NATIONAL QUALITY FORUM

Proposed NQF policy on "Inactive Measures"

TO: Surgery Endorsement Maintenance Steering Committee

FR: Melinda Murphy, MS, RN and Alexis Forman, MPH

SU: Proposed NQF policy on "inactive measures"

DA: April 26, 2011

At the February 28 – March 1, 2011 meeting, the Steering Committee discussed potential for retiring measures when current performance is very high and there seems to be little opportunity for improved performance. Such measures have been successful in driving improvement in performance, but concerns have been raised about possible decline in performance if measurement is discontinued. For measures that otherwise meet all NQF endorsement criteria, NQF is considering approaches to address this issue including the potential of establishing a category of "inactive endorsement" so that performance could be monitored in the future if necessary to ensure that performance does not decline. This status would apply only to highly credible, reliable, and valid measures that have high levels of performance due to quality improvement actions (often facilitated or motivated through public reporting and pay-for-performance programs) rather than problems with the measure specifications. The key issue is the opportunity cost associated with measuring processes at high levels of performance—rather than focusing on areas where there is really a gap in care.

NQF does not want to move measures into inactive status that are really not needed because they are too far from the desired outcome. The proposed policy is open for public comment and NQF's Board of Directors will consider this policy at their meeting in May. In anticipation of this policy, the Steering Committee should determine whether measures it might recommend for retirement could be recommended for inactive endorsement instead.

CONSIDERATIONS FOR INACTIVE ENDORSEMENT STATUS

The data provided in measure submissions are frequently limited. In determining whether there is further opportunity for improvement, the Steering Committee should review data on representation, variation, and disparities:

- What is the representativeness of the data, i.e., is it national data from a majority of hospitals or is the data from a single state or payer group?
- What is the range in performance, particularly in the lowest decile or quartile?
- What is the performance among possible disparities population(s)?
- Is the measure performance data indicating high levels of performance consistent with other evidence (epidemiologic or research)?

- What is the size of the population at risk, effectiveness of an intervention, and consequences of a quality problem (e.g., even modest variation would be sufficient justification for some highly effective, potentially life-saving treatments)?

Other considerations include:

- Is this a measure with strong, direct evidence of a link to a desired health outcome? Generally measures more distal to the desired outcome with only indirect evidence would not qualify, e.g., assessment of blood pressure (BP) measurement rather than the BP value.
- Measures with a focus more distal to a desired outcome are not needed if there is a measure with a focus more proximal to the desired outcome (e.g., venous thromboembolism (VTE) prophylaxis ordered versus VTE prophylaxis administered).
- Is the measure needed if outcomes (i.e., mortality, readmission) of care are being measured?

In summary, the Steering Committee will have the opportunity to consider the questions above and make recommendations to:

- Continue endorsement
- Inactive endorsement
- Remove endorsement

PHASE I MEASURE FOR POSSIBLE INACTIVE ENDORSEMENT STATUS

The Steering Committee preliminarily has recommended one measure from Phase I that may be a candidate for inactive endorsement status. The Committee noted that the following measure may be "topped out" due to high performance and lack of opportunity for improvement:

- 0129 Risk-adjusted prolonged intubation (ventilation) (STS)

DATA ON OPPORTUNITY FOR IMPROVEMENT

0129 Risk-adjusted prolonged intubation (ventilation)

National performance rates: With 1.0 as the median, the data shows an incidence range from 0.3 to 3.4 with 1.4 and 0.7 at the 25^{th} and 75^{th} quartiles respectively.

- Representative: Analysis includes 640 STS Adult Cardiac Surgery Database participants who had at least 100 eligible cases for the measure and reported data (not restricted to this measure) to the STS Adult Cardiac Surgery Database for all 12 months.
  Dates: January 1, 2009-December 31, 2009
- Disparities: See STS disparities tables attached.

COMPLETE MEASURE EVALUATION

ACTION ITEM: Measures under consideration for inactive endorsement must meet all criteria. The measure submission forms are provided again for your reference. The preliminary ratings from the Committee are provided below:

0129 Risk-adjusted prolonged intubation (ventilation)

Description: Percent of patients aged 18 years and older undergoing isolated CABG who require intubation for more than 24 hours.
Numerator Statement: Number of patients undergoing isolated CABG who require intubation > 24 hours.
Denominator Statement: All patients undergoing isolated CABG.
Exclusions: N/A
Adjustment/Stratification: case-mix adjustment No stratification is required for this measure.
Level of Analysis: Clinicians: Group; Facility/Agency; Population: National, regional/network, states, counties or cities
Type of Measure: Outcome
Data Source: Registry data-STS Adult Cardiac Surgery Database, Version 2.73
Measure Steward: Society of Thoracic Surgeons | 633 North Saint Clair Street, Suite 2320 | Chicago | Illinois | 60611

Steering Committee Recommendation for Endorsement: Conditional Y-21; N-1; A-0
Rationale: Intubation is linked to morbidity, and an increase in length-of-stay, cost and resource utilization.

