NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** 0124 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Surgical Volume - a. Isolated Coronary Artery Bypass Graft (CABG) Surgery, b. Valve Surgery, c. CABG+Valve Surgery
De.2 **Brief description of measure:** Annual procedural volume of three surgeries: isolated CABG surgery, valve surgery, and valve + CABG surgery.
1.1-2 **Type of Measure:** Structure De.3 **If included in a composite or paired with another measure, please identify composite or paired measure**
De.4 **National Priority Partners Priority Area:** Safety De.5 **IOM Quality Domain**: Safety De.6 **Consumer Care Need:** Getting better

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** A.3 **Measure Steward Agreement:** Agreement will be signed and submitted prior to or at the time of measure submission A.4 **Measure Steward Agreement attached:** STS Measure Steward Agreement. Fully Executed-634285394550508970.pdf	**A** Y☐ N☐

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:**	**C** Y☐ N☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1**Testing:** Yes, fully developed and tested D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Affects large numbers, High resource use, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Risk-adjusted outcomes are the most reliable cardiac surgery quality metric. However, a volume-outcome association does exist for most procedures, although the strength of the association varies substantially. For less frequently performed procedures, or when program volumes are too low to accurately estimate risk-adjusted outcomes, volumes may provide useful information for consumers. The greatest utility of surgical volumes is to identify extremely low-volume providers, as this group tends to have, on average, the worst outcomes. This information will inform consumer choice, and it may identify programs for targeted outcomes analyses and oversight. 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127. - Carey JS, Robertson JM, Misbach GA, Fisher AL. Relationship of hospital volume to outcome in cardiac surgery programs in California. Am Surg. 2003;69(1):63-68. - Eagle KA, Guyton RA, Davidoff R, Ewy GA, Fonger J, Gardner TJ, Gott JP, Hermann HC, Marlow RA, Nugent W, O'Connor GT, Orszulak TA, Rieselbach RE, Winters WL, Yusuf S. ACC/AHA guidelines for coronary artery bypass graft surgery—executive summary and recommendations: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1991 Guidelines for Coronary Artery Bypass Graft Surgery). Circulation. 1999;100:1464-1480.	**1a** C☐ P☐ M☐ N☐

- Glance LG, Dick AW, et al. Is the hospital volume-mortality relationship in CABG surgery the same for low-risk versus high-risk patients? Ann Thorac Surg. 2003;76:1155-1162.
- Hannan EL, Wu C, Ryan TJ, et al. Do hospitals and surgeons with higher coronary artery bypass graft surgery volumes still have lower risk-adjusted mortality? Circulation. 2003;108(7):795-801.
- Nowicki ER, Weintruab RW, et al. Mitral valve repair and replacement in Northern New England. Am Heart J. 2003;145(6):1058-1062.
- Peterson ED, Coombs LP, et al. Procedural volume as a marker of quality for CABG surgery. JAMA. 2004;291:195-201.
- Shahian DM. Improving cardiac surgery quality--volume, outcome, process? JAMA 2004 Jan 14;291(2):246-8.
- Shahian DM, Normand SL. The volume-outcome relationship: from Luft to Leapfrog. Ann Thorac Surg 2003 Mar;75(3):1048-58.
- Shahian DM, O´Brien SM, Normand SL, Peterson ED, Edwards FH. Association of hospital coronary artery bypass volume with processes of care, mortality, morbidity, and the Society of Thoracic Surgeons composite quality score. J Thorac Cardiovasc Surg 2010 Feb;139(2):273-82.
- Shahian DM, Normand SL. Low-volume coronary artery bypass surgery: Measuring and optimizing performance. J Thorac Cardiovasc Surg 2008 Jun 1;135(6):1202-9.

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** Risk-adjusted outcomes are the most reliable cardiac surgery quality metric. However, a volume-outcome association does exist for most procedures, although the strength of the association varies substantially. For less frequently performed procedures, or when program volumes are too low to accurately estimate risk-adjusted outcomes, volumes may provide useful information for consumers. The greatest utility of surgical volumes is to identify extremely low-volume providers, as this group tends to have, on average, the worst outcomes. This information will inform consumer choice, and it may identify programs for targeted outcomes analyses and oversight.

1b.2 **Summary of data demonstrating performance gap (variation or overall poor performance) across providers:**
Please see attachment

1b.3 **Citations for data on performance gap:**
Dates: January 1, 2009-December 31, 2009

a. Count of Isolated Coronary Artery Bypass Graft (CABG) procedures for each participant
b. Count of valve procedures for each participant
c. Count of CABG + valve procedures for each participant

1b.4 **Summary of Data on disparities by population group:**
N/A

1b.5 **Citations for data on Disparities:**
N/A

1b
C☐
P☐
M☐
N☐

1c. Outcome or Evidence to Support Measure Focus

1c.1 **Relationship to Outcomes** (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** Risk-adjusted outcomes are the most reliable cardiac surgery quality metric. However, a volume-outcome association does exist for most procedures, although the strength of the association varies substantially. For less frequently performed procedures, or when program volumes are too low to accurately estimate risk-adjusted outcomes, volumes may provide useful information for consumers. The greatest utility of surgical volumes is to identify extremely low-volume providers, as this group tends to have, on average, the worst outcomes. This information will inform consumer choice, and it may identify programs for targeted outcomes analyses and oversight.

1c.2-3. **Type of Evidence:** Observational study, Expert opinion, Systematic synthesis of research, Other

1c
C☐
P☐
M☐
N☐

Clinical results from approximately 90% of cardiac surgery centers in the US

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**

1c.6 Method for rating evidence:

1c.7 Summary of Controversy/Contradictory Evidence:

1c.8 Citations for Evidence (*other than guidelines***):**

- Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127.
- Carey JS, Robertson JM, Misbach GA, Fisher AL. Relationship of hospital volume to outcome in cardiac surgery programs in California. Am Surg. 2003;69(1):63-68.
- Eagle KA, Guyton RA, Davidoff R, Ewy GA, Fonger J, Gardner TJ, Gott JP, Hermann HC, Marlow RA, Nugent W, O'Connor GT, Orszulak TA, Rieselbach RE, Winters WL, Yusuf S. ACC/AHA guidelines for coronary artery bypass graft surgery—executive summary and recommendations: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1991 Guidelines for Coronary Artery Bypass Graft Surgery). Circulation. 1999;100:1464-1480.
- Glance LG, Dick AW, et al. Is the hospital volume-mortality relationship in CABG surgery the same for low-risk versus high-risk patients? Ann Thorac Surg. 2003;76:1155-1162.
- Hannan EL, Wu C, Ryan TJ, et al. Do hospitals and surgeons with higher coronary artery bypass graft surgery volumes still have lower risk-adjusted mortality? Circulation. 2003;108(7):795-801.
- Nowicki ER, Weintruab RW, et al. Mitral valve repair and replacement in Northern New England. Am Heart J. 2003;145(6):1058-1062.
- Peterson ED, Coombs LP, et al. Procedural volume as a marker of quality for CABG surgery. JAMA. 2004;291:195-201.
- Shahian DM. Improving cardiac surgery quality--volume, outcome, process? JAMA 2004 Jan 14;291(2):246-8.
- Shahian DM, Normand SL. The volume-outcome relationship: from Luft to Leapfrog. Ann Thorac Surg 2003 Mar;75(3):1048-58.
- Shahian DM, O´Brien SM, Normand SL, Peterson ED, Edwards FH. Association of hospital coronary artery bypass volume with processes of care, mortality, morbidity, and the Society of Thoracic Surgeons composite quality score. J Thorac Cardiovasc Surg 2010 Feb;139(2):273-82.
- Shahian DM, Normand SL. Low-volume coronary artery bypass surgery: Measuring and optimizing performance. J Thorac Cardiovasc Surg 2008 Jun 1;135(6):1202-9.

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number***):**
N/A

1c.10 Clinical Practice Guideline Citation: N/A
1c.11 National Guideline Clearinghouse or other URL: N/A

1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom***):**
N/A

1c.13 Method for rating strength of recommendation (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:**
N/A

1c.14 Rationale for using this guideline over others:
N/A

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to*	**1**

Measure and Report?	
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**
2. 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. (evaluation criteria)	**Eval Rating**
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2 If yes, provide web page URL:** **2a. Precisely Specified**	
2a.1 Numerator Statement (*Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome***):** a. Number of patients undergoing isolated CABG surgery b. Number of patients undergoing heart valve surgery c. Number of patients undergoing CABG + valve surgery **2a.2 Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** 12 months **2a.3 Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):** a. Isolated CABG is determined as a procedure for which all of the following apply: - OpCAB (STS Adult Cardiac Surgery Database Version 2.73) is marked "Yes" - (VADProc is marked "No" or "Missing") or (VADProc is marked "Yes, Implanted" and UnplVAD is marked "yes") - OCarASDTy is marked "PFO" or "missing" - OCarAFibAProc is marked "primarily epicardial" or "missing" and - OpValve, VSAV, VSAVPr, ResectSubA, VSMV, VSMVPr, OpTricus, OpPulm, OpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc, OCTumor, OCPulThromDis, OCarOthr are all marked "no" or "missing" b. Any mitral, aortic, tricuspid, or pulmonary valve surgery without CABG; Valve surgery is determined as a procedure for which OpValve is marked "yes" and any of the following is marked "yes": - Aortic Valve Procedure (VSAV) - Mitral Valve Procedure (VSMV) - Tricuspid Valve Procedure (OpTricus) - Pulmonic Valve Procedure (OpPulm) c. Any mitral, aortic, tricuspid, or pulmonary valve surgery with a CABG; CABG + Valve Surgery is determined as a procedure for which OpCAB is marked "yes," OpValve is marked "yes," and one of the following is marked "yes": - Aortic Valve Procedure (VSAV) - Mitral Valve Procedure (VSMV) - Tricuspid Valve Procedure (OpTricus) - Pulmonic Valve Procedure (OpPulm)	
2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):** N/A **2a.5 Target population gender:** Female, Male **2a.6 Target population age range:** 18 years and older	**2a-specs** **C☐** **P☐** **M☐** **N☐**

2a.7 **Denominator Time Window (***The time period in which cases are eligible for inclusion in the denominator***):**
N/A

2a.8 **Denominator Details (***All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
N/A

2a.9 **Denominator Exclusions (***Brief text description of exclusions from the target population***):** N/A

2a.10 **Denominator Exclusion Details (***All information required to collect exclusions to the denominator, including all codes, logic, and definitions***):**
N/A

2a.11 **Stratification Details/Variables (***All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
N/A

2a.12-13 **Risk Adjustment Type:** No risk adjustment necessary

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

2a.15-17 **Detailed risk model available Web page URL or attachment:**

2a.18-19 **Type of Score:** Categorical
2a.20 **Interpretation of Score:** Passing score defines better quality
2a.21 **Calculation Algorithm (***Describe the calculation of the measure as a flowchart or series of steps***):**
N/A

2a.22 **Describe the method for discriminating performance** *(e.g., significance testing)***:**

2a.23 **Sampling (Survey) Methodology** *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
N/A

2a.24 **Data Source (***Check the source(s) for which the measure is specified and tested)*
Registry data

2a.25 **Data source/data collection instrument (***Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
STS Adult Cardiac Surgery Database – Version 2.73

2a.26-28 **Data source/data collection instrument reference web page URL or attachment:** URL Data Collection Form (an updated version will be made available on the STS Website in mid-January 2011) --- http://www.sts.org/documents/pdf/ndb2010/STSAdultCVDataCollectionForm2_7_Annotated_20101021.pdf

2a.29-31 **Data dictionary/code table web page URL or attachment:** URL http://www.sts.org/documents/pdf/ndb2010/STSAdultCVDataSpecificationsV2_7_20101021.pdf -- an updated version will be made available on the STS Website in mid-January 2011

2a.32-35 **Level of Measurement/Analysis** (*Check the level(s) for which the measure is specified and tested)*
Clinicians : Group, Facility/Agency, Population : Counties or cities, Population : National, Population : Regional/network, Population : states

2a.36-37 **Care Settings (***Check the setting(s) for which the measure is specified and tested)*
Hospital

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
Clinicians: Physicians (MD/DO)

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)*: One of the most important characteristics of test reliability is consistency across measurement platforms. This characteristic may be measured by inter-rater or parallel forms reliability. One should obtain reasonably similar results from different raters or from different forms of testing. In the case of cardiac surgery volumes, data submitted to the STS Adult Cardiac Surgery Database (ACD) are compared to hospital operative logs during our audit process.

Hospital logs of CABG-only and isolated valve cases are compared with a list (provided by the Duke Clinical Research Institute) of CABG-only and isolated valve cases submitted to the Database to determine if the data from these two sources are consistent. For audits conducted in 2009 (N=29), all sites were found to have processes in place to ensure that eligible cases were submitted to the database. The results revealed high percentage agreement between the lists of procedures performed and submitted.

There are, however, concerns with the reliability of cardiac surgical volumes derived from administrative data. Many different algorithms exist for determining which administrative codes to include, and they will thus yield different numbers. Even more importantly, these administratively-derived results are generally not consistent with the numbers of procedures determined by gold-standard clinical databases such as those maintained by the STS. This reflects on both their reliability and their validity [1,2].

1. Mack MJ, Herbert M, Prince S, Dewey TM, Magee MJ, Edgerton JR. Does reporting of coronary artery bypass grafting from administrative databases accurately reflect actual clinical outcomes? J Thorac Cardiovasc Surg 2005 Jun;129(6):1309-17.
2. Shahian DM, Silverstein T, Lovett AF, Wolf RE, Normand S-L. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. Circulation 2007 Mar 27;115(12):1518-27.

This is a serious deficiency of administrative data used for volume or outcomes profiling. In such activities, the goal is to focus on relatively homogeneous sets of procedures such as isolated CABG. Data derived from administrative sources are more likely to erroneously include CABG cases combined with other more complex procedures, and these combined operations have a significantly higher risk. Thus, any analyses of such data will be comparing apples and oranges. It is particularly problematic for tertiary centers that perform disproportionately more of the complex, combined operations. If these cases are inappropriately included among what should be isolated CABG, then their observed mortality will be higher. Even with risk adjustment, they will appear to be performing poorly compared with institutions that perform mostly isolated CABG.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:

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

2b C☐ P☐ M☐ N☐

2c. Validity testing

2c.1 Data/sample *(description of data/sample and size)*: STS Adult Cardiac Surgery Database

Audits conducted in 2010, all cases performed in 2009; N = 40 randomly selected sites participating in the STS Adult Cardiac Surgery Database

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
Participating sites are randomly selected for participation in STS Adult Cardiac Surgery Database Audit, which is designed to evaluate the accuracy, consistency, and comprehensiveness of data collection and

2c C☐ P☐ M☐ N☐

ultimately validate the integrity of the data contained in the database. The Iowa Foundation for Medical Care (IFMC), the quality improvement organization for Iowa and Illinois, has conducted audits on behalf of STS since 2006. Each year, the IFMC conducts audits at randomly selected sites throughout the country and tracks the individual agreement rates by variable and by year. More specifically, for each site, agreement rates are calculated for 73 individual elements. In addition, aggregate agreement rates for each element, variable category (e.g., pre-operative risk factors, previous interventions, etc), and overall for all categories are calculated for all sites. While this is not region specific, it is data point specific and comparison agreement rates confirm the improvement over time as well as the consistency. In addition, validity was confirmed and is regularly assessed by an expert panel of thoracic surgeons assembled by the STS Adult Cardiac Surgery Database Task Force, the STS Task Force on Quality Initiatives and the STS Workforce on National Databases. 2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:	
2d. Exclusions Justified 2d.1 **Summary of Evidence supporting exclusion(s):** N/A 2d.2 **Citations for Evidence:** 2d.3 **Data/sample** *(description of data/sample and size)*: 2d.4 **Analytic Method** *(type analysis & rationale)*: 2d.5 **Testing Results** *(e.g., frequency, variability, sensitivity analyses)*:	**2d** C☐ P☐ M☐ N☐ NA☐
2e. Risk Adjustment for Outcomes/ Resource Use Measures 2e.1 **Data/sample** *(description of data/sample and size)*: N/A 2e.2 **Analytic Method** *(type of risk adjustment, analysis, & rationale)*: 2e.3 **Testing Results** *(risk model performance metrics)*: 2e.4 **If outcome or resource use measure is not risk adjusted, provide rationale:**	**2e** C☐ P☐ M☐ N☐ NA☐
2f. Identification of Meaningful Differences in Performance 2f.1 **Data/sample from Testing or Current Use** *(description of data/sample and size)*: Dates: January 1, 2009-December 31, 2009 a. Count of Isolated Coronary Artery Bypass Graft (CABG) procedures for each participant b. Count of valve procedures for each participant c. Count of CABG + valve procedures for each participant 2f.2 **Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*:	**2f** C☐ P☐ M☐ N☐

2f.3 Provide 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)*: Please see attachment	
2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: N/A **2g.2 Analytic Method** *(type of analysis & rationale)*: **2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*:	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: N/A **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:**	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	***2*** C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	**Eval Rating**
3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** In use **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*): Currently being considered for NQF endorsement, the STS CABG Composite Score is a multidimensional performance measure comprised of four domains consisting of 11 individual NQF-endorsed cardiac surgery metrics: (1) Operative Care--use of the internal mammary artery; (2) Perioperative Medical Care (use of preoperative beta blockade; discharge beta blockade, antiplatelet agents, and lipid-lowering agents—an "all-or-none" measure); (3) Risk-adjusted Operative Mortality; and (4) Risk-Adjusted Postoperative Morbidity (occurrence of postoperative stroke, renal failure, prolonged ventilation, re-exploration, or deep sternal wound infection--an "any-or-none" measure). Composite star ratings are presented in the health section of the Consumers Union website, www.ConsumerReportsHealth.org Currently, there are 221 STS Adult Cardiac Surgery Database participating hospitals who voluntarily participate in this initiative. At present, STS does not support public reporting of volumes, as we believe risk-adjusted outcomes and our composite measure are superior performance metrics **3a.3 If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):**	**3a** C☐ P☐ M☐ N☐

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*) **3a.4 Data/sample** *(description of data/sample and size)*: See 3a.6 below **3a.5 Methods** *(e.g., focus group, survey, QI project)*: **3a.6 Results** *(qualitative and/or quantitative results and conclusions)*: Please see attachment	
3b/3c. Relation to other NQF-endorsed measures **3b.1 NQF # and Title of similar or related measures:** …	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): **3b.2 Are the measure specifications harmonized? If not, why?** N/A; however, data definitions and key elements have been established by a multi-societal writing committee called the "ACCF/AHA Writing Committee to Develop Acute Coronary Syndromes and Coronary Artery Disease Clinical Data Standards" with representatives from each of the following organizations: Agency for Healthcare Research and Quality American College of Cardiology American College of Chest Physicians American College of Emergency Physicians American College of Physicians American College of Preventative Medicine American Heart Association American Medical Association Centers for Disease Control and Prevention Emergency Nurses Association Food and Drug Administration Joint Commission on Accreditation of Healthcare Organizations National Association of Emergency Medical Technicians National Association of EMS Physicians National Heart, Lung, and Blood Institute Preventive Cardiovascular Nurses Association Society for Academic Emergency Medicine Society of Chest Pain Centers and Providers Society of General Internal Medicine Society of Thoracic Surgeons	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value **3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** **5.1 If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:**	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐

	P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Data generated as byproduct of care processes during care delivery (Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition), Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	4a C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) Yes 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	4b C☐ P☐ M☐ N☐
4c. Exclusions 4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** No 4c.2 **If yes, provide justification.**	4c C☐ P☐ M☐ N☐ NA☐
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences 4d.1 **Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.** This measure may be susceptible to human error (i.e., recording the measure inaccurately or not at all). When data collection on this measure is done through participation in the STS Adult Cardiac Surgery Database, an auditing strategy is in place. Both STS and the Duke Clinical Research Institute have a list of database participants making participation in the STS Adult Cardiac Surgery Database easy to track. Each participant is responsible for the quality and accuracy of the data they submit to the database. The participant agrees to the following quality control measures in the participation agreement: i) Participant hereby warrants that all data submitted for inclusion in the STS National Database will be accurate and complete, and acknowledges that such data may be subject to independent audit. Participant will use its best efforts to address any data or related deficiencies identified by the independent data warehouse service provider and agrees to cooperate with and assist STS and its designees in connection with the performance of any independent audit. ii) Participant warrants that it will take all reasonable steps to avoid the submission of duplicative data for inclusion in the STS National Database, including but not limited to apprising the Director of the STS National Database and the independent data warehouse service provider about any other Participation Agreements in which an individual cardiothoracic surgeon named above or on Schedule A attached hereto (as amended from time to time) is also named. STS audited for these potential problems during testing.	4d C☐ P☐ M☐ N☐

4e. Data Collection Strategy/Implementation **4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:** **4e.2 Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Data Collection: There are no direct costs to collect the data for this measure. Costs to develop the measure included volunteer cardiothoracic time, STS staff time, and DCRI statistician and project management time. Other fees: STS Adult Cardiac Surgery Database participants (single cardiothoracic surgeons or a group of surgeons) pay annual participant fees of $2,950 or $3,700, depending on whether participants are STS members (or whether the majority of surgeons in a group are STS members). As a benefit of STS membership, STS members are charged the lesser of the two fees. **4e.3 Evidence for costs:** **4e.4 Business case documentation:**	**4e** **C☐** **P☐** **M☐** **N☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** **C☐** **P☐** **M☐** **N☐**
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited ☐**
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y☐** **N☐** **A☐**
CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) **Co.1 Organization** Society of Thoracic Surgeons, 633 North Saint Clair Street, Suite 2320, Chicago, Illinois, 60611 **Co.2 Point of Contact** Jane, Han, MSW, jhan@sts.org, 312-202-5856-	
Measure Developer If different from Measure Steward **Co.3 Organization** Society of Thoracic Surgeons, 633 North Saint Clair Street, Suite 2320, Chicago, Illinois, 60611 **Co.4 Point of Contact** Jane, Han, MSW, jhan@sts.org, 312-202-5856-	
Co.5 Submitter If different from Measure Steward POC	

Jane, Han, MSW, jhan@sts.org, 312-202-5856-, Society of Thoracic Surgeons
Co.6 Additional organizations that sponsored/participated in measure development
ADDITIONAL INFORMATION
Workgroup/Expert Panel involved in measure development **Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.** Members of the STS Task Force on Quality Initiatives provide clinical expertise as needed. The STS Workforce on National Databases meets at the STS Annual Meeting and reviews the measures on a yearly basis. Changes or updates to the measure will be at the recommendation of the Workforce.
Ad.2 If adapted, provide name of original measure: **Ad.3-5 If adapted, provide original specifications URL or attachment**
Measure Developer/Steward Updates and Ongoing Maintenance **Ad.6 Year the measure was first released:** 2004 **Ad.7 Month and Year of most recent revision:** 12, 2010 **Ad.8 What is your frequency for review/update of this measure?** annually **Ad.9 When is the next scheduled review/update for this measure?** 2011
Ad.10 Copyright statement/disclaimers:
Ad.11 -13 Additional Information web page URL or attachment: Attachment 0124 Sections 1b.2, 2f.3, 3a.6.pdf
Date of Submission (*MM/DD/YY*)**:** 01/12/2011

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.)*

a. 966 STS Adult Cardiac Surgery Database Participants who submitted any data to STS Adult Cardiac database in at least 4 months in 2009.

Measurement	*Surgical Volume - a. Isolated Coronary Artery Bypass Graft (CABG)*
N	966
Mean	173.8
1st	25.0
5th	41.0
10th	54.9
25th	86.0
Median	139.0
75th	222.0
90th	336.0
95th	423.0
99th	654.0

b. 966 STS Adult Cardiac Surgery Database Participants who submitted any data to STS Adult Cardiac database in at least 4 months in 2009.

Measurement	*Surgical Volume – b. Valve Surgery*
N	966
Mean	64.8
1st	1.0
5th	6.0
10th	9.0
25th	20.0
Median	38.0
75th	75.0
90th	144.0
95th	206.0
99th	368.0

c. 966 STS Adult Cardiac Surgery Database Participants who submitted any data to STS Adult Cardiac database in at least 4 months in 2009.

Measurement	*Surgical Volume – c. CABG+Valve Surgery*
N	966
Mean	38.5
1st	0.0
5th	3.0

10th	6.0
25th	13.0
Median	27.0
75th	48.0
90th	84.0
95th	115.0
99th	189.0

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)*

a. 966 STS Adult Cardiac Surgery Database Participants who submitted any data to STS Adult Cardiac database in at least 4 months in 2009.

Measurement	*Surgical Volume - a. Isolated Coronary Artery Bypass Graft (CABG)*
N	966
Mean	173.8
1st	25.0
5th	41.0
10th	54.9
25th	86.0
Median	139.0
75th	222.0
90th	336.0
95th	423.0
99th	654.0

b. 966 STS Adult Cardiac Surgery Database Participants who submitted any data to STS Adult Cardiac database in at least 4 months in 2009.

Measurement	*Surgical Volume – b. Valve Surgery*
N	966
Mean	64.8
1st	1.0
5th	6.0
10th	9.0
25th	20.0
Median	38.0
75th	75.0
90th	144.0
95th	206.0
99th	368.0

c. 966 STS Adult Cardiac Surgery Database Participants who submitted any data to STS Adult Cardiac database in at least 4 months in 2009.

Measurement	*Surgical Volume – c. CABG+Valve Surgery*
N	966
Mean	38.5
1st	0.0

5th	3.0
10th	6.0
25th	13.0
Median	27.0
75th	48.0
90th	84.0
95th	115.0
99th	189.0

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."

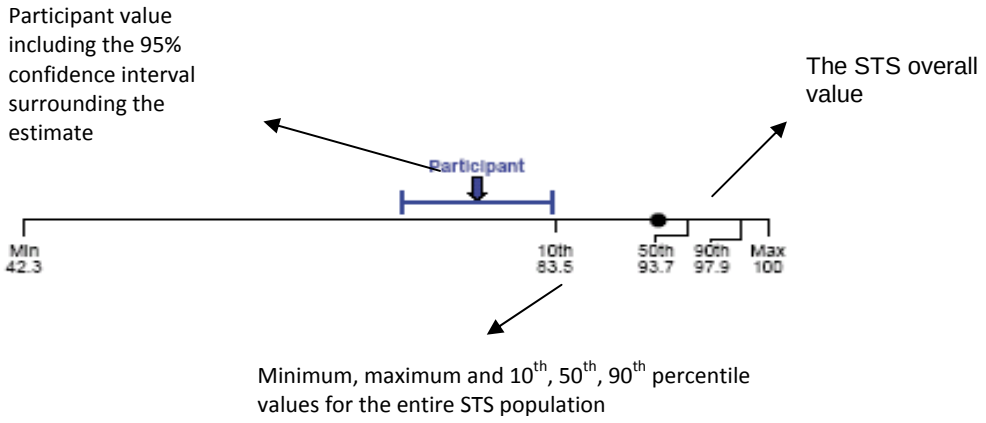
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion 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.

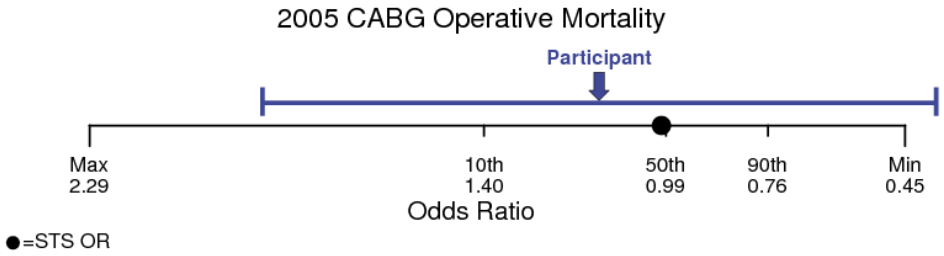
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

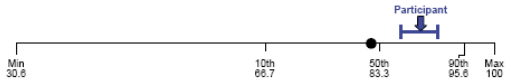
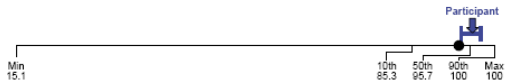
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.