Clinical Prediction Models for Valvular Heart Disease

Benjamin S. Wessler, MD, MS; Christine M. Lundquist, MPH; Benjamin Koethe, MPH; Jinny G. Park, MPH; Kristen Brown, BS; Tatum Williamson, MPH; Muhammad Ajlan, MD; Zuhair Natto, DMD; Jennifer S. Lutz, MA; Jessica K. Paulus, ScD; David M. Kent, MD, MS

Background—While many clinical prediction models (CPMs) exist to guide valvular heart disease treatment decisions, the relative performance of these CPMs is largely unknown. We systematically describe the CPMs available for patients with valvular heart disease with specific attention to performance in external validations.

Methods and Results—A systematic review identified 49 CPMs for patients with valvular heart disease treated with surgery (n=34), percutaneous interventions (n=12), or no intervention (n=3). There were 204 external validations of these CPMs. Only 35 (71%) CPMs have been externally validated. Sixty-five percent (n=133) of the external validations were performed on distantly related populations. There was substantial heterogeneity in model performance and a median percentage change in discrimination of −27.1% (interquartile range, −49.4% to −5.7%). Nearly two-thirds of validations (n=129) demonstrate at least a 10% relative decline in discrimination. Discriminatory performance of EuroSCORE II and Society of Thoracic Surgeons (2009) models (accounting for 73% of external validations) varied widely: EuroSCORE II validation c-statistic range 0.50 to 0.95; Society of Thoracic Surgeons (2009) Models validation c-statistic range 0.50 to 0.86. These models performed well when tested on related populations (median related validation c-statistics: EuroSCORE II, 0.82 [0.76, 0.85]; Society of Thoracic Surgeons [2009], 0.72 [0.67, 0.79]). There remain few (n=9) external validations of transcatheter aortic valve replacement CPMs.

Conclusions—Many CPMs for patients with valvular heart disease have never been externally validated and isolated external validations appear insufficient to assess the trustworthiness of predictions. For surgical valve interventions, there are existing predictive models that perform reasonably well on related populations. For transcatheter aortic valve replacement (CPMs additional external validations are needed to broadly understand the trustworthiness of predictions. (J Am Heart Assoc. 2019;8:e011972. DOI: 10.1161/JAHA.119.011972.)

Key Words: clinical prediction models • risk • valvular heart disease

Treatments for patients with advanced valvular heart disease (VHD) are increasingly offered to patients with advanced age and elevated pre-procedural risk.1-3 Clinical predictive models (CPMs) have assumed a central role in clinical decision making and current guidelines link VHD treatment decisions to predicted risk.4-6 CPMs can potentially enhance shared decision making,7,8 when they perform well and are appropriately matched to the correct decisional context, though there remain major questions about how well CPMs for patients with VHD perform in external validations.

It is well recognized that many of the best known (and most widely used) CPMs for VHD were derived on patients receiving surgical interventions9,10 and do not accurately predict outcomes for patients treated with percutaneous interventions,11 though they continue to be used for this purpose. While there are newer efforts to create CPMs specific to percutaneous valve interventions, the relative performance of these models and their performances in external validations remains largely unknown. Generally, the performance of CPMs has often been underreported and incompletely assessed12 and since CPMs that perform poorly can yield misleading predictions that motivate harmful decision making,13 it is essential that clinicians understand CPM performance before leveraging outputs to inform decisions. This is especially important for VHD treatment decisions, given the importance of these tools.

Here, using the Tufts PACE (Predictive Analytics and Comparative Effectiveness) CPM Registry, we describe the
available CPMs for patients with VHD treated with percutaneous or surgical interventions. This analysis focuses on comparative model performance during external validations.

**Clinical Perspective**

**What Is New?**
- Risk prediction is central to decision making for patients with advanced valvular heart disease; however, the performance of clinical predictive models in external validations is often substantially worse than expected based on derivation data set performance.

**What Are the Clinical Implications?**
- Isolated external validations appear insufficient to broadly understand the performance of valvular heart disease clinical predictive models.
- There are clinical predictive models for surgical valvular heart disease interventions that perform well across multiple external validations.
- The trustworthiness of transcatheter aortic valve replacement predictions is largely unknown as these models have not been widely tested in external validations.

**Methods**

**General Approach**

The data that support the findings of this study are available from the corresponding author upon reasonable request. This study analyzed data from the Tufts PACE (Predictive Analytics and Comparative Effectiveness) CPM Registry, a database created to describe the CPM literature for patients at risk for and with known cardiovascular disease. The registry, which is free and available to the public at http://pace.tuftsmedical-center.org/cpm, encompasses a field synopsis of CPMs for patients with VHD. The methods have been previously reported. Briefly, we had previously searched PubMed for English-language articles containing CPMs for cardiovascular disease published from January 1990 through 2015. We extended the search for VHD CPMs to January 2017 to include more recent CPM development (Table S1, Figure S1). Citations were reviewed to confirm completeness of our review. All citations and data fields were extracted in duplicate to ensure accuracy. Discrepancies were discussed until consensus was achieved.

For inclusion in the registry, articles had to meet the following criteria: (1) develop a CPM as a primary aim, (2) contain a model predicting the development of a specified clinical diagnosis (diagnostic models) or the probability of developing a clinical outcome (prognostic models), (3) contain at least 2 outcome predictors, and (4) present enough information to estimate the probability for an individual patient. Articles were excluded if they did not provide enough information to predict a patient’s risk or if the described models predicted surrogate outcomes. We also excluded non-English reports, pharmacology reports, cost-effectiveness models, decision-analysis models, systematic reviews, and editorials.

**Model Selection**

This report focuses on CPMs predicting outcomes for patients with VHD. CPMs predicting natural history outcomes and outcomes after surgical and percutaneous procedures were included. CPMs were grouped based on underlying valve pathology and procedure. CPMs were also included if they were derived on cardiac surgery cohorts where at least 50% of patients received treatment for VHD. CPMs derived exclusively on coronary artery bypass populations were excluded.

**CPM Reporting**

Information was extracted on CPM derivation and reporting. Collected fields included: index clinical condition, predicted outcome, timeframe of prediction, sample size, cohort size, and number of events. We calculated the events per variable (EPV) based on the number of variables included in the model. We also extracted information on modeling method and performance with specific attention to reporting of discrimination and calibration (Table S2).

**Validation Search**

Citations for each CPM article through September 2017 were identified using Scopus and reviewed for inclusion as external validations. An external validation was defined as any evaluation of CPM performance (assessment of either discrimination or calibration) on a data set distinct from the derivation data set. External validations included validations that were done on the same cohort but temporally or geographically distinct from the derivation cohort or on an entirely separate cohort. Each validation citation was reviewed by 2 investigators for inclusion and discrepancies were reviewed with an additional investigator to arrive at consensus.

**Validation Reporting**

Information on validation reporting was extracted, including sample size, continent of study, number of events, and reporting of measures of discrimination and calibration (Table S3). The validation performance analysis focused on whether CPM discrimination changed when compared with
that seen in the derivation population. Because the c-statistic ranges from 0.5 (no discriminatory ability) to 1.0 (perfect discrimination), it has been rescaled as Somer’s D statistic $2(c−0.5))$ so that discrimination ranges from 0 (no discrimination) to 1.0 (perfect discrimination). We describe changes on this scale because it more intuitively reflects the true changes in discriminatory power. The percentage change in discrimination [(Validation AUC−0.5)/(Derivation AUC−0.5)×100] is presented. We also document whether validations include any assessment of CPM calibration. There is currently no literature standard for assessing calibration. Given this lack of consistency and interpretability, we have only reported on whether this dimension of performance was assessed. Calibration assessment included any comparison of observed versus expected outcomes. Examples include a Hosmer-Lemeshow statistic or calibration plot. For this study we also included measures of global fit, where overall observed event rates are compared with predicted rates (ie, calibration-in-the-large).

