continue to advance.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

As described in the testing attachment, the primary benefit of this measure, while not unexpected, is a decrease in national estimates of central line-associated blood stream infections over time.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0139 : National Healthcare Safety Network (NHSN) Central line-associated Bloodstream Infection (CLABSI) Outcome Measure

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

The measure is specified for a level of analysis that includes the individual practitioner with the intent of providing data to clinicians and other health professionals regarding their individual performance.

Similar measures exist including the Centers for Disease Control and Prevention's Central line-associated Bloodstream Infection measure (NQF measure 0139). This measure is specified and NQF endorsed for analysis at the facility level. That measure, although closely associated with and may touch upon this process measure, is an outcome measure. Although ASA welcomes a conversation on harmonization, we do not believe that this measure conflicts or competes with these measures.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

No

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

The measure is specified for a level of analysis that includes the individual practitioner with the intent of providing data to clinicians and other health professionals regarding their individual performance. Similar measures exist including the Centers for Disease Control and Prevention's Central line-associated Bloodstream Infection measure (NQF measure 0139). This measure is specified and NQF endorsed for analysis at the facility level. That measure, although closely associated with and may touch upon this process measure, is an outcome measure. Although ASA welcomes a conversation on harmonization, we do not believe that this measure conflicts or competes with these measures.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

The measure does not compete with NQF #0139.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

No appendix **Attachment:**

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): American Society of Anesthesiologists

Co.2 Point of Contact: Toni, Kaye, t.kaye@asahq.org, 847-268-9160-

Co.3 Measure Developer if different from Measure Steward: American Society of Anesthesiologists

Co.4 Point of Contact: Matthew, Popovich, qra@asahq.org, 202-289-2222-316

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

The AMA-PCPI-convened the Anesthesiology and Critical Care Workgroup developed the submitted measure. PCPI measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and other health care professional disciplines participating in patient care for the clinical condition or topic under study are invited to participate as equal contributors to the measure development process. In addition, the PCPI strives to include on its work groups individuals representing the perspectives of patients, consumers, private health plans, and employers. This broad-based approach to measure development ensures buy-in on the measures from all stakeholders and minimizes bias toward any individual specialty or stakeholder group. All work groups have at least two co-chairs who have relevant clinical and/or measure development expertise and who are responsible for ensuring that consensus is achieved and that all perspectives are voiced.

The AMA-PCPI Anesthesiology and Critical Care Workgroup consisted of the following experts involved in measure development:

Alexander A. Hannenberg, MD, Co-chair – American Society of Anesthesiologists

Andrew J. Patterson, MD, PhD, Co-chair – American Board of Anesthesiology

William R. Andrews, MD, MS – American College of Chest Physicians

Rebecca A. Aslakson, MD, PhD – American Academy of Hospice and Palliative Medicine

Daniel R. Brown, MD, PhD – Mayo Clinic

Neal H. Cohen, MD, MPH, MS – American Society of Anesthesiologists

Peggy Duke, MD – American Society of Anesthesiologists

Heidi L. Frankel, MD – American College of Surgeons

Lorraine M. Jordan, BSN, MS, PhD – American Association of Nurse Anesthetists

Jeremy M. Kahn, MD, MS – American Thoracic Society

Jason N. Katz, MD, MHS – American College of Cardiology

Gerald A. Maccioli, MD – American Society of Anesthesiologists

Catherine L. Scholl, MD – Texas Medical Association

Todd L. Slesinger, MD – American College of Emergency Physicians

Victoria M. Steelman, PhD, RN – Association of Perioperative Registered Nurses

Avery Tung, MD – Society of Critical Care Medicine

In preparation for measure submission, the ASA also convened a Measure Expert Panel (MEP) who reviewed the measure specifications and were asked to rate their agreement with the following statement: "The scores

obtained from the measure as specified will proved an accurate reflection of quality and can be used to distinguish good and poor quality." The results were displayed in 2b2.3.

John P. Abenstein, M.D., M.S.E.E.

Associate Professor of Anesthesiology

Mayo Clinic

Rochester, MN

Brian Cammarata, M.D.

Old Pueblo Anesthesia/Tucson Medical Center

Tucson, AZ

Christopher John Curatolo, M.D., M.E.M.

