**National Quality Forum—Measure Testing (subcriteria 2a2, 2b2-2b6)**

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

**Date of Submission**: 1/6/2014

**Type of Measure:**

|  |  |
| --- | --- |
| Composite – ***STOP – use composite testing form*** | Outcome (*including PRO-PM*) |
| Cost/resource | Process |
| 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, 2b2, 2b3, and 2b5 must be completed.** * **For outcome and resource use measures**, section **2b4** also must be completed. * If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** 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 (2b2-2b6) 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 20 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](http://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx). |

|  |
| --- |
| **Note: The information provided in this form is intended to aid the Steering 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**](#Note10) 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.  **2b2.** **Validity testing** [**11**](#Note11) 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.    **2b3.** Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [**12**](#Note12)  **AND**  If patient preference (e.g., informed decision making) 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**](#Note13)  **2b4.** **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 that influence the measured outcome (but not factors related to disparities in care or the quality of care) and are present at start of care; [**14**](#Note14)**,**[**15**](#Note15) and has demonstrated adequate discrimination and calibration   **OR**   * rationale/data support no risk adjustment/ stratification.   **2b5.** 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**](#Note16) **differences in performance**;  **OR**  there is evidence of overall less-than-optimal performance.  **2b6.** **If multiple data sources/methods are specified, there is demonstration they produce comparable results**.  **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.  **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.** Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care, such as race, socioeconomic status, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.  **16.** 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.23*)** | **Measure Tested with Data From:** |
| abstracted from paper record | abstracted from paper record |
| administrative claims | administrative claims |
| clinical database/registry | clinical database/registry |
| abstracted from electronic health record | abstracted from electronic health record |
| eMeasure (HQMF) implemented in EHRs | eMeasure (HQMF) implemented in EHRs |
| other: Click here to describe | other: Click here to describe |

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

Medicare Limited Data Set Carrier SAF – 5% File

Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (NACOR)

**1.3. What are the dates of the data used in testing**? Medicare Limited Data Set Carrier SAF – 5% File (2008-2012); Anesthesia Quality Institute (AQI) National Anesthesia Clinical Outcomes Registry (2010-2013)

**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.26*)** | **Measure Tested at Level of:** |
| individual clinician | individual clinician |
| group/practice | group/practice |
| hospital/facility/agency | hospital/facility/agency |
| health plan | health plan |
| other: Click here to describe | other: Click here to describe |

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

The National Anesthesia Clinical Outcomes Registry (NACOR) is a very large sample of all anesthesia care in the United States. NACOR currently collects data from more than 250 anesthesia practices, representing 21,000 providers (12,000 MDs and 9,000 nurse anesthetists) or about 25% of the national work force. Data are reported from about 2,100 healthcare facilities, of which about half are university or community hospitals that would expect to place CVCs for fluid resuscitation, monitoring and medication delivery. NACOR has captured 14,000,000 cases since 2010, of which 220,000 are CVC insertions in one of the CPT code groups in this measure’s denominator.

Because participation in NACOR is voluntary, it is possible that biases exist in which practices and hospitals are represented. Regardless, the measure was tested at the clinician, group, and facility level, the entities included in the NACOR analysis were:

|  |  |
| --- | --- |
| **NACOR Analysis Per Entity (2010-2013)** | |
| **Entity Reporting** | **Total Number** |
| Practices | 157 |
| Facilities | 595 |
| Providers | 9,981 |

Analysis was also conducted using the Medicare Limited Data Set Carrier SAF 5% File.

**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*)

NACOR data has been presented in the tables below, including distribution of 220,000 CVC cases by age, sex, ASA Physical Status (a common measure of patient risk in anesthesia procedures) and average household income of their reported zip code.

**Total Cases Captured (NACOR):**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Year** | **Total number of CVC placements** | **Number (% of column B) of CVC procedures in which Medicare is the primary insurance** | **Number (% of column B) meeting all measure criteria** | **Number (% of column C) meeting all measure criteria in which Medicare is the primary insurance** |
| Cases | 226,092 (100%) | 71,873 (32%) | 8,003 (3.54%) | 6,223 (8.66%) |
| 2010 | 49,422 (100%) | 14,093 (28%) | 1,586 (3.21%) | 1,078 (7.65%) |
| 2011 | 61,865 (100%) | 20,464 (33%) | 2,430 (3.93%) | 1,641 (8.02%) |
| 2012 | 66,366 (100%) | 22,439 (34%) | 2,549 (3.84%) | 2,209 (9.84%) |
| 2013, through September | 48,439 (100%) | 14,877 (31%) | 1,438 (2.97%) | 1,295 (8.7%) |

