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 #: 0139

De.2. Measure Title: National Healthcare Safety Network (NHSN) Central line-associated Bloodstream Infection (CLABSI) Outcome Measure

Co.1.1. Measure Steward: Centers for Disease Control and Prevention

De.3. Brief Description of Measure: Standardized Infection Ratio (SIR) and Adjusted Ranking Metric (ARM) of healthcare-associated, central line-associated bloodstream infections (CLABSI) will be calculated among patients in bedded inpatient care locations.

This includes acute care general hospitals, long-term acute care hospitals, rehabilitation hospitals, oncology hospitals, and behavioral health hospitals.

1b.1. Developer Rationale: It is envisioned the use of this measure will promote CLABSI prevention activities which will lead to improved patient outcomes including reduction of avoidable medical costs, and patient morbidity and mortality.

S.4. Numerator Statement: Total number of observed healthcare-associated CLABSI among patients in bedded inpatient care locations.

S.6. Denominator Statement: Total number of predicted healthcare-associated CLABSI among patients in bedded inpatient care locations, calculated using the facility's number of central line days and the following significant risk factors:

- Acute Care Hospitals: CDC location, facility bed size, medical school affiliation, facility type, birthweight category (NICU locations only)
- Critical Access Hospitals: no significant risk factors, calculation based intercept only model
- Inpatient Rehabilitation Facilities: Proportion of admissions with stroke, proportion of admissions in other non-specific diagnostic categories
- Long Term Acute Care Hospitals: CDC location type , facility bed size, average length of stay, proportion of admissions on a ventilator, proportion of admissions on hemodialysis

S.8. Denominator Exclusions: Data from patients who are not assigned to an inpatient bed are excluded from the denominator counts, including outpatient clinics, 24-hour observation units, and emergency department visits. Inpatient rehab locations and inpatient psychiatric locations that have their own Centers for Medicare and Medicaid Services (CMS) Certification Number (CCN) are excluded.

De.1. Measure Type: Outcome

S.17. Data Source: Electronic Health Data, Electronic Health Records, Other, Paper Medical Records

S.20. Level of Analysis: Facility, Population : Regional and State

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 health outcome measure include providing empirical data that demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service; if these data not available, data demonstrating wide variation in performance, assuming the data are from a robust number of providers and results are not subject to systematic bias. For measures derived from patient report, evidence also should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

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:**

Updates:

- The developer identifies a number of prevention activities that can reduce the incidence of CLABSI. These include:
    - Appropriate central line use: promptly removing non-essential intravascular catheters,
    - Hand hygiene and aseptic technique
    - The use of maximal barrier equipment including a large patient drape, inserter mask, sterile gloves, cap, and sterile gown during aseptic insertion of the central line
    - Appropriate insertion site decontamination before central line insertion
    - Chlorhexidine-impregnated dressings (in patients ≥ 18 years), and (vi) implementing surveillance strategies
- To support these practices, the developer cites a guideline:
    - O'Grady NP, Alexander M, Burns LA, Dellinger PE, Garland J, et al. Guidelines for the prevention of intravascular catheter-related infections. Available at http://www.cdc.gov/hicpac/pdf/guidelines/bsi-guidelines-2011.pdf. .

Question for the Committee:

- *The evidence provided by the developer is updated, directionally the same, and stronger compared to that for the previous NQF review. Does the Committee agree there is no need for repeat discussion and vote on Evidence?*

Guidance from the Evidence Algorithm

Health outcome [Box 1] → Relationship between the measured health outcome and at least one healthcare action is demonstrated by empirical data [Box 2] → Pass

Preliminary rating for evidence: ☒ Pass ☐ No Pass

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.

- The developer provides national Standardized Infection Ratios (SIRs) for CAUTI in 2015, 2016, and 2017:
  - National CLABSI SIR in 2015 is 0.994 = 26,029 observed / 26,183.537 predicted
  - National CLABSI SIR in 2016 is 0.891 = 23,591 observed / 26,472.710 predicted
  - National CLABSI SIR in 2017 is 0.814 = 21,173 observed / 25,993.180 predicted
- The developer also reports that there was a 10% decrease in CLABSI between 2015 and 2016, and a 9% decrease between 2016 and 2017.

Disparities

- The developer reports that, among patients hospitalized with acute cardiovascular disease, pneumonia, and major surgery, Asian and Hispanic patients had significantly higher rates of HAIs than white, non-Hispanic patients.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?

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

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

1a. Evidence

Comments:

**no need for repeat discussion

**Clear links between measure, processes, and outcome. Ongoing data exist on impact and paths to address outcme.

**No need for repeat discussion

**updated evidence

**Not aware of any new studies that changes the evidence base.

1b. Performance Gap

Comments:

**Apparently, there was subgroup disparity, but there was no link to the data.

**Gaps and opportunitie exist despite improvments, and gaps are assymetric across subpopulations. There are opportunities to address disparities in outcomes.

**Moderate gap

**demonstrated gap still exists

**Yes it demonstrates a gap in care. For Disparities, there is scant evidence provided and that evidence is meaningful but rather old. It would be ideal for the field to explore this further as recently there is evidence to suggest that a bias in voluntary reporting exists. Would the same types of bias apply to management of a central lines?

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:

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

Evaluators: NQF Scientific Methods Panel Subgroup

Methods Panel Review (Combined)

Evaluation of Reliability and Validity:

Scientific Methods Panel Votes: Measure passes

- Reliability: H-0, M-4, L-0, I-0
- Validity: H-0, M-3, L-1, I-0

This measure was reviewed by the Scientific Methods Panel and discussed on their call. A summary of the measure is provided below:

Reliability

- Reliability testing was performed at the data element level.

- Data Element
  - Data element validity testing was conducted, which NQF allows to serve as a demonstration of data element reliability.
  - The methods and results of data element validity testing are described in the validity section below.

Validity

- Validity testing was performed at the data element level.
- Data Element
  - The developer notes that the critical data elements of this measure have been validated by a number of state health departments that require mandatory reporting of CLABSI through the NHSN.
  - Data validation is conducted by trained auditors, who review medical records and determine whether facilities' identification of patients meeting or not meeting CLABSI criteria was accurate.
  - Sensitivity, specificity, positive predicted value, and negative predicted value are calculated.
  - Validation results from 5 states are provided—the developers report that these validations indicated a pooled mean sensitivity of 87.5% (range: 80.3%-100%), specificity of 99.3% (range: 98.7% - 100%), positive predictive value of 96.9% (range: 94.2% - 100%) and negative predictive value of 96.9% (range: 93.7% - 100%).
  - One concern by a Methods Panel member was that no statistical results (e.g., c-statistic) of risk adjustment model power were reported. In addition, one Methods Panel member felt risk adjustment should be at score level and not the provider level and thus was not satisfied with approach.

Standing Committee Action Item(s): The Standing Committee can discuss reliability and/or validity or accept the Scientific Methods Panel ratings.

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 Scientific Methods Panel 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:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The Scientific Methods Panel is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

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

Methods Panel Evaluation (Combined): Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0139

Measure Title: National Healthcare Safety Network (NHSN) Central Line -Associated Bloodstream Infection (CLABSI) Outcome Measure

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 Methods Panel member 3**:(additional NHSN forms)

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other Methods Panel member 2:**"Other" is specified in the application, but I don't know if they really intend this to be "Integrated Delivery System"

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.**

   Methods Panel member 1: None

   Methods Panel member 2: None.

   Methods Panel member 4: 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**

   Methods Panel member 1: Per 2a2.1

4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☒ **No**

   Methods Panel member 1: States that reliability testing was conducted (2.1), but I could not find any results of this testing. Section 2a2.2 states to look at section 2b1 for validity testing; 2a2.3 references 2b1.3 for reliability testing; 2a2.4 references 2b1.4 for reliability testing at either the score or date element levels. These referenced sections do provide reliability testing results for data elements but not for performance measure.

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

Methods Panel member 1: Disorganization how reliability information was presented could suggest a misunderstanding of the difference between reliability and validity as assessed at the date element and performance score level.

Methods Panel member 2: Data element reliability testing is described under validity (sections 2b1.2-4). Basically, state state health departments review medical records and look at the concordance with CLABSIinfections documented in the statistics used in the measure. External validation across 5 states consisted of 2,594 chart reviews

Methods Panel member 4: Note that MD states in section sa2.2 that "As per NQF email "…data element validity testing may serve as a demonstration of data element reliability." Please see section 2b1.2 for demonstration of data element reliability.

Methods Panel member 3: Used data collected externally by 5 states to assess the reliability of the data elements.

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

   Submission document: Testing attachment, section 2a2.3

Methods Panel member 2: These validations indicated a pooled mean sensitivity of 87.5% (range: 80.3%-100%), specificity of 99.3% (range: 98.7% - 100%), positive predictive value of 96.9% (range: 94.2% - 100%) and negative predictive value of 96.9% (range: 93.7% - 100%).

Methods Panel member 1: Source data are for all 50 states. However, only data from 5 states were presented. Why? Results presented in section 2b1.2 appear to be only at the Data Element and not the Performance Measure level for state health departments. The measure level specification (section 1.4) states "hospital/facility/agency" level. Additionally, Positive and Negative Predicitve Value operational definitions should be specified.

Methods Panel member 4: Note that MD states in section sa2.2 that "As per NQF email "…data element validity testing may serve as a demonstration of data element reliability." Please see section 2b1.2 for demonstration of data element reliability.

Methods Panel member 3: The sensitivity, specificity, PPV, and NPV were all high for the data collected in the 5 states.

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)

    Methods Panel member 4: Data element validity testing was only performed for the outcome variable. However, the factors used in the risk adjustment model are all facility-level variables. Hence, it is reasonable to assume that these variables are valid.

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Methods Panel member 2: The application refers to an NQF email stating "…data element validity testing may serve as a demonstration of data element reliability." The methods described were appropriate and the results good, but since score-level testing was not conducted, Moderate is the highest eligible rating.