Resident Physician, Anesthesiology

The Mount Sinai Medical Center

New York, NY

Peggy G. Duke, M.D.

Emeritus Professor

Emory Healthcare

Atlanta, GA

Brenda G. Fahy, M.D.

Professor

University of Florida

Gainesville, FL

Chris Giordano, M.D.

Division Chief of Transplant

University of Florida

Gainesville, FL

Alexander A. Hannenberg, M.D.

Associate Chair of Anesthesiology

Newton-Wellesley Hospital

Newton, MA

Aaron Martin Joffe, D.O.

Associate Professor

University of Washington, Department of Anesthesiology and Pain Medicine

Seattle, WA

Meghan B. Lane-Fall, M.D., MSc

Assistant Professor

University of Pennsylvania

Philadelphia, PA

David P. Martin, M.D., Ph.D.

Associate Professor of Anesthesiology

Mayo Clinic
Rochester, MN
Dolores B. Njoku, M.D.
Associate Professor, ACCM, Pediatrics, Pathology
Johns Hopkins University
Baltimore, MD
Andrew Jay Patterson, M.D., Ph.D.
Associate Professor
Stanford University
Stanford, CA
Marjorie Podraza Stiegler, M.D.
Associate Professor
Consortium of Anesthesia Patient Safety and Experiential Learning
UNC – Chapel Hill
Chapel Hill, NC
Erin A. Sullivan, M.D.
Chief, Division of Cardiothoracic Anesthesiology
University of Pittsburgh Medical Center
Pittsburgh, PA
Steven L. Sween, M.D.
Chairman, Medical Director
Emory Saint Joseph's Hospital
Atlanta, GA
Richard D. Urman, M.D., M.B.A.
Associate Professor of Anesthesia
Brigham and Women's Hospital
Boston, MA
Cassie D. Volker, M.D.
Mobile Anesthesia Care
Overland Park, KS
Toby N. Weingarten, M.D.
Associate Professor of Anesthesiology
Mayo Clinic
Rochester, MN
Jeffrey D. White, M.D.
Assistant Clinical Professor
University of Florida
Gainesville, FL

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2008

Ad.3 Month and Year of most recent revision: 04, 2018

Ad.4 What is your frequency for review/update of this measure? Annually

Ad.5 When is the next scheduled review/update for this measure? 04, 2020

Ad.6 Copyright statement: COPYRIGHT:

The Measures are not clinical guidelines, do not establish a standard of medical care, and have not been tested for

all potential applications.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial

purposes, eg, use by health care providers in connection with their practices. Commercial use is defined as the sale,

license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or

service that is sold, licensed or distributed for commercial gain.

Commercial uses of the Measures require a license agreement between the user and the PCPI® Foundation

(PCPI®) or ASA. Neither ASA, nor the American Medical Association (AMA), nor the AMA-convened Physician

Consortium for Performance Improvement® (AMA-PCPI), now known as the PCPI, nor their members shall be

responsible for any use of the Measures.

The AMA's and AMA-PCPI's significant past efforts and contributions to the development and updating of

the Measures is acknowledged. ASA is solely responsible for the review and enhancement ("Maintenance")

of the Measures as of May 15, 2014.

ASA encourages use of the Measures by other health care professionals, where appropriate.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

© 2017 PCPI® Foundation and American Society of Anesthesiologists. All Rights Reserved.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code

sets should obtain all necessary licenses from the owners of these code sets. ASA, the AMA, the PCPI and its

members and former members of the AMA-PCPI disclaim all liability for use or accuracy of any Current Procedural

Terminology (CPT®) or other coding contained in the specifications.

CPT® contained in the Measures specifications is copyright 2004-2018 American Medical Association. LOINC®

copyright 2004-2018 Regenstrief Institute, Inc. SNOMED CLINICAL TERMS (SNOMED CT®) copyright 2004-2018

The International Health Terminology Standards Development Organisation (IHTSDO). ICD-10 is copyright 2018

World Health Organization. All Rights Reserved.

Ad.7 Disclaimers: The Measures are not clinical guidelines, do not establish a standard of medical care, and have not been tested for all potential applications.

Ad.8 Additional Information/Comments: Make sure to cc Matt Popovich alternative email regarding correspondence of this measure (qra@asahq.org)