**Cases Captured by Age:**

| **Age Group** | **Total number of CVC placements** | **Number (% of column total) of CVC procedures in which Medicare is the primary insurance** | **Number (% of column total) meeting all measure criteria** | **Number (% of column total) meeting all measure criteria in which Medicare is the primary insurance** |
| --- | --- | --- | --- | --- |
| TOTAL | 226,092 (100%) | 71,873 (100%) | 8,003 (100%) | 6,223 (100%) |
| < 1 | 4,946 (2.19%) | 4 (0.01%) | (0%) | (0%) |
| 1 - 18 | 10,640 (4.71%) | 111 (0.15%) | 9 (0.11%) | 8 (0.13%) |
| 19 - 49 | 34,277 (15.16%) | 3,906 (5.43%) | 367 (4.59%) | 278 (4.47%) |
| 50 - 64 | 67,098 (29.68%) | 8,624 (12%) | 967 (12.08%) | 651 (10.46%) |
| 65 - 79 | 70,350 (31.12%) | 38,771 (53.94%) | 4,540 (56.73%) | 3,513 (56.45%) |
| 80+ | 20,100 (8.89%) | 12,383 (17.23%) | 1,982 (24.77%) | 1,642 (26.39%) |
| Unknown | 18,681 (8.26%) | 8,074 (11.23%) | 138 (1.72%) | 131 (2.11%) |

**Cases Captured by Gender:**

| **Gender** | **Total number of CVC placements** | **Number (% of column total) of CVC procedures in which Medicare is the primary insurance** | **Number (% of column total) meeting all measure criteria** | **Number (% of column total) meeting all measure criteria in which Medicare is the primary insurance** |
| --- | --- | --- | --- | --- |
| TOTAL | 226,092 (100%) | 71,873 (100%) | 8,003 (100%) | 6,223 (100%) |
| Female | 99,923 (44.2%) | 31,127 (43.31%) | 3,283 (41.02%) | 2,587 (41.57%) |
| Male | 107,546 (47.57%) | 33,176 (46.16%) | 4,582 (57.25%) | 3,505 (56.32%) |
| Unknown | 18,623 (8.24%) | 7,570 (10.53%) | 138 (1.72%) | 131 (2.11%) |

**Cases Captured by ASA Physical Status:**

| **ASA Physical Status** | **Total number of CVC placements** | **Number (% of column total) of CVC procedures in which Medicare is the primary insurance** | **Number (% of column total) meeting all measure criteria** | **Number (% of column total) meeting all measure criteria in which Medicare is the primary insurance** |
| --- | --- | --- | --- | --- |
| TOTAL | 177,361 (100%) | 60,323 (100%) | 7,438 (100%) | 6,096 (100%) |
| I - II | 62,714 (35.36%) | 22,199 (36.8%) | 605 (8.13%) | 544 (8.92%) |
| III | 62,346 (35.15%) | 19,822 (32.86%) | 2,564 (34.47%) | 2,331 (38.24%) |
| IV | 50,947 (28.73%) | 17,902 (29.68%) | 4,160 (55.93%) | 3,136 (51.44%) |
| V | 1,290 (0.73%) | 390 (0.65%) | 109 (1.47%) | 85 (1.39%) |
| VI | 64 (0.04%) | 10 (0.02%) | (0%) | (0%) |

**Cases Captured by Mean Income (estimated from patient ZIP code):**

| **Mean Income** | **Total number of CVC placements** | **Number (% of column total) of CVC procedures in which Medicare is the primary insurance** | **Number (% of column total) meeting all measure criteria** | **Number (% of column total) meeting all measure criteria in which Medicare is the primary insurance** |
| --- | --- | --- | --- | --- |
| TOTAL | 226,092 (100%) | 71,873 (100%) | 8,003 (100%) | 6,223 (100%) |
| < $20,000 | 391 (0.17%) | 132 (0.18%) | 13 (0.16%) | 11 (0.18%) |
| $20,000 - $39,999 | 13,147 (5.81%) | 4,246 (5.91%) | 264 (3.3%) | 183 (2.94%) |
| $40,000 - $59,999 | 84,444 (37.35%) | 28,986 (40.33%) | 2,517 (31.45%) | 1,647 (26.47%) |
| $60,000 - $99,999 | 86,536 (38.27%) | 27,900 (38.82%) | 3,757 (46.94%) | 3,077 (49.45%) |
| $100,000 + | 21,447 (9.49%) | 7,426 (10.33%) | 1,305 (16.31%) | 1,204 (19.35%) |
| Unknown | 20,127 (8.9%) | 3,183 (4.43%) | 147 (1.84%) | 101 (1.62%) |