    Methods Panel member 1: While there is some evidence at the Date Element level, the analysis is performed at a different level (state unit) than is proposed for the measure (facility). Similarly, there is no evidence of performance measure reliability provided by the developer. One could make the case that the overall rating. should be Insufficient.
    Methods Panel member 4: Note that MD sates in section sa2.2 that "As per NQF email "…data element validity testing may serve as a demonstration of data element reliability." Please see section 2b1.2 for demonstration of data element reliability

    Methods Panel member 5: Outcome measure up for maintenance; note reliability testing for elements conducted and to see validity section; however, element validity testing and score face validity (not empirical validity) purportedly conducted.

    Methods Panel member 3: Did not conduct score-level testing; data element testing in 5 states was robust and produced good results.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    Methods Panel member 2: None.

    Methods Panel member 1: This section is labeled as "NA".

    Methods Panel member 4: There are no exclusions.

    Methods Panel member 5:N/A

    Methods Panel member 3: Not applicable.

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

    Submission document: Testing attachment, section 2b4.

    Methods Panel member 1: There were no results presented for the "CLABSI SIR" performance measure. There are statements such as "In some places where large scale CLABSI prevention programs have been implemented over the past several years, significant reductions in the CLABSI SIR have been seen, reflecting better quality. However, there are still facilities with significantly high CLABSI SIRs, indicating that they have

not made progress in reducing CLABSI (high SIRs indicate poor quality)." (section 2b1.4) with no numerical results presented to support this claim.

Methods Panel member 2: None.

Methods Panel member 4: None. 1b.3, CLABSI infection rate means vary from a low of 0.0% per 1000 device days to a high of 3.4% per 1000 device days between all types of reporting locations.

Methods Panel member 5: Note that "~7.21%" of faciltities may have an opportunity for improvement, which means ~92% don't; while CLABSI should never happen, there is not provided rationale of how meaningful this remaining difference is to performance. What impact is this gap having? How does it relate to the effort of this measure? Is this measure valuable enough to close this remaining gap?

Methods Panel member 3: No concerns. See variation in performance across facilities (14.88% were stat sig less than 1.0; 7.21% were stat sig bettern than 1.0)

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.

Methods Panel member 2: None.

Methods Panel member 1: No empirical results submitted.

Methods Panel member 5:N/A

Methods Panel member 3: Not applicable.

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

Submission document: Testing attachment, section 2b6.

Methods Panel member 2: None.

Methods Panel member 1: No empirical results submitted.

Methods Panel member 4: None

Methods Panel member 5: None

Methods Panel member 3: There are no mssing data, as facilities have to submit full numerators and denominators.

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

Methods Panel member 4:MD states, without justification, that there is a "paucity" of evidence to support social risk factors inclusion. Since patient-level data is not collected (other than the outcome), it would be difficult for MD to test whether inclusion of social risk factors has an impact.

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 **Methods Panel member 1:** No statistical results (e.g., c-statistic) of model power were reported

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

Methods Panel member 1: Cannot determine without model results statistics.

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

Methods Panel member 2: The SIR approach is well described and tested in the CDC document referenced.

Methods Panel member 1: Model development process is probably sound. Quality of model to predict results is not possible to evaluate given no empirical information on this question.

Methods Panel member 4: The risk modeling was conducted using negative binomial regression The model adjusts for patient care location type, medical school affiliation, facility bedsize, and central line device utilization. The MD states that patient-level data is not collected in order to minimize data collection burden. Facility peformance is evaluated using Standardized Infection Ratio (SIR) – which is analogous to OE ratio.

The MD also estimates the same model using a hierarchical framework in which facilities are specified as a random effect. Facility performance is then evaluated using the PE ratio (which they designate as an Adjusted Ranking Metric [ARM].

Methods Panel member 5: Provided an empirical approach; however, would have liked to also have seen the rationale for theoretically adding the variables to be tested in the first place.

Methods Panel member 3:I could not find any informaiton on the discrimination and calibration of the risk models.

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

Methods Panel member 2: The face validity of the CLABSI definition and criteria were assessed by the Healthcare Infection Control Practices Advisory Committee (HICPAC) Subject Matter Expert (SME) panel using Delphi process in a previous endorsement.

Methods Panel member 1: Data element level validity is probably sufficient; there were no empirical results provided for performance measure validity

Methods Panel member 4:MD assessed the predictive validity of the risk adjustment model in development data using deviance, log likelihood and Akaike information criterion statistics. No calibration curves were presented.

Methods Panel member 3:5 states validated the data elements used for the CLABSI measure. It appears as if they used a TEP for face validity of the measure score, but this was poorly described.

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

Submission document: Testing attachment, section 2b2.3

Methods Panel member 2: Data element validation tests described above under reliability indicated a pooled mean sensitivity of 87.5% (range: 80.3%-100%), specificity of 99.3% (range: 98.7% - 100%), positive predictive value of 96.9% (range: 94.2% - 100%) and negative predictive value of 96.9% (range: 93.7% - 100%).

Methods Panel member 1: Data element level validity is probably sufficient; there were no empirical results provided for performance measure validity

Methods Panel member 4: The model deviance, log likelihood and Akaike information criterion statistics are not adequate to assess validity of risk adjustment model. No testing was performed in a validation data set.

Methods Panel member 3: The results for the data element validity testing were very good (sensitivity, specificity, NPV, and PPV). It was unclear what the results were for the face validity.

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

    Submission document: Testing attachment, section 2b1.

    ☒ **Yes Methods Panel member 1**: (for data elements)

    ☒ **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 <u>are</u> threats to validity and/or relevant threats to validity were <u>not assessed OR</u> 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 <u>is required</u>; 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.**

    Methods Panel member 2: The methods described were appropriate and the results good, but since score-level testing was not conducted, Moderate is the highest eligible rating.

    Methods Panel member 1: Given that this is a resubmission for a performance measure, the rating probably should be either Low or Insufficient given that no performance measure results (neither general descriptive nor meaningful differences at the proposed measurement level—hospital/facility) were provided.

    Methods Panel member 4:MD tested sensitivity, specificity, PPV and NPF of outcome variable (CLASBI) by re-abstracting data using trained auditors. Validations of critical data elements performed by external validators indicated a pooled mean sensitivity of 87.5% (range: 80.3%-100%), specificity of 99.3% (range: 98.7% - 100%), positive predictive value of 96.9% (range: 94.2% - 100%) and negative predictive value of 96.9% (range: 93.7% - 100%).

As per NQF criteria, a measure can be deemed scientifically valid if data reliability and data validity testing is provided. NQF provided guidance to MD that data validity testing incorporates data reliability testing. This guidance is reasonable. However, this is a risk-adjusted measure, and the risk-adjustment model was not adequately tested. In particular, no testing was performed in a validation data set. This is an important threat to validity. Hence, I assigned "low" to overall measure of validity. In addition, there was no attempt to support the decision to not include any measure of social risk in the risk adjustment model – other than to state that there is a "paucity of evidence." At a minimum, the MD needs to describe the evidence and make an argument as to why there is a "paucity of evidence."

However, the most important threat to validity is the lack of patient-level risk factors in the risk adjustment model. Although I appreciate the MD's intent to minimize data collection burden, these measures are used for reimbursement purposes – and may unintentionally penalize providers with higher-risk patients due to incomplete risk adjustment. Including hospital level characteristics in the risk adj model is not enough. Readmission and mortality models include patient-level risk adjustment. Why should it be acceptable to not include patient risk in complication models? Before leaving out patient-level risk factors, the MD needs to empirically show that they do not make a difference in the evaluation of hospital performance.

Methods Panel member 5: Outcome measure up for maintenance; note reliability testing for elements conducted, and to see validity section; however, element validity testing and score face validity only, no empirical validity conducted. Additionally, no maintenance analyses reported, simply noted that the panel didn't make any changes to the definitions, so they are therefore valid. Some 2015 data from last submission are still included.

Methods Panel member 3: Given the lack of information on discriminaton and calibration, it is not clear about the appropriateness of the risk adjustment models.

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.**

    Methods Panel member 1: Perhaps the measure developer has the missing information (e.g., performance measure and risk adjustment model performance) and was—for some reason—unaware that s/he was supposed to provide this in the form.

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

2a1. Reliability – Specifications

Comments:

**While presentation is a both awkward/out of place, the reliability appears to meet moderate or above thresholds. The adjusting and sampling concerns are modest and well understood.

**No concerns

**No concerns about reliability

**No concerns

2a2. Reliability – Testing

Comments:

**No

**No concerns

**testing is sufficient

**Appreciate the data on independent verification of application of the definitions of the measure.

2b1. Validity –Testing

Comments:

**No concerns

**No - brief reporting is adequate and is consistent with ongoing mesaure that is mature and needed. No analytic concerns.

**No concerns

**No concerns

**No

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

Comments:

**No

**No concerns though interval reporting is modest; given the directness and maturity of the measure and trajectory of outcomes, this is a minor concern.

**No

**No concerns

**No concerns

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

Comments:

**No comment

**No concerns

**Appropriate

**No concerns

** No concerns

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.

- Data for the measure are collected through the National Healthcare Safety Network (NHSN) using a set of standardized forms.
- The developer reports that CLABSI and central line days (the numerator and denominator) must be collected by trained hospital staff from information available in clinical data sources.
- The developer notes that some of the data used in the measure can be mined from electronic sources, adding that NHSN is moving towards an electronically captured CAUTI measure for future use. However, development and testing is not complete at this time.

