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 #:** 0120 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR)
De.2 **Brief description of measure***:* Percent of patients undergoing Aortic Valve Replacement (AVR)who die, including both 1) all deaths occurring during the hospitalization in which the procedure was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure
1.1-2 **Type of Measure:** Outcome 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-634282024404397262.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:** Public reporting, Internal quality improvement	**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, Frequently performed procedure, Leading cause of morbidity/mortality, High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Mortality is probably the single most important negative outcome that can be associated with a surgical procedure. Operative mortality, defined as mortality within 30 days of surgery or on the same hospital admission, should include nearly all deaths that occur as a direct result of the surgery or an immediate postoperative complication. 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892. - Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137. - Mihaljevic T, Nowicki ER, Rajeswaran J, et. al. Survival after valve replacement for aortic stenosis: implications for decision making. J Thorac Cardiovasc Surg. 2008 Jun;135(6):1270-8; discussion 1278-9. - Tabata M, Umakanthan R, Cohn LH, et. al. Early and late outcomes of 1000 minimally invasive	**1a** C☐ P☐ M☐ N☐

aortic valve operations. Eur J Cardiothorac Surg. 2008 Apr;33(4):537-41. Epub 2008 Feb 5.
- Chaliki HP, Mohty D, Avierinos JF, et. al. Outcomes after aortic valve replacement in patients with severe aortic regurgitation and markedly reduced left ventricular function. Circulation. 2002 Nov 19;106(21):2687-93

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** Mortality is probably the single most important negative outcome that can be associated with a surgical procedure. Operative mortality, defined as mortality within 30 days of surgery or on the same hospital admission, should include nearly all deaths that occur as a direct result of the surgery or an immediate postoperative complication. Critical evaluation of operative mortality allows one to evaluate the risk associated with a given procedure for various patient characteristics, and more importantly, aggressively search for ways to minimize that risk.

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, 2005-December 31, 2009

Analysis includes 538 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data (not restricted to this measure) to STS for at least 36 months in 2005-2009.

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

1b.5 **Citations for data on Disparities:**

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*)**:** Mortality is probably the single most important negative outcome that can be associated with a surgical procedure. Operative mortality, defined as mortality within 30 days of surgery or on the same hospital admission, should include nearly all deaths that occur as a direct result of the surgery or an immediate postoperative complication.

1c.2-3. **Type of Evidence:** Observational study, Expert opinion, Systematic synthesis of research, Other
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*)**:**
The published literature on aortic valve replacement, including the aforementioned references, includes multiple examples of services/care processes that impact operative mortality. Pre-operative patient selection, surgical timing, intraoperative conduct of the case, and many aspects to postoperative care have all been shown to have significant impact on the operative mortality over the last few decades.

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 NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease

1c
C☐
P☐
M☐
N☐

Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892. - Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137. - Mihaljevic T, Nowicki ER, Rajeswaran J, et. al. Survival after valve replacement for aortic stenosis: implications for decision making. J Thorac Cardiovasc Surg. 2008 Jun;135(6):1270-8; discussion 1278-9. - Tabata M, Umakanthan R, Cohn LH, et. al. Early and late outcomes of 1000 minimally invasive aortic valve operations. Eur J Cardiothorac Surg. 2008 Apr;33(4):537-41. Epub 2008 Feb 5. - Chaliki HP, Mohty D, Avierinos JF, et. al. Outcomes after aortic valve replacement in patients with severe aortic regurgitation and markedly reduced left ventricular function. Circulation. 2002 Nov 19;106(21):2687-93 1c.9 **Quote the Specific guideline recommendation (***including guideline number and/or page number***):** N/A 1c.10 **Clinical Practice Guideline Citation:** 1c.11 **National Guideline Clearinghouse or other URL:** 1c.12 **Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):** 1c.13 **Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:** 1c.14 **Rationale for using this guideline over others:**	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
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***):** Number of patients undergoing AVR who die, including both 1) all deaths occurring during the hospitalization in which the operation was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure 2a.2 **Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** During hospitalization regardless of length of stay or within 30 days of surgery if discharged 2a.3 **Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):**	**2a-specs** **C☐** **P☐** **M☐** **N☐**

Number of isolated AVR procedures with an operative mortality;

Number of isolated AVR procedures in which Mortality [Mortalty (STS Adult Cardiac Surgery Database Version 2.73)] and Mortality Operative Death (MtOpD) are marked "yes." Operative mortality is further verified by the following variables: Mortality Status at 30 days (Mt30Stat), Mortality Date (MtDate), Mortality Discharge Status (MtDCStat)

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):**
All patients undergoing isolated AVR surgery

2a.5 Target population gender: Female, Male
2a.6 Target population age range: 18 yrs and older

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

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
Number of isolated AVR procedures;

Isolated AVR is determined as a procedure for which all of the following apply:
- OpValve is marked "Yes"
- VSAV is marked "Yes"
- VSAVPr is marked "Replacement"
- (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
- OpCAB , ResectSubA , VSMV, VSMVPr , OpTricus, OpPulmOpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc, OCTumor, OCPulThromDis, OCarOthr are all marked "no" or "missing"

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

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

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

2a.12-13 Risk Adjustment Type: Case-mix adjustment

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

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment 2a.15 Detailed Risk Model-634282025771376018.pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**

2a.22 Describe the method for discriminating performance *(e.g., significance testing)*:
Participant specific OR and their 95% CI were estimated in the hierarchical model. These model-based estimates were used to control variation due to random statistical fluctuations while estimating true signal

variation. A 95% CI excluding zero indicates the participant's performance is significantly lower or higher than an "average" STS participant.

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)*:

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: national, Population: regional/network, Population: states, Population: counties or cities

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)*: STS Adult Cardiac Surgery Database - Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009.

2b.2 **Analytic Method** *(type of reliability & rationale, method for testing)*:
Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009. Excluded from analysis are participants that did not submit results for both time periods. Because database participants can change their underlying care processes at any time, we would not expect perfect correlation between two sets of results from even proximate time periods.

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

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,

2c
C☐
P☐
M☐
N☐

which is designed to evaluate the accuracy, consistency, and comprehensiveness of data collection and 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.

2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Mortality Operative Death: 100.0% agreement rate

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)*: Please see Risk Adjustment Type section above

2e.2 **Analytic Method** *(type of risk adjustment, analysis, & rationale)*:
Detailed information regarding the risk adjustment model can be found in the attachment:

O´Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S23-42.

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)*: 538 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009; January 1, 2005-December 31, 2009

2f.2 **Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*:
We calculated the risk adjusted event rate with the participant's Odds Ratio (OR) estimate and the overall STS event rate. Therefore, the risk adjusted rate is closely related to OR estimate. If OR > 1, then the participant's risk adjusted rate will be greater than the overall STS event rate; if OR < 1, then the participant's risk adjusted rate will be smaller than the overall STS event rate. The statistical significance

2f
C☐
P☐
M☐
N☐

is defined by the 95% confidence interval (CI) or the OR estimate. If the 95% CI for a participant's OR includes the null value 1.0, then we cannot distinguish this participant's performance from the STS average – either the participant's performance was close to average or else the participant's sample size was too small to make a reliable inference. Otherwise, if the 95% CI falls to the right of 1.0, then the participant's performance is considered significantly lower than the average STS results; if the 95% CI falls to the left of 1.0, then the participant's performance is considered significantly higher than the average STS results. 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)*: 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)*: 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 STS plans to publicly report more measures in the future. There is no definite date yet assigned to this measure; however, STS staff and surgeon leadership have engaged in initial internal STS discussions regarding this matter.	**3a** C☐ P☐ M☐ N☐

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***):**

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	**4d** **C☐** **P☐** **M☐** **N☐**

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. Please see IFMC audit results.	
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 0120 Sections 2a.14, 1b.2, 2b.3, 2f.3, 3a.6.pdf
Date of Submission (*MM/DD/YY*)**:** 01/12/2011

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 valve procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest four and a half years of data (January 2006 – 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 Isolated Valve Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
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
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise

Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 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 ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 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
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Stenosis mitral	= 1 if patient has mitral 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; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention;

The final patient-level model was built by step-wise selection method with several variables decided by surgeon panel 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	*Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) Surgery*
N	538
Mean	1.0
1st	2.0
5th	1.5
10th	1.4
25th	1.2
Median	0.9
75th	0.8
90th	0.7
95th	0.6
99th	0.5
Outlier	20 (3.7)
High	12
Low	8

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)

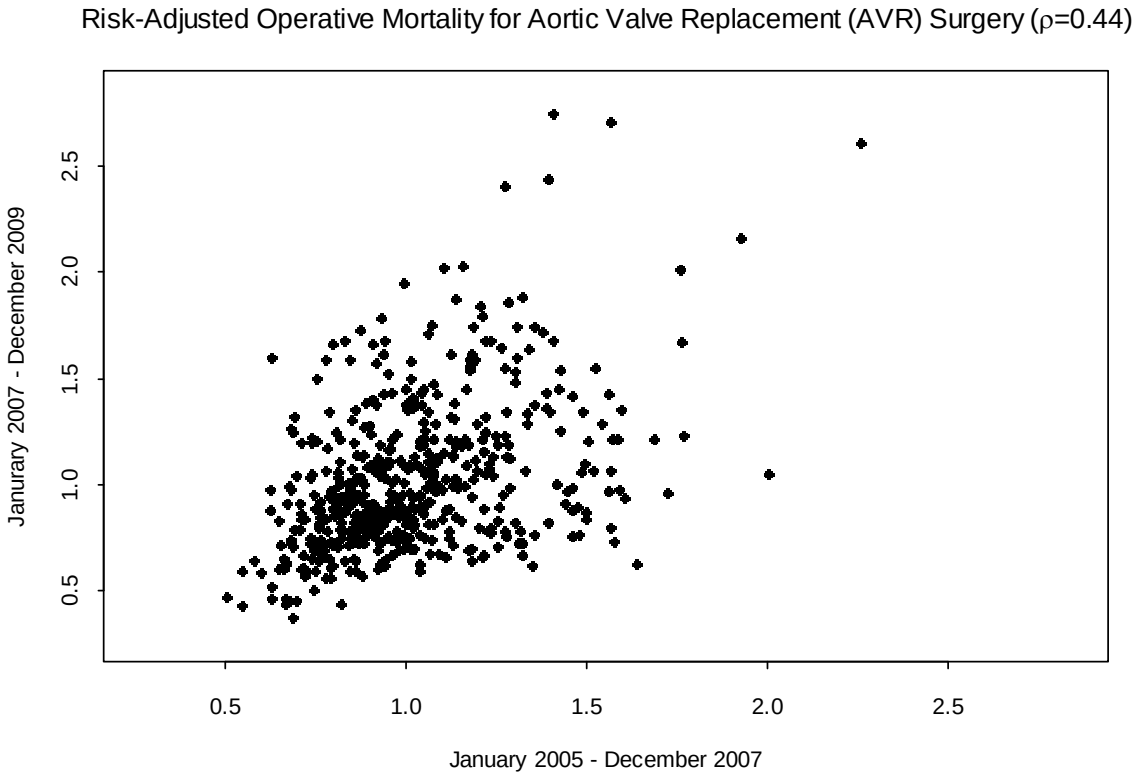
In the above equation, "Participant's Expected Event Rate" is calculated with the participant's actual OR, and "Participant's Expected Event Rate Assuming Its Performance = STS Average Performance" is calculated by assuming the participant's OR = 1 (i.e. no difference in performance from the STS average).

Measurement	*Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) Surgery*
N	538
Mean	3.1
1st	1.6

Measurement	*Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) Surgery*
5th	1.9
10th	2.2
25th	2.5
Median	2.9
75th	3.5
90th	4.1
95th	4.5
99th	5.6
Outlier	20 (3.7)
High	12
Low	8

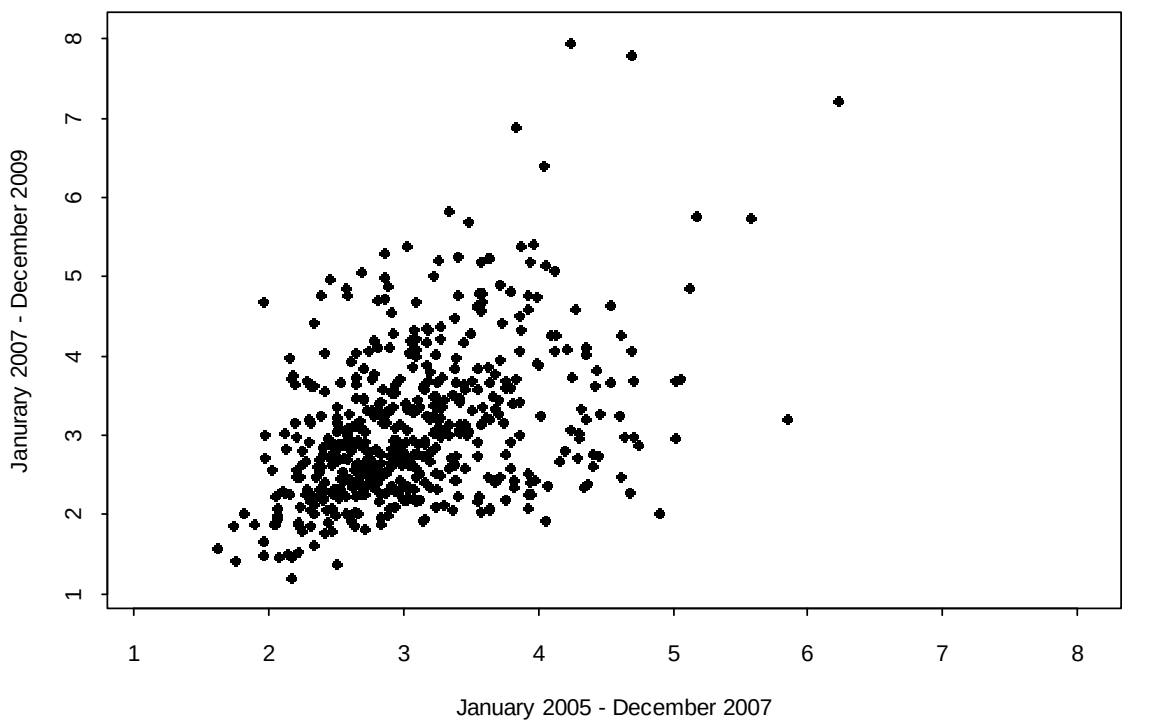
2b.3. Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*

Testing results: $\rho = 0.44$



Risk Adjusted Rate:

Testing results: $\rho = 0.43$



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, 2005-December 31, 2009. Sample contains 538 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009.

Measurement	*Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) Surgery*
N	538
Mean	1.0
1st	2.0
5th	1.5
10th	1.4
25th	1.2
Median	0.9
75th	0.8
90th	0.7
95th	0.6
99th	0.5
Outlier†	20 (3.7)
High	12
Low	8

Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) Surgery*
N	538
Mean	3.1
1st	1.6
5th	1.9
10th	2.2
25th	2.5
Median	2.9
75th	3.5
90th	4.1
95th	4.5
99th	5.6
Outlier†	20 (3.7)
High	12

Measurement	*Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) Surgery*
Low	8

†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."

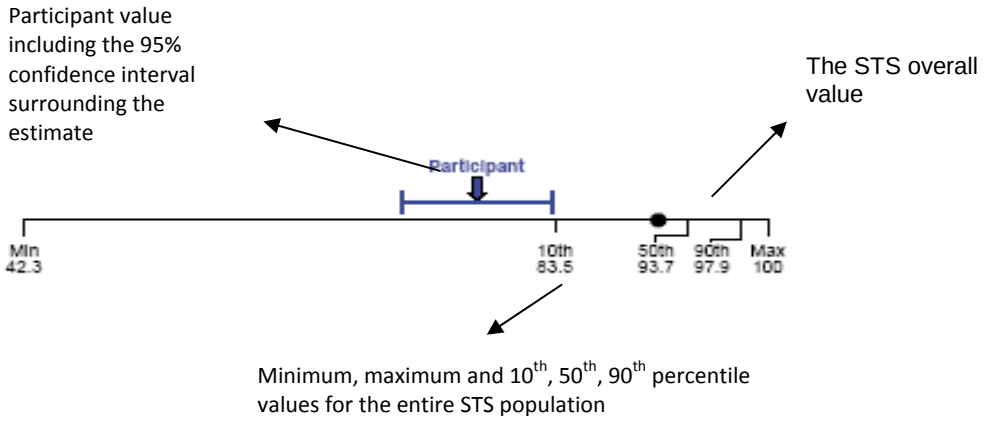
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

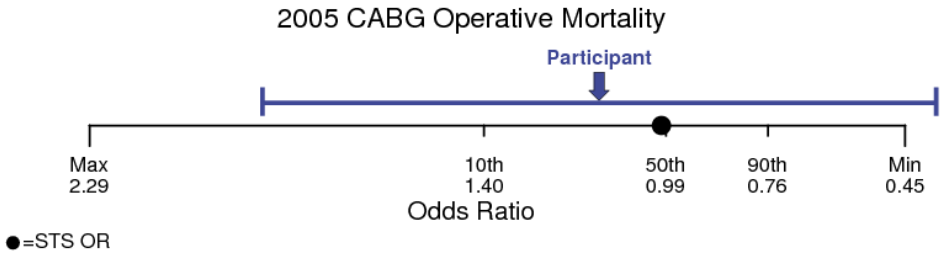
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

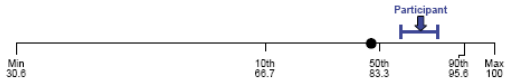
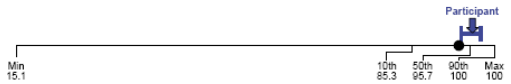
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 2—Isolated Valve Surgery

Sean M. O'Brien, PhD,[a] David M. Shahian, MD,[b] Giovanni Filardo, PhD, MPH,[c] Victor A. Ferraris, MD,[d] Constance K. Haan, MD,[e] Jeffrey B. Rich, MD,[f] Sharon-Lise T. Normand, PhD,[g] Elizabeth R. DeLong, PhD,[a] Cynthia M. Shewan, PhD,[h] Rachel S. Dokholyan, MPH,[a] Eric D. Peterson, MD, MPH,[a] Fred H. Edwards, MD,[e] and Richard P. Anderson, MD[i†]

[a]Duke Clinical Research Institute, Durham, North Carolina; [b]Massachusetts General Hospital, Boston, Massachusetts; [c]Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; [d]University of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; [e]University of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; [f]Sentara Cardiovascular Research Institute, Norfolk, Virginia; [g]Department of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [h]The Society of Thoracic Surgeons, Chicago, Illinois; and [i]Seattle, Washington

Background. **Adjustment for case-mix is essential when using observational data to compare surgical techniques or providers. That is most often accomplished through the use of risk models that account for preoperative patient factors that may impact outcomes. The Society of Thoracic Surgeons (STS) uses such risk models to create risk-adjusted performance reports for participants in the STS National Adult Cardiac Surgery Database (NCD). Although risk models were initially developed for coronary artery bypass surgery, similar models have now been developed for use with heart valve surgery, particularly as the proportion of such procedures has increased. The last published STS model for isolated valve surgery was based on data from 1994 to 1997 and did not include patients undergoing mitral valve repair. STS has developed new valve surgery models using contemporary data that include both valve repair as well as replacement. Expanding upon existing valve models, the new STS models include several nonfatal complications in addition to mortality.**

Methods. **Using STS data from 2002 to 2006, isolated valve surgery risk models were developed for operative mortality, permanent stroke, renal failure, prolonged ventilation (> 24 hours), deep sternal wound infection, reoperation for any reason, a major morbidity or mortality composite endpoint, prolonged postoperative length of stay, and short postoperative length of stay. The study population consisted of adult patients who underwent one of three types of valve surgery: isolated aortic valve replacement (n = 67,292), isolated mitral valve replacement (n = 21,229), or isolated mitral valve repair (n = 21,238). The population was divided into a 60% development sample and a 40% validation sample. After an initial empirical investigation, the three surgery groups were combined into a single logistic regression model with numerous interactions to allow the covariate effects to differ across these groups. Variables were selected based on a combination of automated stepwise selection and expert panel review.**

Results. **Unadjusted operative mortality (in-hospital regardless of timing, and 30-day regardless of venue) for all isolated valve procedures was 3.4%, and unadjusted in-hospital morbidity rates ranged from 0.3% for deep sternal wound infection to 11.8% for prolonged ventilation. The number of predictors in each model ranged from 10 covariates in the sternal infection model to 24 covariates in the composite mortality plus morbidity model. Discrimination as measured by the c-index ranged from 0.639 for reoperation to 0.799 for mortality. When patients in the validation sample were grouped into 10 categories based on deciles of predicted risk, the average absolute difference between observed versus predicted events within these groups ranged from 0.06% for deep sternal wound infection to 1.06% for prolonged postoperative stay.**

Conclusions. **The new STS risk models for valve surgery include mitral valve repair as well as multiple endpoints other than mortality. Model coefficients are provided and an online risk calculator is publicly available from The Society of Thoracic Surgeons website.**

(Ann Thorac Surg 2009;88:S23–42)

Models for predicting surgical outcomes on the basis of patient preoperative characteristics are valuable tools for research, quality improvement, and clinical practice. Such models are used by The Society of Thoracic Surgeons (STS) to produce risk-adjusted performance re-

†This author is deceased. Former Chair, Quality, Research and Patient Safety Council, The Society of Thoracic Surgeons, Chicago, IL.

Address correspondence to Dr Shahian, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

Drs O'Brien, Shahian, Filardo, Ferraris, Haan, Rich, Normand, DeLong, Shewan, Peterson, Edwards, Anderson, and Ms Dokholyan, have no conflicts of interest to declare regarding this work.

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.056

Abbreviations and Acronyms

AVR	= aortic valve replacement
CABG	= coronary artery bypass graft surgery
CI	= confidence interval
MI	= myocardial infarction
MVR	= mitral valve replacement
MVRepair	= mitral valve repair
NCD	= National Adult Cardiac Surgery Database
QMTF	= Quality Measurement Task Force
STS	= The Society for Thoracic Surgeons

ports for providers participating in the STS National Adult Cardiac Surgery Database (NCD). They are also used by STS surgeons and other physicians for counseling patients about the risk of surgery.

The earliest STS risk models were developed nearly 2 decades ago for isolated coronary artery bypass graft surgery (CABG). Subsequently, similar models have been developed for isolated valve replacement and combined CABG plus valve replacement. Because surgical practice and outcomes are changing rapidly, these models are updated periodically to reflect contemporary experience.

The last published STS model for isolated valve surgery was based on STS data from 1994 to 1997. The reference population included aortic and mitral valve replacements but excluded mitral valve repair, and the endpoint was operative mortality. In the decade since this model was published, many aspects of heart surgery have changed. First, as CABG volumes have decreased with the introduction of coronary stents, valve surgery as a proportion of overall heart surgery volume has increased in most practices. Between 2000 and 2006, the percentage of isolated CABG procedures decreased from 73% to 60% and the percentage of isolated valve procedures increased from 18% to 22%. Thus, in assessing provider performance, it is no longer sufficient only to consider isolated CABG surgery. Second, the frequency of mitral repair as a percentage of all isolated mitral operations in the STS NCD increased from 35% in 2000 to 53% in 2006. Third, during the same time period, the average mortality rate for isolated aortic or mitral surgery also decreased. Finally, efforts to measure and compare surgical performance have intensified and expanded. In addition to measuring operative mortality, performance reports increasingly focus on nonfatal complications as well as resource utilization and efficiency. Such outcomes have not historically been risk-adjusted for valve surgery.

The STS Quality Measurement Task Force (QMTF) has undertaken a complete revision of all STS risk models for adult cardiac surgery, and these new models were implemented in January 2008. This report, Part 2 of 3, describes the new STS models for isolated valve surgery (Part 1 describes the STS isolated CABG models, and Part 3 describes the models for CABG plus valve surgery). Authors of this report are the QMTF members who were involved in this initiative.

Two important features have been incorporated into these new models. First, the population includes mitral valve repair as well as aortic and mitral valve replacement. Second, in addition to operative mortality, the new models include six nonfatal in-hospital morbidity endpoints and two length-of-stay endpoints. In comparison with several other valve models that have recently been published [1–6], the STS models are distinguished by the large size of the development population and the broad spectrum of endpoints included.

Study Population and Endpoints

The population for this analysis consisted of operations on adult patients aged 20 to 100 years who underwent isolated single aortic or mitral valve surgery between January 1, 2002, and December 31, 2006. Only patients undergoing one of the following procedures were included: (1) isolated aortic valve replacement (AVR); (2) isolated mitral valve replacement (MVR); and (3) isolated mitral valve repair (MVRepair).

Because of the relatively small number of pulmonic, tricuspid, multiple valve procedures, and aortic repairs, these cases were not included in the current models. Patients undergoing concomitant CABG were excluded from the current analysis, but these were included in the separate STS valve plus CABG models described in Part 3 of this series. Records with missing data on sex (n = 44) were excluded because missing sex is not allowed in the analysis dataset used for creating STS database participant feedback reports. This left a final study population of 109,759 patient operations performed at 809 STS NCD participating groups. Patients on dialysis preoperatively (n = 2,699) were not included when developing the risk model for prediction of postoperative renal failure.

Patient characteristics in the study population are presented in Table 1.

Training and Validation Samples

The study population was randomly divided into a 60% training (development) sample and a 40% test (validation) sample. The development sample was used to identify predictor variables and estimate model coefficients. Data from the validation sample were used to assess model fit, discrimination, and calibration. After choosing variables and assessing model fit, the development and validation samples were subsequently combined, and the final model coefficients were estimated using the combined (development plus validation) data.

Endpoints

Risk models were developed for nine endpoints, identical to those in the STS CABG models. In contrast with the definition of operative mortality, which includes hospital deaths as well as deaths that occur after discharge within 30 days of surgery, the morbidity endpoints only include events that occurred before discharge. However, beginning with version 2.61, sternal infection data will be recorded for as long as 30 days postoperatively. The nine endpoints are as follows: (1) operative mortality: death during the same

Table 1. Distribution of Risk Factors in Overall Study Population Isolated Valve (2002–2006)

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Demographics								
Age, years								
< 55	28,147	25.6	13,227	19.66	6,601	31.09	8,319	39.17
55–64	23,258	21.2	12,987	19.30	4,833	22.77	5,438	25.61
65–74	28,145	25.6	18,299	27.19	5,294	24.94	4,552	21.43
≥75	30,209	27.5	22,779	33.85	4,501	21.20	2,929	13.79
Sex								
Male	60,752	55.4	39,209	58.27	9,055	42.65	12,488	58.80
Female	49,007	44.6	28,083	41.73	12,174	57.35	8,750	41.20
Race								
Caucasian	93,522	85.2	58,656	87.17	16,810	79.18	18,056	85.02
Black	7,630	7.0	3,555	5.28	2,383	11.23	1,692	7.97
Hispanic	3,680	3.4	2,344	3.48	889	4.19	447	2.10
Asian	1,538	1.4	719	1.07	437	2.06	382	1.80
Other	2,493	2.3	1,508	2.24	505	2.38	480	2.26
Missing	896	0.8	510	0.76	205	0.97	181	0.85
Risk factors								
Body surface area, m^2								
< 1.50	4,351	4.0	2,341	3.48	1,234	5.81	776	3.65
1.50–1.74	24,577	22.4	13,713	20.38	6,151	28.97	4,713	22.19
1.75–1.99	40,548	36.9	24,744	36.77	7,914	37.28	7,890	37.15
≥ 2.00	39,517	36.0	26,007	38.65	5,768	27.17	7,742	36.45
Missing	766	0.7	487	0.72	162	0.76	117	0.55
Body mass index, kg/m^2								
< 25	35,526	32.4	18,509	27.51	8,447	39.79	8,570	40.35
25–29	39,074	35.6	24,035	35.72	6,992	32.94	8,047	37.89
30–34	20,534	18.7	14,142	21.02	3,318	15.63	3,074	14.47
≥ 35	13,682	12.5	10,008	14.87	2,280	10.74	1,394	6.56
Missing	943	0.9	598	0.89	192	0.90	153	0.72
Diabetes mellitus								
No diabetes	88,709	80.8	52,052	77.35	17,535	82.60	19,122	90.04
Diabetes, noninsulin	14,900	13.6	11,026	16.39	2,412	11.36	1,462	6.88
Diabetes, insulin	5,788	5.3	3,974	5.91	1,216	5.73	598	2.82
Diabetes missing	138	0.1	91	0.14	34	0.16	13	0.06
Treatment missing	224	0.2	149	0.22	32	0.15	43	0.20
Hypertension								
No	41,649	37.9	22,338	33.20	8,859	41.73	10,452	49.21
Yes	67,886	61.9	44,816	66.60	12,326	58.06	10,744	50.59
Missing	224	0.2	138	0.21	44	0.21	42	0.20
Hypercholesterolemia								
No	59,003	53.8	33,156	49.27	12,857	60.56	12,990	61.16
Yes	50,328	45.9	33,865	50.33	8,286	39.03	8,177	38.50
Missing	428	0.4	271	0.40	86	0.41	71	0.33
Past or present smoker								
No	57,609	52.5	33,953	50.46	11,075	52.17	12,581	59.24
Yes	51,910	47.3	33,191	49.32	10,109	47.62	8,610	40.54
Missing	240	0.2	148	0.22	45	0.21	47	0.22
Chronic lung disease								
None	87,826	80.0	53,503	79.51	16,125	75.96	18,198	85.69
Mild	11,184	10.2	6,991	10.39	2,520	11.87	1,673	7.88
Moderate	6,346	5.8	4,022	5.98	1,494	7.04	830	3.91
Severe	3,332	3.0	2,110	3.14	853	4.02	369	1.74
Missing	1,071	1.0	666	0.99	237	1.12	168	0.79

Table 1. Continued

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Peripheral vascular disease								
No	101,129	92.1	61,222	90.98	19,550	92.09	20,357	95.85
Yes	8,381	7.6	5,909	8.78	1,641	7.73	831	3.91
Missing	249	0.2	161	0.24	38	0.18	50	0.24
Cerebrovascular disease								
No	96,852	88.2	58,983	87.65	18,158	85.53	19,711	92.81
Yes	12,661	11.5	8,147	12.11	3,033	14.29	1,481	6.97
Missing	246	0.2	162	0.24	38	0.18	46	0.22
CVA								
No CVA	101,631	92.6	62,518	92.91	18,833	88.71	20,280	95.49
Remote CVA (> 2 weeks)	6,926	6.3	4,203	6.25	1,912	9.01	811	3.82
Recent CVA (≤ 2 weeks)	818	0.7	325	0.48	409	1.93	84	0.40
CVA–missing timing	100	0.1	60	0.09	29	0.14	11	0.05
Missing	284	0.3	186	0.28	46	0.22	52	0.24
Endocarditis								
No endocarditis	100,998	92.0	63,257	94.00	17,926	84.44	19,815	93.30
Treated endocarditis	4,197	3.8	1,761	2.62	1,445	6.81	991	4.67
Active endocarditis	4,238	3.9	2,068	3.07	1,791	8.44	379	1.78
Endocarditis–missing type	63	0.1	30	0.04	27	0.13	6	0.03
Missing	263	0.2	176	0.26	40	0.19	47	0.22
Renal failure								
No	102,205	93.1	62,873	93.43	19,016	89.58	20,316	95.66
Yes	7,305	6.7	4,251	6.32	2,173	10.24	881	4.15
Missing	249	0.2	168	0.25	40	0.19	41	0.19
Renal function								
Creatinine < 1.00 mg/dL	42,028	38.3	25,679	38.16	7,754	36.53	8,595	40.47
Creatinine 1–1.49 mg/dL	51,939	47.3	32,058	47.64	9,372	44.15	10,509	49.48
Creatinine 1.50–1.99 mg/dL	8,081	7.4	5,078	7.55	1,875	8.83	1,128	5.31
Creatinine 2.00–2.49 mg/dL	1,946	1.8	1,192	1.77	512	2.41	242	1.14
Creatinine ≥ 2.50 mg/dL	1,294	1.2	750	1.11	390	1.84	154	0.73
Dialysis	2,699	2.5	1,464	2.18	900	4.24	335	1.58
Missing	1,772	1.6	1,071	1.59	426	2.01	275	1.29
Immunosuppressive treatment								
No	106,037	96.6	64,953	96.52	20,356	95.89	20,728	97.60
Yes	3,336	3.0	2,074	3.08	819	3.86	443	2.09
Missing	386	0.4	265	0.39	54	0.25	67	0.32
Previous CV interventions								
Previous coronary artery bypass surgery								
No	98,978	90.2	60,351	89.69	18,564	87.45	20,063	94.47
Yes	10,399	9.5	6,713	9.98	2,569	12.10	1,117	5.26
Missing	382	0.3	228	0.34	96	0.45	58	0.27
Previous valve surgery								
No	100,179	91.3	62,898	93.47	16,857	79.41	20,424	96.17
Yes	9,227	8.4	4,186	6.22	4,285	20.18	756	3.56
Missing	353	0.3	208	0.31	87	0.41	58	0.27
Previous other cardiac surgery								
No	105,686	96.3	65,084	96.72	20,034	94.37	20,568	96.85
Yes	3,662	3.3	1,975	2.93	1,077	5.07	610	2.87
Missing	411	0.4	233	0.35	118	0.56	60	0.28
Number of previous CV surgeries								
No prior CV surgery	91,196	83.1	56,629	84.15	15,239	71.78	19,328	91.01
1 prior CV surgery	15,399	14.0	9,122	13.56	4,775	22.49	1,502	7.07
2 or more prior CV surgeries	2,653	2.4	1,260	1.87	1,069	5.04	324	1.53
Missing	511	0.5	281	0.42	146	0.69	84	0.40

Table 1. Continued

Variable	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
	N	%	N	%	N	%	N	%
Prior PCI								
No PCI	101,878	92.8	62,145	92.35	19,573	92.20	20,160	94.92
PCI within 6 hours	122	0.1	58	0.09	51	0.24	13	0.06
PCI not within 6 hours	7,100	6.5	4,678	6.95	1,447	6.82	975	4.59
PCI–missing timing	133	0.1	90	0.13	28	0.13	15	0.07
Missing	526	0.5	321	0.48	130	0.61	75	0.35
Preoperative cardiac status								
Acuity status								
Elective	84,052	76.6	51,734	76.88	14,293	67.33	18,025	84.87
Urgent	23,795	21.7	14,670	21.80	6,071	28.60	3,054	14.38
Emergent	1,555	1.4	685	1.02	747	3.52	123	0.58
Emergent salvage	154	0.1	70	0.10	78	0.37	6	0.03
Missing	203	0.2	133	0.20	40	0.19	30	0.14
MI								
No prior MI	99,416	90.6	60,850	90.43	18,716	88.16	19,850	93.46
MI > 21 days	7,785	7.1	4,770	7.09	1,848	8.71	1,167	5.49
MI 8–21 days	719	0.7	480	0.71	170	0.80	69	0.32
MI 1–7 days	1,247	1.1	863	1.28	315	1.48	69	0.32
MI > 6 and < 24 hours	142	0.1	61	0.09	66	0.31	15	0.07
MI ≤ 6 hours	90	0.1	42	0.06	40	0.19	8	0.04
MI–missing timing	127	0.1	79	0.12	33	0.16	15	0.07
Missing	233	0.2	147	0.22	41	0.19	45	0.21
Angina								
No	85,364	77.8	49,573	73.67	17,598	82.90	18,193	85.66
Yes	24,164	22.0	17,577	26.12	3,591	16.92	2,996	14.11
Missing	231	0.2	142	0.21	40	0.19	49	0.23
Cardiogenic shock								
No	108,163	98.5	66,646	99.04	20,460	96.38	21,057	99.15
Yes	1,329	1.2	485	0.72	725	3.42	119	0.56
Missing	267	0.2	161	0.24	44	0.21	62	0.29
Resuscitation								
No	108,958	99.3	66,832	99.32	20,992	98.88	21,134	99.51
Yes	533	0.5	297	0.44	186	0.88	50	0.24
Missing	268	0.2	163	0.24	51	0.24	54	0.25
Arrhythmia								
No arrhythmia	89,779	81.8	57,451	85.38	14,604	68.79	17,724	83.45
AFib/flutter	16,124	14.7	7,569	11.25	5,721	26.95	2,834	13.34
Heart block	1,598	1.5	1,109	1.65	315	1.48	174	0.82
Sustained VT/VF	984	0.9	486	0.72	290	1.37	208	0.98
Arrhythmia–other	688	0.6	324	0.48	175	0.82	189	0.89
Arrhythmia–missing type	312	0.3	175	0.26	74	0.35	63	0.30
Missing	274	0.2	178	0.26	50	0.24	46	0.22
Preoperative IABP								
No	107,945	98.3	66,733	99.17	20,332	95.77	20,880	98.31
Yes	1,431	1.3	342	0.51	809	3.81	280	1.32
Missing	383	0.3	217	0.32	88	0.41	78	0.37
NYHA class								
I	17,413	15.9	10,222	15.19	2,706	12.75	4,485	21.12
II	32,360	29.5	20,295	30.16	4,915	23.15	7,150	33.67
III	40,321	36.7	25,483	37.87	8,205	38.65	6,633	31.23
IV	14,324	13.1	8,104	12.04	4,256	20.05	1,964	9.25
Missing	5,341	4.9	3,188	4.74	1,147	5.40	1,006	4.74

Table 1. Continued

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Congestive heart failure								
No	64,608	58.9	41,972	62.37	9,341	44.00	13,295	62.60
Yes	44,934	40.9	25,185	37.43	11,849	55.82	7,900	37.20
Missing	217	0.2	135	0.20	39	0.18	43	0.20
Number of diseased coronary vessels								
None	90,281	82.3	55,072	81.84	17,525	82.55	17,684	83.27
One	8,947	8.2	5,393	8.01	1,498	7.06	2,056	9.68
Two	3,386	3.1	2,180	3.24	735	3.46	471	2.22
Three	5,611	5.1	3,766	5.60	1,147	5.40	698	3.29
Missing	1,534	1.4	881	1.31	324	1.53	329	1.55
Left main disease ≥ 50%								
No	106,462	97.0	65,328	97.08	20,495	96.54	20,639	97.18
Yes	1,625	1.5	1,127	1.67	289	1.36	209	0.98
Missing	1,672	1.5	837	1.24	445	2.10	390	1.84
Ejection fraction, %								
< 25	2,694	2.5	1,774	2.64	341	1.61	579	2.73
25–34	5,900	5.4	3,810	5.66	1,052	4.96	1,038	4.89
35–44	10,035	9.1	6,181	9.19	2,208	10.40	1,646	7.75
45–54	20,481	18.7	12,411	18.44	4,382	20.64	3,688	17.37
≥ 55	60,890	55.5	36,584	54.37	11,308	53.27	12,998	61.20
Missing	9,759	8.9	6,532	9.71	1,938	9.13	1,289	6.07
Aortic stenosis								
No	54,457	49.6	13,309	19.78	20,303	95.64	20,845	98.15
Yes	54,681	49.8	53,722	79.83	696	3.28	263	1.24
Missing	621	0.6	261	0.39	230	1.08	130	0.61
Mitral stenosis								
No	100,609	91.7	65,186	96.87	15,383	72.46	20,040	94.36
Yes	8,155	7.4	1,401	2.08	5,676	26.74	1,078	5.08
Missing	995	0.9	705	1.05	170	0.80	120	0.57
Tricuspid stenosis								
No	108,073	98.5	66,243	98.44	20,821	98.08	21,009	98.92
Yes	331	0.3	152	0.23	120	0.57	59	0.28
Missing	1,355	1.2	897	1.33	288	1.36	170	0.80
Pulmonic stenosis								
No	107,512	98.0	65,842	97.85	20,783	97.90	20,887	98.35
Yes	141	0.1	91	0.14	29	0.14	21	0.10
Missing	2,106	1.9	1,359	2.02	417	1.96	330	1.55
Aortic insufficiency								
None	59,905	54.6	25,861	38.43	16,701	78.67	17,343	81.66
Trivial	9,191	8.4	5,916	8.79	1,661	7.82	1,614	7.60
Mild	13,282	12.1	10,014	14.88	1,798	8.47	1,470	6.92
Moderate	9,501	8.7	8,815	13.10	382	1.80	304	1.43
Severe	15,722	14.3	15,529	23.08	109	0.51	84	0.40
Missing	2,158	2.0	1,157	1.72	578	2.72	423	1.99
Mitral insufficiency								
None	43,731	39.8	40,453	60.12	2,283	10.75	995	4.68
Trivial	7,743	7.1	7,285	10.83	388	1.83	70	0.33
Mild	14,455	13.2	13,066	19.42	1,089	5.13	300	1.41
Moderate	10,224	9.3	4,438	6.60	3,246	15.29	2,540	11.96
Severe	31,813	29.0	573	0.85	14,045	66.16	17,195	80.96
Missing	1,793	1.6	1,477	2.19	178	0.84	138	0.65

Table 1. Continued

Variable	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
	N	%	N	%	N	%	N	%
Tricuspid insufficiency								
None	78,472	71.5	49,976	74.27	14,266	67.20	14,230	67.00
Trivial	8,856	8.1	5,612	8.34	1,381	6.51	1,863	8.77
Mild	13,346	12.2	7,333	10.90	2,788	13.13	3,225	15.19
Moderate	5,167	4.7	2,126	3.16	1,753	8.26	1,288	6.06
Severe	974	0.9	297	0.44	460	2.17	217	1.02
Missing	2,944	2.7	1,948	2.89	581	2.74	415	1.95
Pulmonic insufficiency								
None	97,954	89.2	60,463	89.85	18,837	88.73	18,654	87.83
Trivial	4,161	3.8	2,370	3.52	779	3.67	1,012	4.77
Mild	2,541	2.3	1,340	1.99	573	2.70	628	2.96
Moderate	441	0.4	209	0.31	144	0.68	88	0.41
Severe	76	0.1	34	0.05	30	0.14	12	0.06
Missing	4,586	4.2	2,876	4.27	866	4.08	844	3.97

AFib = atrial fibrillation; AVR = aortic valve replacement; CV = cardiovascular; CVA = cerebrovascular accident (stroke); IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repaid; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; VF = ventricular fibrillation; VT = ventricular tachycardia.

hospitalization as surgery, regardless of timing, or within 30 days of surgery regardless of venue; (2) permanent stroke (cerebrovascular accident [CVA]): a central neurologic deficit persisting longer than 72 hours; (3) renal failure: a new requirement for dialysis or an increase of the serum creatinine to greater than 2.0 mg/dL and double the most recent preoperative creatinine level; (4) prolonged ventilation (longer than 24 hours); (5) deep sternal wound infection; (6) reoperation for any reason; (7) major morbidity or mortality: a composite defined as the occurrence of any of the above endpoints; (8) prolonged postoperative length of stay (PLOS): length of stay (LOS) more than 14 days (alive or dead); and (9) short postoperative LOS (SLOS): LOS less than 6 days and patient alive at discharge.

Table 2 summarizes the endpoint frequencies in the study population.

Single Versus Multiple Models

Two issues required particularly careful consideration: whether to construct separate models for the AVR and MVR populations, and how best to further subdivide the mitral population into repair versus replacement.

Because of the large size of the STS NCD, separate

Table 2. Frequency of Endpoints in Overall Study Population 2002 to 2006

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
All isolated valve (AVR, MVR, MVRepair)									
N	109,759	109,759	107,060	109,759	109,759	109,759	109,759	109,759	109,759
Events	3,706	1,751	4,673	12,892	307	9,164	20,074	9,718	41,214
%	3.4	1.6	4.3	11.8	0.3	8.4	18.3	8.9	37.6
AVR									
N	67,292	67,292	65,828	67,292	67,292	67292	67,292	67,292	67,292
Events	2,157	1,007	2,774	7,323	197	5369	11,706	5,308	26,144
%	3.2	1.5	4.1	10.9	0.3	8.0	17.4	7.9	38.9
MVR									
N	21,229	21,229	20,329	21,229	21,229	21229	21,229	21,229	21,229
Events	1,210	447	1,348	4,015	71	2450	5,675	3,244	4,727
%	5.7	2.1	6.4	18.9	0.3	11.5	26.7	15.3	22.3
MVRepair									
N	21,238	21,238	20,903	21,238	21,238	21,238	21,238	21,238	21,238
Events	339	297	551	1,554	39	1,345	2,693	1,166	10,343
%	1.6	1.4	2.6	7.3	0.2	6.3	12.7	5.5	48.7

AVR = aortic valve replacement; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

models for AVR, MVR, and MVRepair initially seemed both feasible and appropriate. However, because the endpoints of interest are rare events, we recognized the possibility that the number of such events would be too small to support reliable estimation of the model coefficients.

To assess this tradeoff, we conducted a pilot study to compare two alternative strategies for developing risk models for isolated valve surgery. The first strategy involved developing models separately for three subpopulations (AVR, MVR, and MVRepair). The second strategy involved modeling all three subpopulations together in a single model; several interaction terms were included to allow the effect of selected risk factors to differ across the subpopulations. Both strategies were pilot tested by developing risk models for two endpoints: operative mortality and permanent stroke. These pilot models were developed in a 60% development sample and tested in a separate 40% validation sample. Each model was assessed by calculating the c-index and the generalized R^2 index of Nagelkerke [7] in the validation sample for each combination of subpopulation and endpoint (3 subpopulations $\times$ 2 endpoints = 6 combinations). With the exception of AVR operative mortality, the combined model with interactions resulted in better discrimination. With the exception of MVR and MVrepair operative mortality, the combined model also captured more variation as measured by the generalized R^2 statistic.

Because the combined model strategy performed better in the majority of cases, and because a single combined model was consistent with the previous STS valve model, the combined model strategy was selected. To avoid assuming that the weighting of each risk factor was exactly constant across the three populations, we included interactions between surgery type and several key predictor variables. In principle, fitting a single model with several interactions is advantageous because it allows for pooling information across related groups without making an a priori assumption that all of the covariate effects are exactly constant across groups.

Selection of Candidate Predictor Variables

Our general approach to variable selection is discussed in Part 1 of this series describing the development of the 2008 STS isolated CABG risk models. Briefly, we initially identified potential candidate variables by reviewing four versions of the STS data collection instrument (data versions 2.35, 2.41, 2.52.1, and 2.61) as well as previously published STS and similar cardiac risk models [1–6]. A panel of cardiac surgeons and health policy experts reviewed the initial variables for face validity and to be certain that no important predictor variables available in (or mappable to) to STS NCD data version 2.61 had been excluded.

Final candidate explanatory variables and their coding are summarized in Table 3. The variables were identical to the CABG model candidate variables with the following differences: (1) percutaneous coronary intervention conducted within 6 hours or less of surgery was not a candidate variable because it was present in only 122 patients (0.1%) in the valve model population; (2) infectious endocarditis was included. This risk factor was rarely present among isolated CABG patients (0.09%), but was not uncommon (7.7%) among patients undergoing valve surgery; (3) mitral stenosis was included; this risk factor was rarely present among isolated CABG patients (0.35%) but was common (7.4%) among patients undergoing valve surgery; and (4) an indicator for surgery type (AVR, MVR, MVRepair) was included in the valve models.

Coding of Explanatory Variables

The coding of continuous and categorical variables was identical to the CABG models, except for the following differences: (1) age was modeled as a linear spline truncated from below at 50 years and with a change of slope at 75; (2) creatinine was modeled as a linear term with values less than 0.5 and greater than 5.0 mapped to those values respectively (approximately the 1st and 99th percentiles of the empirical distribution); (3) previous myocardial infarction (MI) was modeled as three categories ($<$ 24 hours, 1 to 21 days, and $>$ 21 days or no MI); the first two categories were subsequently combined after expert panel review; (4) race was modeled as three categories: black, Hispanic, Caucasian/other; and (5) chronic lung disease was modeled as linear across four categories (none, mild, moderate, severe).

In general, these differences reflect a slightly simpler coding scheme (fewer parameters) for the valve models compared with the isolated CABG models.

Repair Versus Replacement

In addition to a number of variables whose inclusion or coding were noted to be problematic during development of the 2008 STS isolated CABG models (Part 1 of this series), the approach to modeling mitral valve repair versus replacement was of some concern in the valve models. From a methodologic perspective, models used for risk-adjustment should include all patient preoperative risk factors that vary in prevalence between institutions and that substantially impact the probability of an adverse outcome. Such models should include variables that reflect the patient's baseline condition but should not include intraoperative events (eg, unexpected hemorrhage) or discretionary care processes (eg, use of a mechanical versus bioprosthetic valve). Adjusting for intraoperative events is not appropriate because these may be a reflection of the surgeon's performance. Adjusting for discretionary care processes may likewise mask differences in performance if the surgeon's choice of procedures has a substantial impact on outcomes. The same patient may receive valve repair if treated by one surgeon and replacement if treated by another. Adjusting for repair versus replacement will potentially conceal the outcomes of surgeons who achieve excellent results by repairing technically challenging valves that might otherwise be replaced if treated by a surgeon with less skill or tenacity. Importantly, there is considerable evidence to suggest the superiority of valve repair whenever feasible.

However, in addition to such discretionary factors, the decision to repair rather than replace the mitral valve is

Table 3. List of Final Candidate Variables and Their Coding for STS Valve Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline truncated from below at 50 and with knot at 75
Ejection fraction	Linear, values > 50 mapped to 50
Body surface area[a]	Quadratic polynomial modeled separately for males and females. Note: body surface area < 1.4 and > 2.6 mapped to those values, respectively.
Creatinine	Linear (only for patients not on dialysis). Note: creatinine < 0.5 and > 5.0 mapped to those values, respectively.
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval. Modeled as linear across categories.
Binary variables	
Active infectious endocarditis	Yes/no
Dialysis	Yes/no
Preoperative atrial fibrillation	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Preoperative IABP or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no MI < 7 days)	Yes/no
Left main disease	Yes/no
Aortic stenosis	Yes/no
Mitral stenosis	Yes/no
Aortic insufficiency	Defined as at least moderate (yes/no)
Mitral insufficiency	Defined as at least moderate (yes/no)
Tricuspid insufficiency	Defined as at least moderate (yes/no)
Categorical variables	
Chronic lung disease	Modeled as linear across categories (none, mild, moderate, severe)
CVD/CVA	3 groups: no CVD, CVD no CVA, CVD + CVA
Diabetes mellitus	3 groups: insulin diabetes, noninsulin diabetes, other or no diabetes
Number diseased coronary vessels	3 groups: < 2, 2, 3. Modeled as linear across the categories
MI	3 groups: < 24 hr, 1–21 days, > 21 days or no MI (groups 1 and 2 were subsequently collapsed)
Race	3 groups: Black; Hispanic; Other including Caucasian
Status	4 groups: elective, urgent, emergent—no resuscitation, salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, ≥2 previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF+NYHA IV
Surgery type	3 groups: AVR, MVR, MVRepair
Interaction terms	
Age by reoperation[a]	
Age by emergent status[a]	
Surgery type by each of the following:	Age, diabetes, dialysis, creatinine, reoperation, endocarditis, emergent status, CLD, CHF, EF, sex, shock, IABP/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, aortic stenosis

[a] These variables were forced into each model.

AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association.

also dependent upon the patient's preoperative valve disease etiology, anatomy, and pathophysiology. On average, patients amenable to valve repair have less extensive valve pathology and a relatively favorable postoperative prognosis (the mortality rate for valve repair is 1.6% compared with 5.7% for replacement). Ignoring these anatomical differences can introduce bias when comparing institutions, especially because these variables are not captured elsewhere on the STS data collection form.

A related difficulty in adjusting for repair versus re-

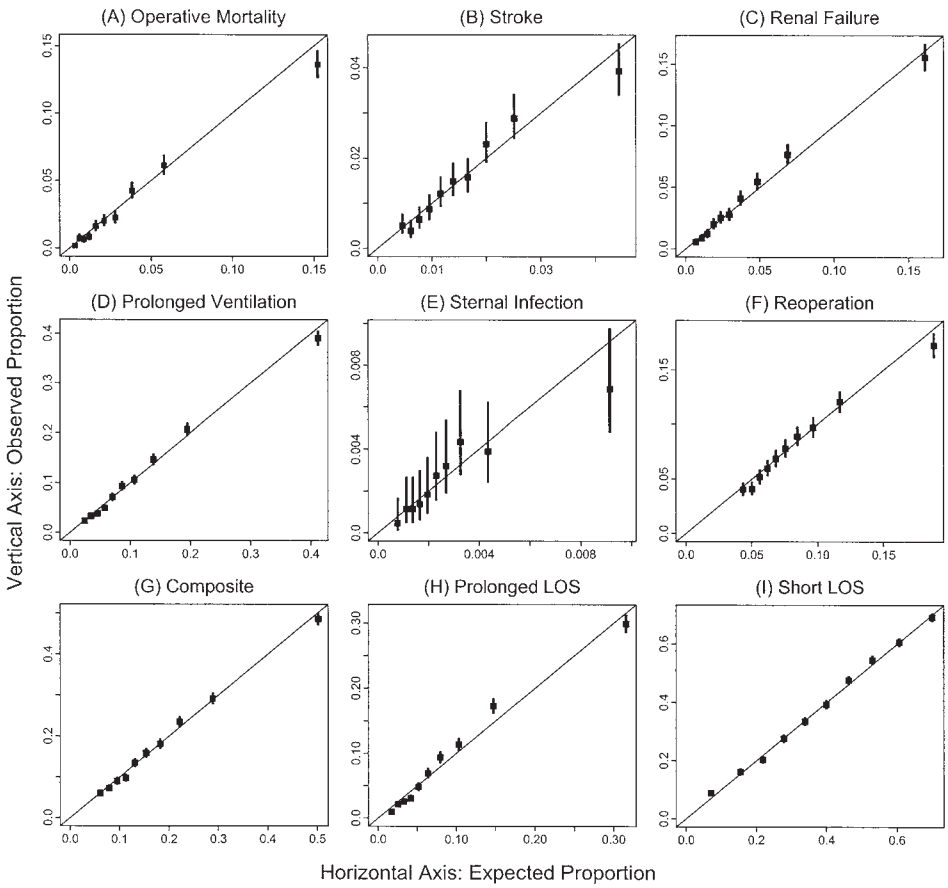


Fig 1. Plots of observed (O) versus expected (E) in validation sample

placement is that the former approach may sometimes be abandoned intraoperatively by the surgeon and converted to MVR. That may sometimes occur because of unforeseen technical problems that would prevent most surgeons from completing the repair, but in other instances, a more skilled surgeon might persist and achieve successful valve repair. Effectively separating these two scenarios is problematic from available data.

Ultimately, it was elected to include an indicator for mitral valve repair versus replacement in the valve risk models, consistent with the approach in a number of existing valve surgery models. We acknowledge that available data make it impossible to determine whether patient differences or surgical skill and judgment are the most important factors in determining between-provider variation in the proportion of valves repaired.

Recognizing the potential limitations of this modeling approach, the decision to adjust for repair versus replacement may be reassessed in future versions of the STS risk models. Beginning with data in version 2.61, the database will capture whether or not repair was attempted, and repair versus replacement may be analyzed based on an intention-to-treat principle.

Missing Data

Model variables with more than 1% missing data in the study sample were ejection fraction (8.9%), NYHA class (4.9%), tricuspid insufficiency (2.7%), aortic insufficiency (2.0%), mitral insufficiency (1.6%), left main disease (1.5%), creatinine/dialysis (1.6%), and number of diseased vessels (1.4%). The method of imputing missing data was identical to that employed in the isolated CABG models and described in Part 1 of this series. Briefly, binary risk factors were modeled as yes versus no or missing (ie, missing values were analyzed as if the endpoint did not occur). Missing data on categorical variables were imputed to the lowest risk value, typically the mode, and outcomes were typically similar for missing data and lowest risk patients. Missing data on continuous variables were imputed by grouping patients into strata and assigning the stratum-specific median value. For example, ejection fraction was imputed by grouping on sex and congestive heart failure and calculating the median ejection fraction among patients with nonmissing ejection fraction in each group.

Although multiple imputation is generally preferable to single imputation [8], single imputation was chosen for this analysis mainly because of practical considerations. Furthermore, because of the small fraction of missing data, the impact of single versus multiple imputation was considered to be inconsequential. Subsequent sensitivity analyses confirmed that the choice between single versus multiple imputation had little impact on the final regression coefficients, risk estimates, and confidence intervals. A summary of these sensitivity analyses, including coef-

Table 4. Discrimination of Models in Development and Validation Samples

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall									
Development sample	0.805	0.694	0.782	0.770	0.704	0.643	0.721	0.770	0.738
Validation sample	0.799	0.691	0.762	0.762	0.659	0.639	0.718	0.773	0.734
AVR									
Development sample	0.779	0.679	0.766	0.748	0.710	0.630	0.698	0.752	0.713
Validation sample	0.759	0.689	0.749	0.736	0.637	0.619	0.694	0.759	0.713
MVR									
Development sample	0.794	0.679	0.767	0.772	0.591	0.642	0.735	0.748	0.726
Validation sample	0.802	0.702	0.748	0.772	0.656	0.634	0.738	0.729	0.710
MVRepair									
Development sample	0.855	0.736	0.813	0.765	0.774	0.616	0.703	0.777	0.733
Validation sample	0.844	0.672	0.788	0.773	0.714	0.646	0.712	0.800	0.725

AVR = aortic valve replacement; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

ficients and covariance matrices, is available at www.sts.org/riskmodels.

Final Variable Selection Procedure

Variables were initially selected using an automated stepwise model selection algorithm. The stepwise procedure began with a model that included all of the final candidate variables except for interaction terms. Age, sex, body surface area, and month of surgery were forced into each model. Other variables were selected in a stepwise fashion using a significance criterion of 0.05 for entry and removal. This criterion was less stringent than that employed in development of the CABG models, because the sample size in the former was so much larger than that which was used for the valve models. The stepwise procedure was performed separately for each endpoint. The results were then reviewed by an expert panel of surgeons, and the following changes were made based on their feedback: (1) "MI less than 24 hours" and "MI 1 to 21 days" were collapsed into a single category; (2) preoperative atrial fibrillation was forced into the model for stroke (CVA); and (3) an indicator variable for dialysis was forced into any model that included creatinine level.

Interaction Terms

In addition to including main effects, we tested the interaction between surgery group (AVR, MVR, MVRepair) and each of the following variables: age, diabetes mellitus, dialysis, creatinine, reoperation, endocarditis, emergent status, chronic lung disease, congestive heart failure, ejection fraction, sex, shock, intra-aortic balloon pump/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, and aortic stenosis. These interaction terms allowed the effect of these selected risk factors to differ across the surgery populations.

Four additional sets of interactions were also included in the models: (1) sex by body surface area (BSA); (2) sex by BSA^2; (3) age by reoperation; and (4) age by emergent status. These interaction terms were preselected and were not tested as part of the backward selection algorithm. Additional technical details are provided in the Appendix. For reasons described in Part 1 of this series (isolated CABG risk models), an extensive automated search for additional interaction terms was not conducted.

Adjustment for Time Trends

Surgery date was included in each model to adjust for changes in the frequency of adverse outcomes over the 5-year study period. Although surgery date is not itself a variable of interest, we adjusted for it to reduce potential confounding by time trends when estimating regression coefficients for the variables that are of primary interest (ie, patient preoperative risk factors). An example is provided in Part 1 of this series.

Surgery date was categorized into 6-month intervals (corresponding to the biannual STS data harvests) and modeled as a linear trend across the ordinal categories. Because it is a nuisance variable, surgery date is not included in the final risk prediction algorithm. Thus, a patient's predicted risk does not depend on the patient's surgery date. As described in the Appendix, the published intercept parameter has been adjusted to incorporate the time trend. The adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Results

Assessment of Model Fit and Discrimination

Because of the relatively large size of our sample, the Hosmer-Lemeshow test is uninformative and would invariably result in a significant p value [9]. As an alternative, model fit was assessed graphically by plotting observed versus predicted rates of each endpoint across deciles of predicted risk in the development and validation samples. This was done in the overall population and in subgroups based on surgery type (AVR, MVR, MVRepair); age ($<$ 60, 60 to 79, $\geq$ 80 years); sex (male, female); diabetes mellitus (yes/no); status (elective, nonelective); and ejection fraction

Table 5. Odds Ratios (95% Confidence Intervals) for the Final Selected Models

A. Odds ratios for variables that do not interact with surgery group

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.10, 1.31)	1.06 (0.93, 1.20)	NA	1.18 (1.11, 1.25)	NA	1.11 (1.04, 1.18)	1.12 (1.07, 1.18)	1.17 (1.10, 1.24)	0.74 (0.70, 0.78)
BSA 1.6 versus 2.0 among females	1.19 (1.09, 1.30)	1.18 (1.03, 1.35)	0.95 (0.87, 1.04)	1.15 (1.08, 1.22)	0.42 (0.27, 0.68)	1.26 (1.18, 1.34)	1.17 (1.12, 1.23)	1.11 (1.04, 1.17)	0.99 (0.95, 1.04)
BSA 1.6 versus 2.0 among males	1.75 (1.48, 2.07)	1.17 (0.92, 1.47)	1.33 (1.12, 1.58)	1.56 (1.41, 1.74)	0.94 (0.49, 1.84)	1.34 (1.21, 1.49)	1.44 (1.33, 1.57)	1.39 (1.25, 1.56)	0.73 (0.68, 0.79)
BSA 1.8 versus 2.0 among females	0.99 (0.95, 1.04)	1.08 (0.99, 1.17)	0.90 (0.86, 0.94)	1.00 (0.97, 1.03)	0.65 (0.54, 0.77)	1.07 (1.03, 1.11)	1.02 (0.99, 1.04)	0.99 (0.96, 1.02)	1.05 (1.03, 1.08)
BSA 1.8 versus 2.0 among males	1.21 (1.14, 1.29)	1.07 (0.98, 1.16)	1.07 (1.00, 1.14)	1.14 (1.10, 1.19)	0.90 (0.70, 1.14)	1.12 (1.08, 1.16)	1.12 (1.09, 1.16)	1.10 (1.06, 1.15)	0.92 (0.89, 0.94)
BSA 2.2 versus 2.0 among females	1.21 (1.11, 1.33)	0.94 (0.80, 1.10)	1.30 (1.21, 1.41)	1.15 (1.09, 1.21)	1.57 (1.26, 1.96)	1.02 (0.95, 1.09)	1.12 (1.07, 1.16)	1.14 (1.08, 1.21)	0.85 (0.81, 0.88)
BSA 2.2 versus 2.0 among males	0.98 (0.93, 1.03)	0.95 (0.88, 1.03)	1.09 (1.05, 1.14)	1.05 (1.02, 1.08)	1.32 (1.17, 1.48)	0.95 (0.93, 0.98)	1.02 (0.99, 1.04)	1.03 (1.00, 1.07)	0.94 (0.93, 0.96)
Creatinine per 1 unit	1.55 (1.46, 1.64)	1.34 (1.22, 1.47)	2.04 (1.93, 2.16)	1.58 (1.51, 1.65)	NA	1.27 (1.20, 1.33)	1.64 (1.57, 1.71)	1.58 (1.51, 1.65)	0.64 (0.61, 0.68)
CVD with CVA	NA	1.81 (1.56, 2.10)	1.22 (1.09, 1.37)	1.28 (1.18, 1.38)	NA	1.14 (1.05, 1.24)	1.20 (1.12, 1.28)	1.40 (1.29, 1.52)	0.77 (0.72, 0.83)
CVD without CVA	NA	1.32 (1.11, 1.57)	1.23 (1.10, 1.37)	1.14 (1.05, 1.23)	NA	1.06 (0.96, 1.17)	1.08 (1.01, 1.15)	NA	0.80 (0.73, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	NA	1.10 (1.01, 1.20)	NA	1.07 (1.02, 1.11)	NA	NA	1.04 (1.00, 1.08)	1.03 (0.98, 1.08)	0.90 (0.86, 0.94)
EF per 10-unit decrease	1.09 (1.05, 1.14)	NA	1.04 (1.00, 1.09)	1.12 (1.09, 1.15)	1.26 (1.12, 1.41)	1.08 (1.04, 1.11)	1.10 (1.07, 1.12)	1.12 (1.08, 1.15)	0.87 (0.85, 0.90)
Hypertension	1.12 (1.03, 1.22)	1.19 (1.07, 1.33)	1.35 (1.25, 1.45)	1.11 (1.06, 1.17)	NA	NA	1.11 (1.07, 1.15)	NA	0.94 (0.91, 0.97)
Immunosuppressive treatment	1.42 (1.21, 1.67)	NA	1.39 (1.19, 1.62)	NA	NA	NA	1.16 (1.06, 1.27)	1.31 (1.17, 1.47)	NA
Left main disease	1.19 (0.98, 1.46)	NA	1.19 (0.98, 1.44)	NA	2.17 (1.13, 4.16)	NA	NA	NA	NA
Active infectious endocarditis	1.95 (1.68, 2.27)	1.87 (1.52, 2.29)	2.17 (1.88, 2.50)	2.15 (1.95, 2.36)	NA	1.55 (1.39, 1.73)	1.97 (1.80, 2.15)	2.79 (2.51, 3.09)	0.34 (0.30, 0.38)
Mitral insufficiency, moderate/severe	NA	1.26 (1.14, 1.39)	NA	NA	NA	NA	NA	NA	NA
Tricuspid insufficiency, moderate/severe	NA	NA	1.14 (1.01, 1.29)	1.14 (1.04, 1.25)	NA	1.09 (1.00, 1.20)	1.21 (1.12, 1.30)	1.17 (1.05, 1.31)	0.82 (0.73, 0.92)
Peripheral vascular disease	1.25 (1.12, 1.38)	1.29 (1.11, 1.49)	NA	NA	NA	1.22 (1.12, 1.32)	1.14 (1.07, 1.21)	1.17 (1.09, 1.25)	0.83 (0.78, 0.88)
Aortic stenosis		NA	NA	0.90 (0.83, 0.97)	NA	0.90 (0.84, 0.96)	0.93 (0.87, 0.98)	0.86 (0.80, 0.92)	1.07 (1.02, 1.13)
Mitral stenosis	1.24 (1.08, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA
MI $\leq$ 21 days	1.14 (0.98, 1.34)	NA	NA	1.37 (1.22, 1.55)	NA	1.04 (0.91, 1.18)	1.28 (1.16, 1.41)	1.21 (1.06, 1.37)	0.81 (0.72, 0.91)
Time trend, per 6-month harvest interval	0.98 (0.97, 0.99)	0.98 (0.96, 1.00)	1.01 (0.99, 1.02)	1.02 (1.01, 1.03)	0.97 (0.93, 1.01)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Race black	NA	1.33 (1.13, 1.57)	1.51 (1.34, 1.69)	1.42 (1.27, 1.58)	NA	1.27 (1.15, 1.40)	1.37 (1.27, 1.49)	1.45 (1.31, 1.60)	0.64 (0.59, 0.70)
Race Hispanic	NA	0.87 (0.64, 1.19)	1.16 (0.97, 1.38)	1.07 (0.94, 1.22)	NA	1.14 (1.00, 1.30)	1.09 (0.98, 1.22)	1.16 (0.98, 1.38)	0.82 (0.72, 0.93)
Status urgent	1.29 (1.19, 1.40)	NA	1.21 (1.11, 1.33)	1.29 (1.20, 1.39)	NA	1.17 (1.10, 1.25)	1.22 (1.15, 1.29)	1.42 (1.33, 1.51)	0.70 (0.66, 0.74)
Unstable angina	1.21 (1.04, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA

Table 5. Continued

B. Odds ratios for aortic valve replacement

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.43 (1.34, 1.52)	1.48 (1.38, 1.59)	1.38 (1.30, 1.47)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.16 (1.12, 1.21)	1.23 (1.19, 1.26)	1.31 (1.25, 1.37)	0.75 (0.73, 0.77)
Age 70 versus 50	2.04 (1.79, 2.32)	2.19 (1.90, 2.52)	1.90 (1.68, 2.16)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.35 (1.25, 1.46)	1.50 (1.42, 1.59)	1.71 (1.55, 1.87)	0.57 (0.54, 0.60)
Age 80 versus 50	3.34 (2.84, 3.93)	3.21 (2.70, 3.81)	2.88 (2.46, 3.37)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.59 (1.44, 1.76)	1.97 (1.82, 2.12)	2.50 (2.24, 2.79)	0.34 (0.32, 0.36)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.33 (1.24, 1.43)	NA	NA	1.20 (1.13, 1.27)	1.25 (1.17, 1.34)	0.86 (0.81, 0.91)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.92 (1.77, 2.08)	NA	1.25 (1.17, 1.35)	1.62 (1.51, 1.73)	1.54 (1.40, 1.68)	0.72 (0.65, 0.79)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.42 (1.31, 1.55)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.39 (1.29, 1.50)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.12 (1.04, 1.20)	NA	NA	1.12 (1.06, 1.18)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	2.85 (2.35, 3.45)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.29 (0.24, 0.34)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	1.78 (1.55, 2.05)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.56 (0.48, 0.66)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	1.74 (1.37, 2.21)	NA
Female versus male (at BSA = 1.8)	1.23 (1.10, 1.36)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.29 (1.21, 1.38)	0.98 (0.72, 1.33)	0.86 (0.81, 0.93)	1.03 (0.98, 1.08)	1.25 (1.16, 1.35)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.27 (1.21, 1.33)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.29 (1.24, 1.34)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	1.83 (1.64, 2.05)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation, ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.49 (2.14, 2.90)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	3.77 (2.75, 5.16)	2.78 (1.85, 4.17)	3.10 (2.21, 4.35)	4.54 (3.54, 5.83)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	7.94 (5.40, 11.66)	2.11 (1.06, 4.19)	3.47 (2.19, 5.51)	3.50 (2.41, 5.08)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

Table 5. Continued

C. Odds ratios for mitral valve replacement

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.65 (1.53, 1.78)	1.48 (1.38, 1.59)	1.35 (1.26, 1.44)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.25 (1.19, 1.31)	1.33 (1.29, 1.39)	1.26 (1.21, 1.33)	0.71 (0.68, 0.74)
Age 70 versus 50	2.71 (2.33, 3.17)	2.19 (1.90, 2.52)	1.81 (1.60, 2.06)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.56 (1.42, 1.71)	1.78 (1.65, 1.92)	1.60 (1.45, 1.76)	0.50 (0.46, 0.55)
Age 80 versus 50	5.14 (4.15, 6.37)	3.21 (2.70, 3.81)	2.67 (2.23, 3.20)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.97 (1.72, 2.26)	2.54 (2.27, 2.84)	2.27 (2.00, 2.58)	0.28 (0.25, 0.32)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.19 (1.07, 1.32)	NA	NA	1.11 (1.01, 1.21)	1.25 (1.17, 1.34)	0.96 (0.87, 1.06)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.72 (1.55, 1.91)	NA	1.25 (1.17, 1.35)	1.49 (1.36, 1.64)	1.54 (1.40, 1.68)	0.80 (0.71, 0.91)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.66 (1.47, 1.86)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.67 (1.52, 1.83)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.30 (1.16, 1.45)	NA	NA	1.34 (1.22, 1.47)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	4.59 (3.65, 5.77)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.23 (0.16, 0.33)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	2.21 (1.90, 2.56)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.63 (0.51, 0.77)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	1.05 (0.85, 1.31)	NA
Female versus male (at BSA=1.8)	1.11 (0.97, 1.27)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.06 (0.98, 1.16)	0.98 (0.72, 1.33)	0.79 (0.72, 0.87)	1.03 (0.98, 1.08)	1.09 (0.99, 1.19)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.08 (1.01, 1.16)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.16 (1.11, 1.22)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	1.50 (1.34, 1.67)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation, ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.03 (1.76, 2.35)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	2.74 (1.99, 3.78)	2.78 (1.85, 4.17)	2.20 (1.59, 3.05)	3.19 (2.41, 4.23)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	5.78 (3.77, 8.85)	2.11 (1.06, 4.19)	2.46 (1.56, 3.88)	2.46 (1.66, 3.65)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

Table 5. Continued

D. Odds ratios for mitral valve repair

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.80 (1.62, 2.00)	1.48 (1.38, 1.59)	1.55 (1.41, 1.71)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.20 (1.13, 1.27)	1.31 (1.26, 1.37)	1.50 (1.41, 1.60)	0.62 (0.60, 0.65)
Age 70 versus 50	3.24 (2.63, 4.00)	2.19 (1.90, 2.52)	2.42 (2.00, 2.92)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.44 (1.29, 1.62)	1.73 (1.58, 1.89)	2.25 (1.98, 2.55)	0.39 (0.36, 0.42)
Age 80 versus 50	6.72 (5.00, 9.04)	3.21 (2.70, 3.81)	4.11 (3.14, 5.38)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.75 (1.48, 2.07)	2.42 (2.12, 2.76)	3.78 (3.17, 4.51)	0.19 (0.17, 0.22)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.16 (0.99, 1.35)	NA	NA	1.11 (0.99, 1.24)	1.25 (1.17, 1.34)	0.92 (0.80, 1.05)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.67 (1.43, 1.95)	NA	1.25 (1.17, 1.35)	1.50 (1.33, 1.68)	1.54 (1.40, 1.68)	0.76 (0.65, 0.90)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.68 (1.42, 1.97)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.57 (1.36, 1.81)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.31 (1.11, 1.55)	NA	NA	1.26 (1.10, 1.45)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	6.24 (4.19, 9.30)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.26 (0.19, 0.37)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	2.90 (2.28, 3.70)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.49 (0.38, 0.64)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	2.50 (1.51, 4.12)	NA
Female versus male (at BSA = 1.8)	0.97 (0.77, 1.21)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.23 (1.10, 1.38)	0.98 (0.72, 1.33)	0.90 (0.80, 1.02)	1.03 (0.98, 1.08)	1.28 (1.12, 1.47)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.23 (1.09, 1.39)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.26 (1.15, 1.40)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	2.06 (1.73, 2.45)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.80 (2.32, 3.37)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	8.73 (4.84, 15.74)	2.78 (1.85, 4.17)	3.03 (1.69, 5.43)	6.12 (3.96, 9.46)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	18.39 (9.68, 34.96)	2.11 (1.06, 4.19)	3.39 (1.76, 6.54)	4.72 (2.71, 8.23)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

[a] Variable interacts with age. Reported odds ratio represents effect of risk factor for patients aged 50 years old.

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; NA = not applicable; NYHA = New York Heart Association; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

(≤ 40, > 40). Calibration plots (observed versus expected) based on the overall validation sample are presented in Figure 1. The average absolute difference between observed versus predicted event rates within deciles of predicted risk ranged from 0.06% for deep sternal wound infection to 1.06% for prolonged postoperative stay. Analogous figures were produced for specific valve procedures and numerous subgroups, and these are available at www.sts.org/riskmodels.

Model fit appeared to be adequate for each endpoint with the possible exception of deep sternal wound infection, which revealed some overfitting within certain subgroups. A modest degree of overfitting was expected for this endpoint given the relatively small number of infections and large number of candidate predictors.

Discrimination was assessed by the c-statistic, also known as the area under the receiver operating characteristic (ROC) curve. Table 4 presents the discrimination of each model in the development and validation samples for all patients combined and for subgroups consisting of AVR, MVR, and MVRepair. In the validation sample, c-statistics for the operative mortality model were 0.799 (overall), 0.759 (AVR), 0.802 (MVR), and 0.844 (MVRepair). C-statistics in the validation sample for other endpoints ranged from 0.619 for reoperation in the AVR subgroup to 0.800 for prolonged length of stay in the MVRepair subgroup.

Final Models

After validating the models in the 40% validation sample, the development and validation samples were then combined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regressions were estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [10]. An independence working correlation matrix was used to apply the generalized estimating equations methodology. With this approach, the estimated regression coefficients were identical to those obtained using ordinary logistic regression, but the standard errors were adjusted to account for the clustered data structure.

Odds Ratios

Odds ratios and 95% confidence intervals (CI) for the final selected models are presented in Table 5. "Not applicable" indicates that the specific predictor was not included in a particular risk model. Because several variables interact with surgery type, the odds ratios for these variables differ depending on the type of surgery (AVR, MVR, MVRepair). For example, in the operative mortality model, the odds ratio for emergent status is 3.77 (95% CI: 2.75, 5.16) for AVR, 2.74 (95% CI: 1.99, 3.78) for MVR, and 8.73 (95% CI: 4.84, 15.74) for MVRepair. Odds ratios that do not interact with surgery type are summarized in Table 5, Part A. Odds ratios that differ by surgery type for at least one endpoint are presented in Table 5, Parts B, C, and D.

Final Model Intercept and Coefficients

The final risk prediction algorithms, including all coefficients and intercepts, are presented in the Appendix.

Limitations

The limitations for these valve models are similar to those for the CABG models and are thoroughly discussed in Part 1 of this series (2008 STS CABG risk models).

Conclusion

The STS Quality Measurement Task Force has developed and tested nine new risk-adjustment models for isolated valve surgery using the STS NCD. This report includes a detailed exposition of the model development process, including not only statistical issues but also the many clinical and pragmatic judgments that were required. An online risk calculator is also available through a link from the STS website.

References

1. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
2. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
3. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
4. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
5. Kuduvalli M, Grayson AD, Au J, Grotte G, Bridgewater B, Fabri BM. A multi-centre additive and logistic risk model for in-hospital mortality following aortic valve replacement. Eur J Cardiothorac Surg 2007;31:607–13.
6. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
7. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
8. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2002.
9. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.
10. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Models

For each endpoint, the formula for calculating a patient's predicted risk of the endpoint has the form:

$$\text{Predicted Risk} = \frac{e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}{1+e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}$$

where $x_1, x_2, \ldots, x_n$ denote patient preoperative risk factors (eg, quantitative variables such as age, and comorbidities coded as 1 = present, 0 = absent); and $\beta_0, \beta_1, \ldots, \beta_n$ denote regression coefficients (numerical constants). Regression coefficients for each endpoint are presented in Appendix Table 1. The variables $x_1, x_2, \ldots, x_n$ are the same for each endpoint and are defined in Appendix Table 2. The regression coefficient for the time trend is not presented. Instead, the intercept has been adjusted to incorporate the time trend. This adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−5.78680	−5.83957	−5.52789	−3.96796	−7.11095	−3.08816	−3.06527	−4.30676	1.25115
Atrial fibrillation	0.18074	0.05524	0.00000	0.16527	0.00000	0.10305	0.11403	0.15530	−0.30247
Age function 1	0.03557	0.03909	0.03219	0.02683	0.04180	0.01512	0.02041	0.02670	−0.02834
Age function 3	0.02804	−0.00132	0.01809	0.00629	−0.05024	0.00218	0.01282	0.02315	−0.04637
Age by reoperation function	−0.01308	−0.02043	−0.00551	−0.00840	−0.00939	−0.00697	−0.00684	−0.00485	0.00927
Age by status function	−0.02495	−0.02987	−0.00721	−0.01377	0.00277	0.00102	−0.00677	−0.00379	−0.00795
Age by MVR function	0.01436	0.00000	−0.00245	0.00000	0.00000	0.00715	0.00848	−0.00324	−0.00603
Age by MVRepair function	0.02326	0.00000	0.01190	0.00000	0.00000	0.00315	0.00685	0.01378	−0.01883
BSA function 1	−1.40168	−0.38619	−0.71012	−1.11750	0.14188	−0.73553	−0.91858	−0.82801	0.77317
BSA function 2	2.16782	0.23148	1.92875	2.29127	2.04603	0.83644	1.65638	1.65423	−1.76728
CHF but not NYHA IV	0.25590	0.00000	0.21233	0.28353	0.00000	0.00000	0.17974	0.22508	−0.15108
CHF and NYHA IV	0.60544	0.00000	0.47812	0.65056	0.00000	0.22686	0.48025	0.42957	−0.33521
CHF by MVR function	0.00000	0.00000	0.00000	−0.11007	0.00000	0.00000	−0.07864	0.00000	0.11503
CHF by MVRepair function	0.00000	0.00000	0.00000	−0.13792	0.00000	0.00000	−0.07731	0.00000	0.06468
CLD function	0.23846	0.00000	0.16629	0.22816	0.23817	0.08406	0.16044	0.25263	−0.21022
CLD by MVR function	−0.15906	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.10092	0.00000
CLD by MVRepair function	−0.03243	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.01795	0.00000
Creatinine function 1	0.43909	0.29230	0.71439	0.45646	0.00000	0.23562	0.49230	0.45631	−0.44178
CVD without prior CVA	0.00000	0.27837	0.20531	0.12726	0.00000	0.05830	0.07684	0.00000	−0.22223
CVD and prior CVA	0.00000	0.59220	0.20018	0.24512	0.00000	0.13200	0.18343	0.33480	−0.25595
Diabetes, noninsulin	0.23563	0.00000	0.37172	0.11040	0.00000	0.00000	0.11355	0.19843	−0.16630
Diabetes, insulin	0.48368	0.00000	0.64648	0.35367	0.44389	0.18293	0.33165	0.51913	−0.45093
Diabetes by MVR function	0.00000	0.00000	0.00000	0.15051	0.00000	0.00000	0.17990	0.00000	0.00000
Diabetes by MVRepair function	0.00000	0.00000	0.00000	0.16260	0.00000	0.00000	0.11734	0.00000	0.00000
Dialysis	1.48666	0.79199	0.00000	1.57690	1.19109	0.81972	1.37741	1.53351	−1.69019
Dialysis by MVR function	0.47550	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.20998
Dialysis by MVRepair function	0.78385	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.07964
Ejection fraction function	0.00904	0.00000	0.00407	0.01107	0.02308	0.00734	0.00925	0.01111	−0.01348
Endocarditis, active	0.66737	0.62434	0.77276	0.76318	0.00000	0.43876	0.67810	1.02521	−1.08299
Female	0.20372	0.21925	−0.03031	0.25668	−0.02355	−0.14567	0.03066	0.22437	−0.36400
Female by MVR function	−0.10089	0.00000	0.00000	−0.19465	0.00000	−0.08773	0.00000	−0.14211	0.00000
Female by MVRepair function	−0.23812	0.00000	0.00000	−0.04564	0.00000	0.04424	0.00000	0.02470	0.00000
Female by BSA function 1	0.96491	−0.02257	0.83074	0.77598	2.00214	0.16707	0.52716	0.57195	−0.75434
Female by BSA function 2	0.18084	−0.07419	0.08397	−0.58460	−1.87036	0.25158	−0.09063	−0.12289	0.35123
Hypertension	0.11372	0.17789	0.29770	0.10799	0.00000	0.00000	0.10361	0.00000	−0.06504
IABP or inotropes	0.38682	0.00000	0.29606	0.57608	0.52474	0.13432	0.56046	0.37621	−0.57115
IABP by MVR function	0.00000	0.00000	0.00000	0.21517	0.00000	0.00000	0.00000	0.00000	0.10760
IABP by MVRepair function	0.00000	0.00000	0.00000	0.48870	0.00000	0.00000	0.00000	0.00000	−0.13850

Appendix Table 1. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Immunosuppressive treatment	0.35022	0.00000	0.32828	0.00000	0.00000	0.00000	0.14887	0.27152	0.00000
Insufficiency mitral	0.00000	0.23253	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Insufficiency tricuspid	0.00000	0.00000	0.13159	0.12973	0.00000	0.08969	0.18929	0.15846	−0.20027
Left main disease	0.17593	0.00000	0.17280	0.00000	0.77557	0.00000	0.00000	0.00000	0.00000
MI ≤ 21 days	0.13276	0.00000	0.00000	0.31706	0.00000	0.03495	0.24687	0.18812	−0.20961
MVR	0.10284	0.00000	0.40455	0.44639	0.00000	0.12852	0.13795	0.58004	−0.61402
MVRepair	−0.65440	0.00000	−0.23666	−0.19726	0.00000	−0.22398	−0.23002	−0.37618	0.25710
No. diseased vessel function	0.00000	0.09556	0.00000	0.06299	0.00000	0.00000	0.03700	0.03312	−0.10126
Peripheral vascular disease	0.21980	0.25236	0.00000	0.00000	0.00000	0.19758	0.13174	0.15342	−0.18903
Race black	0.00000	0.28378	0.40941	0.34795	0.00000	0.23856	0.31567	0.37161	−0.44177
Race Hispanic	0.00000	−0.13774	0.14968	0.06720	0.00000	0.12816	0.08581	0.15128	−0.20068
Reop, 1 previous operation	0.74484	0.73489	0.43804	0.60704	0.00000	0.27365	0.44052	0.35252	−0.40042
Reop, ≥ 2 previous operations	0.90625	0.85841	0.50595	0.91229	0.00000	0.34233	0.67201	0.56294	−0.69765
Reop by MVR function	0.00000	0.00000	0.00000	−0.20333	0.00000	0.00000	0.00000	0.00000	0.00000
Reop by MVRepair function	0.00000	0.00000	0.00000	0.11559	0.00000	0.00000	0.00000	0.00000	0.00000
Shock	0.47961	0.50213	0.00000	0.73670	0.00000	0.28068	0.74786	0.55376	0.00000
Shock by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.50071	0.00000
Shock by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.36096	0.00000
Status urgent	0.25552	0.00000	0.19344	0.25714	0.00000	0.15548	0.19858	0.35184	−0.36106
Status emergent	1.32597	1.02109	1.13199	1.51294	0.00000	0.49075	1.17360	0.89480	−1.12373
Status salvage	2.07144	0.74530	1.24544	1.25342	0.00000	0.00000	1.21823	0.00000	−1.13785
Status by MVR function	−0.31729	0.00000	−0.34380	−0.35206	0.00000	0.00000	0.00000	0.00000	0.00000
Status by MVRepair function	0.84051	0.00000	−0.02373	0.29927	0.00000	0.00000	0.00000	0.00000	0.00000
Stenosis aortic	0.00000	0.00000	0.00000	−0.10782	0.00000	−0.10852	−0.07479	−0.15434	0.06873
Stenosis mitral	0.21309	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Unstable angina	0.18950	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Appendix Table 2. Definition of Variables Appearing in STS 2008 Valve Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
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
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 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 ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise

Appendix Table 2. Continued

Variable	Definition
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 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
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Stenosis mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina, no MI within 7 days of surgery, = 0 otherwise

Note: See www.sts.org for exact definitions of terms used above.

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; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PLOS = prolonged length of stay; Preop = preoperative; Reop = reoperation; Comp = composite adverse event (any); RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

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 #:** 0121 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Operative Mortality for Mitral Valve (MV) Replacement
De.2 **Brief description of measure***:* Percent of patients undergoing MV Replacement who die, including both 1) all deaths occurring during the hospitalization in which the procedure was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure
1.1-2 **Type of Measure:** Outcome 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-634267315118486062.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:** Public reporting, Internal quality improvement	**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, Frequently performed procedure, Leading cause of morbidity/mortality, High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Review of morbidity and mortality results of mitral valve replacement has impact on every center that tracks and reports their performance; the process of publishing and comparing institutional results provides benchmark feedback and triggers institutional programs to further improve their mitral valve procedural outcomes. 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Petyerson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892. - Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137. - Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467.	**1a** C☐ P☐ M☐ N☐

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: The benefit of reporting mitral valve replacement data is to expand the scope of procedures monitored by the STS and local institutions. Whether it takes 3 years or 6 months to reach the statistically significant reporting threshold of 100 cases, the ongoing surveillance tests the hypothesis that valve replacement outcomes parallel CAB performance. If not, quality improvement projects targeting mitral valve disease should be elaborated.

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, 2005-December 31, 2009

Analysis includes 106 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data (not restricted to this measure) to STS for at least 36 months in 2005-2009.

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*)**:** Reporting an individual center's results in quartile relationship to all participating centers gives a program a reference point relative to peers and a sense of how much improvement is necessary to reach a level of performance expected by the center's cardiovascular leadership.

1c.2-3. Type of Evidence: Observational study, Expert opinion, Systematic synthesis of research, Meta-analysis, Other 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*)**:**
All levels and processes of surgical care from pre-op evaluation and preparation to discharge planning influences the procedural outcome. Identifying the processes of care which most influence the outcome requires phase of care analysis so that quality improvement efforts are well-targeted.

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 NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437.
- Edwards FH, Petyerson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892.
- Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137.
- Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467.

1c
C☐
P☐
M☐
N☐

1c.9 Quote the Specific guideline recommendation (*including guideline number and/or page number*): N/A **1c.10 Clinical Practice Guideline Citation:** **1c.11 National Guideline Clearinghouse or other URL:** **1c.12 Rating of strength of recommendation (*also provide narrative description of the rating and by whom*):** **1c.13 Method for rating strength of recommendation (*If different from USPSTF system, also describe rating and how it relates to USPSTF*):** **1c.14 Rationale for using this guideline over others:**	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
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*): Number of patients undergoing MV Replacement who die, including both 1) all deaths occurring during the hospitalization in which the operation was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure **2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*):** During hospitalization regardless of length of stay or within 30 days of surgery if discharged **2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):** Number of isolated MV Replacement procedures with an operative mortality; Number of isolated MV Replacement procedures in which Mortality [Mortalty (STS Adult Cardiac Surgery Database Version 2.73)] and Mortality Operative Death (MtOpD) are marked “yes.” Operative mortality is further verified by the following variables: Mortality Status at 30 days (Mt30Stat), Mortality Date (MtDate), Mortality Discharge Status (MtDCStat)	
2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*): All patients undergoing isolated MV Replacement surgery **2a.5 Target population gender:** Female, Male **2a.6 Target population age range:** 18 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***):**
60 months

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
Number of isolated MV Replacement procedures;

Isolated MV Replacement is determined as a procedure for which all of the following apply:
- OpValve is marked "Yes"
- VSMV is marked "Yes"
- VSMVPr is marked "Replacement"
- (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
- OpCAB, VSAV, VSAVPr, ResectSubA, OpTricus, OpPulm, OpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc, OCTumor, OCPulThromDis, OCarOthr are all marked "no" or "missing"

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***):**

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: Case-mix adjustment

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

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment 2a.15 Detailed Risk Model-634267316854669390.pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
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)*:
Participant specific OR and their 95% CI were estimated in the hierarchical model. These model-based estimates were used to control variation due to random statistical fluctuations while estimating true signal variation. A 95% CI excluding zero indicates the participant's performance is significantly lower or higher than an "average" STS participant.

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-December of 2010)--- 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-December of 2010

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

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)*: STS Adult Cardiac Surgery Database - Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009. Excluded from analysis are participants that did not submit results for both time periods. Because database participants can change their underlying care processes at any time, we would not expect perfect correlation between two sets of results from even proximate time periods.

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

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 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.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test*

2c
C☐ P☐ M☐ N☐

conducted): Mortality Operative Death: 100.0% agreement rate	
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)*: Please see Risk Adjustment Type section above **2e.2 Analytic Method** *(type of risk adjustment, analysis, & rationale)*: Detailed information regarding the risk adjustment model can be found in the attachment: O´Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S23-42. **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)*: 106 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009; January 1, 2005-December 31, 2009 **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: We calculated the risk adjusted event rate with the participant's Odds Ratio (OR) estimate and the overall STS event rate. Therefore, the risk adjusted rate is closely related to OR estimate. If OR > 1, then the participant's risk adjusted rate will be greater than the overall STS event rate; if OR < 1, then the participant's risk adjusted rate will be smaller than the overall STS event rate. The statistical significance is defined by the 95% confidence interval (CI) or the OR estimate. If the 95% CI for a participant's OR includes the null value 1.0, then we cannot distinguish this participant's performance from the STS average – either the participant's performance was close to average or else the participant's sample size was too small to make a reliable inference. Otherwise, if the 95% CI falls to the right of 1.0, then the participant's performance is considered significantly lower than the average STS results; if the 95% CI falls to the left of 1.0, then the participant's performance is considered significantly higher than the average STS results. **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	**2f** C☐ P☐ M☐ N☐
2g. Comparability of Multiple Data Sources/Methods	**2g**

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)*:	**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 STS plans to publicly report more measures in the future. There is no definite date yet assigned to this measure; however, STS staff and surgeon leadership have engaged in initial internal STS discussions regarding this matter. **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***):** **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** **C☐ P☐ M☐ N☐**

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. Please see IFMC audit results.	**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** C☐ P☐ M☐ N☐

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 surgeon 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:**	
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 0121 Sections 2a.14, 1b.2, 2b.3, 2f.3, 3a.6.pdf
Date of Submission (*MM/DD/YY*)**:** 01/12/2011

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 valve procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest four and a half years of data (January 2006 – 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 Isolated Valve Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
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
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 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 ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 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
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Stenosis mitral	= 1 if patient has mitral 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; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention;

The final patient-level model was built by step-wise selection method with several variables decided by surgeon panel 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	*Risk-Adjusted Operative Mortality for Mitral Valve Replacement Surgery*
N	106
Mean	1.0
1st	2.1
5th	1.7
10th	1.4
25th	1.1
Median	0.9
75th	0.7
90th	0.6
95th	0.5
99th	0.4
Outlier	9 (8.5)
High	6
Low	3

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)

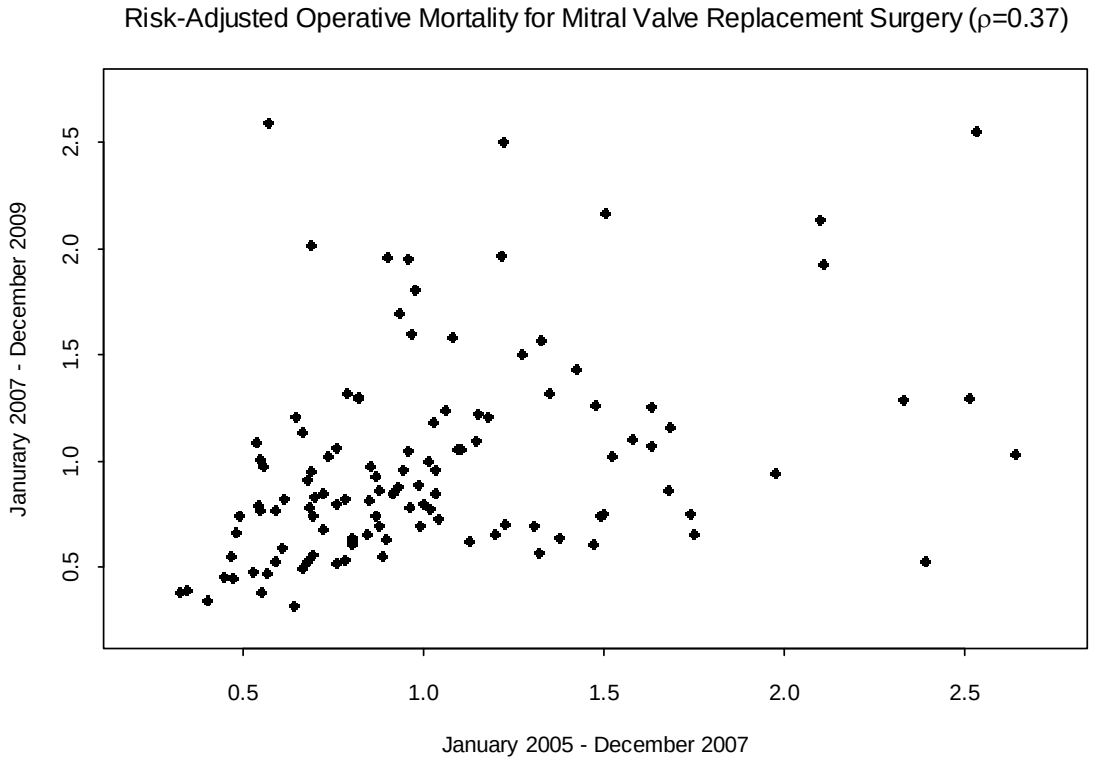
In the above equation, "Participant's Expected Event Rate" is calculated with the participant's actual OR, and "Participant's Expected Event Rate Assuming Its Performance = STS Average Performance" is calculated by assuming the participant's OR = 1 (i.e. no difference in performance from the STS average).

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Replacement Surgery*
N	106
Mean	5.4
1st	2.3
5th	3.1

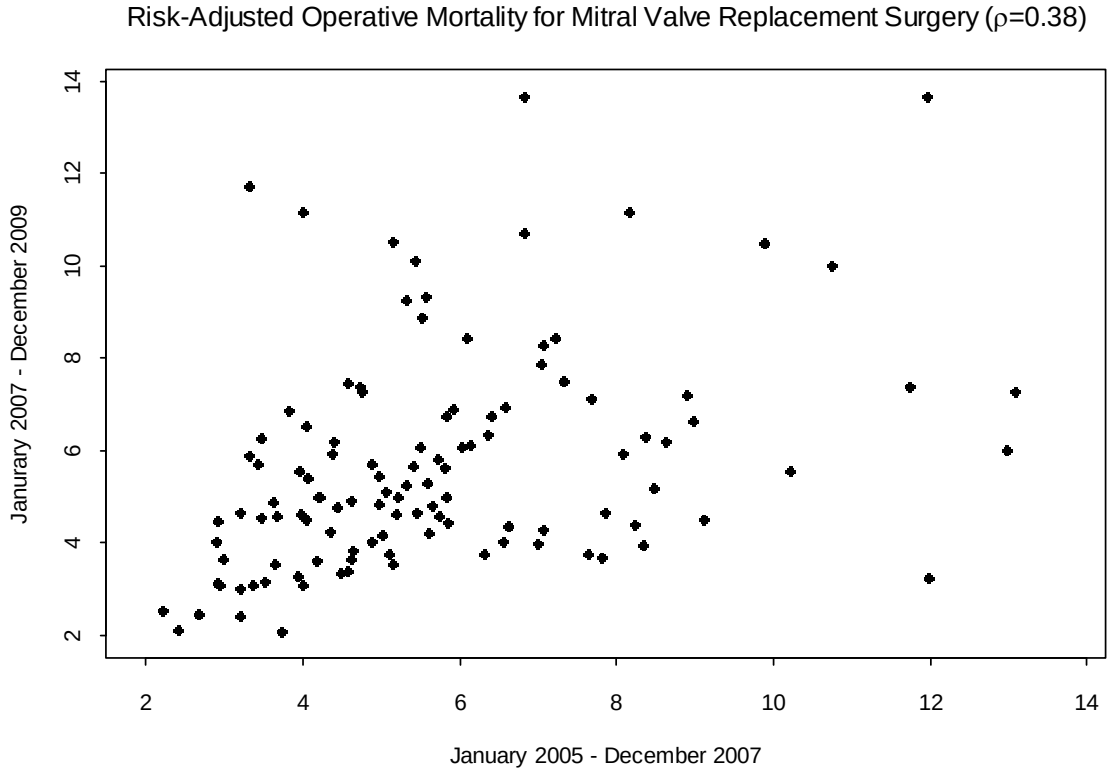
Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Replacement Surgery*
10th	3.5
25th	4.2
Median	5.1
75th	6.3
90th	7.8
95th	8.5
99th	11.1
Outlier	9 (8.5)
High	6
Low	3

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

Testing results: $\rho = 0.37$



Risk Adjusted Rate:
Testing results: $\rho = 0.38$



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, 2005-December 31, 2009. Sample contains 106 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009.

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Replacement Surgery*
N	106
Mean	1.0
1st	2.1
5th	1.7
10th	1.4
25th	1.1
Median	0.9
75th	0.7
90th	0.6
95th	0.5
99th	0.4
Outlier†	9 (8.5)
High	6
Low	3

Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Replacement Surgery*
N	106
Mean	5.4
1st	2.3
5th	3.1
10th	3.5
25th	4.2
Median	5.1
75th	6.3
90th	7.8
95th	8.5
99th	11.1
Outlier†	9 (8.5)
High	6
Low	3

†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."

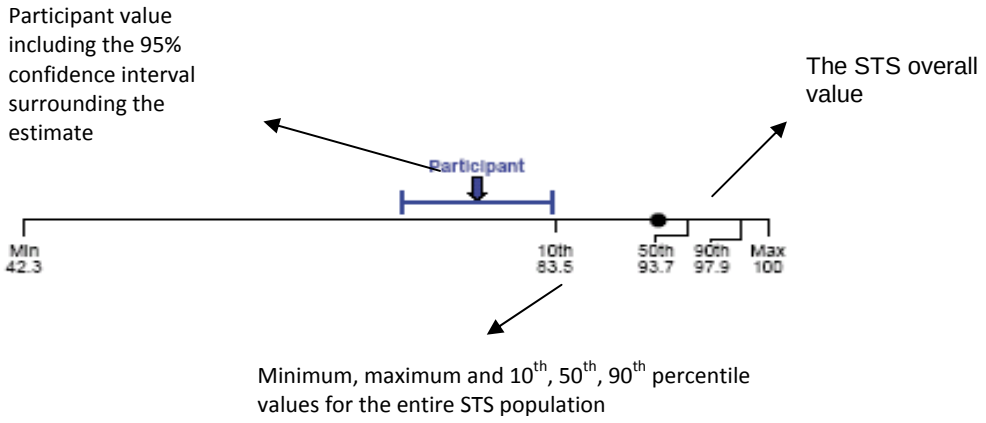
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

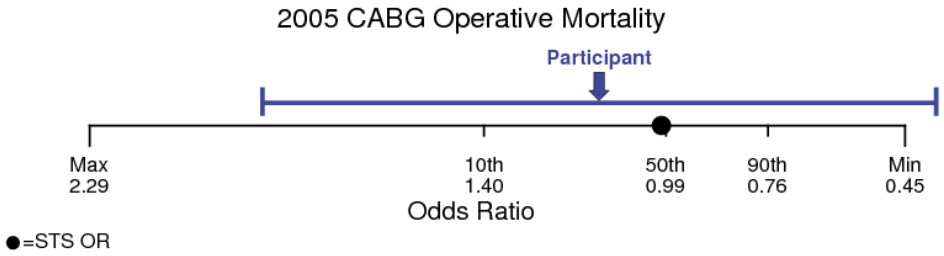
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

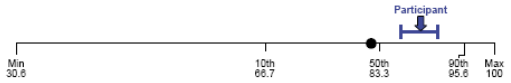
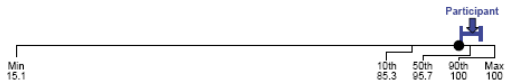
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 2—Isolated Valve Surgery

Sean M. O'Brien, PhD,[a] David M. Shahian, MD,[b] Giovanni Filardo, PhD, MPH,[c] Victor A. Ferraris, MD,[d] Constance K. Haan, MD,[e] Jeffrey B. Rich, MD,[f] Sharon-Lise T. Normand, PhD,[g] Elizabeth R. DeLong, PhD,[a] Cynthia M. Shewan, PhD,[h] Rachel S. Dokholyan, MPH,[a] Eric D. Peterson, MD, MPH,[a] Fred H. Edwards, MD,[e] and Richard P. Anderson, MD[i†]

[a]Duke Clinical Research Institute, Durham, North Carolina; [b]Massachusetts General Hospital, Boston, Massachusetts; [c]Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; [d]University of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; [e]University of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; [f]Sentara Cardiovascular Research Institute, Norfolk, Virginia; [g]Department of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [h]The Society of Thoracic Surgeons, Chicago, Illinois; and [i]Seattle, Washington

Background. **Adjustment for case-mix is essential when using observational data to compare surgical techniques or providers. That is most often accomplished through the use of risk models that account for preoperative patient factors that may impact outcomes. The Society of Thoracic Surgeons (STS) uses such risk models to create risk-adjusted performance reports for participants in the STS National Adult Cardiac Surgery Database (NCD). Although risk models were initially developed for coronary artery bypass surgery, similar models have now been developed for use with heart valve surgery, particularly as the proportion of such procedures has increased. The last published STS model for isolated valve surgery was based on data from 1994 to 1997 and did not include patients undergoing mitral valve repair. STS has developed new valve surgery models using contemporary data that include both valve repair as well as replacement. Expanding upon existing valve models, the new STS models include several nonfatal complications in addition to mortality.**

Methods. **Using STS data from 2002 to 2006, isolated valve surgery risk models were developed for operative mortality, permanent stroke, renal failure, prolonged ventilation (> 24 hours), deep sternal wound infection, reoperation for any reason, a major morbidity or mortality composite endpoint, prolonged postoperative length of stay, and short postoperative length of stay. The study population consisted of adult patients who underwent one of three types of valve surgery: isolated aortic valve replacement (n = 67,292), isolated mitral valve replacement (n = 21,229), or isolated mitral valve repair (n = 21,238). The population was divided into a 60% development sample and a 40% validation sample. After an initial empirical investigation, the three surgery groups were combined into a single logistic regression model with numerous interactions to allow the covariate effects to differ across these groups. Variables were selected based on a combination of automated stepwise selection and expert panel review.**

Results. **Unadjusted operative mortality (in-hospital regardless of timing, and 30-day regardless of venue) for all isolated valve procedures was 3.4%, and unadjusted in-hospital morbidity rates ranged from 0.3% for deep sternal wound infection to 11.8% for prolonged ventilation. The number of predictors in each model ranged from 10 covariates in the sternal infection model to 24 covariates in the composite mortality plus morbidity model. Discrimination as measured by the c-index ranged from 0.639 for reoperation to 0.799 for mortality. When patients in the validation sample were grouped into 10 categories based on deciles of predicted risk, the average absolute difference between observed versus predicted events within these groups ranged from 0.06% for deep sternal wound infection to 1.06% for prolonged postoperative stay.**

Conclusions. **The new STS risk models for valve surgery include mitral valve repair as well as multiple endpoints other than mortality. Model coefficients are provided and an online risk calculator is publicly available from The Society of Thoracic Surgeons website.**

(Ann Thorac Surg 2009;88:S23–42)

Models for predicting surgical outcomes on the basis of patient preoperative characteristics are valuable tools for research, quality improvement, and clinical practice. Such models are used by The Society of Thoracic Surgeons (STS) to produce risk-adjusted performance re-

†This author is deceased. Former Chair, Quality, Research and Patient Safety Council, The Society of Thoracic Surgeons, Chicago, IL.

Address correspondence to Dr Shahian, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

Drs O'Brien, Shahian, Filardo, Ferraris, Haan, Rich, Normand, DeLong, Shewan, Peterson, Edwards, Anderson, and Ms Dokholyan, have no conflicts of interest to declare regarding this work.

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.056

Abbreviations and Acronyms

AVR	= aortic valve replacement
CABG	= coronary artery bypass graft surgery
CI	= confidence interval
MI	= myocardial infarction
MVR	= mitral valve replacement
MVRepair	= mitral valve repair
NCD	= National Adult Cardiac Surgery Database
QMTF	= Quality Measurement Task Force
STS	= The Society for Thoracic Surgeons

ports for providers participating in the STS National Adult Cardiac Surgery Database (NCD). They are also used by STS surgeons and other physicians for counseling patients about the risk of surgery.

The earliest STS risk models were developed nearly 2 decades ago for isolated coronary artery bypass graft surgery (CABG). Subsequently, similar models have been developed for isolated valve replacement and combined CABG plus valve replacement. Because surgical practice and outcomes are changing rapidly, these models are updated periodically to reflect contemporary experience.

The last published STS model for isolated valve surgery was based on STS data from 1994 to 1997. The reference population included aortic and mitral valve replacements but excluded mitral valve repair, and the endpoint was operative mortality. In the decade since this model was published, many aspects of heart surgery have changed. First, as CABG volumes have decreased with the introduction of coronary stents, valve surgery as a proportion of overall heart surgery volume has increased in most practices. Between 2000 and 2006, the percentage of isolated CABG procedures decreased from 73% to 60% and the percentage of isolated valve procedures increased from 18% to 22%. Thus, in assessing provider performance, it is no longer sufficient only to consider isolated CABG surgery. Second, the frequency of mitral repair as a percentage of all isolated mitral operations in the STS NCD increased from 35% in 2000 to 53% in 2006. Third, during the same time period, the average mortality rate for isolated aortic or mitral surgery also decreased. Finally, efforts to measure and compare surgical performance have intensified and expanded. In addition to measuring operative mortality, performance reports increasingly focus on nonfatal complications as well as resource utilization and efficiency. Such outcomes have not historically been risk-adjusted for valve surgery.

The STS Quality Measurement Task Force (QMTF) has undertaken a complete revision of all STS risk models for adult cardiac surgery, and these new models were implemented in January 2008. This report, Part 2 of 3, describes the new STS models for isolated valve surgery (Part 1 describes the STS isolated CABG models, and Part 3 describes the models for CABG plus valve surgery). Authors of this report are the QMTF members who were involved in this initiative.

Two important features have been incorporated into these new models. First, the population includes mitral valve repair as well as aortic and mitral valve replacement. Second, in addition to operative mortality, the new models include six nonfatal in-hospital morbidity endpoints and two length-of-stay endpoints. In comparison with several other valve models that have recently been published [1–6], the STS models are distinguished by the large size of the development population and the broad spectrum of endpoints included.

Study Population and Endpoints

The population for this analysis consisted of operations on adult patients aged 20 to 100 years who underwent isolated single aortic or mitral valve surgery between January 1, 2002, and December 31, 2006. Only patients undergoing one of the following procedures were included: (1) isolated aortic valve replacement (AVR); (2) isolated mitral valve replacement (MVR); and (3) isolated mitral valve repair (MVRepair).

Because of the relatively small number of pulmonic, tricuspid, multiple valve procedures, and aortic repairs, these cases were not included in the current models. Patients undergoing concomitant CABG were excluded from the current analysis, but these were included in the separate STS valve plus CABG models described in Part 3 of this series. Records with missing data on sex (n = 44) were excluded because missing sex is not allowed in the analysis dataset used for creating STS database participant feedback reports. This left a final study population of 109,759 patient operations performed at 809 STS NCD participating groups. Patients on dialysis preoperatively (n = 2,699) were not included when developing the risk model for prediction of postoperative renal failure.

Patient characteristics in the study population are presented in Table 1.

Training and Validation Samples

The study population was randomly divided into a 60% training (development) sample and a 40% test (validation) sample. The development sample was used to identify predictor variables and estimate model coefficients. Data from the validation sample were used to assess model fit, discrimination, and calibration. After choosing variables and assessing model fit, the development and validation samples were subsequently combined, and the final model coefficients were estimated using the combined (development plus validation) data.

Endpoints

Risk models were developed for nine endpoints, identical to those in the STS CABG models. In contrast with the definition of operative mortality, which includes hospital deaths as well as deaths that occur after discharge within 30 days of surgery, the morbidity endpoints only include events that occurred before discharge. However, beginning with version 2.61, sternal infection data will be recorded for as long as 30 days postoperatively. The nine endpoints are as follows: (1) operative mortality: death during the same

Table 1. Distribution of Risk Factors in Overall Study Population Isolated Valve (2002–2006)

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Demographics								
Age, years								
< 55	28,147	25.6	13,227	19.66	6,601	31.09	8,319	39.17
55–64	23,258	21.2	12,987	19.30	4,833	22.77	5,438	25.61
65–74	28,145	25.6	18,299	27.19	5,294	24.94	4,552	21.43
≥75	30,209	27.5	22,779	33.85	4,501	21.20	2,929	13.79
Sex								
Male	60,752	55.4	39,209	58.27	9,055	42.65	12,488	58.80
Female	49,007	44.6	28,083	41.73	12,174	57.35	8,750	41.20
Race								
Caucasian	93,522	85.2	58,656	87.17	16,810	79.18	18,056	85.02
Black	7,630	7.0	3,555	5.28	2,383	11.23	1,692	7.97
Hispanic	3,680	3.4	2,344	3.48	889	4.19	447	2.10
Asian	1,538	1.4	719	1.07	437	2.06	382	1.80
Other	2,493	2.3	1,508	2.24	505	2.38	480	2.26
Missing	896	0.8	510	0.76	205	0.97	181	0.85
Risk factors								
Body surface area, m^2								
< 1.50	4,351	4.0	2,341	3.48	1,234	5.81	776	3.65
1.50–1.74	24,577	22.4	13,713	20.38	6,151	28.97	4,713	22.19
1.75–1.99	40,548	36.9	24,744	36.77	7,914	37.28	7,890	37.15
≥ 2.00	39,517	36.0	26,007	38.65	5,768	27.17	7,742	36.45
Missing	766	0.7	487	0.72	162	0.76	117	0.55
Body mass index, kg/m^2								
< 25	35,526	32.4	18,509	27.51	8,447	39.79	8,570	40.35
25–29	39,074	35.6	24,035	35.72	6,992	32.94	8,047	37.89
30–34	20,534	18.7	14,142	21.02	3,318	15.63	3,074	14.47
≥ 35	13,682	12.5	10,008	14.87	2,280	10.74	1,394	6.56
Missing	943	0.9	598	0.89	192	0.90	153	0.72
Diabetes mellitus								
No diabetes	88,709	80.8	52,052	77.35	17,535	82.60	19,122	90.04
Diabetes, noninsulin	14,900	13.6	11,026	16.39	2,412	11.36	1,462	6.88
Diabetes, insulin	5,788	5.3	3,974	5.91	1,216	5.73	598	2.82
Diabetes missing	138	0.1	91	0.14	34	0.16	13	0.06
Treatment missing	224	0.2	149	0.22	32	0.15	43	0.20
Hypertension								
No	41,649	37.9	22,338	33.20	8,859	41.73	10,452	49.21
Yes	67,886	61.9	44,816	66.60	12,326	58.06	10,744	50.59
Missing	224	0.2	138	0.21	44	0.21	42	0.20
Hypercholesterolemia								
No	59,003	53.8	33,156	49.27	12,857	60.56	12,990	61.16
Yes	50,328	45.9	33,865	50.33	8,286	39.03	8,177	38.50
Missing	428	0.4	271	0.40	86	0.41	71	0.33
Past or present smoker								
No	57,609	52.5	33,953	50.46	11,075	52.17	12,581	59.24
Yes	51,910	47.3	33,191	49.32	10,109	47.62	8,610	40.54
Missing	240	0.2	148	0.22	45	0.21	47	0.22
Chronic lung disease								
None	87,826	80.0	53,503	79.51	16,125	75.96	18,198	85.69
Mild	11,184	10.2	6,991	10.39	2,520	11.87	1,673	7.88
Moderate	6,346	5.8	4,022	5.98	1,494	7.04	830	3.91
Severe	3,332	3.0	2,110	3.14	853	4.02	369	1.74
Missing	1,071	1.0	666	0.99	237	1.12	168	0.79

Table 1. Continued

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Peripheral vascular disease								
No	101,129	92.1	61,222	90.98	19,550	92.09	20,357	95.85
Yes	8,381	7.6	5,909	8.78	1,641	7.73	831	3.91
Missing	249	0.2	161	0.24	38	0.18	50	0.24
Cerebrovascular disease								
No	96,852	88.2	58,983	87.65	18,158	85.53	19,711	92.81
Yes	12,661	11.5	8,147	12.11	3,033	14.29	1,481	6.97
Missing	246	0.2	162	0.24	38	0.18	46	0.22
CVA								
No CVA	101,631	92.6	62,518	92.91	18,833	88.71	20,280	95.49
Remote CVA (> 2 weeks)	6,926	6.3	4,203	6.25	1,912	9.01	811	3.82
Recent CVA (≤ 2 weeks)	818	0.7	325	0.48	409	1.93	84	0.40
CVA–missing timing	100	0.1	60	0.09	29	0.14	11	0.05
Missing	284	0.3	186	0.28	46	0.22	52	0.24
Endocarditis								
No endocarditis	100,998	92.0	63,257	94.00	17,926	84.44	19,815	93.30
Treated endocarditis	4,197	3.8	1,761	2.62	1,445	6.81	991	4.67
Active endocarditis	4,238	3.9	2,068	3.07	1,791	8.44	379	1.78
Endocarditis–missing type	63	0.1	30	0.04	27	0.13	6	0.03
Missing	263	0.2	176	0.26	40	0.19	47	0.22
Renal failure								
No	102,205	93.1	62,873	93.43	19,016	89.58	20,316	95.66
Yes	7,305	6.7	4,251	6.32	2,173	10.24	881	4.15
Missing	249	0.2	168	0.25	40	0.19	41	0.19
Renal function								
Creatinine < 1.00 mg/dL	42,028	38.3	25,679	38.16	7,754	36.53	8,595	40.47
Creatinine 1–1.49 mg/dL	51,939	47.3	32,058	47.64	9,372	44.15	10,509	49.48
Creatinine 1.50–1.99 mg/dL	8,081	7.4	5,078	7.55	1,875	8.83	1,128	5.31
Creatinine 2.00–2.49 mg/dL	1,946	1.8	1,192	1.77	512	2.41	242	1.14
Creatinine ≥ 2.50 mg/dL	1,294	1.2	750	1.11	390	1.84	154	0.73
Dialysis	2,699	2.5	1,464	2.18	900	4.24	335	1.58
Missing	1,772	1.6	1,071	1.59	426	2.01	275	1.29
Immunosuppressive treatment								
No	106,037	96.6	64,953	96.52	20,356	95.89	20,728	97.60
Yes	3,336	3.0	2,074	3.08	819	3.86	443	2.09
Missing	386	0.4	265	0.39	54	0.25	67	0.32
Previous CV interventions								
Previous coronary artery bypass surgery								
No	98,978	90.2	60,351	89.69	18,564	87.45	20,063	94.47
Yes	10,399	9.5	6,713	9.98	2,569	12.10	1,117	5.26
Missing	382	0.3	228	0.34	96	0.45	58	0.27
Previous valve surgery								
No	100,179	91.3	62,898	93.47	16,857	79.41	20,424	96.17
Yes	9,227	8.4	4,186	6.22	4,285	20.18	756	3.56
Missing	353	0.3	208	0.31	87	0.41	58	0.27
Previous other cardiac surgery								
No	105,686	96.3	65,084	96.72	20,034	94.37	20,568	96.85
Yes	3,662	3.3	1,975	2.93	1,077	5.07	610	2.87
Missing	411	0.4	233	0.35	118	0.56	60	0.28
Number of previous CV surgeries								
No prior CV surgery	91,196	83.1	56,629	84.15	15,239	71.78	19,328	91.01
1 prior CV surgery	15,399	14.0	9,122	13.56	4,775	22.49	1,502	7.07
2 or more prior CV surgeries	2,653	2.4	1,260	1.87	1,069	5.04	324	1.53
Missing	511	0.5	281	0.42	146	0.69	84	0.40

Table 1. Continued

Variable	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
	N	%	N	%	N	%	N	%
Prior PCI								
No PCI	101,878	92.8	62,145	92.35	19,573	92.20	20,160	94.92
PCI within 6 hours	122	0.1	58	0.09	51	0.24	13	0.06
PCI not within 6 hours	7,100	6.5	4,678	6.95	1,447	6.82	975	4.59
PCI–missing timing	133	0.1	90	0.13	28	0.13	15	0.07
Missing	526	0.5	321	0.48	130	0.61	75	0.35
Preoperative cardiac status								
Acuity status								
Elective	84,052	76.6	51,734	76.88	14,293	67.33	18,025	84.87
Urgent	23,795	21.7	14,670	21.80	6,071	28.60	3,054	14.38
Emergent	1,555	1.4	685	1.02	747	3.52	123	0.58
Emergent salvage	154	0.1	70	0.10	78	0.37	6	0.03
Missing	203	0.2	133	0.20	40	0.19	30	0.14
MI								
No prior MI	99,416	90.6	60,850	90.43	18,716	88.16	19,850	93.46
MI > 21 days	7,785	7.1	4,770	7.09	1,848	8.71	1,167	5.49
MI 8–21 days	719	0.7	480	0.71	170	0.80	69	0.32
MI 1–7 days	1,247	1.1	863	1.28	315	1.48	69	0.32
MI > 6 and < 24 hours	142	0.1	61	0.09	66	0.31	15	0.07
MI ≤ 6 hours	90	0.1	42	0.06	40	0.19	8	0.04
MI–missing timing	127	0.1	79	0.12	33	0.16	15	0.07
Missing	233	0.2	147	0.22	41	0.19	45	0.21
Angina								
No	85,364	77.8	49,573	73.67	17,598	82.90	18,193	85.66
Yes	24,164	22.0	17,577	26.12	3,591	16.92	2,996	14.11
Missing	231	0.2	142	0.21	40	0.19	49	0.23
Cardiogenic shock								
No	108,163	98.5	66,646	99.04	20,460	96.38	21,057	99.15
Yes	1,329	1.2	485	0.72	725	3.42	119	0.56
Missing	267	0.2	161	0.24	44	0.21	62	0.29
Resuscitation								
No	108,958	99.3	66,832	99.32	20,992	98.88	21,134	99.51
Yes	533	0.5	297	0.44	186	0.88	50	0.24
Missing	268	0.2	163	0.24	51	0.24	54	0.25
Arrhythmia								
No arrhythmia	89,779	81.8	57,451	85.38	14,604	68.79	17,724	83.45
AFib/flutter	16,124	14.7	7,569	11.25	5,721	26.95	2,834	13.34
Heart block	1,598	1.5	1,109	1.65	315	1.48	174	0.82
Sustained VT/VF	984	0.9	486	0.72	290	1.37	208	0.98
Arrhythmia–other	688	0.6	324	0.48	175	0.82	189	0.89
Arrhythmia–missing type	312	0.3	175	0.26	74	0.35	63	0.30
Missing	274	0.2	178	0.26	50	0.24	46	0.22
Preoperative IABP								
No	107,945	98.3	66,733	99.17	20,332	95.77	20,880	98.31
Yes	1,431	1.3	342	0.51	809	3.81	280	1.32
Missing	383	0.3	217	0.32	88	0.41	78	0.37
NYHA class								
I	17,413	15.9	10,222	15.19	2,706	12.75	4,485	21.12
II	32,360	29.5	20,295	30.16	4,915	23.15	7,150	33.67
III	40,321	36.7	25,483	37.87	8,205	38.65	6,633	31.23
IV	14,324	13.1	8,104	12.04	4,256	20.05	1,964	9.25
Missing	5,341	4.9	3,188	4.74	1,147	5.40	1,006	4.74

Table 1. Continued

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Congestive heart failure								
No	64,608	58.9	41,972	62.37	9,341	44.00	13,295	62.60
Yes	44,934	40.9	25,185	37.43	11,849	55.82	7,900	37.20
Missing	217	0.2	135	0.20	39	0.18	43	0.20
Number of diseased coronary vessels								
None	90,281	82.3	55,072	81.84	17,525	82.55	17,684	83.27
One	8,947	8.2	5,393	8.01	1,498	7.06	2,056	9.68
Two	3,386	3.1	2,180	3.24	735	3.46	471	2.22
Three	5,611	5.1	3,766	5.60	1,147	5.40	698	3.29
Missing	1,534	1.4	881	1.31	324	1.53	329	1.55
Left main disease ≥ 50%								
No	106,462	97.0	65,328	97.08	20,495	96.54	20,639	97.18
Yes	1,625	1.5	1,127	1.67	289	1.36	209	0.98
Missing	1,672	1.5	837	1.24	445	2.10	390	1.84
Ejection fraction, %								
< 25	2,694	2.5	1,774	2.64	341	1.61	579	2.73
25–34	5,900	5.4	3,810	5.66	1,052	4.96	1,038	4.89
35–44	10,035	9.1	6,181	9.19	2,208	10.40	1,646	7.75
45–54	20,481	18.7	12,411	18.44	4,382	20.64	3,688	17.37
≥ 55	60,890	55.5	36,584	54.37	11,308	53.27	12,998	61.20
Missing	9,759	8.9	6,532	9.71	1,938	9.13	1,289	6.07
Aortic stenosis								
No	54,457	49.6	13,309	19.78	20,303	95.64	20,845	98.15
Yes	54,681	49.8	53,722	79.83	696	3.28	263	1.24
Missing	621	0.6	261	0.39	230	1.08	130	0.61
Mitral stenosis								
No	100,609	91.7	65,186	96.87	15,383	72.46	20,040	94.36
Yes	8,155	7.4	1,401	2.08	5,676	26.74	1,078	5.08
Missing	995	0.9	705	1.05	170	0.80	120	0.57
Tricuspid stenosis								
No	108,073	98.5	66,243	98.44	20,821	98.08	21,009	98.92
Yes	331	0.3	152	0.23	120	0.57	59	0.28
Missing	1,355	1.2	897	1.33	288	1.36	170	0.80
Pulmonic stenosis								
No	107,512	98.0	65,842	97.85	20,783	97.90	20,887	98.35
Yes	141	0.1	91	0.14	29	0.14	21	0.10
Missing	2,106	1.9	1,359	2.02	417	1.96	330	1.55
Aortic insufficiency								
None	59,905	54.6	25,861	38.43	16,701	78.67	17,343	81.66
Trivial	9,191	8.4	5,916	8.79	1,661	7.82	1,614	7.60
Mild	13,282	12.1	10,014	14.88	1,798	8.47	1,470	6.92
Moderate	9,501	8.7	8,815	13.10	382	1.80	304	1.43
Severe	15,722	14.3	15,529	23.08	109	0.51	84	0.40
Missing	2,158	2.0	1,157	1.72	578	2.72	423	1.99
Mitral insufficiency								
None	43,731	39.8	40,453	60.12	2,283	10.75	995	4.68
Trivial	7,743	7.1	7,285	10.83	388	1.83	70	0.33
Mild	14,455	13.2	13,066	19.42	1,089	5.13	300	1.41
Moderate	10,224	9.3	4,438	6.60	3,246	15.29	2,540	11.96
Severe	31,813	29.0	573	0.85	14,045	66.16	17,195	80.96
Missing	1,793	1.6	1,477	2.19	178	0.84	138	0.65

Table 1. Continued

Variable	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
	N	%	N	%	N	%	N	%
Tricuspid insufficiency								
None	78,472	71.5	49,976	74.27	14,266	67.20	14,230	67.00
Trivial	8,856	8.1	5,612	8.34	1,381	6.51	1,863	8.77
Mild	13,346	12.2	7,333	10.90	2,788	13.13	3,225	15.19
Moderate	5,167	4.7	2,126	3.16	1,753	8.26	1,288	6.06
Severe	974	0.9	297	0.44	460	2.17	217	1.02
Missing	2,944	2.7	1,948	2.89	581	2.74	415	1.95
Pulmonic insufficiency								
None	97,954	89.2	60,463	89.85	18,837	88.73	18,654	87.83
Trivial	4,161	3.8	2,370	3.52	779	3.67	1,012	4.77
Mild	2,541	2.3	1,340	1.99	573	2.70	628	2.96
Moderate	441	0.4	209	0.31	144	0.68	88	0.41
Severe	76	0.1	34	0.05	30	0.14	12	0.06
Missing	4,586	4.2	2,876	4.27	866	4.08	844	3.97

AFib = atrial fibrillation; AVR = aortic valve replacement; CV = cardiovascular; CVA = cerebrovascular accident (stroke); IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repaid; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; VF = ventricular fibrillation; VT = ventricular tachycardia.

hospitalization as surgery, regardless of timing, or within 30 days of surgery regardless of venue; (2) permanent stroke (cerebrovascular accident [CVA]): a central neurologic deficit persisting longer than 72 hours; (3) renal failure: a new requirement for dialysis or an increase of the serum creatinine to greater than 2.0 mg/dL and double the most recent preoperative creatinine level; (4) prolonged ventilation (longer than 24 hours); (5) deep sternal wound infection; (6) reoperation for any reason; (7) major morbidity or mortality: a composite defined as the occurrence of any of the above endpoints; (8) prolonged postoperative length of stay (PLOS): length of stay (LOS) more than 14 days (alive or dead); and (9) short postoperative LOS (SLOS): LOS less than 6 days and patient alive at discharge.

Table 2 summarizes the endpoint frequencies in the study population.

Single Versus Multiple Models

Two issues required particularly careful consideration: whether to construct separate models for the AVR and MVR populations, and how best to further subdivide the mitral population into repair versus replacement.

Because of the large size of the STS NCD, separate

Table 2. Frequency of Endpoints in Overall Study Population 2002 to 2006

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
All isolated valve (AVR, MVR, MVRepair)									
N	109,759	109,759	107,060	109,759	109,759	109,759	109,759	109,759	109,759
Events	3,706	1,751	4,673	12,892	307	9,164	20,074	9,718	41,214
%	3.4	1.6	4.3	11.8	0.3	8.4	18.3	8.9	37.6
AVR									
N	67,292	67,292	65,828	67,292	67,292	67292	67,292	67,292	67,292
Events	2,157	1,007	2,774	7,323	197	5369	11,706	5,308	26,144
%	3.2	1.5	4.1	10.9	0.3	8.0	17.4	7.9	38.9
MVR									
N	21,229	21,229	20,329	21,229	21,229	21229	21,229	21,229	21,229
Events	1,210	447	1,348	4,015	71	2450	5,675	3,244	4,727
%	5.7	2.1	6.4	18.9	0.3	11.5	26.7	15.3	22.3
MVRepair									
N	21,238	21,238	20,903	21,238	21,238	21,238	21,238	21,238	21,238
Events	339	297	551	1,554	39	1,345	2,693	1,166	10,343
%	1.6	1.4	2.6	7.3	0.2	6.3	12.7	5.5	48.7

AVR = aortic valve replacement; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

models for AVR, MVR, and MVRepair initially seemed both feasible and appropriate. However, because the endpoints of interest are rare events, we recognized the possibility that the number of such events would be too small to support reliable estimation of the model coefficients.

To assess this tradeoff, we conducted a pilot study to compare two alternative strategies for developing risk models for isolated valve surgery. The first strategy involved developing models separately for three subpopulations (AVR, MVR, and MVRepair). The second strategy involved modeling all three subpopulations together in a single model; several interaction terms were included to allow the effect of selected risk factors to differ across the subpopulations. Both strategies were pilot tested by developing risk models for two endpoints: operative mortality and permanent stroke. These pilot models were developed in a 60% development sample and tested in a separate 40% validation sample. Each model was assessed by calculating the c-index and the generalized R^2 index of Nagelkerke [7] in the validation sample for each combination of subpopulation and endpoint (3 subpopulations $\times$ 2 endpoints = 6 combinations). With the exception of AVR operative mortality, the combined model with interactions resulted in better discrimination. With the exception of MVR and MVrepair operative mortality, the combined model also captured more variation as measured by the generalized R^2 statistic.

Because the combined model strategy performed better in the majority of cases, and because a single combined model was consistent with the previous STS valve model, the combined model strategy was selected. To avoid assuming that the weighting of each risk factor was exactly constant across the three populations, we included interactions between surgery type and several key predictor variables. In principle, fitting a single model with several interactions is advantageous because it allows for pooling information across related groups without making an a priori assumption that all of the covariate effects are exactly constant across groups.

Selection of Candidate Predictor Variables

Our general approach to variable selection is discussed in Part 1 of this series describing the development of the 2008 STS isolated CABG risk models. Briefly, we initially identified potential candidate variables by reviewing four versions of the STS data collection instrument (data versions 2.35, 2.41, 2.52.1, and 2.61) as well as previously published STS and similar cardiac risk models [1–6]. A panel of cardiac surgeons and health policy experts reviewed the initial variables for face validity and to be certain that no important predictor variables available in (or mappable to) to STS NCD data version 2.61 had been excluded.

Final candidate explanatory variables and their coding are summarized in Table 3. The variables were identical to the CABG model candidate variables with the following differences: (1) percutaneous coronary intervention conducted within 6 hours or less of surgery was not a candidate variable because it was present in only 122 patients (0.1%) in the valve model population; (2) infectious endocarditis was included. This risk factor was rarely present among isolated CABG patients (0.09%), but was not uncommon (7.7%) among patients undergoing valve surgery; (3) mitral stenosis was included; this risk factor was rarely present among isolated CABG patients (0.35%) but was common (7.4%) among patients undergoing valve surgery; and (4) an indicator for surgery type (AVR, MVR, MVRepair) was included in the valve models.

Coding of Explanatory Variables

The coding of continuous and categorical variables was identical to the CABG models, except for the following differences: (1) age was modeled as a linear spline truncated from below at 50 years and with a change of slope at 75; (2) creatinine was modeled as a linear term with values less than 0.5 and greater than 5.0 mapped to those values respectively (approximately the 1st and 99th percentiles of the empirical distribution); (3) previous myocardial infarction (MI) was modeled as three categories ($<$ 24 hours, 1 to 21 days, and $>$ 21 days or no MI); the first two categories were subsequently combined after expert panel review; (4) race was modeled as three categories: black, Hispanic, Caucasian/other; and (5) chronic lung disease was modeled as linear across four categories (none, mild, moderate, severe).

In general, these differences reflect a slightly simpler coding scheme (fewer parameters) for the valve models compared with the isolated CABG models.

Repair Versus Replacement

In addition to a number of variables whose inclusion or coding were noted to be problematic during development of the 2008 STS isolated CABG models (Part 1 of this series), the approach to modeling mitral valve repair versus replacement was of some concern in the valve models. From a methodologic perspective, models used for risk-adjustment should include all patient preoperative risk factors that vary in prevalence between institutions and that substantially impact the probability of an adverse outcome. Such models should include variables that reflect the patient's baseline condition but should not include intraoperative events (eg, unexpected hemorrhage) or discretionary care processes (eg, use of a mechanical versus bioprosthetic valve). Adjusting for intraoperative events is not appropriate because these may be a reflection of the surgeon's performance. Adjusting for discretionary care processes may likewise mask differences in performance if the surgeon's choice of procedures has a substantial impact on outcomes. The same patient may receive valve repair if treated by one surgeon and replacement if treated by another. Adjusting for repair versus replacement will potentially conceal the outcomes of surgeons who achieve excellent results by repairing technically challenging valves that might otherwise be replaced if treated by a surgeon with less skill or tenacity. Importantly, there is considerable evidence to suggest the superiority of valve repair whenever feasible.

However, in addition to such discretionary factors, the decision to repair rather than replace the mitral valve is

Table 3. List of Final Candidate Variables and Their Coding for STS Valve Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline truncated from below at 50 and with knot at 75
Ejection fraction	Linear, values > 50 mapped to 50
Body surface area[a]	Quadratic polynomial modeled separately for males and females. Note: body surface area < 1.4 and > 2.6 mapped to those values, respectively.
Creatinine	Linear (only for patients not on dialysis). Note: creatinine < 0.5 and > 5.0 mapped to those values, respectively.
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval. Modeled as linear across categories.
Binary variables	
Active infectious endocarditis	Yes/no
Dialysis	Yes/no
Preoperative atrial fibrillation	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Preoperative IABP or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no MI < 7 days)	Yes/no
Left main disease	Yes/no
Aortic stenosis	Yes/no
Mitral stenosis	Yes/no
Aortic insufficiency	Defined as at least moderate (yes/no)
Mitral insufficiency	Defined as at least moderate (yes/no)
Tricuspid insufficiency	Defined as at least moderate (yes/no)
Categorical variables	
Chronic lung disease	Modeled as linear across categories (none, mild, moderate, severe)
CVD/CVA	3 groups: no CVD, CVD no CVA, CVD + CVA
Diabetes mellitus	3 groups: insulin diabetes, noninsulin diabetes, other or no diabetes
Number diseased coronary vessels	3 groups: < 2, 2, 3. Modeled as linear across the categories
MI	3 groups: < 24 hr, 1–21 days, > 21 days or no MI (groups 1 and 2 were subsequently collapsed)
Race	3 groups: Black; Hispanic; Other including Caucasian
Status	4 groups: elective, urgent, emergent—no resuscitation, salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, ≥2 previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF+NYHA IV
Surgery type	3 groups: AVR, MVR, MVRepair
Interaction terms	
Age by reoperation[a]	
Age by emergent status[a]	
Surgery type by each of the following:	Age, diabetes, dialysis, creatinine, reoperation, endocarditis, emergent status, CLD, CHF, EF, sex, shock, IABP/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, aortic stenosis

[a] These variables were forced into each model.

AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association.

also dependent upon the patient's preoperative valve disease etiology, anatomy, and pathophysiology. On average, patients amenable to valve repair have less extensive valve pathology and a relatively favorable postoperative prognosis (the mortality rate for valve repair is 1.6% compared with 5.7% for replacement). Ignoring these anatomical differences can introduce bias when comparing institutions, especially because these variables are not captured elsewhere on the STS data collection form.

A related difficulty in adjusting for repair versus re-

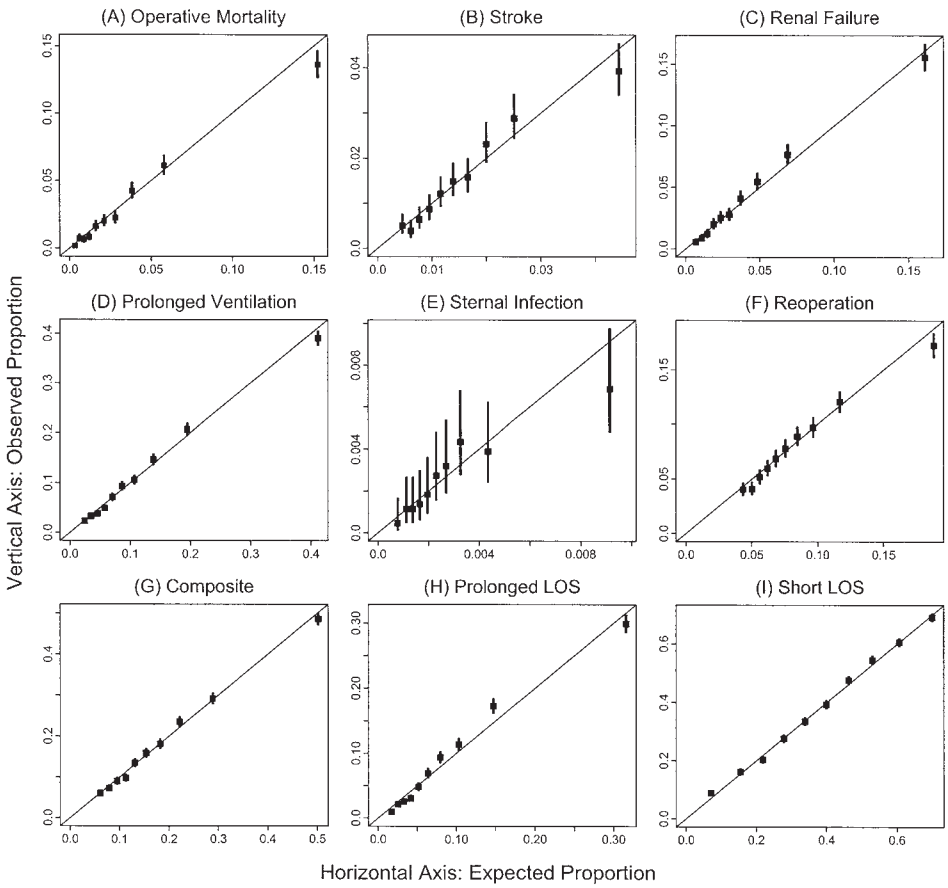


Fig 1. Plots of observed (O) versus expected (E) in validation sample

placement is that the former approach may sometimes be abandoned intraoperatively by the surgeon and converted to MVR. That may sometimes occur because of unforeseen technical problems that would prevent most surgeons from completing the repair, but in other instances, a more skilled surgeon might persist and achieve successful valve repair. Effectively separating these two scenarios is problematic from available data.

Ultimately, it was elected to include an indicator for mitral valve repair versus replacement in the valve risk models, consistent with the approach in a number of existing valve surgery models. We acknowledge that available data make it impossible to determine whether patient differences or surgical skill and judgment are the most important factors in determining between-provider variation in the proportion of valves repaired.

Recognizing the potential limitations of this modeling approach, the decision to adjust for repair versus replacement may be reassessed in future versions of the STS risk models. Beginning with data in version 2.61, the database will capture whether or not repair was attempted, and repair versus replacement may be analyzed based on an intention-to-treat principle.

Missing Data

Model variables with more than 1% missing data in the study sample were ejection fraction (8.9%), NYHA class (4.9%), tricuspid insufficiency (2.7%), aortic insufficiency (2.0%), mitral insufficiency (1.6%), left main disease (1.5%), creatinine/dialysis (1.6%), and number of diseased vessels (1.4%). The method of imputing missing data was identical to that employed in the isolated CABG models and described in Part 1 of this series. Briefly, binary risk factors were modeled as yes versus no or missing (ie, missing values were analyzed as if the endpoint did not occur). Missing data on categorical variables were imputed to the lowest risk value, typically the mode, and outcomes were typically similar for missing data and lowest risk patients. Missing data on continuous variables were imputed by grouping patients into strata and assigning the stratum-specific median value. For example, ejection fraction was imputed by grouping on sex and congestive heart failure and calculating the median ejection fraction among patients with nonmissing ejection fraction in each group.

Although multiple imputation is generally preferable to single imputation [8], single imputation was chosen for this analysis mainly because of practical considerations. Furthermore, because of the small fraction of missing data, the impact of single versus multiple imputation was considered to be inconsequential. Subsequent sensitivity analyses confirmed that the choice between single versus multiple imputation had little impact on the final regression coefficients, risk estimates, and confidence intervals. A summary of these sensitivity analyses, including coef-

Table 4. Discrimination of Models in Development and Validation Samples

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall									
Development sample	0.805	0.694	0.782	0.770	0.704	0.643	0.721	0.770	0.738
Validation sample	0.799	0.691	0.762	0.762	0.659	0.639	0.718	0.773	0.734
AVR									
Development sample	0.779	0.679	0.766	0.748	0.710	0.630	0.698	0.752	0.713
Validation sample	0.759	0.689	0.749	0.736	0.637	0.619	0.694	0.759	0.713
MVR									
Development sample	0.794	0.679	0.767	0.772	0.591	0.642	0.735	0.748	0.726
Validation sample	0.802	0.702	0.748	0.772	0.656	0.634	0.738	0.729	0.710
MVRepair									
Development sample	0.855	0.736	0.813	0.765	0.774	0.616	0.703	0.777	0.733
Validation sample	0.844	0.672	0.788	0.773	0.714	0.646	0.712	0.800	0.725

AVR = aortic valve replacement; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

ficients and covariance matrices, is available at www.sts.org/riskmodels.

Final Variable Selection Procedure

Variables were initially selected using an automated stepwise model selection algorithm. The stepwise procedure began with a model that included all of the final candidate variables except for interaction terms. Age, sex, body surface area, and month of surgery were forced into each model. Other variables were selected in a stepwise fashion using a significance criterion of 0.05 for entry and removal. This criterion was less stringent than that employed in development of the CABG models, because the sample size in the former was so much larger than that which was used for the valve models. The stepwise procedure was performed separately for each endpoint. The results were then reviewed by an expert panel of surgeons, and the following changes were made based on their feedback: (1) "MI less than 24 hours" and "MI 1 to 21 days" were collapsed into a single category; (2) preoperative atrial fibrillation was forced into the model for stroke (CVA); and (3) an indicator variable for dialysis was forced into any model that included creatinine level.

Interaction Terms

In addition to including main effects, we tested the interaction between surgery group (AVR, MVR, MVRepair) and each of the following variables: age, diabetes mellitus, dialysis, creatinine, reoperation, endocarditis, emergent status, chronic lung disease, congestive heart failure, ejection fraction, sex, shock, intra-aortic balloon pump/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, and aortic stenosis. These interaction terms allowed the effect of these selected risk factors to differ across the surgery populations.

Four additional sets of interactions were also included in the models: (1) sex by body surface area (BSA); (2) sex by BSA^2; (3) age by reoperation; and (4) age by emergent status. These interaction terms were preselected and were not tested as part of the backward selection algorithm. Additional technical details are provided in the Appendix. For reasons described in Part 1 of this series (isolated CABG risk models), an extensive automated search for additional interaction terms was not conducted.

Adjustment for Time Trends

Surgery date was included in each model to adjust for changes in the frequency of adverse outcomes over the 5-year study period. Although surgery date is not itself a variable of interest, we adjusted for it to reduce potential confounding by time trends when estimating regression coefficients for the variables that are of primary interest (ie, patient preoperative risk factors). An example is provided in Part 1 of this series.

Surgery date was categorized into 6-month intervals (corresponding to the biannual STS data harvests) and modeled as a linear trend across the ordinal categories. Because it is a nuisance variable, surgery date is not included in the final risk prediction algorithm. Thus, a patient's predicted risk does not depend on the patient's surgery date. As described in the Appendix, the published intercept parameter has been adjusted to incorporate the time trend. The adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Results

Assessment of Model Fit and Discrimination

Because of the relatively large size of our sample, the Hosmer-Lemeshow test is uninformative and would invariably result in a significant p value [9]. As an alternative, model fit was assessed graphically by plotting observed versus predicted rates of each endpoint across deciles of predicted risk in the development and validation samples. This was done in the overall population and in subgroups based on surgery type (AVR, MVR, MVRepair); age ($<$ 60, 60 to 79, $\geq$ 80 years); sex (male, female); diabetes mellitus (yes/no); status (elective, nonelective); and ejection fraction

Table 5. Odds Ratios (95% Confidence Intervals) for the Final Selected Models

A. Odds ratios for variables that do not interact with surgery group

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.10, 1.31)	1.06 (0.93, 1.20)	NA	1.18 (1.11, 1.25)	NA	1.11 (1.04, 1.18)	1.12 (1.07, 1.18)	1.17 (1.10, 1.24)	0.74 (0.70, 0.78)
BSA 1.6 versus 2.0 among females	1.19 (1.09, 1.30)	1.18 (1.03, 1.35)	0.95 (0.87, 1.04)	1.15 (1.08, 1.22)	0.42 (0.27, 0.68)	1.26 (1.18, 1.34)	1.17 (1.12, 1.23)	1.11 (1.04, 1.17)	0.99 (0.95, 1.04)
BSA 1.6 versus 2.0 among males	1.75 (1.48, 2.07)	1.17 (0.92, 1.47)	1.33 (1.12, 1.58)	1.56 (1.41, 1.74)	0.94 (0.49, 1.84)	1.34 (1.21, 1.49)	1.44 (1.33, 1.57)	1.39 (1.25, 1.56)	0.73 (0.68, 0.79)
BSA 1.8 versus 2.0 among females	0.99 (0.95, 1.04)	1.08 (0.99, 1.17)	0.90 (0.86, 0.94)	1.00 (0.97, 1.03)	0.65 (0.54, 0.77)	1.07 (1.03, 1.11)	1.02 (0.99, 1.04)	0.99 (0.96, 1.02)	1.05 (1.03, 1.08)
BSA 1.8 versus 2.0 among males	1.21 (1.14, 1.29)	1.07 (0.98, 1.16)	1.07 (1.00, 1.14)	1.14 (1.10, 1.19)	0.90 (0.70, 1.14)	1.12 (1.08, 1.16)	1.12 (1.09, 1.16)	1.10 (1.06, 1.15)	0.92 (0.89, 0.94)
BSA 2.2 versus 2.0 among females	1.21 (1.11, 1.33)	0.94 (0.80, 1.10)	1.30 (1.21, 1.41)	1.15 (1.09, 1.21)	1.57 (1.26, 1.96)	1.02 (0.95, 1.09)	1.12 (1.07, 1.16)	1.14 (1.08, 1.21)	0.85 (0.81, 0.88)
BSA 2.2 versus 2.0 among males	0.98 (0.93, 1.03)	0.95 (0.88, 1.03)	1.09 (1.05, 1.14)	1.05 (1.02, 1.08)	1.32 (1.17, 1.48)	0.95 (0.93, 0.98)	1.02 (0.99, 1.04)	1.03 (1.00, 1.07)	0.94 (0.93, 0.96)
Creatinine per 1 unit	1.55 (1.46, 1.64)	1.34 (1.22, 1.47)	2.04 (1.93, 2.16)	1.58 (1.51, 1.65)	NA	1.27 (1.20, 1.33)	1.64 (1.57, 1.71)	1.58 (1.51, 1.65)	0.64 (0.61, 0.68)
CVD with CVA	NA	1.81 (1.56, 2.10)	1.22 (1.09, 1.37)	1.28 (1.18, 1.38)	NA	1.14 (1.05, 1.24)	1.20 (1.12, 1.28)	1.40 (1.29, 1.52)	0.77 (0.72, 0.83)
CVD without CVA	NA	1.32 (1.11, 1.57)	1.23 (1.10, 1.37)	1.14 (1.05, 1.23)	NA	1.06 (0.96, 1.17)	1.08 (1.01, 1.15)	NA	0.80 (0.73, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	NA	1.10 (1.01, 1.20)	NA	1.07 (1.02, 1.11)	NA	NA	1.04 (1.00, 1.08)	1.03 (0.98, 1.08)	0.90 (0.86, 0.94)
EF per 10-unit decrease	1.09 (1.05, 1.14)	NA	1.04 (1.00, 1.09)	1.12 (1.09, 1.15)	1.26 (1.12, 1.41)	1.08 (1.04, 1.11)	1.10 (1.07, 1.12)	1.12 (1.08, 1.15)	0.87 (0.85, 0.90)
Hypertension	1.12 (1.03, 1.22)	1.19 (1.07, 1.33)	1.35 (1.25, 1.45)	1.11 (1.06, 1.17)	NA	NA	1.11 (1.07, 1.15)	NA	0.94 (0.91, 0.97)
Immunosuppressive treatment	1.42 (1.21, 1.67)	NA	1.39 (1.19, 1.62)	NA	NA	NA	1.16 (1.06, 1.27)	1.31 (1.17, 1.47)	NA
Left main disease	1.19 (0.98, 1.46)	NA	1.19 (0.98, 1.44)	NA	2.17 (1.13, 4.16)	NA	NA	NA	NA
Active infectious endocarditis	1.95 (1.68, 2.27)	1.87 (1.52, 2.29)	2.17 (1.88, 2.50)	2.15 (1.95, 2.36)	NA	1.55 (1.39, 1.73)	1.97 (1.80, 2.15)	2.79 (2.51, 3.09)	0.34 (0.30, 0.38)
Mitral insufficiency, moderate/severe	NA	1.26 (1.14, 1.39)	NA	NA	NA	NA	NA	NA	NA
Tricuspid insufficiency, moderate/severe	NA	NA	1.14 (1.01, 1.29)	1.14 (1.04, 1.25)	NA	1.09 (1.00, 1.20)	1.21 (1.12, 1.30)	1.17 (1.05, 1.31)	0.82 (0.73, 0.92)
Peripheral vascular disease	1.25 (1.12, 1.38)	1.29 (1.11, 1.49)	NA	NA	NA	1.22 (1.12, 1.32)	1.14 (1.07, 1.21)	1.17 (1.09, 1.25)	0.83 (0.78, 0.88)
Aortic stenosis		NA	NA	0.90 (0.83, 0.97)	NA	0.90 (0.84, 0.96)	0.93 (0.87, 0.98)	0.86 (0.80, 0.92)	1.07 (1.02, 1.13)
Mitral stenosis	1.24 (1.08, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA
MI $\leq$ 21 days	1.14 (0.98, 1.34)	NA	NA	1.37 (1.22, 1.55)	NA	1.04 (0.91, 1.18)	1.28 (1.16, 1.41)	1.21 (1.06, 1.37)	0.81 (0.72, 0.91)
Time trend, per 6-month harvest interval	0.98 (0.97, 0.99)	0.98 (0.96, 1.00)	1.01 (0.99, 1.02)	1.02 (1.01, 1.03)	0.97 (0.93, 1.01)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Race black	NA	1.33 (1.13, 1.57)	1.51 (1.34, 1.69)	1.42 (1.27, 1.58)	NA	1.27 (1.15, 1.40)	1.37 (1.27, 1.49)	1.45 (1.31, 1.60)	0.64 (0.59, 0.70)
Race Hispanic	NA	0.87 (0.64, 1.19)	1.16 (0.97, 1.38)	1.07 (0.94, 1.22)	NA	1.14 (1.00, 1.30)	1.09 (0.98, 1.22)	1.16 (0.98, 1.38)	0.82 (0.72, 0.93)
Status urgent	1.29 (1.19, 1.40)	NA	1.21 (1.11, 1.33)	1.29 (1.20, 1.39)	NA	1.17 (1.10, 1.25)	1.22 (1.15, 1.29)	1.42 (1.33, 1.51)	0.70 (0.66, 0.74)
Unstable angina	1.21 (1.04, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA

Table 5. Continued

B. Odds ratios for aortic valve replacement

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.43 (1.34, 1.52)	1.48 (1.38, 1.59)	1.38 (1.30, 1.47)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.16 (1.12, 1.21)	1.23 (1.19, 1.26)	1.31 (1.25, 1.37)	0.75 (0.73, 0.77)
Age 70 versus 50	2.04 (1.79, 2.32)	2.19 (1.90, 2.52)	1.90 (1.68, 2.16)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.35 (1.25, 1.46)	1.50 (1.42, 1.59)	1.71 (1.55, 1.87)	0.57 (0.54, 0.60)
Age 80 versus 50	3.34 (2.84, 3.93)	3.21 (2.70, 3.81)	2.88 (2.46, 3.37)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.59 (1.44, 1.76)	1.97 (1.82, 2.12)	2.50 (2.24, 2.79)	0.34 (0.32, 0.36)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.33 (1.24, 1.43)	NA	NA	1.20 (1.13, 1.27)	1.25 (1.17, 1.34)	0.86 (0.81, 0.91)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.92 (1.77, 2.08)	NA	1.25 (1.17, 1.35)	1.62 (1.51, 1.73)	1.54 (1.40, 1.68)	0.72 (0.65, 0.79)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.42 (1.31, 1.55)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.39 (1.29, 1.50)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.12 (1.04, 1.20)	NA	NA	1.12 (1.06, 1.18)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	2.85 (2.35, 3.45)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.29 (0.24, 0.34)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	1.78 (1.55, 2.05)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.56 (0.48, 0.66)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	1.74 (1.37, 2.21)	NA
Female versus male (at BSA = 1.8)	1.23 (1.10, 1.36)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.29 (1.21, 1.38)	0.98 (0.72, 1.33)	0.86 (0.81, 0.93)	1.03 (0.98, 1.08)	1.25 (1.16, 1.35)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.27 (1.21, 1.33)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.29 (1.24, 1.34)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	1.83 (1.64, 2.05)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation, ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.49 (2.14, 2.90)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	3.77 (2.75, 5.16)	2.78 (1.85, 4.17)	3.10 (2.21, 4.35)	4.54 (3.54, 5.83)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	7.94 (5.40, 11.66)	2.11 (1.06, 4.19)	3.47 (2.19, 5.51)	3.50 (2.41, 5.08)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

Table 5. Continued

C. Odds ratios for mitral valve replacement

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.65 (1.53, 1.78)	1.48 (1.38, 1.59)	1.35 (1.26, 1.44)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.25 (1.19, 1.31)	1.33 (1.29, 1.39)	1.26 (1.21, 1.33)	0.71 (0.68, 0.74)
Age 70 versus 50	2.71 (2.33, 3.17)	2.19 (1.90, 2.52)	1.81 (1.60, 2.06)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.56 (1.42, 1.71)	1.78 (1.65, 1.92)	1.60 (1.45, 1.76)	0.50 (0.46, 0.55)
Age 80 versus 50	5.14 (4.15, 6.37)	3.21 (2.70, 3.81)	2.67 (2.23, 3.20)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.97 (1.72, 2.26)	2.54 (2.27, 2.84)	2.27 (2.00, 2.58)	0.28 (0.25, 0.32)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.19 (1.07, 1.32)	NA	NA	1.11 (1.01, 1.21)	1.25 (1.17, 1.34)	0.96 (0.87, 1.06)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.72 (1.55, 1.91)	NA	1.25 (1.17, 1.35)	1.49 (1.36, 1.64)	1.54 (1.40, 1.68)	0.80 (0.71, 0.91)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.66 (1.47, 1.86)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.67 (1.52, 1.83)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.30 (1.16, 1.45)	NA	NA	1.34 (1.22, 1.47)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	4.59 (3.65, 5.77)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.23 (0.16, 0.33)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	2.21 (1.90, 2.56)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.63 (0.51, 0.77)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	1.05 (0.85, 1.31)	NA
Female versus male (at BSA=1.8)	1.11 (0.97, 1.27)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.06 (0.98, 1.16)	0.98 (0.72, 1.33)	0.79 (0.72, 0.87)	1.03 (0.98, 1.08)	1.09 (0.99, 1.19)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.08 (1.01, 1.16)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.16 (1.11, 1.22)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	1.50 (1.34, 1.67)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation, ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.03 (1.76, 2.35)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	2.74 (1.99, 3.78)	2.78 (1.85, 4.17)	2.20 (1.59, 3.05)	3.19 (2.41, 4.23)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	5.78 (3.77, 8.85)	2.11 (1.06, 4.19)	2.46 (1.56, 3.88)	2.46 (1.66, 3.65)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

Table 5. Continued

D. Odds ratios for mitral valve repair

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.80 (1.62, 2.00)	1.48 (1.38, 1.59)	1.55 (1.41, 1.71)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.20 (1.13, 1.27)	1.31 (1.26, 1.37)	1.50 (1.41, 1.60)	0.62 (0.60, 0.65)
Age 70 versus 50	3.24 (2.63, 4.00)	2.19 (1.90, 2.52)	2.42 (2.00, 2.92)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.44 (1.29, 1.62)	1.73 (1.58, 1.89)	2.25 (1.98, 2.55)	0.39 (0.36, 0.42)
Age 80 versus 50	6.72 (5.00, 9.04)	3.21 (2.70, 3.81)	4.11 (3.14, 5.38)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.75 (1.48, 2.07)	2.42 (2.12, 2.76)	3.78 (3.17, 4.51)	0.19 (0.17, 0.22)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.16 (0.99, 1.35)	NA	NA	1.11 (0.99, 1.24)	1.25 (1.17, 1.34)	0.92 (0.80, 1.05)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.67 (1.43, 1.95)	NA	1.25 (1.17, 1.35)	1.50 (1.33, 1.68)	1.54 (1.40, 1.68)	0.76 (0.65, 0.90)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.68 (1.42, 1.97)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.57 (1.36, 1.81)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.31 (1.11, 1.55)	NA	NA	1.26 (1.10, 1.45)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	6.24 (4.19, 9.30)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.26 (0.19, 0.37)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	2.90 (2.28, 3.70)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.49 (0.38, 0.64)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	2.50 (1.51, 4.12)	NA
Female versus male (at BSA = 1.8)	0.97 (0.77, 1.21)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.23 (1.10, 1.38)	0.98 (0.72, 1.33)	0.90 (0.80, 1.02)	1.03 (0.98, 1.08)	1.28 (1.12, 1.47)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.23 (1.09, 1.39)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.26 (1.15, 1.40)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	2.06 (1.73, 2.45)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.80 (2.32, 3.37)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	8.73 (4.84, 15.74)	2.78 (1.85, 4.17)	3.03 (1.69, 5.43)	6.12 (3.96, 9.46)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	18.39 (9.68, 34.96)	2.11 (1.06, 4.19)	3.39 (1.76, 6.54)	4.72 (2.71, 8.23)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

[a] Variable interacts with age. Reported odds ratio represents effect of risk factor for patients aged 50 years old.

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; NA = not applicable; NYHA = New York Heart Association; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

(≤ 40, > 40). Calibration plots (observed versus expected) based on the overall validation sample are presented in Figure 1. The average absolute difference between observed versus predicted event rates within deciles of predicted risk ranged from 0.06% for deep sternal wound infection to 1.06% for prolonged postoperative stay. Analogous figures were produced for specific valve procedures and numerous subgroups, and these are available at www.sts.org/riskmodels.

Model fit appeared to be adequate for each endpoint with the possible exception of deep sternal wound infection, which revealed some overfitting within certain subgroups. A modest degree of overfitting was expected for this endpoint given the relatively small number of infections and large number of candidate predictors.

Discrimination was assessed by the c-statistic, also known as the area under the receiver operating characteristic (ROC) curve. Table 4 presents the discrimination of each model in the development and validation samples for all patients combined and for subgroups consisting of AVR, MVR, and MVRepair. In the validation sample, c-statistics for the operative mortality model were 0.799 (overall), 0.759 (AVR), 0.802 (MVR), and 0.844 (MVRepair). C-statistics in the validation sample for other endpoints ranged from 0.619 for reoperation in the AVR subgroup to 0.800 for prolonged length of stay in the MVRepair subgroup.

Final Models

After validating the models in the 40% validation sample, the development and validation samples were then combined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regressions were estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [10]. An independence working correlation matrix was used to apply the generalized estimating equations methodology. With this approach, the estimated regression coefficients were identical to those obtained using ordinary logistic regression, but the standard errors were adjusted to account for the clustered data structure.

Odds Ratios

Odds ratios and 95% confidence intervals (CI) for the final selected models are presented in Table 5. "Not applicable" indicates that the specific predictor was not included in a particular risk model. Because several variables interact with surgery type, the odds ratios for these variables differ depending on the type of surgery (AVR, MVR, MVRepair). For example, in the operative mortality model, the odds ratio for emergent status is 3.77 (95% CI: 2.75, 5.16) for AVR, 2.74 (95% CI: 1.99, 3.78) for MVR, and 8.73 (95% CI: 4.84, 15.74) for MVRepair. Odds ratios that do not interact with surgery type are summarized in Table 5, Part A. Odds ratios that differ by surgery type for at least one endpoint are presented in Table 5, Parts B, C, and D.

Final Model Intercept and Coefficients

The final risk prediction algorithms, including all coefficients and intercepts, are presented in the Appendix.

Limitations

The limitations for these valve models are similar to those for the CABG models and are thoroughly discussed in Part 1 of this series (2008 STS CABG risk models).

Conclusion

The STS Quality Measurement Task Force has developed and tested nine new risk-adjustment models for isolated valve surgery using the STS NCD. This report includes a detailed exposition of the model development process, including not only statistical issues but also the many clinical and pragmatic judgments that were required. An online risk calculator is also available through a link from the STS website.

References

1. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
2. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
3. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
4. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
5. Kuduvalli M, Grayson AD, Au J, Grotte G, Bridgewater B, Fabri BM. A multi-centre additive and logistic risk model for in-hospital mortality following aortic valve replacement. Eur J Cardiothorac Surg 2007;31:607–13.
6. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
7. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
8. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2002.
9. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.
10. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Models

For each endpoint, the formula for calculating a patient's predicted risk of the endpoint has the form:

$$\text{Predicted Risk} = \frac{e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}{1+e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}$$

where $x_1, x_2, \ldots, x_n$ denote patient preoperative risk factors (eg, quantitative variables such as age, and comorbidities coded as 1 = present, 0 = absent); and $\beta_0, \beta_1, \ldots, \beta_n$ denote regression coefficients (numerical constants). Regression coefficients for each endpoint are presented in Appendix Table 1. The variables $x_1, x_2, \ldots, x_n$ are the same for each endpoint and are defined in Appendix Table 2. The regression coefficient for the time trend is not presented. Instead, the intercept has been adjusted to incorporate the time trend. This adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−5.78680	−5.83957	−5.52789	−3.96796	−7.11095	−3.08816	−3.06527	−4.30676	1.25115
Atrial fibrillation	0.18074	0.05524	0.00000	0.16527	0.00000	0.10305	0.11403	0.15530	−0.30247
Age function 1	0.03557	0.03909	0.03219	0.02683	0.04180	0.01512	0.02041	0.02670	−0.02834
Age function 3	0.02804	−0.00132	0.01809	0.00629	−0.05024	0.00218	0.01282	0.02315	−0.04637
Age by reoperation function	−0.01308	−0.02043	−0.00551	−0.00840	−0.00939	−0.00697	−0.00684	−0.00485	0.00927
Age by status function	−0.02495	−0.02987	−0.00721	−0.01377	0.00277	0.00102	−0.00677	−0.00379	−0.00795
Age by MVR function	0.01436	0.00000	−0.00245	0.00000	0.00000	0.00715	0.00848	−0.00324	−0.00603
Age by MVRepair function	0.02326	0.00000	0.01190	0.00000	0.00000	0.00315	0.00685	0.01378	−0.01883
BSA function 1	−1.40168	−0.38619	−0.71012	−1.11750	0.14188	−0.73553	−0.91858	−0.82801	0.77317
BSA function 2	2.16782	0.23148	1.92875	2.29127	2.04603	0.83644	1.65638	1.65423	−1.76728
CHF but not NYHA IV	0.25590	0.00000	0.21233	0.28353	0.00000	0.00000	0.17974	0.22508	−0.15108
CHF and NYHA IV	0.60544	0.00000	0.47812	0.65056	0.00000	0.22686	0.48025	0.42957	−0.33521
CHF by MVR function	0.00000	0.00000	0.00000	−0.11007	0.00000	0.00000	−0.07864	0.00000	0.11503
CHF by MVRepair function	0.00000	0.00000	0.00000	−0.13792	0.00000	0.00000	−0.07731	0.00000	0.06468
CLD function	0.23846	0.00000	0.16629	0.22816	0.23817	0.08406	0.16044	0.25263	−0.21022
CLD by MVR function	−0.15906	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.10092	0.00000
CLD by MVRepair function	−0.03243	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.01795	0.00000
Creatinine function 1	0.43909	0.29230	0.71439	0.45646	0.00000	0.23562	0.49230	0.45631	−0.44178
CVD without prior CVA	0.00000	0.27837	0.20531	0.12726	0.00000	0.05830	0.07684	0.00000	−0.22223
CVD and prior CVA	0.00000	0.59220	0.20018	0.24512	0.00000	0.13200	0.18343	0.33480	−0.25595
Diabetes, noninsulin	0.23563	0.00000	0.37172	0.11040	0.00000	0.00000	0.11355	0.19843	−0.16630
Diabetes, insulin	0.48368	0.00000	0.64648	0.35367	0.44389	0.18293	0.33165	0.51913	−0.45093
Diabetes by MVR function	0.00000	0.00000	0.00000	0.15051	0.00000	0.00000	0.17990	0.00000	0.00000
Diabetes by MVRepair function	0.00000	0.00000	0.00000	0.16260	0.00000	0.00000	0.11734	0.00000	0.00000
Dialysis	1.48666	0.79199	0.00000	1.57690	1.19109	0.81972	1.37741	1.53351	−1.69019
Dialysis by MVR function	0.47550	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.20998
Dialysis by MVRepair function	0.78385	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.07964
Ejection fraction function	0.00904	0.00000	0.00407	0.01107	0.02308	0.00734	0.00925	0.01111	−0.01348
Endocarditis, active	0.66737	0.62434	0.77276	0.76318	0.00000	0.43876	0.67810	1.02521	−1.08299
Female	0.20372	0.21925	−0.03031	0.25668	−0.02355	−0.14567	0.03066	0.22437	−0.36400
Female by MVR function	−0.10089	0.00000	0.00000	−0.19465	0.00000	−0.08773	0.00000	−0.14211	0.00000
Female by MVRepair function	−0.23812	0.00000	0.00000	−0.04564	0.00000	0.04424	0.00000	0.02470	0.00000
Female by BSA function 1	0.96491	−0.02257	0.83074	0.77598	2.00214	0.16707	0.52716	0.57195	−0.75434
Female by BSA function 2	0.18084	−0.07419	0.08397	−0.58460	−1.87036	0.25158	−0.09063	−0.12289	0.35123
Hypertension	0.11372	0.17789	0.29770	0.10799	0.00000	0.00000	0.10361	0.00000	−0.06504
IABP or inotropes	0.38682	0.00000	0.29606	0.57608	0.52474	0.13432	0.56046	0.37621	−0.57115
IABP by MVR function	0.00000	0.00000	0.00000	0.21517	0.00000	0.00000	0.00000	0.00000	0.10760
IABP by MVRepair function	0.00000	0.00000	0.00000	0.48870	0.00000	0.00000	0.00000	0.00000	−0.13850

Appendix Table 1. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Immunosuppressive treatment	0.35022	0.00000	0.32828	0.00000	0.00000	0.00000	0.14887	0.27152	0.00000
Insufficiency mitral	0.00000	0.23253	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Insufficiency tricuspid	0.00000	0.00000	0.13159	0.12973	0.00000	0.08969	0.18929	0.15846	−0.20027
Left main disease	0.17593	0.00000	0.17280	0.00000	0.77557	0.00000	0.00000	0.00000	0.00000
MI ≤ 21 days	0.13276	0.00000	0.00000	0.31706	0.00000	0.03495	0.24687	0.18812	−0.20961
MVR	0.10284	0.00000	0.40455	0.44639	0.00000	0.12852	0.13795	0.58004	−0.61402
MVRepair	−0.65440	0.00000	−0.23666	−0.19726	0.00000	−0.22398	−0.23002	−0.37618	0.25710
No. diseased vessel function	0.00000	0.09556	0.00000	0.06299	0.00000	0.00000	0.03700	0.03312	−0.10126
Peripheral vascular disease	0.21980	0.25236	0.00000	0.00000	0.00000	0.19758	0.13174	0.15342	−0.18903
Race black	0.00000	0.28378	0.40941	0.34795	0.00000	0.23856	0.31567	0.37161	−0.44177
Race Hispanic	0.00000	−0.13774	0.14968	0.06720	0.00000	0.12816	0.08581	0.15128	−0.20068
Reop, 1 previous operation	0.74484	0.73489	0.43804	0.60704	0.00000	0.27365	0.44052	0.35252	−0.40042
Reop, ≥ 2 previous operations	0.90625	0.85841	0.50595	0.91229	0.00000	0.34233	0.67201	0.56294	−0.69765
Reop by MVR function	0.00000	0.00000	0.00000	−0.20333	0.00000	0.00000	0.00000	0.00000	0.00000
Reop by MVRepair function	0.00000	0.00000	0.00000	0.11559	0.00000	0.00000	0.00000	0.00000	0.00000
Shock	0.47961	0.50213	0.00000	0.73670	0.00000	0.28068	0.74786	0.55376	0.00000
Shock by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.50071	0.00000
Shock by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.36096	0.00000
Status urgent	0.25552	0.00000	0.19344	0.25714	0.00000	0.15548	0.19858	0.35184	−0.36106
Status emergent	1.32597	1.02109	1.13199	1.51294	0.00000	0.49075	1.17360	0.89480	−1.12373
Status salvage	2.07144	0.74530	1.24544	1.25342	0.00000	0.00000	1.21823	0.00000	−1.13785
Status by MVR function	−0.31729	0.00000	−0.34380	−0.35206	0.00000	0.00000	0.00000	0.00000	0.00000
Status by MVRepair function	0.84051	0.00000	−0.02373	0.29927	0.00000	0.00000	0.00000	0.00000	0.00000
Stenosis aortic	0.00000	0.00000	0.00000	−0.10782	0.00000	−0.10852	−0.07479	−0.15434	0.06873
Stenosis mitral	0.21309	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Unstable angina	0.18950	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Appendix Table 2. Definition of Variables Appearing in STS 2008 Valve Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
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
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 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 ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise

Appendix Table 2. Continued

Variable	Definition
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 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
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Stenosis mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina, no MI within 7 days of surgery, = 0 otherwise

Note: See www.sts.org for exact definitions of terms used above.

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; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PLOS = prolonged length of stay; Preop = preoperative; Reop = reoperation; Comp = composite adverse event (any); RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

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 #:** 0122 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Operative Mortality MV Replacement + CABG Surgery
De.2 **Brief description of measure:** Percent of patients undergoing combined MV Replacement and CABG who die, including both 1) all deaths occurring during the hospitalization in which the procedure was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure
1.1-2 **Type of Measure:** Outcome 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-634281980937555930.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:** Public reporting, Internal quality improvement	**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, Frequently performed procedure, Leading cause of morbidity/mortality, High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** The incidence of mitral valve incompetence as a result of or coexistent with coronary artery disease is increasing as a result of a progressively older population of patients coming to medical attention and recommendations for earlier surgical intervention among other factors. Patients undergoing combined coronary bypass and mitral valve replacement (CABG/MVR), in addition, increasingly have a larger number and severity of co-morbid risk factors. As a result, these patients have among one of the highest mortality rates for all surgical procedures. As a direct outcome measure, surgical mortality following CABG/MVR is a time-tested and well-accepted gold standard for the quality of this procedure. Given the demographics of heart disease as well as the increasing scientific evidence for earlier intervention, the impact of this measure is expected to increase considerably. 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892.	**1a** C☐ P☐ M☐ N☐

- Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137.
- Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467.
- Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009 Jul; 88(1 Suppl):S43-62.
- Miyata H, Motomura N, Tsukihara H, Takamoto S; Japan Cardiovascular Surgery Database. Risk models including high-risk cardiovascular procedures: clinical predictors of mortality and morbidity. Eur J Cardiothorac Surg. 2010 Nov 1
- Vassileva CM, Boley T, Markwell S, Hazelrigg S. Meta-analysis of short-term and long-term survival following repair versus replacement for ischemic mitral regurgitation. Eur J Cardiothorac Surg. 2010 Aug 18.
- Daneshmand MA, Milano CA, Rankin JS, Honeycutt EF, Shaw LK, Davis RD, Wolfe WG, Glower DD, Smith PK. Influence of patient age on procedural selection in mitral valve surgery. Ann Thorac Surg. 2010 Nov; 90(5):1479-85

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** The reporting of this measure, particularly in a risk-adjusted format, continues to be increasingly beneficial to individual providers and their affiliated healthcare organizations as they assess and compare their own results with those of similar organizations and with best practices nation-wide. Through public reporting of such a measure, the process of quality improvement becomes more transparent in order to reduce the existing variability and improve overall outcomes across all hospitals offering this procedure. In addition, this measure will anchor current outcomes as new technologies and treatment therapies are developed in the future. This includes such devices as endovascular valve procedures as well as newer, non-invasive treatments currently under development.

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, 2005-December 31, 2009

Analysis includes 33 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data (not restricted to this measure) to STS for at least 36 months in 2005-2009.

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

1b.5 **Citations for data on Disparities:**

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*)**:** Mortality following surgery is a direct outcome measure. Risk-adjustment methodology makes this measure more impactful from a quality improvement perspective.

1c.2-3. **Type of Evidence:** Observational study, Expert opinion, Systematic synthesis of research, Other
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*)**:**
Surgery for ischemic heart disease and for mitral valvar insufficiency have evolved over the last several decades and currently represents the best form of treatment for these conditions depending on the stage of

1c
C☐
P☐
M☐
N☐

progression. Initial outcomes, including mortality, have substantially been reduced over this time period through a number of technical innovations, refined selection criteria, and improved preoperative management. While risk-adjusted mortality with CABG/MVR is currently low and declining, ongoing modifications in technique and improved patient selection are necessary. Furthermore, persistent variability in outcomes between providers, even when risk-adjusted, indicates the presence of considerable opportunity for improvement.

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 NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437.

- Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892.
- Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137.
- Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467.
- Vassileva CM, Boley T, Markwell S, Hazelrigg S. Meta-analysis of short-term and long-term survival following repair versus replacement for ischemic mitral regurgitation. Eur J Cardiothorac Surg. 2010 Aug 18.
- Daneshmand MA, Milano CA, Rankin JS, Honeycutt EF, Shaw LK, Davis RD, Wolfe WG, Glower DD, Smith PK. Influence of patient age on procedural selection in mitral valve surgery. Ann Thorac Surg. 2010 Nov; 90(5):1479-85
- Miyata H, Motomura N, Tsukihara H, Takamoto S; Japan Cardiovascular Surgery Database. Risk models including high-risk cardiovascular procedures: clinical predictors of mortality and morbidity. Eur J Cardiothorac Surg. 2010 Nov 1.
- Murphy MO, Rao C, Punjabi PP, Athanasiou T. In patients undergoing mitral surgery for ischaemic mitral regurgitation is it preferable to repair or replace the mitral valve? Interact Cardiovasc Thorac Surg. 2010 Nov 18.
- Bouma W, van der Horst IC, Wijdh-den Hamer IJ, Erasmus ME, Zijlstra F, Mariani MA, Ebels T. Chronic ischaemic mitral regurgitation. Current treatment results and new mechanism-based surgical approaches. Eur J Cardiothorac Surg. 2010 Jan; 37(1):170-85.
- Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009 Jul; 88(1 Suppl):S43-62.
- Sajja LR, Mannam G, Dandu BR, Pathuri S, Sompalli S, Anjaneyulu A. Outcomes of mitral valve repair for chronic ischemic mitral regurgitation. Asian Cardiovasc Thorac Ann. 2009 Jan; 17(1):29-34.

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

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

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

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

1c.14 Rationale for using this guideline over others:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
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*): Number of patients undergoing combined MV Replacement and CABG who die, including both 1) all deaths occurring during the hospitalization in which the operation was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure **2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*):** During hospitalization regardless of length of stay or within 30 days of surgery if discharged **2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):** Number of MV Replacement + CABG procedures with an operative mortality; Number of MV Replacement + CABG procedures in which Mortality [Mortalty (STS Adult Cardiac Surgery Database Version 2.73)] and Mortality Operative Death (MtOpD) are marked "yes." Operative mortality is further verified by the following variables: Mortality Status at 30 days (Mt30Stat), Mortality Date (MtDate), Mortality Discharge Status (MtDCStat)	
2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*): All patients undergoing combined MV Replacement + CABG **2a.5 Target population gender:** Female, Male **2a.6 Target population age range:** 18 and older **2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*):** 60 months **2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):** Number of MV Replacement + CABG procedures; MV Replacement + CABG is determined as a procedure for which all of the following apply: - OpCAB is marked as "Yes" - OpValve is marked "Yes" - VSMV is marked "Yes"	**2a-specs** **C☐** **P☐** **M☐** **N☐**

- VSMVPr is marked “Replacement”
- (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
- VSAV, VSAVPr, ResectSubA, OpTricus, OpPulm, OpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc, OCTumor, OCPulThromDis, OCarOthr are all marked “no” or “missing”

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

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

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

2a.12-13 Risk Adjustment Type: Case-mix adjustment

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

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment 2a.15 Detailed Risk Model-634281986749363998.pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**

2a.22 Describe the method for discriminating performance *(e.g., significance testing)*:
Participant specific OR and their 95% CI were estimated in the hierarchical model. These model-based estimates were used to control variation due to random statistical fluctuations while estimating true signal variation. A 95% CI excluding zero indicates the participant’s performance is significantly lower or higher than an “average” STS participant.

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)*:

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: national, Population: regional/network, Population: states, Population: counties or cities

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)*: STS Adult Cardiac Surgery Database - Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009. Excluded from analysis are participants that did not submit results for both time periods. Because database participants can change their underlying care processes at any time, we would not expect perfect correlation between two sets of results from even proximate time periods.

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

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 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.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Mortality Operative Death: 100.0% agreement rate

2c
C☐
P☐
M☐
N☐

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
C☐
P☐
M☐
N☐
NA☐

2d.4 **Analytic Method** *(type analysis & rationale)*: 2d.5 **Testing Results** *(e.g., frequency, variability, sensitivity analyses)*:	
2e. Risk Adjustment for Outcomes/ Resource Use Measures 2e.1 **Data/sample** *(description of data/sample and size)*: Please see Risk Adjustment Type section above 2e.2 **Analytic Method** *(type of risk adjustment, analysis, & rationale)*: Detailed information regarding the risk adjustment model can be found in the attachment: Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S43-62. 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)*: 33 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009; January 1, 2005-December 31, 2009 2f.2 **Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: We calculated the risk adjusted event rate with the participant's Odds Ratio (OR) estimate and the overall STS event rate. Therefore, the risk adjusted rate is closely related to OR estimate. If OR > 1, then the participant's risk adjusted rate will be greater than the overall STS event rate; if OR < 1, then the participant's risk adjusted rate will be smaller than the overall STS event rate. The statistical significance is defined by the 95% confidence interval (CI) or the OR estimate. If the 95% CI for a participant's OR includes the null value 1.0, then we cannot distinguish this participant's performance from the STS average – either the participant's performance was close to average or else the participant's sample size was too small to make a reliable inference. Otherwise, if the 95% CI falls to the right of 1.0, then the participant's performance is considered significantly lower than the average STS results; if the 95% CI falls to the left of 1.0, then the participant's performance is considered significantly higher than the average STS results. 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	**2f** C☐ P☐ M☐ N☐
2g. Comparability of Multiple Data Sources/Methods 2g.1 **Data/sample** *(description of data/sample and size)*: 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)*:	**2h** C☐ P☐

2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:	**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 STS plans to publicly report more measures in the future. There is no definite date yet assigned to this measure; however, STS staff and surgeon leadership have engaged in initial internal STS discussions regarding this matter. **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***):** **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	**3a** **C☐ P☐ M☐ N☐**
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	**3b** **C☐**

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	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*)	**4b** C☐ P☐ M☐

Yes **4b.2 If not, specify the near-term path to achieve electronic capture by most providers.**	**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. Please see IFMC audit results.	**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	**4e** **C☐** **P☐** **M☐** **N☐**

members are charged the lesser of the two fees. 4e.3 **Evidence for costs:** 4e.4 **Business case documentation:**	
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 0122 Sections 2a.14, 1b.2, 2b.3, 2f.3, 3a.6.pdf
Date of Submission (*MM*/*DD*/*YY*)**:** 01/12/2011

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.

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

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 valve procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest four and a half years of data (January 2006 – 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 Valve plus CABG Models

Variable	Definition
Intercept	= 1 for all patients
Preoperative AFib	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine by MVR function	= Creatinine function 1 if valve operation is MVR, = 0 otherwise
Creatinine by MVRepair function	= Creatinine function 1 if valve operation is MVRepair, = 0 otherwise
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
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
EF by MVR function	= Ejection fraction function if valve operation is MVR, = 0 otherwise

EF by MVRepair function	= Ejection fraction function if valve operation is MVRepair, = 0 otherwise
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Endocarditis by MVR function	= 1 if patient has active endocarditis and valve operation is MVR, = 0 otherwise
Endocarditis by MVRepair function	= 1 if patient has active endocarditis and valve operation is MVRepair, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP or inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP or inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient has received immunosuppressive therapy within 30 days, = 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 1–21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI ≤ 21 days [a]	= 1 if patient has history of MI within 21 days prior to surgery, = 0 otherwise (for CVA and PLOS; coded as < 24 hours and 1–21 days for others)
MI < 24 hours	= 1 if history of MI < 24 hours prior to surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased coronary vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
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
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not rescuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis, mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina and no MI within 7 days of surgery, = 0 otherwise

[a] **MI coded as < 24 hours and 1 to 21 days.**

BSA = body surface area; CABG = coronary artery bypass graft surgery; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease;

The final patient-level model was built by step-wise selection method with several variables decided by surgeon panel 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	*Risk-Adjusted Operative Mortality for MV Replacement + CABG Surgery*
N	33
Mean	1.0
1st	1.6
5th	1.5
10th	1.3
25th	1.1
Median	1.0
75th	0.8
90th	0.7
95th	0.6
99th	0.6
Outlier	0 (0.0)
High	0
Low	0

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).

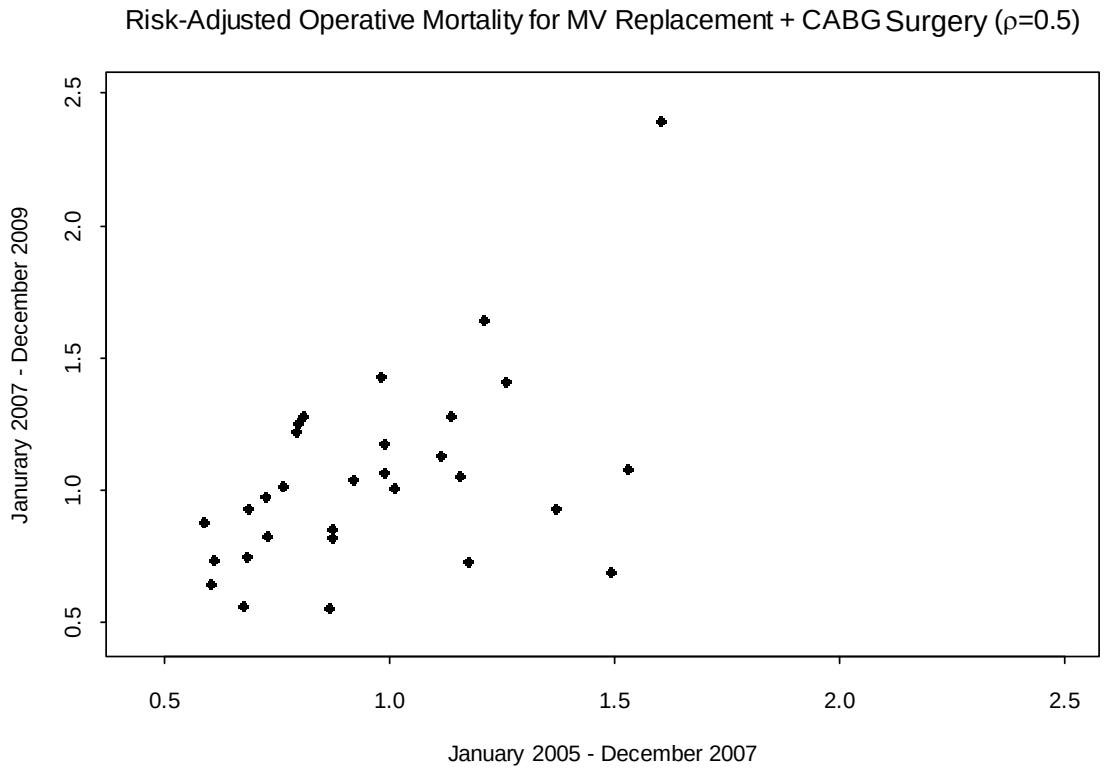
Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for MV Replacement + CABG Surgery*
N	33

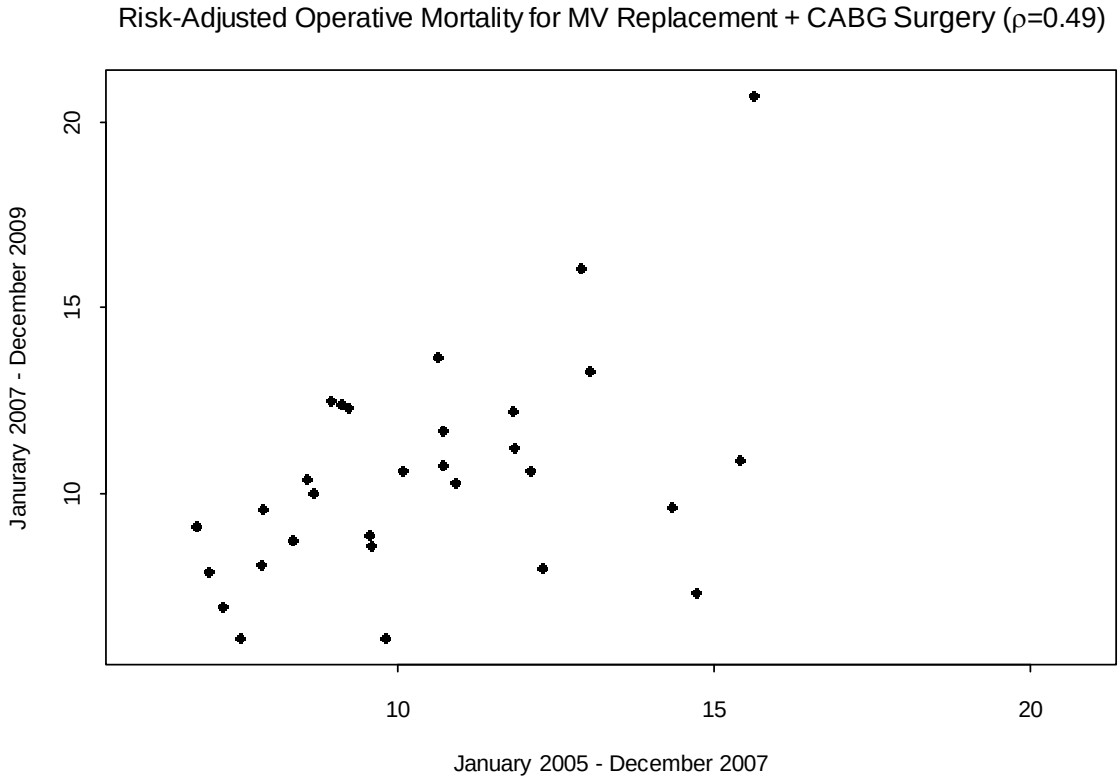
Measurement	*Risk-Adjusted Operative Mortality for MV Replacement + CABG Surgery*
Mean	10.3
1st	6.5
5th	7.0
10th	7.8
25th	8.6
Median	10.4
75th	11.3
90th	13.0
95th	15.0
99th	15.2
Outlier	0 (0.0)
High	0
Low	0

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

Testing results: $\rho = 0.50$



Risk adjusted Rate:
Testing results: $\rho = 0.49$



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, 2005-December 31, 2009. Sample contains 33 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009.

Measurement	*Risk-Adjusted Operative Mortality for MV Replacement + CABG Surgery*
N	33
Mean	1.0
1st	1.6
5th	1.5
10th	1.3
25th	1.1
Median	1.0
75th	0.8
90th	0.7
95th	0.6
99th	0.6
Outlier†	0 (0.0)
High	0
Low	0

Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for MV Replacement + CABG Surgery*
N	33
Mean	10.3
1st	6.5
5th	7.0
10th	7.8
25th	8.6
Median	10.4
75th	11.3
90th	13.0
95th	15.0
99th	15.2
Outlier†	0 (0.0)
High	0

Measurement	*Risk-Adjusted Operative Mortality for MV Replacement + CABG Surgery*
Low	0

†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."

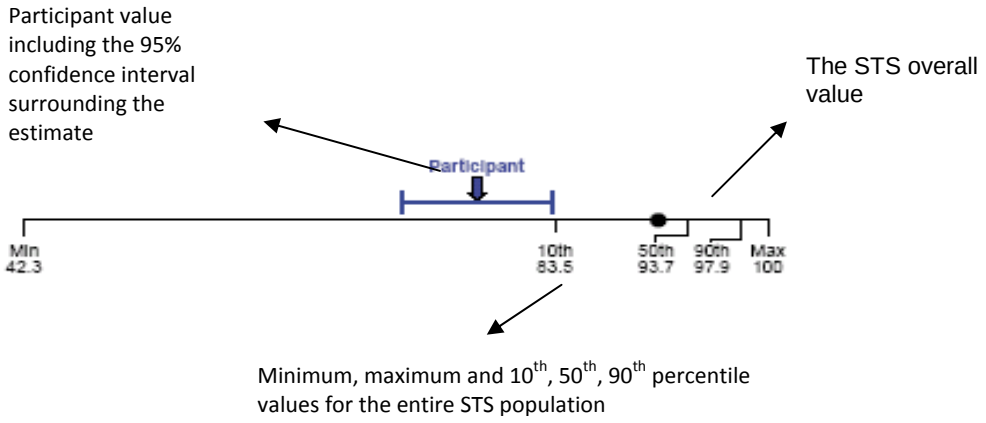
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

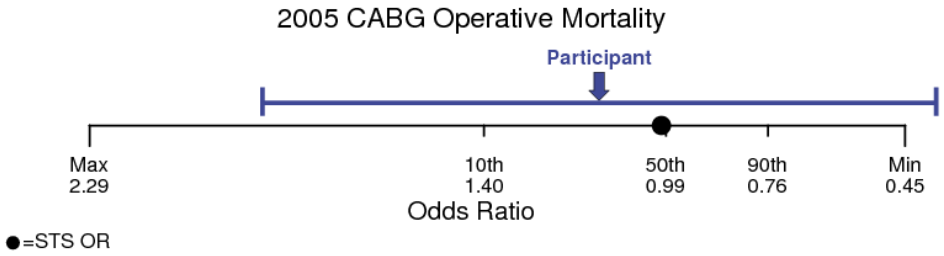
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

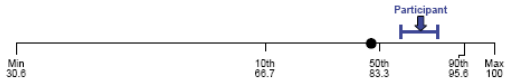
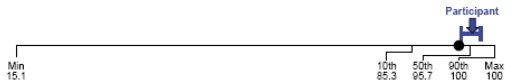
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 3—Valve Plus Coronary Artery Bypass Grafting Surgery

David M. Shahian, MD,[a] Sean M. O'Brien, PhD,[b] Giovanni Filardo, PhD, MPH,[c] Victor A. Ferraris, MD,[d] Constance K. Haan, MD,[e] Jeffrey B. Rich, MD,[f] Sharon-Lise T. Normand, PhD,[g] Elizabeth R. DeLong, PhD,[b] Cynthia M. Shewan, PhD,[h] Rachel S. Dokholyan, MPH,[b] Eric D. Peterson, MD, MPH,[b] Fred H. Edwards, MD,[e] and Richard P. Anderson, MD[i†]

[a]Massachusetts General Hospital, Boston, Massachusetts; [b]Duke Clinical Research Institute, Durham, North Carolina; [c]Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; [d]University of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; [e]University of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; [f]Sentara Cardiovascular Research Institute, Norfolk, Virginia; [g]Department of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [h]The Society of Thoracic Surgeons, Chicago, Illinois; and [i]Seattle, Washington

Background. Since 1999, The Society of Thoracic Surgeons (STS) has published two risk models that can be used to adjust the results of valve surgery combined with coronary artery bypass graft surgery (CABG). The most recent was developed from data for patients who had surgery between 1994 and 1997 using operative mortality as the only endpoint. Furthermore, this model did not specifically consider mitral valve repair plus CABG, an increasingly common procedure. Consistent with STS policy of periodically updating and improving its risk models, new models for valve surgery combined with CABG have been developed. These models specifically address both perioperative morbidity and mitral valve repair, and they are based on contemporary data.

Methods. The final study population consisted of 101,661 procedures, including aortic valve replacement (AVR) plus CABG, mitral valve replacement (MVR) plus CABG, or mitral valve repair (MVRepair) plus CABG between January 1, 2002, and December 31, 2006. Model outcomes included operative mortality, stroke, deep sternal wound infection, reoperation, prolonged ventilation, renal failure, composite major morbidity or mortality, prolonged postoperative length of stay, and short postoperative length of stay. Candidate variables were screened for frequency of missing data, and imputation techniques were used where appropriate. Stepwise variable selection was employed, supplemented by advice from an expert panel of cardiac surgeons and biostatisticians. Several variables were forced into models to insure face validity (eg, atrial fibrillation for the permanent stroke model, sex for all models). Based on preliminary analyses of the data, a single model was employed for valve plus CABG, with indicator variables for the specific type of procedure. Interaction terms were included to allow for differential impact of predictor variables depending on procedure type. After validating the model in the 40% validation sample, the development and validation samples were then combined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regression model was estimated using generalized estimating equations to account for clustering of patients within institutions.

Results. The c-index for mortality prediction for the overall valve plus CABG population was 0.75. Morbidity model c-indices for specific complications (permanent stroke, renal failure, prolonged ventilation > 24 hours, deep sternal wound infection, reoperation for any reason, major morbidity or mortality composite, and prolonged postoperative length of stay) for the overall group of valve plus CABG procedures ranged from 0.622 to 0.724, and calibration was excellent.

Conclusions. New STS risk models have been developed for heart valve surgery combined with CABG. These are the first valve plus CABG models that also include risk prediction for individual major morbidities, composite major morbidity or mortality, and short and prolonged length of stay.

(Ann Thorac Surg 2009;88:S43–62)

Risk models for cardiac surgery were first developed almost 2 decades ago, and most of these early models focused on isolated coronary artery bypass graft surgery (CABG) [1–4]. The results of this frequently performed surgical procedure have often been used as the sole marker to assess the quality of care delivered by cardiac surgical programs. Risk-adjusted results for CABG have been used

Drs Shahian, O'Brien, Filardo, Ferraris, Haan, Rich, Normand, DeLong, Shewan, Peterson, Edwards, Anderson, and Ms Dokholyan, have no conflicts of interest to declare regarding this work.

†This author is deceased. Former Chair, Quality, Research and Patient Safety Council, The Society of Thoracic Surgeons, Chicago, IL.

Address correspondence to Dr Shahian, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.055

Abbreviations and Acronyms

AVR	= aortic valve replacement
CABG	= coronary artery bypass graft surgery
MI	= myocardial infarction
MVR	= mitral valve replacement
MVRepair	= mitral valve repair
NCD	= National Adult Cardiac Surgery Database
QMTF	= Quality Measurement Task Force
STS	= The Society of Thoracic Surgeons

for hospital and regional quality improvement initiatives, public reporting, pay for performance reimbursement programs, decision support, patient counseling, and clinical research. Earlier models focused primarily on mortality prediction, but subsequent models have been developed for both risk-adjusted morbidity and length of stay [5].

The other commonly performed category of cardiac surgery consists of operations on the heart valves, either alone or in combination with CABG. Relative to isolated CABG procedures, which are declining in frequency, the proportion of valve cases is steadily increasing. To better assess the overall performance of cardiac surgery programs, to discern the factors that are most significantly related to patient outcomes, and to aid in physician and patient decision-making, risk models have now also been developed for heart valve surgery [6–18].

Unlike risk models for isolated CABG, a relatively standardized procedure, valve surgery encompasses a much more diverse group of operations. There are four cardiac valves, and they may malfunction in a number of quite different ways (eg, stenosis, regurgitation, infection, and so forth). The valves may be repaired or replaced with a wide range of techniques and prosthetics. In some cases, procedures may be performed on multiple valves, or the valve procedure may be combined with CABG.

Given the heterogeneity of heart valve surgery, it is not surprising that a variety of risk-modeling techniques has been applied. At one extreme, the European System for Cardiac Operative Risk Evaluation (EuroSCORE) algorithm, developed by a European consortium, groups all cardiac operations together in a single risk model with indicator variables included to account for valve procedures [14, 18]. Although this approach is simple and easy to apply, recent studies by van Gameren and associates [19] have suggested that a dedicated valve risk model may have better discrimination and calibration than the EuroSCORE algorithm when applied to valve surgery patients. Combined models for aortic and mitral valve procedures with or without CABG have been developed by Jin and colleagues [12] and by Ambler and associates [13]. The 2001 valve models developed by The Society of Thoracic Surgeons (STS) [6] consisted of one model for all isolated valve procedures and one model for valve procedures combined with CABG, and a 2007 risk model derived from the New York Cardiac Surgery Reporting System used a similar stratification [8].

Unified valve models reflect the fact that many risk factors are common to both aortic and mitral valve surgery. They offer simplicity, and they also permit larger sample sizes for development and validation [12]. However, there are significant differences between aortic and mitral valvular disease in both pathophysiology and outcomes, and both also differ substantially from isolated CABG [11]. Some investigators advocate separate aortic and mitral valve models to have more homogeneous patient populations. Examples include models developed by STS, the New York Cardiac Surgery Reporting System, and the Northern New England Cardiovascular Disease Study Group [7, 9, 10]. Some of these models have been developed solely for isolated valve replacement, some have included CABG as a separate predictor variable in the isolated valve model, and some models have focused specifically on valve plus CABG. All these decisions involve a tradeoff—the more homogeneous the study group, the fewer patients are available for model development and validation [12].

Because of the large number of valve surgery patients available for analysis in the STS National Adult Cardiac Surgery Database (NCD), our approach has favored separate models for valve plus CABG versus isolated valve surgery. The STS Quality Measurement Task Force (QMTF) presumes that when adequate numbers of patients are available for study, relatively homogeneous operative categories result in more accurate risk prediction. Furthermore, recent studies by van Gameren and colleagues [19] suggest that the valve plus CABG group may be the most difficult to model accurately, thus meriting its own algorithm.

Several new features were added to the 2008 valve plus CABG models described in this report. First, recognizing that mitral valve repair is often different in both etiology and outcomes than replacement, the QMTF has included interactions between surgery type and several key predictor variables. Fitting a single model with several such interactions is useful. It allows for pooling information across related groups of valve procedures without making an a priori assumption that the effect of key risk factors is constant across these groups. Finally, new models have been developed for specific major complications of each valve plus CABG procedure, as well as for composite morbidity, mortality, and for both short and prolonged postoperative length of stay.

The authors of this report are members of the STS QMTF who were involved in this risk model development project.

Study Population and Endpoints

Our general approaches to variable selection and risk model development have been described in the companion articles on isolated CABG (Part 1) and isolated valve surgery (Part 2). Details specific to the valve plus CABG models are included in this report.

Study Population

The study population for this analysis consisted of single aortic or mitral valve surgical procedures combined with

CABG performed on adult patients between January 1, 2002, and December 31, 2006. Only the following procedures were included: (1) isolated aortic valve replacement (AVR) plus CABG; (2) isolated mitral valve replacement (MVR) plus CABG; and (3) isolated mitral valve repair (MVRepair) plus CABG.

Because of the relatively small number of pulmonic, tricuspid, multiple valve procedures, and aortic repairs, these cases were not included in the current models. Patients undergoing isolated valve surgery without CABG were excluded from the current analysis, but these cases are the focus of a separate model described in Part 2 of this three-part series. Patients with missing sex data (n = 17) were excluded because these patients are not allowed in the analysis dataset used for creating STS database participant feedback reports. Patients on dialysis preoperatively (n = 2,443) were excluded when developing the risk model for prediction of postoperative renal failure. The final study population comprised 101,661 patient operations (66,074 AVR plus CABG; 13,663 MVR plus CABG; and 21,924 MVRepair plus CABG) from 814 STS NCD participating groups.

Characteristics of the study population are summarized in Table 1.

Training and Validation Samples

The study population was randomly divided into a 60% training (development) sample and a 40% test (validation) sample. The development sample was used to identify predictor variables and estimate model coefficients. Data from the validation sample were used to assess model fit, discrimination, and calibration. After choosing variables and assessing model fit, the development and validation samples were subsequently combined, and the final model coefficients were estimated using the combined (development plus validation) data.

Endpoints

In developing the valve plus CABG risk models, we used the same nine endpoints that were analyzed in the STS isolated CABG (Part 1) and the STS isolated valve (Part 2) models. Morbidities in all three models are recorded only in-hospital, in contrast to the operative mortality endpoint defined below (although beginning with version 2.61, sternal infection will be recorded at 30 days): (1) operative mortality: death during the same hospitalization as surgery, regardless of timing or within 30 days of surgery regardless of venue; (2) permanent stroke (CVA): a central neurologic deficit persisting longer than 72 hours; (3) renal failure: a new requirement for dialysis or an increase of the serum creatinine to more than 2.0 mg/dL and double the most recent preoperative creatinine level; (4) prolonged ventilation (> 24 hours); (5) deep sternal wound infection; (6) reoperation for any reason; (7) major morbidity or mortality, a composite defined as the occurrence of any of the above endpoints; (8) prolonged postoperative length of stay (PLOS): length of stay (LOS) more than 14 days (alive or dead); and (4) short postoperative length of stay (SLOS): LOS less than 6 days and patient alive at discharge.

Endpoint frequencies in the study population are presented in Table 2.

Separate Versus Combined Models

Given the variety of approaches used in previous models by STS and other developers, we investigated the option of developing separate models for the AVR plus CABG and MVR plus CABG populations, and we also studied how best to subdivide the mitral plus CABG population into repair versus replacement. Although we had a large study population available, many of the individual outcomes were relatively rare. We were concerned that the number of events would be too small to permit reliable estimation of the model coefficients in separate models for each valve. Thus, in theory, the development of separate custom models for each valve type could be inferior to a single combined model because the custom models would have a smaller sample size and hence larger variance.

As described in detail in Part 2 of this series (isolated valve surgery), we performed preliminary empirical analyses to compare two alternative strategies (separate versus combined AVR plus CABG and MVR/Repair plus CABG) for developing these risk models. We first developed separate models for the three subpopulations (AVR plus CABG, MVR plus CABG, and MVRepair plus CABG), then modeled all three subpopulations together in a single model. In the latter approach, we included several interaction terms to allow the effect of certain risk factors to differ across the specific valve subpopulations. These strategies were used to develop risk models for operative mortality and permanent stroke, using a 60% development sample and a separate 40% validation sample. The performance of the combined model was then assessed separately within each subpopulation and compared to the model that was developed specifically for that subpopulation. In the case of mortality, the combined model had better discrimination (larger c-index) than the corresponding custom model in each of the three subpopulations (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). For stroke, the combined model had better discrimination in two of the three populations (all except AVR plus CABG). Finally, when explained variation was quantified by the generalized R^2 index of Nagelkerke [20], the combined model had greater explained variation than the custom model in each subpopulation for each endpoint. These results provide empirical support for the use of a single model with several interactions, which allows pooling of information across valve groups without assuming that the effect of risk factors is constant.

Selection of Candidate Predictor Variables

The candidate variables for the STS valve plus CABG models were identical to those in the STS isolated valve models, described in Part 2 of this series. They differed from the isolated CABG model variables in the following specific areas: (1) Percutaneous coronary intervention (PCI) occurring 6 hours or less before surgery was present in only 315 patients (0.3%) in the valve plus CABG study population, and was not included as a candidate variable. (2) Infectious endocarditis was not included in the isolated CABG model but was considered for the valve plus CABG model. Although this risk factor was rarely present (0.8% active

Table 1. Distribution of Risk Factors in Overall Study Population 2002 to 2006

	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
Variable	N	%	N	%	N	%	N	%
Demographics								
Age, years								
< 55	6,693	6.6	2,983	4.51	1,309	9.58	2,401	10.95
55–64	17,188	16.9	9,132	13.82	2,790	20.42	5,266	24.02
65–74	33,628	33.1	21,313	32.26	4,667	34.16	7,648	34.88
≥ 75	44,152	43.4	32,646	49.41	4,897	35.84	6,609	30.15
Sex								
Male	65,588	64.5	44,619	67.53	7,348	53.78	13,621	62.13
Female	36,073	35.5	21,455	32.47	6,315	46.22	8,303	37.87
Race								
Caucasian	90,572	89.1	60,121	90.99	11,765	86.11	18,686	85.23
Black	4,534	4.5	2,094	3.17	914	6.69	1,526	6.96
Hispanic	2,487	2.4	1,487	2.25	354	2.59	646	2.95
Asian	1,083	1.1	542	0.82	191	1.40	350	1.60
Other	2,295	2.3	1,402	2.12	331	2.42	562	2.56
Missing	690	0.7	428	0.65	108	0.79	154	0.70
Risk factors								
Body surface area, m^2								
< 1.50	3,340	3.3	1,985	3.00	638	4.67	717	3.27
1.50–1.74	20,779	20.4	12,580	19.04	3,500	25.62	4,699	21.43
1.75–1.99	40,017	39.4	25,814	39.07	5,440	39.82	8,763	39.97
≥ 2.00	36,956	36.4	25,361	38.38	3,996	29.25	7,599	34.66
Missing	569	0.6	334	0.51	89	0.65	146	0.67
Body mass index, kg/m^2								
< 25	29,353	28.9	17,712	26.81	4,787	35.04	6,854	31.26
25–29	39,345	38.7	25,692	38.88	4,951	36.24	8,702	39.69
30–34	21,063	20.7	14,447	21.86	2,507	18.35	4,109	18.74
≥ 35	11,165	11.0	7,785	11.78	1,299	9.51	2,081	9.49
Missing	735	0.7	438	0.66	119	0.87	178	0.81
Diabetes mellitus								
No diabetes	68,112	67.0	44,489	67.33	9,517	69.66	14,106	64.34
Diabetes, noninsulin	23,383	23.0	15,705	23.77	2,642	19.34	5,036	22.97
Diabetes, insulin	9,848	9.7	5,677	8.59	1,463	10.71	2,708	12.35
Diabetes, missing treatment	167	0.2	105	0.16	20	0.15	42	0.19
Missing	151	0.1	98	0.15	21	0.15	32	0.15
Hypertension								
No	22,709	22.3	13,944	21.10	3,482	25.48	5,283	24.10
Yes	78,823	77.5	52,050	78.78	10,163	74.38	16,610	75.76
Missing	129	0.1	80	0.12	18	0.13	31	0.14
Hypercholesterolemia								
No	33,759	33.2	21,248	32.16	5,324	38.97	7,187	32.78
Yes	67,613	66.5	44,649	67.57	8,280	60.60	14,684	66.98
Missing	289	0.3	177	0.27	59	0.43	53	0.24
Past or present smoker								
No	43,687	43.0	29,123	44.08	5,835	42.71	8,729	39.81
Yes	57,813	56.9	36,849	55.77	7,797	57.07	13,167	60.06
Missing	161	0.2	102	0.15	31	0.23	28	0.13
Chronic lung disease								
None	76,803	75.5	50,632	76.63	9,756	71.40	16,415	74.87
Mild	12,157	12.0	7,658	11.59	1,853	13.56	2,646	12.07
Moderate	7,797	7.7	4,720	7.14	1,269	9.29	1,808	8.25
Severe	4,005	3.9	2,463	3.73	658	4.82	884	4.03
Missing	899	0.9	601	0.91	127	0.93	171	0.78

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Peripheral vascular disease								
No	84,183	82.8	54,658	82.72	11,373	83.24	18,152	82.80
Yes	17,294	17.0	11,296	17.10	2,267	16.59	3,731	17.02
Missing	184	0.2	120	0.18	23	0.17	41	0.19
Cerebrovascular disease								
No	83,284	81.9	53,509	80.98	11,304	82.73	18,471	84.25
Yes	18,202	17.9	12,449	18.84	2,335	17.09	3,418	15.59
Missing	175	0.2	116	0.18	24	0.18	35	0.16
CVA								
No CVA	92,527	91.0	60,141	91.02	12,283	89.90	20,103	91.69
Remote CVA (> 2 weeks)	8,461	8.3	5,545	8.39	1,240	9.08	1,676	7.64
Recent CVA (≤ 2 weeks)	348	0.3	184	0.28	88	0.64	76	0.35
CVA, missing timing	114	0.1	62	0.09	23	0.17	29	0.13
Missing	211	0.2	142	0.21	29	0.21	40	0.18
Endocarditis								
No endocarditis	99,517	97.9	65,023	98.41	12,914	94.52	21,580	98.43
Treated endocarditis	1,091	1.1	525	0.79	356	2.61	210	0.96
Active endocarditis	827	0.8	387	0.59	356	2.61	84	0.38
Endocarditis, missing type	24	0.0	11	0.02	8	0.06	5	0.02
Missing	202	0.2	128	0.19	29	0.21	45	0.21
Renal failure								
No	92,592	91.1	60,880	92.14	12,037	88.10	19,675	89.74
Yes	8,888	8.7	5,072	7.68	1,605	11.75	2,211	10.08
Missing	181	0.2	122	0.18	21	0.15	38	0.17
Renal function								
Creatinine < 1.0 mg/dL	30,178	29.7	20,297	30.72	3,672	26.88	6,209	28.32
Creatinine 1.00–1.49 mg/dL	52,008	51.2	34,054	51.54	6,758	49.46	11,196	51.07
Creatinine 1.50–1.99 mg/dL	11,469	11.3	7,151	10.82	1,732	12.68	2,586	11.80
Creatinine 2.00–2.49 mg/dL	2,711	2.7	1,554	2.35	498	3.64	659	3.01
Creatinine ≥ 2.5 mg/dL	1,602	1.6	844	1.28	319	2.33	439	2.00
Dialysis	2,443	2.4	1,364	2.06	482	3.53	597	2.72
Missing	1,250	1.2	810	1.23	202	1.48	238	1.09
Immunosuppressive treatment								
No	98,421	96.8	63,984	96.84	13,211	96.69	21,226	96.82
Yes	2,975	2.9	1,904	2.88	427	3.13	644	2.94
Missing	265	0.3	186	0.28	25	0.18	54	0.25
Previous CV interventions								
Previous CABG surgery								
No	91,657	90.2	59,583	90.18	12,057	88.25	20,017	91.30
Yes	9,615	9.5	6,257	9.47	1,540	11.27	1,818	8.29
Missing	389	0.4	234	0.35	66	0.48	89	0.41
Previous valve surgery								
No	98,737	97.1	64,265	97.26	12,794	93.64	21,678	98.88
Yes	2,540	2.5	1,567	2.37	813	5.95	160	0.73
Missing	384	0.4	242	0.37	56	0.41	86	0.39
Previous other cardiac surgery								
No	98,538	96.9	64,166	97.11	13,181	96.47	21,191	96.66
Yes	2,683	2.6	1,634	2.47	407	2.98	642	2.93
Missing	440	0.4	274	0.41	75	0.55	91	0.42
Number of previous CV surgeries								
No previous CV surgery	89,419	88.0	58,161	88.02	11,530	84.39	19,728	89.98
1 prior CV surgery	10,453	10.3	6,796	10.29	1,799	13.17	1,858	8.47
≥ 2 prior CV surgeries	1,200	1.2	766	1.16	231	1.69	203	0.93
Missing	589	0.6	351	0.53	103	0.75	135	0.62

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Prior PCI								
No PCI	84,553	83.2	55,581	84.12	11,152	81.62	17,820	81.28
PCI ≤ 6 hours	315	0.3	151	0.23	89	0.65	75	0.34
PCI > 6 hours	16,158	15.9	9,946	15.05	2,321	16.99	3,891	17.75
PCI, missing timing	234	0.2	145	0.22	45	0.33	44	0.20
Missing	401	0.4	251	0.38	56	0.41	94	0.43
Preoperative cardiac status								
Acuity status								
Elective	62,298	61.3	43,682	66.11	7,277	53.26	11,339	51.72
Urgent	36,454	35.9	21,414	32.41	5,315	38.90	9,725	44.36
Emergent	2,479	2.4	763	1.15	945	6.92	771	3.52
Emergent salvage	258	0.3	97	0.15	104	0.76	57	0.26
Missing	172	0.2	118	0.18	22	0.16	32	0.15
MI								
No prior MI	68,332	67.2	49,673	75.18	8,056	58.96	10,603	48.36
MI ≤ 21 days	16,934	16.7	9,308	14.09	2,621	19.18	5,005	22.83
MI 8–21 days	3,751	3.7	1,725	2.61	624	4.57	1,402	6.39
MI 1–7 days	10,458	10.3	4,514	6.83	1,741	12.74	4,203	19.17
MI > 6 and < 24 hours	1,113	1.1	367	0.56	341	2.50	405	1.85
MI ≤ 6 hours	531	0.5	178	0.27	192	1.41	161	0.73
MI, missing timing	355	0.3	184	0.28	59	0.43	112	0.51
Missing	187	0.2	125	0.19	29	0.21	33	0.15
Angina								
No	42,542	41.8	28,032	42.43	6,248	45.73	8,262	37.68
Yes	58,967	58.0	37,945	57.43	7,394	54.12	13,628	62.16
Missing	152	0.1	97	0.15	21	0.15	34	0.16
Cardiogenic shock								
No	98,743	97.1	65,219	98.71	12,590	92.15	20,934	95.48
Yes	2,719	2.7	720	1.09	1,055	7.72	944	4.31
Missing	199	0.2	135	0.20	18	0.13	46	0.21
Resuscitation								
No	100,474	98.8	65,522	99.16	13,359	97.78	21,593	98.49
Yes	971	1.0	405	0.61	281	2.06	285	1.30
Missing	216	0.2	147	0.22	23	0.17	46	0.21
Arrhythmia								
No arrhythmia	83,856	82.5	56,040	84.81	9,992	73.13	17,824	81.30
AFib/flutter	13,386	13.2	7,533	11.40	2,940	21.52	2,913	13.29
Heart block	1,975	1.9	1,311	1.98	289	2.12	375	1.71
Sustained VT/VF	1,513	1.5	614	0.93	299	2.19	600	2.74
Arrhythmia, other	483	0.5	305	0.46	63	0.46	115	0.52
Arrhythmia, missing type	242	0.2	135	0.20	59	0.43	48	0.22
Missing	206	0.2	136	0.21	21	0.15	49	0.22
Preoperative IABP								
No	96,136	94.6	64,597	97.76	11,957	87.51	19,582	89.32
Yes	5,205	5.1	1,275	1.93	1,655	12.11	2,275	10.38
Missing	320	0.3	202	0.31	51	0.37	67	0.31
NYHA class								
I	9,839	9.7	6,934	10.49	1,103	8.07	1,802	8.22
II	24,830	24.4	17,808	26.95	2,524	18.47	4,498	20.52
III	42,593	41.9	28,079	42.50	5,458	39.95	9,056	41.31
IV	20,571	20.2	10,808	16.36	3,882	28.41	5,881	26.82
Missing	3,828	3.8	2,445	3.70	696	5.09	687	3.13

Table 1. Continued

	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
Variable	N	%	N	%	N	%	N	%
Congestive heart failure								
No	58,086	57.1	41,984	63.54	5,797	42.43	10,305	47.00
Yes	43,377	42.7	23,953	36.25	7,845	57.42	11,579	52.81
Missing	198	0.2	137	0.21	21	0.15	40	0.18
Number of diseased coronary vessels								
None	2,362	2.3	1,786	2.70	281	2.06	295	1.35
One	22,718	22.3	16,934	25.63	3,040	22.25	2,744	12.52
Two	27,144	26.7	19,014	28.78	3,655	26.75	4,475	20.41
Three	49,060	48.3	28,107	42.54	6,623	48.47	14,330	65.36
Missing	377	0.4	233	0.35	64	0.47	80	0.36
Left main disease ≥ 50%								
No	84,025	82.7	55,292	83.68	11,503	84.19	17,230	78.59
Yes	17,175	16.9	10,512	15.91	2,072	15.17	4,591	20.94
Missing	461	0.5	270	0.41	88	0.64	103	0.47
Ejection fraction, %								
< 25	5,805	5.7	2,199	3.33	640	4.68	2,966	13.53
25–34	10,988	10.8	4,877	7.38	1,566	11.46	4,545	20.73
35–44	14,928	14.7	8,064	12.20	2,487	18.20	4,377	19.96
45–54	20,398	20.1	13,424	20.32	3,048	22.31	3,926	17.91
≥ 55	43,556	42.8	32,973	49.90	5,209	38.12	5,374	24.51
Missing	5,986	5.9	4,537	6.87	713	5.22	736	3.36
Aortic stenosis								
No	42,831	42.1	8,527	12.91	12,974	94.96	21,330	97.29
Yes	58,317	57.4	57,319	86.75	535	3.92	463	2.11
Missing	513	0.5	228	0.35	154	1.13	131	0.60
Mitral stenosis								
No	95,696	94.1	63,862	96.65	11,166	81.72	20,668	94.27
Yes	4,993	4.9	1,542	2.33	2,366	17.32	1,085	4.95
Missing	972	1.0	670	1.01	131	0.96	171	0.78
Tricuspid stenosis								
No	100,093	98.5	65,060	98.47	13,402	98.09	21,631	98.66
Yes	275	0.3	154	0.23	57	0.42	64	0.29
Missing	1,293	1.3	860	1.30	204	1.49	229	1.04
Pulmonic stenosis								
No	99,484	97.9	64,693	97.91	13,348	97.69	21,443	97.81
Yes	122	0.1	85	0.13	14	0.10	23	0.10
Missing	2,055	2.0	1,296	1.96	301	2.20	458	2.09
Aortic insufficiency								
None	57,561	56.6	28,972	43.85	10,821	79.20	17,768	81.04
Trivial	9,243	9.1	6,573	9.95	1,023	7.49	1,647	7.51
Mild	13,828	13.6	11,082	16.77	1,156	8.46	1,590	7.25
Moderate	10,195	10.0	9,581	14.50	232	1.70	382	1.74
Severe	8,686	8.5	8,580	12.99	49	0.36	57	0.26
Missing	2,148	2.1	1,286	1.95	382	2.80	480	2.19
Mitral insufficiency								
None	41,756	41.1	38,790	58.71	1,297	9.49	1,669	7.61
Trivial	7,467	7.3	7,139	10.80	147	1.08	181	0.83
Mild	15,407	15.2	13,485	20.41	584	4.27	1,338	6.10
Moderate	14,987	14.7	4,842	7.33	2,790	20.42	7,355	33.55
Severe	20,516	20.2	527	0.80	8,743	63.99	11,246	51.30
Missing	1,528	1.5	1,291	1.95	102	0.75	135	0.62

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Tricuspid insufficiency								
None	74,774	73.6	49,614	75.09	9,758	71.42	15,402	70.25
Trivial	7,972	7.8	5,454	8.25	839	6.14	1,679	7.66
Mild	11,505	11.3	7,060	10.68	1,631	11.94	2,814	12.84
Moderate	4,119	4.1	1,919	2.90	874	6.40	1,326	6.05
Severe	636	0.6	237	0.36	186	1.36	213	0.97
Missing	2,655	2.6	1,790	2.71	375	2.74	490	2.23
Pulmonic insufficiency								
None	91,715	90.2	59,891	90.64	12,275	89.84	19,549	89.17
Trivial	3,411	3.4	2,122	3.21	442	3.24	847	3.86
Mild	2,065	2.0	1,215	1.84	306	2.24	544	2.48
Moderate	326	0.3	165	0.25	70	0.51	91	0.42
Severe	49	0.0	25	0.04	11	0.08	13	0.06
Missing	4,095	4.0	2,656	4.02	559	4.09	880	4.01

AFib = atrial fibrillation; AVR = aortic valve replacement; CABG = coronary artery bypass graft; CV = cardiovascular; CVA = cerebrovascular accident (stroke); IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; VF = ventricular fibrillation; VT = ventricular tachycardia.

endocarditis) in the overall valve plus CABG population, it was included for consistency with the isolated valve model. Active endocarditis was present in 2.6% of patients undergoing mitral replacement plus CABG. (3) Mitral stenosis was rarely present among isolated CABG patients (0.35%). However, it was not uncommon (4.9%) among patients undergoing valve plus CABG surgery and was included as a candidate variable. It was present in 17.3% of mitral replacements and 5.0% of mitral repairs.

An indicator for valve procedure (AVR, MVR, MVRepair) was included in the combined valve plus CABG model, as previously noted.

Missing Data

Missing data are uncommon in the STS NCD, with a frequency of less than 1% missing for most variables. Model variables with more than 1% missing were ejection fraction

Table 2. Frequency of Endpoints in Overall Study Population 2002 to 2006

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall (AVR + CABG, MVR + CABG, MVRepair + CABG)									
N	101,661	101,661	99,218	101,661	101,661	101,661	101,661	101,661	101,661
Events	6,919	2,935	9,097	21,561	684	12,117	30,580	15,594	22,534
%	6.8	2.9	9.0	21.2	0.7	11.9	30.1	15.3	22.2
AVR + CABG									
N	66,074	66,074	64,710	66,074	66,074	66,074	66,074	66,074	66,074
Events	3,718	1,751	5,032	11,608	394	7,090	17,343	8,412	16,961
%	5.6	2.7	7.6	17.6	0.6	10.7	26.3	12.7	25.7
MVR + CABG									
N	13,663	13,663	13,181	13,663	13,663	13,663	13,663	13,663	13,663
Events	1,590	499	1,829	4,469	114	2,274	5,897	3,277	1,512
%	11.6	3.7	13.6	32.7	0.8	16.6	43.2	24.0	11.1
MVRepair + CABG									
N	21,924	21,924	21,327	21,924	21,924	21,924	21,924	21,924	21,924
Events	1,611	685	2,236	5,484	176	2,753	7,340	3,905	4,061
%	7.4	3.1	10.3	25.0	0.8	12.6	33.5	17.8	18.5

AVR = aortic valve replacement; CABG = coronary artery bypass graft surgery; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Table 3. List of Candidate Variables and Their Coding for STS Valve Plus CABG Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline truncated from below at 50 with knot at 75.
Ejection fraction	Linear; values > 50 mapped to 50
Body surface area[a]	Quadratic polynomial modeled separately for males and females. *Note: BSA < 1.4 and > 2.6 were mapped to those values, respectively.*
Creatinine	Linear (only for patients not on dialysis). *Note: Creatinine < 0.5 and > 5.0 mapped to those values, respectively.*
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval. Modeled as linear across the categories.
Binary variables	
Active infectious endocarditis	Yes/no
Dialysis	Yes/no
Preoperative atrial fibrillation	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Preop IABP or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no MI < 7 days)	Yes/no
Left main disease	Yes/no
Aortic stenosis	Yes/no
Mitral stenosis	Yes/no
Aortic insufficiency	Defined as at least moderate (yes/no)
Mitral insufficiency	Defined as at least moderate (yes/no)
Tricuspid insufficiency	Defined as at least moderate (yes/no)
Categorical variables	
Surgery type	3 groups: AVR + CABG, MVR + CABG, MVRepair + CABG
Chronic lung disease	Modeled as linear across categories (none, mild, moderate, severe)
CVD/CVA	3 groups: no CVD, CVD no CVA, CVD + CVA
Diabetes mellitus	3 groups: insulin diabetes, noninsulin diabetes, other or no diabetes
No. diseased coronary vessels	3 groups: < 2-vessel disease; 2-vessel disease; 3-vessel disease. Modeled as linear across the categories
MI	3 groups: < 24 hours, 1–21 days, > 21 days or no MI. *Note: groups 1 and 2 were subsequently collapsed for some models.*
Race	3 groups: black, Hispanic, other including Caucasian
Status	4 groups: elective, urgent, emergent no resuscitation, salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, ≥ 2 previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF and NYHA IV
Interaction terms	
Age by reoperation[a]	
Age by emergent status[a]	
Surgery type by each of the following:	Age, diabetes, dialysis, creatinine, reoperation, endocarditis, emergent status, CLD, CHF, EF, sex, shock, IABP/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, aortic stenosis.

[a] These variables were forced into each model.

AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association.

(5.9%), New York Heart Association functional class (3.8%), tricuspid insufficiency (2.6%), aortic insufficiency (2.1%), mitral insufficiency (1.5%), and creatinine/dialysis (1.2%).

To make full use of the available data, binary risk factors were modeled as yes versus no or missing. Thus, missing values were analyzed as if the endpoint did not occur. Missing data on categorical variables were imputed to the lowest risk value, which, in most instances, was the mode. Missing data on continuous variables were imputed to the conditional median. For ejection fraction, we conditioned

on congestive heart failure and sex. For body surface area, we conditioned on sex. For serum creatinine, we conditioned on renal failure.

Although multiple imputation is generally preferred on statistical grounds [21], we chose single imputation for this analysis based largely on practical considerations, including computational intensity. Furthermore, the fraction of missing data was small, and single and multiple imputation would give similar results. Finally, multiple imputation is primarily used for calculating appropriate standard error estimates, but an adjustment to the standard errors would not impact our study results or the published risk algorithms. In a separate sensitivity analysis, we compared predicted risk estimates from our final models to risk estimates that were derived from analogous models using multiple instead of single imputation. For each endpoint, the relative difference in predicted risk was less than 6% (eg, an absolute difference of 5.0% versus 5.3%) for all patients in the development and validation samples, and it was less than 2% (eg, an absolute difference of 5.0% versus 5.1%) for 99% of patients. A summary of these analyses including regression coefficients and covariance matrices is available at www.sts.org/riskmodels.

Final Variable Selection Procedure

Variables were initially selected using an automated stepwise model selection algorithm. The stepwise procedure began with a model that included all of the candidate variables except for interaction terms. Age, body surface area, and month of surgery were forced into each model. As in the isolated CABG and isolated valve models described in Parts 1 and 2 of this series, month of surgery was used only to adjust for time trends in the frequency of adverse outcomes over the 5-year study period. We adjusted for this to reduce potential confounding by time trends when estimating regression coefficients for the variables that are of primary interest (ie, patient preoperative risk factors—see example in Part 1). Surgery date was categorized into 6-month intervals and modeled as a linear trend across the ordinal categories. Surgery date is not included in the final risk prediction algorithm, and a patient's predicted risk does not depend on it. The published intercept parameter has been adjusted to incorporate the time trend, and this adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Other variables were selected in a stepwise fashion using a significance criterion of 0.05 for entry and removal. Ordinal categorical variables were initially coded such that removing an indicator variable caused a category to be combined with the lowest risk category (the reference group). In the case of myocardial infarction (MI), there were two outcomes (permanent stroke, prolonged length of stay) in which "MI 1 to 21 days" was retained but "MI less than 24 hours" was removed. For these two cases, the two MI categories were replaced by the single category "MI 21 days or less." The stepwise procedure was performed separately for each endpoint. Multiple interaction terms consisting of predictor variable and surgery type were also evaluated, and two additional interaction terms (age by reoperation and age by emergent status) were forced into the models (see Tables 3 and 5).

The results of this initial selection process were then reviewed by surgeon members of the QMTF for face validity and consistency with previous STS or other valve models: (1) preoperative atrial fibrillation was forced into the model for permanent stroke; (2) an indicator variable for dialysis was forced into any model that included creatinine (this did not apply to the renal failure model, as patients with preoperative dialysis were excluded); (3) sex was forced into all models; and (4) each variable that interacted with surgery group was also included as a main effect.

After validating the model in the 40% validation sample, the development and validation samples were then rejoined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regression model was estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [22]. An independence working correlation matrix was used to apply the generalized estimating equations. With this approach, the estimated regression coefficients were identical to those obtained using ordinary logistic regression, but the standard errors were adjusted to account for the clustered data structure.

Results

Risk Factors, Outcomes, and Predictor Variables

Table 1 presents the distribution of risk factors and endpoints in the overall 2002 to 2006 study population. Because there are three valve plus CABG categories, space limitations prevent display of the bivariate relationships for each predictor variable, endpoint, and valve plus CABG group. These are available upon request from STS.

Table 2 summarizes the overall frequency of adverse outcomes as well as the outcomes for the three major valve groups. Table 3 lists the candidate predictor variables and their coding schemes.

Assessment of Model Fit and Discrimination

The Hosmer-Lemeshow test was not employed to assess overall calibration. Large sample sizes make a significant p value almost inevitable, as all risk models are only approximations of reality [23]. Rather, we assessed calibration graphically by plotting observed versus predicted event rates within deciles of predicted risk in the development and validation samples (Fig 1). These plots were constructed for the overall sample and for subgroups based on surgery type (AVR plus CABG, MVR plus CABG, MVRepair plus CABG); age ($<$ 60, 60 to 79, $\geq$ 80 years); sex (male, female); diabetes mellitus (yes/no); status (elective, nonelective); and ejection fraction ($\leq$ 40, $>$ 40). Because of space constraints, only the overall sample results in the validation sample are presented. Additional results are available at www.sts.org/riskmodels.

In general, the models were well calibrated in the validation sample. The average absolute difference between observed versus predicted event rates across the decile categories ranged from 0.1% for deep sternal wound infec-

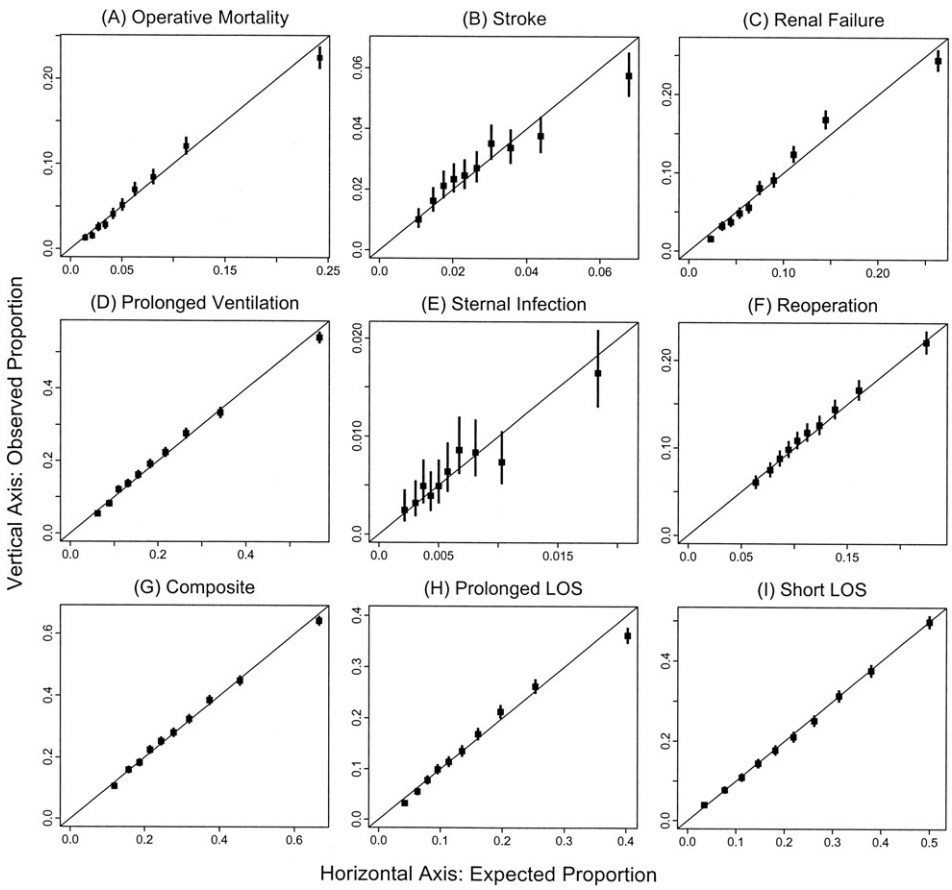


Fig 1. Plots of observed (O) versus expected (E) in validation sample

tion to 0.96% for prolonged length of stay. There was a very slight tendency for the models to overpredict risk in the highest decile. Although perfect prediction would be ideal, a slight overprediction implies that the model will give adequate credit to surgeons who take on patients with several model risk factors.

Discrimination was assessed by determining the c-statistic, also known as the area under the receiver operating characteristic (ROC) curve. Table 4 presents the discrimination of the various models. In the validation sample, the c-index of the overall valve plus CABG operative mortality model was 0.750, and the c-indices of the morbidity models ranged from 0.617 for reoperation to 0.724 for renal failure and short length of stay.

Table 4. Discrimination of Models (C-Index) in Development and Validation Samples

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall									
Development sample	0.754	0.656	0.729	0.730	0.670	0.623	0.704	0.719	0.726
Validation sample	0.750	0.622	0.724	0.720	0.646	0.617	0.698	0.710	0.724
AVR + CABG									
Development sample	0.737	0.648	0.720	0.706	0.639	0.607	0.678	0.705	0.700
Validation sample	0.736	0.609	0.718	0.697	0.657	0.604	0.673	0.699	0.698
MVR + CABG									
Development sample	0.764	0.665	0.712	0.746	0.713	0.608	0.725	0.694	0.726
Validation sample	0.739	0.611	0.701	0.733	0.580	0.599	0.714	0.680	0.733
MVRepair + CABG									
Development sample	0.746	0.650	0.727	0.725	0.692	0.624	0.707	0.712	0.738
Validation sample	0.755	0.652	0.715	0.716	0.644	0.623	0.705	0.702	0.733

AVR = aortic valve replacement; CABG = coronary artery bypass graft; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Table 5. Estimated Odds Ratios for CABG Mortality, Morbidity, and Length of Stay Models

A. Odds ratios for variables that do not interact with surgery group

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.12, 1.29)	1.05 (0.94, 1.17)	1.18 (1.11, 1.26)	1.13 (1.07, 1.19)	NA	1.16 (1.10, 1.22)	1.15 (1.10, 1.20)	1.22 (1.15, 1.28)	0.71 (0.67, 0.75)
BSA 1.6 versus 2.0 among females	1.29 (1.19, 1.39)	1.34 (1.18, 1.52)	0.87 (0.81, 0.94)	1.08 (1.02, 1.14)	0.51 (0.39, 0.67)	1.13 (1.07, 1.23)	1.12 (1.07, 1.18)	0.97 (0.92, 1.03)	1.03 (0.96, 1.10)
BSA 1.6 versus 2.0 among males	1.58 (1.41, 1.77)	1.38 (1.17, 1.64)	1.18 (1.07, 1.31)	1.31 (1.21, 1.41)	0.71 (0.49, 1.03)	1.18 (1.12, 1.34)	1.32 (1.24, 1.41)	1.40 (1.29, 1.52)	0.81 (0.75, 0.88)
BSA 1.8 versus 2.0 among females	1.05 (1.00, 1.10)	1.16 (1.06, 1.26)	0.89 (0.85, 0.93)	0.99 (0.95, 1.02)	0.69 (0.61, 0.77)	1.03 (0.98, 1.06)	1.01 (0.98, 1.04)	0.94 (0.90, 0.97)	1.08 (1.04, 1.12)
BSA 1.8 versus 2.0 among males	1.15 (1.10, 1.20)	1.13 (1.07, 1.20)	1.01 (0.97, 1.05)	1.06 (1.03, 1.09)	0.83 (0.72, 0.95)	1.06 (1.04, 1.11)	1.07 (1.05, 1.10)	1.09 (1.06, 1.12)	0.96 (0.94, 0.99)
BSA 2.2 versus 2.0 among females	1.12 (1.02, 1.22)	0.87 (0.74, 1.02)	1.25 (1.15, 1.35)	1.13 (1.06, 1.20)	1.57 (1.32, 1.89)	1.04 (1.00, 1.17)	1.10 (1.04, 1.17)	1.19 (1.11, 1.27)	0.82 (0.76, 0.89)
BSA 2.2 versus 2.0 among males	1.04 (1.00, 1.09)	0.95 (0.90, 1.01)	1.15 (1.11, 1.18)	1.09 (1.06, 1.11)	1.25 (1.14, 1.37)	1.00 (0.95, 1.01)	1.07 (1.04, 1.09)	1.09 (1.06, 1.12)	0.91 (0.89, 0.93)
CVD with CVA	1.22 (1.11, 1.33)	1.72 (1.52, 1.95)	1.12 (1.04, 1.22)	1.27 (1.19, 1.34)	1.22 (0.95, 1.56)	1.12 (1.04, 1.20)	1.26 (1.20, 1.33)	1.26 (1.18, 1.35)	0.75 (0.70, 0.81)
CVD without CVA	NA	1.28 (1.13, 1.45)	1.14 (1.06, 1.23)	1.10 (1.04, 1.16)	NA	NA	1.11 (1.05, 1.17)	1.11 (1.05, 1.18)	0.85 (0.78, 0.92)
Diabetes, insulin	1.31 (1.20, 1.42)	1.16 (1.03, 1.30)	1.62 (1.52, 1.74)	1.32 (1.25, 1.40)	1.98 (1.59, 2.46)	NA	1.34 (1.28, 1.41)	1.49 (1.40, 1.58)	0.67 (0.62, 0.72)
Diabetes, noninsulin	1.12 (1.05, 1.19)	1.16 (1.06, 1.26)	1.28 (1.21, 1.35)	1.11 (1.07, 1.15)	1.30 (1.10, 1.54)	NA	1.12 (1.08, 1.16)	1.17 (1.12, 1.22)	0.84 (0.81, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	1.15 (1.11, 1.19)	1.20 (1.14, 1.26)	1.17 (1.14, 1.21)	1.19 (1.16, 1.22)	1.28 (1.15, 1.42)	1.09 (1.06, 1.11)	1.16 (1.14, 1.18)	1.13 (1.10, 1.16)	0.82 (0.81, 0.84)
Hypertension	NA	1.19 (1.08, 1.31)	1.25 (1.18, 1.33)	1.10 (1.05, 1.15)	1.33 (1.09, 1.63)	NA	1.12 (1.08, 1.16)	1.08 (1.03, 1.13)	0.92 (0.88, 0.96)
Immunosuppressive treatment	1.35 (1.17, 1.54)	NA	1.30 (1.15, 1.47)	1.28 (1.17, 1.40)	NA	1.27 (1.14, 1.42)	1.26 (1.16, 1.37)	1.22 (1.11, 1.34)	0.75 (0.67, 0.84)
Left main disease	1.12 (1.05, 1.20)	NA	NA	1.06 (1.02, 1.11)	NA	NA	NA	NA	NA
Mitral insufficiency, moderate/severe	NA	NA	NA	NA	NA	NA	1.07 (1.01, 1.12)	NA	NA
Tricuspid insufficiency, moderate/severe	1.27 (1.15, 1.41)	NA	1.25 (1.13, 1.38)	1.15 (1.06, 1.24)	NA	NA	1.14 (1.07, 1.22)	NA	0.79 (0.69, 0.92)
Peripheral vascular disease	1.29 (1.21, 1.37)	1.15 (1.04, 1.27)	1.16 (1.10, 1.23)	1.18 (1.12, 1.24)	NA	1.15 (1.09, 1.22)	1.20 (1.15, 1.25)	1.16 (1.11, 1.22)	NA
Mitral stenosis	1.10 (0.99, 1.24)	NA	NA	NA	NA	NA	NA	1.09 (1.00, 1.18)	NA
MI 1–21 days	1.19 (1.10, 1.28)	NA	1.18 (1.10, 1.26)	1.28 (1.21, 1.35)	NA	NA	1.22 (1.16, 1.28)	NA	NA
MI $\leq$ 21 days[a]	NA	1.22 (1.11, 1.34)	NA	NA	NA	NA	NA	1.16 (1.10, 1.22)	NA
MI $<$ 24 hrs	1.65 (1.42, 1.91)	NA	1.30 (1.10, 1.54)	1.41 (1.23, 1.62)	NA	1.15 (1.00, 1.32)	1.49 (1.30, 1.70)	NA	NA
Time trend per 6-month harvest interval	0.98 (0.96, 0.99)	0.98 (0.97, 1.00)	1.01 (1.00, 1.02)	1.01 (1.00, 1.02)	0.96 (0.93, 0.99)	0.99 (0.98, 1.00)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)
Race black	NA	NA	1.15 (1.03, 1.30)	1.31 (1.19, 1.44)	NA	1.19 (1.06, 1.33)	1.21 (1.11, 1.32)	1.31 (1.19, 1.44)	0.65 (0.58, 0.72)
Race Hispanic	NA	NA	1.20 (1.03, 1.40)	1.17 (1.03, 1.32)	NA	1.08 (0.94, 1.24)	1.15 (1.03, 1.28)	1.13 (0.98, 1.30)	0.85 (0.71, 1.02)
Status, urgent versus elective	1.25 (1.17, 1.34)	NA	1.18 (1.10, 1.26)	1.26 (1.19, 1.33)	NA	1.14 (1.07, 1.21)	1.19 (1.14, 1.25)	1.28 (1.22, 1.35)	0.77 (0.72, 0.81)
Unstable angina	1.11 (1.03, 1.21)	0.89 (0.80, 1.00)	1.12 (1.05, 1.20)	1.06 (0.99, 1.13)	NA	NA	NA	NA	NA

Table 5. Continued

B. Odds ratios for AVR plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.29 (1.20, 1.39)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.20 (1.16, 1.24)	1.37 (1.32, 1.42)	0.74 (0.72, 0.77)
Age 70 versus 50 (no reop, elective)	1.67 (1.45, 1.92)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.44 (1.36, 1.54)	1.86 (1.73, 2.01)	0.55 (0.52, 0.59)
Age 80 versus 50 (no reop, elective)	2.47 (2.08, 2.94)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	1.86 (1.73, 2.01)	2.67 (2.46, 2.91)	0.33 (0.30, 0.36)
CHF, not NYHA IV	1.24 (1.14, 1.34)	0.98 (0.88, 1.09)	1.19 (1.11, 1.28)	1.22 (1.16, 1.29)	NA	NA	1.14 (1.08, 1.19)	1.30 (1.23, 1.38)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.48 (1.34, 1.64)	1.15 (1.00, 1.32)	1.35 (1.24, 1.48)	1.47 (1.36, 1.59)	NA	1.16 (1.08, 1.24)	1.36 (1.27, 1.45)	1.49 (1.39, 1.60)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	2.26 (2.13, 2.40)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.62 (0.58, 0.67)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.65 (1.41, 1.92)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.30 (0.25, 0.37)
EF per 10-unit decrease	1.10 (1.06, 1.15)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.87 (0.84, 0.89)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.41 (1.25, 1.58)	0.56 (0.48, 0.65)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.17 (0.92, 1.50)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	1.79 (1.50, 2.15)	1.45 (1.29, 1.63)	NA
Female versus male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.52 (1.44, 1.61)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.61 (0.57, 0.64)
Active infectious endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	1.81 (1.41, 2.32)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.26 (1.22, 1.30)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.26 (1.22, 1.30)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.83 (1.58, 2.11)	NA	1.39 (1.16, 1.67)	1.50 (1.32, 1.69)	1.55 (1.33, 1.81)	0.67 (0.58, 0.77)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	2.19 (1.80, 2.65)	NA	1.48 (1.15, 1.92)	1.77 (1.51, 2.06)	1.65 (1.34, 2.03)	0.53 (0.43, 0.65)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.33 (0.22, 0.50)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.18 (0.09, 0.34)

Table 5. Continued

C. Odds ratios for MVR plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.51 (1.39, 1.64)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.27 (1.21, 1.32)	1.37 (1.32, 1.42)	0.68 (0.64, 0.72)
Age 70 versus 50 (no reop, elective)	2.28 (1.94, 2.68)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.60 (1.47, 1.75)	1.86 (1.73, 2.01)	0.46 (0.41, 0.52)
Age 80 versus 50 (no reop, elective)	3.95 (3.17, 4.93)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	2.18 (1.92, 2.48)	2.67 (2.46, 2.91)	0.25 (0.21, 0.30)
CHF, not NYHA IV	0.91 (0.80, 1.03)	0.80 (0.64, 0.99)	0.92 (0.82, 1.03)	1.02 (0.93, 1.11)	NA	NA	0.94 (0.87, 1.02)	1.03 (0.94, 1.12)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.09 (0.95, 1.24)	0.93 (0.75, 1.17)	1.04 (0.92, 1.19)	1.22 (1.10, 1.35)	NA	1.16 (1.08, 1.24)	1.13 (1.03, 1.23)	1.17 (1.06, 1.30)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	1.82 (1.66, 2.01)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.66 (0.57, 0.78)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.21 (0.95, 1.55)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.30 (0.18, 0.48)
EF per 10-unit decrease	1.23 (1.16, 1.30)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.89 (0.82, 0.95)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.29 (1.14, 1.46)	0.51 (0.39, 0.65)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.21 (0.97, 1.50)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	2.76 (2.22, 3.42)	1.45 (1.29, 1.63)	NA
Female versus male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.17 (1.08, 1.28)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.66 (0.59, 0.74)
Active infectious endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	2.08 (1.62, 2.67)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.18 (1.12, 1.24)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.20 (1.14, 1.26)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.38 (1.19, 1.61)	NA	1.15 (0.95, 1.38)	1.50 (1.32, 1.69)	1.30 (1.10, 1.53)	0.81 (0.66, 0.99)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	1.66 (1.35, 2.03)	NA	1.22 (0.95, 1.56)	1.77 (1.51, 2.06)	1.38 (1.12, 1.71)	0.64 (0.50, 0.82)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.26 (0.16, 0.43)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.14 (0.07, 0.27)

Table 5. Continued

D. Odds ratios for MVRepair plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.46 (1.36, 1.57)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.28 (1.23, 1.33)	1.37 (1.32, 1.42)	0.66 (0.63, 0.69)
Age 70 versus 50 (no reop, elective)	2.14 (1.86, 2.46)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.63 (1.51, 1.76)	1.86 (1.73, 2.01)	0.44 (0.40, 0.48)
Age 80 versus 50 (no reop, elective)	3.60 (2.97, 4.33)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	2.23 (2.00, 2.49)	2.67 (2.46, 2.91)	0.23 (0.20, 0.27)
CHF, not NYHA IV	0.96 (0.85, 1.09)	1.05 (0.90, 1.23)	0.99 (0.88, 1.10)	1.10 (1.02, 1.19)	NA	NA	1.06 (0.99, 1.14)	1.17 (1.08, 1.26)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.16 (1.02, 1.32)	1.23 (1.04, 1.46)	1.12 (0.99, 1.27)	1.32 (1.21, 1.44)	NA	1.16 (1.08, 1.24)	1.27 (1.17, 1.37)	1.33 (1.22, 1.45)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	1.87 (1.72, 2.04)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.59 (0.53, 0.67)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.88 (1.52, 2.31)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.35 (0.24, 0.49)
EF per 10-unit decrease	1.09 (1.04, 1.15)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.84 (0.81, 0.87)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.56 (1.40, 1.73)	0.52 (0.44, 0.62)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.69 (1.41, 2.01)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	2.17 (1.81, 2.60)	1.45 (1.29, 1.63)	NA
Female vs male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.25 (1.15, 1.36)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.60 (0.55, 0.66)
Active infectious Endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	2.98 (1.86, 4.77)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.21 (1.16, 1.27)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.16 (1.10, 1.21)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.55 (1.32, 1.82)	NA	1.49 (1.23, 1.82)	1.50 (1.32, 1.69)	1.32 (1.10, 1.58)	0.80 (0.68, 0.95)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	1.86 (1.53, 2.26)	NA	1.59 (1.20, 2.11)	1.77 (1.51, 2.06)	1.41 (1.11, 1.79)	0.63 (0.51, 0.79)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.43 (0.29, 0.66)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.23 (0.12, 0.44)

[a] For CVA and PLOS, MI coded ≤ 21 days; for all other endpoints, MI coded < 24 hrs or 1 to 21 days. [b] Variable interacts with age. Reported odds ratio represents effect of risk factor for patients aged 50 years old.

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; Mort = mortality; NA = not applicable; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Odds Ratios

Table 5 presents the odds ratios and 95% confidence intervals (CI) derived from these models. "Not applicable" indicates that those predictors were not included in a particular risk model.

Odds ratios that do not interact with surgery type are summarized in Part A of Table 5. Several variables interact with surgery type, and the odds ratios for these variables differ for some of the endpoints depending on the specific type of surgery, as summarized in Tables 5B, C, and D (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). For example, in the model for prolonged length of stay, the odds ratio for active endocarditis is 1.81 (95% CI: 1.41 to 2.32) for AVR plus CABG; 2.08 (95% CI: 1.62 to 2.67) for MVR plus CABG; and 2.98 (95% CI: 1.86 to 4.77) for MVRepair plus CABG.

Final Model Intercept and Coefficients

The algorithms for calculating predicted risk values, including the intercepts and regression coefficients, are presented in the Appendix.

Limitations

The limitations of the STS valve plus CABG models are similar to those discussed in Part 1 of this series.

Conclusion

A new STS model has been developed for valve surgery combined with CABG. This model includes specific indicator variables for each major type of valve plus CABG procedure (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). Models have been developed for operative mortality, individual morbidity endpoints, a composite morbidity or mortality endpoint, and short and prolonged postoperative length of stay. Overall model performance is excellent.

References

1. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA 1990;264:2768–74.
2. Edwards FH, Clark RE, Schwartz M. Coronary artery bypass grafting: the Society of Thoracic Surgeons National Database experience. Ann Thorac Surg 1994;57:12–9.
3. O'Connor GT, Plume SK, Olmstead EM, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. Circulation 1992;85:2110–8.
4. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. Ann Thorac Surg 2004; 78:1868–77.
5. Shroyer AL, Coombs LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. Ann Thorac Surg 2003;75:1856–64.
6. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
7. Jamieson WR, Edwards FH, Schwartz M, Bero JW, Clark RE, Grover FL. Risk stratification for cardiac valve replacement. National Cardiac Surgery Database. Database Committee of The Society of Thoracic Surgeons. Ann Thorac Surg 1999;67:943–51.
8. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
9. Hannan EL, Racz MJ, Jones RH, et al. Predictors of mortality for patients undergoing cardiac valve replacements in New York State. Ann Thorac Surg 2000;70:1212–8.
10. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
11. Nowicki ER. What is the future of mortality prediction models in heart valve surgery? Ann Thorac Surg 2005;80:396–8.
12. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
13. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
14. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
15. Gardner SC, Grunwald GK, Rumsfeld JS, et al. Comparison of short-term mortality risk factors for valve replacement versus coronary artery bypass graft surgery. Ann Thorac Surg 2004;77:549–56.
16. Grover FL, Edwards FH. Similarity between the STS and New York State databases for valvular heart disease. Ann Thorac Surg 2000;70:1143–4.
17. Rankin JS, Hammill BG, Ferguson TB Jr, et al. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg 2006;131:547–57.
18. Roques F, Nashef SA, Michel P. Risk factors for early mortality after valve surgery in Europe in the 1990s: lessons from the EuroSCORE pilot program. J Heart Valve Dis 2001;10:572–7.
19. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85:921–30.
20. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
21. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2002.
22. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.
23. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Plus CABG Models

For each endpoint, the formula for calculating a patient's predicted risk of the endpoint has the form:

$$\text{Predicted Risk} = \frac{e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}{1+e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}$$

where $x_1, x_2, \ldots, x_n$ denote patient preoperative risk factors (eg, quantitative variables such as age, and comorbidities coded as 1=present, 0=absent); and $\beta_0, \beta_1, \ldots, \beta_n$ denote regression coefficients (numerical constants). Regression coefficients for each endpoint are presented in Appendix Table 1. The variables $x_1, x_2, \ldots, x_n$ are the same for each endpoint and are defined in Appendix Table 2. The regression coefficient for the time trend is not presented. Instead, the intercept has been adjusted to incorporate the time trend. This adjusted intercept reflects the baseline risk for a reference period of July–December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−5.24391	−5.14546	−5.32535	−3.63438	−6.50043	−3.16980	−2.99714	−4.15892	1.18582
Preoperative AFib	0.18430	0.04634	0.16567	0.12059	0.00000	0.14910	0.13766	0.19656	−0.34095
Age function 1	0.02560	0.02487	0.03268	0.02106	0.00545	0.01715	0.01838	0.03115	−0.02970
Age function 3	0.02758	−0.00709	0.00671	0.00791	−0.00985	−0.00021	0.01425	0.00985	−0.04542
Age by reop function	−0.00861	0.00458	0.00077	−0.00673	0.00314	−0.00399	−0.00202	−0.00678	0.00656
Age by status function	−0.00507	−0.01979	−0.00178	−0.00750	0.01627	−0.00029	0.00229	−0.02247	0.00692
Age by MVR function	0.01564	0.00000	0.00000	0.00000	0.00000	0.00000	0.00527	0.00000	−0.00866
Age by MVRepair function	0.01240	0.00000	0.00000	0.00000	0.00000	0.00000	0.00599	0.00000	−0.01159
BSA function 1	−1.14176	−0.81169	−0.41848	−0.66843	0.86401	−0.51266	−0.70411	−0.84204	0.51295
BSA function 2	2.25471	0.94689	1.84088	1.80467	0.42453	0.70024	1.70623	2.10402	−1.66758
CHF but not NYHA IV	0.21206	−0.01726	0.17460	0.20063	0.00000	0.00000	0.12880	0.26291	−0.17652
CHF and NYHA IV	0.39457	0.14109	0.30146	0.38383	0.00000	0.14499	0.30567	0.39791	−0.31077
CHF by MVR function	−0.31077	−0.20917	−0.25767	−0.18455	0.00000	0.00000	−0.18635	−0.23729	0.00000
CHF by MVRepair function	−0.24791	0.06897	−0.18667	−0.10484	0.00000	0.00000	−0.06920	−0.10954	0.00000
CLD function	0.17713	0.00000	0.11379	0.23345	0.27571	0.09280	0.16523	0.22999	−0.19234
CLD by MVR function	0.00000	0.00000	0.00000	−0.06780	0.00000	0.00000	0.00000	−0.04591	0.00000
CLD by MVRepair function	0.00000	0.00000	0.00000	−0.04014	0.00000	0.00000	0.00000	−0.08501	0.00000
Creatinine function 1	0.44794	0.23545	0.81612	0.38147	0.00000	0.24620	0.51256	0.41472	−0.47658
Creatinine by MVR function	0.00000	0.00000	−0.21574	0.00000	0.00000	0.00000	0.00000	0.00000	0.06652
Creatinine by MVRepair function	0.00000	0.00000	−0.18787	0.00000	0.00000	0.00000	0.00000	0.00000	−0.04407
CVD without prior CVA	0.00000	0.24847	0.13299	0.09769	0.00000	0.00000	0.10255	0.10601	−0.16643
CVD and prior CVA	0.19754	0.54344	0.11571	0.23581	0.19686	0.10974	0.23332	0.23319	−0.28560
Diabetes, noninsulin	0.11060	0.14576	0.24490	0.10365	0.26281	0.00000	0.11462	0.15846	−0.17020
Diabetes, insulin	0.26870	0.14582	0.48504	0.27893	0.68330	0.00000	0.29508	0.39583	−0.40448
Dialysis	1.61151	0.58833	0.00000	1.20290	0.61527	0.74332	1.25181	1.29747	−1.67728
Dialysis by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	−0.30339	0.00000	0.00000	0.04745
Dialysis by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.13058	0.00000	0.00000	0.09778
Ejection fraction function	0.00989	0.00000	0.00534	0.01113	0.00000	0.00703	0.01061	0.00995	−0.01440
EF by MVR function	0.01056	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00228
EF by MVRepair function	−0.00117	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.00309
Endocarditis, active	0.71327	0.60657	0.41797	0.67172	0.00000	0.44757	0.74858	0.59333	−1.27854
Endocarditis by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.13795	0.00000
Endocarditis by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.49934	0.00000
Female	0.30852	0.17170	0.16668	0.41874	0.10654	−0.08221	0.18594	0.26947	−0.50044
Female by MVR function	0.00000	0.00000	0.00000	−0.25972	0.00000	0.00000	0.00000	0.00000	0.08895
Female by MVRepair function	0.00000	0.00000	0.00000	−0.19373	0.00000	0.00000	0.00000	0.00000	−0.00229
Female by BSA function 1	0.51233	0.07575	0.76032	0.48032	0.80594	0.16701	0.41581	0.91055	−0.59086
Female by BSA function 2	−0.27980	−0.88628	−0.57622	−0.49740	0.58767	0.52524	−0.40427	−0.78096	0.15748
Hypertension	0.00000	0.17080	0.22638	0.09581	0.28851	0.00000	0.11445	0.07602	−0.08668

Appendix Table 1. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
IABP or inotropes	0.36025	0.00000	0.23674	0.77918	0.00000	0.15075	0.56477	0.34008	−0.58536
IABP by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.08732	−0.09462
IABP by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10281	−0.06743
Immunosuppressive treatment	0.29654	0.00000	0.26400	0.24814	0.00000	0.24041	0.23332	0.19750	−0.28819
Insufficiency, mitral	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.06414	0.00000	0.00000
Insufficiency, tricuspid	0.24006	0.00000	0.22040	0.13606	0.00000	0.00000	0.13318	0.00000	−0.23141
Left main disease	0.11450	0.00000	0.00000	0.06181	0.00000	0.00000	0.00000	0.00000	0.00000
MI 1–21 days	0.17038	0.00000	0.16476	0.24560	0.00000	0.00000	0.19751	0.00000	0.00000
MI ≤ 21 days	0.00000	0.19671	0.00000	0.00000	0.00000	0.00000	0.00000	0.14599	0.00000
MI < 24 hours	0.49918	0.00000	0.26240	0.34321	0.00000	0.13716	0.39731	0.00000	0.00000
MVR	0.14888	0.32659	0.90926	0.76504	0.28437	0.41642	0.41322	0.73530	−0.82339
MVRepair	−0.07374	0.06933	0.51275	0.28204	0.19499	0.07390	−0.03949	0.30384	−0.03552
No. diseased coronary vessel function	0.13746	0.18243	0.15791	0.17277	0.24582	0.08187	0.14767	0.12474	−0.19250
Peripheral vascular disease	0.25173	0.13776	0.14995	0.16591	0.00000	0.14312	0.18062	0.14863	0.00000
Race black	0.00000	0.00000	0.14301	0.26900	0.00000	0.17364	0.19182	0.26856	−0.43385
Race Hispanic	0.00000	0.00000	0.18384	0.15363	0.00000	0.08065	0.13561	0.12286	−0.15901
Reop, 1 previous operation	0.78624	0.00000	0.25782	0.60179	0.00000	0.33209	0.40293	0.43757	−0.39723
Reop, ≥ 2 previous operations	0.90015	0.00000	0.38499	0.78263	0.00000	0.39502	0.56875	0.50334	−0.63237
Reop by MVR function	0.00000	0.00000	0.00000	−0.27846	0.00000	−0.19608	0.00000	−0.17836	0.18262
Reop by MVRepair function	0.00000	0.00000	0.00000	−0.16306	0.00000	0.06985	0.00000	−0.16007	0.17613
Shock	0.51917	0.17321	0.15810	0.65653	0.00000	0.21271	0.58409	0.36987	0.00000
Shock by MVR function	0.00000	0.00000	0.02883	0.00000	0.00000	0.00000	0.43045	0.00000	0.00000
Shock by MVRepair function	0.00000	0.00000	0.36429	0.00000	0.00000	0.00000	0.19084	0.00000	0.00000
Status urgent	0.22591	0.00000	0.16451	0.22905	0.00000	0.12800	0.17511	0.24758	−0.26626
Status emergent	0.75852	0.79460	0.56854	0.99818	0.00000	0.34063	0.77631	1.00162	−1.09633
Status salvage	1.51811	0.95665	0.61798	0.75178	0.00000	0.00000	1.20732	0.56482	−1.72252
Status by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.25083
Status by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.25943
Stenosis, mitral	0.09879	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.08469	0.00000
Unstable angina	0.10722	−0.11292	0.11597	0.05762	0.00000	0.00000	0.00000	0.00000	0.00000

Afib = atrial fibrillation; BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Appendix Table 2. Definition of Variables Appearing in STS 2008 Valve Plus CABG Models

Variable	Definition
Intercept	= 1 for all patients
Preoperative AFib	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine by MVR function	= Creatinine function 1 if valve operation is MVR, = 0 otherwise
Creatinine by MVRepair function	= Creatinine function 1 if valve operation is MVRepair, = 0 otherwise
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
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
EF by MVR function	= Ejection fraction function if valve operation is MVR, = 0 otherwise
EF by MVRepair function	= Ejection fraction function if valve operation is MVRepair, = 0 otherwise
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Endocarditis by MVR function	= 1 if patient has active endocarditis and valve operation is MVR, = 0 otherwise
Endocarditis by MVRepair function	= 1 if patient has active endocarditis and valve operation is MVRepair, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP or inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP or inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient has received immunosuppressive therapy within 30 days, = 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 1–21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI ≤ 21 days[a]	= 1 if patient has history of MI within 21 days prior to surgery, = 0 otherwise (for CVA and PLOS; coded as < 24 hours and 1–21 days for others)
MI < 24 hours	= 1 if history of MI < 24 hours prior to surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased coronary vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 0 otherwise

Appendix Table 2. Continued

Variable	Definition
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
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
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not rescuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis, mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina and no MI within 7 days of surgery, = 0 otherwise

[a] MI coded ≤ 21 days for CVA and PLOS endpoints; for all other endpoints, coded as < 24 hours and 1 to 21 days.

Note: See www.sts.org for exact definitions of terms used above.

BSA = body surface area; CABG = coronary artery bypass graft surgery; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

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 #:** 0123 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Operative Mortality for Aortic Valve Replacement (AVR) + CABG Surgery
De.2 **Brief description of measure:** Percent of patients undergoing combined AVR and CABG who die, including both 1) all deaths occurring during the hospitalization in which the procedure was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure
1.1-2 **Type of Measure:** Outcome 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-634282033057346226.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:** Public reporting, Internal quality improvement	**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, Frequently performed procedure, Leading cause of morbidity/mortality, High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Aortic valve replacement with concomitant coronary artery bypass grafting (AVR and CABG) remains a commonly performed operation. Given the increasing prevalence of both aortic stenosis and coronary artery disease in our aging population, this therapy will remain in high demand. With the establishment of validated risk models (see Shahian below), benchmarks for mortality and morbidity can be created. The lowering of complication rates – especially stroke, re-operation for bleeding, mediastinitis, renal failure, and respiratory failure will be essential to improve the outcomes for patients undergoing AVR and CABG. These complications likely will have a direct impact upon mortality 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892. - Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137.	**1a** C☐ P☐ M☐ N☐

- Chikwe J, Croft LB, Goldstone AB, Castillo JG, Rahmanian PB, Adams DH, et al . Comparison of the results of aortic valve replacement with or without concomitant coronary artery bypass grafting in patients with left ventricular ejection fraction <30%versus patients with ejection fraction > 30%. Am J Cardiol. 2009;104:1717-21.
- Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 3 – Valve Plus Coronary Artery Bypass Grafting Surgery. Ann Thor Surg. 2009;88:S43-62.

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** Establishing benchmarks for mortality following aortic valve replacement and coronary artery bypass grafting will allow for cardiac surgery programs to collectively improve outcomes through continuous quality improvement measures.

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, 2005-December 31, 2009

Analysis includes 517 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data (not restricted to this measure) to STS for at least 36 months in 2005-2009.

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

1b.5 **Citations for data on Disparities:**

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*)**:** Measurement of the risk-adjusted mortality rate for AVR + CABG will provide useful information for patients. Furthermore, feedback of these results to cardiac programs may stimulate quality improvement initiatives.

1c.2-3. **Type of Evidence:** Observational study, Expert opinion, Systematic synthesis of research, Other 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*)**:**
The evidence to date consists of large, single center reports and from the STS National Adult Cardiac Surgery Database. As cited above, the STS Database collects results from 90% of cardiac surgery programs in the US

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 NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437.
- Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892.

1c
C☐
P☐
M☐
N☐

- Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137. - Chikwe J, Croft LB, Goldstone AB, Castillo JG, Rahmanian PB, Adams DH, et al . Comparison of the results of aortic valve replacement with or without concomitant coronary artery bypass grafting in patients with left ventricular ejection fraction <30%versus patients with ejection fraction > 30%. Am J Cardiol. 2009;104:1717-21. - Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 3 – Valve Plus Coronary Artery Bypass Grafting Surgery. Ann Thor Surg. 2009;88:S43-62. **1c.9 Quote the Specific guideline recommendation (***including guideline number and/or page number***):** N/A **1c.10 Clinical Practice Guideline Citation:** **1c.11 National Guideline Clearinghouse or other URL:** **1c.12 Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):** **1c.13 Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:** **1c.14 Rationale for using this guideline over others:**	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
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***):** Number of patients undergoing combined AVR and CABG who die, including both 1) all deaths occurring during the hospitalization in which the operation was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure. **2a.2 Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** During hospitalization regardless of length of stay or within 30 days of surgery if discharged **2a.3 Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):** Number of AVR + CABG procedures with an operative mortality; Number of AVR + CABG procedures in which Mortality [Mortalty (STS Adult Cardiac Surgery Database Version	**2a-specs** **C☐** **P☐** **M☐** **N☐**

2.73)] and Mortality Operative Death (MtOpD) are marked "yes." Operative mortality is further verified by the following variables: Mortality Status at 30 days (Mt30Stat), Mortality Date (MtDate), Mortality Discharge Status (MtDCStat)

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):**
All patients undergoing combined AVR + CABG

2a.5 Target population gender: Female, Male
2a.6 Target population age range: 18 yrs and older

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

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
Number of AVR + CABG procedures;

AVR + CABG is determined as a procedure for which all of the following apply:
- OpCAB is marked "Yes"
- OpValve is marked "Yes"
- VSAV is marked "Yes"
- VSAVPr is marked "Replacement"
- (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
- ResectSubA, VSMV, VSMVPr, OpTricus, OpPulm, OpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc, OCTumor, OCPulThromDis, OCarOthr are all marked "no" or "missing"

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

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

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

2a.12-13 Risk Adjustment Type: Case-mix adjustment

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

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment 2a.15 Detailed Risk Model-634282035059769330.pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**

2a.22 Describe the method for discriminating performance *(e.g., significance testing)***:**
Participant specific OR and their 95% CI were estimated in the hierarchical model. These model-based estimates were used to control variation due to random statistical fluctuations while estimating true signal variation. A 95% CI excluding zero indicates the participant's performance is significantly lower or higher than an "average" STS participant.

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)*:

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: national, Population: regional/network, Population: states, Population: counties or cities

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)***:** STS Adult Cardiac Surgery Database - Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)***:**
Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009. Excluded from analysis are participants that did not submit results for both time periods. Because database participants can change their underlying care processes at any time, we would not expect perfect correlation between two sets of results from even proximate time periods.

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

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 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

2c
C☐
P☐
M☐
N☐

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.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Mortality Operative Death: 100.0% agreement rate

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)*: Please see Risk Adjustment Type section above

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
Detailed information regarding the risk adjustment model can be found in the attachment:

Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S43-62.

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)*: 517 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009; January 1, 2005-December 31, 2009

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:
We calculated the risk adjusted event rate with the participant's Odds Ratio (OR) estimate and the overall STS event rate. Therefore, the risk adjusted rate is closely related to OR estimate. If OR > 1, then the participant's risk adjusted rate will be greater than the overall STS event rate; if OR < 1, then the participant's risk adjusted rate will be smaller than the overall STS event rate. The statistical significance is defined by the 95% confidence interval (CI) or the OR estimate. If the 95% CI for a participant's OR includes the null value 1.0, then we cannot distinguish this participant's performance from the STS average

2f
C☐
P☐
M☐
N☐

– either the participant's performance was close to average or else the participant's sample size was too small to make a reliable inference. Otherwise, if the 95% CI falls to the right of 1.0, then the participant's performance is considered significantly lower than the average STS results; if the 95% CI falls to the left of 1.0, then the participant's performance is considered significantly higher than the average STS results. 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 STS plans to publicly report more measures in the future. There is no definite date yet assigned to this measure; however, STS staff and surgeon leadership have engaged in initial internal STS discussions regarding this matter. 3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives,*	**3a** C☐ P☐ M☐ N☐

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)**:** **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.	**4d** **C☐** **P☐** **M☐** **N☐**

STS audited for these potential problems during testing. Please see IFMC audit results.	
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 0123 Sections 2a.14, 1b.2, 2b.3, 2f.3, 3a.6.pdf
Date of Submission (*MM/DD/YY*)**:** 01/12/2011

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.

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

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 valve procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest four and a half years of data (January 2006 – 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 Valve plus CABG Models

Variable	Definition
Intercept	= 1 for all patients
Preoperative AFib	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine by MVR function	= Creatinine function 1 if valve operation is MVR, = 0 otherwise
Creatinine by MVRepair function	= Creatinine function 1 if valve operation is MVRepair, = 0 otherwise
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
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
EF by MVR function	= Ejection fraction function if valve operation is MVR, = 0 otherwise

EF by MVRepair function	= Ejection fraction function if valve operation is MVRepair, = 0 otherwise
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Endocarditis by MVR function	= 1 if patient has active endocarditis and valve operation is MVR, = 0 otherwise
Endocarditis by MVRepair function	= 1 if patient has active endocarditis and valve operation is MVRepair, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP or inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP or inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient has received immunosuppressive therapy within 30 days, = 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 1–21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI ≤ 21 days [a]	= 1 if patient has history of MI within 21 days prior to surgery, = 0 otherwise (for CVA and PLOS; coded as < 24 hours and 1–21 days for others)
MI < 24 hours	= 1 if history of MI < 24 hours prior to surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased coronary vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
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
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not rescuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis, mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina and no MI within 7 days of surgery, = 0 otherwise

[a] **MI coded as < 24 hours and 1 to 21 days.**

BSA = body surface area; CABG = coronary artery bypass graft surgery; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease;

The final patient-level model was built by step-wise selection method with several variables decided by surgeon panel 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	*Risk-Adjusted Operative Mortality for AV Replacement + CABG Surgery*
N	517
Mean	1.0
1st	1.9
5th	1.5
10th	1.3
25th	1.1
Median	1.0
75th	0.8
90th	0.7
95th	0.6
99th	0.5
Outlier	20 (3.9)
High	8
Low	12

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).

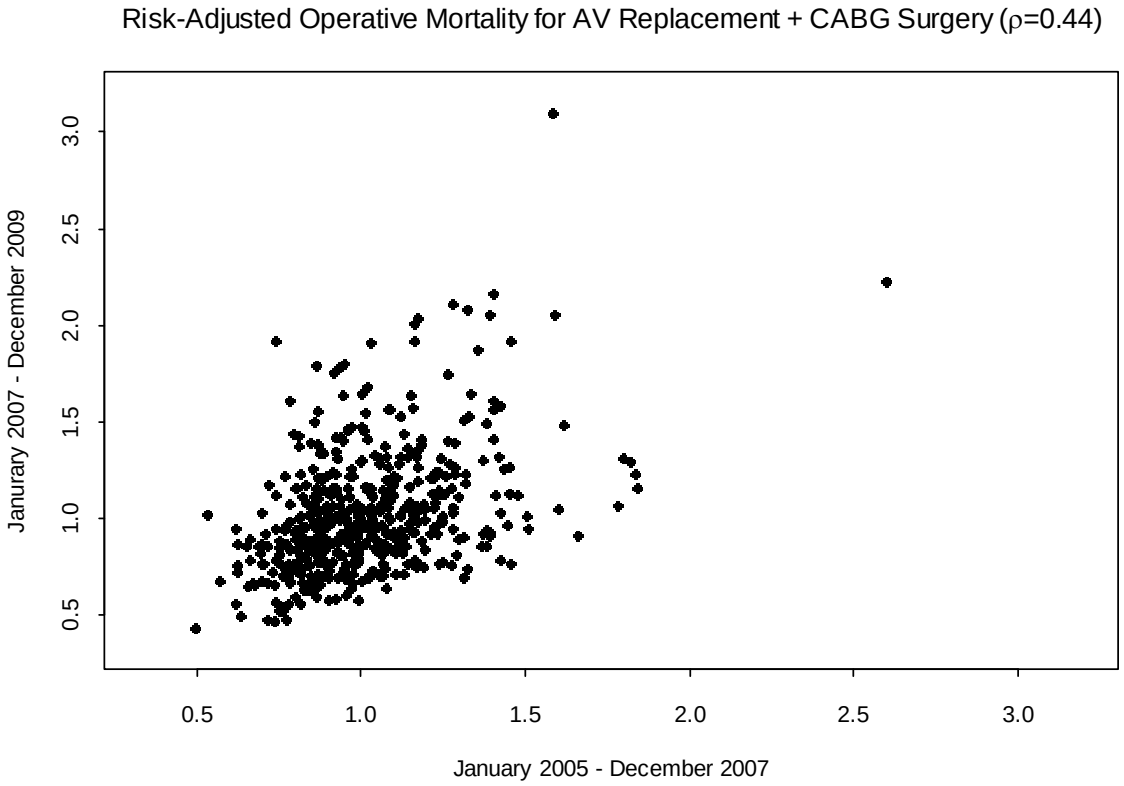
Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for AV Replacement + CABG Surgery*
N	517
Mean	5.0

Measurement	*Risk-Adjusted Operative Mortality for AV Replacement + CABG Surgery*
1st	2.9
5th	3.3
10th	3.7
25th	4.2
Median	4.9
75th	5.6
90th	6.6
95th	7.5
99th	8.7
Outlier	20 (3.9)
High	8
Low	12

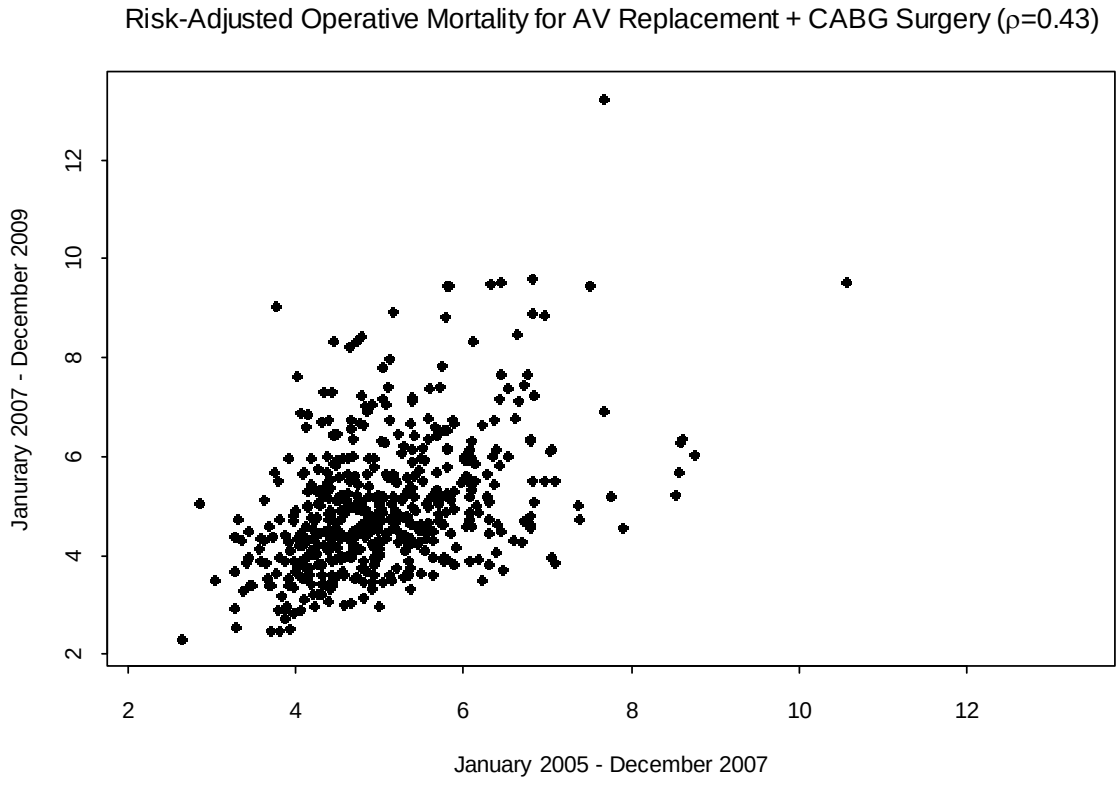
2b.3. Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*

Testing results: $\rho = 0.44$



Risk Adjusted Rate:

Testing results: $\rho = 0.43$



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, 2005-December 31, 2009. Sample contains 517 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009.

Measurement	*Risk-Adjusted Operative Mortality for AV Replacement + CABG Surgery*
N	517
Mean	1.0
1st	1.9
5th	1.5
10th	1.3
25th	1.1
Median	1.0
75th	0.8
90th	0.7
95th	0.6
99th	0.5
Outlier†	20 (3.9)
High	8
Low	12

Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for AV Replacement + CABG Surgery*
N	517
Mean	5.0
1st	2.9
5th	3.3
10th	3.7
25th	4.2
Median	4.9
75th	5.6
90th	6.6
95th	7.5
99th	8.7
Outlier†	20 (3.9)

Measurement	*Risk-Adjusted Operative Mortality for AV Replacement + CABG Surgery*
High	8
Low	12

†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."

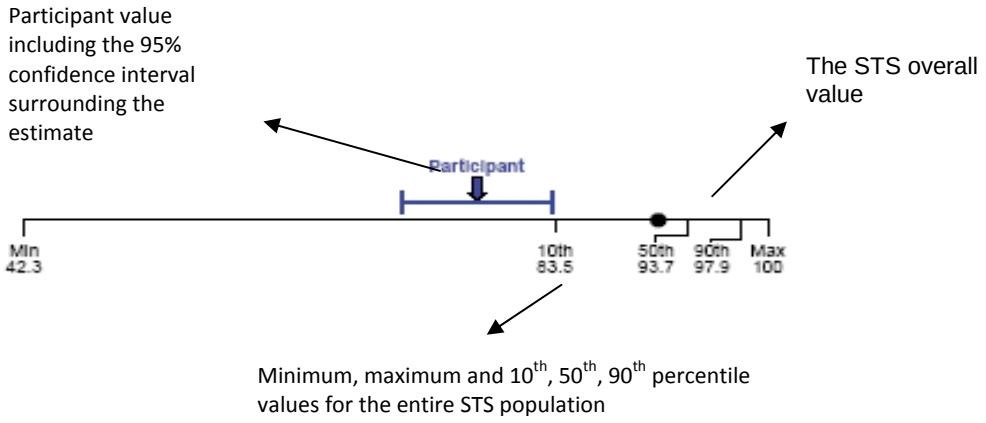
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

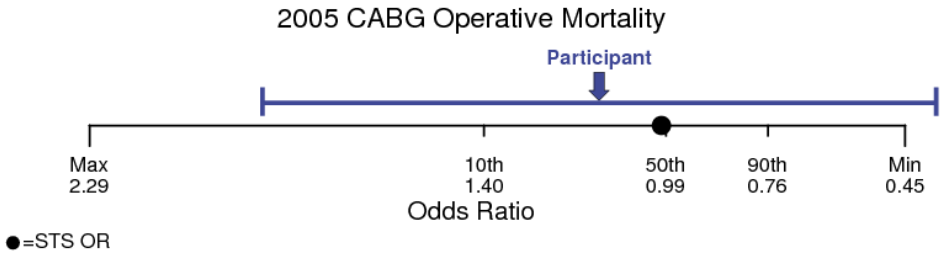
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

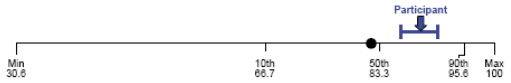
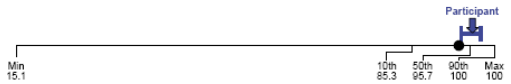
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 3—Valve Plus Coronary Artery Bypass Grafting Surgery

David M. Shahian, MD,[a] Sean M. O'Brien, PhD,[b] Giovanni Filardo, PhD, MPH,[c] Victor A. Ferraris, MD,[d] Constance K. Haan, MD,[e] Jeffrey B. Rich, MD,[f] Sharon-Lise T. Normand, PhD,[g] Elizabeth R. DeLong, PhD,[b] Cynthia M. Shewan, PhD,[h] Rachel S. Dokholyan, MPH,[b] Eric D. Peterson, MD, MPH,[b] Fred H. Edwards, MD,[e] and Richard P. Anderson, MD[i†]

[a]Massachusetts General Hospital, Boston, Massachusetts; [b]Duke Clinical Research Institute, Durham, North Carolina; [c]Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; [d]University of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; [e]University of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; [f]Sentara Cardiovascular Research Institute, Norfolk, Virginia; [g]Department of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [h]The Society of Thoracic Surgeons, Chicago, Illinois; and [i]Seattle, Washington

Background. Since 1999, The Society of Thoracic Surgeons (STS) has published two risk models that can be used to adjust the results of valve surgery combined with coronary artery bypass graft surgery (CABG). The most recent was developed from data for patients who had surgery between 1994 and 1997 using operative mortality as the only endpoint. Furthermore, this model did not specifically consider mitral valve repair plus CABG, an increasingly common procedure. Consistent with STS policy of periodically updating and improving its risk models, new models for valve surgery combined with CABG have been developed. These models specifically address both perioperative morbidity and mitral valve repair, and they are based on contemporary data.

Methods. The final study population consisted of 101,661 procedures, including aortic valve replacement (AVR) plus CABG, mitral valve replacement (MVR) plus CABG, or mitral valve repair (MVRepair) plus CABG between January 1, 2002, and December 31, 2006. Model outcomes included operative mortality, stroke, deep sternal wound infection, reoperation, prolonged ventilation, renal failure, composite major morbidity or mortality, prolonged postoperative length of stay, and short postoperative length of stay. Candidate variables were screened for frequency of missing data, and imputation techniques were used where appropriate. Stepwise variable selection was employed, supplemented by advice from an expert panel of cardiac surgeons and biostatisticians. Several variables were forced into models to insure face validity (eg, atrial fibrillation for the permanent stroke model, sex for all models). Based on preliminary analyses of the data, a single model was employed for valve plus CABG, with indicator variables for the specific type of procedure. Interaction terms were included to allow for differential impact of predictor variables depending on procedure type. After validating the model in the 40% validation sample, the development and validation samples were then combined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regression model was estimated using generalized estimating equations to account for clustering of patients within institutions.

Results. The c-index for mortality prediction for the overall valve plus CABG population was 0.75. Morbidity model c-indices for specific complications (permanent stroke, renal failure, prolonged ventilation > 24 hours, deep sternal wound infection, reoperation for any reason, major morbidity or mortality composite, and prolonged postoperative length of stay) for the overall group of valve plus CABG procedures ranged from 0.622 to 0.724, and calibration was excellent.

Conclusions. New STS risk models have been developed for heart valve surgery combined with CABG. These are the first valve plus CABG models that also include risk prediction for individual major morbidities, composite major morbidity or mortality, and short and prolonged length of stay.

(Ann Thorac Surg 2009;88:S43–62)

Risk models for cardiac surgery were first developed almost 2 decades ago, and most of these early models focused on isolated coronary artery bypass graft surgery (CABG) [1–4]. The results of this frequently performed surgical procedure have often been used as the sole marker to assess the quality of care delivered by cardiac surgical programs. Risk-adjusted results for CABG have been used

Drs Shahian, O'Brien, Filardo, Ferraris, Haan, Rich, Normand, DeLong, Shewan, Peterson, Edwards, Anderson, and Ms Dokholyan, have no conflicts of interest to declare regarding this work.

†This author is deceased. Former Chair, Quality, Research and Patient Safety Council, The Society of Thoracic Surgeons, Chicago, IL.

Address correspondence to Dr Shahian, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.055

Abbreviations and Acronyms

AVR	= aortic valve replacement
CABG	= coronary artery bypass graft surgery
MI	= myocardial infarction
MVR	= mitral valve replacement
MVRepair	= mitral valve repair
NCD	= National Adult Cardiac Surgery Database
QMTF	= Quality Measurement Task Force
STS	= The Society of Thoracic Surgeons

for hospital and regional quality improvement initiatives, public reporting, pay for performance reimbursement programs, decision support, patient counseling, and clinical research. Earlier models focused primarily on mortality prediction, but subsequent models have been developed for both risk-adjusted morbidity and length of stay [5].

The other commonly performed category of cardiac surgery consists of operations on the heart valves, either alone or in combination with CABG. Relative to isolated CABG procedures, which are declining in frequency, the proportion of valve cases is steadily increasing. To better assess the overall performance of cardiac surgery programs, to discern the factors that are most significantly related to patient outcomes, and to aid in physician and patient decision-making, risk models have now also been developed for heart valve surgery [6–18].

Unlike risk models for isolated CABG, a relatively standardized procedure, valve surgery encompasses a much more diverse group of operations. There are four cardiac valves, and they may malfunction in a number of quite different ways (eg, stenosis, regurgitation, infection, and so forth). The valves may be repaired or replaced with a wide range of techniques and prosthetics. In some cases, procedures may be performed on multiple valves, or the valve procedure may be combined with CABG.

Given the heterogeneity of heart valve surgery, it is not surprising that a variety of risk-modeling techniques has been applied. At one extreme, the European System for Cardiac Operative Risk Evaluation (EuroSCORE) algorithm, developed by a European consortium, groups all cardiac operations together in a single risk model with indicator variables included to account for valve procedures [14, 18]. Although this approach is simple and easy to apply, recent studies by van Gameren and associates [19] have suggested that a dedicated valve risk model may have better discrimination and calibration than the EuroSCORE algorithm when applied to valve surgery patients. Combined models for aortic and mitral valve procedures with or without CABG have been developed by Jin and colleagues [12] and by Ambler and associates [13]. The 2001 valve models developed by The Society of Thoracic Surgeons (STS) [6] consisted of one model for all isolated valve procedures and one model for valve procedures combined with CABG, and a 2007 risk model derived from the New York Cardiac Surgery Reporting System used a similar stratification [8].

Unified valve models reflect the fact that many risk factors are common to both aortic and mitral valve surgery. They offer simplicity, and they also permit larger sample sizes for development and validation [12]. However, there are significant differences between aortic and mitral valvular disease in both pathophysiology and outcomes, and both also differ substantially from isolated CABG [11]. Some investigators advocate separate aortic and mitral valve models to have more homogeneous patient populations. Examples include models developed by STS, the New York Cardiac Surgery Reporting System, and the Northern New England Cardiovascular Disease Study Group [7, 9, 10]. Some of these models have been developed solely for isolated valve replacement, some have included CABG as a separate predictor variable in the isolated valve model, and some models have focused specifically on valve plus CABG. All these decisions involve a tradeoff—the more homogeneous the study group, the fewer patients are available for model development and validation [12].

Because of the large number of valve surgery patients available for analysis in the STS National Adult Cardiac Surgery Database (NCD), our approach has favored separate models for valve plus CABG versus isolated valve surgery. The STS Quality Measurement Task Force (QMTF) presumes that when adequate numbers of patients are available for study, relatively homogeneous operative categories result in more accurate risk prediction. Furthermore, recent studies by van Gameren and colleagues [19] suggest that the valve plus CABG group may be the most difficult to model accurately, thus meriting its own algorithm.

Several new features were added to the 2008 valve plus CABG models described in this report. First, recognizing that mitral valve repair is often different in both etiology and outcomes than replacement, the QMTF has included interactions between surgery type and several key predictor variables. Fitting a single model with several such interactions is useful. It allows for pooling information across related groups of valve procedures without making an a priori assumption that the effect of key risk factors is constant across these groups. Finally, new models have been developed for specific major complications of each valve plus CABG procedure, as well as for composite morbidity, mortality, and for both short and prolonged postoperative length of stay.

The authors of this report are members of the STS QMTF who were involved in this risk model development project.

Study Population and Endpoints

Our general approaches to variable selection and risk model development have been described in the companion articles on isolated CABG (Part 1) and isolated valve surgery (Part 2). Details specific to the valve plus CABG models are included in this report.

Study Population

The study population for this analysis consisted of single aortic or mitral valve surgical procedures combined with

CABG performed on adult patients between January 1, 2002, and December 31, 2006. Only the following procedures were included: (1) isolated aortic valve replacement (AVR) plus CABG; (2) isolated mitral valve replacement (MVR) plus CABG; and (3) isolated mitral valve repair (MVRepair) plus CABG.

Because of the relatively small number of pulmonic, tricuspid, multiple valve procedures, and aortic repairs, these cases were not included in the current models. Patients undergoing isolated valve surgery without CABG were excluded from the current analysis, but these cases are the focus of a separate model described in Part 2 of this three-part series. Patients with missing sex data (n = 17) were excluded because these patients are not allowed in the analysis dataset used for creating STS database participant feedback reports. Patients on dialysis preoperatively (n = 2,443) were excluded when developing the risk model for prediction of postoperative renal failure. The final study population comprised 101,661 patient operations (66,074 AVR plus CABG; 13,663 MVR plus CABG; and 21,924 MVRepair plus CABG) from 814 STS NCD participating groups.

Characteristics of the study population are summarized in Table 1.

Training and Validation Samples

The study population was randomly divided into a 60% training (development) sample and a 40% test (validation) sample. The development sample was used to identify predictor variables and estimate model coefficients. Data from the validation sample were used to assess model fit, discrimination, and calibration. After choosing variables and assessing model fit, the development and validation samples were subsequently combined, and the final model coefficients were estimated using the combined (development plus validation) data.

Endpoints

In developing the valve plus CABG risk models, we used the same nine endpoints that were analyzed in the STS isolated CABG (Part 1) and the STS isolated valve (Part 2) models. Morbidities in all three models are recorded only in-hospital, in contrast to the operative mortality endpoint defined below (although beginning with version 2.61, sternal infection will be recorded at 30 days): (1) operative mortality: death during the same hospitalization as surgery, regardless of timing or within 30 days of surgery regardless of venue; (2) permanent stroke (CVA): a central neurologic deficit persisting longer than 72 hours; (3) renal failure: a new requirement for dialysis or an increase of the serum creatinine to more than 2.0 mg/dL and double the most recent preoperative creatinine level; (4) prolonged ventilation (> 24 hours); (5) deep sternal wound infection; (6) reoperation for any reason; (7) major morbidity or mortality, a composite defined as the occurrence of any of the above endpoints; (8) prolonged postoperative length of stay (PLOS): length of stay (LOS) more than 14 days (alive or dead); and (4) short postoperative length of stay (SLOS): LOS less than 6 days and patient alive at discharge.

Endpoint frequencies in the study population are presented in Table 2.

Separate Versus Combined Models

Given the variety of approaches used in previous models by STS and other developers, we investigated the option of developing separate models for the AVR plus CABG and MVR plus CABG populations, and we also studied how best to subdivide the mitral plus CABG population into repair versus replacement. Although we had a large study population available, many of the individual outcomes were relatively rare. We were concerned that the number of events would be too small to permit reliable estimation of the model coefficients in separate models for each valve. Thus, in theory, the development of separate custom models for each valve type could be inferior to a single combined model because the custom models would have a smaller sample size and hence larger variance.

As described in detail in Part 2 of this series (isolated valve surgery), we performed preliminary empirical analyses to compare two alternative strategies (separate versus combined AVR plus CABG and MVR/Repair plus CABG) for developing these risk models. We first developed separate models for the three subpopulations (AVR plus CABG, MVR plus CABG, and MVRepair plus CABG), then modeled all three subpopulations together in a single model. In the latter approach, we included several interaction terms to allow the effect of certain risk factors to differ across the specific valve subpopulations. These strategies were used to develop risk models for operative mortality and permanent stroke, using a 60% development sample and a separate 40% validation sample. The performance of the combined model was then assessed separately within each subpopulation and compared to the model that was developed specifically for that subpopulation. In the case of mortality, the combined model had better discrimination (larger c-index) than the corresponding custom model in each of the three subpopulations (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). For stroke, the combined model had better discrimination in two of the three populations (all except AVR plus CABG). Finally, when explained variation was quantified by the generalized R^2 index of Nagelkerke [20], the combined model had greater explained variation than the custom model in each subpopulation for each endpoint. These results provide empirical support for the use of a single model with several interactions, which allows pooling of information across valve groups without assuming that the effect of risk factors is constant.

Selection of Candidate Predictor Variables

The candidate variables for the STS valve plus CABG models were identical to those in the STS isolated valve models, described in Part 2 of this series. They differed from the isolated CABG model variables in the following specific areas: (1) Percutaneous coronary intervention (PCI) occurring 6 hours or less before surgery was present in only 315 patients (0.3%) in the valve plus CABG study population, and was not included as a candidate variable. (2) Infectious endocarditis was not included in the isolated CABG model but was considered for the valve plus CABG model. Although this risk factor was rarely present (0.8% active

Table 1. Distribution of Risk Factors in Overall Study Population 2002 to 2006

	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
Variable	N	%	N	%	N	%	N	%
Demographics								
Age, years								
< 55	6,693	6.6	2,983	4.51	1,309	9.58	2,401	10.95
55–64	17,188	16.9	9,132	13.82	2,790	20.42	5,266	24.02
65–74	33,628	33.1	21,313	32.26	4,667	34.16	7,648	34.88
≥ 75	44,152	43.4	32,646	49.41	4,897	35.84	6,609	30.15
Sex								
Male	65,588	64.5	44,619	67.53	7,348	53.78	13,621	62.13
Female	36,073	35.5	21,455	32.47	6,315	46.22	8,303	37.87
Race								
Caucasian	90,572	89.1	60,121	90.99	11,765	86.11	18,686	85.23
Black	4,534	4.5	2,094	3.17	914	6.69	1,526	6.96
Hispanic	2,487	2.4	1,487	2.25	354	2.59	646	2.95
Asian	1,083	1.1	542	0.82	191	1.40	350	1.60
Other	2,295	2.3	1,402	2.12	331	2.42	562	2.56
Missing	690	0.7	428	0.65	108	0.79	154	0.70
Risk factors								
Body surface area, m^2								
< 1.50	3,340	3.3	1,985	3.00	638	4.67	717	3.27
1.50–1.74	20,779	20.4	12,580	19.04	3,500	25.62	4,699	21.43
1.75–1.99	40,017	39.4	25,814	39.07	5,440	39.82	8,763	39.97
≥ 2.00	36,956	36.4	25,361	38.38	3,996	29.25	7,599	34.66
Missing	569	0.6	334	0.51	89	0.65	146	0.67
Body mass index, kg/m^2								
< 25	29,353	28.9	17,712	26.81	4,787	35.04	6,854	31.26
25–29	39,345	38.7	25,692	38.88	4,951	36.24	8,702	39.69
30–34	21,063	20.7	14,447	21.86	2,507	18.35	4,109	18.74
≥ 35	11,165	11.0	7,785	11.78	1,299	9.51	2,081	9.49
Missing	735	0.7	438	0.66	119	0.87	178	0.81
Diabetes mellitus								
No diabetes	68,112	67.0	44,489	67.33	9,517	69.66	14,106	64.34
Diabetes, noninsulin	23,383	23.0	15,705	23.77	2,642	19.34	5,036	22.97
Diabetes, insulin	9,848	9.7	5,677	8.59	1,463	10.71	2,708	12.35
Diabetes, missing treatment	167	0.2	105	0.16	20	0.15	42	0.19
Missing	151	0.1	98	0.15	21	0.15	32	0.15
Hypertension								
No	22,709	22.3	13,944	21.10	3,482	25.48	5,283	24.10
Yes	78,823	77.5	52,050	78.78	10,163	74.38	16,610	75.76
Missing	129	0.1	80	0.12	18	0.13	31	0.14
Hypercholesterolemia								
No	33,759	33.2	21,248	32.16	5,324	38.97	7,187	32.78
Yes	67,613	66.5	44,649	67.57	8,280	60.60	14,684	66.98
Missing	289	0.3	177	0.27	59	0.43	53	0.24
Past or present smoker								
No	43,687	43.0	29,123	44.08	5,835	42.71	8,729	39.81
Yes	57,813	56.9	36,849	55.77	7,797	57.07	13,167	60.06
Missing	161	0.2	102	0.15	31	0.23	28	0.13
Chronic lung disease								
None	76,803	75.5	50,632	76.63	9,756	71.40	16,415	74.87
Mild	12,157	12.0	7,658	11.59	1,853	13.56	2,646	12.07
Moderate	7,797	7.7	4,720	7.14	1,269	9.29	1,808	8.25
Severe	4,005	3.9	2,463	3.73	658	4.82	884	4.03
Missing	899	0.9	601	0.91	127	0.93	171	0.78

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Peripheral vascular disease								
No	84,183	82.8	54,658	82.72	11,373	83.24	18,152	82.80
Yes	17,294	17.0	11,296	17.10	2,267	16.59	3,731	17.02
Missing	184	0.2	120	0.18	23	0.17	41	0.19
Cerebrovascular disease								
No	83,284	81.9	53,509	80.98	11,304	82.73	18,471	84.25
Yes	18,202	17.9	12,449	18.84	2,335	17.09	3,418	15.59
Missing	175	0.2	116	0.18	24	0.18	35	0.16
CVA								
No CVA	92,527	91.0	60,141	91.02	12,283	89.90	20,103	91.69
Remote CVA (> 2 weeks)	8,461	8.3	5,545	8.39	1,240	9.08	1,676	7.64
Recent CVA (≤ 2 weeks)	348	0.3	184	0.28	88	0.64	76	0.35
CVA, missing timing	114	0.1	62	0.09	23	0.17	29	0.13
Missing	211	0.2	142	0.21	29	0.21	40	0.18
Endocarditis								
No endocarditis	99,517	97.9	65,023	98.41	12,914	94.52	21,580	98.43
Treated endocarditis	1,091	1.1	525	0.79	356	2.61	210	0.96
Active endocarditis	827	0.8	387	0.59	356	2.61	84	0.38
Endocarditis, missing type	24	0.0	11	0.02	8	0.06	5	0.02
Missing	202	0.2	128	0.19	29	0.21	45	0.21
Renal failure								
No	92,592	91.1	60,880	92.14	12,037	88.10	19,675	89.74
Yes	8,888	8.7	5,072	7.68	1,605	11.75	2,211	10.08
Missing	181	0.2	122	0.18	21	0.15	38	0.17
Renal function								
Creatinine < 1.0 mg/dL	30,178	29.7	20,297	30.72	3,672	26.88	6,209	28.32
Creatinine 1.00–1.49 mg/dL	52,008	51.2	34,054	51.54	6,758	49.46	11,196	51.07
Creatinine 1.50–1.99 mg/dL	11,469	11.3	7,151	10.82	1,732	12.68	2,586	11.80
Creatinine 2.00–2.49 mg/dL	2,711	2.7	1,554	2.35	498	3.64	659	3.01
Creatinine ≥ 2.5 mg/dL	1,602	1.6	844	1.28	319	2.33	439	2.00
Dialysis	2,443	2.4	1,364	2.06	482	3.53	597	2.72
Missing	1,250	1.2	810	1.23	202	1.48	238	1.09
Immunosuppressive treatment								
No	98,421	96.8	63,984	96.84	13,211	96.69	21,226	96.82
Yes	2,975	2.9	1,904	2.88	427	3.13	644	2.94
Missing	265	0.3	186	0.28	25	0.18	54	0.25
Previous CV interventions								
Previous CABG surgery								
No	91,657	90.2	59,583	90.18	12,057	88.25	20,017	91.30
Yes	9,615	9.5	6,257	9.47	1,540	11.27	1,818	8.29
Missing	389	0.4	234	0.35	66	0.48	89	0.41
Previous valve surgery								
No	98,737	97.1	64,265	97.26	12,794	93.64	21,678	98.88
Yes	2,540	2.5	1,567	2.37	813	5.95	160	0.73
Missing	384	0.4	242	0.37	56	0.41	86	0.39
Previous other cardiac surgery								
No	98,538	96.9	64,166	97.11	13,181	96.47	21,191	96.66
Yes	2,683	2.6	1,634	2.47	407	2.98	642	2.93
Missing	440	0.4	274	0.41	75	0.55	91	0.42
Number of previous CV surgeries								
No previous CV surgery	89,419	88.0	58,161	88.02	11,530	84.39	19,728	89.98
1 prior CV surgery	10,453	10.3	6,796	10.29	1,799	13.17	1,858	8.47
≥ 2 prior CV surgeries	1,200	1.2	766	1.16	231	1.69	203	0.93
Missing	589	0.6	351	0.53	103	0.75	135	0.62

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Prior PCI								
No PCI	84,553	83.2	55,581	84.12	11,152	81.62	17,820	81.28
PCI ≤ 6 hours	315	0.3	151	0.23	89	0.65	75	0.34
PCI > 6 hours	16,158	15.9	9,946	15.05	2,321	16.99	3,891	17.75
PCI, missing timing	234	0.2	145	0.22	45	0.33	44	0.20
Missing	401	0.4	251	0.38	56	0.41	94	0.43
Preoperative cardiac status								
Acuity status								
Elective	62,298	61.3	43,682	66.11	7,277	53.26	11,339	51.72
Urgent	36,454	35.9	21,414	32.41	5,315	38.90	9,725	44.36
Emergent	2,479	2.4	763	1.15	945	6.92	771	3.52
Emergent salvage	258	0.3	97	0.15	104	0.76	57	0.26
Missing	172	0.2	118	0.18	22	0.16	32	0.15
MI								
No prior MI	68,332	67.2	49,673	75.18	8,056	58.96	10,603	48.36
MI ≤ 21 days	16,934	16.7	9,308	14.09	2,621	19.18	5,005	22.83
MI 8–21 days	3,751	3.7	1,725	2.61	624	4.57	1,402	6.39
MI 1–7 days	10,458	10.3	4,514	6.83	1,741	12.74	4,203	19.17
MI > 6 and < 24 hours	1,113	1.1	367	0.56	341	2.50	405	1.85
MI ≤ 6 hours	531	0.5	178	0.27	192	1.41	161	0.73
MI, missing timing	355	0.3	184	0.28	59	0.43	112	0.51
Missing	187	0.2	125	0.19	29	0.21	33	0.15
Angina								
No	42,542	41.8	28,032	42.43	6,248	45.73	8,262	37.68
Yes	58,967	58.0	37,945	57.43	7,394	54.12	13,628	62.16
Missing	152	0.1	97	0.15	21	0.15	34	0.16
Cardiogenic shock								
No	98,743	97.1	65,219	98.71	12,590	92.15	20,934	95.48
Yes	2,719	2.7	720	1.09	1,055	7.72	944	4.31
Missing	199	0.2	135	0.20	18	0.13	46	0.21
Resuscitation								
No	100,474	98.8	65,522	99.16	13,359	97.78	21,593	98.49
Yes	971	1.0	405	0.61	281	2.06	285	1.30
Missing	216	0.2	147	0.22	23	0.17	46	0.21
Arrhythmia								
No arrhythmia	83,856	82.5	56,040	84.81	9,992	73.13	17,824	81.30
AFib/flutter	13,386	13.2	7,533	11.40	2,940	21.52	2,913	13.29
Heart block	1,975	1.9	1,311	1.98	289	2.12	375	1.71
Sustained VT/VF	1,513	1.5	614	0.93	299	2.19	600	2.74
Arrhythmia, other	483	0.5	305	0.46	63	0.46	115	0.52
Arrhythmia, missing type	242	0.2	135	0.20	59	0.43	48	0.22
Missing	206	0.2	136	0.21	21	0.15	49	0.22
Preoperative IABP								
No	96,136	94.6	64,597	97.76	11,957	87.51	19,582	89.32
Yes	5,205	5.1	1,275	1.93	1,655	12.11	2,275	10.38
Missing	320	0.3	202	0.31	51	0.37	67	0.31
NYHA class								
I	9,839	9.7	6,934	10.49	1,103	8.07	1,802	8.22
II	24,830	24.4	17,808	26.95	2,524	18.47	4,498	20.52
III	42,593	41.9	28,079	42.50	5,458	39.95	9,056	41.31
IV	20,571	20.2	10,808	16.36	3,882	28.41	5,881	26.82
Missing	3,828	3.8	2,445	3.70	696	5.09	687	3.13

Table 1. Continued

	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
Variable	N	%	N	%	N	%	N	%
Congestive heart failure								
No	58,086	57.1	41,984	63.54	5,797	42.43	10,305	47.00
Yes	43,377	42.7	23,953	36.25	7,845	57.42	11,579	52.81
Missing	198	0.2	137	0.21	21	0.15	40	0.18
Number of diseased coronary vessels								
None	2,362	2.3	1,786	2.70	281	2.06	295	1.35
One	22,718	22.3	16,934	25.63	3,040	22.25	2,744	12.52
Two	27,144	26.7	19,014	28.78	3,655	26.75	4,475	20.41
Three	49,060	48.3	28,107	42.54	6,623	48.47	14,330	65.36
Missing	377	0.4	233	0.35	64	0.47	80	0.36
Left main disease ≥ 50%								
No	84,025	82.7	55,292	83.68	11,503	84.19	17,230	78.59
Yes	17,175	16.9	10,512	15.91	2,072	15.17	4,591	20.94
Missing	461	0.5	270	0.41	88	0.64	103	0.47
Ejection fraction, %								
< 25	5,805	5.7	2,199	3.33	640	4.68	2,966	13.53
25–34	10,988	10.8	4,877	7.38	1,566	11.46	4,545	20.73
35–44	14,928	14.7	8,064	12.20	2,487	18.20	4,377	19.96
45–54	20,398	20.1	13,424	20.32	3,048	22.31	3,926	17.91
≥ 55	43,556	42.8	32,973	49.90	5,209	38.12	5,374	24.51
Missing	5,986	5.9	4,537	6.87	713	5.22	736	3.36
Aortic stenosis								
No	42,831	42.1	8,527	12.91	12,974	94.96	21,330	97.29
Yes	58,317	57.4	57,319	86.75	535	3.92	463	2.11
Missing	513	0.5	228	0.35	154	1.13	131	0.60
Mitral stenosis								
No	95,696	94.1	63,862	96.65	11,166	81.72	20,668	94.27
Yes	4,993	4.9	1,542	2.33	2,366	17.32	1,085	4.95
Missing	972	1.0	670	1.01	131	0.96	171	0.78
Tricuspid stenosis								
No	100,093	98.5	65,060	98.47	13,402	98.09	21,631	98.66
Yes	275	0.3	154	0.23	57	0.42	64	0.29
Missing	1,293	1.3	860	1.30	204	1.49	229	1.04
Pulmonic stenosis								
No	99,484	97.9	64,693	97.91	13,348	97.69	21,443	97.81
Yes	122	0.1	85	0.13	14	0.10	23	0.10
Missing	2,055	2.0	1,296	1.96	301	2.20	458	2.09
Aortic insufficiency								
None	57,561	56.6	28,972	43.85	10,821	79.20	17,768	81.04
Trivial	9,243	9.1	6,573	9.95	1,023	7.49	1,647	7.51
Mild	13,828	13.6	11,082	16.77	1,156	8.46	1,590	7.25
Moderate	10,195	10.0	9,581	14.50	232	1.70	382	1.74
Severe	8,686	8.5	8,580	12.99	49	0.36	57	0.26
Missing	2,148	2.1	1,286	1.95	382	2.80	480	2.19
Mitral insufficiency								
None	41,756	41.1	38,790	58.71	1,297	9.49	1,669	7.61
Trivial	7,467	7.3	7,139	10.80	147	1.08	181	0.83
Mild	15,407	15.2	13,485	20.41	584	4.27	1,338	6.10
Moderate	14,987	14.7	4,842	7.33	2,790	20.42	7,355	33.55
Severe	20,516	20.2	527	0.80	8,743	63.99	11,246	51.30
Missing	1,528	1.5	1,291	1.95	102	0.75	135	0.62

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Tricuspid insufficiency								
None	74,774	73.6	49,614	75.09	9,758	71.42	15,402	70.25
Trivial	7,972	7.8	5,454	8.25	839	6.14	1,679	7.66
Mild	11,505	11.3	7,060	10.68	1,631	11.94	2,814	12.84
Moderate	4,119	4.1	1,919	2.90	874	6.40	1,326	6.05
Severe	636	0.6	237	0.36	186	1.36	213	0.97
Missing	2,655	2.6	1,790	2.71	375	2.74	490	2.23
Pulmonic insufficiency								
None	91,715	90.2	59,891	90.64	12,275	89.84	19,549	89.17
Trivial	3,411	3.4	2,122	3.21	442	3.24	847	3.86
Mild	2,065	2.0	1,215	1.84	306	2.24	544	2.48
Moderate	326	0.3	165	0.25	70	0.51	91	0.42
Severe	49	0.0	25	0.04	11	0.08	13	0.06
Missing	4,095	4.0	2,656	4.02	559	4.09	880	4.01

AFib = atrial fibrillation; AVR = aortic valve replacement; CABG = coronary artery bypass graft; CV = cardiovascular; CVA = cerebrovascular accident (stroke); IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; VF = ventricular fibrillation; VT = ventricular tachycardia.

endocarditis) in the overall valve plus CABG population, it was included for consistency with the isolated valve model. Active endocarditis was present in 2.6% of patients undergoing mitral replacement plus CABG. (3) Mitral stenosis was rarely present among isolated CABG patients (0.35%). However, it was not uncommon (4.9%) among patients undergoing valve plus CABG surgery and was included as a candidate variable. It was present in 17.3% of mitral replacements and 5.0% of mitral repairs.

An indicator for valve procedure (AVR, MVR, MVRepair) was included in the combined valve plus CABG model, as previously noted.

Missing Data

Missing data are uncommon in the STS NCD, with a frequency of less than 1% missing for most variables. Model variables with more than 1% missing were ejection fraction

Table 2. Frequency of Endpoints in Overall Study Population 2002 to 2006

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall (AVR + CABG, MVR + CABG, MVRepair + CABG)									
N	101,661	101,661	99,218	101,661	101,661	101,661	101,661	101,661	101,661
Events	6,919	2,935	9,097	21,561	684	12,117	30,580	15,594	22,534
%	6.8	2.9	9.0	21.2	0.7	11.9	30.1	15.3	22.2
AVR + CABG									
N	66,074	66,074	64,710	66,074	66,074	66,074	66,074	66,074	66,074
Events	3,718	1,751	5,032	11,608	394	7,090	17,343	8,412	16,961
%	5.6	2.7	7.6	17.6	0.6	10.7	26.3	12.7	25.7
MVR + CABG									
N	13,663	13,663	13,181	13,663	13,663	13,663	13,663	13,663	13,663
Events	1,590	499	1,829	4,469	114	2,274	5,897	3,277	1,512
%	11.6	3.7	13.6	32.7	0.8	16.6	43.2	24.0	11.1
MVRepair + CABG									
N	21,924	21,924	21,327	21,924	21,924	21,924	21,924	21,924	21,924
Events	1,611	685	2,236	5,484	176	2,753	7,340	3,905	4,061
%	7.4	3.1	10.3	25.0	0.8	12.6	33.5	17.8	18.5

AVR = aortic valve replacement; CABG = coronary artery bypass graft surgery; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Table 3. List of Candidate Variables and Their Coding for STS Valve Plus CABG Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline truncated from below at 50 with knot at 75.
Ejection fraction	Linear; values > 50 mapped to 50
Body surface area[a]	Quadratic polynomial modeled separately for males and females. *Note: BSA < 1.4 and > 2.6 were mapped to those values, respectively.*
Creatinine	Linear (only for patients not on dialysis). *Note: Creatinine < 0.5 and > 5.0 mapped to those values, respectively.*
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval. Modeled as linear across the categories.
Binary variables	
Active infectious endocarditis	Yes/no
Dialysis	Yes/no
Preoperative atrial fibrillation	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Preop IABP or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no MI < 7 days)	Yes/no
Left main disease	Yes/no
Aortic stenosis	Yes/no
Mitral stenosis	Yes/no
Aortic insufficiency	Defined as at least moderate (yes/no)
Mitral insufficiency	Defined as at least moderate (yes/no)
Tricuspid insufficiency	Defined as at least moderate (yes/no)
Categorical variables	
Surgery type	3 groups: AVR + CABG, MVR + CABG, MVRepair + CABG
Chronic lung disease	Modeled as linear across categories (none, mild, moderate, severe)
CVD/CVA	3 groups: no CVD, CVD no CVA, CVD + CVA
Diabetes mellitus	3 groups: insulin diabetes, noninsulin diabetes, other or no diabetes
No. diseased coronary vessels	3 groups: < 2-vessel disease; 2-vessel disease; 3-vessel disease. Modeled as linear across the categories
MI	3 groups: < 24 hours, 1–21 days, > 21 days or no MI. *Note: groups 1 and 2 were subsequently collapsed for some models.*
Race	3 groups: black, Hispanic, other including Caucasian
Status	4 groups: elective, urgent, emergent no resuscitation, salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, ≥ 2 previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF and NYHA IV
Interaction terms	
Age by reoperation[a]	
Age by emergent status[a]	
Surgery type by each of the following:	Age, diabetes, dialysis, creatinine, reoperation, endocarditis, emergent status, CLD, CHF, EF, sex, shock, IABP/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, aortic stenosis.

[a] These variables were forced into each model.

AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association.

(5.9%), New York Heart Association functional class (3.8%), tricuspid insufficiency (2.6%), aortic insufficiency (2.1%), mitral insufficiency (1.5%), and creatinine/dialysis (1.2%).

To make full use of the available data, binary risk factors were modeled as yes versus no or missing. Thus, missing values were analyzed as if the endpoint did not occur. Missing data on categorical variables were imputed to the lowest risk value, which, in most instances, was the mode. Missing data on continuous variables were imputed to the conditional median. For ejection fraction, we conditioned

on congestive heart failure and sex. For body surface area, we conditioned on sex. For serum creatinine, we conditioned on renal failure.

Although multiple imputation is generally preferred on statistical grounds [21], we chose single imputation for this analysis based largely on practical considerations, including computational intensity. Furthermore, the fraction of missing data was small, and single and multiple imputation would give similar results. Finally, multiple imputation is primarily used for calculating appropriate standard error estimates, but an adjustment to the standard errors would not impact our study results or the published risk algorithms. In a separate sensitivity analysis, we compared predicted risk estimates from our final models to risk estimates that were derived from analogous models using multiple instead of single imputation. For each endpoint, the relative difference in predicted risk was less than 6% (eg, an absolute difference of 5.0% versus 5.3%) for all patients in the development and validation samples, and it was less than 2% (eg, an absolute difference of 5.0% versus 5.1%) for 99% of patients. A summary of these analyses including regression coefficients and covariance matrices is available at www.sts.org/riskmodels.

Final Variable Selection Procedure

Variables were initially selected using an automated stepwise model selection algorithm. The stepwise procedure began with a model that included all of the candidate variables except for interaction terms. Age, body surface area, and month of surgery were forced into each model. As in the isolated CABG and isolated valve models described in Parts 1 and 2 of this series, month of surgery was used only to adjust for time trends in the frequency of adverse outcomes over the 5-year study period. We adjusted for this to reduce potential confounding by time trends when estimating regression coefficients for the variables that are of primary interest (ie, patient preoperative risk factors—see example in Part 1). Surgery date was categorized into 6-month intervals and modeled as a linear trend across the ordinal categories. Surgery date is not included in the final risk prediction algorithm, and a patient's predicted risk does not depend on it. The published intercept parameter has been adjusted to incorporate the time trend, and this adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Other variables were selected in a stepwise fashion using a significance criterion of 0.05 for entry and removal. Ordinal categorical variables were initially coded such that removing an indicator variable caused a category to be combined with the lowest risk category (the reference group). In the case of myocardial infarction (MI), there were two outcomes (permanent stroke, prolonged length of stay) in which "MI 1 to 21 days" was retained but "MI less than 24 hours" was removed. For these two cases, the two MI categories were replaced by the single category "MI 21 days or less." The stepwise procedure was performed separately for each endpoint. Multiple interaction terms consisting of predictor variable and surgery type were also evaluated, and two additional interaction terms (age by reoperation and age by emergent status) were forced into the models (see Tables 3 and 5).

The results of this initial selection process were then reviewed by surgeon members of the QMTF for face validity and consistency with previous STS or other valve models: (1) preoperative atrial fibrillation was forced into the model for permanent stroke; (2) an indicator variable for dialysis was forced into any model that included creatinine (this did not apply to the renal failure model, as patients with preoperative dialysis were excluded); (3) sex was forced into all models; and (4) each variable that interacted with surgery group was also included as a main effect.

After validating the model in the 40% validation sample, the development and validation samples were then rejoined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regression model was estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [22]. An independence working correlation matrix was used to apply the generalized estimating equations. With this approach, the estimated regression coefficients were identical to those obtained using ordinary logistic regression, but the standard errors were adjusted to account for the clustered data structure.

Results

Risk Factors, Outcomes, and Predictor Variables

Table 1 presents the distribution of risk factors and endpoints in the overall 2002 to 2006 study population. Because there are three valve plus CABG categories, space limitations prevent display of the bivariate relationships for each predictor variable, endpoint, and valve plus CABG group. These are available upon request from STS.

Table 2 summarizes the overall frequency of adverse outcomes as well as the outcomes for the three major valve groups. Table 3 lists the candidate predictor variables and their coding schemes.

Assessment of Model Fit and Discrimination

The Hosmer-Lemeshow test was not employed to assess overall calibration. Large sample sizes make a significant p value almost inevitable, as all risk models are only approximations of reality [23]. Rather, we assessed calibration graphically by plotting observed versus predicted event rates within deciles of predicted risk in the development and validation samples (Fig 1). These plots were constructed for the overall sample and for subgroups based on surgery type (AVR plus CABG, MVR plus CABG, MVRepair plus CABG); age ($<$ 60, 60 to 79, $\geq$ 80 years); sex (male, female); diabetes mellitus (yes/no); status (elective, nonelective); and ejection fraction ($\leq$ 40, $>$ 40). Because of space constraints, only the overall sample results in the validation sample are presented. Additional results are available at www.sts.org/riskmodels.

In general, the models were well calibrated in the validation sample. The average absolute difference between observed versus predicted event rates across the decile categories ranged from 0.1% for deep sternal wound infec-

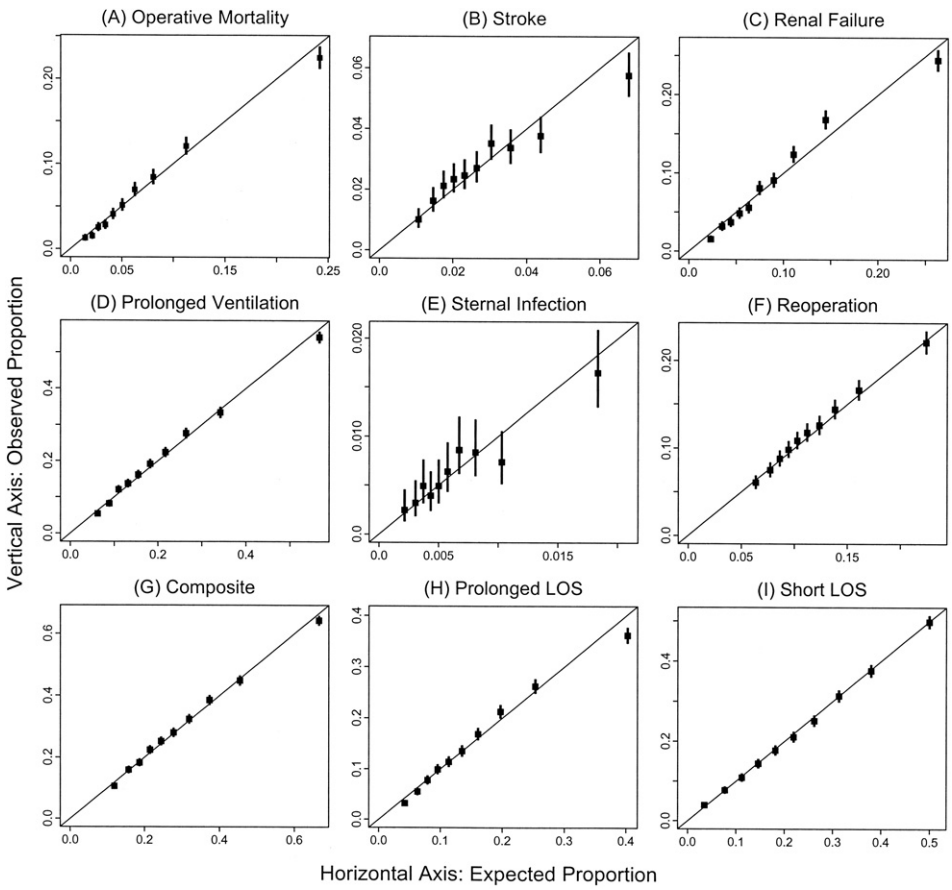


Fig 1. Plots of observed (O) versus expected (E) in validation sample

tion to 0.96% for prolonged length of stay. There was a very slight tendency for the models to overpredict risk in the highest decile. Although perfect prediction would be ideal, a slight overprediction implies that the model will give adequate credit to surgeons who take on patients with several model risk factors.

Discrimination was assessed by determining the c-statistic, also known as the area under the receiver operating characteristic (ROC) curve. Table 4 presents the discrimination of the various models. In the validation sample, the c-index of the overall valve plus CABG operative mortality model was 0.750, and the c-indices of the morbidity models ranged from 0.617 for reoperation to 0.724 for renal failure and short length of stay.

Table 4. Discrimination of Models (C-Index) in Development and Validation Samples

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall									
Development sample	0.754	0.656	0.729	0.730	0.670	0.623	0.704	0.719	0.726
Validation sample	0.750	0.622	0.724	0.720	0.646	0.617	0.698	0.710	0.724
AVR + CABG									
Development sample	0.737	0.648	0.720	0.706	0.639	0.607	0.678	0.705	0.700
Validation sample	0.736	0.609	0.718	0.697	0.657	0.604	0.673	0.699	0.698
MVR + CABG									
Development sample	0.764	0.665	0.712	0.746	0.713	0.608	0.725	0.694	0.726
Validation sample	0.739	0.611	0.701	0.733	0.580	0.599	0.714	0.680	0.733
MVRepair + CABG									
Development sample	0.746	0.650	0.727	0.725	0.692	0.624	0.707	0.712	0.738
Validation sample	0.755	0.652	0.715	0.716	0.644	0.623	0.705	0.702	0.733

AVR = aortic valve replacement; CABG = coronary artery bypass graft; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Table 5. Estimated Odds Ratios for CABG Mortality, Morbidity, and Length of Stay Models

A. Odds ratios for variables that do not interact with surgery group

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.12, 1.29)	1.05 (0.94, 1.17)	1.18 (1.11, 1.26)	1.13 (1.07, 1.19)	NA	1.16 (1.10, 1.22)	1.15 (1.10, 1.20)	1.22 (1.15, 1.28)	0.71 (0.67, 0.75)
BSA 1.6 versus 2.0 among females	1.29 (1.19, 1.39)	1.34 (1.18, 1.52)	0.87 (0.81, 0.94)	1.08 (1.02, 1.14)	0.51 (0.39, 0.67)	1.13 (1.07, 1.23)	1.12 (1.07, 1.18)	0.97 (0.92, 1.03)	1.03 (0.96, 1.10)
BSA 1.6 versus 2.0 among males	1.58 (1.41, 1.77)	1.38 (1.17, 1.64)	1.18 (1.07, 1.31)	1.31 (1.21, 1.41)	0.71 (0.49, 1.03)	1.18 (1.12, 1.34)	1.32 (1.24, 1.41)	1.40 (1.29, 1.52)	0.81 (0.75, 0.88)
BSA 1.8 versus 2.0 among females	1.05 (1.00, 1.10)	1.16 (1.06, 1.26)	0.89 (0.85, 0.93)	0.99 (0.95, 1.02)	0.69 (0.61, 0.77)	1.03 (0.98, 1.06)	1.01 (0.98, 1.04)	0.94 (0.90, 0.97)	1.08 (1.04, 1.12)
BSA 1.8 versus 2.0 among males	1.15 (1.10, 1.20)	1.13 (1.07, 1.20)	1.01 (0.97, 1.05)	1.06 (1.03, 1.09)	0.83 (0.72, 0.95)	1.06 (1.04, 1.11)	1.07 (1.05, 1.10)	1.09 (1.06, 1.12)	0.96 (0.94, 0.99)
BSA 2.2 versus 2.0 among females	1.12 (1.02, 1.22)	0.87 (0.74, 1.02)	1.25 (1.15, 1.35)	1.13 (1.06, 1.20)	1.57 (1.32, 1.89)	1.04 (1.00, 1.17)	1.10 (1.04, 1.17)	1.19 (1.11, 1.27)	0.82 (0.76, 0.89)
BSA 2.2 versus 2.0 among males	1.04 (1.00, 1.09)	0.95 (0.90, 1.01)	1.15 (1.11, 1.18)	1.09 (1.06, 1.11)	1.25 (1.14, 1.37)	1.00 (0.95, 1.01)	1.07 (1.04, 1.09)	1.09 (1.06, 1.12)	0.91 (0.89, 0.93)
CVD with CVA	1.22 (1.11, 1.33)	1.72 (1.52, 1.95)	1.12 (1.04, 1.22)	1.27 (1.19, 1.34)	1.22 (0.95, 1.56)	1.12 (1.04, 1.20)	1.26 (1.20, 1.33)	1.26 (1.18, 1.35)	0.75 (0.70, 0.81)
CVD without CVA	NA	1.28 (1.13, 1.45)	1.14 (1.06, 1.23)	1.10 (1.04, 1.16)	NA	NA	1.11 (1.05, 1.17)	1.11 (1.05, 1.18)	0.85 (0.78, 0.92)
Diabetes, insulin	1.31 (1.20, 1.42)	1.16 (1.03, 1.30)	1.62 (1.52, 1.74)	1.32 (1.25, 1.40)	1.98 (1.59, 2.46)	NA	1.34 (1.28, 1.41)	1.49 (1.40, 1.58)	0.67 (0.62, 0.72)
Diabetes, noninsulin	1.12 (1.05, 1.19)	1.16 (1.06, 1.26)	1.28 (1.21, 1.35)	1.11 (1.07, 1.15)	1.30 (1.10, 1.54)	NA	1.12 (1.08, 1.16)	1.17 (1.12, 1.22)	0.84 (0.81, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	1.15 (1.11, 1.19)	1.20 (1.14, 1.26)	1.17 (1.14, 1.21)	1.19 (1.16, 1.22)	1.28 (1.15, 1.42)	1.09 (1.06, 1.11)	1.16 (1.14, 1.18)	1.13 (1.10, 1.16)	0.82 (0.81, 0.84)
Hypertension	NA	1.19 (1.08, 1.31)	1.25 (1.18, 1.33)	1.10 (1.05, 1.15)	1.33 (1.09, 1.63)	NA	1.12 (1.08, 1.16)	1.08 (1.03, 1.13)	0.92 (0.88, 0.96)
Immunosuppressive treatment	1.35 (1.17, 1.54)	NA	1.30 (1.15, 1.47)	1.28 (1.17, 1.40)	NA	1.27 (1.14, 1.42)	1.26 (1.16, 1.37)	1.22 (1.11, 1.34)	0.75 (0.67, 0.84)
Left main disease	1.12 (1.05, 1.20)	NA	NA	1.06 (1.02, 1.11)	NA	NA	NA	NA	NA
Mitral insufficiency, moderate/severe	NA	NA	NA	NA	NA	NA	1.07 (1.01, 1.12)	NA	NA
Tricuspid insufficiency, moderate/severe	1.27 (1.15, 1.41)	NA	1.25 (1.13, 1.38)	1.15 (1.06, 1.24)	NA	NA	1.14 (1.07, 1.22)	NA	0.79 (0.69, 0.92)
Peripheral vascular disease	1.29 (1.21, 1.37)	1.15 (1.04, 1.27)	1.16 (1.10, 1.23)	1.18 (1.12, 1.24)	NA	1.15 (1.09, 1.22)	1.20 (1.15, 1.25)	1.16 (1.11, 1.22)	NA
Mitral stenosis	1.10 (0.99, 1.24)	NA	NA	NA	NA	NA	NA	1.09 (1.00, 1.18)	NA
MI 1–21 days	1.19 (1.10, 1.28)	NA	1.18 (1.10, 1.26)	1.28 (1.21, 1.35)	NA	NA	1.22 (1.16, 1.28)	NA	NA
MI $\leq$ 21 days[a]	NA	1.22 (1.11, 1.34)	NA	NA	NA	NA	NA	1.16 (1.10, 1.22)	NA
MI $<$ 24 hrs	1.65 (1.42, 1.91)	NA	1.30 (1.10, 1.54)	1.41 (1.23, 1.62)	NA	1.15 (1.00, 1.32)	1.49 (1.30, 1.70)	NA	NA
Time trend per 6-month harvest interval	0.98 (0.96, 0.99)	0.98 (0.97, 1.00)	1.01 (1.00, 1.02)	1.01 (1.00, 1.02)	0.96 (0.93, 0.99)	0.99 (0.98, 1.00)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)
Race black	NA	NA	1.15 (1.03, 1.30)	1.31 (1.19, 1.44)	NA	1.19 (1.06, 1.33)	1.21 (1.11, 1.32)	1.31 (1.19, 1.44)	0.65 (0.58, 0.72)
Race Hispanic	NA	NA	1.20 (1.03, 1.40)	1.17 (1.03, 1.32)	NA	1.08 (0.94, 1.24)	1.15 (1.03, 1.28)	1.13 (0.98, 1.30)	0.85 (0.71, 1.02)
Status, urgent versus elective	1.25 (1.17, 1.34)	NA	1.18 (1.10, 1.26)	1.26 (1.19, 1.33)	NA	1.14 (1.07, 1.21)	1.19 (1.14, 1.25)	1.28 (1.22, 1.35)	0.77 (0.72, 0.81)
Unstable angina	1.11 (1.03, 1.21)	0.89 (0.80, 1.00)	1.12 (1.05, 1.20)	1.06 (0.99, 1.13)	NA	NA	NA	NA	NA

Table 5. Continued

B. Odds ratios for AVR plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.29 (1.20, 1.39)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.20 (1.16, 1.24)	1.37 (1.32, 1.42)	0.74 (0.72, 0.77)
Age 70 versus 50 (no reop, elective)	1.67 (1.45, 1.92)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.44 (1.36, 1.54)	1.86 (1.73, 2.01)	0.55 (0.52, 0.59)
Age 80 versus 50 (no reop, elective)	2.47 (2.08, 2.94)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	1.86 (1.73, 2.01)	2.67 (2.46, 2.91)	0.33 (0.30, 0.36)
CHF, not NYHA IV	1.24 (1.14, 1.34)	0.98 (0.88, 1.09)	1.19 (1.11, 1.28)	1.22 (1.16, 1.29)	NA	NA	1.14 (1.08, 1.19)	1.30 (1.23, 1.38)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.48 (1.34, 1.64)	1.15 (1.00, 1.32)	1.35 (1.24, 1.48)	1.47 (1.36, 1.59)	NA	1.16 (1.08, 1.24)	1.36 (1.27, 1.45)	1.49 (1.39, 1.60)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	2.26 (2.13, 2.40)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.62 (0.58, 0.67)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.65 (1.41, 1.92)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.30 (0.25, 0.37)
EF per 10-unit decrease	1.10 (1.06, 1.15)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.87 (0.84, 0.89)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.41 (1.25, 1.58)	0.56 (0.48, 0.65)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.17 (0.92, 1.50)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	1.79 (1.50, 2.15)	1.45 (1.29, 1.63)	NA
Female versus male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.52 (1.44, 1.61)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.61 (0.57, 0.64)
Active infectious endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	1.81 (1.41, 2.32)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.26 (1.22, 1.30)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.26 (1.22, 1.30)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.83 (1.58, 2.11)	NA	1.39 (1.16, 1.67)	1.50 (1.32, 1.69)	1.55 (1.33, 1.81)	0.67 (0.58, 0.77)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	2.19 (1.80, 2.65)	NA	1.48 (1.15, 1.92)	1.77 (1.51, 2.06)	1.65 (1.34, 2.03)	0.53 (0.43, 0.65)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.33 (0.22, 0.50)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.18 (0.09, 0.34)

Table 5. Continued

C. Odds ratios for MVR plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.51 (1.39, 1.64)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.27 (1.21, 1.32)	1.37 (1.32, 1.42)	0.68 (0.64, 0.72)
Age 70 versus 50 (no reop, elective)	2.28 (1.94, 2.68)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.60 (1.47, 1.75)	1.86 (1.73, 2.01)	0.46 (0.41, 0.52)
Age 80 versus 50 (no reop, elective)	3.95 (3.17, 4.93)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	2.18 (1.92, 2.48)	2.67 (2.46, 2.91)	0.25 (0.21, 0.30)
CHF, not NYHA IV	0.91 (0.80, 1.03)	0.80 (0.64, 0.99)	0.92 (0.82, 1.03)	1.02 (0.93, 1.11)	NA	NA	0.94 (0.87, 1.02)	1.03 (0.94, 1.12)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.09 (0.95, 1.24)	0.93 (0.75, 1.17)	1.04 (0.92, 1.19)	1.22 (1.10, 1.35)	NA	1.16 (1.08, 1.24)	1.13 (1.03, 1.23)	1.17 (1.06, 1.30)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	1.82 (1.66, 2.01)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.66 (0.57, 0.78)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.21 (0.95, 1.55)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.30 (0.18, 0.48)
EF per 10-unit decrease	1.23 (1.16, 1.30)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.89 (0.82, 0.95)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.29 (1.14, 1.46)	0.51 (0.39, 0.65)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.21 (0.97, 1.50)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	2.76 (2.22, 3.42)	1.45 (1.29, 1.63)	NA
Female versus male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.17 (1.08, 1.28)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.66 (0.59, 0.74)
Active infectious endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	2.08 (1.62, 2.67)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.18 (1.12, 1.24)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.20 (1.14, 1.26)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.38 (1.19, 1.61)	NA	1.15 (0.95, 1.38)	1.50 (1.32, 1.69)	1.30 (1.10, 1.53)	0.81 (0.66, 0.99)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	1.66 (1.35, 2.03)	NA	1.22 (0.95, 1.56)	1.77 (1.51, 2.06)	1.38 (1.12, 1.71)	0.64 (0.50, 0.82)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.26 (0.16, 0.43)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.14 (0.07, 0.27)

Table 5. Continued

D. Odds ratios for MVRepair plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.46 (1.36, 1.57)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.28 (1.23, 1.33)	1.37 (1.32, 1.42)	0.66 (0.63, 0.69)
Age 70 versus 50 (no reop, elective)	2.14 (1.86, 2.46)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.63 (1.51, 1.76)	1.86 (1.73, 2.01)	0.44 (0.40, 0.48)
Age 80 versus 50 (no reop, elective)	3.60 (2.97, 4.33)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	2.23 (2.00, 2.49)	2.67 (2.46, 2.91)	0.23 (0.20, 0.27)
CHF, not NYHA IV	0.96 (0.85, 1.09)	1.05 (0.90, 1.23)	0.99 (0.88, 1.10)	1.10 (1.02, 1.19)	NA	NA	1.06 (0.99, 1.14)	1.17 (1.08, 1.26)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.16 (1.02, 1.32)	1.23 (1.04, 1.46)	1.12 (0.99, 1.27)	1.32 (1.21, 1.44)	NA	1.16 (1.08, 1.24)	1.27 (1.17, 1.37)	1.33 (1.22, 1.45)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	1.87 (1.72, 2.04)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.59 (0.53, 0.67)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.88 (1.52, 2.31)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.35 (0.24, 0.49)
EF per 10-unit decrease	1.09 (1.04, 1.15)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.84 (0.81, 0.87)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.56 (1.40, 1.73)	0.52 (0.44, 0.62)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.69 (1.41, 2.01)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	2.17 (1.81, 2.60)	1.45 (1.29, 1.63)	NA
Female vs male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.25 (1.15, 1.36)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.60 (0.55, 0.66)
Active infectious Endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	2.98 (1.86, 4.77)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.21 (1.16, 1.27)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.16 (1.10, 1.21)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.55 (1.32, 1.82)	NA	1.49 (1.23, 1.82)	1.50 (1.32, 1.69)	1.32 (1.10, 1.58)	0.80 (0.68, 0.95)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	1.86 (1.53, 2.26)	NA	1.59 (1.20, 2.11)	1.77 (1.51, 2.06)	1.41 (1.11, 1.79)	0.63 (0.51, 0.79)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.43 (0.29, 0.66)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.23 (0.12, 0.44)

[a] For CVA and PLOS, MI coded ≤ 21 days; for all other endpoints, MI coded < 24 hrs or 1 to 21 days. [b] Variable interacts with age. Reported odds ratio represents effect of risk factor for patients aged 50 years old.

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; Mort = mortality; NA = not applicable; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Odds Ratios

Table 5 presents the odds ratios and 95% confidence intervals (CI) derived from these models. "Not applicable" indicates that those predictors were not included in a particular risk model.

Odds ratios that do not interact with surgery type are summarized in Part A of Table 5. Several variables interact with surgery type, and the odds ratios for these variables differ for some of the endpoints depending on the specific type of surgery, as summarized in Tables 5B, C, and D (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). For example, in the model for prolonged length of stay, the odds ratio for active endocarditis is 1.81 (95% CI: 1.41 to 2.32) for AVR plus CABG; 2.08 (95% CI: 1.62 to 2.67) for MVR plus CABG; and 2.98 (95% CI: 1.86 to 4.77) for MVRepair plus CABG.

Final Model Intercept and Coefficients

The algorithms for calculating predicted risk values, including the intercepts and regression coefficients, are presented in the Appendix.

Limitations

The limitations of the STS valve plus CABG models are similar to those discussed in Part 1 of this series.

Conclusion

A new STS model has been developed for valve surgery combined with CABG. This model includes specific indicator variables for each major type of valve plus CABG procedure (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). Models have been developed for operative mortality, individual morbidity endpoints, a composite morbidity or mortality endpoint, and short and prolonged postoperative length of stay. Overall model performance is excellent.

References

1. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA 1990;264:2768–74.
2. Edwards FH, Clark RE, Schwartz M. Coronary artery bypass grafting: the Society of Thoracic Surgeons National Database experience. Ann Thorac Surg 1994;57:12–9.
3. O'Connor GT, Plume SK, Olmstead EM, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. Circulation 1992;85:2110–8.
4. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. Ann Thorac Surg 2004; 78:1868–77.
5. Shroyer AL, Coombs LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. Ann Thorac Surg 2003;75:1856–64.
6. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
7. Jamieson WR, Edwards FH, Schwartz M, Bero JW, Clark RE, Grover FL. Risk stratification for cardiac valve replacement. National Cardiac Surgery Database. Database Committee of The Society of Thoracic Surgeons. Ann Thorac Surg 1999;67:943–51.
8. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
9. Hannan EL, Racz MJ, Jones RH, et al. Predictors of mortality for patients undergoing cardiac valve replacements in New York State. Ann Thorac Surg 2000;70:1212–8.
10. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
11. Nowicki ER. What is the future of mortality prediction models in heart valve surgery? Ann Thorac Surg 2005;80:396–8.
12. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
13. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
14. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
15. Gardner SC, Grunwald GK, Rumsfeld JS, et al. Comparison of short-term mortality risk factors for valve replacement versus coronary artery bypass graft surgery. Ann Thorac Surg 2004;77:549–56.
16. Grover FL, Edwards FH. Similarity between the STS and New York State databases for valvular heart disease. Ann Thorac Surg 2000;70:1143–4.
17. Rankin JS, Hammill BG, Ferguson TB Jr, et al. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg 2006;131:547–57.
18. Roques F, Nashef SA, Michel P. Risk factors for early mortality after valve surgery in Europe in the 1990s: lessons from the EuroSCORE pilot program. J Heart Valve Dis 2001;10:572–7.
19. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85:921–30.
20. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
21. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2002.
22. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.
23. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Plus CABG Models

For each endpoint, the formula for calculating a patient's predicted risk of the endpoint has the form:

$$\text{Predicted Risk} = \frac{e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}{1+e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}$$

where $x_1, x_2, \ldots, x_n$ denote patient preoperative risk factors (eg, quantitative variables such as age, and comorbidities coded as 1=present, 0=absent); and $\beta_0, \beta_1, \ldots, \beta_n$ denote regression coefficients (numerical constants). Regression coefficients for each endpoint are presented in Appendix Table 1. The variables $x_1, x_2, \ldots, x_n$ are the same for each endpoint and are defined in Appendix Table 2. The regression coefficient for the time trend is not presented. Instead, the intercept has been adjusted to incorporate the time trend. This adjusted intercept reflects the baseline risk for a reference period of July–December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−5.24391	−5.14546	−5.32535	−3.63438	−6.50043	−3.16980	−2.99714	−4.15892	1.18582
Preoperative AFib	0.18430	0.04634	0.16567	0.12059	0.00000	0.14910	0.13766	0.19656	−0.34095
Age function 1	0.02560	0.02487	0.03268	0.02106	0.00545	0.01715	0.01838	0.03115	−0.02970
Age function 3	0.02758	−0.00709	0.00671	0.00791	−0.00985	−0.00021	0.01425	0.00985	−0.04542
Age by reop function	−0.00861	0.00458	0.00077	−0.00673	0.00314	−0.00399	−0.00202	−0.00678	0.00656
Age by status function	−0.00507	−0.01979	−0.00178	−0.00750	0.01627	−0.00029	0.00229	−0.02247	0.00692
Age by MVR function	0.01564	0.00000	0.00000	0.00000	0.00000	0.00000	0.00527	0.00000	−0.00866
Age by MVRepair function	0.01240	0.00000	0.00000	0.00000	0.00000	0.00000	0.00599	0.00000	−0.01159
BSA function 1	−1.14176	−0.81169	−0.41848	−0.66843	0.86401	−0.51266	−0.70411	−0.84204	0.51295
BSA function 2	2.25471	0.94689	1.84088	1.80467	0.42453	0.70024	1.70623	2.10402	−1.66758
CHF but not NYHA IV	0.21206	−0.01726	0.17460	0.20063	0.00000	0.00000	0.12880	0.26291	−0.17652
CHF and NYHA IV	0.39457	0.14109	0.30146	0.38383	0.00000	0.14499	0.30567	0.39791	−0.31077
CHF by MVR function	−0.31077	−0.20917	−0.25767	−0.18455	0.00000	0.00000	−0.18635	−0.23729	0.00000
CHF by MVRepair function	−0.24791	0.06897	−0.18667	−0.10484	0.00000	0.00000	−0.06920	−0.10954	0.00000
CLD function	0.17713	0.00000	0.11379	0.23345	0.27571	0.09280	0.16523	0.22999	−0.19234
CLD by MVR function	0.00000	0.00000	0.00000	−0.06780	0.00000	0.00000	0.00000	−0.04591	0.00000
CLD by MVRepair function	0.00000	0.00000	0.00000	−0.04014	0.00000	0.00000	0.00000	−0.08501	0.00000
Creatinine function 1	0.44794	0.23545	0.81612	0.38147	0.00000	0.24620	0.51256	0.41472	−0.47658
Creatinine by MVR function	0.00000	0.00000	−0.21574	0.00000	0.00000	0.00000	0.00000	0.00000	0.06652
Creatinine by MVRepair function	0.00000	0.00000	−0.18787	0.00000	0.00000	0.00000	0.00000	0.00000	−0.04407
CVD without prior CVA	0.00000	0.24847	0.13299	0.09769	0.00000	0.00000	0.10255	0.10601	−0.16643
CVD and prior CVA	0.19754	0.54344	0.11571	0.23581	0.19686	0.10974	0.23332	0.23319	−0.28560
Diabetes, noninsulin	0.11060	0.14576	0.24490	0.10365	0.26281	0.00000	0.11462	0.15846	−0.17020
Diabetes, insulin	0.26870	0.14582	0.48504	0.27893	0.68330	0.00000	0.29508	0.39583	−0.40448
Dialysis	1.61151	0.58833	0.00000	1.20290	0.61527	0.74332	1.25181	1.29747	−1.67728
Dialysis by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	−0.30339	0.00000	0.00000	0.04745
Dialysis by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.13058	0.00000	0.00000	0.09778
Ejection fraction function	0.00989	0.00000	0.00534	0.01113	0.00000	0.00703	0.01061	0.00995	−0.01440
EF by MVR function	0.01056	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00228
EF by MVRepair function	−0.00117	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.00309
Endocarditis, active	0.71327	0.60657	0.41797	0.67172	0.00000	0.44757	0.74858	0.59333	−1.27854
Endocarditis by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.13795	0.00000
Endocarditis by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.49934	0.00000
Female	0.30852	0.17170	0.16668	0.41874	0.10654	−0.08221	0.18594	0.26947	−0.50044
Female by MVR function	0.00000	0.00000	0.00000	−0.25972	0.00000	0.00000	0.00000	0.00000	0.08895
Female by MVRepair function	0.00000	0.00000	0.00000	−0.19373	0.00000	0.00000	0.00000	0.00000	−0.00229
Female by BSA function 1	0.51233	0.07575	0.76032	0.48032	0.80594	0.16701	0.41581	0.91055	−0.59086
Female by BSA function 2	−0.27980	−0.88628	−0.57622	−0.49740	0.58767	0.52524	−0.40427	−0.78096	0.15748
Hypertension	0.00000	0.17080	0.22638	0.09581	0.28851	0.00000	0.11445	0.07602	−0.08668

Appendix Table 1. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
IABP or inotropes	0.36025	0.00000	0.23674	0.77918	0.00000	0.15075	0.56477	0.34008	−0.58536
IABP by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.08732	−0.09462
IABP by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10281	−0.06743
Immunosuppressive treatment	0.29654	0.00000	0.26400	0.24814	0.00000	0.24041	0.23332	0.19750	−0.28819
Insufficiency, mitral	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.06414	0.00000	0.00000
Insufficiency, tricuspid	0.24006	0.00000	0.22040	0.13606	0.00000	0.00000	0.13318	0.00000	−0.23141
Left main disease	0.11450	0.00000	0.00000	0.06181	0.00000	0.00000	0.00000	0.00000	0.00000
MI 1–21 days	0.17038	0.00000	0.16476	0.24560	0.00000	0.00000	0.19751	0.00000	0.00000
MI ≤ 21 days	0.00000	0.19671	0.00000	0.00000	0.00000	0.00000	0.00000	0.14599	0.00000
MI < 24 hours	0.49918	0.00000	0.26240	0.34321	0.00000	0.13716	0.39731	0.00000	0.00000
MVR	0.14888	0.32659	0.90926	0.76504	0.28437	0.41642	0.41322	0.73530	−0.82339
MVRepair	−0.07374	0.06933	0.51275	0.28204	0.19499	0.07390	−0.03949	0.30384	−0.03552
No. diseased coronary vessel function	0.13746	0.18243	0.15791	0.17277	0.24582	0.08187	0.14767	0.12474	−0.19250
Peripheral vascular disease	0.25173	0.13776	0.14995	0.16591	0.00000	0.14312	0.18062	0.14863	0.00000
Race black	0.00000	0.00000	0.14301	0.26900	0.00000	0.17364	0.19182	0.26856	−0.43385
Race Hispanic	0.00000	0.00000	0.18384	0.15363	0.00000	0.08065	0.13561	0.12286	−0.15901
Reop, 1 previous operation	0.78624	0.00000	0.25782	0.60179	0.00000	0.33209	0.40293	0.43757	−0.39723
Reop, ≥ 2 previous operations	0.90015	0.00000	0.38499	0.78263	0.00000	0.39502	0.56875	0.50334	−0.63237
Reop by MVR function	0.00000	0.00000	0.00000	−0.27846	0.00000	−0.19608	0.00000	−0.17836	0.18262
Reop by MVRepair function	0.00000	0.00000	0.00000	−0.16306	0.00000	0.06985	0.00000	−0.16007	0.17613
Shock	0.51917	0.17321	0.15810	0.65653	0.00000	0.21271	0.58409	0.36987	0.00000
Shock by MVR function	0.00000	0.00000	0.02883	0.00000	0.00000	0.00000	0.43045	0.00000	0.00000
Shock by MVRepair function	0.00000	0.00000	0.36429	0.00000	0.00000	0.00000	0.19084	0.00000	0.00000
Status urgent	0.22591	0.00000	0.16451	0.22905	0.00000	0.12800	0.17511	0.24758	−0.26626
Status emergent	0.75852	0.79460	0.56854	0.99818	0.00000	0.34063	0.77631	1.00162	−1.09633
Status salvage	1.51811	0.95665	0.61798	0.75178	0.00000	0.00000	1.20732	0.56482	−1.72252
Status by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.25083
Status by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.25943
Stenosis, mitral	0.09879	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.08469	0.00000
Unstable angina	0.10722	−0.11292	0.11597	0.05762	0.00000	0.00000	0.00000	0.00000	0.00000

Afib = atrial fibrillation; BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Appendix Table 2. Definition of Variables Appearing in STS 2008 Valve Plus CABG Models

Variable	Definition
Intercept	= 1 for all patients
Preoperative AFib	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine by MVR function	= Creatinine function 1 if valve operation is MVR, = 0 otherwise
Creatinine by MVRepair function	= Creatinine function 1 if valve operation is MVRepair, = 0 otherwise
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
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
EF by MVR function	= Ejection fraction function if valve operation is MVR, = 0 otherwise
EF by MVRepair function	= Ejection fraction function if valve operation is MVRepair, = 0 otherwise
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Endocarditis by MVR function	= 1 if patient has active endocarditis and valve operation is MVR, = 0 otherwise
Endocarditis by MVRepair function	= 1 if patient has active endocarditis and valve operation is MVRepair, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP or inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP or inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient has received immunosuppressive therapy within 30 days, = 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 1–21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI ≤ 21 days[a]	= 1 if patient has history of MI within 21 days prior to surgery, = 0 otherwise (for CVA and PLOS; coded as < 24 hours and 1–21 days for others)
MI < 24 hours	= 1 if history of MI < 24 hours prior to surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased coronary vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 0 otherwise

Appendix Table 2. Continued

Variable	Definition
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
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
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not rescuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis, mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina and no MI within 7 days of surgery, = 0 otherwise

[a] MI coded ≤ 21 days for CVA and PLOS endpoints; for all other endpoints, coded as < 24 hours and 1 to 21 days.

Note: See www.sts.org for exact definitions of terms used above.

BSA = body surface area; CABG = coronary artery bypass graft surgery; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

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/management 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:**	**A** Y☐ N☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and	**B**

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	Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement	**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: national, Population: regional/network, Population: states, Population: counties or cities

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."

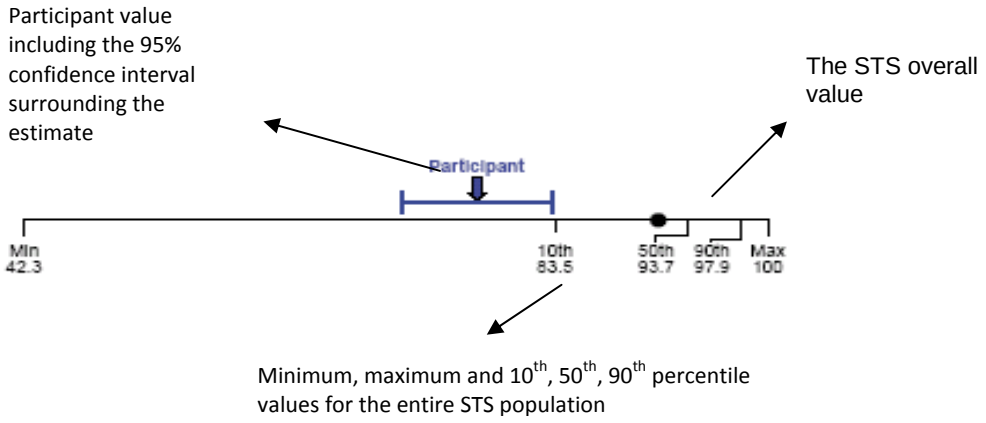
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

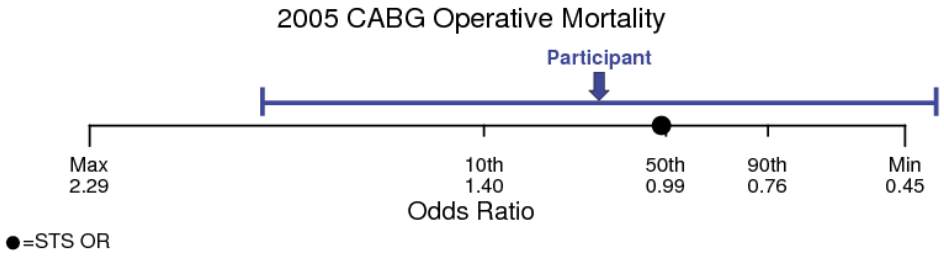
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

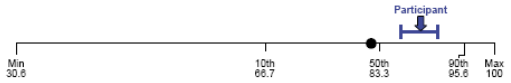
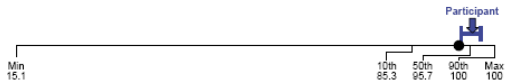
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

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 #:** 1501 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Operative Mortality for Mitral Valve (MV) Repair
De.2 **Brief description of measure***:* Percent of patients undergoing MV Repair who die, including both 1) all deaths occurring during the hospitalization in which the procedure was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure
1.1-2 **Type of Measure:** Outcome 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-634267380312545014.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:** Public reporting, Internal quality improvement	**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**: Safety: All hospitals will reduce preventable and premature hospital-level mortality rates to best-in-class.	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Affects large numbers, Leading cause of morbidity/mortality, High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Monitoring the outcomes of mitral valve repair procedures influences the standard of practice for all participating cardiac surgeons and CV programs by establishing a contemporary, real-life performance benchmarks. The basic requirement of self-reporting encourages all participants to provide optimal patient care and improve outcomes so as to maintain public credibility. 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Petyerson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892. - Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137. - Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467. - Iribarne A, Russo MJ, Easterwood R et al Minimally invasive versus sternotomy approach for mitral	**1a** C☐ P☐ M☐ N☐

valve surgery: a propensity analysis. Ann Thorac Surg. 2010; 90: 1471 – 1477
- LaPar DJ, Hennessy S, Fonner E, et al. Does urgent or emergent status influence choice in mitral valve operations? An analysis of outcomes from the Virginia Cardiac Surgery Quality Initiative. 2010; 90: 153 – 60
- Umakanthan R, Petracek MR, Leacche M et al, Minimally invasive right lateral thoracotomy without aortic cross-clamping: an attractive alternative to repeat sternotomy for reoperative mitral valve surger;y. J Heart Valve Dis. 2010; 19: 236-43

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** The primary benefits of monitoring outcomes of mitral valve repair surgery will be to stimulate technical and surgical care refinements to further reduce the intrinsic operative mortality and morbidity risk of this procedure. Furthermore, defining current thresholds of unacceptable performance will alert both patients and providers to the need for improvement.

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

Dates: January 1, 2005-December 31, 2009

Analysis includes 143 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data (not restricted to this measure) to STS for at least 36 months in 2005-2009.

1b.3 **Citations for data on performance gap:**

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

1b.5 **Citations for data on Disparities:**

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*)**:** Separating institutional performance into quartiles allows comparison of best and worst performing institutions with outcomes published in the current literature. As the new references cited above point out, the general performance of mitral valve surgery in the US today is congruent with results reported in the literature.

1c.2-3. **Type of Evidence:** Observational study, Expert opinion, Systematic synthesis of research, Other
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*)**:**
Institutional care processes and technical competence both influence the success and survival of mitral valve repair patients. Core competence in intra-operative echocardiography, staff surgeon ability to safely perform mitral valve repair and post-op care capabilities all determine measured outcomes.

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
C☐
P☐
M☐
N☐

1c.8 Citations for Evidence (*other than guidelines*): - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Petyerson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892. - Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137. - Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467. - Iribarne A, Russo MJ, Easterwood R et al Minimally invasive versus sternotomy approach for mitral valve surgery: a propensity analysis. Ann Thorac Surg. 2010; 90: 1471 – 1477 - LaPar DJ, Hennessy S, Fonner E, et al. Does urgent or emergent status influence choice in mitral valve operations? An analysis of outcomes from the Virginia Cardiac Surgery Quality Initiative. 2010; 90: 153 – 60 - Umakanthan R, Petracek MR, Leacche M et al, Minimally invasive right lateral thoracotomy without aortic cross-clamping: an attractive alternative to repeat sternotomy for reoperative mitral valve surger;y. J Heart Valve Dis. 2010; 19: 236-43 **1c.9 Quote the Specific guideline recommendation** (*including guideline number and/or page number*): N/A **1c.10 Clinical Practice Guideline Citation:** **1c.11 National Guideline Clearinghouse or other URL:** **1c.12 Rating of strength of recommendation** (*also provide narrative description of the rating and by whom*): **1c.13 Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)*: **1c.14 Rationale for using this guideline over others:**	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
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*): Number of patients undergoing MV Repair who die, including both 1) all deaths occurring during the hospitalization in which the operation was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure	**2a-specs** **C☐** **P☐** **M☐** **N☐**

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*):
During hospitalization regardless of length of stay or within 30 days of surgery if discharged

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):
Number of isolated MV Repair procedures with an operative mortality;

Number of isolated MV Repair procedures in which Mortality [Mortalty (STS Adult Cardiac Surgery Database Version 2.73)] and Mortality Operative Death (MtOpD) are marked "yes." Operative mortality is further verified by the following variables: Mortality Status at 30 days (Mt30Stat), Mortality Date (MtDate), Mortality Discharge Status (MtDCStat)

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*):
All patients undergoing isolated MV Repair surgery

2a.5 Target population gender: Female, Male
2a.6 Target population age range: 18 and older

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

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):
Number of isolated MV Repair procedures;

Isolated MV Repair is determined as a procedure for which all of the following apply:
- OpValve is marked "Yes"
- VSMV is marked "Yes"
- VSMVPr is marked "Repair"
- (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
- OpCAB, VSAV, VSAVPr, ResectSubA, OpTricus, OpPulm, OpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc , OCTumor, OCPulThromDis, OCarOthr are all marked "no" or "missing"

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*):

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: Case-mix adjustment

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

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment 2a.15 Detailed Risk Model-634267381711241302.pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
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)*:
Participant specific OR and their 95% CI were estimated in the hierarchical model. These model-based estimates were used to control variation due to random statistical fluctuations while estimating true signal variation. A 95% CI excluding zero indicates the participant's performance is significantly lower or higher than an "average" STS participant.

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-December of 2010)---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-December of 2010

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

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)*: STS Adult Cardiac Surgery Database - Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009. Excluded from analysis are participants that did not submit results for both time periods. Because database participants can change their underlying care processes at any time, we would not expect perfect correlation between two sets of results from even proximate time periods.

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

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

2c C☐ P☐ M☐ N☐

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 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. **2c.3 Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*: Mortality Operative Death: 100.0% agreement rate	
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)*: Please see Risk Adjustment Type section above **2e.2 Analytic Method** *(type of risk adjustment, analysis, & rationale)*: Detailed information regarding the risk adjustment model can be found in the attachment: O´Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S23-42. **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)*: 143 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009; January 1, 2005-December 31, 2009 **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*:	**2f** C☐ P☐ M☐ N☐

We calculated the risk adjusted event rate with the participant's Odds Ratio (OR) estimate and the overall STS event rate. Therefore, the risk adjusted rate is closely related to OR estimate. If OR > 1, then the participant's risk adjusted rate will be greater than the overall STS event rate; if OR < 1, then the participant's risk adjusted rate will be smaller than the overall STS event rate. The statistical significance is defined by the 95% confidence interval (CI) or the OR estimate. If the 95% CI for a participant's OR includes the null value 1.0, then we cannot distinguish this participant's performance from the STS average - either the participant's performance was close to average or else the participant's sample size was too small to make a reliable inference. Otherwise, if the 95% CI falls to the right of 1.0, then the participant's performance is considered significantly lower than the average STS results; if the 95% CI falls to the left of 1.0, then the participant's performance is considered significantly higher than the average STS results. 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)*: 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)*: 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	**3a** C☐ P☐ M☐ N☐

STS plans to publicly report more measures in the future. There is no definite date yet assigned to this measure; however, STS staff and surgeon leadership have engaged in initial internal STS discussions regarding this matter. 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***):** **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:**	**3c** C☐ P☐ M☐

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:	**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	**4d** **C☐** **P☐** **M☐** **N☐**

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. Please see IFMC audit results.	
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 surgeon 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: This measure has been separated from NQF #121.
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?

Ad.10 Copyright statement/disclaimers:

Ad.11 -13 Additional Information web page URL or attachment: Attachment 1501 Sections 2a.14, 1b.2, 2b.3, 2f.3, 3a.6.pdf

Date of Submission (*MM/DD/YY*)**:** 01/12/2011

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 valve procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest four and a half years of data (January 2006 – 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 Isolated Valve Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
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
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 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 ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 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
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Stenosis mitral	= 1 if patient has mitral 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; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention;

The final patient-level model was built by step-wise selection method with several variables decided by surgeon panel 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	*Risk-Adjusted Operative Mortality for Mitral Valve Repair Surgery*
N	143
Mean	1.0
1st	4.2
5th	2.2
10th	1.8
25th	1.4
Median	0.9
75th	0.5
90th	0.4
95th	0.3
99th	0.3
Outlier	10 (7.0)
High	3
Low	7

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)

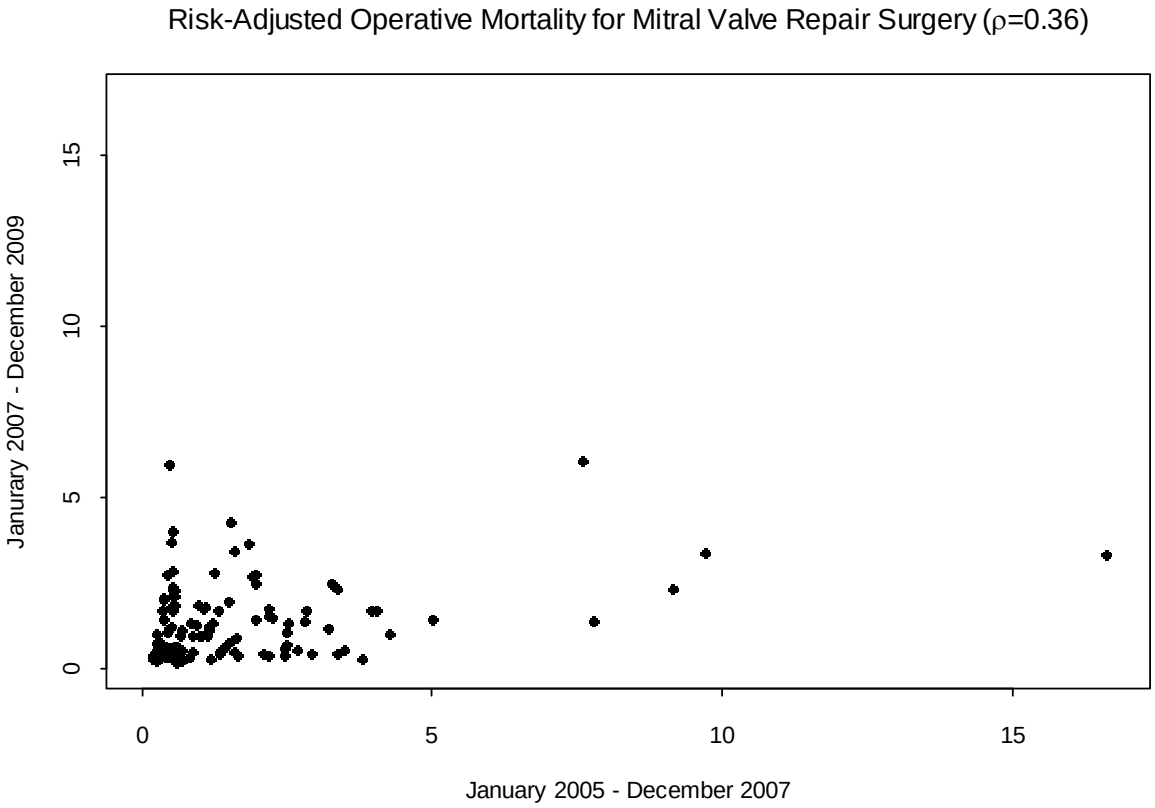
In the above equation, "Participant's Expected Event Rate" is calculated with the participant's actual OR, and "Participant's Expected Event Rate Assuming Its Performance = STS Average Performance" is calculated by assuming the participant's OR = 1 (i.e. no difference in performance from the STS average).

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Repair Surgery*
N	143
Mean	1.6
1st	0.4
5th	0.5
10th	0.6
25th	0.8

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Repair Surgery*
Median	1.4
75th	2.1
90th	2.7
95th	3.2
99th	5.9
Outlier	10 (7.0)
High	3
Low	7

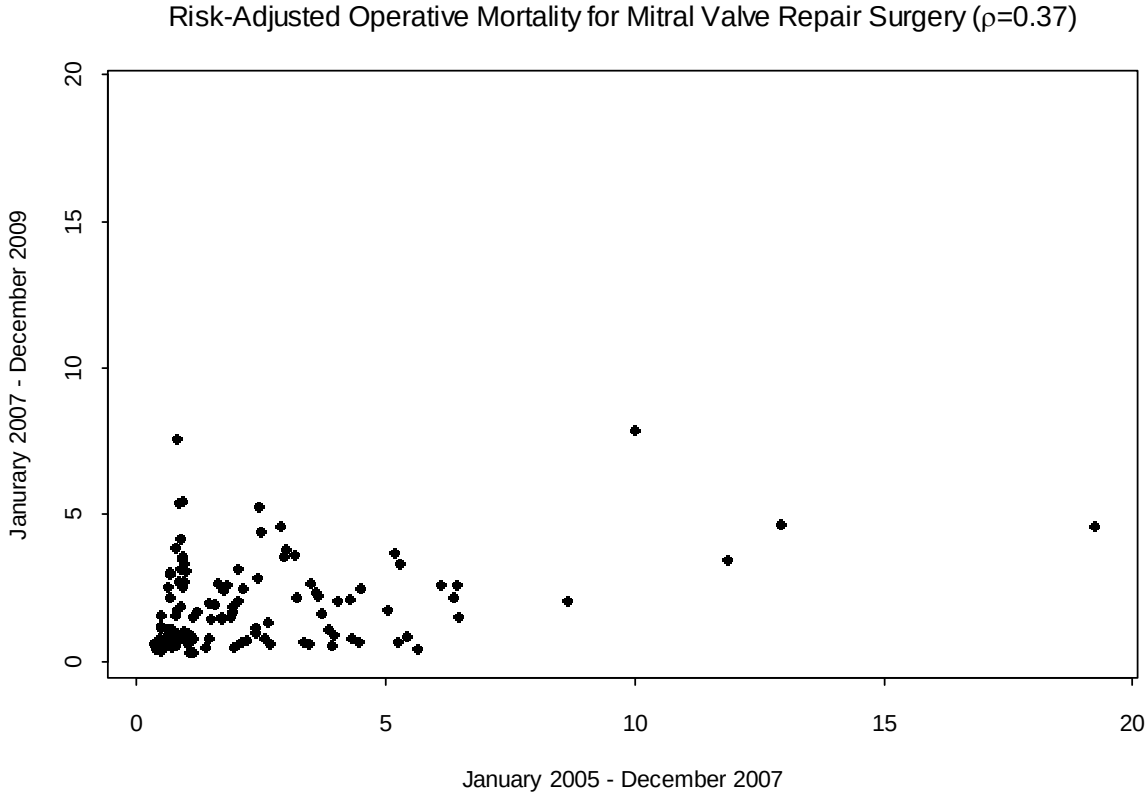
2b.3. Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*

Testing results: $\rho = 0.36$



Risk Adjusted Rate:

Testing results: $\rho = 0.37$



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, 2005-December 31, 2009. Sample contains 143 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009.

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Repair Surgery*
N	143
Mean	1.0
1st	4.2
5th	2.2
10th	1.8
25th	1.4
Median	0.9
75th	0.5
90th	0.4
95th	0.3
99th	0.3
Outlier†	10 (7.0)
High	3
Low	7

Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for Mitral Valve Repair Surgery*
N	143
Mean	1.6
1st	0.4
5th	0.5
10th	0.6
25th	0.8
Median	1.4
75th	2.1
90th	2.7
95th	3.2
99th	5.9
Outlier†	10 (7.0)
High	3
Low	7

†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."

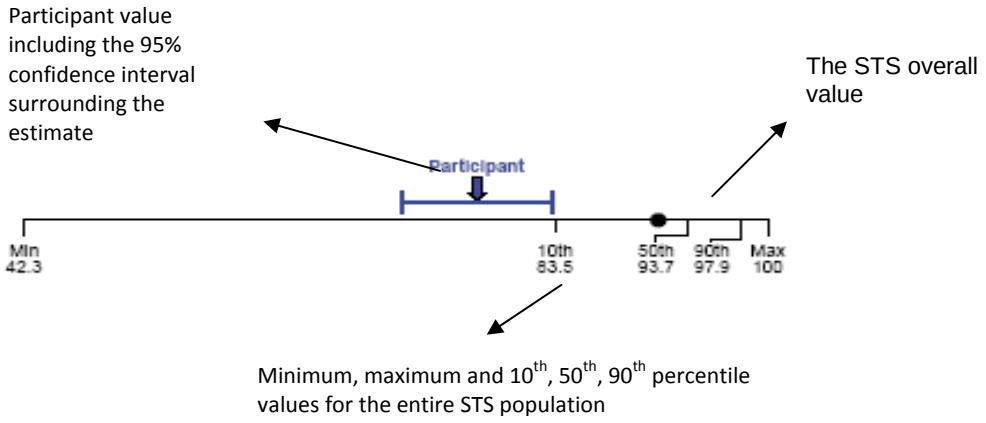
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

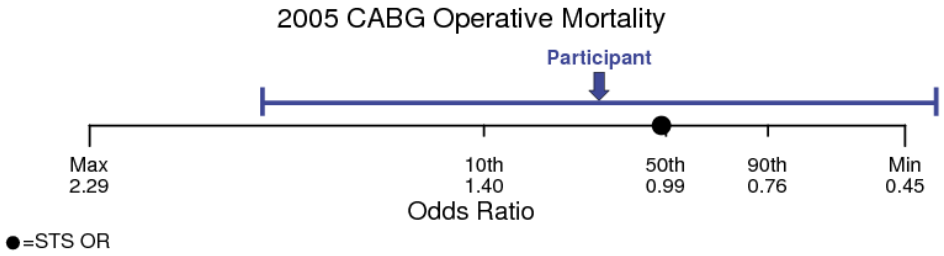
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

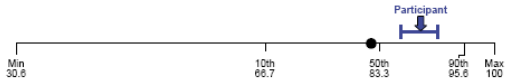
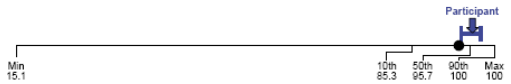
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 2—Isolated Valve Surgery

Sean M. O'Brien, PhD,[a] David M. Shahian, MD,[b] Giovanni Filardo, PhD, MPH,[c] Victor A. Ferraris, MD,[d] Constance K. Haan, MD,[e] Jeffrey B. Rich, MD,[f] Sharon-Lise T. Normand, PhD,[g] Elizabeth R. DeLong, PhD,[a] Cynthia M. Shewan, PhD,[h] Rachel S. Dokholyan, MPH,[a] Eric D. Peterson, MD, MPH,[a] Fred H. Edwards, MD,[e] and Richard P. Anderson, MD[i†]

[a]Duke Clinical Research Institute, Durham, North Carolina; [b]Massachusetts General Hospital, Boston, Massachusetts; [c]Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; [d]University of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; [e]University of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; [f]Sentara Cardiovascular Research Institute, Norfolk, Virginia; [g]Department of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [h]The Society of Thoracic Surgeons, Chicago, Illinois; and [i]Seattle, Washington

Background. **Adjustment for case-mix is essential when using observational data to compare surgical techniques or providers. That is most often accomplished through the use of risk models that account for preoperative patient factors that may impact outcomes. The Society of Thoracic Surgeons (STS) uses such risk models to create risk-adjusted performance reports for participants in the STS National Adult Cardiac Surgery Database (NCD). Although risk models were initially developed for coronary artery bypass surgery, similar models have now been developed for use with heart valve surgery, particularly as the proportion of such procedures has increased. The last published STS model for isolated valve surgery was based on data from 1994 to 1997 and did not include patients undergoing mitral valve repair. STS has developed new valve surgery models using contemporary data that include both valve repair as well as replacement. Expanding upon existing valve models, the new STS models include several nonfatal complications in addition to mortality.**

Methods. **Using STS data from 2002 to 2006, isolated valve surgery risk models were developed for operative mortality, permanent stroke, renal failure, prolonged ventilation (> 24 hours), deep sternal wound infection, reoperation for any reason, a major morbidity or mortality composite endpoint, prolonged postoperative length of stay, and short postoperative length of stay. The study population consisted of adult patients who underwent one of three types of valve surgery: isolated aortic valve replacement (n = 67,292), isolated mitral valve replacement (n = 21,229), or isolated mitral valve repair (n = 21,238). The population was divided into a 60% development sample and a 40% validation sample. After an initial empirical investigation, the three surgery groups were combined into a single logistic regression model with numerous interactions to allow the covariate effects to differ across these groups. Variables were selected based on a combination of automated stepwise selection and expert panel review.**

Results. **Unadjusted operative mortality (in-hospital regardless of timing, and 30-day regardless of venue) for all isolated valve procedures was 3.4%, and unadjusted in-hospital morbidity rates ranged from 0.3% for deep sternal wound infection to 11.8% for prolonged ventilation. The number of predictors in each model ranged from 10 covariates in the sternal infection model to 24 covariates in the composite mortality plus morbidity model. Discrimination as measured by the c-index ranged from 0.639 for reoperation to 0.799 for mortality. When patients in the validation sample were grouped into 10 categories based on deciles of predicted risk, the average absolute difference between observed versus predicted events within these groups ranged from 0.06% for deep sternal wound infection to 1.06% for prolonged postoperative stay.**

Conclusions. **The new STS risk models for valve surgery include mitral valve repair as well as multiple endpoints other than mortality. Model coefficients are provided and an online risk calculator is publicly available from The Society of Thoracic Surgeons website.**

(Ann Thorac Surg 2009;88:S23–42)

Models for predicting surgical outcomes on the basis of patient preoperative characteristics are valuable tools for research, quality improvement, and clinical practice. Such models are used by The Society of Thoracic Surgeons (STS) to produce risk-adjusted performance re-

†This author is deceased. Former Chair, Quality, Research and Patient Safety Council, The Society of Thoracic Surgeons, Chicago, IL.

Address correspondence to Dr Shahian, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

Drs O'Brien, Shahian, Filardo, Ferraris, Haan, Rich, Normand, DeLong, Shewan, Peterson, Edwards, Anderson, and Ms Dokholyan, have no conflicts of interest to declare regarding this work.

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.056

Abbreviations and Acronyms

AVR	= aortic valve replacement
CABG	= coronary artery bypass graft surgery
CI	= confidence interval
MI	= myocardial infarction
MVR	= mitral valve replacement
MVRepair	= mitral valve repair
NCD	= National Adult Cardiac Surgery Database
QMTF	= Quality Measurement Task Force
STS	= The Society for Thoracic Surgeons

ports for providers participating in the STS National Adult Cardiac Surgery Database (NCD). They are also used by STS surgeons and other physicians for counseling patients about the risk of surgery.

The earliest STS risk models were developed nearly 2 decades ago for isolated coronary artery bypass graft surgery (CABG). Subsequently, similar models have been developed for isolated valve replacement and combined CABG plus valve replacement. Because surgical practice and outcomes are changing rapidly, these models are updated periodically to reflect contemporary experience.

The last published STS model for isolated valve surgery was based on STS data from 1994 to 1997. The reference population included aortic and mitral valve replacements but excluded mitral valve repair, and the endpoint was operative mortality. In the decade since this model was published, many aspects of heart surgery have changed. First, as CABG volumes have decreased with the introduction of coronary stents, valve surgery as a proportion of overall heart surgery volume has increased in most practices. Between 2000 and 2006, the percentage of isolated CABG procedures decreased from 73% to 60% and the percentage of isolated valve procedures increased from 18% to 22%. Thus, in assessing provider performance, it is no longer sufficient only to consider isolated CABG surgery. Second, the frequency of mitral repair as a percentage of all isolated mitral operations in the STS NCD increased from 35% in 2000 to 53% in 2006. Third, during the same time period, the average mortality rate for isolated aortic or mitral surgery also decreased. Finally, efforts to measure and compare surgical performance have intensified and expanded. In addition to measuring operative mortality, performance reports increasingly focus on nonfatal complications as well as resource utilization and efficiency. Such outcomes have not historically been risk-adjusted for valve surgery.

The STS Quality Measurement Task Force (QMTF) has undertaken a complete revision of all STS risk models for adult cardiac surgery, and these new models were implemented in January 2008. This report, Part 2 of 3, describes the new STS models for isolated valve surgery (Part 1 describes the STS isolated CABG models, and Part 3 describes the models for CABG plus valve surgery). Authors of this report are the QMTF members who were involved in this initiative.

Two important features have been incorporated into these new models. First, the population includes mitral valve repair as well as aortic and mitral valve replacement. Second, in addition to operative mortality, the new models include six nonfatal in-hospital morbidity endpoints and two length-of-stay endpoints. In comparison with several other valve models that have recently been published [1–6], the STS models are distinguished by the large size of the development population and the broad spectrum of endpoints included.

Study Population and Endpoints

The population for this analysis consisted of operations on adult patients aged 20 to 100 years who underwent isolated single aortic or mitral valve surgery between January 1, 2002, and December 31, 2006. Only patients undergoing one of the following procedures were included: (1) isolated aortic valve replacement (AVR); (2) isolated mitral valve replacement (MVR); and (3) isolated mitral valve repair (MVRepair).

Because of the relatively small number of pulmonic, tricuspid, multiple valve procedures, and aortic repairs, these cases were not included in the current models. Patients undergoing concomitant CABG were excluded from the current analysis, but these were included in the separate STS valve plus CABG models described in Part 3 of this series. Records with missing data on sex (n = 44) were excluded because missing sex is not allowed in the analysis dataset used for creating STS database participant feedback reports. This left a final study population of 109,759 patient operations performed at 809 STS NCD participating groups. Patients on dialysis preoperatively (n = 2,699) were not included when developing the risk model for prediction of postoperative renal failure.

Patient characteristics in the study population are presented in Table 1.

Training and Validation Samples

The study population was randomly divided into a 60% training (development) sample and a 40% test (validation) sample. The development sample was used to identify predictor variables and estimate model coefficients. Data from the validation sample were used to assess model fit, discrimination, and calibration. After choosing variables and assessing model fit, the development and validation samples were subsequently combined, and the final model coefficients were estimated using the combined (development plus validation) data.

Endpoints

Risk models were developed for nine endpoints, identical to those in the STS CABG models. In contrast with the definition of operative mortality, which includes hospital deaths as well as deaths that occur after discharge within 30 days of surgery, the morbidity endpoints only include events that occurred before discharge. However, beginning with version 2.61, sternal infection data will be recorded for as long as 30 days postoperatively. The nine endpoints are as follows: (1) operative mortality: death during the same

Table 1. Distribution of Risk Factors in Overall Study Population Isolated Valve (2002–2006)

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Demographics								
Age, years								
< 55	28,147	25.6	13,227	19.66	6,601	31.09	8,319	39.17
55–64	23,258	21.2	12,987	19.30	4,833	22.77	5,438	25.61
65–74	28,145	25.6	18,299	27.19	5,294	24.94	4,552	21.43
≥75	30,209	27.5	22,779	33.85	4,501	21.20	2,929	13.79
Sex								
Male	60,752	55.4	39,209	58.27	9,055	42.65	12,488	58.80
Female	49,007	44.6	28,083	41.73	12,174	57.35	8,750	41.20
Race								
Caucasian	93,522	85.2	58,656	87.17	16,810	79.18	18,056	85.02
Black	7,630	7.0	3,555	5.28	2,383	11.23	1,692	7.97
Hispanic	3,680	3.4	2,344	3.48	889	4.19	447	2.10
Asian	1,538	1.4	719	1.07	437	2.06	382	1.80
Other	2,493	2.3	1,508	2.24	505	2.38	480	2.26
Missing	896	0.8	510	0.76	205	0.97	181	0.85
Risk factors								
Body surface area, m^2								
< 1.50	4,351	4.0	2,341	3.48	1,234	5.81	776	3.65
1.50–1.74	24,577	22.4	13,713	20.38	6,151	28.97	4,713	22.19
1.75–1.99	40,548	36.9	24,744	36.77	7,914	37.28	7,890	37.15
≥ 2.00	39,517	36.0	26,007	38.65	5,768	27.17	7,742	36.45
Missing	766	0.7	487	0.72	162	0.76	117	0.55
Body mass index, kg/m^2								
< 25	35,526	32.4	18,509	27.51	8,447	39.79	8,570	40.35
25–29	39,074	35.6	24,035	35.72	6,992	32.94	8,047	37.89
30–34	20,534	18.7	14,142	21.02	3,318	15.63	3,074	14.47
≥ 35	13,682	12.5	10,008	14.87	2,280	10.74	1,394	6.56
Missing	943	0.9	598	0.89	192	0.90	153	0.72
Diabetes mellitus								
No diabetes	88,709	80.8	52,052	77.35	17,535	82.60	19,122	90.04
Diabetes, noninsulin	14,900	13.6	11,026	16.39	2,412	11.36	1,462	6.88
Diabetes, insulin	5,788	5.3	3,974	5.91	1,216	5.73	598	2.82
Diabetes missing	138	0.1	91	0.14	34	0.16	13	0.06
Treatment missing	224	0.2	149	0.22	32	0.15	43	0.20
Hypertension								
No	41,649	37.9	22,338	33.20	8,859	41.73	10,452	49.21
Yes	67,886	61.9	44,816	66.60	12,326	58.06	10,744	50.59
Missing	224	0.2	138	0.21	44	0.21	42	0.20
Hypercholesterolemia								
No	59,003	53.8	33,156	49.27	12,857	60.56	12,990	61.16
Yes	50,328	45.9	33,865	50.33	8,286	39.03	8,177	38.50
Missing	428	0.4	271	0.40	86	0.41	71	0.33
Past or present smoker								
No	57,609	52.5	33,953	50.46	11,075	52.17	12,581	59.24
Yes	51,910	47.3	33,191	49.32	10,109	47.62	8,610	40.54
Missing	240	0.2	148	0.22	45	0.21	47	0.22
Chronic lung disease								
None	87,826	80.0	53,503	79.51	16,125	75.96	18,198	85.69
Mild	11,184	10.2	6,991	10.39	2,520	11.87	1,673	7.88
Moderate	6,346	5.8	4,022	5.98	1,494	7.04	830	3.91
Severe	3,332	3.0	2,110	3.14	853	4.02	369	1.74
Missing	1,071	1.0	666	0.99	237	1.12	168	0.79

Table 1. Continued

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Peripheral vascular disease								
No	101,129	92.1	61,222	90.98	19,550	92.09	20,357	95.85
Yes	8,381	7.6	5,909	8.78	1,641	7.73	831	3.91
Missing	249	0.2	161	0.24	38	0.18	50	0.24
Cerebrovascular disease								
No	96,852	88.2	58,983	87.65	18,158	85.53	19,711	92.81
Yes	12,661	11.5	8,147	12.11	3,033	14.29	1,481	6.97
Missing	246	0.2	162	0.24	38	0.18	46	0.22
CVA								
No CVA	101,631	92.6	62,518	92.91	18,833	88.71	20,280	95.49
Remote CVA (> 2 weeks)	6,926	6.3	4,203	6.25	1,912	9.01	811	3.82
Recent CVA (≤ 2 weeks)	818	0.7	325	0.48	409	1.93	84	0.40
CVA–missing timing	100	0.1	60	0.09	29	0.14	11	0.05
Missing	284	0.3	186	0.28	46	0.22	52	0.24
Endocarditis								
No endocarditis	100,998	92.0	63,257	94.00	17,926	84.44	19,815	93.30
Treated endocarditis	4,197	3.8	1,761	2.62	1,445	6.81	991	4.67
Active endocarditis	4,238	3.9	2,068	3.07	1,791	8.44	379	1.78
Endocarditis–missing type	63	0.1	30	0.04	27	0.13	6	0.03
Missing	263	0.2	176	0.26	40	0.19	47	0.22
Renal failure								
No	102,205	93.1	62,873	93.43	19,016	89.58	20,316	95.66
Yes	7,305	6.7	4,251	6.32	2,173	10.24	881	4.15
Missing	249	0.2	168	0.25	40	0.19	41	0.19
Renal function								
Creatinine < 1.00 mg/dL	42,028	38.3	25,679	38.16	7,754	36.53	8,595	40.47
Creatinine 1–1.49 mg/dL	51,939	47.3	32,058	47.64	9,372	44.15	10,509	49.48
Creatinine 1.50–1.99 mg/dL	8,081	7.4	5,078	7.55	1,875	8.83	1,128	5.31
Creatinine 2.00–2.49 mg/dL	1,946	1.8	1,192	1.77	512	2.41	242	1.14
Creatinine ≥ 2.50 mg/dL	1,294	1.2	750	1.11	390	1.84	154	0.73
Dialysis	2,699	2.5	1,464	2.18	900	4.24	335	1.58
Missing	1,772	1.6	1,071	1.59	426	2.01	275	1.29
Immunosuppressive treatment								
No	106,037	96.6	64,953	96.52	20,356	95.89	20,728	97.60
Yes	3,336	3.0	2,074	3.08	819	3.86	443	2.09
Missing	386	0.4	265	0.39	54	0.25	67	0.32
Previous CV interventions								
Previous coronary artery bypass surgery								
No	98,978	90.2	60,351	89.69	18,564	87.45	20,063	94.47
Yes	10,399	9.5	6,713	9.98	2,569	12.10	1,117	5.26
Missing	382	0.3	228	0.34	96	0.45	58	0.27
Previous valve surgery								
No	100,179	91.3	62,898	93.47	16,857	79.41	20,424	96.17
Yes	9,227	8.4	4,186	6.22	4,285	20.18	756	3.56
Missing	353	0.3	208	0.31	87	0.41	58	0.27
Previous other cardiac surgery								
No	105,686	96.3	65,084	96.72	20,034	94.37	20,568	96.85
Yes	3,662	3.3	1,975	2.93	1,077	5.07	610	2.87
Missing	411	0.4	233	0.35	118	0.56	60	0.28
Number of previous CV surgeries								
No prior CV surgery	91,196	83.1	56,629	84.15	15,239	71.78	19,328	91.01
1 prior CV surgery	15,399	14.0	9,122	13.56	4,775	22.49	1,502	7.07
2 or more prior CV surgeries	2,653	2.4	1,260	1.87	1,069	5.04	324	1.53
Missing	511	0.5	281	0.42	146	0.69	84	0.40

Table 1. Continued

Variable	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
	N	%	N	%	N	%	N	%
Prior PCI								
No PCI	101,878	92.8	62,145	92.35	19,573	92.20	20,160	94.92
PCI within 6 hours	122	0.1	58	0.09	51	0.24	13	0.06
PCI not within 6 hours	7,100	6.5	4,678	6.95	1,447	6.82	975	4.59
PCI–missing timing	133	0.1	90	0.13	28	0.13	15	0.07
Missing	526	0.5	321	0.48	130	0.61	75	0.35
Preoperative cardiac status								
Acuity status								
Elective	84,052	76.6	51,734	76.88	14,293	67.33	18,025	84.87
Urgent	23,795	21.7	14,670	21.80	6,071	28.60	3,054	14.38
Emergent	1,555	1.4	685	1.02	747	3.52	123	0.58
Emergent salvage	154	0.1	70	0.10	78	0.37	6	0.03
Missing	203	0.2	133	0.20	40	0.19	30	0.14
MI								
No prior MI	99,416	90.6	60,850	90.43	18,716	88.16	19,850	93.46
MI > 21 days	7,785	7.1	4,770	7.09	1,848	8.71	1,167	5.49
MI 8–21 days	719	0.7	480	0.71	170	0.80	69	0.32
MI 1–7 days	1,247	1.1	863	1.28	315	1.48	69	0.32
MI > 6 and < 24 hours	142	0.1	61	0.09	66	0.31	15	0.07
MI ≤ 6 hours	90	0.1	42	0.06	40	0.19	8	0.04
MI–missing timing	127	0.1	79	0.12	33	0.16	15	0.07
Missing	233	0.2	147	0.22	41	0.19	45	0.21
Angina								
No	85,364	77.8	49,573	73.67	17,598	82.90	18,193	85.66
Yes	24,164	22.0	17,577	26.12	3,591	16.92	2,996	14.11
Missing	231	0.2	142	0.21	40	0.19	49	0.23
Cardiogenic shock								
No	108,163	98.5	66,646	99.04	20,460	96.38	21,057	99.15
Yes	1,329	1.2	485	0.72	725	3.42	119	0.56
Missing	267	0.2	161	0.24	44	0.21	62	0.29
Resuscitation								
No	108,958	99.3	66,832	99.32	20,992	98.88	21,134	99.51
Yes	533	0.5	297	0.44	186	0.88	50	0.24
Missing	268	0.2	163	0.24	51	0.24	54	0.25
Arrhythmia								
No arrhythmia	89,779	81.8	57,451	85.38	14,604	68.79	17,724	83.45
AFib/flutter	16,124	14.7	7,569	11.25	5,721	26.95	2,834	13.34
Heart block	1,598	1.5	1,109	1.65	315	1.48	174	0.82
Sustained VT/VF	984	0.9	486	0.72	290	1.37	208	0.98
Arrhythmia–other	688	0.6	324	0.48	175	0.82	189	0.89
Arrhythmia–missing type	312	0.3	175	0.26	74	0.35	63	0.30
Missing	274	0.2	178	0.26	50	0.24	46	0.22
Preoperative IABP								
No	107,945	98.3	66,733	99.17	20,332	95.77	20,880	98.31
Yes	1,431	1.3	342	0.51	809	3.81	280	1.32
Missing	383	0.3	217	0.32	88	0.41	78	0.37
NYHA class								
I	17,413	15.9	10,222	15.19	2,706	12.75	4,485	21.12
II	32,360	29.5	20,295	30.16	4,915	23.15	7,150	33.67
III	40,321	36.7	25,483	37.87	8,205	38.65	6,633	31.23
IV	14,324	13.1	8,104	12.04	4,256	20.05	1,964	9.25
Missing	5,341	4.9	3,188	4.74	1,147	5.40	1,006	4.74

Table 1. Continued

	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
Variable	N	%	N	%	N	%	N	%
Congestive heart failure								
No	64,608	58.9	41,972	62.37	9,341	44.00	13,295	62.60
Yes	44,934	40.9	25,185	37.43	11,849	55.82	7,900	37.20
Missing	217	0.2	135	0.20	39	0.18	43	0.20
Number of diseased coronary vessels								
None	90,281	82.3	55,072	81.84	17,525	82.55	17,684	83.27
One	8,947	8.2	5,393	8.01	1,498	7.06	2,056	9.68
Two	3,386	3.1	2,180	3.24	735	3.46	471	2.22
Three	5,611	5.1	3,766	5.60	1,147	5.40	698	3.29
Missing	1,534	1.4	881	1.31	324	1.53	329	1.55
Left main disease ≥ 50%								
No	106,462	97.0	65,328	97.08	20,495	96.54	20,639	97.18
Yes	1,625	1.5	1,127	1.67	289	1.36	209	0.98
Missing	1,672	1.5	837	1.24	445	2.10	390	1.84
Ejection fraction, %								
< 25	2,694	2.5	1,774	2.64	341	1.61	579	2.73
25–34	5,900	5.4	3,810	5.66	1,052	4.96	1,038	4.89
35–44	10,035	9.1	6,181	9.19	2,208	10.40	1,646	7.75
45–54	20,481	18.7	12,411	18.44	4,382	20.64	3,688	17.37
≥ 55	60,890	55.5	36,584	54.37	11,308	53.27	12,998	61.20
Missing	9,759	8.9	6,532	9.71	1,938	9.13	1,289	6.07
Aortic stenosis								
No	54,457	49.6	13,309	19.78	20,303	95.64	20,845	98.15
Yes	54,681	49.8	53,722	79.83	696	3.28	263	1.24
Missing	621	0.6	261	0.39	230	1.08	130	0.61
Mitral stenosis								
No	100,609	91.7	65,186	96.87	15,383	72.46	20,040	94.36
Yes	8,155	7.4	1,401	2.08	5,676	26.74	1,078	5.08
Missing	995	0.9	705	1.05	170	0.80	120	0.57
Tricuspid stenosis								
No	108,073	98.5	66,243	98.44	20,821	98.08	21,009	98.92
Yes	331	0.3	152	0.23	120	0.57	59	0.28
Missing	1,355	1.2	897	1.33	288	1.36	170	0.80
Pulmonic stenosis								
No	107,512	98.0	65,842	97.85	20,783	97.90	20,887	98.35
Yes	141	0.1	91	0.14	29	0.14	21	0.10
Missing	2,106	1.9	1,359	2.02	417	1.96	330	1.55
Aortic insufficiency								
None	59,905	54.6	25,861	38.43	16,701	78.67	17,343	81.66
Trivial	9,191	8.4	5,916	8.79	1,661	7.82	1,614	7.60
Mild	13,282	12.1	10,014	14.88	1,798	8.47	1,470	6.92
Moderate	9,501	8.7	8,815	13.10	382	1.80	304	1.43
Severe	15,722	14.3	15,529	23.08	109	0.51	84	0.40
Missing	2,158	2.0	1,157	1.72	578	2.72	423	1.99
Mitral insufficiency								
None	43,731	39.8	40,453	60.12	2,283	10.75	995	4.68
Trivial	7,743	7.1	7,285	10.83	388	1.83	70	0.33
Mild	14,455	13.2	13,066	19.42	1,089	5.13	300	1.41
Moderate	10,224	9.3	4,438	6.60	3,246	15.29	2,540	11.96
Severe	31,813	29.0	573	0.85	14,045	66.16	17,195	80.96
Missing	1,793	1.6	1,477	2.19	178	0.84	138	0.65

Table 1. Continued

Variable	Overall Valve (n = 109,759)		AVR (n = 67,292)		MVR (n = 21,229)		MVRepair (n = 21,238)	
	N	%	N	%	N	%	N	%
Tricuspid insufficiency								
None	78,472	71.5	49,976	74.27	14,266	67.20	14,230	67.00
Trivial	8,856	8.1	5,612	8.34	1,381	6.51	1,863	8.77
Mild	13,346	12.2	7,333	10.90	2,788	13.13	3,225	15.19
Moderate	5,167	4.7	2,126	3.16	1,753	8.26	1,288	6.06
Severe	974	0.9	297	0.44	460	2.17	217	1.02
Missing	2,944	2.7	1,948	2.89	581	2.74	415	1.95
Pulmonic insufficiency								
None	97,954	89.2	60,463	89.85	18,837	88.73	18,654	87.83
Trivial	4,161	3.8	2,370	3.52	779	3.67	1,012	4.77
Mild	2,541	2.3	1,340	1.99	573	2.70	628	2.96
Moderate	441	0.4	209	0.31	144	0.68	88	0.41
Severe	76	0.1	34	0.05	30	0.14	12	0.06
Missing	4,586	4.2	2,876	4.27	866	4.08	844	3.97

AFib = atrial fibrillation; AVR = aortic valve replacement; CV = cardiovascular; CVA = cerebrovascular accident (stroke); IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repaid; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; VF = ventricular fibrillation; VT = ventricular tachycardia.

hospitalization as surgery, regardless of timing, or within 30 days of surgery regardless of venue; (2) permanent stroke (cerebrovascular accident [CVA]): a central neurologic deficit persisting longer than 72 hours; (3) renal failure: a new requirement for dialysis or an increase of the serum creatinine to greater than 2.0 mg/dL and double the most recent preoperative creatinine level; (4) prolonged ventilation (longer than 24 hours); (5) deep sternal wound infection; (6) reoperation for any reason; (7) major morbidity or mortality: a composite defined as the occurrence of any of the above endpoints; (8) prolonged postoperative length of stay (PLOS): length of stay (LOS) more than 14 days (alive or dead); and (9) short postoperative LOS (SLOS): LOS less than 6 days and patient alive at discharge.

Table 2 summarizes the endpoint frequencies in the study population.

Single Versus Multiple Models

Two issues required particularly careful consideration: whether to construct separate models for the AVR and MVR populations, and how best to further subdivide the mitral population into repair versus replacement.

Because of the large size of the STS NCD, separate

Table 2. Frequency of Endpoints in Overall Study Population 2002 to 2006

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
All isolated valve (AVR, MVR, MVRepair)									
N	109,759	109,759	107,060	109,759	109,759	109,759	109,759	109,759	109,759
Events	3,706	1,751	4,673	12,892	307	9,164	20,074	9,718	41,214
%	3.4	1.6	4.3	11.8	0.3	8.4	18.3	8.9	37.6
AVR									
N	67,292	67,292	65,828	67,292	67,292	67292	67,292	67,292	67,292
Events	2,157	1,007	2,774	7,323	197	5369	11,706	5,308	26,144
%	3.2	1.5	4.1	10.9	0.3	8.0	17.4	7.9	38.9
MVR									
N	21,229	21,229	20,329	21,229	21,229	21229	21,229	21,229	21,229
Events	1,210	447	1,348	4,015	71	2450	5,675	3,244	4,727
%	5.7	2.1	6.4	18.9	0.3	11.5	26.7	15.3	22.3
MVRepair									
N	21,238	21,238	20,903	21,238	21,238	21,238	21,238	21,238	21,238
Events	339	297	551	1,554	39	1,345	2,693	1,166	10,343
%	1.6	1.4	2.6	7.3	0.2	6.3	12.7	5.5	48.7

AVR = aortic valve replacement; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

models for AVR, MVR, and MVRepair initially seemed both feasible and appropriate. However, because the endpoints of interest are rare events, we recognized the possibility that the number of such events would be too small to support reliable estimation of the model coefficients.

To assess this tradeoff, we conducted a pilot study to compare two alternative strategies for developing risk models for isolated valve surgery. The first strategy involved developing models separately for three subpopulations (AVR, MVR, and MVRepair). The second strategy involved modeling all three subpopulations together in a single model; several interaction terms were included to allow the effect of selected risk factors to differ across the subpopulations. Both strategies were pilot tested by developing risk models for two endpoints: operative mortality and permanent stroke. These pilot models were developed in a 60% development sample and tested in a separate 40% validation sample. Each model was assessed by calculating the c-index and the generalized R^2 index of Nagelkerke [7] in the validation sample for each combination of subpopulation and endpoint (3 subpopulations $\times$ 2 endpoints = 6 combinations). With the exception of AVR operative mortality, the combined model with interactions resulted in better discrimination. With the exception of MVR and MVrepair operative mortality, the combined model also captured more variation as measured by the generalized R^2 statistic.

Because the combined model strategy performed better in the majority of cases, and because a single combined model was consistent with the previous STS valve model, the combined model strategy was selected. To avoid assuming that the weighting of each risk factor was exactly constant across the three populations, we included interactions between surgery type and several key predictor variables. In principle, fitting a single model with several interactions is advantageous because it allows for pooling information across related groups without making an a priori assumption that all of the covariate effects are exactly constant across groups.

Selection of Candidate Predictor Variables

Our general approach to variable selection is discussed in Part 1 of this series describing the development of the 2008 STS isolated CABG risk models. Briefly, we initially identified potential candidate variables by reviewing four versions of the STS data collection instrument (data versions 2.35, 2.41, 2.52.1, and 2.61) as well as previously published STS and similar cardiac risk models [1–6]. A panel of cardiac surgeons and health policy experts reviewed the initial variables for face validity and to be certain that no important predictor variables available in (or mappable to) to STS NCD data version 2.61 had been excluded.

Final candidate explanatory variables and their coding are summarized in Table 3. The variables were identical to the CABG model candidate variables with the following differences: (1) percutaneous coronary intervention conducted within 6 hours or less of surgery was not a candidate variable because it was present in only 122 patients (0.1%) in the valve model population; (2) infectious endocarditis was included. This risk factor was rarely present among isolated CABG patients (0.09%), but was not uncommon (7.7%) among patients undergoing valve surgery; (3) mitral stenosis was included; this risk factor was rarely present among isolated CABG patients (0.35%) but was common (7.4%) among patients undergoing valve surgery; and (4) an indicator for surgery type (AVR, MVR, MVRepair) was included in the valve models.

Coding of Explanatory Variables

The coding of continuous and categorical variables was identical to the CABG models, except for the following differences: (1) age was modeled as a linear spline truncated from below at 50 years and with a change of slope at 75; (2) creatinine was modeled as a linear term with values less than 0.5 and greater than 5.0 mapped to those values respectively (approximately the 1st and 99th percentiles of the empirical distribution); (3) previous myocardial infarction (MI) was modeled as three categories ($<$ 24 hours, 1 to 21 days, and $>$ 21 days or no MI); the first two categories were subsequently combined after expert panel review; (4) race was modeled as three categories: black, Hispanic, Caucasian/other; and (5) chronic lung disease was modeled as linear across four categories (none, mild, moderate, severe).

In general, these differences reflect a slightly simpler coding scheme (fewer parameters) for the valve models compared with the isolated CABG models.

Repair Versus Replacement

In addition to a number of variables whose inclusion or coding were noted to be problematic during development of the 2008 STS isolated CABG models (Part 1 of this series), the approach to modeling mitral valve repair versus replacement was of some concern in the valve models. From a methodologic perspective, models used for risk-adjustment should include all patient preoperative risk factors that vary in prevalence between institutions and that substantially impact the probability of an adverse outcome. Such models should include variables that reflect the patient's baseline condition but should not include intraoperative events (eg, unexpected hemorrhage) or discretionary care processes (eg, use of a mechanical versus bioprosthetic valve). Adjusting for intraoperative events is not appropriate because these may be a reflection of the surgeon's performance. Adjusting for discretionary care processes may likewise mask differences in performance if the surgeon's choice of procedures has a substantial impact on outcomes. The same patient may receive valve repair if treated by one surgeon and replacement if treated by another. Adjusting for repair versus replacement will potentially conceal the outcomes of surgeons who achieve excellent results by repairing technically challenging valves that might otherwise be replaced if treated by a surgeon with less skill or tenacity. Importantly, there is considerable evidence to suggest the superiority of valve repair whenever feasible.

However, in addition to such discretionary factors, the decision to repair rather than replace the mitral valve is

Table 3. List of Final Candidate Variables and Their Coding for STS Valve Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline truncated from below at 50 and with knot at 75
Ejection fraction	Linear, values > 50 mapped to 50
Body surface area[a]	Quadratic polynomial modeled separately for males and females. Note: body surface area < 1.4 and > 2.6 mapped to those values, respectively.
Creatinine	Linear (only for patients not on dialysis). Note: creatinine < 0.5 and > 5.0 mapped to those values, respectively.
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval. Modeled as linear across categories.
Binary variables	
Active infectious endocarditis	Yes/no
Dialysis	Yes/no
Preoperative atrial fibrillation	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Preoperative IABP or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no MI < 7 days)	Yes/no
Left main disease	Yes/no
Aortic stenosis	Yes/no
Mitral stenosis	Yes/no
Aortic insufficiency	Defined as at least moderate (yes/no)
Mitral insufficiency	Defined as at least moderate (yes/no)
Tricuspid insufficiency	Defined as at least moderate (yes/no)
Categorical variables	
Chronic lung disease	Modeled as linear across categories (none, mild, moderate, severe)
CVD/CVA	3 groups: no CVD, CVD no CVA, CVD + CVA
Diabetes mellitus	3 groups: insulin diabetes, noninsulin diabetes, other or no diabetes
Number diseased coronary vessels	3 groups: < 2, 2, 3. Modeled as linear across the categories
MI	3 groups: < 24 hr, 1–21 days, > 21 days or no MI (groups 1 and 2 were subsequently collapsed)
Race	3 groups: Black; Hispanic; Other including Caucasian
Status	4 groups: elective, urgent, emergent—no resuscitation, salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, ≥2 previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF+NYHA IV
Surgery type	3 groups: AVR, MVR, MVRepair
Interaction terms	
Age by reoperation[a]	
Age by emergent status[a]	
Surgery type by each of the following:	Age, diabetes, dialysis, creatinine, reoperation, endocarditis, emergent status, CLD, CHF, EF, sex, shock, IABP/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, aortic stenosis

[a] These variables were forced into each model.

AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association.

also dependent upon the patient's preoperative valve disease etiology, anatomy, and pathophysiology. On average, patients amenable to valve repair have less extensive valve pathology and a relatively favorable postoperative prognosis (the mortality rate for valve repair is 1.6% compared with 5.7% for replacement). Ignoring these anatomical differences can introduce bias when comparing institutions, especially because these variables are not captured elsewhere on the STS data collection form.

A related difficulty in adjusting for repair versus re-

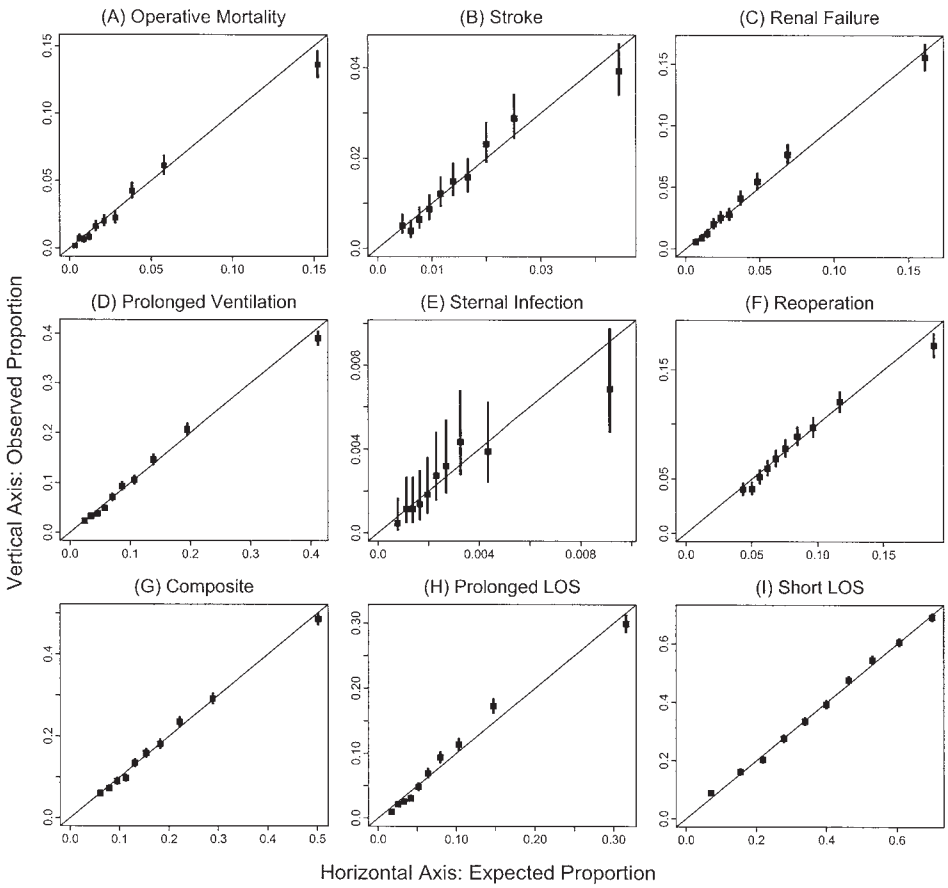


Fig 1. Plots of observed (O) versus expected (E) in validation sample

placement is that the former approach may sometimes be abandoned intraoperatively by the surgeon and converted to MVR. That may sometimes occur because of unforeseen technical problems that would prevent most surgeons from completing the repair, but in other instances, a more skilled surgeon might persist and achieve successful valve repair. Effectively separating these two scenarios is problematic from available data.

Ultimately, it was elected to include an indicator for mitral valve repair versus replacement in the valve risk models, consistent with the approach in a number of existing valve surgery models. We acknowledge that available data make it impossible to determine whether patient differences or surgical skill and judgment are the most important factors in determining between-provider variation in the proportion of valves repaired.

Recognizing the potential limitations of this modeling approach, the decision to adjust for repair versus replacement may be reassessed in future versions of the STS risk models. Beginning with data in version 2.61, the database will capture whether or not repair was attempted, and repair versus replacement may be analyzed based on an intention-to-treat principle.

Missing Data

Model variables with more than 1% missing data in the study sample were ejection fraction (8.9%), NYHA class (4.9%), tricuspid insufficiency (2.7%), aortic insufficiency (2.0%), mitral insufficiency (1.6%), left main disease (1.5%), creatinine/dialysis (1.6%), and number of diseased vessels (1.4%). The method of imputing missing data was identical to that employed in the isolated CABG models and described in Part 1 of this series. Briefly, binary risk factors were modeled as yes versus no or missing (ie, missing values were analyzed as if the endpoint did not occur). Missing data on categorical variables were imputed to the lowest risk value, typically the mode, and outcomes were typically similar for missing data and lowest risk patients. Missing data on continuous variables were imputed by grouping patients into strata and assigning the stratum-specific median value. For example, ejection fraction was imputed by grouping on sex and congestive heart failure and calculating the median ejection fraction among patients with nonmissing ejection fraction in each group.

Although multiple imputation is generally preferable to single imputation [8], single imputation was chosen for this analysis mainly because of practical considerations. Furthermore, because of the small fraction of missing data, the impact of single versus multiple imputation was considered to be inconsequential. Subsequent sensitivity analyses confirmed that the choice between single versus multiple imputation had little impact on the final regression coefficients, risk estimates, and confidence intervals. A summary of these sensitivity analyses, including coef-

Table 4. Discrimination of Models in Development and Validation Samples

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall									
Development sample	0.805	0.694	0.782	0.770	0.704	0.643	0.721	0.770	0.738
Validation sample	0.799	0.691	0.762	0.762	0.659	0.639	0.718	0.773	0.734
AVR									
Development sample	0.779	0.679	0.766	0.748	0.710	0.630	0.698	0.752	0.713
Validation sample	0.759	0.689	0.749	0.736	0.637	0.619	0.694	0.759	0.713
MVR									
Development sample	0.794	0.679	0.767	0.772	0.591	0.642	0.735	0.748	0.726
Validation sample	0.802	0.702	0.748	0.772	0.656	0.634	0.738	0.729	0.710
MVRepair									
Development sample	0.855	0.736	0.813	0.765	0.774	0.616	0.703	0.777	0.733
Validation sample	0.844	0.672	0.788	0.773	0.714	0.646	0.712	0.800	0.725

AVR = aortic valve replacement; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

ficients and covariance matrices, is available at www.sts.org/riskmodels.

Final Variable Selection Procedure

Variables were initially selected using an automated stepwise model selection algorithm. The stepwise procedure began with a model that included all of the final candidate variables except for interaction terms. Age, sex, body surface area, and month of surgery were forced into each model. Other variables were selected in a stepwise fashion using a significance criterion of 0.05 for entry and removal. This criterion was less stringent than that employed in development of the CABG models, because the sample size in the former was so much larger than that which was used for the valve models. The stepwise procedure was performed separately for each endpoint. The results were then reviewed by an expert panel of surgeons, and the following changes were made based on their feedback: (1) "MI less than 24 hours" and "MI 1 to 21 days" were collapsed into a single category; (2) preoperative atrial fibrillation was forced into the model for stroke (CVA); and (3) an indicator variable for dialysis was forced into any model that included creatinine level.

Interaction Terms

In addition to including main effects, we tested the interaction between surgery group (AVR, MVR, MVRepair) and each of the following variables: age, diabetes mellitus, dialysis, creatinine, reoperation, endocarditis, emergent status, chronic lung disease, congestive heart failure, ejection fraction, sex, shock, intra-aortic balloon pump/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, and aortic stenosis. These interaction terms allowed the effect of these selected risk factors to differ across the surgery populations.

Four additional sets of interactions were also included in the models: (1) sex by body surface area (BSA); (2) sex by BSA^2; (3) age by reoperation; and (4) age by emergent status. These interaction terms were preselected and were not tested as part of the backward selection algorithm. Additional technical details are provided in the Appendix. For reasons described in Part 1 of this series (isolated CABG risk models), an extensive automated search for additional interaction terms was not conducted.

Adjustment for Time Trends

Surgery date was included in each model to adjust for changes in the frequency of adverse outcomes over the 5-year study period. Although surgery date is not itself a variable of interest, we adjusted for it to reduce potential confounding by time trends when estimating regression coefficients for the variables that are of primary interest (ie, patient preoperative risk factors). An example is provided in Part 1 of this series.

Surgery date was categorized into 6-month intervals (corresponding to the biannual STS data harvests) and modeled as a linear trend across the ordinal categories. Because it is a nuisance variable, surgery date is not included in the final risk prediction algorithm. Thus, a patient's predicted risk does not depend on the patient's surgery date. As described in the Appendix, the published intercept parameter has been adjusted to incorporate the time trend. The adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Results

Assessment of Model Fit and Discrimination

Because of the relatively large size of our sample, the Hosmer-Lemeshow test is uninformative and would invariably result in a significant p value [9]. As an alternative, model fit was assessed graphically by plotting observed versus predicted rates of each endpoint across deciles of predicted risk in the development and validation samples. This was done in the overall population and in subgroups based on surgery type (AVR, MVR, MVRepair); age ($<$ 60, 60 to 79, $\geq$ 80 years); sex (male, female); diabetes mellitus (yes/no); status (elective, nonelective); and ejection fraction

Table 5. Odds Ratios (95% Confidence Intervals) for the Final Selected Models

A. Odds ratios for variables that do not interact with surgery group

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.10, 1.31)	1.06 (0.93, 1.20)	NA	1.18 (1.11, 1.25)	NA	1.11 (1.04, 1.18)	1.12 (1.07, 1.18)	1.17 (1.10, 1.24)	0.74 (0.70, 0.78)
BSA 1.6 versus 2.0 among females	1.19 (1.09, 1.30)	1.18 (1.03, 1.35)	0.95 (0.87, 1.04)	1.15 (1.08, 1.22)	0.42 (0.27, 0.68)	1.26 (1.18, 1.34)	1.17 (1.12, 1.23)	1.11 (1.04, 1.17)	0.99 (0.95, 1.04)
BSA 1.6 versus 2.0 among males	1.75 (1.48, 2.07)	1.17 (0.92, 1.47)	1.33 (1.12, 1.58)	1.56 (1.41, 1.74)	0.94 (0.49, 1.84)	1.34 (1.21, 1.49)	1.44 (1.33, 1.57)	1.39 (1.25, 1.56)	0.73 (0.68, 0.79)
BSA 1.8 versus 2.0 among females	0.99 (0.95, 1.04)	1.08 (0.99, 1.17)	0.90 (0.86, 0.94)	1.00 (0.97, 1.03)	0.65 (0.54, 0.77)	1.07 (1.03, 1.11)	1.02 (0.99, 1.04)	0.99 (0.96, 1.02)	1.05 (1.03, 1.08)
BSA 1.8 versus 2.0 among males	1.21 (1.14, 1.29)	1.07 (0.98, 1.16)	1.07 (1.00, 1.14)	1.14 (1.10, 1.19)	0.90 (0.70, 1.14)	1.12 (1.08, 1.16)	1.12 (1.09, 1.16)	1.10 (1.06, 1.15)	0.92 (0.89, 0.94)
BSA 2.2 versus 2.0 among females	1.21 (1.11, 1.33)	0.94 (0.80, 1.10)	1.30 (1.21, 1.41)	1.15 (1.09, 1.21)	1.57 (1.26, 1.96)	1.02 (0.95, 1.09)	1.12 (1.07, 1.16)	1.14 (1.08, 1.21)	0.85 (0.81, 0.88)
BSA 2.2 versus 2.0 among males	0.98 (0.93, 1.03)	0.95 (0.88, 1.03)	1.09 (1.05, 1.14)	1.05 (1.02, 1.08)	1.32 (1.17, 1.48)	0.95 (0.93, 0.98)	1.02 (0.99, 1.04)	1.03 (1.00, 1.07)	0.94 (0.93, 0.96)
Creatinine per 1 unit	1.55 (1.46, 1.64)	1.34 (1.22, 1.47)	2.04 (1.93, 2.16)	1.58 (1.51, 1.65)	NA	1.27 (1.20, 1.33)	1.64 (1.57, 1.71)	1.58 (1.51, 1.65)	0.64 (0.61, 0.68)
CVD with CVA	NA	1.81 (1.56, 2.10)	1.22 (1.09, 1.37)	1.28 (1.18, 1.38)	NA	1.14 (1.05, 1.24)	1.20 (1.12, 1.28)	1.40 (1.29, 1.52)	0.77 (0.72, 0.83)
CVD without CVA	NA	1.32 (1.11, 1.57)	1.23 (1.10, 1.37)	1.14 (1.05, 1.23)	NA	1.06 (0.96, 1.17)	1.08 (1.01, 1.15)	NA	0.80 (0.73, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	NA	1.10 (1.01, 1.20)	NA	1.07 (1.02, 1.11)	NA	NA	1.04 (1.00, 1.08)	1.03 (0.98, 1.08)	0.90 (0.86, 0.94)
EF per 10-unit decrease	1.09 (1.05, 1.14)	NA	1.04 (1.00, 1.09)	1.12 (1.09, 1.15)	1.26 (1.12, 1.41)	1.08 (1.04, 1.11)	1.10 (1.07, 1.12)	1.12 (1.08, 1.15)	0.87 (0.85, 0.90)
Hypertension	1.12 (1.03, 1.22)	1.19 (1.07, 1.33)	1.35 (1.25, 1.45)	1.11 (1.06, 1.17)	NA	NA	1.11 (1.07, 1.15)	NA	0.94 (0.91, 0.97)
Immunosuppressive treatment	1.42 (1.21, 1.67)	NA	1.39 (1.19, 1.62)	NA	NA	NA	1.16 (1.06, 1.27)	1.31 (1.17, 1.47)	NA
Left main disease	1.19 (0.98, 1.46)	NA	1.19 (0.98, 1.44)	NA	2.17 (1.13, 4.16)	NA	NA	NA	NA
Active infectious endocarditis	1.95 (1.68, 2.27)	1.87 (1.52, 2.29)	2.17 (1.88, 2.50)	2.15 (1.95, 2.36)	NA	1.55 (1.39, 1.73)	1.97 (1.80, 2.15)	2.79 (2.51, 3.09)	0.34 (0.30, 0.38)
Mitral insufficiency, moderate/severe	NA	1.26 (1.14, 1.39)	NA	NA	NA	NA	NA	NA	NA
Tricuspid insufficiency, moderate/severe	NA	NA	1.14 (1.01, 1.29)	1.14 (1.04, 1.25)	NA	1.09 (1.00, 1.20)	1.21 (1.12, 1.30)	1.17 (1.05, 1.31)	0.82 (0.73, 0.92)
Peripheral vascular disease	1.25 (1.12, 1.38)	1.29 (1.11, 1.49)	NA	NA	NA	1.22 (1.12, 1.32)	1.14 (1.07, 1.21)	1.17 (1.09, 1.25)	0.83 (0.78, 0.88)
Aortic stenosis		NA	NA	0.90 (0.83, 0.97)	NA	0.90 (0.84, 0.96)	0.93 (0.87, 0.98)	0.86 (0.80, 0.92)	1.07 (1.02, 1.13)
Mitral stenosis	1.24 (1.08, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA
MI $\leq$ 21 days	1.14 (0.98, 1.34)	NA	NA	1.37 (1.22, 1.55)	NA	1.04 (0.91, 1.18)	1.28 (1.16, 1.41)	1.21 (1.06, 1.37)	0.81 (0.72, 0.91)
Time trend, per 6-month harvest interval	0.98 (0.97, 0.99)	0.98 (0.96, 1.00)	1.01 (0.99, 1.02)	1.02 (1.01, 1.03)	0.97 (0.93, 1.01)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Race black	NA	1.33 (1.13, 1.57)	1.51 (1.34, 1.69)	1.42 (1.27, 1.58)	NA	1.27 (1.15, 1.40)	1.37 (1.27, 1.49)	1.45 (1.31, 1.60)	0.64 (0.59, 0.70)
Race Hispanic	NA	0.87 (0.64, 1.19)	1.16 (0.97, 1.38)	1.07 (0.94, 1.22)	NA	1.14 (1.00, 1.30)	1.09 (0.98, 1.22)	1.16 (0.98, 1.38)	0.82 (0.72, 0.93)
Status urgent	1.29 (1.19, 1.40)	NA	1.21 (1.11, 1.33)	1.29 (1.20, 1.39)	NA	1.17 (1.10, 1.25)	1.22 (1.15, 1.29)	1.42 (1.33, 1.51)	0.70 (0.66, 0.74)
Unstable angina	1.21 (1.04, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA

Table 5. Continued

B. Odds ratios for aortic valve replacement

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.43 (1.34, 1.52)	1.48 (1.38, 1.59)	1.38 (1.30, 1.47)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.16 (1.12, 1.21)	1.23 (1.19, 1.26)	1.31 (1.25, 1.37)	0.75 (0.73, 0.77)
Age 70 versus 50	2.04 (1.79, 2.32)	2.19 (1.90, 2.52)	1.90 (1.68, 2.16)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.35 (1.25, 1.46)	1.50 (1.42, 1.59)	1.71 (1.55, 1.87)	0.57 (0.54, 0.60)
Age 80 versus 50	3.34 (2.84, 3.93)	3.21 (2.70, 3.81)	2.88 (2.46, 3.37)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.59 (1.44, 1.76)	1.97 (1.82, 2.12)	2.50 (2.24, 2.79)	0.34 (0.32, 0.36)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.33 (1.24, 1.43)	NA	NA	1.20 (1.13, 1.27)	1.25 (1.17, 1.34)	0.86 (0.81, 0.91)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.92 (1.77, 2.08)	NA	1.25 (1.17, 1.35)	1.62 (1.51, 1.73)	1.54 (1.40, 1.68)	0.72 (0.65, 0.79)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.42 (1.31, 1.55)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.39 (1.29, 1.50)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.12 (1.04, 1.20)	NA	NA	1.12 (1.06, 1.18)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	2.85 (2.35, 3.45)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.29 (0.24, 0.34)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	1.78 (1.55, 2.05)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.56 (0.48, 0.66)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	1.74 (1.37, 2.21)	NA
Female versus male (at BSA = 1.8)	1.23 (1.10, 1.36)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.29 (1.21, 1.38)	0.98 (0.72, 1.33)	0.86 (0.81, 0.93)	1.03 (0.98, 1.08)	1.25 (1.16, 1.35)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.27 (1.21, 1.33)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.29 (1.24, 1.34)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	1.83 (1.64, 2.05)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation, ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.49 (2.14, 2.90)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	3.77 (2.75, 5.16)	2.78 (1.85, 4.17)	3.10 (2.21, 4.35)	4.54 (3.54, 5.83)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	7.94 (5.40, 11.66)	2.11 (1.06, 4.19)	3.47 (2.19, 5.51)	3.50 (2.41, 5.08)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

Table 5. Continued

C. Odds ratios for mitral valve replacement

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.65 (1.53, 1.78)	1.48 (1.38, 1.59)	1.35 (1.26, 1.44)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.25 (1.19, 1.31)	1.33 (1.29, 1.39)	1.26 (1.21, 1.33)	0.71 (0.68, 0.74)
Age 70 versus 50	2.71 (2.33, 3.17)	2.19 (1.90, 2.52)	1.81 (1.60, 2.06)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.56 (1.42, 1.71)	1.78 (1.65, 1.92)	1.60 (1.45, 1.76)	0.50 (0.46, 0.55)
Age 80 versus 50	5.14 (4.15, 6.37)	3.21 (2.70, 3.81)	2.67 (2.23, 3.20)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.97 (1.72, 2.26)	2.54 (2.27, 2.84)	2.27 (2.00, 2.58)	0.28 (0.25, 0.32)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.19 (1.07, 1.32)	NA	NA	1.11 (1.01, 1.21)	1.25 (1.17, 1.34)	0.96 (0.87, 1.06)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.72 (1.55, 1.91)	NA	1.25 (1.17, 1.35)	1.49 (1.36, 1.64)	1.54 (1.40, 1.68)	0.80 (0.71, 0.91)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.66 (1.47, 1.86)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.67 (1.52, 1.83)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.30 (1.16, 1.45)	NA	NA	1.34 (1.22, 1.47)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	4.59 (3.65, 5.77)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.23 (0.16, 0.33)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	2.21 (1.90, 2.56)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.63 (0.51, 0.77)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	1.05 (0.85, 1.31)	NA
Female versus male (at BSA=1.8)	1.11 (0.97, 1.27)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.06 (0.98, 1.16)	0.98 (0.72, 1.33)	0.79 (0.72, 0.87)	1.03 (0.98, 1.08)	1.09 (0.99, 1.19)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.08 (1.01, 1.16)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.16 (1.11, 1.22)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	1.50 (1.34, 1.67)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation, ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.03 (1.76, 2.35)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	2.74 (1.99, 3.78)	2.78 (1.85, 4.17)	2.20 (1.59, 3.05)	3.19 (2.41, 4.23)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	5.78 (3.77, 8.85)	2.11 (1.06, 4.19)	2.46 (1.56, 3.88)	2.46 (1.66, 3.65)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

Table 5. Continued

D. Odds ratios for mitral valve repair

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50	1.80 (1.62, 2.00)	1.48 (1.38, 1.59)	1.55 (1.41, 1.71)	1.31 (1.26, 1.36)	1.52 (1.31, 1.76)	1.20 (1.13, 1.27)	1.31 (1.26, 1.37)	1.50 (1.41, 1.60)	0.62 (0.60, 0.65)
Age 70 versus 50	3.24 (2.63, 4.00)	2.19 (1.90, 2.52)	2.42 (2.00, 2.92)	1.71 (1.59, 1.84)	2.31 (1.72, 3.10)	1.44 (1.29, 1.62)	1.73 (1.58, 1.89)	2.25 (1.98, 2.55)	0.39 (0.36, 0.42)
Age 80 versus 50	6.72 (5.00, 9.04)	3.21 (2.70, 3.81)	4.11 (3.14, 5.38)	2.31 (2.12, 2.52)	2.73 (1.95, 3.80)	1.75 (1.48, 2.07)	2.42 (2.12, 2.76)	3.78 (3.17, 4.51)	0.19 (0.17, 0.22)
CHF, not NYHA IV	1.29 (1.18, 1.42)	NA	1.24 (1.14, 1.34)	1.16 (0.99, 1.35)	NA	NA	1.11 (0.99, 1.24)	1.25 (1.17, 1.34)	0.92 (0.80, 1.05)
CHF, NYHA IV	1.83 (1.62, 2.07)	NA	1.61 (1.44, 1.81)	1.67 (1.43, 1.95)	NA	1.25 (1.17, 1.35)	1.50 (1.33, 1.68)	1.54 (1.40, 1.68)	0.76 (0.65, 0.90)
Diabetes, insulin	1.62 (1.43, 1.83)	NA	1.91 (1.70, 2.14)	1.68 (1.42, 1.97)	1.56 (1.05, 2.31)	1.20 (1.10, 1.31)	1.57 (1.36, 1.81)	1.68 (1.55, 1.83)	0.64 (0.59, 0.69)
Diabetes, noninsulin	1.27 (1.15, 1.39)	NA	1.45 (1.34, 1.57)	1.31 (1.11, 1.55)	NA	NA	1.26 (1.10, 1.45)	1.22 (1.15, 1.30)	0.85 (0.81, 0.88)
Dialysis versus no dialysis and creatinine = 1.0	6.24 (4.19, 9.30)	1.65 (1.34, 2.03)	NA	3.07 (2.74, 3.43)	NA	1.79 (1.60, 2.01)	2.42 (2.21, 2.66)	2.94 (2.64, 3.27)	0.26 (0.19, 0.37)
Preoperative IABP/ inotropes	1.47 (1.26, 1.71)	NA	1.34 (1.15, 1.57)	2.90 (2.28, 3.70)	1.69 (1.08, 2.65)	1.14 (1.02, 1.29)	1.75 (1.59, 1.94)	1.46 (1.30, 1.63)	0.49 (0.38, 0.64)
Shock	1.62 (1.29, 2.03)	1.65 (1.21, 2.25)	NA	2.09 (1.77, 2.47)	NA	1.32 (1.11, 1.58)	2.11 (1.80, 2.49)	2.50 (1.51, 4.12)	NA
Female versus male (at BSA = 1.8)	0.97 (0.77, 1.21)	1.25 (1.09, 1.43)	0.97 (0.88, 1.07)	1.23 (1.10, 1.38)	0.98 (0.72, 1.33)	0.90 (0.80, 1.02)	1.03 (0.98, 1.08)	1.28 (1.12, 1.47)	0.69 (0.66, 0.73)
CLD (moderate versus mild, or severe versus moderate)	1.23 (1.09, 1.39)	NA	1.18 (1.13, 1.23)	1.26 (1.22, 1.30)	1.27 (1.13, 1.42)	1.09 (1.06, 1.12)	1.17 (1.14, 1.20)	1.26 (1.15, 1.40)	0.81 (0.79, 0.83)
Reoperation, 1 previous operation[a]	2.11 (1.78, 2.49)	2.09 (1.64, 2.65)	1.55 (1.31, 1.84)	2.06 (1.73, 2.45)	NA	1.31 (1.16, 1.49)	1.55 (1.42, 1.70)	1.42 (1.27, 1.59)	0.67 (0.62, 0.72)
Reoperation ≥ 2 previous operations[a]	2.48 (1.99, 3.08)	2.36 (1.76, 3.16)	1.66 (1.33, 2.07)	2.80 (2.32, 3.37)	NA	1.41 (1.19, 1.67)	1.96 (1.73, 2.22)	1.76 (1.52, 2.03)	0.50 (0.43, 0.58)
Status emergent, no resuscitation[a]	8.73 (4.84, 15.74)	2.78 (1.85, 4.17)	3.03 (1.69, 5.43)	6.12 (3.96, 9.46)	NA	1.63 (1.31, 2.03)	3.23 (2.66, 3.93)	2.45 (2.02, 2.97)	0.33 (0.25, 0.42)
Status emergent, with resuscitation or salvage[a]	18.39 (9.68, 34.96)	2.11 (1.06, 4.19)	3.39 (1.76, 6.54)	4.72 (2.71, 8.23)	NA	NA	3.38 (2.36, 4.84)	NA	0.32 (0.19, 0.54)

[a] Variable interacts with age. Reported odds ratio represents effect of risk factor for patients aged 50 years old.

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; NA = not applicable; NYHA = New York Heart Association; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

(≤ 40, > 40). Calibration plots (observed versus expected) based on the overall validation sample are presented in Figure 1. The average absolute difference between observed versus predicted event rates within deciles of predicted risk ranged from 0.06% for deep sternal wound infection to 1.06% for prolonged postoperative stay. Analogous figures were produced for specific valve procedures and numerous subgroups, and these are available at www.sts.org/riskmodels.

Model fit appeared to be adequate for each endpoint with the possible exception of deep sternal wound infection, which revealed some overfitting within certain subgroups. A modest degree of overfitting was expected for this endpoint given the relatively small number of infections and large number of candidate predictors.

Discrimination was assessed by the c-statistic, also known as the area under the receiver operating characteristic (ROC) curve. Table 4 presents the discrimination of each model in the development and validation samples for all patients combined and for subgroups consisting of AVR, MVR, and MVRepair. In the validation sample, c-statistics for the operative mortality model were 0.799 (overall), 0.759 (AVR), 0.802 (MVR), and 0.844 (MVRepair). C-statistics in the validation sample for other endpoints ranged from 0.619 for reoperation in the AVR subgroup to 0.800 for prolonged length of stay in the MVRepair subgroup.

Final Models

After validating the models in the 40% validation sample, the development and validation samples were then combined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regressions were estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [10]. An independence working correlation matrix was used to apply the generalized estimating equations methodology. With this approach, the estimated regression coefficients were identical to those obtained using ordinary logistic regression, but the standard errors were adjusted to account for the clustered data structure.

Odds Ratios

Odds ratios and 95% confidence intervals (CI) for the final selected models are presented in Table 5. "Not applicable" indicates that the specific predictor was not included in a particular risk model. Because several variables interact with surgery type, the odds ratios for these variables differ depending on the type of surgery (AVR, MVR, MVRepair). For example, in the operative mortality model, the odds ratio for emergent status is 3.77 (95% CI: 2.75, 5.16) for AVR, 2.74 (95% CI: 1.99, 3.78) for MVR, and 8.73 (95% CI: 4.84, 15.74) for MVRepair. Odds ratios that do not interact with surgery type are summarized in Table 5, Part A. Odds ratios that differ by surgery type for at least one endpoint are presented in Table 5, Parts B, C, and D.

Final Model Intercept and Coefficients

The final risk prediction algorithms, including all coefficients and intercepts, are presented in the Appendix.

Limitations

The limitations for these valve models are similar to those for the CABG models and are thoroughly discussed in Part 1 of this series (2008 STS CABG risk models).

Conclusion

The STS Quality Measurement Task Force has developed and tested nine new risk-adjustment models for isolated valve surgery using the STS NCD. This report includes a detailed exposition of the model development process, including not only statistical issues but also the many clinical and pragmatic judgments that were required. An online risk calculator is also available through a link from the STS website.

References

1. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
2. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
3. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
4. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
5. Kuduvalli M, Grayson AD, Au J, Grotte G, Bridgewater B, Fabri BM. A multi-centre additive and logistic risk model for in-hospital mortality following aortic valve replacement. Eur J Cardiothorac Surg 2007;31:607–13.
6. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
7. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
8. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2002.
9. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.
10. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Models

For each endpoint, the formula for calculating a patient's predicted risk of the endpoint has the form:

$$\text{Predicted Risk} = \frac{e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}{1+e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}$$

where $x_1, x_2, \ldots, x_n$ denote patient preoperative risk factors (eg, quantitative variables such as age, and comorbidities coded as 1 = present, 0 = absent); and $\beta_0, \beta_1, \ldots, \beta_n$ denote regression coefficients (numerical constants). Regression coefficients for each endpoint are presented in Appendix Table 1. The variables $x_1, x_2, \ldots, x_n$ are the same for each endpoint and are defined in Appendix Table 2. The regression coefficient for the time trend is not presented. Instead, the intercept has been adjusted to incorporate the time trend. This adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−5.78680	−5.83957	−5.52789	−3.96796	−7.11095	−3.08816	−3.06527	−4.30676	1.25115
Atrial fibrillation	0.18074	0.05524	0.00000	0.16527	0.00000	0.10305	0.11403	0.15530	−0.30247
Age function 1	0.03557	0.03909	0.03219	0.02683	0.04180	0.01512	0.02041	0.02670	−0.02834
Age function 3	0.02804	−0.00132	0.01809	0.00629	−0.05024	0.00218	0.01282	0.02315	−0.04637
Age by reoperation function	−0.01308	−0.02043	−0.00551	−0.00840	−0.00939	−0.00697	−0.00684	−0.00485	0.00927
Age by status function	−0.02495	−0.02987	−0.00721	−0.01377	0.00277	0.00102	−0.00677	−0.00379	−0.00795
Age by MVR function	0.01436	0.00000	−0.00245	0.00000	0.00000	0.00715	0.00848	−0.00324	−0.00603
Age by MVRepair function	0.02326	0.00000	0.01190	0.00000	0.00000	0.00315	0.00685	0.01378	−0.01883
BSA function 1	−1.40168	−0.38619	−0.71012	−1.11750	0.14188	−0.73553	−0.91858	−0.82801	0.77317
BSA function 2	2.16782	0.23148	1.92875	2.29127	2.04603	0.83644	1.65638	1.65423	−1.76728
CHF but not NYHA IV	0.25590	0.00000	0.21233	0.28353	0.00000	0.00000	0.17974	0.22508	−0.15108
CHF and NYHA IV	0.60544	0.00000	0.47812	0.65056	0.00000	0.22686	0.48025	0.42957	−0.33521
CHF by MVR function	0.00000	0.00000	0.00000	−0.11007	0.00000	0.00000	−0.07864	0.00000	0.11503
CHF by MVRepair function	0.00000	0.00000	0.00000	−0.13792	0.00000	0.00000	−0.07731	0.00000	0.06468
CLD function	0.23846	0.00000	0.16629	0.22816	0.23817	0.08406	0.16044	0.25263	−0.21022
CLD by MVR function	−0.15906	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.10092	0.00000
CLD by MVRepair function	−0.03243	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.01795	0.00000
Creatinine function 1	0.43909	0.29230	0.71439	0.45646	0.00000	0.23562	0.49230	0.45631	−0.44178
CVD without prior CVA	0.00000	0.27837	0.20531	0.12726	0.00000	0.05830	0.07684	0.00000	−0.22223
CVD and prior CVA	0.00000	0.59220	0.20018	0.24512	0.00000	0.13200	0.18343	0.33480	−0.25595
Diabetes, noninsulin	0.23563	0.00000	0.37172	0.11040	0.00000	0.00000	0.11355	0.19843	−0.16630
Diabetes, insulin	0.48368	0.00000	0.64648	0.35367	0.44389	0.18293	0.33165	0.51913	−0.45093
Diabetes by MVR function	0.00000	0.00000	0.00000	0.15051	0.00000	0.00000	0.17990	0.00000	0.00000
Diabetes by MVRepair function	0.00000	0.00000	0.00000	0.16260	0.00000	0.00000	0.11734	0.00000	0.00000
Dialysis	1.48666	0.79199	0.00000	1.57690	1.19109	0.81972	1.37741	1.53351	−1.69019
Dialysis by MVR function	0.47550	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.20998
Dialysis by MVRepair function	0.78385	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.07964
Ejection fraction function	0.00904	0.00000	0.00407	0.01107	0.02308	0.00734	0.00925	0.01111	−0.01348
Endocarditis, active	0.66737	0.62434	0.77276	0.76318	0.00000	0.43876	0.67810	1.02521	−1.08299
Female	0.20372	0.21925	−0.03031	0.25668	−0.02355	−0.14567	0.03066	0.22437	−0.36400
Female by MVR function	−0.10089	0.00000	0.00000	−0.19465	0.00000	−0.08773	0.00000	−0.14211	0.00000
Female by MVRepair function	−0.23812	0.00000	0.00000	−0.04564	0.00000	0.04424	0.00000	0.02470	0.00000
Female by BSA function 1	0.96491	−0.02257	0.83074	0.77598	2.00214	0.16707	0.52716	0.57195	−0.75434
Female by BSA function 2	0.18084	−0.07419	0.08397	−0.58460	−1.87036	0.25158	−0.09063	−0.12289	0.35123
Hypertension	0.11372	0.17789	0.29770	0.10799	0.00000	0.00000	0.10361	0.00000	−0.06504
IABP or inotropes	0.38682	0.00000	0.29606	0.57608	0.52474	0.13432	0.56046	0.37621	−0.57115
IABP by MVR function	0.00000	0.00000	0.00000	0.21517	0.00000	0.00000	0.00000	0.00000	0.10760
IABP by MVRepair function	0.00000	0.00000	0.00000	0.48870	0.00000	0.00000	0.00000	0.00000	−0.13850

Appendix Table 1. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Immunosuppressive treatment	0.35022	0.00000	0.32828	0.00000	0.00000	0.00000	0.14887	0.27152	0.00000
Insufficiency mitral	0.00000	0.23253	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Insufficiency tricuspid	0.00000	0.00000	0.13159	0.12973	0.00000	0.08969	0.18929	0.15846	−0.20027
Left main disease	0.17593	0.00000	0.17280	0.00000	0.77557	0.00000	0.00000	0.00000	0.00000
MI ≤ 21 days	0.13276	0.00000	0.00000	0.31706	0.00000	0.03495	0.24687	0.18812	−0.20961
MVR	0.10284	0.00000	0.40455	0.44639	0.00000	0.12852	0.13795	0.58004	−0.61402
MVRepair	−0.65440	0.00000	−0.23666	−0.19726	0.00000	−0.22398	−0.23002	−0.37618	0.25710
No. diseased vessel function	0.00000	0.09556	0.00000	0.06299	0.00000	0.00000	0.03700	0.03312	−0.10126
Peripheral vascular disease	0.21980	0.25236	0.00000	0.00000	0.00000	0.19758	0.13174	0.15342	−0.18903
Race black	0.00000	0.28378	0.40941	0.34795	0.00000	0.23856	0.31567	0.37161	−0.44177
Race Hispanic	0.00000	−0.13774	0.14968	0.06720	0.00000	0.12816	0.08581	0.15128	−0.20068
Reop, 1 previous operation	0.74484	0.73489	0.43804	0.60704	0.00000	0.27365	0.44052	0.35252	−0.40042
Reop, ≥ 2 previous operations	0.90625	0.85841	0.50595	0.91229	0.00000	0.34233	0.67201	0.56294	−0.69765
Reop by MVR function	0.00000	0.00000	0.00000	−0.20333	0.00000	0.00000	0.00000	0.00000	0.00000
Reop by MVRepair function	0.00000	0.00000	0.00000	0.11559	0.00000	0.00000	0.00000	0.00000	0.00000
Shock	0.47961	0.50213	0.00000	0.73670	0.00000	0.28068	0.74786	0.55376	0.00000
Shock by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.50071	0.00000
Shock by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.36096	0.00000
Status urgent	0.25552	0.00000	0.19344	0.25714	0.00000	0.15548	0.19858	0.35184	−0.36106
Status emergent	1.32597	1.02109	1.13199	1.51294	0.00000	0.49075	1.17360	0.89480	−1.12373
Status salvage	2.07144	0.74530	1.24544	1.25342	0.00000	0.00000	1.21823	0.00000	−1.13785
Status by MVR function	−0.31729	0.00000	−0.34380	−0.35206	0.00000	0.00000	0.00000	0.00000	0.00000
Status by MVRepair function	0.84051	0.00000	−0.02373	0.29927	0.00000	0.00000	0.00000	0.00000	0.00000
Stenosis aortic	0.00000	0.00000	0.00000	−0.10782	0.00000	−0.10852	−0.07479	−0.15434	0.06873
Stenosis mitral	0.21309	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Unstable angina	0.18950	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Appendix Table 2. Definition of Variables Appearing in STS 2008 Valve Models

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
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
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 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 ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise

Appendix Table 2. Continued

Variable	Definition
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 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
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Stenosis mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina, no MI within 7 days of surgery, = 0 otherwise

Note: See www.sts.org for exact definitions of terms used above.

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; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PLOS = prolonged length of stay; Preop = preoperative; Reop = reoperation; Comp = composite adverse event (any); RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

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 #:** 1502 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery
De.2 **Brief description of measure:** Percent of patients undergoing combined MV Repair and CABG who die, including both 1) all deaths occurring during the hospitalization in which the procedure was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure
1.1-2 **Type of Measure:** Outcome 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-634282066928233034.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:** Public reporting, Internal quality improvement	**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, Frequently performed procedure, Leading cause of morbidity/mortality, High resource use, Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** The incidence of mitral valve incompetence as a result of or coexistent with coronary artery disease is increasing as a result of a progressively older population of patients coming to medical attention and recommendations for earlier surgical intervention among other factors. Patients undergoing combined coronary bypass and mitral valve repair (CABG/MV Replace), in addition, increasingly have a larger number and severity of co-morbid risk factors. As a result, these patients have among one of the highest mortality rates for all surgical procedures. As a direct outcome measure, surgical mortality following CABG/MV Repair is a time-tested and well-accepted gold standard for the quality of this procedure. Given the demographics of heart disease as well as the increasing scientific evidence for earlier intervention, the impact of this measure is expected to increase considerably. 1a.4 **Citations for Evidence of High Impact:** - Birkmeyer NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437. - Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892.	**1a** C☐ P☐ M☐ N☐

- Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137.
- Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467.
- Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009 Jul; 88(1 Suppl):S43-62.
- Miyata H, Motomura N, Tsukihara H, Takamoto S; Japan Cardiovascular Surgery Database. Risk models including high-risk cardiovascular procedures: clinical predictors of mortality and morbidity. Eur J Cardiothorac Surg. 2010 Nov 1
- Vassileva CM, Boley T, Markwell S, Hazelrigg S. Meta-analysis of short-term and long-term survival following repair versus replacement for ischemic mitral regurgitation. Eur J Cardiothorac Surg. 2010 Aug 18.
- Daneshmand MA, Milano CA, Rankin JS, Honeycutt EF, Shaw LK, Davis RD, Wolfe WG, Glower DD, Smith PK. Influence of patient age on procedural selection in mitral valve surgery. Ann Thorac Surg. 2010 Nov; 90(5):1479-85

1b. Opportunity for Improvement

1b.1 **Benefits (improvements in quality) envisioned by use of this measure:** The reporting of this measure, particularly in a risk-adjusted format, continues to be increasingly beneficial to individual providers and their affiliated healthcare organizations as they assess and compare their own results with those of similar organizations and with best practices nation-wide. Through public reporting of such a measure, the process of quality improvement becomes more transparent in order to reduce the existing variability and improve overall outcomes across all hospitals offering this procedure. In addition, this measure will anchor current outcomes as new technologies and treatment therapies are developed in the future. This includes such devices as endovascular valve procedures as well as newer, non-invasive treatments currently under development.

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, 2005-December 31, 2009

Analysis includes 130 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data (not restricted to this measure) to STS for at least 36 months in 2005-2009.

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

1b.5 **Citations for data on Disparities:**

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*)**:** Mortality following surgery is a direct outcome measure. Risk-adjustment methodology makes this measure more impactful from a quality improvement perspective.

1c.2-3. **Type of Evidence:** Observational study, Expert opinion, Systematic synthesis of research, Other
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*)**:**
Surgery for ischemic heart disease and for mitral valve insufficiency have evolved over the last several decades and currently represents the best form of treatment for these conditions depending on the stage of

1c
C☐
P☐
M☐
N☐

progression. Initial outcomes, including mortality, have substantially been reduced over this time period through a number of technical innovations, refined selection criteria, and improved preoperative management. While risk-adjusted mortality with CABG/MV Repair is currently low and declining, ongoing modifications in technique and improved patient selection are necessary. Furthermore, persistent variability in outcomes between providers, even when risk-adjusted, indicates the presence of considerable opportunity for improvement.

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 NJ, Marrin CA, et al. Decreasing mortality for aortic and mitral valve surgery in Northern New England. Northern New England Cardiovascular Disease Study Group. Ann Thorac Surg. 2000;70(2):432-437.

- Edwards FH, Peterson ED, et al. Prediction of operative mortality following valve replacement surgery. JACC. 37:3:885-892.
- Goodney PP, O'Connor GT, et al. Do hospitals with low mortality rates in coronary artery bypass also perform well in valve replacement? Ann Thorac Surg. 2003;76:1131-1137.
- Mehta RH, Eagle KA, et al. Influence of age on outcomes in patients undergoing mitral valve replacement. Ann Thorac Surg. 2002;74:1459-1467.
- Vassileva CM, Boley T, Markwell S, Hazelrigg S. Meta-analysis of short-term and long-term survival following repair versus replacement for ischemic mitral regurgitation. Eur J Cardiothorac Surg. 2010 Aug 18.
- Daneshmand MA, Milano CA, Rankin JS, Honeycutt EF, Shaw LK, Davis RD, Wolfe WG, Glower DD, Smith PK. Influence of patient age on procedural selection in mitral valve surgery. Ann Thorac Surg. 2010 Nov; 90(5):1479-85
- Miyata H, Motomura N, Tsukihara H, Takamoto S; Japan Cardiovascular Surgery Database. Risk models including high-risk cardiovascular procedures: clinical predictors of mortality and morbidity. Eur J Cardiothorac Surg. 2010 Nov 1.
- Murphy MO, Rao C, Punjabi PP, Athanasiou T. In patients undergoing mitral surgery for ischaemic mitral regurgitation is it preferable to repair or replace the mitral valve? Interact Cardiovasc Thorac Surg. 2010 Nov 18.
- Bouma W, van der Horst IC, Wijdh-den Hamer IJ, Erasmus ME, Zijlstra F, Mariani MA, Ebels T. Chronic ischaemic mitral regurgitation. Current treatment results and new mechanism-based surgical approaches. Eur J Cardiothorac Surg. 2010 Jan; 37(1):170-85.
- Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg. 2009 Jul; 88(1 Suppl):S43-62.
- Sajja LR, Mannam G, Dandu BR, Pathuri S, Sompalli S, Anjaneyulu A. Outcomes of mitral valve repair for chronic ischemic mitral regurgitation. Asian Cardiovasc Thorac Ann. 2009 Jan; 17(1):29-34.

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

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

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

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

1c.14 **Rationale for using this guideline over others:**	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
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***):** Number of patients undergoing combined MV Repair and CABG who die, including both 1) all deaths occurring during the hospitalization in which the operation was performed, even if after 30 days, and 2) those deaths occurring after discharge from the hospital, but within 30 days of the procedure 2a.2 **Numerator Time Window (***The time period in which cases are eligible for inclusion in the numerator***):** During hospitalization regardless of length of stay or within 30 days of surgery if discharged 2a.3 **Numerator Details (***All information required to collect/calculate the numerator, including all codes, logic, and definitions***):** Number of MV Repair + CABG procedures with an operative mortality; Number of MV Repair + CABG procedures in which Mortality [Mortalty (STS Adult Cardiac Surgery Database Version 2.73)] and Mortality Operative Death (MtOpD) are marked "yes." Operative mortality is further verified by the following variables: Mortality Status at 30 days (Mt30Stat), Mortality Date (MtDate), Mortality Discharge Status (MtDCStat)	
2a.4 **Denominator Statement (***Brief, text description of the denominator - target population being measured***):** All patients undergoing combined MV Repair + CABG 2a.5 **Target population gender:** Female, Male 2a.6 **Target population age range:** 18 yrs and older 2a.7 **Denominator Time Window (***The time period in which cases are eligible for inclusion in the denominator***):** 60 months 2a.8 **Denominator Details (***All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):** Number of MV Repair + CABG procedures; MV Repair + CABG is determined as a procedure for which all of the following apply: - OpCAB is marked as "Yes" - OpValve is marked "Yes" - VSMV is marked "Yes"	**2a-specs** **C☐** **P☐** **M☐** **N☐**

- VSMVPr is marked "Repair"
- (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
- VSAV, VSAVPr, ResectSubA, OpTricus, OpPulm, OpONCard, OCarLVA, OCarVSD, OCarSVR, OCarCong, OCarTrma, OCarCrTx, OCAoProcType, EndoProc, OCTumor, OCPulThromDis, OCarOthr are all marked "no" or "missing"

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

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

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

2a.12-13 Risk Adjustment Type: Case-mix adjustment

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

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment 2a.15 Detailed Risk Model-634282068151467310.pdf

2a.18-19 Type of Score: Rate/proportion
2a.20 Interpretation of Score: Better quality = Lower score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**

2a.22 Describe the method for discriminating performance *(e.g., significance testing)*:
Participant specific OR and their 95% CI were estimated in the hierarchical model. These model-based estimates were used to control variation due to random statistical fluctuations while estimating true signal variation. A 95% CI excluding zero indicates the participant's performance is significantly lower or higher than an "average" STS participant.

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)*:

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: national, Population: regional/network, Population: states, Population: counties or cities

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)*: STS Adult Cardiac Surgery Database - Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
Compared results between two proximate time periods with one-year overlap: January 2005-December 2007 and January 2007-December 2009. Excluded from analysis are participants that did not submit results for both time periods. Because database participants can change their underlying care processes at any time, we would not expect perfect correlation between two sets of results from even proximate time periods.

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

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 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.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Mortality Operative Death: 100.0% agreement rate

2c
C☐
P☐
M☐
N☐

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):

2d.2 Citations for Evidence:

2d.3 Data/sample *(description of data/sample and size)*:

2d
C☐
P☐
M☐
N☐
NA☐

2d.4 Analytic Method *(type analysis & rationale)*:

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*:

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample *(description of data/sample and size)*: Please see Risk Adjustment Type section above

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
Detailed information regarding the risk adjustment model can be found in the attachment:

Shahian DM, O´Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3--valve plus coronary artery bypass grafting surgery. Ann Thorac Surg 2009 Jul;88(1 Suppl):S43-62.

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)*: 130 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009; January 1, 2005-December 31, 2009

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:
We calculated the risk adjusted event rate with the participant's Odds Ratio (OR) estimate and the overall STS event rate. Therefore, the risk adjusted rate is closely related to OR estimate. If OR > 1, then the participant's risk adjusted rate will be greater than the overall STS event rate; if OR < 1, then the participant's risk adjusted rate will be smaller than the overall STS event rate. The statistical significance is defined by the 95% confidence interval (CI) or the OR estimate. If the 95% CI for a participant's OR includes the null value 1.0, then we cannot distinguish this participant's performance from the STS average – either the participant's performance was close to average or else the participant's sample size was too small to make a reliable inference. Otherwise, if the 95% CI falls to the right of 1.0, then the participant's performance is considered significantly lower than the average STS results; if the 95% CI falls to the left of 1.0, then the participant's performance is considered significantly higher than the average STS results.

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

2f
C ☐
P ☐
M ☐
N ☐

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)*:

2h
C ☐
P ☐

2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:	**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 STS plans to publicly report more measures in the future. There is no definite date yet assigned to this measure; however, STS staff and surgeon leadership have engaged in initial internal STS discussions regarding this matter. **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***):** **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	**3a** **C☐ P☐ M☐ N☐**
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	**3b** **C☐**

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	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*)	**4b** C☐ P☐ M☐

Yes **4b.2 If not, specify the near-term path to achieve electronic capture by most providers.**	**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. Please see IFMC audit results.	**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	**4e** **C** ☐ **P** ☐ **M** ☐ **N** ☐

members are charged the lesser of the two fees. 4e.3 **Evidence for costs:** 4e.4 **Business case documentation:**	
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:** This measure has been separated from NQF #122. In addition, age is now 18 years and older
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 1502 Sections 2a.14, 1b.2, 2b.3, 2f.3, 3a.6.pdf
Date of Submission (*MM/DD/YY*)**:** 01/12/2011

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.

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

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 valve procedure in the time period January 2002 – December 2006. For this measurement we re-fit the patient-level model using the latest four and a half years of data (January 2006 – 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 Valve plus CABG Models

Variable	Definition
Intercept	= 1 for all patients
Preoperative AFib	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine by MVR function	= Creatinine function 1 if valve operation is MVR, = 0 otherwise
Creatinine by MVRepair function	= Creatinine function 1 if valve operation is MVRepair, = 0 otherwise
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
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
EF by MVR function	= Ejection fraction function if valve operation is MVR, = 0 otherwise

EF by MVRepair function	= Ejection fraction function if valve operation is MVRepair, = 0 otherwise
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Endocarditis by MVR function	= 1 if patient has active endocarditis and valve operation is MVR, = 0 otherwise
Endocarditis by MVRepair function	= 1 if patient has active endocarditis and valve operation is MVRepair, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP or inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP or inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient has received immunosuppressive therapy within 30 days, = 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 1–21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI ≤ 21 days [a]	= 1 if patient has history of MI within 21 days prior to surgery, = 0 otherwise (for CVA and PLOS; coded as < 24 hours and 1–21 days for others)
MI < 24 hours	= 1 if history of MI < 24 hours prior to surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased coronary vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 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
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
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not rescuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis, mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina and no MI within 7 days of surgery, = 0 otherwise

[a] **MI coded as < 24 hours and 1 to 21 days.**

BSA = body surface area; CABG = coronary artery bypass graft surgery; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease;

The final patient-level model was built by step-wise selection method with several variables decided by surgeon panel 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	*Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery*
N	130
Mean	1.0
1st	1.5
5th	1.4
10th	1.2
25th	1.1
Median	0.9
75th	0.8
90th	0.7
95th	0.7
99th	0.5
Outlier	3 (2.3)
High	3
Low	0

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).

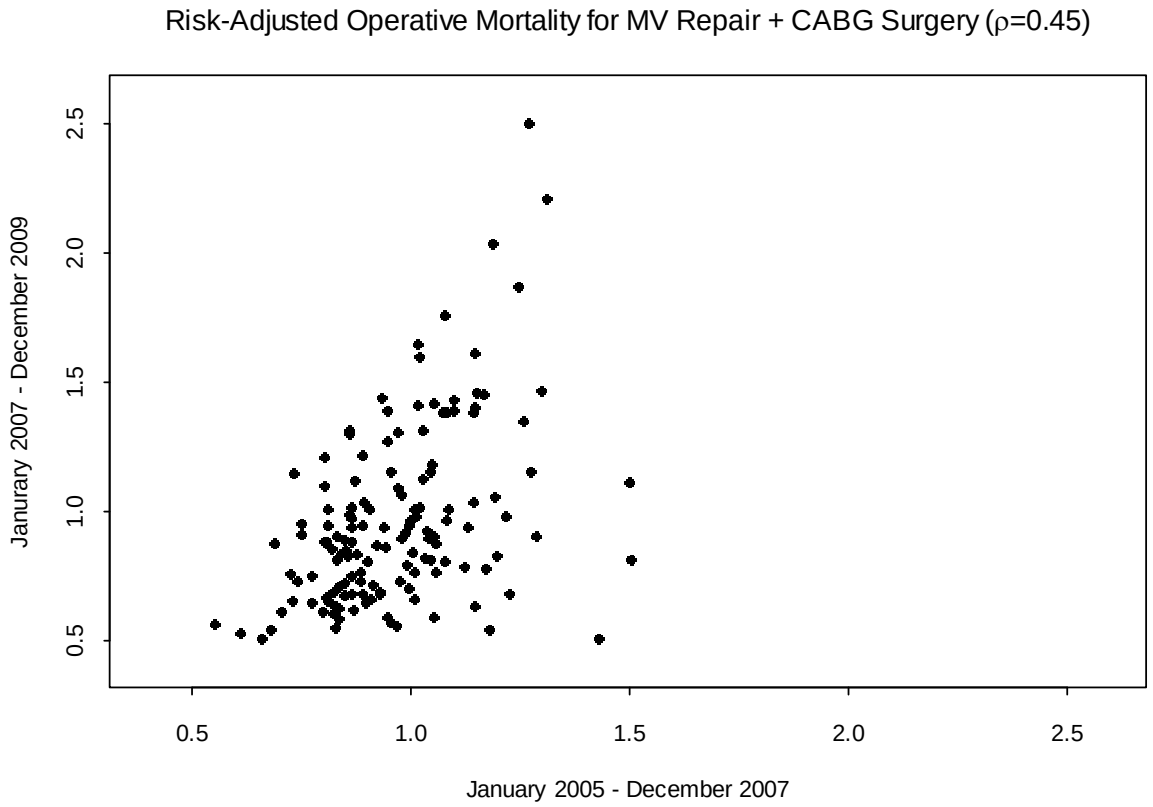
Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery*
N	130
Mean	5.9
1st	3.5

Measurement	*Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery*
5th	4.4
10th	4.5
25th	5.0
Median	5.7
75th	6.5
90th	7.4
95th	8.1
99th	8.7
Outlier	3 (2.3)
High	3
Low	0

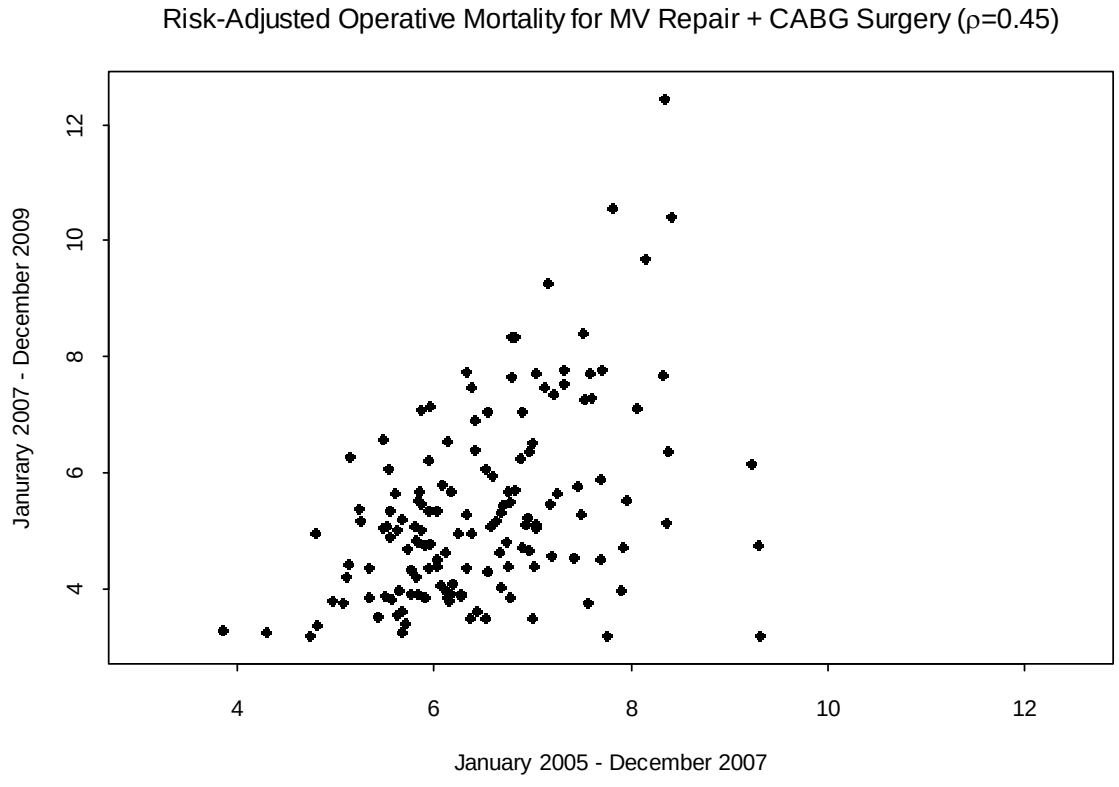
2b.3. Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*

Testing results: $\rho = 0.45$



Risk Adjusted Rate:

Testing results: $\rho = 0.45$



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, 2005-December 31, 2009. Sample contains 130 STS Adult Cardiac Surgery Database Participants who had at least 50 eligible cases for the measure and reported data to STS for at least 36 months in 2005-2009.

Measurement	*Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery*
N	130
Mean	1.0
1st	1.5
5th	1.4
10th	1.2
25th	1.1
Median	0.9
75th	0.8
90th	0.7
95th	0.7
99th	0.5
Outlier†	3 (2.3)
High	3
Low	0

Risk Adjusted Rate:

Measurement	*Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery*
N	130
Mean	5.9
1st	3.5
5th	4.4
10th	4.5
25th	5.0
Median	5.7
75th	6.5
90th	7.4
95th	8.1
99th	8.7
Outlier†	3 (2.3)

Measurement	*Risk-Adjusted Operative Mortality for MV Repair + CABG Surgery*
High	3
Low	0

†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."

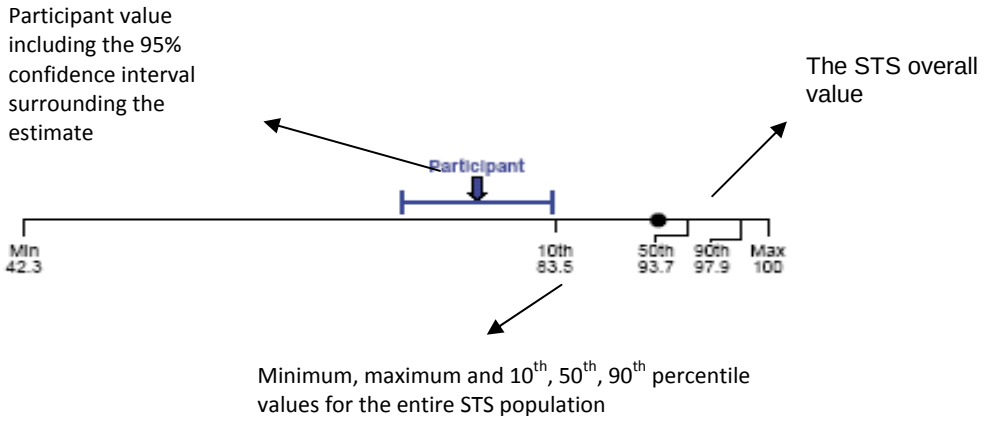
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

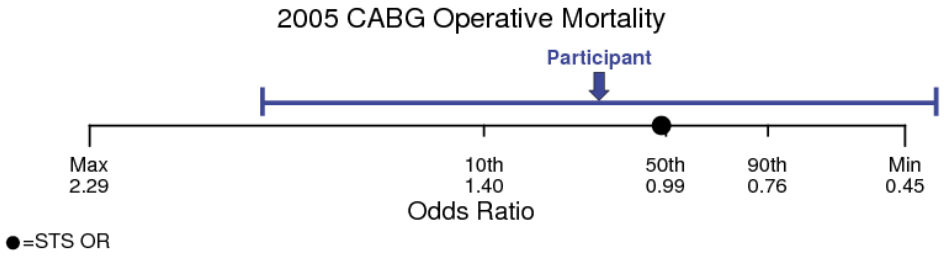
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

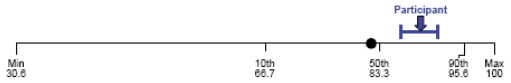
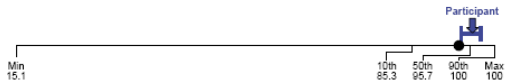
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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.

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 3—Valve Plus Coronary Artery Bypass Grafting Surgery

David M. Shahian, MD,[a] Sean M. O'Brien, PhD,[b] Giovanni Filardo, PhD, MPH,[c] Victor A. Ferraris, MD,[d] Constance K. Haan, MD,[e] Jeffrey B. Rich, MD,[f] Sharon-Lise T. Normand, PhD,[g] Elizabeth R. DeLong, PhD,[b] Cynthia M. Shewan, PhD,[h] Rachel S. Dokholyan, MPH,[b] Eric D. Peterson, MD, MPH,[b] Fred H. Edwards, MD,[e] and Richard P. Anderson, MD[i†]

[a]Massachusetts General Hospital, Boston, Massachusetts; [b]Duke Clinical Research Institute, Durham, North Carolina; [c]Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; [d]University of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; [e]University of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; [f]Sentara Cardiovascular Research Institute, Norfolk, Virginia; [g]Department of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; [h]The Society of Thoracic Surgeons, Chicago, Illinois; and [i]Seattle, Washington

Background. Since 1999, The Society of Thoracic Surgeons (STS) has published two risk models that can be used to adjust the results of valve surgery combined with coronary artery bypass graft surgery (CABG). The most recent was developed from data for patients who had surgery between 1994 and 1997 using operative mortality as the only endpoint. Furthermore, this model did not specifically consider mitral valve repair plus CABG, an increasingly common procedure. Consistent with STS policy of periodically updating and improving its risk models, new models for valve surgery combined with CABG have been developed. These models specifically address both perioperative morbidity and mitral valve repair, and they are based on contemporary data.

Methods. The final study population consisted of 101,661 procedures, including aortic valve replacement (AVR) plus CABG, mitral valve replacement (MVR) plus CABG, or mitral valve repair (MVRepair) plus CABG between January 1, 2002, and December 31, 2006. Model outcomes included operative mortality, stroke, deep sternal wound infection, reoperation, prolonged ventilation, renal failure, composite major morbidity or mortality, prolonged postoperative length of stay, and short postoperative length of stay. Candidate variables were screened for frequency of missing data, and imputation techniques were used where appropriate. Stepwise variable selection was employed, supplemented by advice from an expert panel of cardiac surgeons and biostatisticians. Several variables were forced into models to insure face validity (eg, atrial fibrillation for the permanent stroke model, sex for all models). Based on preliminary analyses of the data, a single model was employed for valve plus CABG, with indicator variables for the specific type of procedure. Interaction terms were included to allow for differential impact of predictor variables depending on procedure type. After validating the model in the 40% validation sample, the development and validation samples were then combined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regression model was estimated using generalized estimating equations to account for clustering of patients within institutions.

Results. The c-index for mortality prediction for the overall valve plus CABG population was 0.75. Morbidity model c-indices for specific complications (permanent stroke, renal failure, prolonged ventilation > 24 hours, deep sternal wound infection, reoperation for any reason, major morbidity or mortality composite, and prolonged postoperative length of stay) for the overall group of valve plus CABG procedures ranged from 0.622 to 0.724, and calibration was excellent.

Conclusions. New STS risk models have been developed for heart valve surgery combined with CABG. These are the first valve plus CABG models that also include risk prediction for individual major morbidities, composite major morbidity or mortality, and short and prolonged length of stay.

(Ann Thorac Surg 2009;88:S43–62)

Risk models for cardiac surgery were first developed almost 2 decades ago, and most of these early models focused on isolated coronary artery bypass graft surgery (CABG) [1–4]. The results of this frequently performed surgical procedure have often been used as the sole marker to assess the quality of care delivered by cardiac surgical programs. Risk-adjusted results for CABG have been used

Drs Shahian, O'Brien, Filardo, Ferraris, Haan, Rich, Normand, DeLong, Shewan, Peterson, Edwards, Anderson, and Ms Dokholyan, have no conflicts of interest to declare regarding this work.

†This author is deceased. Former Chair, Quality, Research and Patient Safety Council, The Society of Thoracic Surgeons, Chicago, IL.

Address correspondence to Dr Shahian, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.055

Abbreviations and Acronyms

AVR	= aortic valve replacement
CABG	= coronary artery bypass graft surgery
MI	= myocardial infarction
MVR	= mitral valve replacement
MVRepair	= mitral valve repair
NCD	= National Adult Cardiac Surgery Database
QMTF	= Quality Measurement Task Force
STS	= The Society of Thoracic Surgeons

for hospital and regional quality improvement initiatives, public reporting, pay for performance reimbursement programs, decision support, patient counseling, and clinical research. Earlier models focused primarily on mortality prediction, but subsequent models have been developed for both risk-adjusted morbidity and length of stay [5].

The other commonly performed category of cardiac surgery consists of operations on the heart valves, either alone or in combination with CABG. Relative to isolated CABG procedures, which are declining in frequency, the proportion of valve cases is steadily increasing. To better assess the overall performance of cardiac surgery programs, to discern the factors that are most significantly related to patient outcomes, and to aid in physician and patient decision-making, risk models have now also been developed for heart valve surgery [6–18].

Unlike risk models for isolated CABG, a relatively standardized procedure, valve surgery encompasses a much more diverse group of operations. There are four cardiac valves, and they may malfunction in a number of quite different ways (eg, stenosis, regurgitation, infection, and so forth). The valves may be repaired or replaced with a wide range of techniques and prosthetics. In some cases, procedures may be performed on multiple valves, or the valve procedure may be combined with CABG.

Given the heterogeneity of heart valve surgery, it is not surprising that a variety of risk-modeling techniques has been applied. At one extreme, the European System for Cardiac Operative Risk Evaluation (EuroSCORE) algorithm, developed by a European consortium, groups all cardiac operations together in a single risk model with indicator variables included to account for valve procedures [14, 18]. Although this approach is simple and easy to apply, recent studies by van Gameren and associates [19] have suggested that a dedicated valve risk model may have better discrimination and calibration than the EuroSCORE algorithm when applied to valve surgery patients. Combined models for aortic and mitral valve procedures with or without CABG have been developed by Jin and colleagues [12] and by Ambler and associates [13]. The 2001 valve models developed by The Society of Thoracic Surgeons (STS) [6] consisted of one model for all isolated valve procedures and one model for valve procedures combined with CABG, and a 2007 risk model derived from the New York Cardiac Surgery Reporting System used a similar stratification [8].

Unified valve models reflect the fact that many risk factors are common to both aortic and mitral valve surgery. They offer simplicity, and they also permit larger sample sizes for development and validation [12]. However, there are significant differences between aortic and mitral valvular disease in both pathophysiology and outcomes, and both also differ substantially from isolated CABG [11]. Some investigators advocate separate aortic and mitral valve models to have more homogeneous patient populations. Examples include models developed by STS, the New York Cardiac Surgery Reporting System, and the Northern New England Cardiovascular Disease Study Group [7, 9, 10]. Some of these models have been developed solely for isolated valve replacement, some have included CABG as a separate predictor variable in the isolated valve model, and some models have focused specifically on valve plus CABG. All these decisions involve a tradeoff—the more homogeneous the study group, the fewer patients are available for model development and validation [12].

Because of the large number of valve surgery patients available for analysis in the STS National Adult Cardiac Surgery Database (NCD), our approach has favored separate models for valve plus CABG versus isolated valve surgery. The STS Quality Measurement Task Force (QMTF) presumes that when adequate numbers of patients are available for study, relatively homogeneous operative categories result in more accurate risk prediction. Furthermore, recent studies by van Gameren and colleagues [19] suggest that the valve plus CABG group may be the most difficult to model accurately, thus meriting its own algorithm.

Several new features were added to the 2008 valve plus CABG models described in this report. First, recognizing that mitral valve repair is often different in both etiology and outcomes than replacement, the QMTF has included interactions between surgery type and several key predictor variables. Fitting a single model with several such interactions is useful. It allows for pooling information across related groups of valve procedures without making an a priori assumption that the effect of key risk factors is constant across these groups. Finally, new models have been developed for specific major complications of each valve plus CABG procedure, as well as for composite morbidity, mortality, and for both short and prolonged postoperative length of stay.

The authors of this report are members of the STS QMTF who were involved in this risk model development project.

Study Population and Endpoints

Our general approaches to variable selection and risk model development have been described in the companion articles on isolated CABG (Part 1) and isolated valve surgery (Part 2). Details specific to the valve plus CABG models are included in this report.

Study Population

The study population for this analysis consisted of single aortic or mitral valve surgical procedures combined with

CABG performed on adult patients between January 1, 2002, and December 31, 2006. Only the following procedures were included: (1) isolated aortic valve replacement (AVR) plus CABG; (2) isolated mitral valve replacement (MVR) plus CABG; and (3) isolated mitral valve repair (MVRepair) plus CABG.

Because of the relatively small number of pulmonic, tricuspid, multiple valve procedures, and aortic repairs, these cases were not included in the current models. Patients undergoing isolated valve surgery without CABG were excluded from the current analysis, but these cases are the focus of a separate model described in Part 2 of this three-part series. Patients with missing sex data (n = 17) were excluded because these patients are not allowed in the analysis dataset used for creating STS database participant feedback reports. Patients on dialysis preoperatively (n = 2,443) were excluded when developing the risk model for prediction of postoperative renal failure. The final study population comprised 101,661 patient operations (66,074 AVR plus CABG; 13,663 MVR plus CABG; and 21,924 MVRepair plus CABG) from 814 STS NCD participating groups.

Characteristics of the study population are summarized in Table 1.

Training and Validation Samples

The study population was randomly divided into a 60% training (development) sample and a 40% test (validation) sample. The development sample was used to identify predictor variables and estimate model coefficients. Data from the validation sample were used to assess model fit, discrimination, and calibration. After choosing variables and assessing model fit, the development and validation samples were subsequently combined, and the final model coefficients were estimated using the combined (development plus validation) data.

Endpoints

In developing the valve plus CABG risk models, we used the same nine endpoints that were analyzed in the STS isolated CABG (Part 1) and the STS isolated valve (Part 2) models. Morbidities in all three models are recorded only in-hospital, in contrast to the operative mortality endpoint defined below (although beginning with version 2.61, sternal infection will be recorded at 30 days): (1) operative mortality: death during the same hospitalization as surgery, regardless of timing or within 30 days of surgery regardless of venue; (2) permanent stroke (CVA): a central neurologic deficit persisting longer than 72 hours; (3) renal failure: a new requirement for dialysis or an increase of the serum creatinine to more than 2.0 mg/dL and double the most recent preoperative creatinine level; (4) prolonged ventilation (> 24 hours); (5) deep sternal wound infection; (6) reoperation for any reason; (7) major morbidity or mortality, a composite defined as the occurrence of any of the above endpoints; (8) prolonged postoperative length of stay (PLOS): length of stay (LOS) more than 14 days (alive or dead); and (4) short postoperative length of stay (SLOS): LOS less than 6 days and patient alive at discharge.

Endpoint frequencies in the study population are presented in Table 2.

Separate Versus Combined Models

Given the variety of approaches used in previous models by STS and other developers, we investigated the option of developing separate models for the AVR plus CABG and MVR plus CABG populations, and we also studied how best to subdivide the mitral plus CABG population into repair versus replacement. Although we had a large study population available, many of the individual outcomes were relatively rare. We were concerned that the number of events would be too small to permit reliable estimation of the model coefficients in separate models for each valve. Thus, in theory, the development of separate custom models for each valve type could be inferior to a single combined model because the custom models would have a smaller sample size and hence larger variance.

As described in detail in Part 2 of this series (isolated valve surgery), we performed preliminary empirical analyses to compare two alternative strategies (separate versus combined AVR plus CABG and MVR/Repair plus CABG) for developing these risk models. We first developed separate models for the three subpopulations (AVR plus CABG, MVR plus CABG, and MVRepair plus CABG), then modeled all three subpopulations together in a single model. In the latter approach, we included several interaction terms to allow the effect of certain risk factors to differ across the specific valve subpopulations. These strategies were used to develop risk models for operative mortality and permanent stroke, using a 60% development sample and a separate 40% validation sample. The performance of the combined model was then assessed separately within each subpopulation and compared to the model that was developed specifically for that subpopulation. In the case of mortality, the combined model had better discrimination (larger c-index) than the corresponding custom model in each of the three subpopulations (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). For stroke, the combined model had better discrimination in two of the three populations (all except AVR plus CABG). Finally, when explained variation was quantified by the generalized R^2 index of Nagelkerke [20], the combined model had greater explained variation than the custom model in each subpopulation for each endpoint. These results provide empirical support for the use of a single model with several interactions, which allows pooling of information across valve groups without assuming that the effect of risk factors is constant.

Selection of Candidate Predictor Variables

The candidate variables for the STS valve plus CABG models were identical to those in the STS isolated valve models, described in Part 2 of this series. They differed from the isolated CABG model variables in the following specific areas: (1) Percutaneous coronary intervention (PCI) occurring 6 hours or less before surgery was present in only 315 patients (0.3%) in the valve plus CABG study population, and was not included as a candidate variable. (2) Infectious endocarditis was not included in the isolated CABG model but was considered for the valve plus CABG model. Although this risk factor was rarely present (0.8% active

Table 1. Distribution of Risk Factors in Overall Study Population 2002 to 2006

	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
Variable	N	%	N	%	N	%	N	%
Demographics								
Age, years								
< 55	6,693	6.6	2,983	4.51	1,309	9.58	2,401	10.95
55–64	17,188	16.9	9,132	13.82	2,790	20.42	5,266	24.02
65–74	33,628	33.1	21,313	32.26	4,667	34.16	7,648	34.88
≥ 75	44,152	43.4	32,646	49.41	4,897	35.84	6,609	30.15
Sex								
Male	65,588	64.5	44,619	67.53	7,348	53.78	13,621	62.13
Female	36,073	35.5	21,455	32.47	6,315	46.22	8,303	37.87
Race								
Caucasian	90,572	89.1	60,121	90.99	11,765	86.11	18,686	85.23
Black	4,534	4.5	2,094	3.17	914	6.69	1,526	6.96
Hispanic	2,487	2.4	1,487	2.25	354	2.59	646	2.95
Asian	1,083	1.1	542	0.82	191	1.40	350	1.60
Other	2,295	2.3	1,402	2.12	331	2.42	562	2.56
Missing	690	0.7	428	0.65	108	0.79	154	0.70
Risk factors								
Body surface area, m^2								
< 1.50	3,340	3.3	1,985	3.00	638	4.67	717	3.27
1.50–1.74	20,779	20.4	12,580	19.04	3,500	25.62	4,699	21.43
1.75–1.99	40,017	39.4	25,814	39.07	5,440	39.82	8,763	39.97
≥ 2.00	36,956	36.4	25,361	38.38	3,996	29.25	7,599	34.66
Missing	569	0.6	334	0.51	89	0.65	146	0.67
Body mass index, kg/m^2								
< 25	29,353	28.9	17,712	26.81	4,787	35.04	6,854	31.26
25–29	39,345	38.7	25,692	38.88	4,951	36.24	8,702	39.69
30–34	21,063	20.7	14,447	21.86	2,507	18.35	4,109	18.74
≥ 35	11,165	11.0	7,785	11.78	1,299	9.51	2,081	9.49
Missing	735	0.7	438	0.66	119	0.87	178	0.81
Diabetes mellitus								
No diabetes	68,112	67.0	44,489	67.33	9,517	69.66	14,106	64.34
Diabetes, noninsulin	23,383	23.0	15,705	23.77	2,642	19.34	5,036	22.97
Diabetes, insulin	9,848	9.7	5,677	8.59	1,463	10.71	2,708	12.35
Diabetes, missing treatment	167	0.2	105	0.16	20	0.15	42	0.19
Missing	151	0.1	98	0.15	21	0.15	32	0.15
Hypertension								
No	22,709	22.3	13,944	21.10	3,482	25.48	5,283	24.10
Yes	78,823	77.5	52,050	78.78	10,163	74.38	16,610	75.76
Missing	129	0.1	80	0.12	18	0.13	31	0.14
Hypercholesterolemia								
No	33,759	33.2	21,248	32.16	5,324	38.97	7,187	32.78
Yes	67,613	66.5	44,649	67.57	8,280	60.60	14,684	66.98
Missing	289	0.3	177	0.27	59	0.43	53	0.24
Past or present smoker								
No	43,687	43.0	29,123	44.08	5,835	42.71	8,729	39.81
Yes	57,813	56.9	36,849	55.77	7,797	57.07	13,167	60.06
Missing	161	0.2	102	0.15	31	0.23	28	0.13
Chronic lung disease								
None	76,803	75.5	50,632	76.63	9,756	71.40	16,415	74.87
Mild	12,157	12.0	7,658	11.59	1,853	13.56	2,646	12.07
Moderate	7,797	7.7	4,720	7.14	1,269	9.29	1,808	8.25
Severe	4,005	3.9	2,463	3.73	658	4.82	884	4.03
Missing	899	0.9	601	0.91	127	0.93	171	0.78

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Peripheral vascular disease								
No	84,183	82.8	54,658	82.72	11,373	83.24	18,152	82.80
Yes	17,294	17.0	11,296	17.10	2,267	16.59	3,731	17.02
Missing	184	0.2	120	0.18	23	0.17	41	0.19
Cerebrovascular disease								
No	83,284	81.9	53,509	80.98	11,304	82.73	18,471	84.25
Yes	18,202	17.9	12,449	18.84	2,335	17.09	3,418	15.59
Missing	175	0.2	116	0.18	24	0.18	35	0.16
CVA								
No CVA	92,527	91.0	60,141	91.02	12,283	89.90	20,103	91.69
Remote CVA (> 2 weeks)	8,461	8.3	5,545	8.39	1,240	9.08	1,676	7.64
Recent CVA (≤ 2 weeks)	348	0.3	184	0.28	88	0.64	76	0.35
CVA, missing timing	114	0.1	62	0.09	23	0.17	29	0.13
Missing	211	0.2	142	0.21	29	0.21	40	0.18
Endocarditis								
No endocarditis	99,517	97.9	65,023	98.41	12,914	94.52	21,580	98.43
Treated endocarditis	1,091	1.1	525	0.79	356	2.61	210	0.96
Active endocarditis	827	0.8	387	0.59	356	2.61	84	0.38
Endocarditis, missing type	24	0.0	11	0.02	8	0.06	5	0.02
Missing	202	0.2	128	0.19	29	0.21	45	0.21
Renal failure								
No	92,592	91.1	60,880	92.14	12,037	88.10	19,675	89.74
Yes	8,888	8.7	5,072	7.68	1,605	11.75	2,211	10.08
Missing	181	0.2	122	0.18	21	0.15	38	0.17
Renal function								
Creatinine < 1.0 mg/dL	30,178	29.7	20,297	30.72	3,672	26.88	6,209	28.32
Creatinine 1.00–1.49 mg/dL	52,008	51.2	34,054	51.54	6,758	49.46	11,196	51.07
Creatinine 1.50–1.99 mg/dL	11,469	11.3	7,151	10.82	1,732	12.68	2,586	11.80
Creatinine 2.00–2.49 mg/dL	2,711	2.7	1,554	2.35	498	3.64	659	3.01
Creatinine ≥ 2.5 mg/dL	1,602	1.6	844	1.28	319	2.33	439	2.00
Dialysis	2,443	2.4	1,364	2.06	482	3.53	597	2.72
Missing	1,250	1.2	810	1.23	202	1.48	238	1.09
Immunosuppressive treatment								
No	98,421	96.8	63,984	96.84	13,211	96.69	21,226	96.82
Yes	2,975	2.9	1,904	2.88	427	3.13	644	2.94
Missing	265	0.3	186	0.28	25	0.18	54	0.25
Previous CV interventions								
Previous CABG surgery								
No	91,657	90.2	59,583	90.18	12,057	88.25	20,017	91.30
Yes	9,615	9.5	6,257	9.47	1,540	11.27	1,818	8.29
Missing	389	0.4	234	0.35	66	0.48	89	0.41
Previous valve surgery								
No	98,737	97.1	64,265	97.26	12,794	93.64	21,678	98.88
Yes	2,540	2.5	1,567	2.37	813	5.95	160	0.73
Missing	384	0.4	242	0.37	56	0.41	86	0.39
Previous other cardiac surgery								
No	98,538	96.9	64,166	97.11	13,181	96.47	21,191	96.66
Yes	2,683	2.6	1,634	2.47	407	2.98	642	2.93
Missing	440	0.4	274	0.41	75	0.55	91	0.42
Number of previous CV surgeries								
No previous CV surgery	89,419	88.0	58,161	88.02	11,530	84.39	19,728	89.98
1 prior CV surgery	10,453	10.3	6,796	10.29	1,799	13.17	1,858	8.47
≥ 2 prior CV surgeries	1,200	1.2	766	1.16	231	1.69	203	0.93
Missing	589	0.6	351	0.53	103	0.75	135	0.62

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Prior PCI								
No PCI	84,553	83.2	55,581	84.12	11,152	81.62	17,820	81.28
PCI ≤ 6 hours	315	0.3	151	0.23	89	0.65	75	0.34
PCI > 6 hours	16,158	15.9	9,946	15.05	2,321	16.99	3,891	17.75
PCI, missing timing	234	0.2	145	0.22	45	0.33	44	0.20
Missing	401	0.4	251	0.38	56	0.41	94	0.43
Preoperative cardiac status								
Acuity status								
Elective	62,298	61.3	43,682	66.11	7,277	53.26	11,339	51.72
Urgent	36,454	35.9	21,414	32.41	5,315	38.90	9,725	44.36
Emergent	2,479	2.4	763	1.15	945	6.92	771	3.52
Emergent salvage	258	0.3	97	0.15	104	0.76	57	0.26
Missing	172	0.2	118	0.18	22	0.16	32	0.15
MI								
No prior MI	68,332	67.2	49,673	75.18	8,056	58.96	10,603	48.36
MI ≤ 21 days	16,934	16.7	9,308	14.09	2,621	19.18	5,005	22.83
MI 8–21 days	3,751	3.7	1,725	2.61	624	4.57	1,402	6.39
MI 1–7 days	10,458	10.3	4,514	6.83	1,741	12.74	4,203	19.17
MI > 6 and < 24 hours	1,113	1.1	367	0.56	341	2.50	405	1.85
MI ≤ 6 hours	531	0.5	178	0.27	192	1.41	161	0.73
MI, missing timing	355	0.3	184	0.28	59	0.43	112	0.51
Missing	187	0.2	125	0.19	29	0.21	33	0.15
Angina								
No	42,542	41.8	28,032	42.43	6,248	45.73	8,262	37.68
Yes	58,967	58.0	37,945	57.43	7,394	54.12	13,628	62.16
Missing	152	0.1	97	0.15	21	0.15	34	0.16
Cardiogenic shock								
No	98,743	97.1	65,219	98.71	12,590	92.15	20,934	95.48
Yes	2,719	2.7	720	1.09	1,055	7.72	944	4.31
Missing	199	0.2	135	0.20	18	0.13	46	0.21
Resuscitation								
No	100,474	98.8	65,522	99.16	13,359	97.78	21,593	98.49
Yes	971	1.0	405	0.61	281	2.06	285	1.30
Missing	216	0.2	147	0.22	23	0.17	46	0.21
Arrhythmia								
No arrhythmia	83,856	82.5	56,040	84.81	9,992	73.13	17,824	81.30
AFib/flutter	13,386	13.2	7,533	11.40	2,940	21.52	2,913	13.29
Heart block	1,975	1.9	1,311	1.98	289	2.12	375	1.71
Sustained VT/VF	1,513	1.5	614	0.93	299	2.19	600	2.74
Arrhythmia, other	483	0.5	305	0.46	63	0.46	115	0.52
Arrhythmia, missing type	242	0.2	135	0.20	59	0.43	48	0.22
Missing	206	0.2	136	0.21	21	0.15	49	0.22
Preoperative IABP								
No	96,136	94.6	64,597	97.76	11,957	87.51	19,582	89.32
Yes	5,205	5.1	1,275	1.93	1,655	12.11	2,275	10.38
Missing	320	0.3	202	0.31	51	0.37	67	0.31
NYHA class								
I	9,839	9.7	6,934	10.49	1,103	8.07	1,802	8.22
II	24,830	24.4	17,808	26.95	2,524	18.47	4,498	20.52
III	42,593	41.9	28,079	42.50	5,458	39.95	9,056	41.31
IV	20,571	20.2	10,808	16.36	3,882	28.41	5,881	26.82
Missing	3,828	3.8	2,445	3.70	696	5.09	687	3.13

Table 1. Continued

	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
Variable	N	%	N	%	N	%	N	%
Congestive heart failure								
No	58,086	57.1	41,984	63.54	5,797	42.43	10,305	47.00
Yes	43,377	42.7	23,953	36.25	7,845	57.42	11,579	52.81
Missing	198	0.2	137	0.21	21	0.15	40	0.18
Number of diseased coronary vessels								
None	2,362	2.3	1,786	2.70	281	2.06	295	1.35
One	22,718	22.3	16,934	25.63	3,040	22.25	2,744	12.52
Two	27,144	26.7	19,014	28.78	3,655	26.75	4,475	20.41
Three	49,060	48.3	28,107	42.54	6,623	48.47	14,330	65.36
Missing	377	0.4	233	0.35	64	0.47	80	0.36
Left main disease ≥ 50%								
No	84,025	82.7	55,292	83.68	11,503	84.19	17,230	78.59
Yes	17,175	16.9	10,512	15.91	2,072	15.17	4,591	20.94
Missing	461	0.5	270	0.41	88	0.64	103	0.47
Ejection fraction, %								
< 25	5,805	5.7	2,199	3.33	640	4.68	2,966	13.53
25–34	10,988	10.8	4,877	7.38	1,566	11.46	4,545	20.73
35–44	14,928	14.7	8,064	12.20	2,487	18.20	4,377	19.96
45–54	20,398	20.1	13,424	20.32	3,048	22.31	3,926	17.91
≥ 55	43,556	42.8	32,973	49.90	5,209	38.12	5,374	24.51
Missing	5,986	5.9	4,537	6.87	713	5.22	736	3.36
Aortic stenosis								
No	42,831	42.1	8,527	12.91	12,974	94.96	21,330	97.29
Yes	58,317	57.4	57,319	86.75	535	3.92	463	2.11
Missing	513	0.5	228	0.35	154	1.13	131	0.60
Mitral stenosis								
No	95,696	94.1	63,862	96.65	11,166	81.72	20,668	94.27
Yes	4,993	4.9	1,542	2.33	2,366	17.32	1,085	4.95
Missing	972	1.0	670	1.01	131	0.96	171	0.78
Tricuspid stenosis								
No	100,093	98.5	65,060	98.47	13,402	98.09	21,631	98.66
Yes	275	0.3	154	0.23	57	0.42	64	0.29
Missing	1,293	1.3	860	1.30	204	1.49	229	1.04
Pulmonic stenosis								
No	99,484	97.9	64,693	97.91	13,348	97.69	21,443	97.81
Yes	122	0.1	85	0.13	14	0.10	23	0.10
Missing	2,055	2.0	1,296	1.96	301	2.20	458	2.09
Aortic insufficiency								
None	57,561	56.6	28,972	43.85	10,821	79.20	17,768	81.04
Trivial	9,243	9.1	6,573	9.95	1,023	7.49	1,647	7.51
Mild	13,828	13.6	11,082	16.77	1,156	8.46	1,590	7.25
Moderate	10,195	10.0	9,581	14.50	232	1.70	382	1.74
Severe	8,686	8.5	8,580	12.99	49	0.36	57	0.26
Missing	2,148	2.1	1,286	1.95	382	2.80	480	2.19
Mitral insufficiency								
None	41,756	41.1	38,790	58.71	1,297	9.49	1,669	7.61
Trivial	7,467	7.3	7,139	10.80	147	1.08	181	0.83
Mild	15,407	15.2	13,485	20.41	584	4.27	1,338	6.10
Moderate	14,987	14.7	4,842	7.33	2,790	20.42	7,355	33.55
Severe	20,516	20.2	527	0.80	8,743	63.99	11,246	51.30
Missing	1,528	1.5	1,291	1.95	102	0.75	135	0.62

Table 1. Continued

Variable	Overall Valve + CABG (n = 101,661)		AVR + CABG (n = 66,074)		MVR + CABG (n = 13,663)		MVRepair + CABG (n = 21,924)	
	N	%	N	%	N	%	N	%
Tricuspid insufficiency								
None	74,774	73.6	49,614	75.09	9,758	71.42	15,402	70.25
Trivial	7,972	7.8	5,454	8.25	839	6.14	1,679	7.66
Mild	11,505	11.3	7,060	10.68	1,631	11.94	2,814	12.84
Moderate	4,119	4.1	1,919	2.90	874	6.40	1,326	6.05
Severe	636	0.6	237	0.36	186	1.36	213	0.97
Missing	2,655	2.6	1,790	2.71	375	2.74	490	2.23
Pulmonic insufficiency								
None	91,715	90.2	59,891	90.64	12,275	89.84	19,549	89.17
Trivial	3,411	3.4	2,122	3.21	442	3.24	847	3.86
Mild	2,065	2.0	1,215	1.84	306	2.24	544	2.48
Moderate	326	0.3	165	0.25	70	0.51	91	0.42
Severe	49	0.0	25	0.04	11	0.08	13	0.06
Missing	4,095	4.0	2,656	4.02	559	4.09	880	4.01

AFib = atrial fibrillation; AVR = aortic valve replacement; CABG = coronary artery bypass graft; CV = cardiovascular; CVA = cerebrovascular accident (stroke); IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; VF = ventricular fibrillation; VT = ventricular tachycardia.

endocarditis) in the overall valve plus CABG population, it was included for consistency with the isolated valve model. Active endocarditis was present in 2.6% of patients undergoing mitral replacement plus CABG. (3) Mitral stenosis was rarely present among isolated CABG patients (0.35%). However, it was not uncommon (4.9%) among patients undergoing valve plus CABG surgery and was included as a candidate variable. It was present in 17.3% of mitral replacements and 5.0% of mitral repairs.

An indicator for valve procedure (AVR, MVR, MVRepair) was included in the combined valve plus CABG model, as previously noted.

Missing Data

Missing data are uncommon in the STS NCD, with a frequency of less than 1% missing for most variables. Model variables with more than 1% missing were ejection fraction

Table 2. Frequency of Endpoints in Overall Study Population 2002 to 2006

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall (AVR + CABG, MVR + CABG, MVRepair + CABG)									
N	101,661	101,661	99,218	101,661	101,661	101,661	101,661	101,661	101,661
Events	6,919	2,935	9,097	21,561	684	12,117	30,580	15,594	22,534
%	6.8	2.9	9.0	21.2	0.7	11.9	30.1	15.3	22.2
AVR + CABG									
N	66,074	66,074	64,710	66,074	66,074	66,074	66,074	66,074	66,074
Events	3,718	1,751	5,032	11,608	394	7,090	17,343	8,412	16,961
%	5.6	2.7	7.6	17.6	0.6	10.7	26.3	12.7	25.7
MVR + CABG									
N	13,663	13,663	13,181	13,663	13,663	13,663	13,663	13,663	13,663
Events	1,590	499	1,829	4,469	114	2,274	5,897	3,277	1,512
%	11.6	3.7	13.6	32.7	0.8	16.6	43.2	24.0	11.1
MVRepair + CABG									
N	21,924	21,924	21,327	21,924	21,924	21,924	21,924	21,924	21,924
Events	1,611	685	2,236	5,484	176	2,753	7,340	3,905	4,061
%	7.4	3.1	10.3	25.0	0.8	12.6	33.5	17.8	18.5

AVR = aortic valve replacement; CABG = coronary artery bypass graft surgery; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Table 3. List of Candidate Variables and Their Coding for STS Valve Plus CABG Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline truncated from below at 50 with knot at 75.
Ejection fraction	Linear; values > 50 mapped to 50
Body surface area[a]	Quadratic polynomial modeled separately for males and females. *Note: BSA < 1.4 and > 2.6 were mapped to those values, respectively.*
Creatinine	Linear (only for patients not on dialysis). *Note: Creatinine < 0.5 and > 5.0 mapped to those values, respectively.*
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval. Modeled as linear across the categories.
Binary variables	
Active infectious endocarditis	Yes/no
Dialysis	Yes/no
Preoperative atrial fibrillation	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Preop IABP or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no MI < 7 days)	Yes/no
Left main disease	Yes/no
Aortic stenosis	Yes/no
Mitral stenosis	Yes/no
Aortic insufficiency	Defined as at least moderate (yes/no)
Mitral insufficiency	Defined as at least moderate (yes/no)
Tricuspid insufficiency	Defined as at least moderate (yes/no)
Categorical variables	
Surgery type	3 groups: AVR + CABG, MVR + CABG, MVRepair + CABG
Chronic lung disease	Modeled as linear across categories (none, mild, moderate, severe)
CVD/CVA	3 groups: no CVD, CVD no CVA, CVD + CVA
Diabetes mellitus	3 groups: insulin diabetes, noninsulin diabetes, other or no diabetes
No. diseased coronary vessels	3 groups: < 2-vessel disease; 2-vessel disease; 3-vessel disease. Modeled as linear across the categories
MI	3 groups: < 24 hours, 1–21 days, > 21 days or no MI. *Note: groups 1 and 2 were subsequently collapsed for some models.*
Race	3 groups: black, Hispanic, other including Caucasian
Status	4 groups: elective, urgent, emergent no resuscitation, salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, ≥ 2 previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF and NYHA IV
Interaction terms	
Age by reoperation[a]	
Age by emergent status[a]	
Surgery type by each of the following:	Age, diabetes, dialysis, creatinine, reoperation, endocarditis, emergent status, CLD, CHF, EF, sex, shock, IABP/inotropes, mitral insufficiency, aortic insufficiency, mitral stenosis, aortic stenosis.

[a] These variables were forced into each model.

AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association.

(5.9%), New York Heart Association functional class (3.8%), tricuspid insufficiency (2.6%), aortic insufficiency (2.1%), mitral insufficiency (1.5%), and creatinine/dialysis (1.2%).

To make full use of the available data, binary risk factors were modeled as yes versus no or missing. Thus, missing values were analyzed as if the endpoint did not occur. Missing data on categorical variables were imputed to the lowest risk value, which, in most instances, was the mode. Missing data on continuous variables were imputed to the conditional median. For ejection fraction, we conditioned

on congestive heart failure and sex. For body surface area, we conditioned on sex. For serum creatinine, we conditioned on renal failure.

Although multiple imputation is generally preferred on statistical grounds [21], we chose single imputation for this analysis based largely on practical considerations, including computational intensity. Furthermore, the fraction of missing data was small, and single and multiple imputation would give similar results. Finally, multiple imputation is primarily used for calculating appropriate standard error estimates, but an adjustment to the standard errors would not impact our study results or the published risk algorithms. In a separate sensitivity analysis, we compared predicted risk estimates from our final models to risk estimates that were derived from analogous models using multiple instead of single imputation. For each endpoint, the relative difference in predicted risk was less than 6% (eg, an absolute difference of 5.0% versus 5.3%) for all patients in the development and validation samples, and it was less than 2% (eg, an absolute difference of 5.0% versus 5.1%) for 99% of patients. A summary of these analyses including regression coefficients and covariance matrices is available at www.sts.org/riskmodels.

Final Variable Selection Procedure

Variables were initially selected using an automated stepwise model selection algorithm. The stepwise procedure began with a model that included all of the candidate variables except for interaction terms. Age, body surface area, and month of surgery were forced into each model. As in the isolated CABG and isolated valve models described in Parts 1 and 2 of this series, month of surgery was used only to adjust for time trends in the frequency of adverse outcomes over the 5-year study period. We adjusted for this to reduce potential confounding by time trends when estimating regression coefficients for the variables that are of primary interest (ie, patient preoperative risk factors—see example in Part 1). Surgery date was categorized into 6-month intervals and modeled as a linear trend across the ordinal categories. Surgery date is not included in the final risk prediction algorithm, and a patient's predicted risk does not depend on it. The published intercept parameter has been adjusted to incorporate the time trend, and this adjusted intercept reflects the baseline risk for a reference period of July to December 2006.

Other variables were selected in a stepwise fashion using a significance criterion of 0.05 for entry and removal. Ordinal categorical variables were initially coded such that removing an indicator variable caused a category to be combined with the lowest risk category (the reference group). In the case of myocardial infarction (MI), there were two outcomes (permanent stroke, prolonged length of stay) in which "MI 1 to 21 days" was retained but "MI less than 24 hours" was removed. For these two cases, the two MI categories were replaced by the single category "MI 21 days or less." The stepwise procedure was performed separately for each endpoint. Multiple interaction terms consisting of predictor variable and surgery type were also evaluated, and two additional interaction terms (age by reoperation and age by emergent status) were forced into the models (see Tables 3 and 5).

The results of this initial selection process were then reviewed by surgeon members of the QMTF for face validity and consistency with previous STS or other valve models: (1) preoperative atrial fibrillation was forced into the model for permanent stroke; (2) an indicator variable for dialysis was forced into any model that included creatinine (this did not apply to the renal failure model, as patients with preoperative dialysis were excluded); (3) sex was forced into all models; and (4) each variable that interacted with surgery group was also included as a main effect.

After validating the model in the 40% validation sample, the development and validation samples were then rejoined, and the final model coefficients were estimated using the overall 100% combined sample. The final logistic regression model was estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [22]. An independence working correlation matrix was used to apply the generalized estimating equations. With this approach, the estimated regression coefficients were identical to those obtained using ordinary logistic regression, but the standard errors were adjusted to account for the clustered data structure.

Results

Risk Factors, Outcomes, and Predictor Variables

Table 1 presents the distribution of risk factors and endpoints in the overall 2002 to 2006 study population. Because there are three valve plus CABG categories, space limitations prevent display of the bivariate relationships for each predictor variable, endpoint, and valve plus CABG group. These are available upon request from STS.

Table 2 summarizes the overall frequency of adverse outcomes as well as the outcomes for the three major valve groups. Table 3 lists the candidate predictor variables and their coding schemes.

Assessment of Model Fit and Discrimination

The Hosmer-Lemeshow test was not employed to assess overall calibration. Large sample sizes make a significant p value almost inevitable, as all risk models are only approximations of reality [23]. Rather, we assessed calibration graphically by plotting observed versus predicted event rates within deciles of predicted risk in the development and validation samples (Fig 1). These plots were constructed for the overall sample and for subgroups based on surgery type (AVR plus CABG, MVR plus CABG, MVRepair plus CABG); age ($<$ 60, 60 to 79, $\geq$ 80 years); sex (male, female); diabetes mellitus (yes/no); status (elective, nonelective); and ejection fraction ($\leq$ 40, $>$ 40). Because of space constraints, only the overall sample results in the validation sample are presented. Additional results are available at www.sts.org/riskmodels.

In general, the models were well calibrated in the validation sample. The average absolute difference between observed versus predicted event rates across the decile categories ranged from 0.1% for deep sternal wound infec-

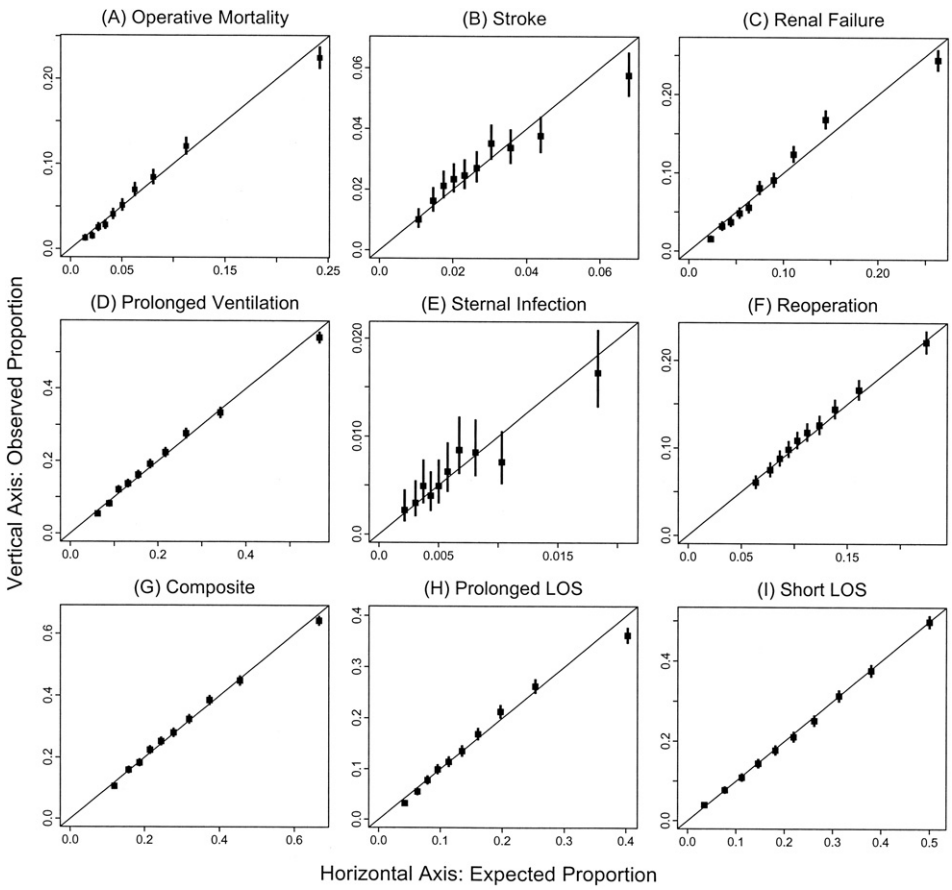


Fig 1. Plots of observed (O) versus expected (E) in validation sample

tion to 0.96% for prolonged length of stay. There was a very slight tendency for the models to overpredict risk in the highest decile. Although perfect prediction would be ideal, a slight overprediction implies that the model will give adequate credit to surgeons who take on patients with several model risk factors.

Discrimination was assessed by determining the c-statistic, also known as the area under the receiver operating characteristic (ROC) curve. Table 4 presents the discrimination of the various models. In the validation sample, the c-index of the overall valve plus CABG operative mortality model was 0.750, and the c-indices of the morbidity models ranged from 0.617 for reoperation to 0.724 for renal failure and short length of stay.

Table 4. Discrimination of Models (C-Index) in Development and Validation Samples

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall									
Development sample	0.754	0.656	0.729	0.730	0.670	0.623	0.704	0.719	0.726
Validation sample	0.750	0.622	0.724	0.720	0.646	0.617	0.698	0.710	0.724
AVR + CABG									
Development sample	0.737	0.648	0.720	0.706	0.639	0.607	0.678	0.705	0.700
Validation sample	0.736	0.609	0.718	0.697	0.657	0.604	0.673	0.699	0.698
MVR + CABG									
Development sample	0.764	0.665	0.712	0.746	0.713	0.608	0.725	0.694	0.726
Validation sample	0.739	0.611	0.701	0.733	0.580	0.599	0.714	0.680	0.733
MVRepair + CABG									
Development sample	0.746	0.650	0.727	0.725	0.692	0.624	0.707	0.712	0.738
Validation sample	0.755	0.652	0.715	0.716	0.644	0.623	0.705	0.702	0.733

AVR = aortic valve replacement; CABG = coronary artery bypass graft; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Table 5. Estimated Odds Ratios for CABG Mortality, Morbidity, and Length of Stay Models

A. Odds ratios for variables that do not interact with surgery group

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.12, 1.29)	1.05 (0.94, 1.17)	1.18 (1.11, 1.26)	1.13 (1.07, 1.19)	NA	1.16 (1.10, 1.22)	1.15 (1.10, 1.20)	1.22 (1.15, 1.28)	0.71 (0.67, 0.75)
BSA 1.6 versus 2.0 among females	1.29 (1.19, 1.39)	1.34 (1.18, 1.52)	0.87 (0.81, 0.94)	1.08 (1.02, 1.14)	0.51 (0.39, 0.67)	1.13 (1.07, 1.23)	1.12 (1.07, 1.18)	0.97 (0.92, 1.03)	1.03 (0.96, 1.10)
BSA 1.6 versus 2.0 among males	1.58 (1.41, 1.77)	1.38 (1.17, 1.64)	1.18 (1.07, 1.31)	1.31 (1.21, 1.41)	0.71 (0.49, 1.03)	1.18 (1.12, 1.34)	1.32 (1.24, 1.41)	1.40 (1.29, 1.52)	0.81 (0.75, 0.88)
BSA 1.8 versus 2.0 among females	1.05 (1.00, 1.10)	1.16 (1.06, 1.26)	0.89 (0.85, 0.93)	0.99 (0.95, 1.02)	0.69 (0.61, 0.77)	1.03 (0.98, 1.06)	1.01 (0.98, 1.04)	0.94 (0.90, 0.97)	1.08 (1.04, 1.12)
BSA 1.8 versus 2.0 among males	1.15 (1.10, 1.20)	1.13 (1.07, 1.20)	1.01 (0.97, 1.05)	1.06 (1.03, 1.09)	0.83 (0.72, 0.95)	1.06 (1.04, 1.11)	1.07 (1.05, 1.10)	1.09 (1.06, 1.12)	0.96 (0.94, 0.99)
BSA 2.2 versus 2.0 among females	1.12 (1.02, 1.22)	0.87 (0.74, 1.02)	1.25 (1.15, 1.35)	1.13 (1.06, 1.20)	1.57 (1.32, 1.89)	1.04 (1.00, 1.17)	1.10 (1.04, 1.17)	1.19 (1.11, 1.27)	0.82 (0.76, 0.89)
BSA 2.2 versus 2.0 among males	1.04 (1.00, 1.09)	0.95 (0.90, 1.01)	1.15 (1.11, 1.18)	1.09 (1.06, 1.11)	1.25 (1.14, 1.37)	1.00 (0.95, 1.01)	1.07 (1.04, 1.09)	1.09 (1.06, 1.12)	0.91 (0.89, 0.93)
CVD with CVA	1.22 (1.11, 1.33)	1.72 (1.52, 1.95)	1.12 (1.04, 1.22)	1.27 (1.19, 1.34)	1.22 (0.95, 1.56)	1.12 (1.04, 1.20)	1.26 (1.20, 1.33)	1.26 (1.18, 1.35)	0.75 (0.70, 0.81)
CVD without CVA	NA	1.28 (1.13, 1.45)	1.14 (1.06, 1.23)	1.10 (1.04, 1.16)	NA	NA	1.11 (1.05, 1.17)	1.11 (1.05, 1.18)	0.85 (0.78, 0.92)
Diabetes, insulin	1.31 (1.20, 1.42)	1.16 (1.03, 1.30)	1.62 (1.52, 1.74)	1.32 (1.25, 1.40)	1.98 (1.59, 2.46)	NA	1.34 (1.28, 1.41)	1.49 (1.40, 1.58)	0.67 (0.62, 0.72)
Diabetes, noninsulin	1.12 (1.05, 1.19)	1.16 (1.06, 1.26)	1.28 (1.21, 1.35)	1.11 (1.07, 1.15)	1.30 (1.10, 1.54)	NA	1.12 (1.08, 1.16)	1.17 (1.12, 1.22)	0.84 (0.81, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	1.15 (1.11, 1.19)	1.20 (1.14, 1.26)	1.17 (1.14, 1.21)	1.19 (1.16, 1.22)	1.28 (1.15, 1.42)	1.09 (1.06, 1.11)	1.16 (1.14, 1.18)	1.13 (1.10, 1.16)	0.82 (0.81, 0.84)
Hypertension	NA	1.19 (1.08, 1.31)	1.25 (1.18, 1.33)	1.10 (1.05, 1.15)	1.33 (1.09, 1.63)	NA	1.12 (1.08, 1.16)	1.08 (1.03, 1.13)	0.92 (0.88, 0.96)
Immunosuppressive treatment	1.35 (1.17, 1.54)	NA	1.30 (1.15, 1.47)	1.28 (1.17, 1.40)	NA	1.27 (1.14, 1.42)	1.26 (1.16, 1.37)	1.22 (1.11, 1.34)	0.75 (0.67, 0.84)
Left main disease	1.12 (1.05, 1.20)	NA	NA	1.06 (1.02, 1.11)	NA	NA	NA	NA	NA
Mitral insufficiency, moderate/severe	NA	NA	NA	NA	NA	NA	1.07 (1.01, 1.12)	NA	NA
Tricuspid insufficiency, moderate/severe	1.27 (1.15, 1.41)	NA	1.25 (1.13, 1.38)	1.15 (1.06, 1.24)	NA	NA	1.14 (1.07, 1.22)	NA	0.79 (0.69, 0.92)
Peripheral vascular disease	1.29 (1.21, 1.37)	1.15 (1.04, 1.27)	1.16 (1.10, 1.23)	1.18 (1.12, 1.24)	NA	1.15 (1.09, 1.22)	1.20 (1.15, 1.25)	1.16 (1.11, 1.22)	NA
Mitral stenosis	1.10 (0.99, 1.24)	NA	NA	NA	NA	NA	NA	1.09 (1.00, 1.18)	NA
MI 1–21 days	1.19 (1.10, 1.28)	NA	1.18 (1.10, 1.26)	1.28 (1.21, 1.35)	NA	NA	1.22 (1.16, 1.28)	NA	NA
MI $\leq$ 21 days[a]	NA	1.22 (1.11, 1.34)	NA	NA	NA	NA	NA	1.16 (1.10, 1.22)	NA
MI $<$ 24 hrs	1.65 (1.42, 1.91)	NA	1.30 (1.10, 1.54)	1.41 (1.23, 1.62)	NA	1.15 (1.00, 1.32)	1.49 (1.30, 1.70)	NA	NA
Time trend per 6-month harvest interval	0.98 (0.96, 0.99)	0.98 (0.97, 1.00)	1.01 (1.00, 1.02)	1.01 (1.00, 1.02)	0.96 (0.93, 0.99)	0.99 (0.98, 1.00)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)
Race black	NA	NA	1.15 (1.03, 1.30)	1.31 (1.19, 1.44)	NA	1.19 (1.06, 1.33)	1.21 (1.11, 1.32)	1.31 (1.19, 1.44)	0.65 (0.58, 0.72)
Race Hispanic	NA	NA	1.20 (1.03, 1.40)	1.17 (1.03, 1.32)	NA	1.08 (0.94, 1.24)	1.15 (1.03, 1.28)	1.13 (0.98, 1.30)	0.85 (0.71, 1.02)
Status, urgent versus elective	1.25 (1.17, 1.34)	NA	1.18 (1.10, 1.26)	1.26 (1.19, 1.33)	NA	1.14 (1.07, 1.21)	1.19 (1.14, 1.25)	1.28 (1.22, 1.35)	0.77 (0.72, 0.81)
Unstable angina	1.11 (1.03, 1.21)	0.89 (0.80, 1.00)	1.12 (1.05, 1.20)	1.06 (0.99, 1.13)	NA	NA	NA	NA	NA

Table 5. Continued

B. Odds ratios for AVR plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.29 (1.20, 1.39)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.20 (1.16, 1.24)	1.37 (1.32, 1.42)	0.74 (0.72, 0.77)
Age 70 versus 50 (no reop, elective)	1.67 (1.45, 1.92)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.44 (1.36, 1.54)	1.86 (1.73, 2.01)	0.55 (0.52, 0.59)
Age 80 versus 50 (no reop, elective)	2.47 (2.08, 2.94)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	1.86 (1.73, 2.01)	2.67 (2.46, 2.91)	0.33 (0.30, 0.36)
CHF, not NYHA IV	1.24 (1.14, 1.34)	0.98 (0.88, 1.09)	1.19 (1.11, 1.28)	1.22 (1.16, 1.29)	NA	NA	1.14 (1.08, 1.19)	1.30 (1.23, 1.38)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.48 (1.34, 1.64)	1.15 (1.00, 1.32)	1.35 (1.24, 1.48)	1.47 (1.36, 1.59)	NA	1.16 (1.08, 1.24)	1.36 (1.27, 1.45)	1.49 (1.39, 1.60)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	2.26 (2.13, 2.40)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.62 (0.58, 0.67)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.65 (1.41, 1.92)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.30 (0.25, 0.37)
EF per 10-unit decrease	1.10 (1.06, 1.15)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.87 (0.84, 0.89)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.41 (1.25, 1.58)	0.56 (0.48, 0.65)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.17 (0.92, 1.50)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	1.79 (1.50, 2.15)	1.45 (1.29, 1.63)	NA
Female versus male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.52 (1.44, 1.61)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.61 (0.57, 0.64)
Active infectious endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	1.81 (1.41, 2.32)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.26 (1.22, 1.30)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.26 (1.22, 1.30)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.83 (1.58, 2.11)	NA	1.39 (1.16, 1.67)	1.50 (1.32, 1.69)	1.55 (1.33, 1.81)	0.67 (0.58, 0.77)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	2.19 (1.80, 2.65)	NA	1.48 (1.15, 1.92)	1.77 (1.51, 2.06)	1.65 (1.34, 2.03)	0.53 (0.43, 0.65)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.33 (0.22, 0.50)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.18 (0.09, 0.34)

Table 5. Continued

C. Odds ratios for MVR plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.51 (1.39, 1.64)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.27 (1.21, 1.32)	1.37 (1.32, 1.42)	0.68 (0.64, 0.72)
Age 70 versus 50 (no reop, elective)	2.28 (1.94, 2.68)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.60 (1.47, 1.75)	1.86 (1.73, 2.01)	0.46 (0.41, 0.52)
Age 80 versus 50 (no reop, elective)	3.95 (3.17, 4.93)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	2.18 (1.92, 2.48)	2.67 (2.46, 2.91)	0.25 (0.21, 0.30)
CHF, not NYHA IV	0.91 (0.80, 1.03)	0.80 (0.64, 0.99)	0.92 (0.82, 1.03)	1.02 (0.93, 1.11)	NA	NA	0.94 (0.87, 1.02)	1.03 (0.94, 1.12)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.09 (0.95, 1.24)	0.93 (0.75, 1.17)	1.04 (0.92, 1.19)	1.22 (1.10, 1.35)	NA	1.16 (1.08, 1.24)	1.13 (1.03, 1.23)	1.17 (1.06, 1.30)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	1.82 (1.66, 2.01)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.66 (0.57, 0.78)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.21 (0.95, 1.55)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.30 (0.18, 0.48)
EF per 10-unit decrease	1.23 (1.16, 1.30)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.89 (0.82, 0.95)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.29 (1.14, 1.46)	0.51 (0.39, 0.65)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.21 (0.97, 1.50)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	2.76 (2.22, 3.42)	1.45 (1.29, 1.63)	NA
Female versus male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.17 (1.08, 1.28)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.66 (0.59, 0.74)
Active infectious endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	2.08 (1.62, 2.67)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.18 (1.12, 1.24)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.20 (1.14, 1.26)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.38 (1.19, 1.61)	NA	1.15 (0.95, 1.38)	1.50 (1.32, 1.69)	1.30 (1.10, 1.53)	0.81 (0.66, 0.99)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	1.66 (1.35, 2.03)	NA	1.22 (0.95, 1.56)	1.77 (1.51, 2.06)	1.38 (1.12, 1.71)	0.64 (0.50, 0.82)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.26 (0.16, 0.43)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.14 (0.07, 0.27)

Table 5. Continued

D. Odds ratios for MVRepair plus CABG

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reop, elective)	1.46 (1.36, 1.57)	1.28 (1.19, 1.38)	1.39 (1.32, 1.45)	1.23 (1.20, 1.27)	1.06 (0.92, 1.21)	1.19 (1.15, 1.23)	1.28 (1.23, 1.33)	1.37 (1.32, 1.42)	0.66 (0.63, 0.69)
Age 70 versus 50 (no reop, elective)	2.14 (1.86, 2.46)	1.64 (1.42, 1.91)	1.92 (1.75, 2.11)	1.52 (1.43, 1.62)	1.11 (0.85, 1.46)	1.41 (1.31, 1.51)	1.63 (1.51, 1.76)	1.86 (1.73, 2.01)	0.44 (0.40, 0.48)
Age 80 versus 50 (no reop, elective)	3.60 (2.97, 4.33)	2.03 (1.71, 2.42)	2.76 (2.47, 3.08)	1.96 (1.82, 2.11)	1.12 (0.82, 1.53)	1.67 (1.54, 1.82)	2.23 (2.00, 2.49)	2.67 (2.46, 2.91)	0.23 (0.20, 0.27)
CHF, not NYHA IV	0.96 (0.85, 1.09)	1.05 (0.90, 1.23)	0.99 (0.88, 1.10)	1.10 (1.02, 1.19)	NA	NA	1.06 (0.99, 1.14)	1.17 (1.08, 1.26)	0.84 (0.79, 0.89)
CHF, NYHA IV	1.16 (1.02, 1.32)	1.23 (1.04, 1.46)	1.12 (0.99, 1.27)	1.32 (1.21, 1.44)	NA	1.16 (1.08, 1.24)	1.27 (1.17, 1.37)	1.33 (1.22, 1.45)	0.73 (0.66, 0.82)
Creatinine per 1 unit	1.57 (1.49, 1.65)	1.27 (1.18, 1.36)	1.87 (1.72, 2.04)	1.46 (1.41, 1.52)	NA	1.28 (1.23, 1.34)	1.67 (1.60, 1.74)	1.51 (1.45, 1.58)	0.59 (0.53, 0.67)
Dialysis vs no dialysis and creatinine = 1.0	3.20 (2.84, 3.61)	1.42 (1.17, 1.73)	NA	2.27 (2.06, 2.51)	NA	1.88 (1.52, 2.31)	2.09 (1.91, 2.30)	2.42 (2.19, 2.67)	0.35 (0.24, 0.49)
EF per 10-unit decrease	1.09 (1.04, 1.15)	NA	1.06 (1.03, 1.08)	1.12 (1.10, 1.14)	NA	1.08 (1.05, 1.10)	1.11 (1.09, 1.13)	1.10 (1.08, 1.13)	0.84 (0.81, 0.87)
Preoperative IABP/ inotropes	1.43 (1.30, 1.58)	NA	1.27 (1.15, 1.39)	2.18 (2.01, 2.36)	NA	1.16 (1.06, 1.27)	1.76 (1.63, 1.90)	1.56 (1.40, 1.73)	0.52 (0.44, 0.62)
Shock	1.68 (1.45, 1.94)	1.19 (0.94, 1.50)	1.69 (1.41, 2.01)	1.93 (1.72, 2.16)	NA	1.24 (1.09, 1.41)	2.17 (1.81, 2.60)	1.45 (1.29, 1.63)	NA
Female vs male (at BSA = 1.8)	1.36 (1.26, 1.47)	1.19 (1.07, 1.32)	1.18 (1.10, 1.26)	1.25 (1.15, 1.36)	1.11 (0.88, 1.40)	0.92 (0.87, 0.97)	1.20 (1.15, 1.26)	1.31 (1.24, 1.38)	0.60 (0.55, 0.66)
Active infectious Endocarditis	2.04 (1.66, 2.50)	1.83 (1.37, 2.46)	1.52 (1.21, 1.91)	1.96 (1.69, 2.27)	NA	1.56 (1.28, 1.91)	2.11 (1.83, 2.44)	2.98 (1.86, 4.77)	0.28 (0.20, 0.38)
CLD (moderate vs mild or severe vs moderate)	1.19 (1.16, 1.23)	NA	1.12 (1.09, 1.15)	1.21 (1.16, 1.27)	1.32 (1.22, 1.42)	1.10 (1.07, 1.13)	1.18 (1.15, 1.21)	1.16 (1.10, 1.21)	0.83 (0.80, 0.85)
Reop, 1 previous operation[b]	2.20 (1.81, 2.67)	NA	1.29 (1.08, 1.55)	1.55 (1.32, 1.82)	NA	1.49 (1.23, 1.82)	1.50 (1.32, 1.69)	1.32 (1.10, 1.58)	0.80 (0.68, 0.95)
Reop, ≥ 2 previous operations[b]	2.46 (1.87, 3.24)	NA	1.47 (1.15, 1.89)	1.86 (1.53, 2.26)	NA	1.59 (1.20, 2.11)	1.77 (1.51, 2.06)	1.41 (1.11, 1.79)	0.63 (0.51, 0.79)
Status emergent, no resuscitation[b]	2.14 (1.62, 2.81)	2.21 (1.45, 3.37)	1.77 (1.31, 2.37)	2.71 (2.14, 3.44)	NA	1.41 (1.16, 1.70)	2.17 (1.74, 2.72)	2.72 (2.19, 3.38)	0.43 (0.29, 0.66)
Status emergent, with resuscitation or salvage[b]	4.56 (3.31, 6.29)	2.60 (1.53, 4.43)	1.86 (1.30, 2.65)	2.12 (1.54, 2.92)	NA	NA	3.34 (2.43, 4.61)	1.76 (1.31, 2.37)	0.23 (0.12, 0.44)

[a] For CVA and PLOS, MI coded ≤ 21 days; for all other endpoints, MI coded < 24 hrs or 1 to 21 days. [b] Variable interacts with age. Reported odds ratio represents effect of risk factor for patients aged 50 years old.

BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; Mort = mortality; NA = not applicable; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Odds Ratios

Table 5 presents the odds ratios and 95% confidence intervals (CI) derived from these models. "Not applicable" indicates that those predictors were not included in a particular risk model.

Odds ratios that do not interact with surgery type are summarized in Part A of Table 5. Several variables interact with surgery type, and the odds ratios for these variables differ for some of the endpoints depending on the specific type of surgery, as summarized in Tables 5B, C, and D (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). For example, in the model for prolonged length of stay, the odds ratio for active endocarditis is 1.81 (95% CI: 1.41 to 2.32) for AVR plus CABG; 2.08 (95% CI: 1.62 to 2.67) for MVR plus CABG; and 2.98 (95% CI: 1.86 to 4.77) for MVRepair plus CABG.

Final Model Intercept and Coefficients

The algorithms for calculating predicted risk values, including the intercepts and regression coefficients, are presented in the Appendix.

Limitations

The limitations of the STS valve plus CABG models are similar to those discussed in Part 1 of this series.

Conclusion

A new STS model has been developed for valve surgery combined with CABG. This model includes specific indicator variables for each major type of valve plus CABG procedure (AVR plus CABG, MVR plus CABG, MVRepair plus CABG). Models have been developed for operative mortality, individual morbidity endpoints, a composite morbidity or mortality endpoint, and short and prolonged postoperative length of stay. Overall model performance is excellent.

References

1. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA 1990;264:2768–74.
2. Edwards FH, Clark RE, Schwartz M. Coronary artery bypass grafting: the Society of Thoracic Surgeons National Database experience. Ann Thorac Surg 1994;57:12–9.
3. O'Connor GT, Plume SK, Olmstead EM, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. Circulation 1992;85:2110–8.
4. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. Ann Thorac Surg 2004; 78:1868–77.
5. Shroyer AL, Coombs LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. Ann Thorac Surg 2003;75:1856–64.
6. Edwards FH, Peterson ED, Coombs LP, et al. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol 2001;37:885–92.
7. Jamieson WR, Edwards FH, Schwartz M, Bero JW, Clark RE, Grover FL. Risk stratification for cardiac valve replacement. National Cardiac Surgery Database. Database Committee of The Society of Thoracic Surgeons. Ann Thorac Surg 1999;67:943–51.
8. Hannan EL, Wu C, Bennett EV, et al. Risk index for predicting in-hospital mortality for cardiac valve surgery. Ann Thorac Surg 2007;83:921–9.
9. Hannan EL, Racz MJ, Jones RH, et al. Predictors of mortality for patients undergoing cardiac valve replacements in New York State. Ann Thorac Surg 2000;70:1212–8.
10. Nowicki ER, Birkmeyer NJ, Weintraub RW, et al. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg 2004;77:1966–77.
11. Nowicki ER. What is the future of mortality prediction models in heart valve surgery? Ann Thorac Surg 2005;80:396–8.
12. Jin R, Grunkemeier GL, Starr A. Validation and refinement of mortality risk models for heart valve surgery. Ann Thorac Surg 2005;80:471–9.
13. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. Circulation 2005;112:224–31.
14. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
15. Gardner SC, Grunwald GK, Rumsfeld JS, et al. Comparison of short-term mortality risk factors for valve replacement versus coronary artery bypass graft surgery. Ann Thorac Surg 2004;77:549–56.
16. Grover FL, Edwards FH. Similarity between the STS and New York State databases for valvular heart disease. Ann Thorac Surg 2000;70:1143–4.
17. Rankin JS, Hammill BG, Ferguson TB Jr, et al. Determinants of operative mortality in valvular heart surgery. J Thorac Cardiovasc Surg 2006;131:547–57.
18. Roques F, Nashef SA, Michel P. Risk factors for early mortality after valve surgery in Europe in the 1990s: lessons from the EuroSCORE pilot program. J Heart Valve Dis 2001;10:572–7.
19. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85:921–30.
20. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
21. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken, NJ: Wiley-Interscience, 2002.
22. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.
23. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Plus CABG Models

For each endpoint, the formula for calculating a patient's predicted risk of the endpoint has the form:

$$\text{Predicted Risk} = \frac{e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}{1+e^{(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}}$$

where $x_1, x_2, \ldots, x_n$ denote patient preoperative risk factors (eg, quantitative variables such as age, and comorbidities coded as 1=present, 0=absent); and $\beta_0, \beta_1, \ldots, \beta_n$ denote regression coefficients (numerical constants). Regression coefficients for each endpoint are presented in Appendix Table 1. The variables $x_1, x_2, \ldots, x_n$ are the same for each endpoint and are defined in Appendix Table 2. The regression coefficient for the time trend is not presented. Instead, the intercept has been adjusted to incorporate the time trend. This adjusted intercept reflects the baseline risk for a reference period of July–December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−5.24391	−5.14546	−5.32535	−3.63438	−6.50043	−3.16980	−2.99714	−4.15892	1.18582
Preoperative AFib	0.18430	0.04634	0.16567	0.12059	0.00000	0.14910	0.13766	0.19656	−0.34095
Age function 1	0.02560	0.02487	0.03268	0.02106	0.00545	0.01715	0.01838	0.03115	−0.02970
Age function 3	0.02758	−0.00709	0.00671	0.00791	−0.00985	−0.00021	0.01425	0.00985	−0.04542
Age by reop function	−0.00861	0.00458	0.00077	−0.00673	0.00314	−0.00399	−0.00202	−0.00678	0.00656
Age by status function	−0.00507	−0.01979	−0.00178	−0.00750	0.01627	−0.00029	0.00229	−0.02247	0.00692
Age by MVR function	0.01564	0.00000	0.00000	0.00000	0.00000	0.00000	0.00527	0.00000	−0.00866
Age by MVRepair function	0.01240	0.00000	0.00000	0.00000	0.00000	0.00000	0.00599	0.00000	−0.01159
BSA function 1	−1.14176	−0.81169	−0.41848	−0.66843	0.86401	−0.51266	−0.70411	−0.84204	0.51295
BSA function 2	2.25471	0.94689	1.84088	1.80467	0.42453	0.70024	1.70623	2.10402	−1.66758
CHF but not NYHA IV	0.21206	−0.01726	0.17460	0.20063	0.00000	0.00000	0.12880	0.26291	−0.17652
CHF and NYHA IV	0.39457	0.14109	0.30146	0.38383	0.00000	0.14499	0.30567	0.39791	−0.31077
CHF by MVR function	−0.31077	−0.20917	−0.25767	−0.18455	0.00000	0.00000	−0.18635	−0.23729	0.00000
CHF by MVRepair function	−0.24791	0.06897	−0.18667	−0.10484	0.00000	0.00000	−0.06920	−0.10954	0.00000
CLD function	0.17713	0.00000	0.11379	0.23345	0.27571	0.09280	0.16523	0.22999	−0.19234
CLD by MVR function	0.00000	0.00000	0.00000	−0.06780	0.00000	0.00000	0.00000	−0.04591	0.00000
CLD by MVRepair function	0.00000	0.00000	0.00000	−0.04014	0.00000	0.00000	0.00000	−0.08501	0.00000
Creatinine function 1	0.44794	0.23545	0.81612	0.38147	0.00000	0.24620	0.51256	0.41472	−0.47658
Creatinine by MVR function	0.00000	0.00000	−0.21574	0.00000	0.00000	0.00000	0.00000	0.00000	0.06652
Creatinine by MVRepair function	0.00000	0.00000	−0.18787	0.00000	0.00000	0.00000	0.00000	0.00000	−0.04407
CVD without prior CVA	0.00000	0.24847	0.13299	0.09769	0.00000	0.00000	0.10255	0.10601	−0.16643
CVD and prior CVA	0.19754	0.54344	0.11571	0.23581	0.19686	0.10974	0.23332	0.23319	−0.28560
Diabetes, noninsulin	0.11060	0.14576	0.24490	0.10365	0.26281	0.00000	0.11462	0.15846	−0.17020
Diabetes, insulin	0.26870	0.14582	0.48504	0.27893	0.68330	0.00000	0.29508	0.39583	−0.40448
Dialysis	1.61151	0.58833	0.00000	1.20290	0.61527	0.74332	1.25181	1.29747	−1.67728
Dialysis by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	−0.30339	0.00000	0.00000	0.04745
Dialysis by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.13058	0.00000	0.00000	0.09778
Ejection fraction function	0.00989	0.00000	0.00534	0.01113	0.00000	0.00703	0.01061	0.00995	−0.01440
EF by MVR function	0.01056	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00228
EF by MVRepair function	−0.00117	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.00309
Endocarditis, active	0.71327	0.60657	0.41797	0.67172	0.00000	0.44757	0.74858	0.59333	−1.27854
Endocarditis by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.13795	0.00000
Endocarditis by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.49934	0.00000
Female	0.30852	0.17170	0.16668	0.41874	0.10654	−0.08221	0.18594	0.26947	−0.50044
Female by MVR function	0.00000	0.00000	0.00000	−0.25972	0.00000	0.00000	0.00000	0.00000	0.08895
Female by MVRepair function	0.00000	0.00000	0.00000	−0.19373	0.00000	0.00000	0.00000	0.00000	−0.00229
Female by BSA function 1	0.51233	0.07575	0.76032	0.48032	0.80594	0.16701	0.41581	0.91055	−0.59086
Female by BSA function 2	−0.27980	−0.88628	−0.57622	−0.49740	0.58767	0.52524	−0.40427	−0.78096	0.15748
Hypertension	0.00000	0.17080	0.22638	0.09581	0.28851	0.00000	0.11445	0.07602	−0.08668

Appendix Table 1. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
IABP or inotropes	0.36025	0.00000	0.23674	0.77918	0.00000	0.15075	0.56477	0.34008	−0.58536
IABP by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.08732	−0.09462
IABP by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10281	−0.06743
Immunosuppressive treatment	0.29654	0.00000	0.26400	0.24814	0.00000	0.24041	0.23332	0.19750	−0.28819
Insufficiency, mitral	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.06414	0.00000	0.00000
Insufficiency, tricuspid	0.24006	0.00000	0.22040	0.13606	0.00000	0.00000	0.13318	0.00000	−0.23141
Left main disease	0.11450	0.00000	0.00000	0.06181	0.00000	0.00000	0.00000	0.00000	0.00000
MI 1–21 days	0.17038	0.00000	0.16476	0.24560	0.00000	0.00000	0.19751	0.00000	0.00000
MI ≤ 21 days	0.00000	0.19671	0.00000	0.00000	0.00000	0.00000	0.00000	0.14599	0.00000
MI < 24 hours	0.49918	0.00000	0.26240	0.34321	0.00000	0.13716	0.39731	0.00000	0.00000
MVR	0.14888	0.32659	0.90926	0.76504	0.28437	0.41642	0.41322	0.73530	−0.82339
MVRepair	−0.07374	0.06933	0.51275	0.28204	0.19499	0.07390	−0.03949	0.30384	−0.03552
No. diseased coronary vessel function	0.13746	0.18243	0.15791	0.17277	0.24582	0.08187	0.14767	0.12474	−0.19250
Peripheral vascular disease	0.25173	0.13776	0.14995	0.16591	0.00000	0.14312	0.18062	0.14863	0.00000
Race black	0.00000	0.00000	0.14301	0.26900	0.00000	0.17364	0.19182	0.26856	−0.43385
Race Hispanic	0.00000	0.00000	0.18384	0.15363	0.00000	0.08065	0.13561	0.12286	−0.15901
Reop, 1 previous operation	0.78624	0.00000	0.25782	0.60179	0.00000	0.33209	0.40293	0.43757	−0.39723
Reop, ≥ 2 previous operations	0.90015	0.00000	0.38499	0.78263	0.00000	0.39502	0.56875	0.50334	−0.63237
Reop by MVR function	0.00000	0.00000	0.00000	−0.27846	0.00000	−0.19608	0.00000	−0.17836	0.18262
Reop by MVRepair function	0.00000	0.00000	0.00000	−0.16306	0.00000	0.06985	0.00000	−0.16007	0.17613
Shock	0.51917	0.17321	0.15810	0.65653	0.00000	0.21271	0.58409	0.36987	0.00000
Shock by MVR function	0.00000	0.00000	0.02883	0.00000	0.00000	0.00000	0.43045	0.00000	0.00000
Shock by MVRepair function	0.00000	0.00000	0.36429	0.00000	0.00000	0.00000	0.19084	0.00000	0.00000
Status urgent	0.22591	0.00000	0.16451	0.22905	0.00000	0.12800	0.17511	0.24758	−0.26626
Status emergent	0.75852	0.79460	0.56854	0.99818	0.00000	0.34063	0.77631	1.00162	−1.09633
Status salvage	1.51811	0.95665	0.61798	0.75178	0.00000	0.00000	1.20732	0.56482	−1.72252
Status by MVR function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.25083
Status by MVRepair function	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.25943
Stenosis, mitral	0.09879	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.08469	0.00000
Unstable angina	0.10722	−0.11292	0.11597	0.05762	0.00000	0.00000	0.00000	0.00000	0.00000

Afib = atrial fibrillation; BSA = body surface area; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Appendix Table 2. Definition of Variables Appearing in STS 2008 Valve Plus CABG Models

Variable	Definition
Intercept	= 1 for all patients
Preoperative AFib	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 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
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 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
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine by MVR function	= Creatinine function 1 if valve operation is MVR, = 0 otherwise
Creatinine by MVRepair function	= Creatinine function 1 if valve operation is MVRepair, = 0 otherwise
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
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
EF by MVR function	= Ejection fraction function if valve operation is MVR, = 0 otherwise
EF by MVRepair function	= Ejection fraction function if valve operation is MVRepair, = 0 otherwise
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Endocarditis by MVR function	= 1 if patient has active endocarditis and valve operation is MVR, = 0 otherwise
Endocarditis by MVRepair function	= 1 if patient has active endocarditis and valve operation is MVRepair, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 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
IABP by MVR function	= 1 if patient requires preop IABP or inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP or inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient has received immunosuppressive therapy within 30 days, = 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 1–21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI ≤ 21 days[a]	= 1 if patient has history of MI within 21 days prior to surgery, = 0 otherwise (for CVA and PLOS; coded as < 24 hours and 1–21 days for others)
MI < 24 hours	= 1 if history of MI < 24 hours prior to surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased coronary vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 0 otherwise

Appendix Table 2. Continued

Variable	Definition
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
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
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Shock by MVR function	= 1 if shock and operation is MVR, = 0 otherwise
Shock by MVRepair function	= 1 if shock and operation is MVRepair, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not rescuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Status by MVR function	= 1 if status is emergent or salvage and operation is MVR, = 0 otherwise
Status by MVRepair function	= 1 if status is emergent or salvage and operation is MVRepair, = 0 otherwise
Stenosis, mitral	= 1 if patient has mitral stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina and no MI within 7 days of surgery, = 0 otherwise

[a] MI coded ≤ 21 days for CVA and PLOS endpoints; for all other endpoints, coded as < 24 hours and 1 to 21 days.

Note: See www.sts.org for exact definitions of terms used above.

BSA = body surface area; CABG = coronary artery bypass graft surgery; CHF = congestive heart failure; CLD = chronic lung disease; Comp = composite adverse event (any); CVA = cerebrovascular accident (stroke); CVD = cerebrovascular disease; DSWI = deep sternal wound infection; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; MVR = mitral valve replacement; MVRepair = mitral valve repair; NYHA = New York Heart Association; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.