**CVC Placement by Region:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **US Region** | **Total number of Cases** | **Count Medicare** | **Count Measure Met** | **Count Medicare and Measure Met** | **Count Exempt** | **Count Medicare and Exempt** | **Count Exempt Not Specified** | **Count Medicare and Exempt Not Specified** |
| Region 4: West | 37,810 | 12,958 | 2,438 | 2,384 | 7 | 2 | 38 | 18 |
| Region 3: South | 95,207 | 31,606 | 2,057 | 1,394 | 30 | 11 | 318 | 58 |
| Region 1: Northeast | 18,397 | 4,696 | 809 | 326 | 7 | 3 | 111 | 105 |
| Region 2: Midwest | 56,868 | 20,150 | 2,642 | 2,087 | 53 | 23 | 158 | 120 |
| Unknown | 9 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |

**CVC Placement by Anesthesia Practice:**

Using NACOR data, CVC placements by anesthesia practice:

Min = 1

Max = 14,597

Mean = 1,440.08

Median = 98

1st Quartile = 18

3rd Quartile = 381.5

The following chart displays CVC placements by anesthesia practice:

**CVC Placement by Facility:**

Using NACOR data, CVC placements by facility:

Min = 1

Max = 12,619

Mean = 402.44

Median = 98

1st Quartile = 18

3rd Quartile = 381.5

The following chart displays CVC placements by facility:

Using NACOR data, CVC placements by facility type:

Small Community Hospital (<100 beds) = 3,706

Other = 16,269

University Hospital = 38,595

Large Community Hospital (500+ beds) = 57,113

Median Community Hospital (100-500 beds) = 83,611

Average = 39,858.8

The following chart displays CVC placements by facility type:

**CVC Placement by Provider:**

Using NACOR data, CVC placements by provider:

Min = 1

Max = 1,931

Mean = 29

Median = 74

1st Quartile = 3

3rd Quartile = 27

For the Medicare Limited Data Set Carrier SAF 5% File, 13,355 patients were analyzed in 2008, 12,985 patients were analyzed in 2009, 12,250 patients were analyzed in 2010, 11,515 patients were analyzed in 2011 and 10,212 patients were analyzed in 2012.

**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**.

The NACOR data and the CMS data differ in average patient age; this difference can be assessed in the tables provided above, which also include the Medicare subset of the NACOR population. This fundamental difference likely colors other assessments as well: more female than male patients are alive and in the healthcare system at later ages, and the case mix between older and younger patients is different as well.

Neither NACOR nor CMS data can accurately determine the indication for CVC placement in any individual cases. Generic reasons for anesthesiologists to place these lines include the need for reliable intravenous access in patients who will require large volumes of fluid, the need for hemodynamic monitoring in patients undergoing high risk procedures, the need for vascular access to facilitate laboratory assay of the blood, and the need for central delivery of dangerous medications. The majority of CVC placements in NACOR (> 90%) were in association with anesthesia for an associated surgical procedure.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**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*)

At present, the only measure testing which has occurred is gross comparison of results in the CMS and NACOR data sets. As indicated above, this comparison has revealed substantial effort is still needed to ensure that all anesthesiologists who are correctly documenting performance at the local level can transmit this information to NACOR, where national-level aggregation and validation will be possible.

**2a2.3. For each level 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*)

The question is not applicable since the measure score confirms accurate translation of the provider’s documentation into a digital capture system. Data from the Medicare Limited Data Set Carrier SAF – 5% File and NACOR confirm that the measure score is precise.

**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?*)

This measure is not reliably reported by all eligible professionals. Early development of the NACOR model suggests that it is possible to document and report performance in practices with the appropriate information technology or on-site abstraction in place, and that universal reporting is therefore an attainable goal.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b2. VALIDITY TESTING**

**2b2.1. 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*)

**2b2.2. For each level 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)*

**CRITICAL DATA ELEMENTS**

NACOR data is imported directly from existing electronic sources. As data files are translated into the system they undergo a series of quality checks, as well as both automated and manual review. Key metrics are compared to previous reports from the same practice as well as to benchmarks established in the broader registry, and outliers are identified and followed up with the sending practice.

While these steps can insure the integrity of electronic transmission of existing information, they do not validate the primary creation of those data or their abstraction into a billing or quality management system. The first step is inherently challenging, as performance on this measure is typically documented by a checkbox in the anesthesia or surgical record, either on paper or in an electronic form. Validation would require direct observation of the procedure itself (or review of video images), and will be prohibitively expensive and intrusive in all except focused research studies. Even then, direct observation may strongly bias performance. Like many measures, this one will depend for some time on attestation of performance by the providers involved.

The second stage of validation, confirming accurate translation of the provider’s documentation into a digital capture system, is commonly assessed by retrospective random audits which compare the local medical record to the information captured in the national registry. NACOR has provisions in place for just such auditing, which will begin in the coming year. NACOR audits may take on a different form than those conducted by older registries, as NACOR is designed to move electronic data directly from the medical record to the registry, without the need for human abstractors.

**PERFORMANCE MEASURE SCORE – SYSTEMATIC REVIEW OF FACE VALIDITY**

As part of the measure maintenance for NQF #0464, American Society of Anesthesiologists solicited a Measure Expert Panel (MEP) that consisted of 37 experts to perform face validity testing on the measure. The names of MEP members have been provided in application question Ad.2.

Face validity of the measure score as an indicator of quality was systematically assessed as follows:

The MEP reviewed NQF #0464 measure specifications and members were asked to rate their agreement with the following statement: “The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality.”

MEP members were able to rate their agreement on a five-point rating scale: Strongly Disagree, Disagree, Neither Agree nor Disagree, Agree, Strongly Agree.

**2b2.3. What were the statistical results from validity testing**? (*e.g., correlation; t-test*)

**PERFORMANCE MEASURE SCORE – VALIDITY TESTING**

Using the Medicare Limited Data Set Carrier SAF – 5% File for 2012, a t-test was conducted to see if there was a significant risk between those that reported the insertion protocol versus that did not submit the protocol. Those with the CRBSI protocol reported were slightly older, less likely to have End Stage Renal Disease, and less likely to be dead. Gender and race were not different between the two groups.

**Patients with CVC Insertion: Comparison of Mean Characteristics  
for Those With and Without CRBSI Protocol Reported**

Based on Medicare 5% Sample Claims Files

**2012**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Mean or Proportion** | | |  |
| **Patient Characteristic** | **All CVC**  **Insertions** | **CRBSI Protocol Reported** | **CRBSI Protocol Not Reported** | ***P*-Value** |
| Age (years) | 71.1 | 71.4 | 70.8 | < .01 |
| Proportion: |  |  |  |  |
| Male | .558 | .559 | .556 | .75 |
| Non-White | .158 | .155 | .161 | .40 |
| ESRD | .078 | .072 | .083 | .03 |
| Died | .187 | .176 | .198 | < .01 |

**PERFORMANCE MEASURE SCORE – SYSTEMATIC REVIEW OF FACE VALIDITY**

The results of the face validity assessment indicate that members had high levels of agreement with the statement: “The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality.”

The results of the expert panel rating of the validity statement were as follows: N = 37; Mean rating = 4.21 and 86% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality.

|  |  |
| --- | --- |
| Rating Scale | Number who selected the rating |
| 1 – Strongly Disagree | 0 |
| 2 – Disagree | 4 |
| 3 – Neither Agree nor Disagree | 1 |
| 4 – Agree | 15 |
| 5 – Strongly Agree | 17 |
| Total Members: | 37 |
| Mean Score: | 4.21 |

**2b2.4. 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 t-test, using the Medicare Limited Data Set Carrier SAF – 5% File for 2012, demonstrated that those with the CRBSI protocol reported were slightly older, less likely to have End Stage Renal Disease, and less likely to be dead. Gender and race were not different between the two groups.

The results of the face validity assessment indicate that out of 37 experts, over 85% (32) agreed that the measure provides an accurate reflection of quality and can be used to distinguish good and poor quality.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b3. EXCLUSIONS ANALYSIS**

**NA**  **no exclusions — *skip to section*** [***2b4***](#section2b4)

**2b3.1. 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*)

NQF #0464, as specified in NQF and the PCPI Anesthesiology and Critical Workgroup, instructs for the patient with appropriate medical reason denominator exception criteria, report 6030F with modifier 1P; modifier 2P and 3P may not be used.

**2b3.2. 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*)

Using the Medicare Limited Data Set Carrier SAF – 5% File produced the following results:

**CVC Insertions and CRBSI Protocol Reported and Followed**

Based on Medicare 5% Sample Claims Files

**2008 – 2012**

**Modifier 1P as Exception**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Indicator** | **2008** | **2009** | **2010** | **2011** | **2012** |
| CVC Insertions | 15,248 | 14,718 | 13,722 | 12,971 | 11,372 |
| Protocol Reported1 | 1,231 | 3,512 | 5,221 | 5,945 | 5,508 |
| Percentage reported | 8.1% | 23.9% | 38.0% | 45.8% | 48.4% |
| Followed | 1,223 | 3,471 | 5,159 | 5,895 | 5,493 |
| Percentage followed | 99.4% | 98.8% | 98.8% | 99.2% | 99.7% |
|  |  |  |  |  |  |
| Patients in Sample | 13,355 | 12,985 | 12,250 | 11,513 | 10,212 |
| Estimated total Medicare | 267,100 | 259,700 | 245,000 | 230,260 | 204,240 |
| Percentage of patients  with any claims in year | 0.80% | 0.77% | 0.72% | 0.67% | 0.63% |

1CRBSI Protocol reported indicated by CPT 6030F.

Using NACOR data produced the following results:

**CVC Insertions and CRBSI Protocol Reported and Followed**

Based on NACOR data (all cases)

**2010 – 2013**

**Modifier 1P as Exception**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Indicator** | **2010** | **2011** | **2012** | **2013 (to 9/30)** |
| CVC Insertions | 49,422 | 61,865 | 66,366 | 48,439 |
| Protocol Reported1 | 1,586 | 2,430 | 2,549 | 1,438 |
| Percentage reported | 3.21% | 3.93% | 3.84% | 2.97% |
| Followed | 1,577 | 2,384 | 2,515 | 1,428 |
| Percentage followed | 99.4% | 98.1% | 98.7% | 99.3% |

----/30/2013)ollowing results.dave End-Stage Renal Disease, and less likely to be dead. Gender and race were not different bet

**2b3.3. 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*)

Systematic exclusions to this measure are not necessary; evidence-based sterile precautions are indicated for every CVC insertion in clinical practice. Exceptions, based largely on the acuity of need for the CVC in trauma or emergency cases, are most appropriately documented on an individual basis, as provided for in the measure specifications.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b4. 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***](#section2b5)***.***

**2b4.1. What method of controlling for differences in case mix is used?**

**No risk adjustment or stratification**

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

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

**Other,** Click here to enter description

**2b4.2. If an outcome or resource use 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**.

**2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient 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 and not related to disparities*)

**2b4.4. What were the statistical results of the analyses used to select risk factors?**

**2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach** (*describe the steps―do not just name a method; what statistical analysis was used*)

*Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below*.  
***if stratified, skip to*** [***2b4.9***](#question2b49)

**2b4.6. Statistical Risk Model Discrimination Statistics** (*e.g., c-statistic, R-squared*)**:**

**2b4.7. Statistical Risk Model Calibration Statistics** (*e.g., Hosmer-Lemeshow statistic*):

**2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves**:

**2b4.9. Results of Risk Stratification Analysis**:

**2b4.10. 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*)

\***2b4.11.** **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*)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE**

**2b5.1. 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)*

NACOR is working to report performance on this measure to participating groups on the practice, facility and individual provider level, showing both trends over time and comparison to national benchmarks. A preliminary review of the data shows variation in performance at the individual level, even within the same practice, but it is not known if these differences represent failure of compliance with the standards or failure to document. The purpose of the NACOR reports will be to provide data for the local quality managers to investigate and address these discrepancies. Anesthesia practices which have engaged in this sort of private reporting and local improvement have demonstrated strong gains in performance.

One open question for quality measurement in this area is the number of CVC insertions required to demonstrate meaningful variations in performance at the level of individual providers. At the facility and practice level this will not be an issue: CVC placement is a common-enough procedure that team-based reporting should generate adequate numbers for performance assessment.

**2b5.2. 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*)

To date, assessment of performance on this measure remains incomplete. The issue at present is to increase the number of providers reporting, and to improve the mechanisms which bring this data to NACOR. Facilitating this sort of data collection and analysis was the American Society of Anesthesiologist’s purpose in creating the Anesthesia Quality Institute and launching NACOR. While enormous progress has occurred, this project is still in its infancy. Taking on aggregation and reporting of this measure is one early step.

**2b5.3. 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 number of CVCs placed by anesthesia professionals in a year is certainly large enough to support meaningful performance assessment, once mechanisms are in place to facilitate provider documentation, electronic abstraction, and transmission to NACOR or similar national aggregators.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS**

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

**Note***: This criterion is directed 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).* ***If comparability is not demonstrated, the different specifications should be submitted as separate measures.***

**2b6.1. Describe the method of testing conducted to demonstrate comparability of performance scores for the same entities across the different datasources/specifications** (*describe the steps―do not just name a method; what statistical analysis was used*)

**2b6.2. 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*)

**2b6.3. What is your interpretation of the results in terms of demonstrating comparability of 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*)