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

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility

Comments:

**CAUTI or CLABSI at 3rd bullet? I have concerns about data capture form paper records.

**Easily extracted and collated.

**Feasible

**labor intensive however working towards electonic measure which is positive

**In use today, no issues

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**

OR

Planned use in an accountability program? ☐ **Yes** ☐ **No**

Accountability program details

- The measure is used in several accountability programs, including:
  - Hospital Inpatient Quality Reporting Program (HIQR)
  - Hospital Value-Based Purchasing
  - Hospital-Acquired Condition Reduction Program (HACRP)

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

Additional Feedback:

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

- The developer refers to trend data provided in section 1b
- The developer states that, combined with declining SIRs, which change in relation to the number of CLABSIs per central line days, declines in the device (i.e. central line) use ratio on wards and especially neonatal intensive care units (NICUs) highlight the net benefit to patients afforded by both the safer and reduced central line use

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

- The developer notes that The NHSN reporting tool includes business logic to minimize misclassification of CLABSI and inaccurate reporting of catheter days. In addition, site visits can be conducted to audit data validity and this has been done for other infection types by some of the states using NHSN as their mandatory reporting tool.

Additional Feedback:

- The developer reports that, based on results from a polling survey, hospitals have indicated that they are running SIR analysis reports within NHSN on a monthly basis, and that they use SIRs for prevention activities in their hospital. State health departments are using the SIR for public reporting purposes and to help target facilities for additional prevention. Feedback was received via email regarding the extent of risk adjustment and the limitations.

Questions for the Committee:

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- Do the benefits of the measure outweigh any potential unintended consequences?

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:

**There seems to be lack of quantitative data on answers to these questions.

**Reported widely, used by multiple groups for assessing care, and triggering action and improvement - all suggests high transparency and good feedback loop.

**Appropriate

**already publicly reported

**Well implemented measure that has led to meaningful improvement.

4b1. Usability – Improvement

Comments:

**There seems to be no evidence that consumers are using the data.

**Limited harm data shared; benefits implicit and clear.

**Benefits outweigh any harms

**no concerns

**Well established and meaningful measure.

Criterion 5: Related and Competing Measures

Related or competing measures

N/A

Harmonization

N/A

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

5. Related and Competing

Comments:

**None

**N/A

**None

**No

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.**

1. Brief Measure Information

NQF #: 0139

Corresponding Measures:

De.2. Measure Title: National Healthcare Safety Network (NHSN) Central line-associated Bloodstream Infection (CLABSI) Outcome Measure

Co.1.1. Measure Steward: Centers for Disease Control and Prevention

De.3. Brief Description of Measure: Standardized Infection Ratio (SIR) and Adjusted Ranking Metric (ARM) of healthcare-associated, central line-associated bloodstream infections (CLABSI) will be calculated among patients in bedded inpatient care locations.

This includes acute care general hospitals, long-term acute care hospitals, rehabilitation hospitals, oncology hospitals, and behavioral health hospitals.

1b.1. Developer Rationale: It is envisioned the use of this measure will promote CLABSI prevention activities which will lead to improved patient outcomes including reduction of avoidable medical costs, and patient morbidity and mortality.

S.4. Numerator Statement: Total number of observed healthcare-associated CLABSI among patients in bedded inpatient care locations.

S.6. Denominator Statement: Total number of predicted healthcare-associated CLABSI among patients in bedded inpatient care locations, calculated using the facility's number of central line days and the following significant risk factors:

- Acute Care Hospitals: CDC location, facility bed size, medical school affiliation, facility type, birthweight category (NICU locations only)
- Critical Access Hospitals: no significant risk factors, calculation based intercept only model
- Inpatient Rehabilitation Facilities: Proportion of admissions with stroke, proportion of admissions in other non-specific diagnostic categories
- Long Term Acute Care Hospitals: CDC location type , facility bed size, average length of stay, proportion of admissions on a ventilator, proportion of admissions on hemodialysis

S.8. Denominator Exclusions: Data from patients who are not assigned to an inpatient bed are excluded from the denominator counts, including outpatient clinics, 24-hour observation units, and emergency department visits. Inpatient rehab locations and inpatient psychiatric locations that have their own Centers for Medicare and Medicaid Services (CMS) Certification Number (CCN) are excluded.

De.1. Measure Type: Outcome

S.17. Data Source: Electronic Health Data, Electronic Health Records, Other, Paper Medical Records

S.20. Level of Analysis: Facility, Population : Regional and State

IF Endorsement Maintenance – Original Endorsement Date: Aug 10, 2009 **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?

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

CLABSI_NQF_Evidence_final_review.docx

1a.1 For Maintenance of Endorsement: 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.

Yes

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** NQF 0139

Measure Title: National Healthcare Safety Network Central Line-associated Bloodstream Infection (CLABSI)

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: National Healthcare Safety Network Central Line-associated Bloodstream Infection (CLABSI)

☐ 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: Click here to name what is being measured

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

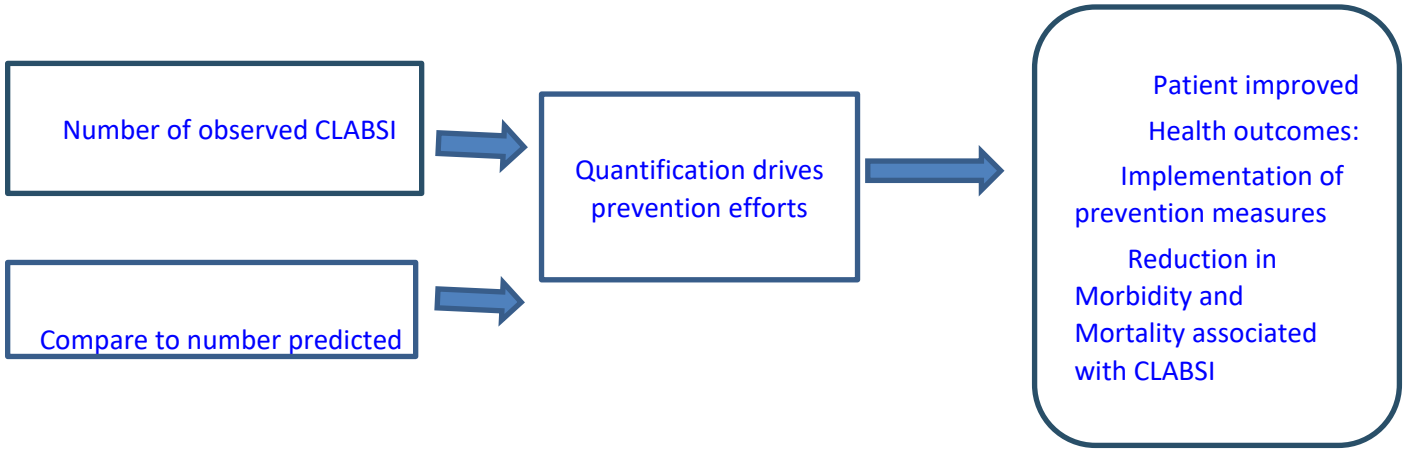
☐ Structure: Click here to name the structure

☐ 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.

An array of prevention efforts have been identified to reduce the incidence of CLABSI. These interventions include (i) Appropriate central line use: promptly removing non-essential intravascular catheters, (ii) hand hygiene and aseptic technique, (iii) the use of maximal barrier equipment including a large patient

drape, inserter mask, sterile gloves, cap, and sterile gown during aseptic insertion of the central line, (iv) appropriate insertion site decontamination before central line insertion, (v) chlorhexidine-impregnated dressings (in patients ≥ 18 years), and (vi) implementing surveillance strategies. The SIR is used as a quality measure and to drive infection prevention efforts in patient care locations, hospitals, health systems, and collaboratives. The SIR describes a healthcare facility´s performance compared to a national baseline. Facilities are able to compare the number of HAI CLABSI events to the number predicted, given national data.



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.)

****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.

O'Grady NP, Alexander M, Burns LA, Dellinger PE, Garland J, et al. Guidelines for the prevention of intravascular catheter-related infections. Available at http://www.cdc.gov/hicpac/pdf/guidelines/bsi-guidelines-2011.pdf. Accessed 3/14/19

See Guideline pages 9-19

Central Venous Catheters

10. Promptly remove any intravascular catheter that is no longer essential. (Category IA)

Hand Hygiene and Aseptic Technique

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

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)
1. Prepare clean skin with an antiseptic (70% alcohol, tincture of iodine, or alcoholic chlorhexidine gluconate solution) before peripheral venous catheter insertion. (Category IB)

Skin Preparation

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)

Catheter Site Dressing Regimens

12. For patients aged 18 years and older: a. Chlorhexidine-impregnated dressings with an FDA-cleared label that specifies a clinical indication for reducing catheter-related bloodstream infection (CRBSI) or catheter-associated bloodstream infection (CABSI) are recommended to protect the insertion site of short-term, non-tunneled central venous catheters. (Category IA).

Definitions of strength of recommendation category:
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.

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**	
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.	
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	

Grade assigned to the **recommendation** with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

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.

It is envisioned the use of this measure will promote CLABSI prevention activities which will lead to improved patient outcomes including reduction of avoidable medical costs, and patient morbidity and mortality.

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.*

When SIRs are compared over time, assessment of performance can be made. Hospitals have made significant progress in preventing CLABSIs–nationally, there has been a roughly 50% drop in CLABSIs between 2008 and 2016. For figures related to data please see https://www.cdc.gov/hai/surveillance/data-reports/data-summary-assessing-progress.html

CLABSI using the 2015 baseline:

National CLABSI SIR in 2015 is 0.994 = 26,029 observed / 26,183.537 predicted

National % change vs. baseline in 2015 < 1%

National CLABSI SIR in 2016 is 0.891 = 23,591 observed / 26,472.710 predicted

National % change vs. baseline in 2016 is 10%

National CLABSI SIR in 2017 is 0.814 = 21,173 observed / 25,993.180 predicted

There was about a 9% statistically significant decrease in CLABSI in acute care hospitals between 2016 and 2017, with the largest decrease seen in wards (10%).

Percent Change 2016 v. 2015 10% decrease

2015-

facilities: 3,550

Median: 0.868

Range, at 5% and 95%: (0.000 – 2.440)

2016-

facilities: 3,531

Median: 0.783

Range, at 5% and 95%: (0.000 – 2.239)

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.