**Relatedness**

To assess the similarity between the derivation population and the validation population for each validation, we created a relatedness rubric to divide validations into 2 categories — "related" and "distantly related." The rubric contained 3 domains: (1) type of intervention (ie, percutaneous or surgical), (2) percentage of the population undergoing isolated valve procedures (as opposed to valve procedures in combination with revascularization), and (3) calendar years of enrollment. We considered a validation population to be “related” if all of the following criteria were met: (1) same type of intervention (eg, both surgical populations), (2) ± 10% absolute difference in the proportion of isolated valve procedure (eg, derivation population was 100% isolated valve and validation population was 95% isolated valve), and (3) overlapping years of enrollment. Matches that did not meet all 3 criteria were deemed “distantly related.”

**Results**

**VHD CPMs**

We identified 49 CPMs predicting clinical outcomes for patients with VHD, which were cited a total of 1296 times (Table 1, Table S2). Thirty-four (69%) predict outcomes following surgical interventions, 12 (24%) predict outcomes following percutaneous interventions, and 3 (6%) predict outcomes in the absence of intervention (Table 2). Overall, the most commonly predicted outcomes were 30-day mortality (n=14, 29%) and in-hospital mortality (n=14, 29%). Twenty-four models (46%) were derived from patients in North America, followed by 12 (23%) from Europe and 8 (15%) from Asia (Figure 1). The median derivation sample size was 4510 (interquartile range [IQR], 1087–18 686), median event rate was 8.3% (IQR, 4.5%–14.8%), and median EPV was 40 (IQR, 20–92) (Table 2). The median number of covariates was 10 (IQR, 7–19).

Among models that reported a c-statistic (n=37, 76%), the overall median ROC was 0.76 (IQR, 0.72–0.78) (Table 2). When stratified by intervention type, the median c-statistic was 0.77 (IQR, 0.75–0.79) for CPMs predicting outcomes following surgical interventions, 0.68 (IQR, 0.67–0.74) for CPMs for percutaneous interventions, and 0.81 (IQR, 0.77–0.86) for CPMs predicting outcomes in the absence of intervention (Table 2).

**CPMs for Isolated Valve Disease**

There are 31 CPMs for isolated valve disease. Sixteen (52%) predict outcomes following surgical intervention and 12 (39%) predict outcomes following percutaneous interventions (transcatheter aortic valve replacement [TAVR], balloon aortic valvuloplasty, and percutaneous mitral balloon valvuloplasty). Three CPMs (10%) predict outcomes for patients with aortic stenosis in the absence of intervention. The median derivation sample size was 2552 (IQR, 1064–108 410) and the median age was 70 (IQR, 64–82). The median number of events was 360 (IQR, 104–2021) and the median EPV was 55 (IQR, 18–112). The median event rate was 10% (IQR, 4.6%–18.3%). For the 27 (87%) models reporting discrimination, the median c-statistic was 0.74 (IQR, 0.69–0.78).

**CPMs for Isolated or Multiple Valve Disease**

There are 11 CPMs that predict outcomes for patients undergoing either single or multiple valve surgical procedures. The median derivation sample size was 3544 (IQR, 2297–12 079) and the median age was 60 (IQR, 54–65). The median number of events was 303 (IQR, 139–507) and the median EPV was 26 (IQR, 20–40). The median event rate was 5.1% (IQR, 4.1%–9.5%). For the 10 (91%) models reporting discrimination, the median c-statistic was 0.78 (IQR, 0.76–0.79).

**CPMs for Multiple Valve Disease**

There are 7 CPMs that predict outcomes specifically for multiple valve surgical interventions. These CPMs include the Society of Thoracic Surgeons (STS) Multi-Valve Risk Models and the derivation sample size was 18 686 (IQR, 4510–22 861). The median number of events was 1420 (IQR, 591–1981) and the median EPV was 71 (IQR, 48–92). Median age was 70 (IQR, 70–71). The median event rate was 9.4% (IQR, 7.6%–11.3%). Median number of covariates was 20 (IQR, 14–23).
### Table 1. De Novo VHD CPMs Overview

| Author, Model Name | Publication, y | Valve | Standardized Type of Intervention | Outcome | Model Method | C-Statistic | Calibration Measure | Externally Validated? |
|--------------------|----------------|-------|-----------------------------------|---------|--------------|-------------|---------------------|----------------------|
| **Isolated valve** |                |       |                                   |         |              |             |                     |                      |
| Edwards,16 STS (original) Isolated Valve | 2001 | Aortic/Mitral | Surgery | 30 d operative mortality | Logistic regression | 0.766 | HL statistic, Calibration plot | Yes |
| Nowicki,17 NNE Aortic and Mitral Models | 2004 | Aortic | Surgery | In-hospital mortality | Logistic regression, score | 0.75 | HL statistic | Yes |
| Kuduvalli,18 NWQIP | 2007 | Aortic | Surgery | In-hospital mortality | Logistic regression, score | 0.79 | HL statistic | Yes |
| Cruz-Gonzalez,19 PMV Score | 2009 | Mitral | Percutaneous | Procedural success | Logistic regression, score | NR | HL statistic | Yes |
| Monin20 | 2009 | Aortic stenosis | Natural History | Composite (Non-MACE) | Logistic regression, score | 0.90 | HL statistic | Yes |
| O’Brien,9 STS (2009)—Composite AEs | 2009 | Aortic/Mitral | Surgery | DSWI | Logistic regression | 0.704 | None | Yes |
| O’Brien,9 STS (2009)—Mortality | 2009 | Aortic/Mitral | Surgery | DSWI | Logistic regression | 0.805 | None | Yes |
| O’Brien,9 STS (2009)—Prolonged LOS | 2009 | Aortic/Mitral | Surgery | Prolonged LOS | Logistic regression | 0.77 | None | Yes |
| O’Brien,9 STS (2009)—Prolonged Ventilation | 2009 | Aortic/Mitral | Surgery | Prolonged ventilation | Logistic regression | 0.77 | None | Yes |
| O’Brien,9 STS (2009)—Renal Failure | 2009 | Aortic/Mitral | Surgery | Renal failure | Logistic regression | 0.782 | None | Yes |
| O’Brien,9 STS (2009)—Reoperation | 2009 | Aortic/Mitral | Surgery | Reoperation | Logistic regression | 0.643 | None | Yes |
| O’Brien,9 STS (2009)—Short LOS | 2009 | Aortic/Mitral | Surgery | Prolonged LOS | Logistic regression | 0.738 | None | No |
| O’Brien,9 STS (2009)—Stroke | 2009 | Aortic/Mitral | Surgery | Stroke | Logistic regression | 0.694 | None | Yes |
| Guaragna,21 GuaragnaSCORE | 2010 | Aortic/Mitral | Surgery | In-hospital mortality | Logistic regression, score | 0.82 | HL statistic, Calibration plot | Yes |
| Guo22 | 2010 | Aortic | Surgery | In-hospital mortality | Logistic regression | NR | HL statistic | No |
| Elmariah,23 CRRAC the AV Score | 2011 | Aortic | Percutaneous | 30 d mortality | Cox regression, score | 0.754 | HL statistic | No |
| Boulet24 | 2012 | Mitral | Percutaneous | Composite (MACE) | Cox regression, score | 0.74 | Calibration plot | No |
| Cioffi25 | 2012 | Aortic stenosis | Natural History | Composite (MACE) | Cox regression, score | NR | None | No |

Continued
| Author, Model Name | Publication, y | Valve | Standardized Type of Intervention | Outcome | Model Method | C-Statistic | Calibration Measure | Externally Validated? |
|--------------------|----------------|-------|-----------------------------------|---------|--------------|-------------|---------------------|---------------------|
| Holme,26 SEAS Score | 2012           | Aortic stenosis | Natural History                  | 5 y mortality | Cox regression | 0.722       | HL statistic, Calibration plot, Brier score | No                  |
| Kötting,27 German Aortic Valve Score | 2013          | Aortic | Percutaneous                      | In-Hospital Mortality | Logistic regression, score | 0.808 | HL statistic | Yes                  |
| Arnold,28 6 mo and 1 y Models | 2014       | Aortic stenosis | Percutaneous                      | Composite (Non-MACE) | Logistic regression | 0.66 | HL statistics, Calibration plot | Yes                  |
| Capodanno,29 OBSERVANT Score | 2014 | Aortic stenosis | Percutaneous                      | 30 d mortality | Logistic regression, score | 0.73 | HL statistic, Calibration plot | Yes                  |
| D'Ascenzo,30 Survival Post-TAVI (STT)—30 d and 1 y Models | 2014 | Aortic | Percutaneous                      | 30 d mortality | Logistic regression, score | 0.66 | HL statistic | Yes                  |
| Iung31 | 2014 | Aortic | Percutaneous                      | 30 d mortality | Logistic regression, score | 0.67 | HL statistic, Calibration in the large, Calibration plot | No                  |
| Debonnaire,32 TAVI2-SCORe | 2015          | Aortic | Percutaneous                      | 1 y mortality | Logistic regression, score | 0.715 | HL statistic, Calibration in the large | Yes                  |
| Edwards7 | 2016 | Aortic | Percutaneous                      | In-hospital mortality | Logistic regression | 0.67 | HL statistics, Calibration in the large, Calibration plot | Yes                  |
| Isolated or multiple valve |                |       |                                   |         |               |             |                     |                     |
| Kaplan33 | 2003 | All | Surgery                           | Pacemaker placement | Logistic regression, score | NR | None | No                  |
| Ambler34 | 2005 | Aortic, mitral | Surgery                          | In-hospital mortality | Logistic regression, Score | 0.77 | HL statistic, Calibration plot | Yes                  |
| Xu35      | 2006 | All | Surgery                           | Prolonged LOS      | Logistic regression | 0.81 | Calibration in table form | No                  |
| Hannan36 | 2007 | Aortic, mitral | Surgery                          | In-hospital mortality | Logistic regression, Score | 0.794 | HL statistic, Calibration plot | Yes                  |
### Table 1. Continued

| Author, Model Name | Publication, y | Valve | Standardized Type of Intervention | Outcome | Model Method | C-Statistic | Calibration Measure | Externally Validated? |
|--------------------|---------------|-------|-----------------------------------|---------|--------------|-------------|---------------------|----------------------|
| Xu, Fuwai Score    | 2007          | All   | Surgery                           | Prolonged LOS | Logistic regression, Score | 0.76        | HL statistic, Calibration plot | Yes |
| Shi, EuroSCORE II  | 2010          | Aortic, mitral | Surgery                        | In-hospital mortality | Logistic regression | 0.7358     | None                | No |
| Aryanaratne, Aus-AVR Score | 2011          | Aortic, mitral | Surgery                        | 30 d mortality    | Logistic regression, Score | 0.78      | HL statistic, Calibration in the large | Yes |
| Nashef, EuroSCORE II | 2012        | All   | Surgery                           | In-hospital mortality | Logistic regression | 0.8095     | None                | Yes |
| Hannan, NY Operative Mortality Risk Score | 2013        | Aortic, mitral | Surgery                        | 30 d mortality    | Logistic regression, Score | 0.781     | HL statistic       | Yes |
| Wang               | 2013          | All   | Surgery                           | Prolonged ventilation | Logistic regression | 0.789      | HL statistic        | No |
| Zheng              | 2013          | Aortic, mitral | Surgery                        | In-hospital mortality | Logistic regression, Score | 0.76      | HL statistic, Chi-square statistic, Calibration plot | No |
| Multiple valve     |               |       |                                   |                     |                            |            |                     |  |
| Guo                | 2010          | Aortic, mitral | Surgery                        | In-hospital mortality | Logistic regression | NR         | HL statistic        | No |
| Rankin, AM Preop   | 2013          | Aortic, mitral | Surgery                        | 30 d mortality    | Logistic regression | NR         | Calibration plot   | Yes |
| Rankin, MT Preop   | 2013          | mitral, tricuspid | Surgery                        | 30 d mortality    | Logistic regression | NR         | Calibration plot   | Yes |
| Rankin, AM Preop   | 2013          | Aortic, mitral, tricuspid | Surgery                        | 30 d mortality    | Logistic regression | NR         | Calibration plot   | Yes |
| Rankin, AM Preop + Intraop | 2013      | Aortic, mitral | Surgery                        | 30 d mortality    | Logistic regression | NR         | Calibration plot   | Yes |
| Rankin, MT Preop + Intraop | 2013       | Mitral, tricuspid | Surgery                        | 30 d mortality    | Logistic regression | NR         | Calibration plot   | Yes |
| Rankin, AM Preop + Intraop | 2013       | Aortic, mitral, tricuspid | Surgery                        | 30 d mortality    | Logistic regression | NR         | Calibration plot   | Yes |

AEs indicates adverse events; AM, aortic, mitral; AMT, aortic, mitral, tricuspid; AV, aortic valvuloplasty; Aus-AVR, Australian aortic valve replacement; CRRAC, critical status, renal dysfunction, eugical atrial pressure, and cardiac output; DSWI, deep sternal wound infections; EuroSCORE, European System for Cardiac Operative Risk Evaluation; HL, Hosmer-Lemeshow; LOS, length of stay; MACE, major adverse cardiovascular events; MT, mitral, tricuspid; NNE, Northern New England; NR, not reported; NWQIP, North West Quality Improvement Programme in Cardiac Interventions; NY, New York; OBSERVANT, Observational Study of Appropriateness, Efficacy and Effectiveness of AVR-TAVR Procedures for the Treatment of Severe Symptomatic Aortic Stenosis; PMV, percutaneous mitral valvuloplasty; SEAS, Simvastatin and Ezetimibe in Aortic Stenosis; STS, Society of Thoracic Surgeons; STT, Survival post-TAVI; TAVI, transcatheter aortic valve implantation; TAVR, Transcatheter Aortic Valve Replacement.
Table 2. Reported Characteristics of De Novo Valvular Heart Disease CPMs

| Characteristic* | Overall (n=49) | Surgical (n=34) | Percutaneous (n=12) | Natural History (n=3) |
|----------------|---------------|----------------|---------------------|----------------------|
| Publication range | 2001 to 2016 | 2001 to 2013 | 2009 to 2016 | 2009 to 2012 |
| Age, y | 69 (61–79) | 65 (58–70) | 82 (82–83) | 68 (67–70) |
| Sample size | 4510 (1087–18 686) | 12 079 (3125–92 563) | 1160 (752–2241) | 772 (440–1169) |
| Event rate | 0.08 (0.05–0.15) | 0.07 (0.04–0.11) | 0.14 (0.06–0.37) | 0.35 (0.23–0.47) |
| Events per variable | 40 (20–92) | 46 (25–110) | 42 (18–81) | 11 (10–45) |
| C-statistic | 0.76 (0.72–0.78) | 0.77 (0.75–0.79) | 0.68 (0.67–0.74) | 0.81 (0.77–0.86) |
| % Externally validated | 71.4 | 73.5 | 83.3 | 33.3 |

CPM indicates clinical predictive models.
*Values are reported as median (interquartile range), unless otherwise specified.
†De novo CPM search spans January 1, 1990 to January 1, 2017.

CPMs for Percutaneous VHD Interventions

Since 2009, there have been 12 CPMs presented that predict outcomes following percutaneous VHD interventions. Two models predict outcomes following mitral percutaneous mitral balloon valvuloplasty.24,44 There were 9 CPMs that predict outcomes following TAVR with a median derivation sample size of 2130 (IQR, 10 642–552). The median age of patients in the TAVR CPMs was 82 (IQR, 82–83). TAVR CPMs had a median number of events of 253 (IQR, 80–180) and a median EPV was 28 (IQR, 20–70). The median event rate was 9.9% (IQR, 5.6%–15.7%). All of the CPMs predicting outcomes following TAVR reported discrimination with a median c-statistic of 0.67 (IQR, 0.66–0.72).

External Validations

Two hundred and four external validations of these CPMs were identified, of which 190 (93%) report a c-statistic. Overall, 35 (71%) of the VHD CPMs have been externally validated and 20 (37%) have been externally validated more than once. External validations were most commonly done in cohorts of patients from Europe (n=93, 46%), Asia (n=38, 19%), and North America (n=37, 18%) (Figure 1). Fifty-three (26%) validations were performed on populations from the same continent as the derivation population, with a median c-statistic of 0.71 (IQR, 0.66–0.77). Seventy-one (35%) were done on populations from a different continent, with a median c-statistic of 0.68 (IQR, 0.64–0.73). External validations overall had a median c-statistic of 0.71 (IQR, 0.65–0.77) (Table 3). For the models that were externally validated, we noted an overall median percentage change in discrimination of –27.1% (IQR, –49.4–5.7). Just under two-thirds of validations (n=129) demonstrate at least a 10% relative decline in discriminatory power, and 18 (9%) showed a decline of >80%. Thirty-three (16%) validations showed CPM discrimination at or above that seen in the derivation cohort.

The distribution of number of validations was skewed towards a small number of CPMs. Two CPMs (EuroSCORE II and STS [2009] Models†) accounted for 73% of the external validations. EuroSCORE II has been validated 78 times across 5 continents (Table 4). Validation c-statistics ranged from 0.50 to 0.95 with a median percentage change of –23.4% (range –100%–46.7%). For the STS (2009) Models, validation c-statistics ranged from 0.50 to 0.86 (Table 5). The median percentage change was –31.8% and ranged from –100% to +18%. The STS (2009) Models have been validated 70 times across 5 continents.

Of CPMs that have been validated at least 2 times (n=17) in related populations, the highest median validation c-statistic was seen for EuroSCORE II (0.82 [IQR, 0.76–0.85]), followed by the North West Quality Improvement Programme in Cardiac Interventions model (0.78 [IQR, 0.77–0.78]), and the Northern New England Aortic model (0.76 [IQR, 0.75–0.77]) (Table 6). Forty-five (22%) external validations did not report any measure of calibration. Of the 159 validations that did report calibration, 103 (65%) reported the Hosmer-Lemeshow statistic, 87 (55%) reported calibration-in-the-large, and 46 (29%) included a calibration plot. Median c-statistic was 0.71 (IQR, 0.65–0.77) for validations that reported some measure of calibration and 0.68 (IQR, 0.63–0.74) for validations that did not report any calibration.

Clinical relatedness between the development and validation populations was assessed using a novel rubric. Seventy-one validations (35%) were performed on related populations, while the remaining 133 (65%) were performed on distantly-related populations. The median validation c-statistic was 0.73 (IQR, 0.67–0.79) for related validations and 0.70 (IQR, 0.62–0.76) for distantly-related validations (P=0.009). There was a significant difference in percentage change in discrimination: the median change in c-statistic was –12.2% (IQR,
−28.3% to +2.5%) for related validations and −32.1% (IQR, −54.9% to −12.8%) for distantly-related validations (Figure 2, \( P<0.0001 \)).

CPMs that were derived on percutaneously-treated VHD populations and externally validated (\( n=9 \)) underwent a total of 19 validations, almost all of which were on percutaneous populations. CPMs that were derived on surgical VHD populations and externally validated (\( n=25 \)) underwent a total of 184 validations, of which 130 (71%) were on surgical populations, 52 (28%) were on percutaneous populations, and

Figure 1. Geography of derivation and validation cohorts. Country of origin for derivation (A) and validation (B) populations. Maps created in Tableau Public.
2 (1%) were on populations including both surgical and percutaneous interventions. For validations of surgical VHD models discrimination was better when CPMs were tested on cohorts treated with surgical versus percutaneous interventions (median c-statistic 0.74 versus 0.63, \(P< 0.001\)).

Of the surgical VHD CPMs validated on percutaneous populations (n = 52 validations), the CPM most often validated was the STS (2009) model predicting mortality (n = 27, median c-statistic 0.64 [IQR, 0.58–0.67]). EuroSCORE II (n = 20) had the highest discrimination in this setting, with a median c-statistic of 0.67 (IQR, 0.55–0.71).

**Discussion**

Here we show that there are many CPMs available for patients with VHD and that many of these CPMs have not been externally validated. For the CPMs that have been externally

| Table 3. Reported Characteristics of Valvular Heart Disease External Validations‡ |
|---------------------------------------------------------------|
| Characteristic | Overall (n=204)† | Surgical (n=131) | Percutaneous (n=70) |
|----------------|------------------|-----------------|-------------------|
| Sample size    | 450 (249–1495)   | 809 (407–3306)  | 304 (180–453)     |
| Number of events | 38 (15–95)      | 48 (14–119)     | 38 (15–56)        |
| Event rate     | 0.06 (0.03–0.12) | 0.04 (0.03–0.07)| 0.11 (0.08–0.17) |
| % Men          | 53 (47–62)       | 57 (53–66)      | 47 (43–52)        |
| C-statistic    | 0.71 (0.65–0.77) | 0.74 (0.70–0.79)| 0.63 (0.57–0.68) |

*Values are reported as median (interquartile range).
†Validations done on populations treated with surgical and percutaneous interventions that did not disaggregate results (n=2) are only included in the overall count.
‡Validation search includes citations through September 8, 2017.

**Table 4. EuroSCORE II Population Compared With External Validation Populations, Stratified by Relatedness**

| Statistic        | EuroSCORE II | Validation Populations§ | Related | Distantly Related |
|------------------|--------------|-------------------------|---------|------------------|
| Total patients (n) | 16 828       | 14 382                  | 98 744  |                  |
| Total validations (n) | NA           | 5                       | 73      |                  |
| Age, y           | Mean (SD): 64.6 (12.5) | 63.4 (62.7–67.0) | 67.1 (61.1–80.5) |
| Number of events (n) | 656          | 123 (53–215)           | 27 (12–57) |
| Event rate, %    | 3.9          | 5.7 (5.7–6.1)          | 6.3 (3.0–10.5) |
| Sex reported, n (%) | NA           | 5 (100%)               | 50 (68%) |
| Men, %           | 69.1         | 65.2 (62.5–66.5)       | 52.5 (46.8–64.1) |
| Type of intervention, n (%) |             |                         |         |
| Surgery          | 1 (100%)     | 5 (100%)               | 52 (71.2%) |
| Percutaneous     | 0 (0%)       | 0 (0%)                 | 20 (27.4%) |
| Both             | 0 (0%)       | 0 (0%)                 | 1 (1.4%)  |
| Valve-related, % | 53.3         | 56 (54.6–56.1)         | 100 (100–100) |
| