

2561 – STS Aortic Valve Replacement (AVR) Composite Score

Performance Gap - Opportunity for Improvement (Measure evaluation criterion 1b)

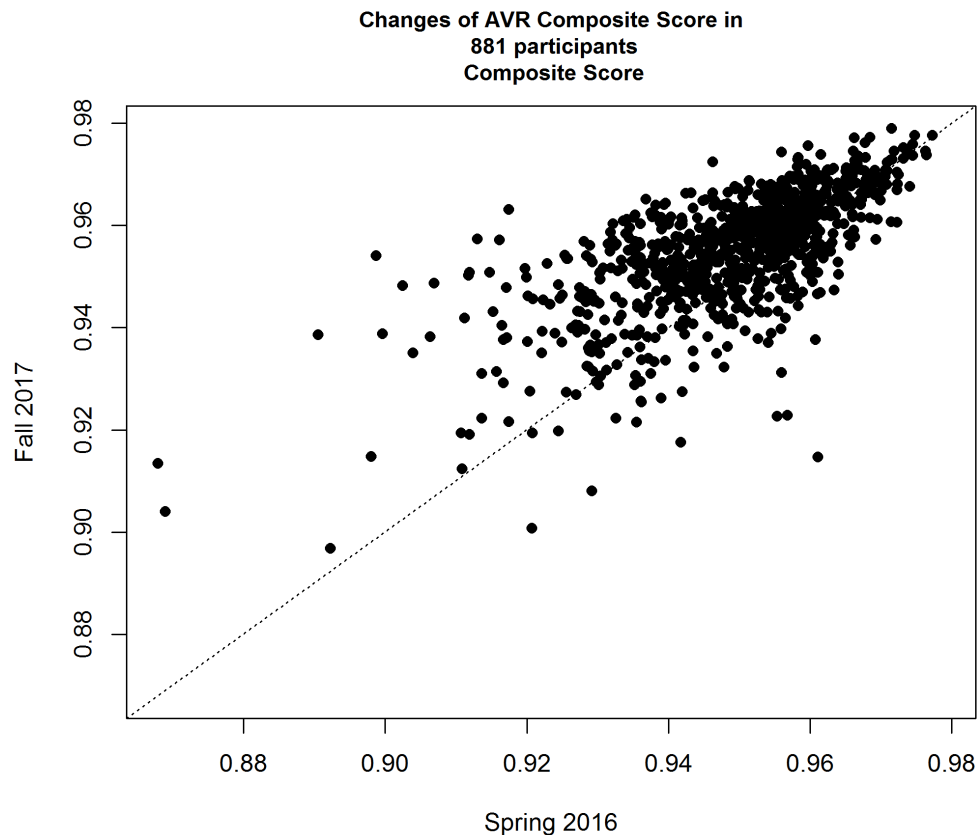
1b.2. Provide performance scores on the measure as specified (**current and over time**) at the specified level of analysis. (This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include). This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.

Distribution of STS isolated AVR composite measure in the latest four STS harvests for which the measure was reported

Stat	STS Harvests*			
	Latest	Spring 2017	Fall 2016	Spring 2016
# Participant	923	921	840	986
# Operations	81531	83549	73792	89392
Mean	0.955	0.953	0.953	0.948
STD	0.0117	0.013	0.0133	0.0148
IQR	0.0136	0.017	0.0164	0.0172
Percentiles				
0%	0.897	0.872	0.880	0.859
10%	0.940	0.936	0.936	0.930
20%	0.946	0.943	0.944	0.938
30%	0.951	0.948	0.948	0.943
40%	0.954	0.952	0.952	0.947
50%	0.956	0.955	0.955	0.951
60%	0.959	0.958	0.958	0.954
70%	0.961	0.961	0.961	0.957
80%	0.964	0.964	0.964	0.960
90%	0.968	0.967	0.968	0.964
100%	0.979	0.979	0.980	0.977
US Geographic Region				
Midwest	258	258	244	270
Northeast	130	128	114	137
Other	10	6	4	1
South	335	338	293	361
West	190	191	185	217

* Composite measure analysis of each harvest uses the most recent three year data until the end of last quarter. For example Spring 2013 harvest uses data until December 2012.

Changes of scores between measures calculated with data from Fall 2017 (July 2014 - June 2017) and Spring 2016 (Jan 2013 - Dec 2015)



The Spearman rank correlation of the measure between the two time periods is 0.67. The Pearson correlation is 0.68.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by region, race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. (This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., “topped out”, disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.

This composite measure gauges the performance of STS participant (typically a hospital, a hospital group, or a surgeon group). It is not a patient or operation level measure. Therefore we do not provide data stratified by patient characteristics. Instead, we provide results stratified by participant characteristics.

Distribution of isolated AVR composite measures by regions, Fall 2017 harvest, July 2014 - June 2017.

Stat	Midwest	Northeast	South	West	Other
# Participant	258	130	335	190	10
# Operations	20157	18238	26573	15672	891

Mean	0.955	0.959	0.953	0.954	0.952
STD	0.0106	0.0102	0.0127	0.0114	0.0113
IQR	0.0119	0.0124	0.0141	0.0158	0.0104
Percentiles					
0%	0.915	0.927	0.897	0.919	0.927
10%	0.941	0.947	0.939	0.939	0.942
20%	0.947	0.952	0.946	0.945	0.944
30%	0.951	0.956	0.949	0.950	0.950
40%	0.954	0.958	0.953	0.953	0.953
50%	0.957	0.961	0.955	0.956	0.954
60%	0.958	0.963	0.958	0.958	0.955
70%	0.961	0.965	0.960	0.961	0.956
80%	0.963	0.968	0.963	0.965	0.957
90%	0.966	0.971	0.967	0.967	0.959
100%	0.976	0.978	0.977	0.979	0.970

Distribution of isolated AVR composite measures by regions, Spring 2016 harvest, Jan 2013 - Dec 2015

Stat	Midwest	Northeast	South	West	Other
# Participant	270	137	361	217	1
# Operations	21608	19590	29614	18556	24
Mean	0.948	0.954	0.947	0.949	0.953
STD	0.0129	0.0122	0.0169	0.014	NA
IQR	0.0157	0.0158	0.0189	0.0185	NA
Percentiles					
0%	0.869	0.907	0.859	0.900	NA
10%	0.932	0.940	0.927	0.933	NA
20%	0.938	0.945	0.935	0.938	NA
30%	0.943	0.949	0.941	0.943	NA
40%	0.947	0.952	0.946	0.947	NA
50%	0.950	0.956	0.950	0.949	NA
60%	0.953	0.957	0.954	0.954	NA
70%	0.955	0.961	0.957	0.958	NA
80%	0.958	0.964	0.960	0.960	NA
90%	0.961	0.968	0.964	0.965	NA
100%	0.974	0.977	0.971	0.975	NA

S.4. – S.11. Measure Specifications

Due to the complex methodology used to construct the composite measure, it is impractical to separately discuss the numerator and denominator. The following discussion describes how each domain score is calculated and how these are combined into an overall composite score.

The STS AVR Composite Score comprises two domains consisting of six individual measures:

1. Absence of Operative Mortality
NQF # 0120 Risk-Adjusted Operative Mortality for AVR
2. Absence of Major Morbidity, scored any-or-none. The measures used are the same morbidity outcomes included in NQF #0696 STS CABG Composite Score.
Risk-Adjusted Postoperative Stroke/Cerebrovascular Accident
Risk-Adjusted Postoperative Surgical Re-exploration
Risk-Adjusted Postoperative Deep Sternal Wound Infection Rate
Risk-Adjusted Postoperative Renal Failure
Risk-Adjusted Postoperative Prolonged Intubation (Ventilation)

Participants receive a score for each of the two domains, plus an overall composite score. The overall composite score is created by “rolling up” the domain scores into a single number. In addition to receiving a numeric score, participants are assigned to rating categories designated by one star (below average performance), two stars (average performance), or three stars (above average performance).

Patient Population: The analysis population consists of adult patients aged 18 years or older who undergo isolated AVR surgery

Time Period: 3 years

Data Completeness Requirement: Participants are excluded from the analysis if they have fewer than 10 isolated AVR procedures in the patient population.

Technical Details

The unit of measurement for the STS AVR Composite Score can be either a participant (most often a cardiac surgical practice but occasionally an individual surgeon) or a hospital.

Domain	Numerator	Denominator
Absence of Operative Mortality	Number of patients undergoing isolated AVR who survived until after discharge and >30 days post-surgery	Number of patients undergoing isolated AVR during the measurement period
Absence of Major Morbidity	Number of patients undergoing isolated AVR who did not experience any of the five specified major morbidity endpoints ¹ .	Number of patients undergoing isolated AVR during the measurement period

1. Morbidity endpoints consist of postoperative stroke/cerebrovascular accident, surgical re-exploration, deep sternal wound infection, renal failure, prolonged intubation (ventilation). Patients with documented history of renal failure (i.e., dialysis or baseline serum creatinine of 4.0 or higher) are excluded when counting renal failure outcomes.

STS AVR risk models are used to estimate expected rates of mortality and any-or-none morbidity (Reference: O'Brien SM, Shahian DM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2—isolated valve surgery. *Ann Thorac Surg* 2009;88(1 Suppl):S23–42). To enhance interpretation, mortality rates are converted to survival rates (risk-standardized survival rate = 100 – risk-standardized mortality rate), and morbidity rates are converted to “absence of morbidity” rates (risk-standardized absence of morbidity rate = 100 – risk-standardized morbidity rate). Defining scores in this manner ensures that increasingly positive values reflect better performance, which is easier for consumers to interpret.

The overall composite score is calculated for each unit by using the following formula:

$$\text{composite}_j = \frac{1}{c} \left(\frac{\text{score}_{\text{mort},j}}{\text{sd}_{\text{mort}}} + \frac{\text{score}_{\text{morb},j}}{\text{sd}_{\text{morb}}} \right)$$

Where

$$c = \left(\frac{1}{\text{sd}_{\text{mort}}} + \frac{1}{\text{sd}_{\text{morb}}} \right)$$

where “score_{mort,j}” denotes the *j*-th participant’s estimated risk-standardized survival rate and “score_{morb,j}” denotes the *j*-th participant’s estimated risk-standardized absence of morbidity rate, and “sd_{mort}” and “sd_{morb}” denote the respective estimated standard deviations (SD) across participants. An equivalent formula is:

$$\text{composite}_j = \text{wt}_{\text{mort}} \text{score}_{\text{mort},j} + \text{wt}_{\text{morb}} \text{score}_{\text{morb},j}$$

where

$$\text{wt}_{\text{mort}} = (1/\text{sd}_{\text{mort}}) \times (1/c)$$

and

$$\text{wt}_{\text{morb}} = (1/\text{sd}_{\text{morb}}) \times (1/c).$$

Thus the method is equivalent to calculating a weighted average, with weights proportional to the inverse of the SD. In the most recent production of the STS AVR Composite Score based on data from July 2010 – June 2013, wt_{mort}=0.79 and wt_{morb} = 0.21.

Star Rating: Star ratings are derived by testing whether the participant's composite or domain score is significantly different from the overall STS average. For instance, if for each of the 2 composite score domains, a participant’s estimated score is lower than the overall STS average, but the difference between the participant and STS is not statistically significant, the ratings would each be 2 stars. If however, for the overall composite, the point estimate is lower than the STS average, AND this difference is statistically significant, the overall participant star rating is 1 star. The fact that statistical significance was achieved for the composite score but not the individual domains reflects the greater precision of the composite score compared to individual endpoints. This precision is achieved by aggregating information across multiple endpoints instead of a single endpoint.

Additional details regarding the AVR Composite Score are provided in the attached manuscript:

Shahian DM, He X, Jacobs JP, et al. The Society of Thoracic Surgeons Isolated Aortic Valve Replacement (AVR) Composite Score: a report of the STS Quality Measurement Task Force. *Ann Thorac Surg* 2012;94:2166-71.

SQL code to create function to identify procedures.txt

BEGIN

```
-- Start by identifying the cases where procedures were performed that definitively put the case into the
Other category. ProcID=null.
  if (VSTCV=1 or EndoProc=1 or OCarACDLE=1 or ResectSubA=1 or OCarCrTx=1 or OCarSVR=1 or CCancCase=1) or
(OCTumor<>1 and OCTumor is not null) or (OCPulThromDis<>1 and OCPulThromDis is not null) then
    Return null;
  else
    if (VADProc=2 and (UnplVAD=2 or UnplVAD is null)) or VADProc=3 or VADProc=4 then
      Return null;
    else
      if OCarASD=1 and (OCarASDTy=1 or OCarASDTy=2 or OCarASDTy is null) then
        Return null;
      else
        if OCarAFibSur=1 and OCarAFibAProc=2 then
          Return null;
        else
          if (OpTricus is not null and OpTricus<>1) or (OpPulm is not null and OpPulm<>1) then
            if UnplProc=1 or UnplProc=2 or UnplProc is null then
              Return null;
            else
              if UnplCABG=1 or UnplAV=1 or UnplMV=1 or UnplAo=1 or UnplVAD=1 then
                Return null;
              end if;
            end if;
          end if;
          if (UnplOth=2 or UnplOth is null) or UnplProc=2 then
            if OpONCard=1 or OCarLVA=1 or OCarVSD=1 or OCarTrma=1 or OCarOthr=1 then
              Return null;
            end if;
          end if;
          if (OCAoProcType is not null and OCAoProcType<>1) then
            if (UnplAo=2 or UnplAo is null) or (UnplAo=1 and UnplProc=2) then
              Return null;
            end if;
          end if;
        end if;
      end if;
    end if;
  end if;
end if;
```

SQL code to create function to identify procedures.txt

```
-- Now determine whether the procedure is an isolated CAB. ProcID=1.
if OpCAB=1 and (UnplCABG=2 or UnplCABG is null) then
    if OpValve=2 or OpValve is null then
        if (OCarCongProc1 is null or OCarCongProc1=10 or OCarCongProc1=1291 or OCarCongProc1=1305) and
            (OCarCongProc2 is null or OCarCongProc2=10 or OCarCongProc2=1291 or
OCarCongProc2=1305) and
            (OCarCongProc3 is null or OCarCongProc3=10 or OCarCongProc3=1291 or
OCarCongProc3=1305) then
            Return 1; -- Isolated CAB procedure.
        else
            Return null;
        end if;
    else
        -- OpValve can only be 1 at this point.
        if UnplProc=3 then
            If (VSAV=2 or VSAV is null) or (VSAV=1 and UnplAV=1) then
                if (VSMV=2 or VSMV is null) or (VSMV=1 and UnplMV=1) then
                    if (OCarCongProc1 is null or OCarCongProc1=10 or OCarCongProc1=1291 or
OCarCongProc1=1305) and
                        (OCarCongProc2 is null or OCarCongProc2=10 or OCarCongProc2=1291 or
OCarCongProc2=1305) and
                        (OCarCongProc3 is null or OCarCongProc3=10 or OCarCongProc3=1291 or
OCarCongProc3=1305) then
                        Return 1; -- Isolated CAB procedure.
                    else
                        Return null;
                    end if;
                end if;
            end if;
        end if;
    end if;
end if;

-- Procedure is not an isolated CABG, but could still be a valve or combination CAB + Valve procedure.

-- Determine whether the procedure is an isolated AVR or AVR + CAB. ProcID=2 or 4.
If OpValve=2 or OpValve is null then
    Return null; -- If procedure is not an isolated CAB and no valve procedures were done, it is an
Other procedure.
else
    if VSAV=1 and (VSAVPr=1 or VSAVPr=9) then
        if (VSMV=2 or VSMV is null) or (VSMV=1 and UnplProc=3 and UnplMV=1) then
            if (OpCAB=2 or OpCAB is null) or (OpCAB=1 and UnplProc=3 and UnplCABG=1) then
                if (OCarCongProc1 is null or OCarCongProc1=10) and (OCarCongProc2 is null or
OCarCongProc2=10) and (OCarCongProc3 is null or OCarCongProc3=10) then
                    Return 2; -- Isolated AVR procedure.
                else
```

SQL code to create function to identify procedures.txt

```

        Return null;
    end if;
else
    -- OpCAB can only be 1 at this point.
    If (Unpl Proc=3 and (Unpl CABG=2 or Unpl CABG is null)) or (Unpl Proc=1 or Unpl Proc=2 or
Unpl Proc is null) then
        if (OCarCongProc1 is null or OCarCongProc1=10 or OCarCongProc1=1291 or
OCarCongProc1=1305) and
            (OCarCongProc2 is null or OCarCongProc2=10 or OCarCongProc2=1291 or
OCarCongProc2=1305) and
            (OCarCongProc3 is null or OCarCongProc3=10 or OCarCongProc3=1291 or
OCarCongProc3=1305) then
            Return 4;    -- AVR + CAB procedure.
        else
            Return null;
        end if;
    end if;
end if;
end if;
end if;
end if;

-- Determine whether the procedure is an isolated MVR or MVR + CAB.  ProcID=3 or 5.
if VSMV=1 and (VSMVPr=2) then
    if (VSAV=2 or VSAV is null) or (VSAV=1 and Unpl Proc=3 and Unpl AV=1) then
        if (OpCAB=2 or OpCAB is null) or (OpCAB=1 and Unpl Proc=3 and Unpl CABG=1) then
            if (OCarCongProc1 is null or OCarCongProc1=10) and (OCarCongProc2 is null or
OCarCongProc2=10) and (OCarCongProc3 is null or OCarCongProc3=10) then
                Return 3;    -- Isolated MVR procedure.
            else
                Return null;
            end if;
        else
            -- OpCAB can only be 1 at this point.
            If (Unpl Proc=3 and (Unpl CABG=2 or Unpl CABG is null)) or (Unpl Proc=1 or Unpl Proc=2 or
Unpl Proc is null) then
                if (OCarCongProc1 is null or OCarCongProc1=10 or OCarCongProc1=1291 or
OCarCongProc1=1305) and
                    (OCarCongProc2 is null or OCarCongProc2=10 or OCarCongProc2=1291 or
OCarCongProc2=1305) and
                    (OCarCongProc3 is null or OCarCongProc3=10 or OCarCongProc3=1291 or
OCarCongProc3=1305) then
                        Return 5;    -- MVR + CAB procedure.
                    else
                        Return null;
                    end if;
                end if;
            end if;
        end if;
    end if;
end if;

```


SQL code to create function to identify procedures.txt

```

    end if;
end if;

-- Determine whether the procedure is an AVR + MVR.   ProcID=6.
if VSAV=1 and (VSAVPr=1 or VSAVPr=9) and VSMV=1 and VSMVPr=2 then
    if (OpCAB=2 or OpCAB is null) or (OpCAB=1 and UnplProc=3 and UnplCABG=1) then
        if (OCarCongProc1 is null or OCarCongProc1=10) and (OCarCongProc2 is null or OCarCongProc2=10)
and (OCarCongProc3 is null or OCarCongProc3=10) then
            Return 6;    -- AVR + MVR procedure.
        else
            Return null;
        end if;
    end if;
end if;

-- Determine whether the procedure is an MV Repair or MV Repair + CAB.   ProcID=7 or 8.
if VSMV=1 and VSMVPr=1 then
    if (VSAV=2 or VSAV is null) or (VSAV=1 and UnplProc=3 and UnplAV=1) then
        if (OpCAB=2 or OpCAB is null) or (OpCAB=1 and UnplProc=3 and UnplCABG=1) then
            if (OCarCongProc1 is null or OCarCongProc1=10) and (OCarCongProc2 is null or
OCarCongProc2=10) and (OCarCongProc3 is null or OCarCongProc3=10) then
                Return 7;    -- MV Repair procedure.
            else
                Return null;
            end if;
        else
            -- OpCAB can only be 1 at this point.
            if (UnplProc=3 and (UnplCABG=2 or UnplCABG is null)) or (UnplProc=1 or UnplProc=2 or
UnplProc is null) then
                if (OCarCongProc1 is null or OCarCongProc1=10 or OCarCongProc1=1291 or
OCarCongProc1=1305) and
                    (OCarCongProc2 is null or OCarCongProc2=10 or OCarCongProc2=1291 or
OCarCongProc2=1305) and
                    (OCarCongProc3 is null or OCarCongProc3=10 or OCarCongProc3=1291 or
OCarCongProc3=1305) then
                    Return 8;    -- MV Repair + CAB procedure.
                else
                    Return null;
                end if;
            end if;
        end if;
    end if;
end if;

-- If ProcID still has not been determined, then it is an Other procedure.   ProcID = null.
return null;

```

SQL code to create function to identify procedures.txt

```
EXCEPTION  
  WHEN NO_DATA_FOUND THEN  
    NULL;  
  WHEN OTHERS THEN  
    Null;  
    RAISE;  
END getProclD;  
/
```

The Society of Thoracic Surgeons Isolated Aortic Valve Replacement (AVR) Composite Score: A Report of the STS Quality Measurement Task Force

David M. Shahian, MD, Xia He, MS, Jeffrey P. Jacobs, MD, J. Scott Rankin, MD, Karl F. Welke, MD, Giovanni Filardo, PhD, MPH, Cynthia M. Shewan, PhD, and Sean M. O'Brien, PhD

Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts (DMS); Duke Clinical Research Institute, Durham, North Carolina (XH, SMO); Congenital Heart Institute of Florida, St. Petersburg, Florida (JPI); Centennial Medical Center, Vanderbilt University, Nashville, Tennessee (JSR); Mary Bridge Children's Hospital, Tacoma, Washington (KFW); the Institute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas (GF); and The Society of Thoracic Surgeons, Chicago, Illinois (CMS)

Background. Risk-standardized mortality rates provide a valuable but incomplete assessment of provider performance. Consequently, The Society of Thoracic Surgeons (STS) previously developed a multidimensional composite quality measure for coronary artery bypass grafting, the most frequently performed cardiac surgical procedure. The current study creates a similar composite measure for isolated aortic valve replacement (AVR).

Methods. Because there are few widely accepted process measures for AVR, the STS AVR composite score is based solely on outcomes, including risk-standardized mortality and any-or-none risk-standardized morbidity (occurrence of sternal infection, reoperation, stroke, renal failure, or prolonged ventilation). Isolated AVR is performed less frequently than coronary artery bypass grafting, and 1 year of data provided inadequate sample sizes for profiling. Therefore, we investigated observation periods of 3 years (July 1, 2007, to June 30, 2010: 67,138 records, 2,082 deaths, and 11,962 morbidity events) and 5

years (July 1, 2005, to June 30, 2010: 101,269 records, 3,123 deaths, and 17,514 morbidity events). We also compared results using 90%, 95%, and 98% credible intervals, corresponding to 95%, 97.5%, and 99% Bayesian probabilities, to determine "star ratings."

Results. Differences between 3-year and 5-year results were small; the former was chosen because this time frame provides more current and relevant data. Using 3 years of data and 95% credible intervals, adjusted mortality and morbidity rates varied threefold from highest performing (3 stars) to lowest performing (1 star) programs. Approximately 3% of participants were 1-star, 6% were 3-star, and 91% were 2-star programs.

Conclusions. STS has developed a composite mortality and morbidity outcomes measure for isolated AVR to be used in quality assessment, provider feedback, public reporting, and performance improvement.

(Ann Thorac Surg 2012;94:2166–71)

© 2012 by The Society of Thoracic Surgeons

Profiling of provider performance is a central feature of health care reform. Cardiac surgery has always been a leader in assessing and improving outcomes, and numerous cardiac surgical registries and risk models have been developed over the past 2 decades. Early initiatives focused on coronary artery bypass grafting (CABG), because this was the most common, signature cardiac surgical procedure [1]. Risk models for valve and combined valve and CABG operations have now also been developed [2, 3]. Based on these registry-derived data and analyses, nationally benchmarked results are provided to surgeons and hospitals, and CABG results have recently been incorporated into voluntary public reporting programs [4].

Risk-standardized mortality rates have historically been the dominant outcomes metrics for isolated CABG and isolated valve procedures, but they do have limitations. In

an era when the average mortality rates for these procedures have declined to very low levels, differentiating performance based on mortality alone is difficult. In addition, mortality rates fail to discriminate among patients who have survived their operations but who nevertheless have had different quality of care. For example, based on mortality alone, a patient who received all recommended care and had no postoperative complications would be categorized the same as another patient who sustained permanent dialysis-dependent renal failure.

For these reasons, STS has developed and implemented a composite CABG measure that encompasses four broad domains of care and 11 individual National Quality Forum-endorsed metrics [5, 6]. For each domain and overall, providers receive a numeric score with confidence intervals as well as a star rating, which for CABG is based on a 99% true Bayesian probability (corresponding to a 98% Bayesian credible interval [CrI] + 1% tail) that the provider's score differs from that of the average STS provider. Because of the increased number of end points constituting this composite, more providers are

Address correspondence to Dr Shahian, Department of Surgery and Center for Quality and Safety, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114; e-mail: dshahian@partners.org.

classified as high-performing or low-performing by this measure than would be identified using mortality rate alone. Over the past 4 years of implementation, approximately 10% to 15% of STS participants have been identified as high-performing (3-star) CABG programs, and about the same percentage have been classified as 1-star or low-performing. This is a substantially greater proportion of outlying programs than identified by profiling initiatives that base performance only on mortality rate.

The STS Quality Measurement Task Force now describes the development of a similar composite performance metric for isolated aortic valve replacement (AVR). This new composite is timely, because the proportion of isolated AVR has increased while that of CABG has correspondingly declined. By measuring performance for both procedures, the overall quality of a typical cardiac surgical practice will be more comprehensively assessed.

Methods

General

We used the same basic statistical approach as previously described for development of the STS composite CABG measure [5, 6], but with several modifications that reflect the differences between isolated CABG and isolated AVR populations. First, in aortic valve surgery, there is no widely accepted surgical analog to use of the internal mammary artery in CABG surgery, the latter having been shown to substantially improve short-term and long-term outcomes. We considered features such as optimal age ranges for bioprosthetic vs mechanical valves and the avoidance of patient–prosthesis mismatch. Although guidelines are being developed that will help to clarify these issues, specific cutoff values for both metrics remain controversial and have not as yet been widely implemented. Because of this current lack of consensus, these potential metrics do not meet the criteria for a nationally endorsed performance measure (eg, by the National Quality Forum), which should be based on the highest level of evidence.

Aortic valve surgery also has no clear analog for the perioperative medication domain included in the CABG composite. Several of the drugs in the CABG composite medication domain are primarily indicated for patients with coronary artery disease and are not necessarily applicable to isolated AVR patients. We considered the use of specific anticoagulants for bioprosthetic and mechanical valves, but this area is also undergoing evolution with the introduction of direct thrombin inhibitors. Antiarrhythmic therapy was also considered, but there is no clear consensus as yet regarding the use of agents such as amiodarone, despite their apparent efficacy and safety in randomized controlled trials [7, 8]. For all these reasons, and also because there is an increasing preference for outcomes measures to assess performance [9, 10], we limited the isolated AVR composite to risk-standardized mortality and any-or-none risk-standardized morbidity domains.

There were other differences in the development of the AVR composite measure compared with the CABG composite measure, the most important of which resulted

from the smaller number of patients. To address this, we explored the possibility of aggregating multiple years of data and potentially using different Bayesian CrIs. Regarding the latter, the selection of the ideal CrI (eg, 90%, 95%, or 98%, which correspond to 95%, 97.5%, and 99% Bayesian probabilities) for a particular cohort should not be made in a routine, “one size fits all” manner because the most appropriate probability criterion is dependent on many factors, including sample size and the potential implications of misclassification.

Inclusion and Exclusion Criteria

The analysis population included all patients undergoing isolated AVR for whom an STS AVR risk score could be calculated (ie, nonmissing age and gender data, which STS requires and are never imputed). We performed identical analyses for two potential measurement periods: 3 years and 5 years. In the rare event that a patient was discharged and readmitted for AVR within 30 days of a previous AVR operation, only the first operation was included in the analysis of mortality. This was done to prevent potentially counting the same patient’s death twice. Only 33 records were excluded for this reason in the 3-year analysis and 58 in the 5-year analysis. This exclusion was only applied to the mortality component; all records were included when analyzing morbidity.

Estimation of Risk-Standardized Outcome Measures

The AVR composite score is a combination of two individual measures: (1) risk-standardized mortality and (2) risk-standardized any-or-none morbidity (ie, the patient experienced any of the following: renal failure, stroke, cardiac reoperation, sternal infection or mediastinitis, or prolonged ventilation). Methods of estimating these quantities were similar to the published STS CABG composite and are described in detail in the Technical Appendix of our prior publication [5]. Briefly, risk-standardized mortality and morbidity were jointly estimated in a Bayesian multivariate hierarchical logistic regression model using aggregated data with STS participant as the unit of analysis. To adjust for case mix, we first created summary measures of case mix for each participant by calculating their expected rates of mortality and any-or-none morbidity according to the published STS AVR model [2]. These expected rates were then entered as covariates in our subsequent Bayesian model to estimate participant-specific risk-standardized rates of mortality and any-or-none morbidity.

Calculation of Composite Scores and Star Ratings

Our approach to calculate the AVR composite score was identical to that used for the published STS CABG composite [5], except that the STS AVR models rather than the STS CABG models were used to estimate expected rates of mortality and any-or-none morbidity, and no process measures were included. To enhance interpretation, mortality rates were converted to survival rates (risk-standardized survival rate = 100 – risk-standardized mortality rate), and morbidity rates were converted to “absence of morbidity” rates (risk-standardized absence of morbidity rate = 100 – risk-

Table 1. Number of Participants, Records, and Events

Variable	3-Year Sample	5-Year Sample
Dates	July 1, 2007–June 30, 2010	July 1, 2005–June 30, 2010
Number of		
Participants	899	942
Records	67,138	101,269
Mortality events	2,082	3,123
Morbidity events ^a	11,962	17,514

^a Number of patients experiencing at least one of the morbidities in the any-or-none morbidity domain.

standardized morbidity rate). Defining scores in this manner ensures that increasingly positive values reflect better performance. Thus, a provider's score for the mortality domain was defined as the provider's risk-standardized survival rate. Similarly, the provider's score for the morbidity domain was defined as the risk-standardized absence of morbidity rate.

The overall composite score for the j -th STS participant was calculated as:

$$\text{composite}_j = \frac{1}{c} \left(\frac{\text{score}_{\text{mort},j}}{\text{sd}_{\text{mort}}} + \frac{\text{score}_{\text{morb},j}}{\text{sd}_{\text{morb}}} \right)$$

where

$$c = \left(\frac{1}{\text{sd}_{\text{mort}}} + \frac{1}{\text{sd}_{\text{morb}}} \right)$$

where “score_{mort,j}” denotes the j -th participant's estimated risk-standardized survival rate and “score_{morb,j}”

denotes the j -th participant's estimated risk-standardized absence of morbidity rate, and “sd_{mort}” and “sd_{morb}” denote the respective estimated standard deviations (SD) across participants. An equivalent formula is

$$\text{composite}_j = wt_{\text{mort}} \text{score}_{\text{mort},j} + wt_{\text{morb}} \text{score}_{\text{morb},j}$$

where

$$wt_{\text{mort}} = (1/\text{sd}_{\text{mort}}) \times (1/c)$$

and

$$wt_{\text{morb}} = (1/\text{sd}_{\text{morb}}) \times (1/c).$$

Thus, the method is equivalent to calculating a weighted average, with weights proportional to the inverse of the SD.

The current CABG composite star rating method classifies STS participants as 1 star (below average performance), 2 stars (average performance), and 3 stars (above average performance). Assignment of 1 star or 3 stars requires 99% Bayesian probability that a provider's score differs from the STS average for their particular case mix. It achieves this by assigning 3 stars if the entire 98% Bayesian CrI plus the 1% upper tail fall entirely above the STS average score, 1 star if the 98% Bayesian CrI and lower 1% tail fall entirely below the STS average score, and 2 stars otherwise. For different choices of time periods and credible intervals, we calculated the observed mortality and any-or-none morbidity rates (number of events/number of operations) and risk-standardized rates using standardized event ratio methods (risk-standardized rate = population average rate \times observed rate/expected rate) in each of the star rating groups. We also assessed how the number of outliers for the isolated AVR composite star rating would

Table 2. Mortality and Morbidity Rates by Composite Star Category

Variable	Mortality, %			Morbidity, %		
	1 Star	2 Stars	3 Stars	1 Star	2 Stars	3 Stars
Observed results						
3 years						
98% CrI	5.3	3.3	1.5	31.6	18.3	11.0
95% CrI	5.2	3.3	1.7	32.2	18.5	11.0
90% CrI	5.2	3.3	1.9	30.4	18.6	11.4
5 years						
98% CrI	4.9	3.2	1.6	28.2	17.7	11.4
95% CrI	4.9	3.3	1.7	28.3	17.9	11.5
90% CrI	5.0	3.3	1.8	28.5	18.1	11.8
Risk-standardized results						
3 years						
98% CrI	5.4	3.3	1.3	31.5	18.3	10.8
95% CrI	5.3	3.3	1.6	31.7	18.5	10.9
90% CrI	5.2	3.3	1.7	29.9	18.6	11.2
5-years						
98% CrI	4.9	3.3	1.5	27.5	17.8	11.1
95% CrI	4.9	3.3	1.6	27.8	17.9	11.4
90% CrI	5.0	3.4	1.8	27.5	18.2	11.7

CrI = credible interval.

differ if using 90%, 95%, or 98% CrI (95%, 97.5%, and 99% Bayesian probabilities).

Reliability of the AVR composite score, defined in this study as the squared correlation between the calculated AVR composite score and the “true” score, was estimated using Markov Chain Monte Carlo simulations, as described in the [Appendix](#).

Results

Table 1 lists the number of STS participants, patient records, and mortality and morbidity events in the 3- and 5-year samples.

Table 2 summarizes the observed and risk-standardized mortality and any-or-none risk-standardized morbidity rates for 1-star, 2-star, and 3-star programs. The data are presented using 3 and 5 year study cohorts and using 90%, 95%, and 98% Bayesian CrIs to determine star rating category. Notably, the directionality was consistent and appropriate in all instances. Lower overall performance (1 star) was associated with higher risk-standardized mortality rate and any-or-none morbidity, and higher overall performance (3 stars) was associated with lower mortality and morbidity rates. Additional analyses (available upon request) showed that programs with higher mortality rates also tended to have higher morbidity rates, and programs with lower mortality rates generally had lower morbidity rates. For risk-standardized mortality and morbidity, there was a roughly threefold increase in incidence moving from 3-star to 1-star categories. Differences between results using 3 and 5 years of data were relatively small, as were the differences resulting from using 90%, 95%, and 98% Bayesian CrIs. Separate analyses, available upon request, show that in most instances, the additional 2 years of data simply affirmed the performance classification derived from 3 years of data. However, in a small percentage of cases, the additional data resulted in reclassification of an outlier to the average (2-star) category, or visa versa.

Table 3 uses 95% and 98% CrI and 3 or 5 years of data to summarize the number of STS participants in each star rating category.

Table 4 reports the SDs of the mortality and morbidity domain scores and the relative weights resulting from standardization. There is very little difference between

Table 4. Estimated Composite Score Weights

Variable	3-Year Sample		5-Year Sample	
	Mortality	Morbidity	Mortality	Morbidity
Standard deviation	1.3	5.5	1.2	5.1
Weight	0.81	0.19	0.81	0.19

the results for 3- and 5-year samples. As a result of this weighting, a 1-percentage point difference in the mortality rate has the same effect as a 4-percentage point difference in the morbidity rate.

Based on the data in Tables 2, 3, and 4, we elected to proceed with rolling 3-year data to determine scores. This decision reflected several considerations. First, although 5 years of data did produce some changes in outlier status, these were not dramatic. We balanced these relatively modest changes in performance classification against the timeliness of the data. Specifically, we were concerned that 5 years of data were not sufficiently current. Older data might reflect staffing or practices that were no longer relevant when the most recent reports were provided back to participants or publicly reported. Accordingly, we elected to base the STS AVR composite measure on a sample derived from a rolling 3 years of data.

We selected 95% Bayesian CrIs (corresponding to 97.5% Bayesian probability) to assign star ratings because

Table 3. Number of Participants by Star Rating Category

Variable	3-Year Sample	5-Year Sample
	No. (%)	No. (%)
98% credible interval		
1 star	14 (1.6)	23 (2.4)
2 stars	851 (94.7)	870 (92.4)
3 stars	34 (3.8)	49 (5.2)
95% credible interval		
1 star	26 (2.9)	41 (4.4)
2 stars	818 (91.0)	819 (86.9)
3 stars	55 (6.1)	82 (8.7)

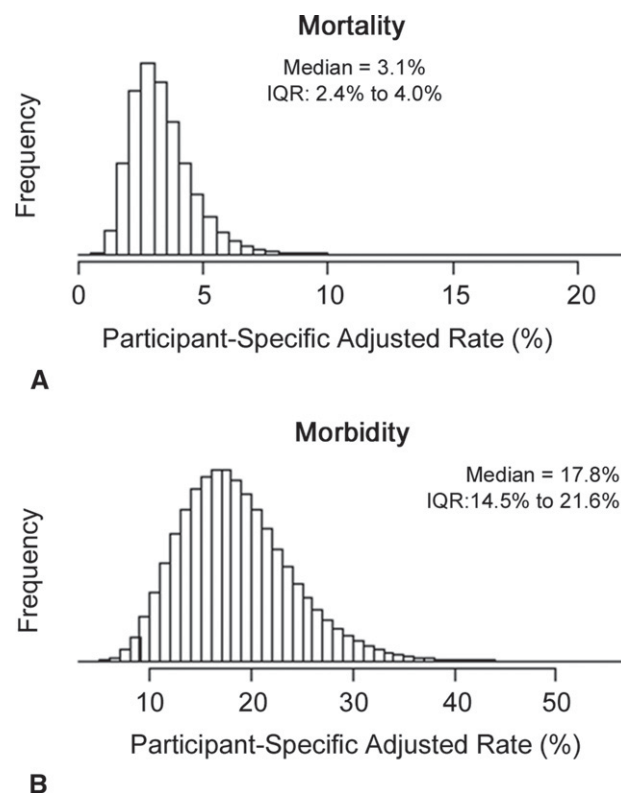


Fig 1. Estimated distribution of true risk-standardized (A) mortality and (B) morbidity, based on 3 years of data. (IQR = interquartile range.)

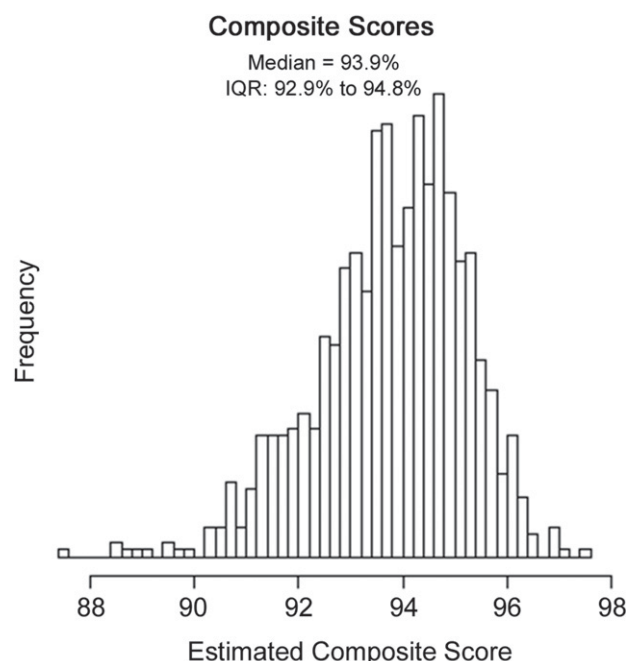


Fig 2. Distribution of estimated aortic valve replacement composite scores, based on 3 years of data. (IQR = interquartile range.)

they produced nearly twice the number of outliers compared with 98% CrIs (99% Bayesian probability). Although 90% CrIs (95% Bayesian probability) produced even more outliers (results available upon request), this was thought to provide inadequate certainty regarding outlier status. Reliability (95% CrI) was 0.50 (0.46, 0.55) using 3 years of data and 0.54 (0.50, 0.59) using 5 years of data (Appendix).

Figure 1 depicts the Bayesian-estimated distribution of risk-standardized mortality and morbidity rates among STS participants. Figure 2 presents the distribution of calculated composite scores (point estimates) based on 3 years of data.

Comment

We have described the development and operating characteristics of the STS isolated AVR composite measure. It is fundamentally similar to our previously described, National Quality Forum-endorsed STS CABG composite measure, with several notable exceptions. First, because we could not identify AVR analogs for the surgical and medical process measures previously used in our CABG composite, and because there is a growing shift towards the use of outcomes measures, the AVR composite consists of only two domains—risk-standardized mortality and any-or-none risk-standardized morbidity.

Second, because of the smaller number of AVR patients compared with CABG, we decreased the Bayesian CrI for star rating determination from 98% to 95% and increased the data collection period to 3 years (rolling) as opposed to 1 year for CABG. These adjustments assured adequate sample size and the ability to discriminate a reasonable number of providers as high-performing or

low-performing. However, even with these modifications, the number of high and low outliers is still less than identified with the STS CABG composite, both because of sample sizes and because there are fewer individual measures within the composite.

Results for the AVR composite measure will be provided to STS Adult Cardiac Surgery Database participants beginning in 2012, and voluntary public reporting of this measure will begin in early 2013.

References

- Shahian DM, O'Brien SM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. *Ann Thorac Surg* 2009;88(1 Suppl):S2–22.
- O'Brien SM, Shahian DM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2—isolated valve surgery. *Ann Thorac Surg* 2009;88(1 Suppl):S23–42.
- Shahian DM, O'Brien SM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3—valve plus coronary artery bypass grafting surgery. *Ann Thorac Surg* 2009;88(1 Suppl):S43–62.
- Ferris TG, Torchiana DF. Public release of clinical outcomes data—online CABG report cards. *N Engl J Med* 2010;363:1593–5.
- O'Brien SM, Shahian DM, DeLong ER, et al. Quality measurement in adult cardiac surgery: part 2—Statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg* 2007;83(4 Suppl):S13–26.
- Shahian DM, Edwards FH, Ferraris VA, et al. Quality measurement in adult cardiac surgery: part 1—conceptual framework and measure selection. *Ann Thorac Surg* 2007;83(4 Suppl):S3–12.
- Giri S, White CM, Dunn AB, et al. Oral amiodarone for prevention of atrial fibrillation after open heart surgery, the Atrial Fibrillation Suppression Trial (AFIST): a randomised placebo-controlled trial. *Lancet* 2001;357:830–6.
- Mitchell LB, Exner DV, Wyse DG, et al. Prophylactic Oral Amiodarone for the Prevention of Arrhythmias that Begin Early After Revascularization, Valve Replacement, or Repair: PAPABEAR: a randomized controlled trial. *JAMA* 2005;294:3093–100.
- National Quality Forum Report. http://www.qualityforum.org/Masuring_Performance/Improving_NQF_Process/Evidence_Task_Force_Final_Report.aspx. Accessed: July 13, 2011.
- Porter ME. What is value in health care? *N Engl J Med* 2010;363:2477–81.

Appendix: Statistical Reliability of the STS AVR Composite Score Definition of Reliability

Reliability is conventionally defined as the proportion of variation in a performance measure that is due to true between-hospital differences (ie, signal) as opposed to random statistical fluctuations (ie, noise). A mathematically equivalent definition is the squared correlation between a measurement and the true value. This quantity cannot be calculated directly because the “true” composite measure values are unknown, but may be estimated, as described below.

Technical Details

Let θ_j denote the true unknown composite measure value for the j -th of J hospitals. Before estimating reliability, the numeric value of θ_j was estimated for each hospital under the assumed hierar-

chical model. Estimation was done using Markov Chain Monte Carlo (MCMC) simulations and involved the following steps:

1. For each j , we randomly generated a large number (N) of possible numeric values of θ_j by sampling from the Bayesian posterior probability distribution of θ_j . Let $\theta_j^{(i)}$ denote the i -th of these N randomly sampled numeric values for the j -th hospital.
2. For each j , a Bayesian estimate $\hat{\theta}_j$ of θ_j was calculated as the arithmetic average of the randomly sampled values $\theta_j^{(1)}, \dots, \theta_j^{(N)}$; in other words $\hat{\theta}_j = 1/N \sum_{i=1}^N \theta_j^{(i)}$.

Our reliability measure was defined as the estimated squared correlation between the set of hospital-specific estimates $\hat{\theta}_1, \dots, \hat{\theta}_J$ and the corresponding unknown true values $\theta_1, \dots, \theta_J$. Let ρ^2 denote the *unknown true* squared correlation of interest and let $\hat{\rho}^2$ denote an *estimate* of this quantity. The estimate was calculated as

$$\hat{\rho}^2 = \frac{1}{N} \sum_{i=1}^N \rho_{(i)}^2$$

where

$$\rho_{(i)}^2 = \frac{\left[\sum_{j=1}^J (\theta_j^{(i)} - \bar{\theta}^{(i)}) (\hat{\theta}_j - \bar{\theta}) \right]^2}{\sum_{j=1}^J (\theta_j^{(i)} - \bar{\theta}^{(i)})^2 \sum_{j=1}^J (\hat{\theta}_j - \bar{\theta})^2} \bar{\theta} = \frac{1}{JN} \sum_{j=1}^J \sum_{i=1}^N \theta_j^{(i)}$$

and

$$\bar{\theta}^{(i)} = \frac{1}{J} \sum_{j=1}^J \theta_j^{(i)}.$$

A 95% Bayesian probability interval for ρ^2 was obtained calculating the 2.5th and 97.5th percentiles of the set of numbers $\rho_{(1)}^2, \dots, \rho_{(N)}^2$.

Results

Using the above method, we obtained the estimates:

Time Span	No. of Hospitals Included	No. of Patients Included	Reliability $\hat{\rho}^2$ (95% PrI)
3 years	899	67,138	0.50 (0.46, 0.55)
5 years	942	101,269	0.54 (0.50, 0.59)

PrI = probability interval.

When 3 years of data were used, the Bayesian estimate of ρ^2 was 0.50 and the 95% Bayesian probability interval for ρ^2 was (0.46, 0.55). Because the lower limit is 0.46, we may be highly confident (probability = 97.5%) that the true reliability ρ^2 is at least 0.46.

For a comparison with the existing measure, we applied the same method to the isolated CABG composite measure during the same time period. We used 1 year of data to be consistent with the current CABG composite measure definition (number of participants = 919, number of patients = ~158,000). The results were $\hat{\rho}^2 = 0.79$ (95% probability interval, 0.76, 0.82). Not surprisingly, the isolated CABG measure is more reliable because (1) there are more CABG patients; (2) there are more individual component measures in the CABG composite; and (3) the procedure components have greater true variations.

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

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

^aDuke Clinical Research Institute, Durham, North Carolina; ^bMassachusetts General Hospital, Boston, Massachusetts; ^cInstitute for Health Care Research and Improvement, Baylor Health Care System, Dallas, Texas; ^dUniversity of Kentucky Chandler Medical Center, Division of Cardiovascular and Thoracic Surgery, Lexington, Kentucky; ^eUniversity of Florida, Division of Cardiothoracic Surgery, Jacksonville, Florida; ^fSentara Cardiovascular Research Institute, Norfolk, Virginia; ^gDepartment of Health Care Policy, Harvard Medical School, and Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; ^hThe Society of Thoracic Surgeons, Chicago, Illinois; and ⁱSeattle, Washington

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

Methods. Using STS data from 2002 to 2006, isolated valve surgery risk models were developed for operative mortality, permanent stroke, renal failure, prolonged ventilation (> 24 hours), deep sternal wound infection, reoperation for any reason, a major morbidity or mortality composite endpoint, prolonged postoperative length of stay, and short postoperative length of stay. The study population consisted of adult patients who underwent one of three types of valve surgery: isolated aortic valve replacement (n = 67,292), isolated mitral valve replacement (n = 21,229), or isolated mitral valve repair (n = 21,238). The

population was divided into a 60% development sample and a 40% validation sample. After an initial empirical investigation, the three surgery groups were combined into a single logistic regression model with numerous interactions to allow the covariate effects to differ across these groups. Variables were selected based on a combination of automated stepwise selection and expert panel review.

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

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

(Ann Thorac Surg 2009;88:S23–42)

© 2009 by The Society of Thoracic Surgeons

Models for predicting surgical outcomes on the basis of patient preoperative characteristics are valuable tools for research, quality improvement, and clinical prac-

tice. Such models are used by The Society of Thoracic Surgeons (STS) to produce risk-adjusted performance re-

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

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

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

Abbreviations and Acronyms

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

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

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

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

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

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

Study Population and Endpoints

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

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

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

Training and Validation Samples

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

Endpoints

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

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

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

Table 1. Continued

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

Table 1. Continued

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

Table 1. Continued

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

Table 1. Continued

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

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

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

dead); and (9) short postoperative LOS (SLOS): LOS less than 6 days and patient alive at discharge.

Table 2 summarizes the endpoint frequencies in the study population.

Single Versus Multiple Models

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

Because of the large size of the STS NCD, separate

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

	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
All isolated valve (AVR, MVR, MVRepair)									
N	109,759	109,759	107,060	109,759	109,759	109,759	109,759	109,759	109,759
Events	3,706	1,751	4,673	12,892	307	9,164	20,074	9,718	41,214
%	3.4	1.6	4.3	11.8	0.3	8.4	18.3	8.9	37.6
AVR									
N	67,292	67,292	65,828	67,292	67,292	67,292	67,292	67,292	67,292
Events	2,157	1,007	2,774	7,323	197	5,369	11,706	5,308	26,144
%	3.2	1.5	4.1	10.9	0.3	8.0	17.4	7.9	38.9
MVR									
N	21,229	21,229	20,329	21,229	21,229	21,229	21,229	21,229	21,229
Events	1,210	447	1,348	4,015	71	2,450	5,675	3,244	4,727
%	5.7	2.1	6.4	18.9	0.3	11.5	26.7	15.3	22.3
MVRepair									
N	21,238	21,238	20,903	21,238	21,238	21,238	21,238	21,238	21,238
Events	339	297	551	1,554	39	1,345	2,693	1,166	10,343
%	1.6	1.4	2.6	7.3	0.2	6.3	12.7	5.5	48.7

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

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

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

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

Selection of Candidate Predictor Variables

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

Final candidate explanatory variables and their coding are summarized in Table 3. The variables were identical to the CABG model candidate variables with the following differences: (1) percutaneous coronary intervention conducted within 6 hours or less of surgery was not a candidate variable because it was present in only 122 patients (0.1%) in the valve model population; (2) infec-

tious endocarditis was included. This risk factor was rarely present among isolated CABG patients (0.09%), but was not uncommon (7.7%) among patients undergoing valve surgery; (3) mitral stenosis was included; this risk factor was rarely present among isolated CABG patients (0.35%) but was common (7.4%) among patients undergoing valve surgery; and (4) an indicator for surgery type (AVR, MVR, MVRrepair) was included in the valve models.

Coding of Explanatory Variables

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

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

Repair Versus Replacement

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

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

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

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

^a These variables were forced into each model.

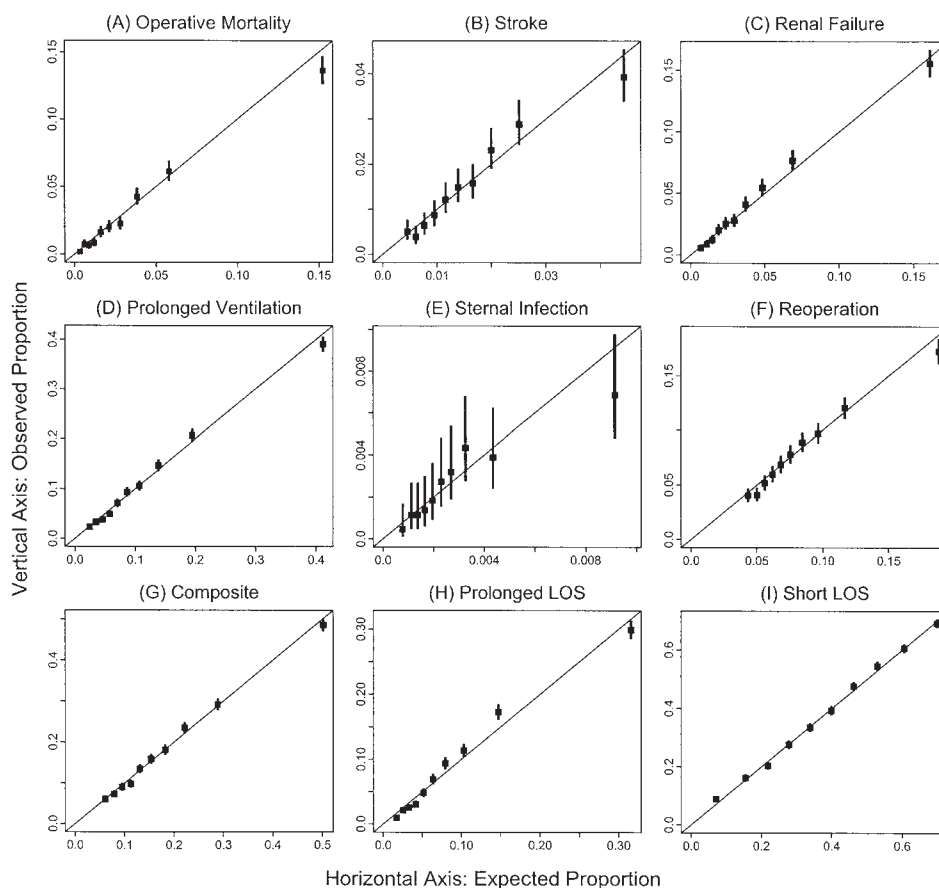
AVR = aortic valve replacement; CHF = congestive heart failure; CLD = chronic lung disease; CVA = cerebrovascular accident (stroke); CVD = cardiovascular disease; EF = ejection fraction; IABP = intra-aortic balloon pump; MI = myocardial infarction; MVR = mitral valve replacement; MVRRepair = mitral valve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

also dependent upon the patient's preoperative valve disease etiology, anatomy, and pathophysiology. On average, patients amenable to valve repair have less extensive valve pathology and a relatively favorable postoperative prognosis (the mortality rate for valve repair is 1.6%

compared with 5.7% for replacement). Ignoring these anatomical differences can introduce bias when comparing institutions, especially because these variables are not captured elsewhere on the STS data collection form.

A related difficulty in adjusting for repair versus re-

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



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

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

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

Missing Data

Model variables with more than 1% missing data in the study sample were ejection fraction (8.9%), NYHA class

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

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

Table 4. Discrimination of Models in Development and Validation Samples

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

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

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

Final Variable Selection Procedure

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

Interaction Terms

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

Four additional sets of interactions were also included in the models: (1) sex by body surface area (BSA); (2) sex by BSA²; (3) age by reoperation; and (4) age by emergent status. These interaction terms were preselected and were

not tested as part of the backward selection algorithm. Additional technical details are provided in the Appendix. For reasons described in Part 1 of this series (isolated CABG risk models), an extensive automated search for additional interaction terms was not conducted.

Adjustment for Time Trends

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

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

Results

Assessment of Model Fit and Discrimination

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

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

A. Odds ratios for variables that do not interact with surgery group									
Variable	Mort	CVA	RF	Vent	DSWI	Reop	Comp	PLOS	SLOS
Preoperative AFib	1.20 (1.10, 1.31)	1.06 (0.93, 1.20)	NA	1.18 (1.11, 1.25)	NA	1.11 (1.04, 1.18)	1.12 (1.07, 1.18)	1.17 (1.10, 1.24)	0.74 (0.70, 0.78)
BSA 1.6 versus 2.0 among females	1.19 (1.09, 1.30)	1.18 (1.03, 1.35)	0.95 (0.87, 1.04)	1.15 (1.08, 1.22)	0.42 (0.27, 0.68)	1.26 (1.18, 1.34)	1.17 (1.12, 1.23)	1.11 (1.04, 1.17)	0.99 (0.95, 1.04)
BSA 1.6 versus 2.0 among males	1.75 (1.48, 2.07)	1.17 (0.92, 1.47)	1.33 (1.12, 1.58)	1.56 (1.41, 1.74)	0.94 (0.49, 1.84)	1.34 (1.21, 1.49)	1.44 (1.33, 1.57)	1.39 (1.25, 1.56)	0.73 (0.68, 0.79)
BSA 1.8 versus 2.0 among females	0.99 (0.95, 1.04)	1.08 (0.99, 1.17)	0.90 (0.86, 0.94)	1.00 (0.97, 1.03)	0.65 (0.54, 0.77)	1.07 (1.03, 1.11)	1.02 (0.99, 1.04)	0.99 (0.96, 1.02)	1.05 (1.03, 1.08)
BSA 1.8 versus 2.0 among males	1.21 (1.14, 1.29)	1.07 (0.98, 1.16)	1.07 (1.00, 1.14)	1.14 (1.10, 1.19)	0.90 (0.70, 1.14)	1.12 (1.08, 1.16)	1.12 (1.09, 1.16)	1.10 (1.06, 1.15)	0.92 (0.89, 0.94)
BSA 2.2 versus 2.0 among females	1.21 (1.11, 1.33)	0.94 (0.80, 1.10)	1.30 (1.21, 1.41)	1.15 (1.09, 1.21)	1.57 (1.26, 1.96)	1.02 (0.95, 1.09)	1.12 (1.07, 1.16)	1.14 (1.08, 1.21)	0.85 (0.81, 0.88)
BSA 2.2 versus 2.0 among males	0.98 (0.93, 1.03)	0.95 (0.88, 1.03)	1.09 (1.05, 1.14)	1.05 (1.02, 1.08)	1.32 (1.17, 1.48)	0.95 (0.93, 0.98)	1.02 (0.99, 1.04)	1.03 (1.00, 1.07)	0.94 (0.93, 0.96)
Creatinine per 1 unit	1.55 (1.46, 1.64)	1.34 (1.22, 1.47)	2.04 (1.93, 2.16)	1.58 (1.51, 1.65)	NA	1.27 (1.20, 1.33)	1.64 (1.57, 1.71)	1.58 (1.51, 1.65)	0.64 (0.61, 0.68)
CVD with CVA	NA	1.81 (1.56, 2.10)	1.22 (1.09, 1.37)	1.28 (1.18, 1.38)	NA	1.14 (1.05, 1.24)	1.20 (1.12, 1.28)	1.40 (1.29, 1.52)	0.77 (0.72, 0.83)
CVD without CVA	NA	1.32 (1.11, 1.57)	1.23 (1.10, 1.37)	1.14 (1.05, 1.23)	NA	1.06 (0.96, 1.17)	1.08 (1.01, 1.15)	NA	0.80 (0.73, 0.88)
No. diseased coronary vessels (2 versus 1 or 3 versus 2)	NA	1.10 (1.01, 1.20)	NA	1.07 (1.02, 1.11)	NA	NA	1.04 (1.00, 1.08)	1.03 (0.98, 1.08)	0.90 (0.86, 0.94)
EF per 10-unit decrease	1.09 (1.05, 1.14)	NA	1.04 (1.00, 1.09)	1.12 (1.09, 1.15)	1.26 (1.12, 1.41)	1.08 (1.04, 1.11)	1.10 (1.07, 1.12)	1.12 (1.08, 1.15)	0.87 (0.85, 0.90)
Hypertension	1.12 (1.03, 1.22)	1.19 (1.07, 1.33)	1.35 (1.25, 1.45)	1.11 (1.06, 1.17)	NA	NA	1.11 (1.07, 1.15)	NA	0.94 (0.91, 0.97)
Immunosuppressive treatment	1.42 (1.21, 1.67)	NA	1.39 (1.19, 1.62)	NA	NA	NA	1.16 (1.06, 1.27)	1.31 (1.17, 1.47)	NA
Left main disease	1.19 (0.98, 1.46)	NA	1.19 (0.98, 1.44)	NA	2.17 (1.13, 4.16)	NA	NA	NA	NA
Active infectious endocarditis	1.95 (1.68, 2.27)	1.87 (1.52, 2.29)	2.17 (1.88, 2.50)	2.15 (1.95, 2.36)	NA	1.55 (1.39, 1.73)	1.97 (1.80, 2.15)	2.79 (2.51, 3.09)	0.34 (0.30, 0.38)
Mitral insufficiency, moderate/severe	NA	1.26 (1.14, 1.39)	NA	NA	NA	NA	NA	NA	NA
Tricuspid insufficiency, moderate/severe	NA	NA	1.14 (1.01, 1.29)	1.14 (1.04, 1.25)	NA	1.09 (1.00, 1.20)	1.21 (1.12, 1.30)	1.17 (1.05, 1.31)	0.82 (0.73, 0.92)
Peripheral vascular disease	1.25 (1.12, 1.38)	1.29 (1.11, 1.49)	NA	NA	NA	1.22 (1.12, 1.32)	1.14 (1.07, 1.21)	1.17 (1.09, 1.25)	0.83 (0.78, 0.88)
Aortic stenosis		NA	NA	0.90 (0.83, 0.97)	NA	0.90 (0.84, 0.96)	0.93 (0.87, 0.98)	0.86 (0.80, 0.92)	1.07 (1.02, 1.13)
Mitral stenosis	1.24 (1.08, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA
MI \leq 21 days	1.14 (0.98, 1.34)	NA	NA	1.37 (1.22, 1.55)	NA	1.04 (0.91, 1.18)	1.28 (1.16, 1.41)	1.21 (1.06, 1.37)	0.81 (0.72, 0.91)
Time trend, per 6-month harvest interval	0.98 (0.97, 0.99)	0.98 (0.96, 1.00)	1.01 (0.99, 1.02)	1.02 (1.01, 1.03)	0.97 (0.93, 1.01)	1.00 (0.99, 1.01)	1.01 (1.00, 1.02)	1.00 (0.99, 1.01)	1.00 (0.99, 1.01)
Race black	NA	1.33 (1.13, 1.57)	1.51 (1.34, 1.69)	1.42 (1.27, 1.58)	NA	1.27 (1.15, 1.40)	1.37 (1.27, 1.49)	1.45 (1.31, 1.60)	0.64 (0.59, 0.70)
Race Hispanic	NA	0.87 (0.64, 1.19)	1.16 (0.97, 1.38)	1.07 (0.94, 1.22)	NA	1.14 (1.00, 1.30)	1.09 (0.98, 1.22)	1.16 (0.98, 1.38)	0.82 (0.72, 0.93)
Status urgent	1.29 (1.19, 1.40)	NA	1.21 (1.11, 1.33)	1.29 (1.20, 1.39)	NA	1.17 (1.10, 1.25)	1.22 (1.15, 1.29)	1.42 (1.33, 1.51)	0.70 (0.66, 0.74)
Unstable angina	1.21 (1.04, 1.41)	NA	NA	NA	NA	NA	NA	NA	NA

Table 5. Continued

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

Table 5. Continued

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

Table 5. Continued

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

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

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

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

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

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

Final Models

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

Odds Ratios

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

Final Model Intercept and Coefficients

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

Limitations

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

Conclusion

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

References

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

Appendix

Regression Coefficients and Variable Definitions for STS 2008 Valve Models

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

$$\text{Predicted Risk} = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

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

Appendix Table 1. Regression Coefficients

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

Appendix Table 1. Continued

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

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

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

Variable	Definition
Intercept	= 1 for all patients
Atrial fibrillation	= 1 if patient has history of preop atrial fibrillation, = 0 otherwise
Age function 1	= max (age – 50, 0)
Age function 3	= max (age – 75, 0)
Age by reoperation function	= Age function 1 if surgery is a reoperation, = 0 otherwise
Age by status function	= Age function 1 if status is emergent or salvage, = 0 otherwise
Age by MVR function	= Age function 1 if operation is MVR, = 0 otherwise
Age by MVRepair function	= Age function 1 if operation is MVRepair, = 0 otherwise
BSA function 1	= max (1.4, min [2.6, BSA]) – 1.8
BSA function 2	= (BSA function 1) ²
CHF but not NYHA IV	= 1 if patient has CHF and is not NYHA class IV, = 0 otherwise
CHF and NYHA IV	= 1 if patient has CHF and is NYHA class IV, = 0 otherwise
CHF by MVR function	= 1 if patient has CHF and operation is MVR, = 0 otherwise
CHF by MVRepair function	= 1 if patient has CHF and operation is MVRepair, = 0 otherwise
CLD function	= 0 if no CLD, = 1 if mild CLD, = 2 if moderate CLD, = 3 if severe CLD
CLD by MVR function	= CLD function if operation is MVR, = 0 otherwise
CLD by MVRepair function	= CLD function if operation is MVRepair, = 0 otherwise
Creatinine function 1	= max (0.5, min [creatinine, 5.0]) if patient is not on dialysis, = 0 otherwise
CVD without prior CVA	= 1 if patient has history of CVD and no prior CVA, = 0 otherwise
CVD and prior CVA	= 1 if patient has history of CVD and a prior CVA, = 0 otherwise
Diabetes, noninsulin	= 1 if patient has diabetes not treated with insulin, = 0 otherwise
Diabetes, insulin	= 1 if patient has diabetes treated with insulin, = 0 otherwise
Diabetes by MVR function	= 1 if patient has diabetes and operation is MVR, = 0 otherwise
Diabetes by MVRepair function	= 1 if patient has diabetes and operation is MVRepair, = 0 otherwise
Dialysis	= 1 if patient requires dialysis preoperatively, = 0 otherwise
Dialysis by MVR function	= 1 if patient has history of dialysis and operation is MVR, = 0 otherwise
Dialysis by MVRepair function	= 1 if patient has history of dialysis and operation is MVRepair, = 0 otherwise
Ejection fraction function	= max (50–ejection fraction, 0)
Endocarditis, active	= 1 if patient has active endocarditis, = 0 otherwise
Female	= 1 if patient is female, = 0 otherwise
Female by MVR function	= 1 if female and operation is MVR, = 0 otherwise
Female by MVRepair function	= 1 if female and operation is MVRepair, = 0 otherwise
Female by BSA function 1	= BSA function 1 if female, = 0 otherwise
Female by BSA function 2	= BSA function 2 if female, = 0 otherwise
Hypertension	= 1 if patient has hypertension, = 0 otherwise
IABP or inotropes	= 1 if patient requires IABP or inotropes preoperatively, = 0 otherwise
IABP by MVR function	= 1 if patient requires preop IABP/inotropes and operation is MVR, = 0 otherwise
IABP by MVRepair function	= 1 if patient requires preop IABP/inotropes and operation is MVRepair, = 0 otherwise
Immunosuppressive treatment	= 1 if patient received immunosuppressive therapy within 30 days, = 0 otherwise
Insufficiency mitral	= 1 if patient has at least moderate mitral insufficiency, = 0 otherwise
Insufficiency tricuspid	= 1 if patient has at least moderate tricuspid insufficiency, = 0 otherwise
Left main disease	= 1 if patient has left main disease, = 0 otherwise
MI ≤ 21 days	= 1 if patient has history of MI within 21 days of surgery, = 0 otherwise
MVR	= 1 if valve operation is mitral valve replacement, = 0 otherwise
MVRepair	= 1 if valve operation is mitral valve repair, = 0 otherwise
No. diseased vessel function	= 2 if triple-vessel disease, = 1 if double-vessel disease, = 0 otherwise
Peripheral vascular disease	= 1 if patient has peripheral vascular disease, = 0 otherwise
Race black	= 1 if patient is black, = 0 otherwise
Race Hispanic	= 1 if patient is nonblack Hispanic, = 0 otherwise
Reop, 1 prior operation	= 1 if patient has had exactly 1 previous CV surgery, = 0 otherwise
Reop, ≥ 2 prior operations	= 1 if patient has had 2 or more previous CV surgeries, = 0 otherwise
Reop by MVR function	= 1 if surgery is a reoperation and operation is MVR, = 0 otherwise
Reop by MVRepair function	= 1 if surgery is a reoperation and operation is MVRepair, = 0 otherwise

Appendix Table 2. Continued

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

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

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