The 2017 Current HAI Progress Report:

https://www.cdc.gov/hai/data/portal/progress-report.html

The Healthcare-associated Infections in the United States, 2006-2016: A Story of Progress located here: https://www.cdc.gov/hai/surveillance/data-reports/data-summary-assessing-progress.html

The 2016 National and State Healthcare-associated Infection Data Report:

https://www.cdc.gov/hai/data/portal/progress-report.html

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.*

No studies provide evidence of a direct relationship between social risk and HAIs. Instead, they provide evidence that social risk factors are associated with an increased risk of chronic disease conditions, suboptimal care for those conditions, compromised functional status, exposure to nursing homes, and colonization with bacterial pathogens. While these associations may be meaningful, they do not establish a direct relationship between social risk factors and HAIs.

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

Bakullari, Anila, Mark L. Metersky, Yun Wang, Noel Eldridge, Sheila Eckenrode, Michelle M. Pandolfi, Lisa Jaser, Deron Galusha, and Ernest Moy. "Racial and Ethnic Disparities in Healthcare-Associated Infections in the United States, 2009–2011." Infection Control and Hospital Epidemiology 35, no. S3 (2014): S10-16. doi:10.1086/677827

Among patients hospitalized with acute cardiovascular disease, pneumonia, and major surgery, Asian and Hispanic patients had significantly higher rates of HAIs than white, non-Hispanic patients.

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):*

Infectious Diseases (ID)

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

Primary Prevention, 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 : Dual eligible beneficiaries, 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://www.cdc.gov/nhsn/pdfs/pscmanual/4psc_clabscurrent.pdf (p.6, p.13)

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)*

Attachment **Attachment:** Copy_of_nhsn-data-dictionary-636893674617277045.xlsx

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.

S.3.1. <u>For maintenance of endorsement:</u> 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.

Yes

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

- Starting in 2018, we removed the exclusion for non-accessed central lines and we now include all central lines in the denominator device day counts regardless of access. There was confusion related to counting non-accessed permanent central lines and removal of this exclusion reduced confusion related to denominator device day counts.
- Intravenous Drug Abuse (IVDA): Bloodstream Infections (BSI) accompanied by documentation of observed or suspected injection into an IV line by the patient during the BSI Infection Window Period are excluded as CLABSIs regardless of presence of a central line. Patients meeting this CLABSI exclusion are at an increased risk for a CLABSI event due to the lack of aseptic technique when injecting their own line.
- In 2019, LCBI definition will no longer be met with the identification of a virus or parasite. These pathogens were excluded from the measure since they are not common causes of HAIS.
- Pus at the vascular access site: Bloodstream Infections (BSIs) accompanied by documentation of required elements during the BSI Infection Window Period are excluded as CLABSIs regardless of presence of a central line.
- Munchausen Syndrome by Proxy (MSBP): Bloodstream infections (BSIs) accompanied by documentation of a confirmed or suspected diagnosis of MSBP during the current admission are excluded as CLABSIs regardless of the presence of a central line.
- Epidermolysis bullosa (EB): Bloodstream infections (BSIs) accompanied by documentation during the current admission are excluded as CLABSIs regardless of the presence of a central line. Patients meeting this CLABSI exclusion are more likely to develop wounds that are heavily colonized with bacteria placing them at an increased risk for BSIs. Oftentimes skin cultures are not collected because the culturing process is painful for patients.
- Group B streptococcus: Group B Streptococcus identified from blood, with a date of event during the first 6 days of life, will not be reported as a CLABSI. Patients positive for GBS in the first 6 days of life are excluded because of vertical transmission.

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.*

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

Total number of observed healthcare-associated CLABSI among patients in bedded inpatient care locations.

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)*

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

Numbers of CLABSIs attributed to each location are counted for each month using the definitions below. CLABSIs attributed to neonatal ICUs are stratified by birth weight category. CLABSIs attributed to Specialty Care Areas or Oncology Locations are stratified by association with temporary vs. permanent central line.

1. Definition of infection that is Present on Admission (POA): An infection is considered Present on Admission (POA) if the date of event of the NHSN site-specific infection criterion occurs during the POA time period, which is defined as the day of admission to an inpatient location (calendar day 1), the 2 days before admission, and the calendar day after admission. For purposes of NHSN surveillance and determination of the Repeat Infection Timeframe (as defined below) if the date of event is determined to be either of the two days prior to inpatient admission, then the date of event will be hospital day 1. POA events are excluded

2. Definition of Healthcare-associated Infection (HAI): An infection is considered a Healthcare-associated Infection (HAI) if the date of event of the NHSN site-specific infection criterion occurs on or after the 3rd calendar day of admission to an inpatient location where day of admission is calendar day 1.

3. Definition of Eligible Central Line: A CL that has been in place for more than two consecutive calendar days (on or after CL day 3), following the first access of the central line, in an inpatient location, during the current admission. Such lines are eligible for CLABSI events and remain eligible for CLABSI events until the day after removal from the body or patient discharge, whichever comes first.

4. Definition of Central line: An intravascular catheter that terminates at or close to the heart or in one of the great vessels that is used for infusion, withdrawal of blood, or hemodynamic monitoring. The following are considered great vessels for the purpose of reporting CLABSI events and counting central-line device days in the NHSN system: Aorta, pulmonary artery, superior vena cava, inferior vena cava, brachiocephalic veins, internal jugular veins, subclavian veins, external iliac veins, common iliac veins, femoral veins, and in neonates, the umbilical artery/vein.

Neither the type of device nor the insertion site are used to determine if a device is considered a central line for NHSN reporting purposes.

The following devices are not considered central lines for NHSN Reporting Purposes:

- Non-lumened Pacemaker wires and other non-lumened devices inserted into central blood vessels or the heart
- Arterial catheters
- Arteriovenous fistula
- Arteriovenous graft
- Atrial catheters (also known as transthoracic intra-cardiac catheters, those catheters inserted directly into the right or left atrium via the heart wall)
- Extracorporeal membrane oxygenation (ECMO)
- Hemodialysis reliable outflow (HERO) dialysis catheters
- Intra-aortic balloon pump (IABP) devices
- Peripheral IV or Midlines
- Ventricular Assist Device (VAD)

5. Definition of CLABSI: A laboratory confirmed bloodstream infection which meets LCBI Criterion 1, 2, or 3, and where an eligible BSI organism is identified and an eligible central line is present on the LCBI DOE or the day before. Access definition: The performance of any of the following activities during the current inpatient admission

6. Definition of Infusion: The administration of any solution through the lumen of a catheter into a blood vessel. Infusions include continuous infusion (for example, nutritional fluids or medications), intermittent infusion (for example, IV flush), IV antimicrobial administration, and blood transfusion or hemodialysis treatment.

7. Definition of Temporary Central Line: A non-tunneled, non-implanted catheter.

8. Definition of Permanent Central Line: Tunneled catheters, (including tunneled dialysis catheters) and implanted catheters (including ports)

9. Definition of Laboratory Confirmed Bloodstream Infection (LCBI):

For all LCBI definitions, the following resources may be referenced:

- Appendix B: Secondary BSI Guide of the CLABSI Surveillance protocol can be found at www.cdc.gov/nhsn/PDFs/pscManual/4PSC_CLABScurrent.pdf (p.32)
- NHSN Common Commensals from the NHSN Organism List can be found at https://www.cdc.gov/nhsn/xls/master-organism-com-commensals-lists.xlsx

LCBI must meet one of the following criteria:

LCBI Criterion 1: Patient of any age has a recognized bacterial or fungal pathogen not included on the NHSN common commensal list, identified from one or more blood specimens obtained by a culture or non-culture based microbiologic testing methods

AND

Organism(s) identified in blood is not related to an infection at another site

(See Appendix B [p.32] Secondary BSI Guide)

LCBI Criterion 2: Patient of any age has at least one of the following signs or symptoms: fever (>38 degrees C), chills, or hypotension and positive Organism(s) identified in blood

AND

Organism(s) identified in blood is not related to an infection at another site

AND

The same NHSN common commensal is identified by a culture or non-culture based microbiologic testing method, from two or more blood specimens collected on separate occasions not related to an infection at another site and the same NHSN common commensal is identified from two or more blood specimens drawn on separate occasions, by a culture or non-culture based microbiologic testing method.

Common Commensal organisms include, but not are not limited to, diphtheroids (Corynebacterium spp. not C. diphtheria), Bacillus spp. (not B. anthracis), Propionibacterium spp., coagulase-negative staphylococci (including S. epidermidis), viridans group streptococci, Aerococcus spp. Micrococcus spp, and Rhodococcus spp.

For a full list of Common Commensals see the Common Commensal tab of the NHSN organisms list. Criterion elements must occur within the Infection Window Period, the seven-day time period which includes the date the positive blood culture was collected, the 3 calendar days before and the 3 calendar days after. Note: The matching common commensals represent a single element; therefore, the collection date of the first common commensal is the date of the element used to determine the Date of Event.

LCBI Criterion 3: Patient 1 year of age or less has at least one of the following signs or symptoms: fever (>38 degrees C), hypothermia (<36 degrees C), apnea, or bradycardia and organism identified in blood not related to an infection at another site (See Appendix B Secondary BSI Guide) and the same NHSN common commensal is identified from two or more blood specimens drawn on separate occasions, by a culture or non-culture based microbiologic testing.

10. Criteria for meeting Mucosal Barrier Injury (MBI) Laboratory Confirmed Bloodstream Infection (LCBI)

For all MBI-LCBI definitions, the following resources may be referenced:

- Appendix B: Secondary BSI Guide of the CLABSI Surveillance protocol can be found at www.cdc.gov/nhsn/PDFs/pscManual/4PSC_CLABScurrent.pdf (p.32)
- NHSN Common Commensals from the NHSN Organism List can be found at https://www.cdc.gov/nhsn/xls/master-organism-com-commensals-lists.xlsx
- NHSN MBI Organism List can be found at https://www.cdc.gov/nhsn/xls/analysis/nhsn-data-dictionary.xlsx

MBI-LCBI Criterion1: Patient of any age fully meets criterion 1 for LCBI with at least one blood specimen identified by a culture or non-culture based microbiologic testing method, with ONLY intestinal organisms from the NHSN MBI organism list and patient meets at least one of the following:

a)Is an allogeneic hematopoietic stem cell transplant recipient within the past year with one of the following documented during same hospitalization as positive blood culture:

i.) Grade III or IV gastrointestinal graft versus host disease [GI GVHD]

ii.)1 liter or more diarrhea in a 24-hour period (or 20 or more mL/kg in a 24-hour period for patients <18 years of age) with onset on or within 7 calendar days before the date the positive blood specimen was collected

b)Is neutropenic, defined as at least 2 separate days with values of absolute neutrophil count (ANC) and/or white blood cell (WBC) values <500 cells/mm3 within a seven-day time period which includes the collection date of the positive blood specimen (Day 1), the 3 calendar days before and the 3 calendar days after.

MBI-LCBI Criterion 2: Patient of any age meets criterion 2 for LCBI when the blood specimens identify only viridans group streptococci or Rothia spp and patient meets at least one of the following:

a)Is an allogeneic hematopoietic stem cell transplant recipient within the past year with one of the following documented during same hospitalization as positive blood culture:

i.) Grade III or IV gastrointestinal graft versus host disease [GI GVHD]

ii.)1 liter or more diarrhea in a 24-hour period (or 20 or more mL/kg in a 24-hour period for patients <18 years of age) with onset on or within 7 calendar days before the date the positive blood specimen was collected

b)Is neutropenic, defined as at least 2 separate days with values of absolute neutrophil count (ANC) and/or white blood cell (WBC) values <500 cells/mm3 within a seven-day time period which includes the collection date of the positive blood specimen (Day 1), the 3 calendar days before and the 3 calendar days after

MBI-LCBI Criterion 3: Patient 1 year of age or less meets criterion 3 for LCBI when the blood specimens identify only viridans group streptococci or Rothia spp and patient meets at least one of the following:

a)Is an allogeneic hematopoietic stem cell transplant recipient within the past year with one of the following documented during same hospitalization as positive blood culture:

i.) Grade III or IV gastrointestinal graft versus host disease [GI GVHD]

ii.)1 liter or more diarrhea in a 24-hour period (or 20 or more mL/kg in a 24-hour period for patients <18 years of age) with onset on or within 7 calendar days before the date the positive blood specimen was collected

b)Is neutropenic, defined as at least 2 separate days with values of absolute neutrophil count (ANC) and/or white blood cell (WBC) values <500 cells/mm3 within a seven-day time period which includes the collection date of the positive blood specimen (Day 1), the 3 calendar days before and the 3 calendar days after

11. Definition of CDC Location: The patient care area to which a patient is assigned while receiving care in the healthcare facility. NOTE: Only locations where patients are housed overnight (i.e., inpatient locations) and where denominator data are collected can be used for reporting CLABSI data. Operating rooms (including cardiac cath labs, c-section rooms, and interventional radiology) and outpatient locations are not valid locations for this type of surveillance. See attached list of CDC/NHSN Location Types to identify Special Care Areas or Oncology Locations. https://www.cdc.gov/nhsn/xls/analysis/nhsn-data-dictionary.xlsx

12. Definition of Infection Window Period: Infection Window Period is defined as the 7-days during which all site-specific infection criteria must be met. It includes the day the first positive diagnostic test that is an element of the site-specific infection criterion, was obtained, the 3 calendar days before and the 3 calendar days after. For purposes of defining the Infection Window Period the following are considered diagnostic tests:

- laboratory specimen collection
- imaging test
- procedure or exam

13. Definition of Repeat Infection Timeframe (RIT): The RIT is a 14-day timeframe during which no new infections of the same type are reported. The date of event is Day 1 of the 14-day RIT. Additional pathogens recovered during the RIT from the same type of infection are added to the event.

The RIT will apply at the level of specific type of infection with the exception of BSI, UTI, and PNEU where the RIT will apply at the major type of infection.

14. Definition of Date of Event (DOE): The Date of Event is the date the first element used to meet an NHSN site-specific infection criterion occurs for the first time within the seven-day infection window period.

15. Definition of Location of Attribution: The location to which the CLABSI is attributed.

16. Definition of birthweight: Birthweight is the weight of the infant at the time of birth and should not be changed as the infant gains weight. The birthweight categories are as follows:

A = 750 g or less; B = 751-1000 g; C = 1001-1500 g; D = 1501-2500 g; E = >2500 g.

17. Definitions for facility physician education status: Teaching statuses: major, graduate, undergraduate - Major: Facility has a program for medical students and post-graduate medical training; Graduate: Facility has a program for post-graduate medical training (i.e., residency and/or fellowships); Undergraduate: Facility has a program for medical students only.

Exclusions from CLABSI:

1. Bloodstream Infections (BSI) accompanied by documentation of observed or suspected injection into an IV line by the patient during the BSI Infection Window Period are excluded as CLABSIs regardless of presence of central line.

2. Group B Streptococcus identified from blood, with a date of event during the first 6 days of life, are excluded as CLABSIs regardless of presence of central line.

3. Occasionally, a patient with both a central line and another vascular access device* will have pus at the other access site. If there is pus at the site of one of the following vascular access devices and a specimen collected from that site has at least one matching organism to an organism identified in blood this will be considered an LCBI but not a CLABSI for NHSN reporting purposes.

*Vascular access devices included in this exception are limited to:

• Arterial catheters

• Arteriovenous fistulae

• Arteriovenous grafts

• Atrial catheters (also known as transthoracic intra-cardiac catheters, those catheters inserted directly into the right or left atrium via the heart wall)

• Hemodialysis reliable outflow (HERO) dialysis catheters

• Intra-aortic balloon pump (IABP) devices

• Non-accessed CL (those neither inserted nor used during current admission)

• Peripheral IV or Midlines

4. CLABSIs in which any of the following organisms are the only pathogens identified are excluded:

- Blastomyces spp.
- Histoplasma spp.
- Coccidioides spp.
- Paracoccidioides spp.
- Cryptococcus spp.
- Pneumocystis spp.

- Any virus
- Parasites

5. If the date of blood specimen collection is on or after the date of documentation of evidence of consent AND the patient is being supported for organ donation purposes, an event identified using the blood specimen result should not be reported as CLABSI.

6. MBI CLABSI events will be excluded from the CLABSI measure

7. Munchausen Syndrome by Proxy (MSBP): If during the current admission, there is documentation of known or suspected (MSBP), also known as factitious disorder imposed on another and a CL has been in place for more than 2 days on a BSI DOE, these events are considered LCBIs but are NOT considered central line associated.

8. Epidermolysis bullosa (EB): If during the current admission, there is a diagnosis of and a CL has been in place for more than 2 days on a BSI DOE, these events are considered LCBIs but are NOT considered central line associated.

9. Extracorporeal life support (ECMO): A BSI meeting LCBI criteria with an eligible central line where ECMO is present for more than 2 days on the BSI DOE, and is still in place on the DOE or the day before, will be considered an LCBI but not a CLABSI for NHSN reporting purposes.

10. Ventricular assist device (VAD): A BSI meeting LCBI criteria with an eligible central line where ECMO is present for more than 2 days on the BSI DOE, and is still in place on the DOE or the day before, will be considered an LCBI but not a CLABSI for NHSN reporting purposes.

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

Total number of predicted healthcare-associated CLABSI among patients in bedded inpatient care locations, calculated using the facility's number of central line days and the following significant risk factors:

- Acute Care Hospitals: CDC location, facility bed size, medical school affiliation, facility type, birthweight category (NICU locations only)
- Critical Access Hospitals: no significant risk factors, calculation based intercept only model
- Inpatient Rehabilitation Facilities: Proportion of admissions with stroke, proportion of admissions in other non-specific diagnostic categories
- Long Term Acute Care Hospitals: CDC location type , facility bed size, average length of stay, proportion of admissions on a ventilator, proportion of admissions on hemodialysis

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).

Methodologies for counting central line days differ according to the location of the patients being monitored. Numbers of central line days attributed to each location are counted for each data period utilizing the following definitions and guidelines. In locations that are not neonatal ICUs, SCA or oncology locations, all CL days for that location and data period are summed. For neonatal ICU central line days counts are stratified by birthweight category. CL day counts for Special Care Areas or Oncology Locations are stratified by temporary vs. permanent central line type.

For locations other than specialty care areas/oncology (SCA/ONC) and NICUs (e.g., ICUs, step-down units, wards), the denominator sampling method can be used. (Refer to sampling method in the Device-Associated BSI protocol available at www.cdc.gov/nhsn/PDFs/pscManual/4PSC_CLABScurrent.pdf)

1. Definition of central line day: For each patient, a day that at least one central line was present at the time of the CL day count.

2. Definition of CDC Location (acute care hospitals, long term acute care hospitals): Each patient care area in a facility that is monitored in NHSN is "mapped" to one or more CDC Locations. The specific CDC Location code is determined by the type of patients cared for in that area according to the 80% Rule. That is, if 80% of patients are of a certain type (e.g., pediatric patients with orthopedic problems) then that area is designated as that type of location (in this case, an Inpatient Pediatric Orthopedic Ward).
https://www.cdc.gov/nhsn/pdfs/pscmanual/15locationsdescriptions_current.pdf

3. Definition of Medical school affiliation categories:

a. Major – facility has a program for medical students and post-graduate medical training

b. Graduate – facility has a program for post-graduate medical training (i.e., residency and/or fellowships)

c. Undergraduate: facility has a program for medical students only

4. Definition of Facility bed size: Number of beds set up and staffed in the healthcare facility

5. Setting (Freestanding or Within a Hospital): Describes physical placement of LTACH or IRF and does not define financial or administrative relationship with other healthcare facility types.

