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 #:** 0129 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Risk-Adjusted Prolonged Intubation (Ventilation)
De.2 **Brief description of measure:** Percent of patients aged 18 years and older undergoing isolated CABG who require intubation for more than 24 hours
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** OT1-013-09 – The STS CABG Composite Score
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-634267338401654454.pdf	**A** Y☐ N☐

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

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. ***Measures must be judged to be important to measure and report*** *in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	**Eval Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Affects large numbers, 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:** Prolonged ventilation has been shown to substantially increase post-CABG length of stay, the costs of care, and is associated with higher rates of respiratory failure, stroke, renal failure, and death. 1a.4 **Citations for Evidence of High Impact:** - Bardell T, Legare JF, Buth KJ, et al. ICU readmission after cardiac surgery. Eur J Cardiothorac Surg. 2003;23(3):354-359. - Meade MO, Guyatt G, Butler R, et al. Trials comparing early vs late extubation following cardiovascular surgery. Chest. 2001:120(6 Suppl):445S-453S. - Naughton C, Reilly N, Powroznyk A, et al. Factors determining the duration of tracheal intubation in cardiac surgery: a single-centre sequential patient audit. Eur J Anaesthesiol. 2003;20(3):225-233. - 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;75:1856-1865. - Welke KF, Ferguson TB, Coombs LP, et al. Validity of the Society of Thoracic Surgeons National Adult Cardiac Surgery Database. Ann Thorac Surg. 2004;77:1137-1139. - Engel AM, McDonough S, Smith JM. Does an obese body mass index affect hospital outcomes after coronary artery bypass graft surgery? Ann Thorac Surg. 2009 Dec;88(6):1793-800.	**1a** C☐ P☐ M☐ N☐

- Speir AM, Kasirajan V, Barnett SD, Fonner E Jr. Additive costs of postoperative complications for isolated coronary artery bypass grafting patients in Virginia. Ann Thorac Surg. 2009 Jul;88(1):40-5; discussion 45-6. PubMed PMID:19559186.
- Brown PP, Kugelmass AD, Cohen DJ, Reynolds MR, Culler SD, Dee AD, Simon AW. The frequency and cost of complications associated with coronary artery bypass grafting surgery: results from the United States Medicare program. Ann Thorac Surg. 2008 Jun;85(6):1980-6. PubMed PMID: 18498806.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: Modalities to decrease the rate of prolonged intubation include physician supervised protocols for extubation implemented by nurses and respiratory therapists, improved pre-operative preparation of patients, reduction of post-operative bleeding, and intra-operative protocolized anesthesia care. Current implementation is highly variable and great opportunities to increase the implementation of evidence based care exist. Cardiac surgery programs with high implementation have lower than average rates of prolonged ventilation and significantly lower rates of adverse events.

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

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

Analysis includes 640 STS Adult Cardiac Surgery Database Participants who had at least 100 eligible cases for the measure and reported data (not restricted to this measure) to the STS Adult Cardiac Surgery Database for all 12 months.

1b.4 Summary of Data on disparities by population group:
Please see attachment

1b.5 Citations for data on Disparities:
Analysis includes STS Adult Cardiac Surgery Database Participants that had more than 50 eligible cases in 2008 and 2009, and reported data for at least 15 months

238263 Patients from 897 Participants were included in the Gender = Male sub-group.
79685 Patients from 653 Participants were included in the Gender = Female sub-group.
13328 Patients from 135 Participants were included in the Race = Black sub-group.
280505 Patients from 887 Participants were included in the Race = White sub-group.
12932 Patients from 121 Participants were included in the Race = Other sub-group.
9710 Patients from 94 Participants were included in the Ethnicity = Hispanic sub-group.
309709 Patients from 902 Participants were included in the Ethnicity = Non-Hispanic sub-group.

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*)**:** Prolonged ventilation has been shown to substantially increase post-CABG length of stay, the costs of care, and is associated with higher rates of respiratory failure, stroke, renal failure, and death.

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*)**:**
Prolonged ventilation has been shown to substantially increase post CABG length of stay, the costs of care, and is associated with higher rates of respiratory failure, stroke, renal failure, and death.

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

1c C☐ P☐ M☐ N☐

whom): 1c.6 **Method for rating evidence:** 1c.7 **Summary of Controversy/Contradictory Evidence:** 1c.8 **Citations for Evidence** (*other than guidelines*): - Bardell T, Legare JF, Buth KJ, et al. ICU readmission after cardiac surgery. Eur J Cardiothorac Surg. 2003;23(3):354-359. - Meade MO, Guyatt G, Butler R, et al. Trials comparing early vs late extubation following cardiovascular surgery. Chest. 2001:120(6 Suppl):445S-453S. - Naughton C, Reilly N, Powroznyk A, et al. Factors determining the duration of tracheal intubation in cardiac surgery: a single-centre sequential patient audit. Eur J Anaesthesiol. 2003;20(3):225-233. - 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;75:1856-1865. - Welke KF, Ferguson TB, Coombs LP, et al. Validity of the Society of Thoracic Surgeons National Adult Cardiac Surgery Database. Ann Thorac Surg. 2004;77:1137-1139. - Engel AM, McDonough S, Smith JM. Does an obese body mass index affect hospital outcomes after coronary artery bypass graft surgery? Ann Thorac Surg. 2009 Dec;88(6):1793-800. - Speir AM, Kasirajan V, Barnett SD, Fonner E Jr. Additive costs of postoperative complications for isolated coronary artery bypass grafting patients in Virginia. Ann Thorac Surg. 2009 Jul;88(1):40-5; discussion 45-6. PubMed PMID:19559186. - Brown PP, Kugelmass AD, Cohen DJ, Reynolds MR, Culler SD, Dee AD, Simon AW. The frequency and cost of complications associated with coronary artery bypass grafting surgery: results from the United States Medicare program. Ann Thorac Surg. 2008 Jun;85(6):1980-6. PubMed PMID: 18498806. 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 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 isolated CABG who require intubation > 24 hours

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*)**:**

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*)**:**
Number of isolated CABG procedures in which Complications-Pulmonary_Vent Prolonged (CPVntLng) is marked "yes"

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

Isolated CABG is determined as a procedure for which all of the following apply:
- OpCAB 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”

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*)**:**

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-634267353926995758.pdf

2a-specs
C☐
P☐
M☐
N☐

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-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 : Counties or cities, Population : National, Population : Regional/network, Population : states

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

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

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)*: STS Adult Cardiac Surgery Database - Compared results between two proximate time periods: January 2008-December 2008 and January 2009-December 2009.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
Compared results between two proximate time periods: January 2008-December 2008 and January 2009-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
C☐

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)*:
Re-intubated During Hospital: 99.8% agreement rate
Additional # of Hours Ventilated: 97.5% agreement rate

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.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 1--coronary artery bypass grafting surgery. Ann Thorac Surg. 2009 Jul;88(1 Suppl):S2-22.

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)*: 640 STS Adult

2f C☐ P☐

Cardiac Surgery Database Participants who had at least 100 eligible cases for the measure and reported data to STS for all 12 months; January 1, 2009-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	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)*: 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*): This measure is one of eleven component measures of the STS CABG Composite Score. Composite star ratings are presented on the STS website, www.sts.org/publicreporting and in the health section of the Consumers Union website, www.ConsumerReportsHealth.org.	**3a** C☐ P☐ M☐ N☐

There are approximately 330 STS Adult Cardiac Surgery Database Participants who voluntarily participate in the Consumer's Union public reporting initiative. In addition, approximately 352 STS Adult Cardiac Surgery Database Participants voluntarily take part in STS Public Reporting Online.

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***):**
CMS Physician Quality Reporting Initiative (PQRI), www.cms.hhs.gov/pqri

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:**
OT1-013-09 – The STS CABG Composite Score; Component measures: 0114 Risk-Adjusted Post-Operative Renal Failure, 0115 Risk-Adjusted Surgical Re-exploration, 0116 Anti-Platelet Medication at Discharge, 0117 Beta Blockade at Discharge, 0118 Anti-Lipid Treatment at Discharge, 0119 Risk-Adjusted Operative Mortality for CABG, 0127 Pre-Operative Beta Blockade, 0129 Risk-Adjusted Prolonged Intubation (ventilation), 0130 Risk-Adjusted Deep Sternal Wound Infection Rate, 0131 Risk-Adjusted Stroke/Cerebrovascular Accident, 0134 Use of Internal Mammary Artery (IMA) in Coronary Artery Bypass Graft (CABG)

(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:** n/a **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/a	**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:	**4d** C☐ P☐ M☐ N☐

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.	
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:
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 0129 Sections 2a.14, 1b.2, 1b.4, 2b.3, 2f.3, 3a.6.pdf

Date of Submission (*MM/DD/YY*)**:** 03/28/2011

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

Variable Definitions and Selection

All variables for consideration are listed in the table below.

Definition of Variables Appearing in STS 2008 CABG Models

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

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

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

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

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

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

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

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

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

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

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

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

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

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

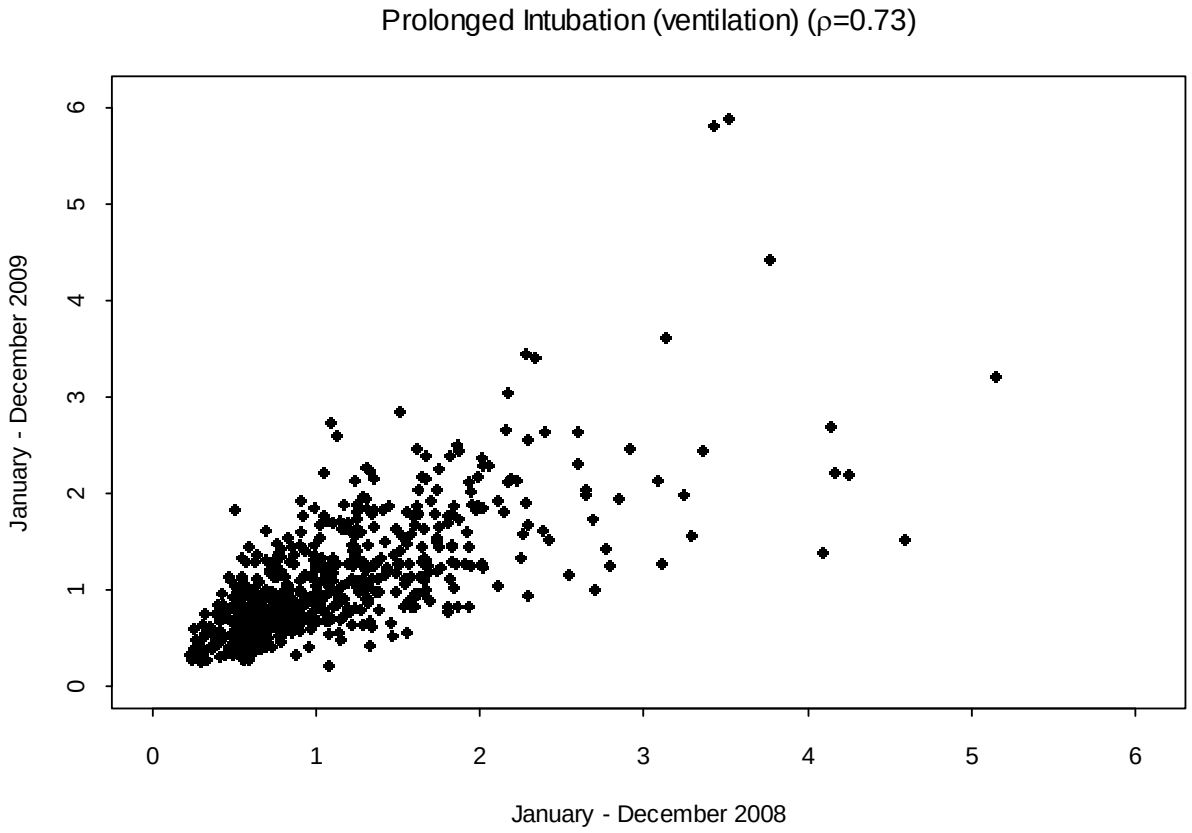
	Prolonged Intubation (ventilation) - Risk Adjusted Rate		
	Population Group		
	Black	*White*	*Other*
Measurement			
N	135	887	121
Mean	15.8	11.3	13.5
1st	6.7	3.6	5.4
5th	8.6	4.9	5.9
10th	9.2	5.9	7.2
25th	12.0	7.7	9.3
Median	14.8	10.3	13.2
75th	19.0	14.2	16.1
90th	23.3	18.0	19.5
95th	24.7	20.3	23.2
99th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

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

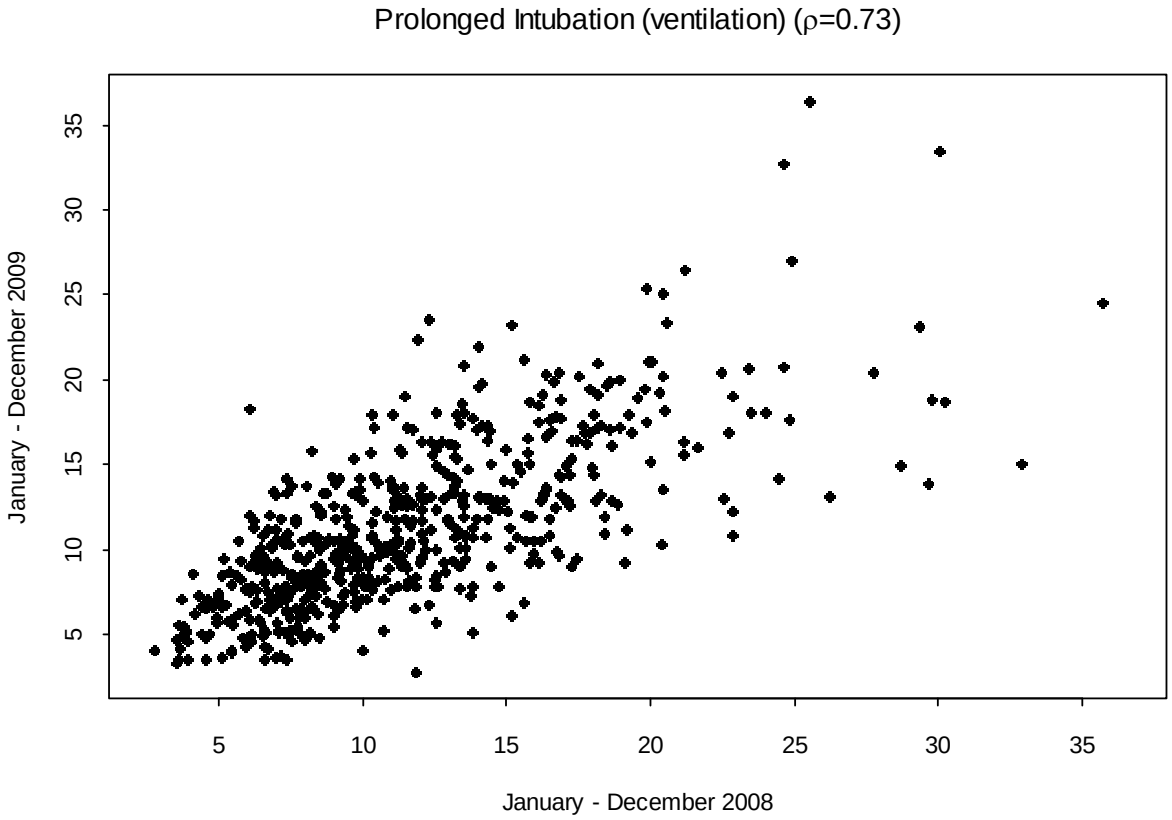
	Prolonged Intubation (ventilation) - Risk Adjusted Rate	
	Population Group	
	Hispanic	*Non-Hispanic*
Measurement		
N	94	902
Mean	14.4	11.8
1st	4.1	3.9
5th	5.6	5.2
10th	6.2	6.0
25th	8.9	8.0
Median	13.6	10.9
75th	18.5	14.7
90th	23.0	18.4
95th	26.3	20.8
99th	36.2	27.2
Outlier	38 (40.4%)	431 (47.8%)
High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

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

Testing results: $\rho = 0.73$



Risk Adjusted Rate



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

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

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

Risk Adjusted Rate:

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

Measurement	*Prolonged Intubation (ventilation)*
Low	150

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

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

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

1) Report Overview and Interpretation Manual:

The NQF Measures Report

a. Organization

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

b. Statistical Calculation and Details – NQF Measures

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

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

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

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

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

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

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

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

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

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

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

c. Technical Notes

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

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

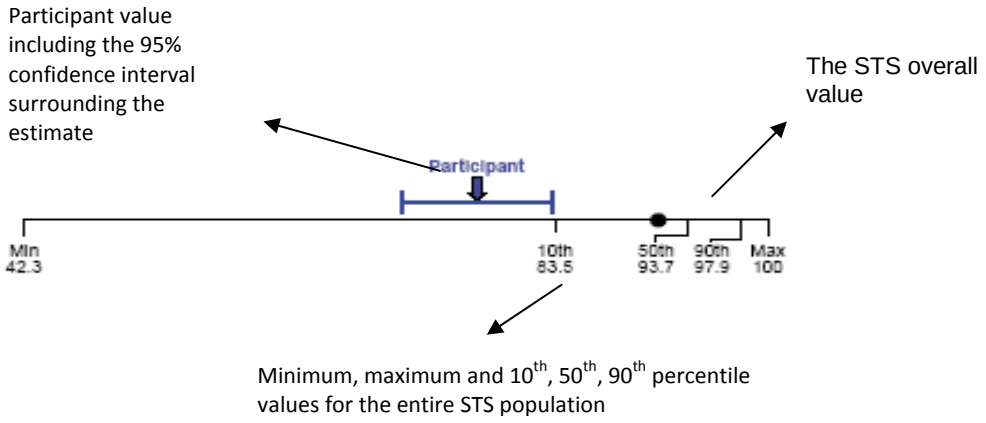
NQF/STS Results Comparison: Participants may see some differences between summaries of their data provided in the NQF section of the report and summaries of their data reported elsewhere in the STS report. These differences are due to subtle variations in variable definitions, patient inclusion and exclusion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

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

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

Interpretation Manual

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



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

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

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

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

a. NQF Measures Interpretation Example

Sample CABG Operative Mortality results – tabular and figure representation.

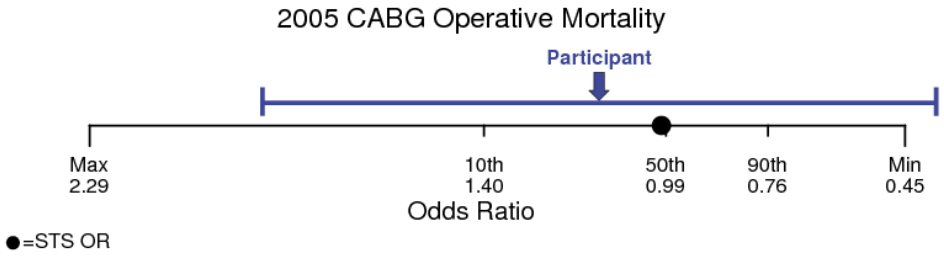
NQF Measure	Eligible Procedures	Participant Estimated OR	Participant Percentile	Participant Observed Rate
2005 CABG Operative Mortality	74	1.14	26.3	5.4%

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

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

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

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



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

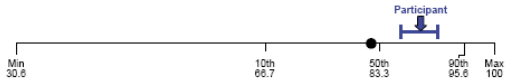
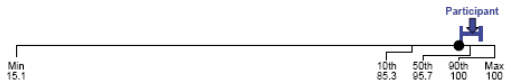
The 95% confidence interval for the participant's OR spans from <0.45 to ~1.90

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

NQF Measures
Process Measures

Participant 99999
STS Period Ending 12/31/2008

Duke Clinical Research Institute
DUKE UNIVERSITY MEDICAL CENTER

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

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

The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 1—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. **The first version of The Society of Thoracic Surgeons National Adult Cardiac Surgery Database (STS NCD) was developed nearly 2 decades ago. Since its inception, the number of participants has grown dramatically, patient acuity has increased, and overall outcomes have consistently improved. To adjust for these and other changes, all STS risk models have undergone periodic revisions. This report provides a detailed description of the 2008 STS risk model for coronary artery bypass grafting surgery (CABG).**

Methods. **The study population consisted of 774,881 isolated CABG procedures performed on adult patients aged 20 to 100 years between January 1, 2002, and December 31, 2006, at 819 STS NCD participating centers. This cohort 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. The validation sample was used to assess model calibration and discrimination. Model outcomes included operative mortality, renal failure, stroke, reoperation for any cause, prolonged ventilation, deep sternal wound infection, composite major morbidity or mortality, prolonged length of stay (> 14 days), and short length of stay (< 6 days and alive). Candidate predictor variables were selected based on their availability in versions 2.35, 2.41, and 2.52.1 of the STS NCD and their presence in (or ability to be mapped to) version 2.61. Potential predictor variables were screened for overall prevalence in the study population, missing data frequency, coding concerns, bivariate relationships with outcomes, and their presence in previous STS or other CABG risk models. Supervised backwards selection was then performed with input from an expert panel of cardiac surgeons and biostatisticians. After successfully validating the fit of the models, the development and validation samples were subsequently combined, and the final regression coefficients were estimated using the overall combined (development plus validation) sample.**

Results. **The c-index for the mortality model was 0.812, and the c-indices for other endpoints ranged from 0.653 for reoperation to 0.793 for renal failure in the validation sample. Plots of observed versus predicted event rates revealed acceptable calibration in the overall population and in numerous subgroups. When patients were grouped into categories of predicted risk, the absolute difference between the observed and expected event rates was less than 1.5% for each endpoint. The final model intercept and coefficients are provided.**

Conclusions. **New STS risk models have been developed for CABG mortality and eight other endpoints. Detailed descriptions of model development and testing are provided, together with the final algorithm. Overall model performance is excellent.**

(Ann Thorac Surg 2009;88:S2–22)

In 1986, The Society of Thoracic Surgeons (STS) convened an Ad Hoc Committee on Risk Factors for Coronary Artery Bypass Graft Surgery (CABG) [1] and an Ad Hoc Committee to Develop a National Database for Cardiothoracic Surgery [2]. This was prompted by the

†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 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.

Published by Elsevier Inc

0003-4975/09/$36.00
doi:10.1016/j.athoracsur.2009.05.053

Abbreviations and Acronyms

BSA = body surface area
CABG = coronary artery bypass graft surgery
CHF = congestive heart failure
EF = ejection fraction
GFR = glomerular filtration rate
HCFA = Health Care Financing Administration
IABP = intra-aortic balloon pump
NYHA = New York Heart Association
NCD = National Adult Cardiac Surgery Database
O/E = observed to expected ratio
QMTF = Quality Measurement Task Force
STS = The Society of Thoracic Surgeons

release earlier that year of inadequately risk-adjusted hospital mortality data by the Health Care Financing Administration (HCFA), now the Centers for Medicare and Medicaid Services. Although the HCFA analytical methodology was widely criticized, STS leadership recognized that the underlying principle of collecting and analyzing data to improve patient outcomes was valid, particularly for complex and costly procedures such as coronary artery bypass grafting surgery (CABG). They believed that it was the responsibility of professional organizations to develop credible clinical data registries for their own specialties, and that risk models derived from such registries would circumvent many of the concerns resulting from the use of unadjusted administrative data. Such clinical registries would be used as credible data sources for quality assessment and improvement activities as well as for research.

These early activities ultimately led to the development of the STS National Adult Cardiac Surgery Database (NCD) [3, 4]. Since its release to members in 1990, the STS NCD has evolved to become one of the largest specialty-specific clinical data registries in the world. It currently has more than 950 participants enrolled, representing just under 90% of the cardiac surgery providers in the United States, with data on more than 3.6 million procedures. Similar STS data registries have now been developed for congenital heart surgery and general thoracic surgery, and future plans include the development of specialty modules (eg, quality metrics, atrial fibrillation surgery, thoracic aortic surgery). Recent enhancements, including the addition of unique physician and patient identifiers, will facilitate linkages with other registries and greatly expand the potential of the STS NCD for longitudinal follow-up, comparative effectiveness, and cost efficiency studies.

In addition to the development of the STS NCD as a comprehensive, nationally representative data registry, the second major goal of the STS was to assure that analyses derived from this registry would be appropriately adjusted for preoperative patient severity, a major deficiency of the HCFA reports that were initially published in 1986. This was accomplished by first identifying risk factors for specific procedures and outcomes, beginning with isolated CABG, then using these predictor variables to develop risk models. With statistical risk models, which are most often based on logistic regression, the expected outcome for a patient with a given set of risk factors can be determined, and that can be compared with the observed outcome. The observed (O) and expected (E) outcomes are summed over all patients of a particular surgeon or hospital to yield the risk-standardized mortality ratio (O/E), which can then be multiplied by the average rate in the reference population to calculate risk-standardized mortality rates [5–7].

STS CABG risk models have undergone periodic updates and revisions, the most recent of which was based upon 2000 to 2002 STS NCD data. In 2007, the STS Database Modernization Task Force completed a major specification upgrade of the STS NCD data collection instrument from version 2.52.1 to version 2.61. This included refinement, modification, consolidation, or elimination of some data elements, as well as an attempt to harmonize definitions with those of the American College of Cardiology National Cardiovascular Data Registry whenever possible. Given these changes, as well as the number of years since the last risk model update, the STS Quality Measurement Task Force (QMTF) was asked to develop new risk models for isolated CABG, isolated valve repair or replacement, and combined CABG plus valve procedures. The authors of this report include the QMTF members who participated in this initiative.

Implementation of these new models in January 2008 coincided with the release of STS NCD version 2.61. This report, Part 1 of 3, describes the development of the new mortality and morbidity models for isolated CABG surgery.

Study Purpose

The primary goal of this study was to develop risk-prediction algorithms for patients undergoing isolated CABG surgery. As the major intended use of these algorithms was to compare participant outcomes to the overall STS national experience, risk factors were generally restricted to patient and clinical characteristics present preoperatively.

Risk Model Development and Transparency

The availability of user-friendly statistical software programs and the exponential increase in computing speed have greatly facilitated statistical analyses such as logistic regression, the basis for many risk models. However, despite these technological advances, clinical judgment, experience, intuition, and practicality still play a critical role in risk model development. There are many points in model development at which legitimate differences in approach may lead to substantial differences in the resulting statistical models and the inferences derived from them [8].

We believe the degree of transparency provided in this report regarding the development of the STS CABG risk

models is essential in today's health care environment. In an era when society demands full transparency regarding health care performance, the methodologies used to evaluate that performance should be just as transparent [9, 10]. This fundamental principle is among the standards established by the American Heart Association and American College of Cardiology for statistical models used for public reporting [11].

Study Population and Endpoints

All isolated CABG procedures performed on adult patients aged 20 to 100 years between January 1, 2000, and December 31, 2006, were initially considered for inclusion, although the final development and validation samples were derived from 2002 to 2006 data. Patients missing data on sex (n = 195) were excluded, as these patients are not included in STS performance feedback reports to database participants. That left a study population of 774,881 surgical procedures from 819 database participants. Patients on dialysis preoperatively (n = 12,415) were excluded when developing the risk model for postoperative renal failure.

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 the nine endpoints listed below. Only mortality was recorded beyond the index hospitalization. Morbidity data included only in-hospital complications, although beginning in STS NCD version 2.61, sternal infections will be recorded for up to 30 days postoperatively. The nine endpoints are as follows: (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 (cerebrovascular accident): 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 (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 (this SLOS definition differs from the previous STS risk models, which did not exclude patients who died in-hospital; patients who died within 5 days of surgery are included in the new models but are treated as not having a short stay).

Table 1 summarizes the frequencies of these endpoints in the study population for each predictor variable category (ie, the bivariate relationships).

Selection of Candidate Predictor Variables

Initial Data Screening of Candidate Predictor Variables

We began by considering all possible candidate variables from the development set (Table 2). Because the primary goal of the STS risk models is to adjust surgical outcomes, in general only preoperative patient variables are included. However, because these models are also used for other purposes such as individual patient prediction and counseling, there were a few modifications (which are discussed in the relevant sections) in the application of this general principle.

As there were a large number of procedures and endpoints available, we were not statistically constrained to highly parsimonious models, nor is such an approach generally favored in regression modeling [12–14]. Discarding valid data elements can waste valuable information that has been collected at substantial effort and cost. Furthermore, although much of the discrimination of a predictive model may be contained in a relatively small number of variables [15, 16], some predictor variables that add only modestly to discrimination may still be important predictors of outcomes at the patient level [17, 18].

Expert Panel Review for Clinical Relevance and Face Validity

All candidate variables available in version 2.61 were individually discussed by a panel of cardiac surgeons and health policy experts to assure that clinical relevance as well as multiple aspects of validity (face, construct, and content) had been considered.

Data Version for Model Development

Although these new risk models were to be introduced in conjunction with the release of STS NCD version 2.61, they were developed with data collected under the three previous data versions (2.35, 2.41, and 2.52.1) because no 2.61 data were yet available. The QMTF began its predictor selection process with two caveats. First, any candidate variable had to be collected consistently across the three previous data versions. Second, it had to also be available in version 2.61 or have the ability to be mapped to this new version. For example, history of smoking and renal failure were not candidate variables as they were either not included in, or were unable to be mapped to, version 2.61. Renal function is now assessed by the last preoperative serum creatinine value, which is collected in all data versions. Because the definition of hypercholesterolemia has changed substantially over successive STS data versions, and because counterintuitive results have been observed in some previous analyses of hyper-

Table 1. Distribution of Risk Factors and Frequency of Adverse Outcomes in Overall Study Population, Isolated Coronary Artery Bypass Graft Surgery (2002–2006)

	Number of Patients		Percent of Patients Experiencing Endpoint								
Variable	N	%	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Overall											
Total	774,881	100.0	2.3	1.4	3.6	9.7	0.4	5.2	14.4	5.6	51.2
Age, years											
< 55	137,318	17.72	1.0	0.5	1.7	7.1	0.3	3.7	10.0	2.7	67.1
55–64	221,697	28.61	1.3	0.9	2.4	7.8	0.4	4.2	11.4	3.8	59.4
65–74	245,132	31.63	2.4	1.6	3.9	10.0	0.5	5.5	14.9	5.9	47.7
≥ 75	170,734	22.03	4.7	2.3	6.4	13.9	0.5	7.5	20.9	9.6	33.0
Sex											
Male	560,006	72.27	2.0	1.2	3.4	8.7	0.4	5.1	13.4	4.9	55.0
Female	214,875	27.73	3.4	1.9	4.1	12.2	0.5	5.6	17.0	7.2	41.5
Race											
Caucasian	665,941	85.94	2.3	1.3	3.5	9.3	0.4	5.1	13.9	5.3	52.2
Black	44,405	5.73	2.7	2.0	5.2	13.5	0.7	6.3	19.0	8.2	41.3
Hispanic	25,103	3.24	2.6	1.5	4.3	11.3	0.5	5.6	16.1	6.1	48.4
Asian	12,509	1.61	2.7	1.9	3.8	12.6	0.3	7.2	17.4	7.0	45.2
Other	21,222	2.74	2.3	1.3	3.6	10.4	0.5	5.5	14.8	6.0	48.7
Missing	5,701	0.74	2.3	1.4	4.1	9.4	0.4	5.2	14.5	6.1	48.9
Body surface area (m^2)											
< 1.50	14,339	1.85	6.2	2.4	4.6	16.2	0.3	8.3	22.1	9.8	36.5
1.50–1.74	111,458	14.38	3.8	2.0	4.0	12.6	0.3	6.5	17.7	7.4	42.5
1.75–1.99	280,677	36.22	2.4	1.5	3.5	9.6	0.4	5.4	14.4	5.6	50.7
≥ 2.00	363,817	46.95	1.7	1.0	3.6	8.6	0.5	4.6	13.1	4.8	55.0
Missing	4,590	0.59	3.7	1.4	4.0	7.6	0.3	4.7	13.9	6.8	46.0
Body mass index (kg/m^2)											
< 25	169,091	21.82	3.3	1.7	3.5	11.0	0.3	6.7	16.3	6.7	47.6
25–29	303,371	39.15	2.1	1.3	3.1	8.6	0.3	4.9	13.1	4.8	54.2
30–34	186,148	24.02	1.8	1.2	3.6	9.0	0.5	4.5	13.4	5.0	53.1
≥ 35	110,213	14.22	2.3	1.2	5.2	12.0	0.8	4.9	16.8	6.8	45.7
Missing	6,058	0.78	3.7	1.4	4.2	8.6	0.3	4.8	14.5	6.7	47.2
Diabetes mellitus											
No diabetes	492,800	63.60	2.1	1.2	2.8	8.8	0.3	5.0	13.0	4.7	54.8
Diabetes–noninsulin	195,421	25.22	2.3	1.6	4.3	10.1	0.5	5.2	15.2	6.0	48.2
Diabetes–insulin	84,406	10.89	3.6	1.8	7.1	13.9	1.0	6.5	20.6	9.7	37.5
Diabetes–missing treatment	1,439	0.19	3.1	2.2	4.6	11.1	0.7	4.6	15.7	8.8	41.9
Missing	815	0.11	3.8	0.7	2.6	8.8	0.5	4.5	12.3	6.9	43.7
Hypertension											
No	167,260	21.59	1.9	0.9	2.2	8.1	0.3	4.6	11.7	4.2	58.2
Yes	606,813	78.31	2.5	1.5	4.0	10.1	0.5	5.4	15.1	5.9	49.3
Missing	808	0.10	3.8	0.7	2.4	9.3	0.5	5.2	12.7	6.7	43.9
Hypercholesterolemia											
No	199,894	25.80	3.0	1.6	3.9	11.0	0.5	5.8	16.1	6.5	48.7
Yes	573,257	73.98	2.1	1.3	3.5	9.2	0.4	5.0	13.8	5.2	52.1
Missing	1,730	0.22	4.1	1.6	3.5	10.3	0.3	4.7	13.9	7.3	47.5
Past or present smoker											
No	295,999	38.20	2.4	1.4	3.7	9.0	0.4	5.1	13.9	5.3	50.1
Yes	477,911	61.68	2.3	1.3	3.6	10.1	0.5	5.3	14.7	5.7	51.9
Missing	971	0.13	3.4	0.7	3.1	9.9	0.4	5.3	13.5	9.1	41.0
Chronic lung disease											
None	612,211	79.01	2.0	1.3	3.3	8.4	0.3	4.9	13.0	4.7	53.7
Mild	85,005	10.97	2.8	1.5	4.2	12.0	0.6	5.8	16.9	7.0	45.7
Moderate	47,745	6.16	3.8	1.6	5.3	15.8	0.8	6.8	20.8	9.6	39.5
Severe	22,302	2.88	7.0	2.0	7.7	22.8	1.1	9.5	29.0	15.3	29.2
Missing	7,618	0.98	2.6	1.4	3.2	8.2	0.2	3.9	12.6	5.7	53.4

Table 1. Continued

Variable	Number of Patients		Percent of Patients Experiencing Endpoint								
	N	%	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Peripheral vascular disease											
No	653,260	84.30	2.0	1.2	3.2	8.8	0.4	4.8	13.1	4.8	53.5
Yes	120,480	15.55	4.4	2.3	6.1	14.4	0.7	7.5	21.2	9.6	38.7
Missing	1,141	0.15	3.9	1.1	3.1	11.8	0.3	5.5	14.4	7.4	43.9
Cerebrovascular disease											
No	668,073	86.22	2.1	1.1	3.3	9.0	0.4	4.9	13.4	5.0	53.4
Yes	105,792	13.65	4.0	2.9	5.8	14.0	0.6	7.2	20.7	9.3	37.7
Missing	1,016	0.13	3.2	0.6	2.4	8.9	0.3	4.2	11.4	6.8	43.2
CVA											
No CVA	717,721	92.62	2.2	1.2	3.4	9.3	0.4	5.1	13.8	5.2	52.5
Remote CVA (> 2 weeks)	53,341	6.88	4.2	3.1	6.1	15.3	0.7	7.4	22.0	10.3	35.5
Recent CVA (≤ 2 weeks)	1,763	0.23	5.0	4.9	6.5	18.8	0.9	8.7	25.4	13.0	32.5
CVA–missing timing	745	0.10	3.8	3.5	6.9	14.0	0.5	5.9	21.7	10.7	34.2
Missing	1,311	0.17	3.3	1.0	2.6	7.4	0.2	4.3	11.4	5.9	47.8
Endocarditis											
No endocarditis	773,002	99.76	2.3	1.4	3.6	9.7	0.4	5.2	14.4	5.5	51.2
Treated endocarditis	472	0.06	4.4	0.8	5.3	15.3	0.6	8.5	19.9	8.9	33.7
Active endocarditis	110	0.01	2.7	1.8	6.3	20.0	1.8	11.8	24.5	20.0	41.8
Endocarditis–missing type	90	0.01	4.4	4.4	5.8	11.1	1.1	3.3	15.6	2.2	55.6
Missing	1,207	0.16	4.1	1.0	3.5	9.0	0.2	4.6	12.8	7.0	46.2
Renal failure											
No	731,626	94.42	2.1	1.3	3.2	9.0	0.4	5.0	13.4	5.0	52.8
Yes	42,153	5.44	7.2	2.7	14.7	22.5	1.0	9.9	31.9	15.8	23.4
Missing	1,102	0.14	3.1	0.8	2.8	7.4	0.3	3.4	10.8	6.4	46.6
Renal function											
Creatinine < 1.00 mg/dL	274,197	35.39	1.6	1.1	1.5	8.0	0.3	4.4	11.2	4.0	55.6
Creatinine 1–1.49 mg/dL	398,833	51.47	2.0	1.3	3.4	8.9	0.4	5.0	13.5	4.9	53.1
Creatinine 1.5–1.99 mg/dL	57,779	7.46	4.5	2.3	10.8	16.1	0.7	7.8	25.2	10.6	34.5
Creatinine 2.0–2.49 mg/dL	12,463	1.61	6.9	2.9	14.3	21.3	0.9	9.4	31.5	15.3	24.7
Creatinine ≥ 2.5 mg/dL	7,906	1.02	8.2	3.2	20.4	23.4	0.9	11.1	37.9	18.6	20.4
Dialysis	12,415	1.60	8.4	2.7	*NA	25.3	1.2	10.5	31.5	16.4	19.6
Missing	11,288	1.46	3.3	1.2	3.1	7.6	0.3	4.3	12.9	5.9	50.1
Immunosuppressive treatment											
No	758,368	97.87	2.3	1.4	3.6	9.6	0.4	5.2	14.2	5.4	51.5
Yes	14,976	1.93	5.4	1.8	6.3	15.6	0.8	8.7	22.5	10.8	37.0
Missing	1,537	0.20	3.3	0.8	2.8	6.5	0.4	4.6	11.4	6.1	46.8
Prior CABG Surgery											
No	735,033	94.86	2.2	1.4	3.5	9.4	0.4	5.1	14.1	5.4	51.7
Yes	36,693	4.74	5.3	1.6	5.8	14.7	0.5	7.5	20.9	7.8	42.6
Missing	3,155	0.41	2.7	1.1	3.3	8.9	0.5	4.8	12.9	6.8	48.9
Prior valve surgery											
No	769,434	99.30	2.3	1.4	3.6	9.7	0.4	5.2	14.4	5.5	51.3
Yes	2,280	0.29	5.9	1.9	6.8	15.3	0.7	8.6	22.5	11.1	32.0
Missing	3,167	0.41	2.9	1.2	3.5	8.6	0.5	4.4	12.7	6.3	50.0
Prior other cardiac surgery											
No	755,653	97.52	2.3	1.4	3.6	9.6	0.4	5.2	14.3	5.5	51.3
Yes	15,218	1.96	3.9	1.5	4.9	13.1	0.6	6.6	18.6	7.6	45.5
Missing	4,010	0.52	2.8	1.0	2.9	8.5	0.4	4.4	12.2	5.9	50.5
Number of previous CV surgeries											
No previous CV surgery	723,623	93.39	2.2	1.4	3.5	9.4	0.4	5.1	14.0	5.4	51.7
One prior CV surgery	40,474	5.22	4.7	1.6	5.4	13.8	0.5	7.3	19.9	7.7	44.1
Two or more prior CV Surgeries	4,840	0.62	6.2	1.4	5.6	14.7	0.6	8.0	22.0	8.4	41.5
Missing	5,944	0.77	2.9	1.1	3.2	9.8	0.6	5.5	14.3	6.1	51.2

Table 1. Continued

Variable	Number of Patients		Percent of Patients Experiencing Endpoint								
	N	%	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Prior PCI											
No PCI	606,824	78.31	2.3	1.4	3.6	9.5	0.4	5.1	14.2	5.6	51.1
PCI ≤ 6 hours	7,373	0.95	8.9	2.1	7.4	25.8	0.6	10.5	32.6	11.3	35.5
PCI > 6 hours	155,161	20.02	2.3	1.2	3.5	9.6	0.4	5.4	14.3	5.1	52.4
PCI–missing timing	2,456	0.32	3.0	0.6	3.3	8.0	0.8	5.1	13.4	6.8	48.9
Missing	3,067	0.40	3.3	1.0	3.0	9.8	0.5	4.7	13.7	6.2	47.3
Acuity status											
Elective	381,116	49.18	1.5	1.1	2.9	6.6	0.4	4.3	11.1	4.1	55.6
Urgent	356,287	45.98	2.4	1.5	3.9	10.8	0.5	5.6	15.7	6.2	48.5
Emergent	34,513	4.45	8.1	2.6	8.3	29.6	0.7	10.4	34.1	13.3	33.2
Emergent salvage	1,967	0.25	38.6	4.9	17.4	52.7	0.7	18.4	70.0	23.5	12.6
Missing	998	0.13	3.2	1.3	3.6	9.2	0.5	4.8	13.7	6.5	43.6
MI											
No prior MI	424,599	54.80	1.5	1.1	2.8	6.9	0.3	4.5	11.3	4.1	55.4
MI > 21 days	137,522	17.75	2.1	1.3	3.5	8.9	0.5	5.2	13.9	5.4	50.4
MI 8–21 days	26,205	3.38	4.0	1.8	6.0	14.4	0.8	7.5	20.7	10.2	38.1
MI 1–7 days	148,659	19.18	3.4	1.8	4.8	14.0	0.5	6.1	18.9	7.5	45.7
MI > 6 and < 24 hours	21,044	2.72	6.0	2.4	6.7	23.6	0.5	8.1	28.1	10.4	39.1
MI ≤ 6 hours	11,539	1.49	10.4	2.6	8.6	31.2	0.6	10.6	36.8	13.3	33.5
MI–missing timing	4,064	0.52	3.6	1.6	4.6	11.3	0.5	5.6	17.5	7.2	43.9
Missing	1,249	0.16	2.1	1.1	2.4	6.6	0.2	3.8	10.2	6.7	49.4
Angina											
No	130,143	16.80	2.5	1.4	3.8	9.5	0.4	5.6	14.7	6.2	48.5
Yes	643,815	83.09	2.3	1.3	3.6	9.7	0.4	5.2	14.3	5.4	51.8
Missing	923	0.12	2.3	1.0	2.6	8.8	0.5	4.0	11.2	8.2	43.9
Cardiogenic shock											
No	758,766	97.92	2.0	1.3	3.4	8.9	0.4	5.0	13.6	5.2	51.9
Yes	14,919	1.93	18.0	3.6	14.6	49.6	1.0	15.3	55.7	23.1	18.3
Missing	1,196	0.15	2.7	1.1	3.2	8.0	0.4	4.4	12.0	7.4	44.7
Resuscitation											
No	766,674	98.94	2.2	1.3	3.5	9.4	0.4	5.1	14.1	5.4	51.5
Yes	6,939	0.90	17.1	3.0	11.4	37.5	0.9	14.0	46.1	18.2	24.3
Missing	1,268	0.16	2.2	0.8	3.4	8.0	0.6	3.9	11.5	7.3	45.0
Arrhythmia											
No arrhythmia	706,709	91.20	2.0	1.3	3.3	8.9	0.4	4.9	13.4	5.0	53.1
AFib/flutter	39,125	5.05	5.4	2.3	7.1	16.4	0.7	8.5	23.8	11.9	29.4
Heart block	10,026	1.29	5.8	1.9	6.8	16.8	0.6	9.2	24.2	9.4	36.4
Sustained VT/VF	14,336	1.85	8.2	2.0	6.8	23.8	0.6	11.1	31.5	12.0	33.0
Arrhythmia–other	1,853	0.24	3.8	1.6	5.3	12.9	0.6	6.5	19.1	7.3	39.2
Arrhythmia–missing type	1,344	0.17	3.9	1.7	4.4	11.9	0.7	7.3	17.5	8.3	37.9
Missing	1,488	0.19	2.7	1.0	3.0	7.1	0.5	3.8	11.1	6.5	45.5
Preoperative IABP											
No	714,824	92.25	2.0	1.3	3.3	8.0	0.4	4.9	12.8	4.9	52.8
Yes	58,134	7.50	6.9	2.2	7.7	30.8	0.6	9.6	34.4	12.9	32.0
Missing	1,923	0.25	4.2	1.7	4.3	10.9	0.6	5.8	16.0	7.2	45.7
NYHA class											
I	97,812	12.62	1.5	1.1	2.5	6.3	0.3	4.4	10.6	3.9	57.0
II	187,947	24.25	1.3	1.1	2.6	6.5	0.3	4.2	10.7	3.8	56.5
III	287,760	37.14	2.0	1.3	3.6	9.0	0.4	5.0	13.9	5.4	51.3
IV	165,325	21.34	4.5	1.9	5.5	16.6	0.6	7.1	21.9	8.8	42.5
Missing	36,037	4.65	2.4	1.2	3.6	8.9	0.3	5.3	13.6	5.8	47.7
Congestive heart failure											
No	666,592	86.03	1.8	1.2	2.9	7.9	0.3	4.7	12.2	4.3	54.7
Yes	106,700	13.77	5.9	2.4	8.5	21.0	0.9	8.7	28.0	13.2	29.5
Missing	1,589	0.21	3.3	1.4	3.1	9.2	0.4	4.5	13.0	7.4	49.6

Table 1. Continued

Variable	Number of Patients		Percent of Patients Experiencing Endpoint								
	N	%	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Number of diseased coronary vessels											
None	2,012	0.26	2.3	0.4	2.8	8.9	0.4	4.6	12.6	5.5	53.1
One	32,311	4.17	1.5	0.6	1.9	6.1	0.2	4.4	9.8	3.1	66.3
Two	150,881	19.47	1.8	1.0	2.7	8.0	0.4	4.5	12.0	4.5	56.3
Three	586,658	75.71	2.5	1.5	4.0	10.4	0.4	5.5	15.3	6.0	49.1
Missing	3,019	0.39	2.6	0.6	1.8	5.5	0.2	4.5	10.8	5.9	54.2
Left main disease ≥ 50%											
No	554,355	71.54	2.1	1.3	3.4	8.8	0.4	5.0	13.5	5.1	52.7
Yes	217,548	28.08	3.0	1.5	4.3	11.9	0.5	5.9	16.8	6.6	47.6
Missing	2,978	0.38	2.3	1.4	2.7	6.3	0.3	5.5	11.9	5.9	45.4
Ejection fraction (%)											
< 25	25,323	3.27	7.2	2.2	8.0	25.2	0.8	10.5	31.9	13.7	27.8
25–34	57,460	7.42	4.6	2.1	6.1	17.6	0.6	7.6	23.8	10.3	36.8
35–44	108,623	14.02	3.0	1.7	4.7	12.4	0.6	6.0	17.5	7.2	45.7
45–54	189,478	24.45	1.9	1.3	3.4	8.7	0.4	4.8	13.2	5.0	53.1
≥ 55	351,455	45.36	1.5	1.1	2.7	6.8	0.3	4.4	11.1	3.9	56.1
Missing	42,542	5.49	3.4	1.4	4.1	10.8	0.4	5.6	16.0	6.1	50.0
Mitral stenosis											
No	756,609	97.64	2.3	1.4	3.6	9.7	0.4	5.2	14.4	5.5	51.2
Yes	2,703	0.35	5.5	2.4	6.4	17.0	0.7	7.5	22.9	10.5	35.7
Missing	15,569	2.01	2.1	1.3	3.4	8.3	0.4	4.6	13.1	5.0	53.1
Aortic stenosis											
No	750,185	96.81	2.3	1.4	3.6	9.6	0.4	5.2	14.3	5.5	51.4
Yes	11,386	1.47	4.7	2.1	6.5	14.8	0.7	7.9	21.5	9.7	36.6
Missing	13,310	1.72	2.3	1.3	3.3	8.5	0.4	4.7	13.1	5.0	52.5
Tricuspid stenosis											
No	756,574	97.64	2.3	1.4	3.6	9.7	0.4	5.2	14.4	5.6	51.2
Yes	597	0.08	3.4	2.3	6.6	14.9	0.7	6.0	20.9	10.1	43.6
Missing	17,710	2.29	2.1	1.3	3.6	8.5	0.4	4.7	13.4	5.0	53.2
Pulmonic stenosis											
No	753,975	97.30	2.3	1.4	3.6	9.7	0.4	5.2	14.4	5.6	51.2
Yes	445	0.06	3.4	2.2	3.9	12.6	0.0	6.3	20.2	6.3	49.4
Missing	20,461	2.64	2.2	1.4	3.8	8.7	0.4	5.0	13.9	5.3	52.5
Mitral insufficiency											
None	622,173	80.29	2.1	1.2	3.3	8.9	0.4	4.9	13.4	5.0	53.2
Trivial	49,152	6.34	2.4	1.6	4.2	10.5	0.4	5.7	15.7	6.2	47.9
Mild	60,811	7.85	3.7	2.0	5.7	14.3	0.5	6.9	20.3	8.6	40.3
Moderate	16,723	2.16	6.7	2.7	7.9	20.1	0.7	9.6	28.0	12.5	30.3
Severe	2,143	0.28	8.7	3.1	8.9	24.1	0.6	11.0	32.6	15.1	28.2
Missing	23,879	3.08	2.1	1.2	3.0	7.5	0.4	4.7	12.0	5.2	51.5
Aortic insufficiency											
None	705,771	91.08	2.3	1.3	3.5	9.5	0.4	5.1	14.1	5.4	51.9
Trivial	17,988	2.32	3.6	2.1	5.6	13.4	0.5	7.0	19.4	8.3	40.9
Mild	18,571	2.40	4.1	2.2	5.9	14.3	0.4	7.3	20.8	9.0	37.9
Moderate	3,576	0.46	5.3	2.6	7.0	16.2	0.5	7.9	23.2	10.2	32.8
Severe	411	0.05	7.1	1.9	6.7	15.6	0.7	9.5	25.8	10.9	37.2
Missing	28,564	3.69	2.1	1.2	3.2	7.9	0.4	4.6	12.5	5.3	51.4
Tricuspid insufficiency											
None	675,778	87.21	2.2	1.3	3.4	9.4	0.4	5.1	13.9	5.3	52.1
Trivial	32,856	4.24	2.5	1.6	4.5	11.1	0.4	6.1	16.6	6.7	47.4
Mild	29,611	3.82	3.9	2.2	5.9	14.7	0.5	7.4	21.0	9.1	39.3
Moderate	5,753	0.74	7.6	3.0	9.0	22.7	0.5	9.8	30.2	13.7	26.9
Severe	728	0.09	9.1	2.9	10.5	24.2	0.4	10.9	33.0	17.2	26.2
Missing	30,155	3.89	2.2	1.2	3.2	7.9	0.4	4.6	12.6	5.2	51.7

Table 1. Continued

Variable	Number of Patients		Percent of Patients Experiencing Endpoint								
	N	%	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Pulmonic insufficiency											
None	724,258	93.47	2.3	1.4	3.6	9.7	0.4	5.2	14.3	5.5	51.4
Trivial	10,726	1.38	2.8	1.5	4.3	12.3	0.5	6.5	17.3	7.3	44.8
Mild	4,867	0.63	3.8	2.1	5.6	14.1	0.4	7.4	21.0	9.1	39.7
Moderate	546	0.07	6.6	3.1	7.8	17.6	0.2	7.9	24.4	11.5	29.7
Severe	217	0.03	5.1	0.5	5.1	9.7	0.5	6.5	15.7	8.8	50.7
Missing	34,267	4.42	2.2	1.3	3.5	8.3	0.4	4.8	13.2	5.5	50.7

AFib = atrial fibrillation; CABG = coronary artery bypass graft surgery; Comp = composite adverse outcome (any); CV = cardiovascular; CVA = cerebrovascular accident (stroke); DSWI = deep sternal wound infection; IABP = intra-aortic balloon pump; MI = myocardial infarction; Mort = mortality; Na = not applicable; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation; VF = ventricular fibrillation; VT = ventricular tachycardia.

cholesterolemia, a decision was made not to include this variable in the new models.

Predictor Frequency

For each variable, the QMTF explored the overall prevalence and missing data frequency per year. Predictor variables that are rarely present in the development sample are difficult to model. For this reason, mitral (0.35%), tricuspid (0.08%), and pulmonic stenosis (0.06%), pulmonic insufficiency (0.10%), and endocarditis (0.09%) were not considered as variables in the new isolated CABG models.

Inconsistently Coded Variables

A few variables have been collected inconsistently or with questionable reliability, often for clinically unavoidable reasons. For example, pulmonary artery mean pressure data were missing for 70% of patients during 2002 to 2006. Furthermore, the value of this continuous variable may vary substantially depending on the clinical state and volume-loading status of the patient when the measurement is obtained. Because of these concerns, pulmonary artery pressure was not included in the models.

Derived or Redundant Variables

Several derived variables were considered for inclusion in the models. For example, body mass index (BMI) is a useful measure of overall body habitus. However, because BMI is highly correlated with body surface area (BSA), the more commonly used anthropometric measure in most previous STS models, the latter was retained in the new models. Similarly, there is a theoretical superiority to inclusion of glomerular filtration rate (GFR) rather than serum creatinine as a measure of renal function. However, the Modification of Diet in Renal Disease formula for estimating GFR is a complex function of creatinine, race, sex, and age, and not all laboratories perform this calculation automatically. Furthermore, as age, sex, and race are already model covariates, using GFR would complicate the interpretation of their regression coefficients. Some of the prognostic value of GFR comes from these variables that are already included in the model. Finally, previous studies suggest that various measures of renal function used in CABG mortality risk models have similar performance [19]. For all these reasons, serum creatinine was retained as the measure of renal function.

Controversial Variables

RACE. Several variables raised particular clinical, statistical, or health policy issues. For example, race was an obvious candidate variable because it was a significant predictor ($p < 0.001$) of each endpoint except mortality and because the proportion of nonwhite patients varied substantially across institutions. In exploratory analyses, the association between race and outcomes persisted after adjusting for hospital identity, suggesting that this association is not explained by differences in hospital quality.

However, general principles of risk model development complicated the decision as to whether or not to include race in the models. When the dominant purpose of a risk model is adjustment of provider results, it is advisable to include only biological and clinical patient variables that are present before a patient's first contact with the provider. In this context, race is clearly a fixed biological characteristic, but its impact on patient outcomes may be mediated through other mechanisms. It is possible that certain racial and ethnic groups have worse outcomes not because of inherent biological characteristics but because of differences in the quality of care delivered to them. In this case, including race and ethnicity in a risk model could essentially select out or obscure the very disparity issues that society wishes to identify and correct. Inclusion of race and ethnicity in a risk model would say, in effect, that we expect nonwhites to have inferior results and would make an allowance for providers who care for such patients, just as we would for providers who care for patients in cardiogenic shock.

After deliberation regarding the pros and cons, the QMTF ultimately elected to retain race and ethnicity in the new models because of their impact on outcomes,

Table 2. Initial List of Potential Candidate Variables

Demographics

1. Age
2. Sex
3. Race (black, Caucasian, Hispanic, Asian, Native American, other)

Note: Data collection changed in v2.61. New version allows for multiple races (check all that apply). Added Hawaiian/Pacific Islander category. Hispanic ethnicity is a separate variable.

Anthropometric

4. Height
5. Weight

Status

6. Status (elective, urgent, emergent, salvage)
7. Shock
8. Resuscitation

Cardiac variables

9. Angina, angina type (STS categories are unstable, stable, no angina)

Note: Angina was removed on v2.61 data collection form. The new form has a variable called "cardiac presentation on admission." Angina is one of possible response categories to that field.

10. New York Heart Association functional class

Note: In v2.61, NYHA class is only collected if patient has congestive heart failure.

11. Arrhythmia and arrhythmia type (sustained VT/VF; heart block; AFib/flutter, None)
12. Myocardial infarction timing: ($\leq$ 6, $>$ 6 and $<$ 24 hours; 1–7, 8–21, $>$ 21 days)

Hemodynamic/catheterization variables

13. Ejection fraction
14. Number of diseased vessels (0, 1, 2, 3)
15. Left main disease
16. Pulmonary artery mean pressure
17. Mitral stenosis
18. Aortic stenosis
19. Tricuspid stenosis
20. Pulmonic stenosis
21. Mitral insufficiency (none, trivial, mild, moderate, severe)
22. Aortic insufficiency (none, trivial, mild, moderate, severe)
23. Tricuspid insufficiency (none, trivial, mild, moderate, severe)
24. Pulmonic insufficiency (none, trivial, mild, moderate, severe)

Comorbidities

25. Serum creatinine
26. Dialysis
27. Renal failure

Note: This variable was removed in v2.61.

28. Endocarditis (active, treated, none)
29. Diabetes and treatment (insulin, oral, diet, untreated, no diabetes)
30. Chronic lung disease (none, mild, moderate, severe)
31. Congestive heart failure
32. Peripheral vascular disease
33. Cerebrovascular disease
34. CVA and CVA timing (recent, remote, none)

Note: CVA is a child field of cerebrovascular disease in v2.61.

35. Hypercholesterolemia (v2.35, v2.41) and Dyslipidemia (v2.52)

Note: Data from all 3 versions were merged and analyzed under the variable name "hypercholesterolemia."

36. Hypertension
37. Smoker

Note: Major definition change in v2.61.

Preoperative interventions

38. Preoperative intra-aortic balloon pump
39. Preoperative inotropes
40. Immunosuppressive treatment
41. Prior percutaneous coronary intervention and timing ($\leq$ 6 hours, $>$ 6 hours, none)

Previous Interventions

42. Prior coronary artery bypass graft surgery
43. Prior valve surgery
44. Prior other cardiac surgery
45. Number of previous cardiovascular surgeries

Afib = atrial fibrillation; CVA = cerebrovascular accident; STS = The Society of Thoracic Surgeons; VT/VF = ventricular tachycardia/fibrillation.

while recognizing the potential limitations of this decision.

PREOPERATIVE INTRA-AORTIC BALLOON PUMP. Preoperative intra-aortic balloon pump (IABP) is a proxy for more serious preoperative status of the patient (eg, unstable angina, ventricular dysfunction). It captures information that may not be present in other data elements, and it is associated with higher risk of postoperative morbidity and mortality. For these reasons, most CABG risk models include preoperative IABP as a risk predictor. However, placement of an IABP is also a highly discretionary care process the frequency of which varies widely among participating institutions. Indications are subjective and are often dictated by the cardiologist before even referring the patient for cardiac surgery. Based on CABG risk models, an institution that liberally utilizes IABPs will have a higher expected risk of morbidity and mortality (according to the model) compared with another institution with a similar case-mix but a more restrictive IABP policy. That would impact their relative O/E ratios and risk-adjusted outcomes.

Despite its discretionary nature (and the potential for gaming), the QMTF decided to retain IABP use in the models because it is such an important predictor. Ultimately, it was elected to model preoperative IABP as a joint variable with preoperative inotrope use as an overall measure of preoperative acuity/severity.

Review of External Sources

The QMTF also reviewed multiple external resources to aid in the selection of potential candidate variables [15, 16, 20]. First, all previous versions of the STS CABG risk models were reviewed. The QMTF also examined other CABG risk models including the European System for Cardiac Operative Risk Evaluation (EuroSCORE) [21], the New York Cardiac Surgery Reporting System [22], the Veterans Affairs Administration cardiac surgery models [23, 24], and the Northern New England Cardiovascular Disease Study Group model [25, 26]. We particularly wanted to identify variables that were found in some form across all the risk models. Subject to the constraints of version 2.61 data specifications, we made a special effort to include such variables in the new STS risk models, in some instances requiring us to "force" them into the models, as described in the section on the final variable selection procedure.

Missing Data

Missing data in the STS NCD are rare, having a frequency of less than 1% for most variables. Candidate predictor variables missing most commonly were ejection fraction (5.5%), New York Heart Association (NYHA) class (4.7%), tricuspid insufficiency (3.9%), aortic insufficiency (3.7%), mitral insufficiency (3.1%), aortic stenosis (1.7%), and creatinine/dialysis (1.5%).

Missing predictor values in the STS NCD were managed using imputation. Multiple imputation is the generally preferred statistical method [27], but single imputation was also considered based on the following practical considerations: (a) the fraction of missing data in the STS NCD was small and, hence, single and multiple imputation would likely give similar point estimates; (b) a slight adjustment to the standard errors would not impact the study conclusions or the published risk algorithms; (c) the large sample size would make multiple imputation less practical to implement because of long computational times.

Prior to selecting an imputation strategy, exploratory analyses were performed using CABG data from 2002 to 2003 to compare single versus multiple imputation results for predicting mortality. These analyses confirmed that the choice between single versus multiple imputation would have only a slight impact on regression coefficients. For example, the estimated odds ratio for a 5-unit increase in ejection fraction was 0.90 (with a 95% confidence interval extending from 0.83 to 0.97) under single imputation and was 0.92 (with a confidence interval extending from 0.85 to 0.99) under multiple imputation. Other variables were missing less frequently than ejection fraction and were even less sensitive to the choice between single versus multiple imputation. Additional analyses of missing data consisted of reestimating the final model coefficients using single versus multiple imputation and comparing results. A summary of these investigations, as well as model coefficients and covariance matrices, are available at www.sts.org/riskmodels. For most patients, if risk were calculated using the multiple imputation model instead of single imputation, the relative change in their risk estimate would only be 1% to 2% (eg, 5% to 5.1% is a 2% change).

Based on the considerations described above, single imputation was used with the following specific rules: (1) binary (yes/no) risk factors were modeled as yes versus no or missing. Missing data for such variables usually implies their absence, and for most binary variables the composite event rates were similar for "no" and "missing" categories; (2) missing data on categorical predictor variables were imputed to the lowest risk value, which, in most instances, was the mode. In most instances, composite event rates for patients with missing data were among the lowest. It is the policy of the STS Data Warehouse and Analysis Center to discourage missing data through this default coding practice; and (3) missing data on continuous predictor variables were imputed to the conditional median. For ejection fraction, we conditioned on congestive heart failure (CHF) and sex. For BSA, we conditioned on sex. For serum creatinine, we conditioned on renal failure (although this approach will be modified when the model is ultimately applied to version 2.61 data, as renal failure has been removed).

For model endpoints (eg, mortality), missing data were handled by modeling yes versus no or missing. Thus, cases with missing data for an endpoint were analyzed as if the endpoint did not occur. Complete case analysis was not used because "missing" was not considered to be consistently coded for these variables. For example, some STS data managers have reported that they set complications to "no" unless there is explicit documentation in the medical record that the complication occurred. Other

data managers may leave the field missing unless there is explicit documentation that the complication did not occur. Thus, missing data may reflect differences in coding practices rather than truly unknown or missing data.

Preliminary Analyses for Ordinal Categorical Variables and Continuous Variables

The QMTF conducted preliminary analyses to determine how best to model ordinal categorical variables and continuous variables. Categorical variables were entered into a logistic regression model by including a separate parameter for each category. Continuous variables were entered as piecewise linear functions (splines) with several changes of slope (knots). Terms were then removed one at a time using backward selection based on the Wald statistic. At each iteration, either two adjacent categories were collapsed into a single category or else two adjacent line segments were collapsed into a single line with no change of slope. The backward selection terminated when all adjacent categories and slopes were statistically different from one another at $p < 0.001$. This variable selection routine was performed separately for each endpoint. An expert panel determined the final coding based on the results of the backwards selection algorithm, supplemented by their clinical judgment and practical considerations. Table 3 summarizes these coding decisions.

Specific Coding Decisions

RACE AND ETHNICITY. In versions 2.35, 2.41, and 2.52.1, race was collected by choosing one of the following mutually exclusive response categories: Caucasian, black, Hispanic, Asian, Native American, and other. In version 2.61, the data collection form was modified to conform to standards adopted by the US Census Bureau. It allows for selecting one or more races per patient (ie, select all that apply), and treats ethnicity (Hispanic versus non-Hispanic) as a separate variable. Because of these differences, the mapping of race among data versions is not straightforward.

Ultimately, the QMTF decided to model race as black, Asian, Hispanic, and Caucasian/other (collapsed). Initially, these categories will be mapped to version 2.61 as follows: (1) black will include all black patients, regardless of ethnicity or additional races; (2) Hispanic will include all nonblack Hispanic patients; (3) Asian will include all Asian patients who are not also identified as black or Hispanic; and (4) all remaining patients will be placed in the Caucasian/other category. The validity of this mapping will be assessed once 2.61 data become available and future versions could employ race "bridging" methodologies.

BODY SURFACE AREA. Height and weight were replaced by BSA, which was modeled as a quadratic trend to allow for a possible U-shaped relationship with outcomes (eg, extreme obesity and cachexia). This quadratic polynomial was modeled separately for males and females. Any BSA values below 1.4 or above 2.6 were mapped to these values respectively, which represent the approximate 1st and 99th percentiles of the empirical distribution.

ANGINA. Version 2.61 of the data collection form eliminates angina and substitutes a new variable called "cardiac presentation on admission," within which unstable angina is one of the possible response categories. The QMTF believed that unstable angina would be coded more consistently than any other angina class, and also that this was the most important type of angina presentation to include in the models. Angina coding was therefore restricted in the new risk models to "unstable angina without MI $<$ 7 days (yes/no)." It was necessary to exclude patients with myocardial infarction less than 7 days because the new version 2.61 does not permit simultaneous coding of angina and acute myocardial infarction.

REOPERATIVE STATUS. The most important consideration with regard to reoperative status is the number of prior sternotomies, irrespective of the specific type of procedure performed. The revised models replaced prior CABG, prior valve, and prior "other" cardiac surgery with simply the number of previous cardiovascular surgeries.

ACUITY STATUS. The new models combine resuscitation with salvage status. By definition, all salvage patients should have resuscitation coded "yes."

NUMBER OF DISEASED CORONARY VESSELS. Outcomes are modeled using the number of diseased vessels (grouped as 0 or 1 versus 2 versus 3), as a linear effect across the three categories. This approach is consistent with the previous STS CABG models and was supported by the data.

NYHA CLASS. Version 2.61 uses NYHA class as a subfield of CHF. The grouping of NYHA IV versus less than IV (I–III) classes is consistent with all existing STS models. The final categories were no CHF, CHF not NYHA IV, and CHF plus NYHA IV.

AGE. Age was modeled as a linear spline with knots at ages 50 and 60 years.

EJECTION FRACTION. Ejection fraction (EF) was modeled linearly, and EFs below 10% and above 50% were mapped to these values respectively. Only 0.03% of patients have EFs lower than 10%; such values are considered invalid and are treated like missing data. The coding decision regarding EF values above 50% was based on preliminary analyses in which the data were used to suggest the functional form of continuous variables.

CREATININE. Creatinine was modeled as a linear spline with knots at 1.0 and 1.5. Creatinine levels less than 0.5 or greater than 5.0 were mapped to these values respectively, which represent the approximate 1st and 99th percentiles of the empirical distribution.

MORTALITY AND LENGTH OF STAY. The QMTF changed the previous STS definition of the "short postoperative length of stay (SLOS)" endpoint. The original definition did not specifically exclude early postoperative deaths, and such patients could have been inappropriately included with the remaining SLOS patients who had a particularly short and uncomplicated postoperative course. In the new models, patients who die within 5 days of surgery are included in the analysis but are not counted as a short stay.

Table 3. Final List of Candidate Variables and Coding For STS Risk Models

Candidate Variables	Coding
Continuous variables	
Age[a]	Linear spline with knots at 50 and 60.
Ejection fraction[a]	Linear; values > 50 are mapped to 50. Only 0.03% of patients have ejection fraction < 10, and that is presumed to be a data entry error; these values are considered invalid and are treated like missing data. The decision to consolidate values > 50 was based on initial exploratory analyses in which data were used to suggest the functional form of continuous variables.
Body surface area[a]	Quadratic polynomial modeled separately for males and females. Note: body surface areas < 1.4 and > 2.6 were mapped to these values, respectively.[c]
Creatinine[a]	Linear spline with knots at 1.0 and 1.5. (Only for patients not on dialysis.) Note: Creatinine values < 0.5 and > 5.0 were mapped to these values, respectively.[d]
Time trend[a]	Ordinal categorical variable with separate category for each 6-month harvest interval.
Binary variables	
Dialysis[a]	Yes/no
Preoperative atrial fibrillation[b]	Yes/no
Shock	Yes/no
Female[a]	Yes/no
Hypertension	Yes/no
Immunosuppressive treatment	Yes/no
Percutaneous coronary intervention ≤ 6 hours	Yes/no
Preoperative intra-aortic balloon pump or inotropes	Yes/no
Peripheral vascular disease	Yes/no
Unstable angina (no myocardial infarction < 7 days)	Yes/no
Left main disease	Yes/no
Aortic 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	4 groups: (1) none, (2) mild, (3) moderate, (4) severe
CVD/CVA	3 groups: (1) no CVD, (2) CVD no CVA, (3) CVD + CVA
Diabetes mellitus	3 groups: (1) insulin diabetes, (2) noninsulin diabetes, (3) other or no diabetes
Number diseased coronary vessels	3 groups: (1) fewer than 2 diseased vessels, (2) 2 disease vessels, (3) 3 diseased vessels; modeled as linear across the categories.
Myocardial infarction	4 groups: (1) ≤ 6 hours, (2) > 6 and < 24 hours, (3) 1 to 21 days, (4) > 21 days or no myocardial infarction.
Race	4 groups: (1) black, (2) Asian, (3) Hispanic, (4) other, including Caucasian
Status	4 groups: (1) elective, (2) urgent, (3) emergent, no resuscitation, (4) salvage or emergent with resuscitation
Previous cardiovascular operations	3 groups: 0 previous, 1 previous, 2 or more previous
CHF and NYHA class	3 groups: no CHF, CHF not NYHA IV, CHF + NYHA IV
Interactions	
Age by reoperation[a]	
Age by emergent status[a]	

[a] These variables were forced into each model. [b] Preoperative atrial fibrillation was forced into the model for stroke. [c] These are the approximate 1st and 99th percentiles of the empirical distribution. Values less than 1.4 were mapped to 1.4. Values greater than 2.6 were mapped to 2.6. Estimates in the extreme tails of the body surface area distribution are highly influenced by data from other regions of the body surface area distribution (owing to use of a parametric, quadratic model) and may not be reliable. [d] These are approximately the 1st and 99th percentiles of the empirical distribution. Although we used a flexible spline model, linear splines can have unreliable extreme results in the tails due to the assumption that the effect is linear above the largest knot and below the smallest knot.

CHF = congestive heart failure; CVA = cerebrovascular accident; CVD = cerebrovascular disease; NYHA = New York Heart Association.

Table 4. Discrimination of Models (C-Index)

New STS models—development sample (C-index)								
Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
0.810	0.716	0.795	0.756	0.706	0.657	0.724	0.769	0.727
New STS models—validation sample (C-index)								
Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
0.812	0.720	0.793	0.754	0.689	0.653	0.725	0.767	0.726
Old STS models—validation sample (C-index)								
Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
0.807	0.713	0.750	0.742	0.672	0.645	0.711	0.754	0.713

Comp = composite adverse outcome (any); CVA = stroke; DSWI = deep sternal wound infection; Mort = mortality; PLOS = prolonged length of stay; Reop = reoperation; RF = renal failure; SLOS = short length of stay; STS = The Society of Thoracic Surgeons; Vent = prolonged ventilation.

Final Variable Selection Procedure

Backward Selection

Using the remaining candidate variables and the coding schemes described previously, a supervised backward selection approach was then performed. Initial variable selection used the Wald χ^2 statistic with a significance criterion of 0.001. This high level of significance was chosen because of the very large sample size that resulted in quite small p values. An expert panel of cardiothoracic surgeons and biostatisticians then reviewed the selected variables and made several modifications. Measures of model performance (discrimination and calibration) were similar when all variables were retained in the models regardless of statistical significance or expert panel review.

Forced Variables

Several variables were included in the models regardless of statistical significance. These included all of the continuous variables (age, BSA, date of surgery [in 6-month intervals], creatinine, ejection fraction), plus sex and dialysis. In addition, atrial fibrillation was included a priori in the model for permanent stroke.

The rationale for including surgery date, a nonmodifiable variable of no intrinsic interest, was to adjust for changes in the frequency of adverse outcomes over the 5-year study period. We adjusted for surgery date to reduce potential confounding by time trends when estimating regression coefficients for variables that are of primary interest, such as preoperative clinical characteristics. For example, temporal changes in the frequency of coding for dyslipidemia, if they occur coincidentally with a secular declining trend in mortality rates, may lead to the unwarranted causal inferences unless there is adjustment for surgery date.

Date of surgery was categorized by 6-month intervals (corresponding to STS data harvests) and modeled as a linear trend across the ordinal categories. Surgery date is not included in the final risk algorithm and a patient's predicted risk is not dependent upon it. The intercept parameter published in the Appendix has been adjusted to incorporate the time trend, and it reflects the baseline risk for a reference period of July to December 2006.

Interaction Terms

These models focused on main effects, and the final models included only four sets of preselected variable interactions: (1) sex by BSA; (2) sex by BSA squared; (3) age by reoperation; (4) age by emergent status. More extensive investigation for interactions was considered, including nonlinear, machine-learning approaches. However, the incremental value of such approaches remains uncertain [28], and interpretability can also become more problematic with numerous interaction terms.

Although multiple terms were allotted for modeling the main effects of age and reoperation, only a single degree of freedom was allotted for their interaction. The models defined a single variable interaction term for age and reoperation. It was equal to the patient's age minus 50 if the patient was at least 50 years old and had a previous CV surgery; otherwise it was equal to zero. This term represents the difference in the change of the slope of age at age 50 for patients who have had at least one previous CV surgery compared with patients who have not had a previous CV surgery. Similarly, only one degree of freedom was allotted for the interaction between age and status. The interaction represents the difference in the change of the slope of age at age 50 for patients with emergent or salvage status compared with patients with elective or urgent status. Although these interaction terms complicate the interpretation of other model variables, this was considered to be acceptable because the main focus of the analysis was prediction, not effect estimation.

Results

Model Performance

Table 4 presents the discrimination of each of the isolated CABG models as well as a comparison with the previous

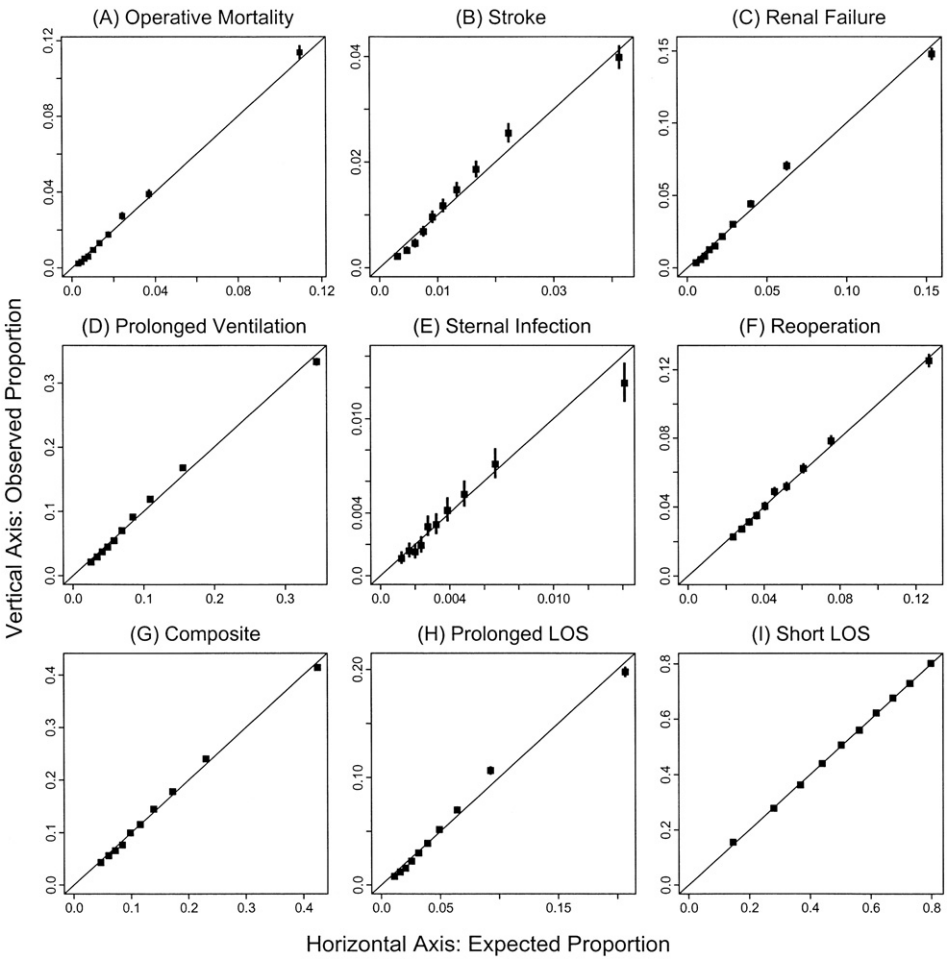


Fig 1. Plots of observed (O) versus expected (E) in validation sample

STS CABG risk models. For the new CABG models, discrimination ranged from 0.657 to 0.810 in the development sample and from 0.653 to 0.812 in the validation sample. The close agreement between c-indices from the development and validation samples reflects the large sample size and suggests that the models did not overfit the data. When the discrimination of the new and previous STS models were compared using the validation sample, the c-index of the new model was larger for each endpoint.

The Hosmer-Lemeshow test is not reported as an overall measure of calibration for these models because of its sensitivity to sample size. With samples as large as those used to develop these models, the null hypothesis will inevitably be proven false, given that all such models are only approximations [29]. As an alternative to such global measures of calibration, Figure 1 shows plots of observed versus expected event proportions within deciles of predicted risk for a variety of endpoints. For each endpoint, the absolute difference between the observed and expected proportions was less than 1.5% in each decile category. Additional analyses of model fit and discrimination are available online at www.sts.org/riskmodels.

Final Models

After calculating these measures of model performance, the final regression coefficients were estimated from the combined training and validation samples. Odds ratios for each predictor variable and model endpoint are summarized in Table 5. "Not applicable" indicates that the specific predictor was not included in a particular risk model. These final models were estimated using generalized estimating equations with empirical (sandwich) standard error estimates to account for clustering of patients within institutions [30]. An independence working correlation matrix was used to apply the generalized estimating equations method. 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 correlated observations within hospitals.

Final Model Intercept and Coefficients

The Appendix contains the algorithm, intercept and coefficients for the final STS 2008 CABG risk models. The variance/covariance matrix is available on the web at www.sts.org/riskmodels. An on-line risk calculator is available at http://209.220.160.181/STSWebRiskCalc261/.

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

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Age 60 versus 50 (no reoperation, elective)	1.36 (1.24, 1.49)	1.78 (1.58, 1.99)	1.24 (1.16, 1.33)	1.06 (1.02, 1.10)	1.43 (1.23, 1.67)	1.14 (1.09, 1.19)	1.08 (1.05, 1.11)	1.35 (1.27, 1.43)	0.77 (0.75, 0.78)
Age 70 versus 50 (no reoperation, elective)	2.53 (2.31, 2.76)	2.43 (2.19, 2.71)	1.93 (1.81, 2.07)	1.42 (1.37, 1.47)	1.70 (1.47, 1.97)	1.45 (1.39, 1.51)	1.49 (1.44, 1.53)	2.17 (2.05, 2.29)	0.44 (0.43, 0.45)
Age 80 versus 50 (no reoperation, elective)	4.70 (4.29, 5.15)	3.34 (2.99, 3.72)	3.01 (2.80, 3.24)	1.90 (1.82, 1.99)	2.02 (1.73, 2.36)	1.85 (1.76, 1.94)	2.05 (1.98, 2.12)	3.48 (3.28, 3.69)	0.25 (0.24, 0.26)
BSA 1.6 versus 2.0 among females	1.26 (1.19, 1.32)	1.15 (1.08, 1.23)	0.84 (0.80, 0.89)	1.03 (1.00, 1.06)	0.49 (0.43, 0.57)	1.23 (1.18, 1.28)	1.03 (1.01, 1.06)	0.94 (0.90, 0.97)	1.17 (1.14, 1.20)
BSA 1.6 versus 2.0 among males	1.75 (1.64, 1.86)	1.19 (1.08, 1.31)	1.24 (1.17, 1.32)	1.40 (1.34, 1.46)	0.77 (0.63, 0.93)	1.40 (1.33, 1.46)	1.35 (1.30, 1.40)	1.43 (1.36, 1.50)	0.79 (0.77, 0.82)
BSA 1.8 versus 2.0 among females	1.02 (0.99, 1.05)	1.07 (1.03, 1.11)	0.86 (0.84, 0.88)	0.95 (0.94, 0.97)	0.67 (0.63, 0.71)	1.06 (1.04, 1.08)	0.96 (0.94, 0.97)	0.90 (0.88, 0.92)	1.14 (1.13, 1.16)
BSA 1.8 versus 2.0 among males	1.20 (1.17, 1.23)	1.09 (1.05, 1.13)	1.02 (1.00, 1.04)	1.09 (1.07, 1.10)	0.85 (0.79, 0.91)	1.13 (1.11, 1.15)	1.08 (1.07, 1.09)	1.10 (1.08, 1.12)	0.96 (0.95, 0.97)
BSA 2.2 versus 2.0 among females	1.20 (1.14, 1.27)	0.95 (0.88, 1.02)	1.32 (1.27, 1.37)	1.18 (1.15, 1.22)	1.62 (1.51, 1.74)	1.03 (1.00, 1.07)	1.19 (1.16, 1.21)	1.28 (1.24, 1.32)	0.78 (0.77, 0.80)
BSA 2.2 versus 2.0 among males	1.01 (0.99, 1.03)	0.92 (0.90, 0.95)	1.17 (1.15, 1.19)	1.10 (1.08, 1.11)	1.27 (1.22, 1.32)	0.97 (0.96, 0.99)	1.07 (1.06, 1.08)	1.08 (1.07, 1.10)	0.90 (0.89, 0.91)
Creatinine 1.5 versus 1.0	1.66 (1.57, 1.76)	1.39 (1.30, 1.49)	3.36 (3.16, 3.58)	1.56 (1.51, 1.62)	1.44 (1.28, 1.62)	1.33 (1.28, 1.38)	1.76 (1.70, 1.82)	1.65 (1.59, 1.72)	0.69 (0.67, 0.71)
Creatinine 2.0 versus 1.0	1.94 (1.84, 2.04)	1.49 (1.39, 1.58)	4.06 (3.83, 4.31)	1.73 (1.68, 1.79)	1.47 (1.30, 1.65)	1.44 (1.40, 1.49)	2.05 (1.98, 2.11)	1.92 (1.86, 2.00)	0.55 (0.53, 0.57)
Creatinine 2.5 versus 1.0	2.26 (2.14, 2.39)	1.59 (1.47, 1.71)	4.90 (4.61, 5.21)	1.92 (1.85, 1.99)	1.50 (1.30, 1.72)	1.57 (1.51, 1.64)	2.39 (2.30, 2.48)	2.24 (2.15, 2.34)	0.44 (0.42, 0.46)
Dialysis versus no dialysis and creatinine = 1.0	3.84 (3.54, 4.16)	1.67 (1.48, 1.88)	NA	2.85 (2.68, 3.03)	2.13 (1.78, 2.56)	1.86 (1.73, 2.00)	2.46 (2.33, 2.60)	2.80 (2.63, 2.98)	0.27 (0.25, 0.29)
EF per 10-unit decrease	1.19 (1.17, 1.22)	1.14 (1.11, 1.16)	1.08 (1.06, 1.10)	1.18 (1.16, 1.20)	1.11 (1.07, 1.16)	1.11 (1.09, 1.13)	1.16 (1.15, 1.18)	1.17 (1.15, 1.19)	0.84 (0.83, 0.85)
Preoperative atrial fibrillation	1.36 (1.28, 1.44)	1.21 (1.12, 1.30)	1.24 (1.18, 1.30)	1.20 (1.16, 1.24)	NA	1.26 (1.21, 1.31)	1.24 (1.21, 1.28)	1.42 (1.37, 1.48)	0.61 (0.59, 0.63)
CHF not NYHA IV	1.21 (1.15, 1.28)	NA	1.36 (1.30, 1.43)	1.31 (1.26, 1.35)	1.33 (1.19, 1.48)	1.16 (1.11, 1.21)	1.27 (1.23, 1.31)	1.43 (1.38, 1.48)	0.72 (0.70, 0.75)
CHF NYHA IV	1.39 (1.31, 1.47)	NA	1.35 (1.28, 1.42)	1.52 (1.45, 1.59)	1.45 (1.25, 1.67)	1.26 (1.20, 1.32)	1.48 (1.42, 1.54)	1.50 (1.44, 1.57)	0.65 (0.61, 0.68)
Chronic lung disease, mild	1.22 (1.16, 1.29)	NA	1.14 (1.08, 1.21)	1.36 (1.31, 1.41)	1.56 (1.40, 1.73)	1.11 (1.07, 1.15)	1.23 (1.19, 1.27)	1.34 (1.29, 1.39)	0.79 (0.76, 0.82)
Chronic lung disease, moderate	1.40 (1.32, 1.49)	NA	1.25 (1.18, 1.33)	1.65 (1.57, 1.73)	1.80 (1.58, 2.06)	1.20 (1.14, 1.26)	1.42 (1.36, 1.47)	1.65 (1.58, 1.73)	0.68 (0.65, 0.71)
Chronic lung disease, severe	2.35 (2.19, 2.52)	NA	1.66 (1.54, 1.79)	2.37 (2.24, 2.51)	2.40 (2.06, 2.79)	1.54 (1.44, 1.64)	1.98 (1.90, 2.07)	2.46 (2.34, 2.60)	0.48 (0.45, 0.51)
CVD with CVA	1.31 (1.24, 1.38)	2.09 (1.96, 2.22)	1.18 (1.12, 1.23)	1.35 (1.31, 1.39)	NA	1.21 (1.17, 1.26)	1.32 (1.29, 1.36)	1.45 (1.40, 1.51)	0.70 (0.68, 0.72)

Table 5. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
CVD without CVA	1.14 (1.08, 1.20)	1.65 (1.54, 1.75)	1.11 (1.06, 1.17)	1.15 (1.11, 1.18)	NA	1.12 (1.08, 1.17)	1.17 (1.14, 1.20)	1.14 (1.10, 1.18)	0.85 (0.81, 0.89)
Diabetes, insulin dependent	1.30 (1.24, 1.37)	1.19 (1.12, 1.27)	1.80 (1.72, 1.87)	1.22 (1.18, 1.26)	2.24 (2.02, 2.48)	1.14 (1.10, 1.18)	1.30 (1.27, 1.34)	1.59 (1.53, 1.64)	0.64 (0.62, 0.66)
Diabetes, noninsulin dependent	1.01 (0.97, 1.06)	1.16 (1.11, 1.22)	1.32 (1.28, 1.36)	1.04 (1.02, 1.07)	1.38 (1.27, 1.49)	0.98 (0.96, 1.01)	1.08 (1.06, 1.10)	1.15 (1.12, 1.17)	0.87 (0.86, 0.88)
Diseased vessels (2 versus 1, or 3 versus 2)	1.17 (1.12, 1.23)	1.35 (1.29, 1.42)	1.23 (1.19, 1.27)	1.19 (1.16, 1.22)	1.15 (1.07, 1.24)	1.07 (1.05, 1.10)	1.16 (1.14, 1.18)	1.15 (1.11, 1.18)	0.81 (0.80, 0.82)
Preoperative IABP/ inotropes	1.41 (1.33, 1.49)	NA	1.43 (1.36, 1.51)	2.56 (2.42, 2.72)	NA	1.37 (1.31, 1.43)	1.96 (1.86, 2.06)	1.60 (1.53, 1.67)	0.60 (0.57, 0.63)
Shock	2.29 (2.12, 2.47)	1.38 (1.23, 1.55)	1.65 (1.54, 1.77)	2.08 (1.96, 2.21)	NA	1.43 (1.34, 1.52)	2.10 (1.99, 2.23)	1.73 (1.62, 1.84)	0.58 (0.54, 0.62)
Female versus male (at BSA = 1.8)	1.31 (1.25, 1.36)	1.32 (1.24, 1.39)	1.25 (1.21, 1.31)	1.33 (1.29, 1.36)	1.19 (1.06, 1.35)	0.90 (0.87, 0.93)	1.18 (1.15, 1.21)	1.24 (1.20, 1.28)	0.65 (0.63, 0.66)
Hypertension	NA	1.29 (1.22, 1.37)	1.25 (1.20, 1.30)	1.10 (1.08, 1.13)	NA	NA	1.12 (1.10, 1.15)	1.07 (1.04, 1.11)	0.92 (0.90, 0.94)
Immunosuppressive treatment	1.48 (1.37, 1.60)	NA	1.21 (1.12, 1.31)	1.11 (1.05, 1.18)	NA	1.32 (1.24, 1.41)	1.20 (1.14, 1.26)	1.28 (1.20, 1.37)	0.80 (0.76, 0.84)
Aortic insufficiency, moderate/severe	NA	NA	NA	NA	NA	NA	NA	NA	0.82 (0.75, 0.89)
Mitral insufficiency, moderate/severe	1.31 (1.21, 1.41)	NA	NA	1.12 (1.06, 1.18)	NA	1.24 (1.16, 1.32)	1.20 (1.15, 1.26)	1.15 (1.09, 1.22)	0.85 (0.80, 0.91)
Tricuspid insufficiency, moderate/severe	NA	NA	1.31 (1.17, 1.45)	1.28 (1.18, 1.39)	NA	NA	1.24 (1.16, 1.33)	NA	0.78 (0.71, 0.87)
PCI ≤ 6 hours	1.37 (1.24, 1.50)	NA	1.29 (1.16, 1.43)	1.21 (1.13, 1.29)	NA	1.30 (1.19, 1.42)	1.31 (1.23, 1.39)	1.17 (1.07, 1.27)	0.79 (0.74, 0.84)
Peripheral vascular disease	1.42 (1.36, 1.48)	1.32 (1.26, 1.39)	1.21 (1.17, 1.26)	1.22 (1.19, 1.26)	1.36 (1.24, 1.48)	1.24 (1.20, 1.28)	1.25 (1.22, 1.28)	1.31 (1.28, 1.35)	0.82 (0.81, 0.84)
Aortic stenosis	NA	NA	NA	1.18 (1.11, 1.26)	NA	NA	1.16 (1.10, 1.22)	1.15 (1.07, 1.23)	0.87 (0.82, 0.92)
Left main disease	NA	NA	NA	1.07 (1.04, 1.09)	NA	NA	1.04 (1.02, 1.06)	NA	NA
MI 1–21 days	1.37 (1.32, 1.44)	1.31 (1.25, 1.37)	1.27 (1.22, 1.32)	1.34 (1.29, 1.38)	NA	NA	1.23 (1.20, 1.25)	1.22 (1.18, 1.25)	0.88 (0.86, 0.90)
MI > 6 and < 24 hours	1.59 (1.46, 1.74)	1.59 (1.43, 1.76)	1.48 (1.36, 1.60)	1.59 (1.49, 1.68)	NA	NA	1.43 (1.37, 1.50)	1.31 (1.24, 1.39)	0.80 (0.76, 0.84)
MI ≤ 6 hours	1.70 (1.53, 1.89)	1.49 (1.31, 1.68)	1.43 (1.29, 1.57)	1.56 (1.45, 1.67)	NA	NA	1.44 (1.35, 1.53)	1.30 (1.21, 1.40)	0.82 (0.77, 0.87)
Time trend, per 6-month harvest interval	0.97 (0.97, 0.98)	0.97 (0.96, 0.98)	1.01 (1.00, 1.02)	1.01 (1.01, 1.02)	0.97 (0.95, 0.99)	0.99 (0.99, 1.00)	1.00 (1.00, 1.01)	1.00 (0.99, 1.01)	0.99 (0.98, 1.00)
Race Asian	NA	1.33 (1.14, 1.55)	1.08 (0.96, 1.22)	1.33 (1.21, 1.47)	1.00 (0.66, 1.51)	1.31 (1.17, 1.46)	1.23 (1.15, 1.31)	1.26 (1.13, 1.40)	0.70 (0.61, 0.81)
Race black	NA	1.41 (1.30, 1.54)	1.24 (1.16, 1.33)	1.37 (1.27, 1.48)	1.30 (1.13, 1.51)	1.21 (1.14, 1.30)	1.31 (1.24, 1.38)	1.43 (1.34, 1.51)	0.69 (0.65, 0.73)
Race Hispanic	NA	1.12 (0.98, 1.27)	1.24 (1.11, 1.39)	1.16 (1.07, 1.26)	1.30 (1.07, 1.58)	1.05 (0.97, 1.13)	1.12 (1.05, 1.19)	1.09 (0.99, 1.20)	0.85 (0.77, 0.94)

Table 5. Continued

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Reoperation, 1 previous operation[a]	3.13 (2.74, 3.57)	NA	1.52 (1.35, 1.71)	1.72 (1.58, 1.86)	NA	1.57 (1.43, 1.74)	1.61 (1.50, 1.72)	1.62 (1.47, 1.80)	0.72 (0.67, 0.78)
Reoperation, ≥ 2 previous operations[a]	4.19 (3.45, 5.09)	NA	1.58 (1.33, 1.87)	1.86 (1.62, 2.14)	NA	1.71 (1.44, 2.03)	1.84 (1.65, 2.05)	1.79 (1.53, 2.08)	0.64 (0.56, 0.73)
Status urgent[a]	1.16 (1.10, 1.22)	1.11 (1.06, 1.17)	1.12 (1.05, 1.19)	1.24 (1.18, 1.31)	1.20 (1.10, 1.32)	1.18 (1.13, 1.23)	1.18 (1.14, 1.22)	1.20 (1.15, 1.25)	0.86 (0.83, 0.90)
Status emergent, no resuscitation[a]	2.83 (2.52, 3.18)	2.12 (1.82, 2.48)	1.68 (1.49, 1.89)	2.14 (1.96, 2.34)	1.87 (1.46, 2.40)	1.83 (1.68, 1.99)	1.77 (1.64, 1.91)	2.12 (1.93, 2.32)	0.62 (0.58, 0.67)
Status emergent with resuscitation or salvage[a]	8.00 (6.91, 9.26)	2.51 (1.98, 3.18)	2.16 (1.82, 2.55)	3.01 (2.68, 3.38)	2.09 (1.45, 3.01)	2.34 (2.06, 2.65)	3.65 (3.26, 4.09)	2.39 (2.10, 2.72)	0.34 (0.30, 0.38)
Unstable angina	1.12 (1.07, 1.17)	NA	1.11 (1.05, 1.17)	1.05 (1.01, 1.10)	NA	NA	NA	NA	NA

[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; Comp = composite adverse outcome (any); 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; NA = not applicable; NYHA = New York Heart Association; PCI = percutaneous coronary intervention; PLOS = prolonged length of stay; PVD = peripheral vascular disease; Reop = reoperation; RF = renal failure; SLOS = short length of stay; Vent = prolonged ventilation.

Previously, the STS risk models were completely upgraded every 3 years, with annual recalibration in the interim to assure that the benchmark O/E ratio is always 1. In the near future, annual upgrades of the models are planned.

Limitations

Regardless of sample size or degree of statistical sophistication, all risk models are imperfect representations of reality. Although the STS risk models are based upon excellent clinical data and large sample sizes, there are some risk factors that are rare in the overall population but, when present, may be important predictors of outcome for specific patients. Some such variables, such as liver disease, are not included in the risk models, and the mortality risk for patients with these risk factors may be underestimated. Addition of a number of such variables will be considered at the next major specification upgrade.

There are other variables whose specifications undergo small but important changes over time, often in response to comments from STS database participants. These refinements are discussed on regular biweekly conference calls open to database participants, and suggested changes are regularly communicated to participants through a variety of means including FAQ's. With each major specification upgrade, they are incorporated into the new software specifications.

Audit is extremely important to assure the accuracy of any data registry. For the STS database and the risk models derived from it, robust audit is particularly critical as this registry is increasingly used for public reporting of outcomes and pay for performance. Studies suggest that the accuracy of the STS database is high for most important variables [31–35], although these audits are currently restricted to a limited number of sites annually because of budgetary constraints. In these audits, one of the most problematic variables has been 30-day mortality status (as opposed to in-hospital mortality). This is often a difficult endpoint to ascertain and may require more substantial investment of time and effort by participants, particularly for patients referred from outside their own institutions. Analysis of STS data suggests that approximately 90% of 30-day deaths occur in-hospital. Thus, if some patients recorded as being alive at 30 days have actually had their status ascertained only during the index hospitalization, the impact of this misclassification on the risk models should be negligible. This hypothesis was confirmed by comparing the odds ratios of all model variables for in-hospital versus 30-day mortality. Differences between the two were quite small, and these data are available on the web at www.sts.org/riskmodels. A new risk model for in-hospital mortality has been developed and placed on the same STS website. Furthermore, an aggressive program is in place to further enhance the accuracy of 30-day follow-up. In 2009, STS instituted a requirement that participants maintain documentation of the method by which they ascertained 30-day status, and that has become part of our routine audit. Linkage of the STS database with external death registries, such as the Social Security Death Master File, will

further support this capability. Finally, plans are being developed to expand the audit of certain key variables such as 30-day mortality to a significantly greater number of sites annually.

Conclusions

Risk-adjustment models account for the effect of patient comorbidities on outcomes. STS risk models are based upon clinical data from the STS NCD, one of the oldest and largest of all specialty registries. The value of such clinical registries is particularly evident in today's health care environment, where accreditation, regulatory compliance, reimbursement, and referrals are increasingly based upon objective data. Organizations such as the AQA and the National Quality Forum that evaluate and endorse performance measures strongly advocate the use of risk-adjusted outcomes measures.

STS believes that clinical data are superior to those derived from administrative sources. Furthermore, given the substantial implications of risk-adjusted outcomes, we believe that all risk models used for profiling quality of care should be transparent to permit comprehensive peer review and to foster credibility among stakeholders.

We present a detailed exposition of the development and validation of the 2008 STS CABG risk model. This describes not only the statistical considerations but, just as importantly, the many clinical and pragmatic judgments that are always necessary in risk model development.

References

1. Kouchoukos NT, Ebert PA, Grover FL, Lindesmith GG. Report of the Ad Hoc Committee on Risk Factors for Coronary Artery Bypass Surgery. Ann Thorac Surg 1988;45:348–9.
2. Clark RE. It is time for a national cardiothoracic surgical data base. Ann Thorac Surg 1989;48:755–6.
3. Edwards FH. Evolution of the Society of Thoracic Surgeons National Cardiac Surgery Database. J Invasive Cardiol 1998; 10:485–8.
4. Grover FL, Shroyer AL, Hammermeister K, et al. A decade's experience with quality improvement in cardiac surgery using the Veterans Affairs and Society of Thoracic Surgeons national databases. Ann Surg 2001;234:464–72.
5. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. Ann Thorac Surg 2004;78:1868–77.
6. Shahian DM, Normand SL, Torchiana DF, et al. Cardiac surgery report cards: comprehensive review and statistical critique. Ann Thorac Surg 2001;72:2155–68.
7. Normand S-LT, Shahian DM. Statistical and clinical aspects of hospital outcomes profiling. Stat Sci 2007;22:206–26.
8. Naftel DC. Do different investigators sometimes produce different multivariable equations from the same data? J Thorac Cardiovasc Surg 1994;107:1528–9.
9. Iezzoni LI. "Black box" medical information systems. A technology needing assessment. JAMA 1991;265:3006–7.
10. Shahian DM, Hutter MM, Torchiana DF, Iezzoni LI. Transparency: a mandatory requirement for risk models. J Am Coll Surg 2008;206:1240–2.
11. Krumholz HM, Brindis RG, Brush JE, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association scientific statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. Circulation 2006;113:456–62.
12. Breiman L. Statistical modeling: the two cultures. Stat Sci 2001;16:199–231.
13. Harrell FE Jr. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag, 2001.
14. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE. Regression methods in biostatistics linear, logistic, survival, and repeated measures models. New York: Springer-Verlag, 2005.
15. Jones RH, Hannan EL, Hammermeister KE, et al. Identification of preoperative variables needed for risk adjustment of short-term mortality after coronary artery bypass graft surgery. The Working Group Panel on the Cooperative CABG Database Project. J Am Coll Cardiol 1996;28:1478–87.
16. Tu JV, Sykora K, Naylor CD. Assessing the outcomes of coronary artery bypass graft surgery: how many risk factors are enough? Steering Committee of the Cardiac Care Network of Ontario. J Am Coll Cardiol 1997;30:1317–23.
17. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928–35.
18. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–72.
19. Cooper WA, O'Brien SM, Thourani VH, et al. Impact of renal dysfunction on outcomes of coronary artery bypass surgery: results from the Society of Thoracic Surgeons National Adult Cardiac Database. Circulation 2006;113:1063–70.
20. Grunkemeier GL, Zerr KJ, Jin R. Cardiac surgery report cards: making the grade. Ann Thorac Surg 2001;72:1845–8.
21. 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.
22. 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.
23. Grover FL, Johnson RR, Marshall G, Hammermeister KE. Factors predictive of operative mortality among coronary artery bypass subsets. Ann Thorac Surg 1993;56:1296–306.
24. Grover FL, Shroyer AL, Hammermeister KE. Calculating risk and outcome: the Veterans Affairs database. Ann Thorac Surg 1996;62(Suppl):6–11.
25. O'Connor GT, Plume SK, Olmstead EM, et al. A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. The Northern New England Cardiovascular Disease Study Group. JAMA 1991;266:803–9.
26. 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.
27. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd ed. Hoboken: Wiley-Interscience, 2002.
28. Austin PC. A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AMI mortality. Stat Med 2007;26:2937–57.
29. Marcin JP, Romano PS. Size matters to a model's fit. Crit Care Med 2007;35:2212–3.
30. Liang KY, Zeger SL. Longitudinal data-analysis using generalized linear-models. Biometrika 1986;73:13–22.
31. Grover FL, Shroyer AL, Edwards FH, et al. Data quality review program: the Society of Thoracic Surgeons Adult Cardiac National Database. Ann Thorac Surg 1996;62:1229–31.
32. Shroyer AL, Edwards FH, Grover FL. Updates to the Data Quality Review Program: the Society of Thoracic Surgeons Adult Cardiac National Database. Ann Thorac Surg 1998;65:1494–7.
33. Herbert MA, Prince SL, Williams JL, Magee MJ, Mack MJ. Are unaudited records from an outcomes registry database accurate? Ann Thorac Surg 2004;77:1960–4.
34. Welke KF, Peterson ED, Vaughan-Sarrazin MS, et al. Comparison of cardiac surgery volumes and mortality rates

between the Society of Thoracic Surgeons and Medicare databases from 1993 through 2001. Ann Thorac Surg 2007;84:1538–46.
35. Welke KF, Ferguson TB Jr, Coombs LP, et al. Validity of the Society of Thoracic Surgeons National Adult Cardiac Surgery Database. Ann Thorac Surg 2004;77:1137–9.

Appendix

Regression Coefficients and Variable Definitions for STS 2008 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 to December 2006.

Appendix Table 1. Regression Coefficients

Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Intercept	−6.34090	−7.18174	−7.94605	−4.15175	−6.75378	−3.84861	−3.71671	−5.35975	2.84959
Atrial fibrillation	0.30830	0.18935	0.21351	0.17871	0.00000	0.23031	0.21565	0.35322	−0.49309
Age	−0.00259	0.00996	0.00678	0.00170	−0.00665	−0.00013	0.00247	0.00914	−0.01781
Age function 1	0.03325	0.04742	0.01496	0.00393	0.04270	0.01305	0.00515	0.02085	−0.00895
Age function 2	0.03140	−0.02582	0.02249	0.02366	−0.01895	0.01133	0.02441	0.01734	−0.02904
Age by reoperation function	−0.01714	−0.00098	−0.00291	−0.00459	0.00304	−0.00720	−0.00444	−0.00809	0.00449
Age by status function	−0.01366	−0.01363	−0.00022	−0.00106	−0.00352	−0.00435	0.00270	−0.00833	−0.00266
BSA function 1	−1.39342	−0.44041	−0.53672	−0.83950	0.65513	−0.83758	−0.75006	−0.89037	0.57952
BSA function 2	2.41303	0.06122	2.19879	2.15647	0.90025	1.16543	1.81770	2.15270	−1.83776
CHF but not NYHA IV	0.19229	0.00000	0.30971	0.26853	0.28272	0.14692	0.23695	0.35623	−0.32350
CHF and NYHA IV	0.32663	0.00000	0.30013	0.41599	0.36909	0.22846	0.39005	0.40757	−0.43827
Chronic lung disease mild	0.20273	0.00000	0.13488	0.30473	0.44371	0.10432	0.20878	0.29051	−0.23600
Chronic lung disease moderate	0.33843	0.00000	0.22530	0.50235	0.59021	0.18071	0.34720	0.50246	−0.39085
Chronic lung disease severe	0.85513	0.00000	0.50645	0.86175	0.87366	0.43034	0.68538	0.90211	−0.73862
Creatinine function 1	0.19353	0.02822	1.91934	−0.02712	−0.37465	0.01583	0.13361	−0.09060	0.00773
Creatinine function 2	0.82140	0.63174	0.50685	0.92120	1.09976	0.55107	0.99190	1.09571	−0.75781
Creatinine function 3	−0.70646	−0.52856	−2.04970	−0.68907	−0.68466	−0.39956	−0.81791	−0.70069	0.30449
CVD without prior CVA	0.13177	0.49807	0.10637	0.13792	0.00000	0.11403	0.15561	0.13271	−0.16385
CVD and prior CVA	0.26877	0.73600	0.16135	0.29946	0.00000	0.19208	0.28099	0.37248	−0.35706
Diabetes noninsulin dependent	0.01375	0.14992	0.27443	0.04283	0.31888	−0.01929	0.07453	0.13541	−0.13813
Diabetes insulin dependent	0.26312	0.17483	0.58581	0.19735	0.80627	0.12930	0.26525	0.46226	−0.44725
Dialysis	1.53777	0.54158	0.00000	1.01943	0.38312	0.63691	1.03466	0.93792	−1.30294
Ejection fraction function	0.01765	0.01274	0.00754	0.01669	0.01081	0.01063	0.01496	0.01542	−0.01756
Female	0.26801	0.27414	0.22704	0.28338	0.17792	−0.10270	0.16434	0.21488	−0.43658
Female by BSA function 1	0.82285	0.08974	0.96428	0.76954	1.11546	0.31901	0.66663	1.05623	−0.96846
Female by BSA function 2	0.05606	0.06490	−0.61086	−0.62558	0.17399	−0.02390	−0.25077	−0.35160	0.46088
Hypertension	0.00000	0.25718	0.22126	0.09930	0.00000	0.00000	0.11674	0.07200	−0.08155
IABP or inotropes	0.34193	0.00000	0.36023	0.94050	0.00000	0.31326	0.67253	0.47092	−0.51444
Immunosuppressive treatment	0.39159	0.00000	0.18881	0.10686	0.00000	0.27802	0.18030	0.24833	−0.22718
Insufficiency, aortic	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	−0.19889
Insufficiency, mitral	0.26631	0.00000	0.00000	0.11169	0.00000	0.21170	0.18225	0.14174	−0.15962
Insufficiency, tricuspid	0.00000	0.00000	0.26729	0.24834	0.00000	0.00000	0.21893	0.00000	−0.24548
Left main disease	0.00000	0.00000	0.00000	0.06629	0.00000	0.00000	0.03570	0.00000	0.00000
MI 1 to 21 days	0.31810	0.27134	0.23962	0.28925	0.00000	0.00000	0.20524	0.19517	−0.12752
MI $>$ 6 and $<$ 24 hours	0.46614	0.46063	0.38917	0.46158	0.00000	0.00000	0.35859	0.27109	−0.22557
MI $\leq$ 6 hours	0.53242	0.39601	0.35421	0.44230	0.00000	0.00000	0.36337	0.26311	−0.19946
No. diseased vessel function	0.16120	0.30339	0.20729	0.17622	0.13869	0.06895	0.15075	0.13589	−0.21043
PCI $\leq$ 6 hours	0.31149	0.00000	0.25189	0.18695	0.00000	0.26256	0.26774	0.15633	−0.23860
Peripheral vascular disease	0.34951	0.27985	0.19308	0.20240	0.30529	0.21306	0.22277	0.27380	−0.19321
Race black	0.00000	0.34423	0.21696	0.31563	0.26572	0.19456	0.26634	0.35426	−0.37515
Race Hispanic	0.00000	0.11002	0.21645	0.14802	0.26330	0.04798	0.11289	0.08968	−0.16091
Race Asian	0.00000	0.28567	0.07579	0.28561	−0.00145	0.26855	0.20484	0.23064	−0.35049
Reop, 1 previous operation	1.13997	0.00000	0.41962	0.53987	0.00000	0.45372	0.47614	0.48534	−0.32375
Reop, $\geq$ 2 previous operations	1.43250	0.00000	0.45592	0.62211	0.00000	0.53695	0.61014	0.57945	−0.44745
Shock	0.82667	0.32434	0.50003	0.73290	0.00000	0.35800	0.74320	0.54575	−0.54475
Status urgent	0.14608	0.10671	0.11226	0.21738	0.18496	0.16500	0.16492	0.18202	−0.14608
Status emergent	1.04010	0.75216	0.51857	0.76090	0.62665	0.60549	0.56983	0.75083	−0.47745
Status salvage	2.07934	0.91950	0.76808	1.10085	0.73651	0.84873	1.29422	0.87072	−1.08265
Stenosis aortic	0.00000	0.00000	0.00000	0.16529	0.00000	0.00000	0.14706	0.13988	−0.14173
Unstable angina	0.11217	0.00000	0.10287	0.05060	0.00000	0.00000	0.00000	0.00000	0.00000

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

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

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preoperative atrial fibrillation, = 0 otherwise
Age	= Patient age in years
Age function 1	= max (age–50, 0)
Age function 2	= max (age–60, 0)
Age by reop function	= Age function 1 if surgery is a reoperation, = 0 otherwise
Age by status function	= Age function 1 if status is emergent or salvage, = 0 otherwise
BSA function 1	= max (1.4, min [2.6, BSA]) – 1.8
BSA function 2	= (BSA function 1)^2
CHF but not NYHA IV	= 1 if patient has CHF and is not NYHA class IV, = 0 otherwise
CHF and NYHA IV	= 1 if patient has CHF and is NYHA class IV, = 0 otherwise
CLD mild	= 1 if patient has mild chronic lung disease, = 0 otherwise
CLD moderate	= 1 if patient has moderate chronic lung disease, = 0 otherwise
CLD severe	= 1 if patient has severe chronic lung disease, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
Creatinine function 2	= max ([creatinine function 1] – 1.0, 0)
Creatinine function 3	= max ([creatinine function 1] – 1.5, 0)
CVD without prior CVA	= 1 if patient has history of CVD and no prior CVA, = 0 otherwise
CVD and prior CVA	= 1 if patient has history of CVD and a prior CVA, = 0 otherwise
Diabetes, noninsulin	= 1 if patient has diabetes not treated with insulin, = 0 otherwise
Diabetes, insulin	= 1 if patient has diabetes treated with insulin, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Ejection fraction function	= max (50 – ejection fraction, 0)
Female	= 1 if patient is female, = 0 otherwise
Female by BSA function 1	= BSA function 1 if female, = 0 otherwise
Female by BSA function 2	= BSA function 2 if female, = 0 otherwise
Hypertension	= 1 if patient has hypertension, = 0 otherwise
IABP or inotropes	= 1 if patient requires IABP or inotropes preoperatively, = 0 otherwise
Immunosuppressive treatment	= 1 if patient given immunosuppressive therapy within 30 days, = 0 otherwise
Insufficiency, aortic	= 1 if patient has at least moderate aortic insufficiency, = 0 otherwise
Insufficiency, mitral	= 1 if patient has at least moderate mitral insufficiency, = 0 otherwise
Insufficiency, tricuspid	= 1 if patient has at least moderate tricuspid insufficiency, = 0 otherwise
Left main disease	= 1 if patient has left main disease, = 0 otherwise
MI 1 to 21 days	= 1 if history of MI 1 to 21 days prior to surgery, = 0 otherwise
MI $>$ 6 and $<$ 24 hours	= 1 if history of MI $>$6 and $<$24 hours prior to surgery, = 0 otherwise
MI $\leq$ 6 hours	= 1 if history of MI $\leq$ 6 hours prior to surgery, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 0 otherwise
PCI $\leq$ 6 hours	= 1 if patient had PCI $\leq$ 6 hours prior to surgery, = 0 otherwise
Peripheral vascular disease	= 1 if patient has peripheral vascular disease, = 0 otherwise
Race black	= 1 if patient is black, = 0 otherwise
Race Hispanic	= 1 if patient is nonblack Hispanic, = 0 otherwise
Race Asian	= 1 if patient is nonblack, non-Hispanic, and is Asian, = 0 otherwise
Reop, 1 previous operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, $\geq$ 2 previous operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Shock	= 1 if patient was in shock at time of procedure, = 0 otherwise
Status urgent	= 1 if status is urgent, = 0 otherwise
Status emergent	= 1 if status is emergent (but not resuscitation), = 0 otherwise
Status salvage	= 1 if status is salvage (or emergent plus resuscitation), = 0 otherwise
Stenosis aortic	= 1 if patient has aortic stenosis, = 0 otherwise
Unstable angina	= 1 if patient has unstable angina, no MI within 7 days of surgery, = 0 otherwise

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