If applicable, Conditions/Questions for Developer:

1. De.2 Measure Description: Please consider change in time limit to a period that is less than 24 hours
2. 1b.4 Summary of Data on Disparities by Population Group: Please provide data on disparities.

Developer Response:

1. Considering the increased complexity of current CT patients, a time period significantly less than 24 hrs (e.g. 6 or 12 hours) would not be appropriate as a *routine performance measure*, even though that is achievable in many patients. In some patients, such a measure could result in the adverse unintended consequences of premature extubation, subsequent ventilatory failure, and re-intubation.
2. Data on disparities are provided in the form.

If applicable, Questions to the Steering Committee:

1. Importance to Measure and Report: Y-22; N-0
(1a. Impact; 1b. Performance gap; 1c. Outcome or Evidence)
Rationale: Although the measure compliance is above 90 percent, the Committee felt compliance should be closer to 100 percent.

2. Scientific Acceptability of Measure Properties: C-17; P-5; M-0; N-0
(2a. Precise specifications; 2b. Reliability testing; 2c. Validity testing; 2d. Exclusions justified; 2e. Risk adjustment/stratification; 2f. Meaningful differences; 2g. Comparability; 2h. Disparities)
Rationale: One potential confounder is the post-CABG patient who is extubatable by clinical criteria but is kept intubated beyond 24 hours due an unrelated unscheduled second surgery the next day. The Committee questioned the developer as to why 24 hours was selected as the standard as opposed to a shorter time period. The literature identifies a range of times, associated with length of stay in ICU and hospital as well as relationship to anesthesia. One study reported that 39 percent of all patients were extubated within 6 hours, 89 percent within 24 hours and 95 percent within 48 hours. Committee members indicated that in their experience the majority of patients are off ventilators sooner than 24 hours.

3. Usability: C-20; P-2; M-0; N-0
(3a. Meaningful/useful for public reporting and quality improvement; 3b. Harmonized; 3c. Distinctive or additive value to existing measures)
Rationale: The measure is meaningful for public reporting and quality improvement.

4. Feasibility: C-20; P-1; M-1; N-0
(4a. Clinical data generated during care process; 4b. Electronic sources; 4c. Exclusions – no additional data source; 4d. Susceptibility to inaccuracies/ unintended consequences identified 4e. Data collection strategy can be implemented)
Rationale: Easily captured and derived from electronic sources.

2a.14. Risk Adjustment Methodology/Variables *(List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method)*

The risk adjusted model is a hierarchical logistic regression model with participant level intercept.

$$\text{logit(outcome)} \sim X\beta + (\gamma \mid \text{participant})$$

where X is the patient's risk factors, β is the regression coefficients of patient-level risk factors and γ is the participant level regression coefficient.

Inclusion Criteria

The patient level risk adjusted model was developed using a population of patients undergoing isolated CABG procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest two and a half years of data (January 2008 – June 2010) from the STS Adult Cardiac Surgery Database.

Variable Definitions and Selection

All variables for consideration are listed in the table below.

Definition of Variables Appearing in STS 2008 CABG Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age	= Patient age in years
Age function 1	= max (age–50, 0)
Age function 2	= max (age–60, 0)
Age by reop function	= Age function 1 if surgery is a reoperation, = 0 otherwise
Age by status function	= Age function 1 if status is emergent or salvage, = 0 otherwise
BSA function 1	= max (1.4, min [2.6, BSA]) – 1.8
BSA function 2	= (BSA function 1)^2
CHF but not NYHA IV	= 1 if patient has CHF and is not NYHA class IV, = 0 otherwise
CHF and NYHA IV	= 1 if patient has CHF and is NYHA class IV, = 0 otherwise
CLD mild	= 1 if patient has mild chronic lung disease, = 0 otherwise
CLD moderate	= 1 if patient has moderate chronic lung disease, = 0 otherwise
CLD severe	= 1 if patient has severe chronic lung disease, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine function 2	= max ([creatinine function 1] – 1.0, 0)
Creatinine function 3	= max ([creatinine function 1] – 1.5, 0)
CVD without prior CVA	= 1 if patient has history of CVD and no prior CVA, = 0 otherwise
CVD and prior CVA	= 1 if patient has history of CVD and a prior CVA, = 0 otherwise
Diabetes, noninsulin	= 1 if patient has diabetes not treated with insulin, = 0 otherwise
Diabetes, insulin	= 1 if patient has diabetes treated with insulin, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
Female	= 1 if patient is female, = 0 otherwise
Female by BSA function 1	= BSA function 1 if female, = 0 otherwise
Female by BSA function 2	= BSA function 2 if female, = 0 otherwise
Hypertension	= 1 if patient has hypertension, = 0 otherwise
IABP or inotropes	= 1 if patient requires IABP or inotropes preoperatively, = 0 otherwise
Immunosuppressive treatment	= 1 if patient given immunosuppressive therapy within 30 days, = 0 otherwise
Insufficiency, aortic	= 1 if patient has at least moderate aortic insufficiency, = 0 otherwise
Insufficiency, mitral	= 1 if patient has at least moderate mitral insufficiency, = 0 otherwise
Insufficiency, tricuspid	= 1 if patient has at least moderate tricuspid insufficiency, = 0 otherwise
Left main disease	= 1 if patient has left main disease, = 0 otherwise
MI 1 to 21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI > 6 and < 24 hours	= 1 if history of MI >6 and <24 hours prior to surgery, = 0 otherwise
MI ≤ 6 hours	= 1 if history of MI ≤ 6 hours prior to surgery, = 0 otherwise

No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 0 otherwise
PCI ≤ 6 hours	= 1 if patient had PCI ≤ 6 hours prior to surgery, = 0 otherwise
Peripheral vascular disease	= 1 if patient has peripheral vascular disease, = 0 otherwise
Race black	= 1 if patient is black, = 0 otherwise
Race Hispanic	= 1 if patient is nonblack Hispanic, = 0 otherwise
Race Asian	= 1 if patient is nonblack, non-Hispanic, and is Asian, = 0 otherwise
Reop, 1 previous operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 previous operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not resuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina, no MI within 7 days of surgery, = 0 otherwise

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident, or stroke; CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PLOS = prolonged length of stay; Reop = reoperation; Comp = composite adverse event (any); RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

The final patient-level model was built by backward selection method with several variables forced into the model. For the final patient-level model, please see the attachment.

1b.2. Summary of Measure Results Demonstrating Performance Gap *(Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.)*

The summary statistic provided is the Participant's Estimated Odds Ratio (OR) based on a hierarchical logistic regression analysis. The OR measures the impact that a participant's performance level has on a patient's probability of experiencing an adverse outcome. An OR greater than 1.0 implies that the participant increases a patient's risk of experiencing the outcome, relative to an "average" STS participant. An OR less than 1.0 implies that the participant decreases a patient's risk of experiencing the outcome, relative to an "average" STS participant. A high OR is undesirable and we define the percentiles with decreasing OR. For example, 90% of STS participants have an OR greater than the value indicated by the "90th percentile" below.

Measurement	*Prolonged Intubation (ventilation)*
N	640
Mean	1.1
1st	3.4
5th	2.3
10th	1.9
25th	1.4
Median	1.0
75th	0.7
90th	0.5
95th	0.4
99th	0.3
Outlier	270 (42.2)
High	120
Low	150

Also provided is the distribution of the risk adjusted event rate (see below). The risk adjusted rate is an estimate of the participant's event rate if, hypothetically, the case-mix of the patients treated by the participants is the same as the overall STS case-mix. It is calculated by the OR of the participant, other patient level parameter estimates from the hierarchical logistic model, and the overall STS event rate, by:

STS event rate * (Participant's Expected Event Rate) / (Participant's Expected Event Rate Assuming Its Performance = STS Average Performance)

In the above equation, "Participant's Expected Event Rate" is calculated with the participant's actual OR, and "Participant's Expected Event Rate Assuming Its Performance = STS Average Performance" is calculated by assuming the participant's OR = 1 (*i.e.* no difference in performance from the STS average).

Measurement	*Prolonged Intubation (ventilation)*
N	640
Mean	11.5
1st	3.5
5th	5.0
10th	5.9

Measurement	*Prolonged Intubation (ventilation)*
25th	7.8
Median	10.6
75th	14.2
90th	18.0
95th	20.4
99th	26.4
Outlier	270 (42.2)
High	120
Low	150

1b.4. Summary of Measure Results on Disparities by Population Group *(Descriptive statistics for performance results for this measure by population group)*

	Prolonged Intubation (ventilation) - Risk Adjusted Rate	
	Population Group	
	Men	*Women*
Measurement		
N	897	653
Mean	10.9	14.1
1st	3.5	5.3
5th	4.8	7.0
10th	5.6	8.1
25th	7.4	10.2
Median	9.8	13.1
75th	13.6	16.9
90th	17.3	21.3
95th	19.3	24.8
99th	26.8	32.1
Outlier	380 (42.4%)	176 (27.0%)
High	172	76
Low	208	100

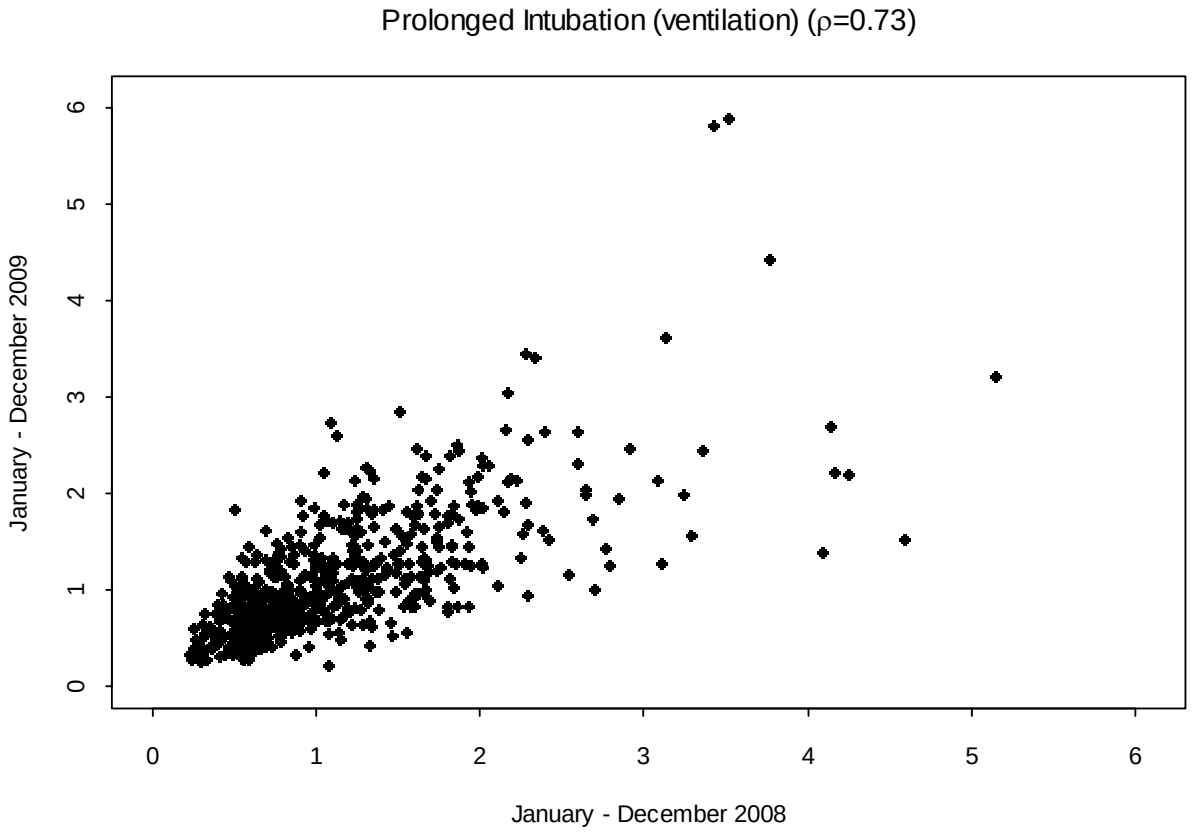
	Prolonged Intubation (ventilation) - Risk Adjusted Rate		
	Population Group		
	Black	*White*	*Other*
Measurement			
N	135	887	121
Mean	15.8	11.3	13.5
1st	6.7	3.6	5.4
5th	8.6	4.9	5.9
10th	9.2	5.9	7.2
25th	12.0	7.7	9.3
Median	14.8	10.3	13.2
75th	19.0	14.2	16.1
90th	23.3	18.0	19.5
95th	24.7	20.3	23.2
99th	34.1	26.9	30.7

	Prolonged Intubation (ventilation) - Risk Adjusted Rate		
	Population Group		
	Black	*White*	*Other*
Measurement			
Outlier	30 (22.2%)	415 (46.8%)	30 (24.8%)
High	13	193	16
Low	17	222	14

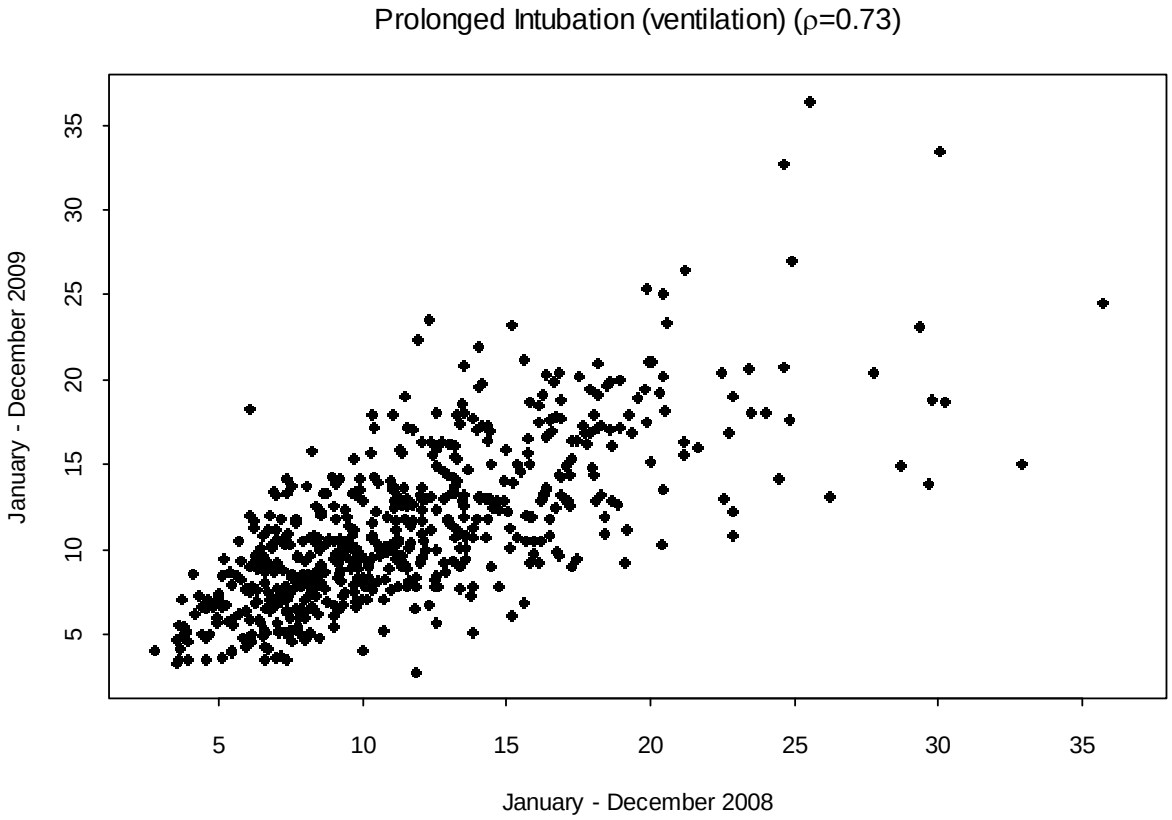
	Prolonged Intubation (ventilation) - Risk Adjusted Rate	
	Population Group	
	Hispanic	*Non-Hispanic*
Measurement		
N	94	902
Mean	14.4	11.8
1st	4.1	3.9
5th	5.6	5.2
10th	6.2	6.0
25th	8.9	8.0
Median	13.6	10.9
75th	18.5	14.7
90th	23.0	18.4
95th	26.3	20.8
99th	36.2	27.2
Outlier	38 (40.4%)	431 (47.8%)
High	13	206
Low	25	225

2b.3. Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*

Testing results: $\rho = 0.73$



Risk Adjusted Rate



2f.3. Measure Scores from Testing or Current Use *(Description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*

Results below are from January 1, 2009-December 31, 2009. Sample contains 640 STS Adult Cardiac Surgery Database Participants who had at least 100 eligible cases for the measure and reported data to STS for all 12 months.

Measurement	*Prolonged Intubation (ventilation)*
N	640
Mean	1.1
1st	3.4
5th	2.3
10th	1.9
25th	1.4
Median	1.0
75th	0.7
90th	0.5
95th	0.4
99th	0.3
Outlier†	270 (42.2)
High	120
Low	150

Risk Adjusted Rate:

Measurement	*Prolonged Intubation (ventilation)*
N	640
Mean	11.5
1st	3.5
5th	5.0
10th	5.9
25th	7.8
Median	10.6
75th	14.2
90th	18.0
95th	20.4
99th	26.4
Outlier†	270 (42.2)
High	120

Measurement	*Prolonged Intubation (ventilation)*
Low	150

†Represents the number of participants that are outliers according to two-sided 95% confidence interval of odds ratio.

3a.6. Results *(Qualitative or quantitative results and conclusions)*

Although formal testing of interpretability has not been performed, this measure has been used and reported for STS Adult Cardiac Surgery database participants since 2007. Current report presentation and interpretation manuals are presented below. These materials are updated as needed based upon feedback from database participants.

1) Report Overview and Interpretation Manual:

The NQF Measures Report

a. Organization

This report section is separated into three areas corresponding to: 1) NQF volume measures, 2) NQF process measures, and 3) NQF outcomes measures, in that order. The header at the top of each page references the report section for that page. Each NQF measure is presented on a single row in the section. Tabular data are on the left-hand side of each page and a standard graphic representation is shown on the right-hand side.

b. Statistical Calculation and Details – NQF Measures

Time period: This report section contains information on the individual STS participant and overall STS performance for the most recent 12 months for volume, process and CABG outcomes measures and the most recent 60 months for Valve and Valve + CABG outcomes. The 5 years (60 months) of performance for outcomes involving Valve procedures is necessary due to smaller sample sizes.

Volume Measures: The NQF report provides average annual case volumes data for three surgery categories: i) Isolated CABG, ii) Valve without CABG, and iii) combined CABG + Valve. Definitions of the three surgery categories are provided in Table 2 of this NQF Report Overview. For each type of surgery, the participant's annualized volume is calculated as:

$$\text{Participant Annualized Volume} = 12 \times (\#\text{ of surgeries}) / (\#\text{ of months})$$

where (# of surgeries) denotes the number of surgeries of the specified type performed by the participant during the specified time period, and (# of months) is the number of months during the specified time period for which the participant submitted at least one cardiac surgery of any type. The intent of calculating "annualized" volumes is to adjust for participants who participated in the database for fewer months than the time period specified. For participants who participated in the database and submitted cases every month during 2006, the annualized volume for 2006 is simply the total number of cases.

The STS Average Annualized Volume is the average value of all of the participant annualized volumes across the entire population of STS participants. The Participant Percentile indicates the percent of STS participants whose annualized volumes are less than, or equal to, your own. Higher percentiles indicate higher volumes in relation to other STS participant sites. The Distribution of Participant Values shows the range and percentiles of the distribution of participant annualized volumes across all database participants. For example, 90% of participants have annualized volumes less than or equal to the value marked "90th percentile." Confidence intervals are not provided for volume measures, as volume is known with certainty and is not estimated.

Process Measures: The NQF process measures provide data on the frequency of usage of five therapies among subsets of Isolated CABG patients. The therapies are: i) preoperative beta blockade therapy, ii) use of IMA, iii) discharge anti-platelet medication, iv) discharge beta blockade therapy, and v) discharge anti-lipid medication. The patient population for each measure differs, in accordance with the NQF specifications (see Table 2 of this NQF Report Overview for details). The number of Eligible

Procedures is the number of cases performed by the participant during the specified time period who meet the eligibility requirements to be included in the calculations when summarizing the participant's data. ***Beginning with the 2008 Harvest 3 report (covering the procedure time period through 6/30/2008), STS implementation of NQF medication process measures using data version 2.61 excludes records for which the medication was contraindicated/not indicated from the eligible population.*** The main summary statistic, Participant Usage, is the percent of eligible Isolated CABG cases during the specified time period for which the patient received the specified therapy. The Overall STS Usage is the percent of all eligible patients in the entire STS population during the specified time period who received the specified therapy. ***In calculating these percentages, missing data are treated as a "No", emphasizing the importance of having complete data in these fields***.

The Participant Percentile indicates the percent of STS participants who applied the therapy in their respective populations less frequently than or as frequently as did your institution. The Distribution of Participant Values shows the range and percentiles of the distribution of participant usage across all participants in the database. For example, 90% of participants use the therapy less frequently than the amount indicated by the "90th percentile". A bar identified as "Participant" indicates the point estimate and limits of a 95% Confidence Interval (CI) for the participant's usage of therapy. The underlying parameter being estimated is the long-run usage rate that would be observed in a large sample of patients. The 95% CI indicates the range of usage rates that are consistent with the data in light of sampling variability.

Outcomes Measures: The NQF outcomes data provide risk-adjusted analyses of mortality and morbidity for Isolated CABG surgery as well as risk-adjusted operative mortality for Isolated AVR, Isolated MVR, AVR+CABG, and MVR+CABG. The main summary statistic provided is the Participant's Estimated Odds Ratio (OR) based on a hierarchical logistic regression analysis. The OR measures the impact that a participant's performance level has on a patient's probability of experiencing an adverse outcome. The interpretation is similar to that of an O/E ratio (see the Risk-Adjusted Results: Overview portion of the General Report Overview for details on STS risk adjustment). An OR greater than 1.0 implies that the participant increases a patient's risk of experiencing the outcome, relative to an "average" STS participant. An OR less than 1.0 implies that the participant decreases a patient's risk of experiencing the outcome, relative to an "average" STS participant. Each measure is calculated among patients undergoing surgery of the type specified during the time period specified who additionally meet certain eligibility requirements. The column labeled Eligible Procedures indicates the number of patients who met the inclusion criteria to be included in the analysis for the indicated measure. The Participant Percentile is the percent of STS participants who have an estimated OR that is greater than or equal to your estimated OR. Note that this is different than performance percentiles for process measures, where the percentile indicates the percentage of STS participants with performance that is *less than* the specified number. This simply reflects the fact that high process compliance is desirable, whereas a high OR is undesirable.

The Observed Participant Rate is the percent of eligible patients who experienced the specified outcome. Unlike the participant estimated OR, the observed participant rate is not risk-adjusted. The estimated OR is the main summary statistic for summarizing the NQF measure in this report.

The Distribution of Participant Values shows the range and percentiles of the distribution of estimated Odds Ratios across all STS participants. For example, 90% of STS participants have an OR greater than the value indicated by the "90th percentile." The line that extends to the left and right of the Participant Value indicates the lower and upper limits of a 95% Confidence Interval (CI) surrounding the participant's estimated OR.

c. Technical Notes

Calculation of Percentiles for the Distribution of Participant Values: The graph provided for each measure contains information about the distribution of the value of the measure across all STS

participants, namely the minimum, maximum, 10th percentile, 50th percentile, and 90th percentile. The "X^{th}" percentile, denoted P_x, is loosely defined as the number having the property that X% of the participant values are less than P_x, and (100 – X)% of the participant values are greater than P_x. **For process measures, participants with greater than 5% missing data were excluded when calculating percentiles of the STS distribution and do not have a calculated participant percentile**. For participants having less than 5% missing data on a process measure, the missing values on the process measure were converted to "No" before calculating percentiles. For outcomes measures, all participants submitting at least one eligible case were included when calculating percentiles of the STS distribution. Missing data on outcomes variables were treated as "No."

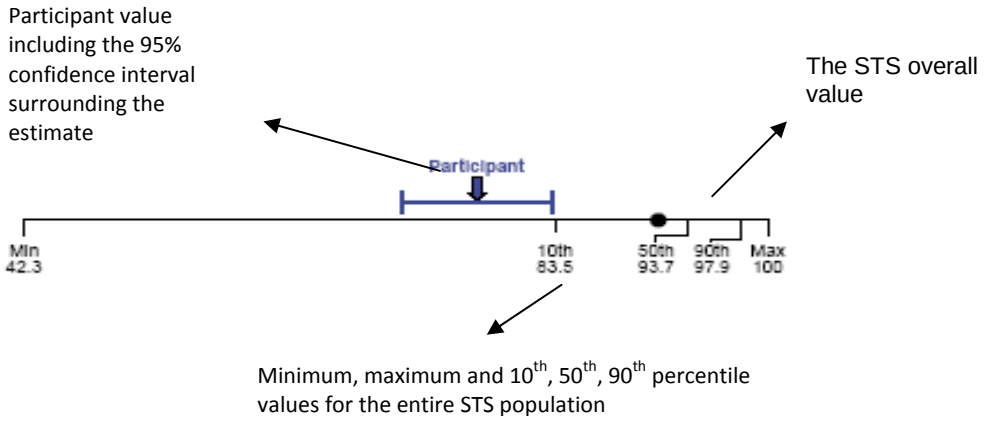
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion criteria, and rules for handling missing data in the NQF section versus the rest of the report. Definitions used in the NQF report were designed to match current NQF specifications as closely as possible. It is expected that these differences will eventually disappear as the NQF measures are refined. Some important differences are:

Case Volumes – The NQF report section presents "annualized" volumes. These are case volumes that have been adjusted for the number of months that a participant was an active contributor to the database. Elsewhere in the STS report, total case volumes are presented without adjustment for the length of participation.

Eligible Cases - The NQF report also presents the number of "eligible cases" for each measure. Separate inclusion criteria are applied to each measure, and these inclusion criteria do not always match the definitions used elsewhere in the STS report. Please refer to the footnotes in each section for specific details.

Interpretation Manual

In addition to the statistics provided for each of the STS Composite Quality Domains and NQF measures, a figure representing the distribution of values for the entire STS population is provided.



The figure allows participants to quickly judge their performance relative to the overall STS. The scale of the figure is set up such that the right side of the distribution represents the most favorable performance and the left side represents the least favorable performance (Note that in some cases smaller numbers will be on the left; in other instances, smaller numbers will be on the right. For example, for the Pre-operative Beta Blockade Therapy measure, the far left side of the distribution will contain the *lowest* percentage Beta Blockade Therapy for an STS participant – this corresponds to least

favorable performance. Alternatively, for the Operative Mortality Measure, the far left side of the distribution will contain the *highest* Estimated Odds Ratio – this also corresponds to least favorable performance). If a participant's value for a given measure is to the left of the STS overall value, the participant is performing worse on that measure than the overall STS. Conversely, if the participant's value for a given measure is located to the right of the overall STS value, the participant is performing better than the overall STS.

NOTE! Care should be given to reading these figures. In some instances, the various percentiles presented cluster very close together in the data. In such cases, the label for the percentile is not necessarily located immediately at the point on the distribution where the percentile occurs. An example of this is apparent in the figure above: The 50th percentile corresponds to a value of 93.7 and looks to align fairly closely with the STS overall value as represented by the large black dot. However, the expandable figure marking actually points to a place somewhere to the right of the STS overall value for the 50th percentile marking. So the STS overall value would be some amount less than 93.7.

Also, please note that in some cases, small sample sizes preclude valid comparisons between the participant and the STS overall. Such instances are clearly noted in the report output.

a. NQF Measures Interpretation Example

Sample CABG Operative Mortality results – tabular and figure representation.

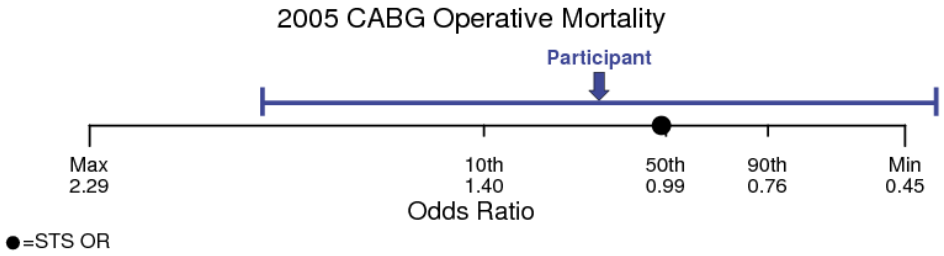
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

Eligible Procedures: 74 patients met the inclusion criteria for the indicated measure.

Participant Estimated OR (Odds Ratio): The main summary statistic measuring the impact that a participant's performance has on a patient's probability of experiencing an adverse outcome has a value of 1.14 indicating worse than expected performance.

Participant Percentile: 26.3% of STS participants had an estimated OR greater than or equal to your estimated OR. In other words, 26.3% had the same or worse performance.

Participant Observed Rate: 5.4% of the 74 eligible patients experienced the specified outcome.



The highest OR among all STS participants = 2.29
The lowest OR among all STS participants = 0.45
The STS average OR is 1.00

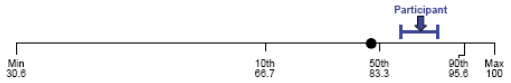
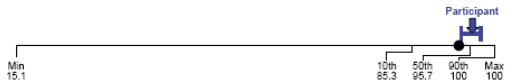
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

2) Sample page from section of the report that contains NQF measure results:

NQF Measures
Process Measures

Participant 99999
STS Period Ending 12/31/2008

Duke Clinical Research Institute
DUKE UNIVERSITY MEDICAL CENTER

NQF Measure	Eligible Procedures	Participant Usage (95% CI)	Participant Percentile	Overall STS Usage	Distribution of Participant Values ● = Overall STS Usage
Jan 2008 - Dec 2008 Preoperative Beta Blockade Therapy[1]	541	89.3% (86.4 , 91.8)	69.9	82.1%	
Jan 2008 - Dec 2008 Use of IMA[2]	536	96.5% (94.5 , 97.9)	63.3	94.2%	
Jan 2008 - Dec 2008 Discharge Anti-Platelet Medication[3]	536	98.7% (97.3 , 99.5)	68.7	96.1%	
Jan 2008 - Dec 2008 Discharge Beta Blockade Therapy[4]	538	96.1% (94.1 , 97.6)	53.4	93.7%	
Jan 2008 - Dec 2008 Discharge Anti-Lipid Treatment[4]	535	91.8% (89.1 , 94.0)	40.7	91.4%	

[1]Excludes v2.61 contranindicated / not indicated records.
[2]Excludes patients with prior CABG surgery
[3]Anti-platelet use includes Aspirin and ADP Inhibitors, and excludes in-hospital mortalities. Excludes v2.61 contranindicated / not indicated records.
[4]Excludes in-hospital mortalities. Excludes v2.61 contranindicated / not indicated records.