6. Average Length of Stay: number of patient days during the calendar year divided by the number of admissions during the calendar year

7. Proportion of admissions within a diagnostic category: number of admissions during the calendar year where the primary diagnosis is of that type (e.g., traumatic spinal cord dysfunction) divided by the total number of admissions during the calendar year

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

Data from patients who are not assigned to an inpatient bed are excluded from the denominator counts, including outpatient clinics, 24-hour observation units, and emergency department visits. Inpatient rehab locations and inpatient psychiatric locations that have their own Centers for Medicare and Medicaid Services (CMS) Certification Number (CCN) are excluded.

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.)*

See S.8. Definition of inpatient - A patient who is located in an inpatient location for care and treatment at the time of the daily inpatient census count.

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 final risk model for the CLABSI SIR in Acute Care Hospitals includes: CDC locations, facility bed size, medical school affiliation, and facility type. For NICU locations the risk factor included in the final model was birthweight category. See S7 above

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

Statistical risk model

If other:

S.12. Type of score:

Ratio

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 = Lower 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.)*

The Standardized Infection Ratio (SIR) for annual and quarterly data aggregation and analysis of CLABSI events is calculated for each healthcare facility for a specified time period. The SIR is an indirect standardization method for summarizing healthcare associated infection (HAI) experience, including CLABSI events, in a single group of data or across any number of stratified groups of data. To produce the SIR:

1. Identify number of observed healthcare-associated CLABSIs for a given time period by adding the total number of observed CLABSIs across the facility.

2. Calculate the number of predicted healthcare-associated CLABSIs for each CDC location using a negative binomial regression model and the risk factors described above.

3. Calculate the number of predicted healthcare-associated CLABSIs for the facility and time period by adding the predicted number of CLABSIs for each location across the facility.

4. Divide the number of observed healthcare-associated CLABSIs (1 above) by the number of predicted healthcare-associated CLABSIs (3 above) to obtain the SIR.

5. Perform a Poisson test to compare the SIR obtained in 4 above to the nominal value of 1. P-value and confidence interval will be calculated, which can be used to assess significance of SIR.

(The NHSN analysis tool will perform the calculations once the patient infection data and denominator information are entered into the system.)

The Adjusted Ranking Metric (ARM) for annual data aggregation and analysis of HAI events, including CLABSI events, combines the method of indirect standardization used to calculate the unadjusted SIR described above with a Bayesian random effects hierarchical model to account for the potentially low precision and/or reliability inherent in the unadjusted SIR. A Bayesian posterior distribution constructed through Monte Carlo Markov Chain sampling is used to produce the adjusted numerator. The ARM enables more meaningful statistical differentiation between hospitals by accounting for differences in patient case-mix, exposure volume (e.g. patient days, central line-days, surgical procedure volume), and unmeasured factors that are not reflected in the unadjusted SIR and that cause variation between healthcare facilities. Accounting for these sources of variability enables better measure discrimination between facilities and leads to more reliable performance rankings. To produce the ARM:

1. Identify the number of CLABSI in each location

2. Obtain the adjusted number of observed CLABSIs by using a Bayesian posterior distribution constructed through Monte Carlo Markov Chain sampling which results from a Bayesian random effects model.

3. Total these numbers for an observed number of CLABSIs

4. Obtain the predicted number of CLABSIs in the same locations by multiplying the observed central line days according to the factors significantly associated with predicting CLABSI incidence as identified through a Log-linear Negative Binomial Regression Model.

5. Divide the total number of adjusted CLABSI events (“3” above) by the predicted number of CLABSIs (“5” above).

6. Result = ARM

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.

No sampling methodology is used in calculating the metric.

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.

Not PRO-PM

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

If other, please describe in S.18.

Electronic Health Data, Electronic Health Records, Other, Paper Medical Records

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.

NHSN Primary BSI collection form

NHSN Denominator for ICU form

NHSN Denominator for NICU form

NHSN Denominator for Specialty Care Area/Oncology Form

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)*

Available at measure-specific web page URL identified in S.1

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

Facility, Population : Regional and State

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

Inpatient/Hospital, Other, Post-Acute Care

If other: Oncology Hospital; IRF; LTACH; Inpatient Psych

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.)*

2. Validity – See attached Measure Testing Submission Form

NQF_CLABSI_Scientific_testing_attachment_Revision.2.26.19_-003-.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.

Yes

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.

Yes - Updated information is included

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

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

Measure Number (*if previously endorsed*)**:** NQF 0139
Measure Title: National Healthcare Safety Network Central Line Associated Bloodstream Infection
Date of Submission: 1/23/2019

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 (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). *Contact NQF staff if more pages are needed.***
- **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**
- **For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.**

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

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.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]

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: NHSN Primary BSI collection form; NHSN Denominator for ICU form; NHSN Denominator for NICU form; NHSN Denominator for Specialty Care Area/Oncology Form	☐ 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*).

The NHSN data set that was used was drawn from the NHSN database, which is an aggregation of data that healthcare facilities throughout the US submit, much of which is data required for submission to NHSN by state and federal mandates or both.

CDC NHSN used 2015 **healthcare-associated infection (**HAI) incidence and risk factor data to develop new predictive models for CLABSI and other HAI's. The number of facilities in 2015 reporting CLABSI data includes: 3550 acute care hospitals (ACH), 489 long term acute care hospitals (LTACH), 662 inpatient rehabilitation facilities (IRF) throughout the US national database.

Please refer to the SIR Guide (p.4) at:

https://www.cdc.gov/nhsn/pdfs/ps-analysis-resources/nhsn-sir-guide.pdf

1.3. What are the dates of the data used in testing? January 2015 – December 31, 2015

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:	☐ other:

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*)

CLABSI data is reported to NHSN from over 4701 facilities in all 50 states, the District of Columbia, and several US territories. In 2015, 26% of hospitals reporting CLABSI data have fewer than 50 beds, 39% have between 51 and 200 beds, and 34% have more 200 beds.

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*)
Facilities reporting CLABSI data to NHSN do not report a count of patients under surveillance. The number of central line days is reported, as described in the measure submission. In 2015, 23,367,134 central line days were reported by acute care participating facilities. The central line day counts are reported by patient care location in the hospital and are not stratified by patient level factors such as age, race, and sex. All ages, all races and all diagnoses are included.

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.

Reliability testing is provided via 2015-2016 state studies; validity testing provided via **working group of the Healthcare Infection Control Practices Advisory Committee** (HICPAC) which developed a definition of the Mucosal Barrier Injury (MBI) instituted as part of the CLABSI surveillance definitions to improve construct validity; sample used to test CLABSI risk models consist 4701 number of facilities reporting CLABSIs in 2015.

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.

No patient-level sociodemographic variables are used in the measure and none were available for analysis. There is no compelling evidence available supporting association between social risk factors and CLABSIs. Instead, they provide evidence that social risk factors are associated with an increased risk of chronic disease conditions, suboptimal care for those conditions, compromised functional status, exposure to nursing homes, and colonization with bacterial pathogens. While these associations may be meaningful, they do not establish a direct relationship between social risk factors and HAIs.

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*)

See section 2b1 for validity testing of data elements.

As per NQF email "…data element validity testing may serve as a demonstration of data element reliability." Please see section 2b1.2 for demonstration of data element reliability.

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*)

As per NQF email "…data element validity testing may serve as a demonstration of data element reliability." Please see section 2b1.3 for statistical results from reliability testing.

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?*)

As per NQF email "…data element validity testing may serve as a demonstration of data element reliability." Please see section 2b1.4 for interpretation of data element reliability.

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.

CLABSI definition and criteria are unchanged from prior submission, with the exception of the addition of the MBI exclusion, and definition and criteria were reviewed by the Healthcare Infection Control Practices Advisory Committee (HICPAC) panel using Delphi process which culminated with definition and criteria.

MBI exclusion was established using Delphi process by **the Healthcare Infection Control Practices Advisory Committee (HICPAC)** Subject Matter Expert (SME) panel with 100 % consensus on the MBI definition and criteria.

Validity of inclusionary criteria, namely lab confirmed BSI was not retested for validity purposes.

The Healthcare Infection Control Practices Advisory Committee (HICPAC) is a federal advisory committee chartered to provide advice and guidance to the Centers for Disease Control and Prevention (CDC) and the Secretary of the Department of Health and Human Services (HHS) regarding the practice of infection control and strategies for surveillance, prevention, and control of healthcare-associated infections, antimicrobial resistance and related events in United States healthcare settings.

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)*

Validity testing of critical data elements is performed by many of the state health departments that have implemented mandatory reporting of CLABSI data to the state using NHSN as the data entry system and the source of case definitions and surveillance methodology. Trained state health department validators apply NHSN CLABSI definition criteria in medical record reviews of records that were compiled during the stay in which patients reportedly met criteria of the CLABSI definition. The validator's determination of whether or not the patient in question had a CLABSI is compared to the facility's determination. Sensitivity, specificity, positive predictive value, and negative predictive value are then calculated. As part of the validation process, some state health departments validate counts of central line days through structured interviews with personnel who collect and report these data to NHSN to ensure that correct data collection methodology is used.

	Year of data validated	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
New Hampshire	2014-2015	80.6	98.9	96.2	93.7
Texas	2015	83.2	99.1	95.9	95.8
Vermont	2015	100	100	100	100
Alabama	2017	96.7	99.7	98.9	99.2
Georgia	2017	80.3	98.7	94.2	94.9

External validation of NHSN CLABSI data has been conducted by at least 5 states since 2015 (NH, TX, VT, AL, GA), using different sampling methods. These validations indicated a pooled mean sensitivity of 87.5% (range: 80.3%-100%), specificity of 99.3% (range: 98.7% - 100%), positive predictive value of 96.9% (range: 94.2% - 100%) and negative predictive value of 96.9% (range: 93.7% - 100%). External validation across the 5 states consisted of 2,594 chart reviews and of these 81 cases were incorrectly classified, yielding an overall classification error rate of 3.1%.

Chart reviews were conducted by trained auditors across the 5 state health departments. These audits identified 520 CLABSI events that should have been reported and among those 463 events were correctly reported by healthcare facilities (57 missed events). Three states (GA, NH, TX) identified incorrect secondary BSI attribution as a reason for failure to report CLABSI events - the cases did not met criteria for a primary source of infection so the BSI should have been reported as a primary CLABSI. This resulted in at least 5 underreported events. One state (NH) noted that one CLABSI was missed due to a symptom being missed by the facility during chart review. One state (AL) identified that underreporting occurred as a result of positive test results not being reported for CLABSI surveillance. One state (GA) identified a failure to report CLABSI events that occurred outside the repeat infection timeframe from a prior CLABSI event.

Among the 2,074 charts that were identified as not meeting the NHSN CLABSI definition, 2,050 charts were correctly called as "CLABSI negative" by the healthcare facilities, thereby leading to 24 over reported CLABSIs. Two states (NH, TX) identified incorrect secondary BSI attribution as a reason for overcalling CLABSI events - the BSI should not have been reported because the cases met criteria for a primary source of infection to which the BSI should have been attributed as secondary. This resulted in at least one over reported CLABSI event. States also reported reasons for errors in reporting both under- and over reported CLABSI events that included misidentifying the date of event or location of attribution (GA), inexperienced infection preventionist (AL), and "misinterpreting rules" of surveillance (AL).

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

Validity of exclusionary criteria, mainly MBI, was established using Delphi process of SME panel with 100% consensus on the MBI definition and criteria.

Validity of inclusionary criteria, namely lab confirmed BSI was not retested for validity purposes.

Validations of critical data elements performed by external validators indicated a pooled mean sensitivity of 87.5% (range: 80.3%-100%), specificity of 99.3% (range: 98.7% - 100%), positive predictive value of 96.9% (range: 94.2% - 100%) and negative predictive value of 96.9% (range: 93.7% - 100%).

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?*)

The results of the Delphi process and consensus of exclusionary criteria, via expert review reaffirm the CLABSI measure is valid for use as a quality measurement. The SIR is based on the standardized mortality ratio, which is a widely accepted method for summarizing the predicted to observed ratio. The CLABSI SIR can distinguish good from poor quality. In some places where large scale CLABSI prevention programs have been implemented over the past several years, significant reductions in the CLABSI SIR have been seen. However, there are still facilities with significantly high CLABSI SIRs, indicating that they have not made progress in reducing CLABSI (high SIRs indicating poor quality) while other facilities have made progress in reducing CLABSI and see significantly low CLABSI SIRs (low SIRs indicate better quality). The CLABSI SIR is used by many state health departments in public reporting of HAI data, and the Centers for Medicare and Medicaid Services (CMS) has included the CLABSI SIR in its Hospital Inpatient Quality Reporting Program and Hospital Value Based Purchasing Program.

The CLABSI SIR is only calculated when sufficient denominator data has been reported, i.e. when the number of predicted CLABSIs is greater than 1. In order to allow for an assessment of CLABSI experience in facilities with lower exposure to urinary catheters, the ARM is used. The ARM uses statistical techniques to

adjust for lower exposure to urinary catheters, in addition to other risk factors, and produces a measure that is interpreted similarly to the SIR.

Norms have not been established however we have a series of studies that show agreement. Very High specificity, PPV and NPV and high sensitivity.

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section 2b4*

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*)

Not applicable.

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*)

Not applicable.

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*)

Not applicable.

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 2b5.

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

☐ **No risk adjustment or stratification**

☒ **Statistical risk model with ___risk factors relevant to the facility type**

☐ **Stratification by ___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.

The risk modeling was conducted using negative binomial regression, in which risk factors were evaluated by both univariate and multivariate modeling steps. Univariate models were fist constructed to evaluate the relationship between each risk factor and the CLABSI incidence rate.

For detailed specifications of the risk model please refer to the SIR Guide (p. 5) at:

https://www.cdc.gov/nhsn/pdfs/ps-analysis-resources/nhsn-sir-guide.pdf

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.

Not applicable.

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?

In the interest of minimizing reporting burden, denominator data are aggregate data at the patient care level. As a result, the candidate risk factor data available are descriptive characteristics for patient care locations and healthcare facility. To risk adjust the CLABSI SIR, national NHSN data is analyzed to assess for differences in rates between different patient care locations (ICU, ward, different specialty types, etc) within the data. Additional facility level characteristics (bedsize, affiliation with a medical school, etc) are included in the analysis.

In the risk adjustment for the CLABSI ARM, national NHSN data is used to produce a negative binomial risk model that includes patient care location type, medical school affiliation, facility bedsize, and central line device utilization.

Model selection was used with variables added if p <0.05. Order of variables were included in the model was based on a combination of deviance, log likelihood and Akaike information criterion statistics.

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)**

No social risk factors applied in the modeling.

Due to the paucity of evidence to support social risk factors and data burden data collection for risk adjustment purposes, social risk factors are not collected in NHSN for all patients in the patient population; therefore, these variables are not available in NHSN to be used for risk adjustment modeling.

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

Variables were eligible for entering the model at p-value=0.25 and retaining in the model at p-value=0.05 significance level. Factors were entered into a multivariate model using forward selection, based on the lowest Wald Chi-square value. Goodness of fit was assessed at each modeling step using the Akaike Information Criterion (AIC) statistics. The final model resulting from forward selection was confirmed via backwards elimination, in which each variable was sequentially removed based on the highest p-value.

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.**

Did not include social risk factors.

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)

Model validation steps:

1. **For each negative binomial regression model to be validated, produce a table of the regression parameters from the final model**
2. **Generate at least 100 new replicate samples using "sampling with replacement" from the original dataset so that each replicate sample contains the same number of observations as the original dataset**
3. **Fit the final model to each of those new replicate samples and store the regression parameters**
4. **This will produce a set of regression parameters as defined in the original final model for each model fit to each new replicate sample**
5. **Produce a distribution of each regression parameter across all the results from the at least 100 new replicate samples**
6. **Construct an empirical "percentile-based" confidence range using the 2.5 and 97.5 percentile for each parameter**
7. **Consider the model validated among all parameters if the respective confidence range does not include null value**

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*)**:**

See 2b3.7

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

Negative binomial model discrimination and calibration were performed using a combination of deviance, log likelihood and Akaike information criterion statistics. Markov chain Monte Carlo sampling methods inherently rely on large scale simulation to produce posterior parameter estimates evaluated using trace plots and highest probability density intervals. In addition, Markov chain convergence, sampling, and stationarity were assessed using Geweke, Raferty-Lewis and Heidelberger-Welch diagnostics, respectively.

Negative Binomial model calibration was further assessed by calculating the root mean squared error (RMSE) between the observed and model predicted values for the final versus null model across 1000 bootstrap samples. The average RMSE for the final model was 1.432 compared to 1.504 for the null model and demonstrates a 5% improvement.

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

See 2b3.7

2b3.9. Results of Risk Stratification Analysis:

See 2b3.4a

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*)

Both risk adjustment methodologies (stratification based on patient care location type and facility-level factors

for the CLABSI SIR and risk modeling using similar factors for the ARM) allow for adequate controlling of factors that can lead to differences in CLABSI risk for patients in acute care hospitals.

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*)

Not Applicable

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)*

NHSN uses the mid p exact test for determining statistically significant differences in performance measurement. This test is applied to facility-specific performance summary statistics. CDC calculates these summary statistics i.e. the CLABSI SIR and ARM, to identify variation from a predicted occurrence of CLABSI based on the experience of a standard population, as well as an assessment of the magnitude of that variation (for example, an SIR of 2.0 indicates a level of occurrence two times higher than what would be predicted). The measures are produced with a confidence interval that can be used to assess the likelihood that the SIR or ARM occurs within a specified range. The confidence interval can be used to assess the SIR or ARM compared to its nominal value of 1.0 (where the number of observed equals the number of predicted CLABSIs).

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*)

Published data from the CDC national and state 2016 HAI progress report shows that 2,345 ACH reported sufficient data to generate a CLABSI SIR in 2016. Approximately 349 healthcare facilities (~14.88%) had SIRs that were statistically significantly less than 1.0, indicating that the facility reported fewer CLABSIIs than predicted. Approximately 169 healthcare facilities (~7.21%) had SIRs that were statistically significantly greater than 1.0, indicating that the facility reported more CLABSIs than predicted.

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?*)

The SIR and the ARM can both be used to assess differences in performance. Facilities that have SIRs significantly lower than 1 are succeeding in preventing CLABSI. Facilities with SIRs that are significantly higher than 1 may not have implemented CLABSI prevention efforts and are potential targets for interventions to improve prevention practices.

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*)

Not applicable.

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*)

Not applicable.

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*)

Not applicable.

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 of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

Healthcare facilities that submit quality measure data to NHSN for CLABSI and other HAIs must submit all data required for measure calculation; otherwise their data cannot be successfully submitted to NHSN. Within the NHSN reporting system, facilities are prompted each month that they have entered infection (numerator) data but no central line count (denominator) data and vice versa to ensure that monthly data submission is complete for each location that is reported.

Facilities are required to verify if no CLABSI events occurred for an inpatient care unit and month.

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*)

All CLABSI numerator and denominator data submitted to NHSN must be complete or the data submission is not accepted by NHSN. As a result there is no missing data for which distributions or other characteristics can be tested.

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*)

NHSN does not produce results pertaining to systematic missing data because the system requires that all data submissions include data used to calculate measure results.

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), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry), Other

If other: CLABSI and central line device days must be collected by trained hospital staff from information available in clinical data sources. The standard population's CLABSI rates are available from the NHSN Report. The NHSN analysis tool will automatically calculate SIRs. Some of the data used in the measure can be mined from electronic data sources.

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).

NHSN is moving towards an electronically captured CLABSI measure for future use. However, development and testing is not complete at this time.

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.

CLABSI surveillance in hospitals participating in CDC surveillance systems began in 1986, and the CLABSI measure has been endorsed by NQF in 2 measure sets since 2004. The criteria for CLABSI have been updated routinely to reflect user input and findings from data analysis, and changes have been made to decrease the difficulty and burden of data collection for the users. Such changes include removal of non-culture confirmed types of bloodstream infection from the types included in measure, addition of mucosal-barrier-injury laboratory confirmed bloodstream infection to identify infections which may not be preventable with previously identified CLABSI prevention tools, as well as reduction of data collection burden informed by data analysis (i.e. discontinuation of NICU central line type stratification).

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)*.

There are no fees to utilize NQF measure 0139. Participants must sign a Rules of Behavior document which states that they will follow the CLABSI surveillance protocol in its entirety and report data that is in accordance with this manual.

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 Hospital Inpatient Quality Reporting Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html The Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (PCHQR) Program http://www.qualitynet.org/dcs/ContentServer?pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772356060 Hospital Inpatient Quality Reporting Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html The Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (PCHQR) Program http://www.qualitynet.org/dcs/ContentServer?pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772356060 Public Health/Disease Surveillance National Healthcare Safety Network http://www.cdc.gov/nhsn/ Payment Program Hospital Inpatient Quality Reporting Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html The Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (PCHQR) Program http://www.qualitynet.org/dcs/ContentServer?pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772356060 Hospital Inpatient Quality Reporting Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html The Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (PCHQR) Program http://www.qualitynet.org/dcs/ContentServer?pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772356060 Regulatory and Accreditation Programs Hospital Inpatient Quality Reporting Program https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html The Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (PCHQR) Program http://www.qualitynet.org/dcs/ContentServer?pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772356060 Quality Improvement (Internal to the specific organization) Regulatory and Accreditation Programs https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalRHQDAPU.html The Prospective Payment System (PPS)-Exempt Cancer Hospital Quality Reporting (PCHQR) Program http://www.qualitynet.org/dcs/ContentServer?pagename=QnetPublic%2FPage%2FQnetTier2&cid=1228772356060

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

1) Name: Hospital Inpatient Quality Reporting Program (HIQR)
Sponsor: Centers for Medicare and Medicaid Services
Purpose: To improve health, improve care and lower cost (triple aims) of Medicare beneficiaries.
Geographic area and number and percentage of accountable entities and patients included: Nationwide, currently covers all acute care hospitals with ICUs (approximately 3300).*
Level of measurement and setting: Facility-Level, acute inpatient hospital
2) Name: Prospective Payment System Exempt Cancer Hospital Quality Reporting Program
Sponsor: Centers for Medicare and Medicaid Services
Purpose: To establish a quality reporting program for PPS-Exempt Cancer Hospital to improve health, improve care and lower cost (triple aims)of Medicare beneficiaries.
Geographic area and number and percentage of accountable entities and patients: 11 Patient Prospective Payment Exempt Cancer Hospitals in 7 U.S. states with 19,203 average discharges each in FY 2012*.
Level of measurement and setting: Facility-Level, PPS-Exempt cancer hospital
3) Name: Long Term Care Hospital (LTCH) Quality Reporting Program
Sponsor: Centers for Medicare and Medicaid Services
Purpose: To establish a quality reporting program for LTCHs to improve health, improve care and lower cost (triple aims) of Medicare beneficiaries.
Geographic area and number and percentage of accountable entities and patients included: All 442 Medicare certified long-term care hospitals are required to participate to receive 100% of reimbursement money due. In 2012, this included 202,050 patient discharges*.
Level of measurement and setting: Facility-Level, LTAC inpatient
4) Name: Hospital Value-Based Purchasing
Sponsor: Centers for Medicare and Medicaid Services
Purpose: To establish a quality reporting program to improve health, improve care and lower cost (triple aims) of Medicare beneficiaries.
Geographic area and number and percentage of accountable entities and patients included: 2808 entities*
Level of measurement and setting: Facility-Level, acute inpatient hospital
5) Name: Hospital-Acquired Condition Reduction Program (HACRP)
Sponsor: Centers for Medicare and Medicaid Services
Purpose: To establish a quality reporting program to improve health, improve care and lower cost (triple aims) of Medicare beneficiaries.
Geographic area and number and percentage of accountable entities and patients included: 3,216 entities*
Level of measurement and setting: Facility-Level, acute inpatient hospital
*provided by Centers for Medicare and Medicaid Services

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?*)
N/A

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.*)

N/A

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.

NHSN has developed numerous training resources to assist users with the proper understanding and interpretation of this measure. Several webinars and written training materials have been provided. Annual in-person trainings are held to discuss the SIR calculations, risk adjustment, and proper interpretation. Training materials are available online to all hospitals enrolled in NHSN, as well as external partners such as state health departments, quality improvement organizations, and healthcare corporations. NHSN users can run monthly analysis reports within NHSN to view their SIR data. On an annual basis, NHSN publishes national and state-level SIRs in the National and State HAI Progress Report.

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.

SIR results are available to NHSN users at any time, based on their current data entry. Data provided within the analysis report includes numerator, denominator, SIR, p-value, and 95% confidence interval. Educational materials are available on the NHSN website that explain each data element

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.

Feedback on measure performance and implementation is obtained via email to the NHSN helpdesk email system. Feedback is provided to us by hospital staff, physicians, epidemiologists, statisticians, state and local health department staff, quality improvement staff, and other personnel. An online survey is provided to all live-training attendees who provide feedback on whether objectives were met, usefulness of the training, and whether additional training is needed.

In 2019, NHSN is piloting an opportunity for facilities, groups and individuals to identify issues and areas for potential improvement for consideration as CDC updates and maintains the Bloodstream Infection (BSI) surveillance protocol for 2020. Comments may be submitted for consideration via the Federal Register, beginning Thursday February 14, 2019 through Monday April 15, 2019. This will be the format for submitting suggested modifications or comments regarding BSI surveillance for 2019.

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

Feedback from Hospitals and states: Based on results from a polling survey, hospitals have indicated that they are running SIR analysis reports within NHSN on a monthly basis, and that they use SIRs for prevention activities in their hospital. State health departments are using the SIR for public reporting purposes and to help target facilities for additional prevention. Feedback was received via email regarding the extent of risk adjustment and the limitations.

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

Feedback from consumers, media, policy, etc. on measure performance and implementation is obtained via email to the NHSN helpdesk email system. Feedback is provided to us by hospital staff, physicians, epidemiologists, statisticians, state and local health department staff, professional organizations, quality improvement staff, infection prevention and other personnel. See 4.a2.2.1.

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.

Protocol changes such as numerator and denominator exclusions which have been incorporated into the protocol are the result of feedback that is received from users. Feedback from all stakeholders is considered when developing and implementing the SIR. Different risk factor variables were analyzed for potential inclusion

in the statistical model due to input from users. Additional training formats, such as live chats and "quick learn" videos, were created in order to address different training environments that best meet the needs of our audience. We have also provided live demonstrations to users showing how to generate their SIRs in NHSN based on earlier feedback we had received.

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.

To a substantial extent the quality measure is a driver of patient care practices and particularly decisions on central line insertion. The trend data in section 1b. display the reductions in central line utilization over time and the reduction in the SIR for this measure before and after the 2015 rebaseline. Combined with declining SIRs, which change in relation to the number of CLABSIs per central line days, declines in the device (i.e. central line) use ratio on wards and especially neonatal intensive care units (NICUs) highlight the net benefit to patients afforded by both the safer and reduced central line use. Carefully determining the necessity of central lines before insertion is a CLABSI prevention strategy.

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.

Patient medical records and other sources of patient data must be reviewed to determine if the patient meets the necessary criteria for a healthcare-associated CLABSI. It is possible that reviewers may miss symptoms or fail to identify that patients meet criteria thereby under-reporting CLABSI events. Data collectors might also intentionally under-report CLABSIs. Both of these actions would result in an SIR that is calculated to be lower than actual. Alternatively, patients may be identified as having a CLABSI when in fact they do not meet CLABSI criteria and thereby calculate an SIR that is higher than actual. In addition, it is possible SIRs may be miscalculated. The NHSN reporting tool includes business logic to minimize misclassification of CLABSI and inaccurate reporting of catheter days. In addition, site visits can be conducted to audit data validity and this has been done for other infection types by some of the states using NHSN as their mandatory reporting tool (for example, see New York's audit process summary: http://www.health.state.ny.us/statistics/facilities/hospital/hospital_acquired_infections/2008/docs/hospital-acquired_infection.pdf, p20).

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

Surveillance for CLABSI uses the results of cultures of blood specimens. Suboptimal blood culture collection and testing technique can result in not only false-CLABSI reporting but also unnecessary antibiotic administration to patients. Unnecessary antibiotics can result in unnecessary adverse reactions, antibiotic resistance and Clostridioides difficile infection and its complications. Facilities may be motivated to assess and improve blood culture collection and testing techniques, to avoid identifying false-CLABSIs and in the process may prevent unnecessary patient complications.

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.

No

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

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

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

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.)

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.

Attachment **Attachment:** CDC_Master_Locations_and_Descriptions_2014-635218543933008267.docx

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Disease Control and Prevention

Co.2 Point of Contact: Daniel, Pollock, dpollock@cdc.gov, 404-639-4237-

Co.3 Measure Developer if different from Measure Steward: Centers for Disease Control and Prevention

Co.4 Point of Contact: Daniel, Pollock, dpollock@cdc.gov, 404-639-4237-

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 Healthcare Infection Control Practices Advisory Committee (HICPAC) consists of experts in the field of HAI surveillance, prevention, and control to provide advice and guidance to CDC. The measure was vetted through the technical panel of HICPAC that informed subsequent changes to measure development.

https://www.cdc.gov/hicpac/about.html

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2004

Ad.3 Month and Year of most recent revision: 12, 2013

Ad.4 What is your frequency for review/update of this measure? annually and when needed

Ad.5 When is the next scheduled review/update for this measure? 04, 2019

Ad.6 Copyright statement:

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: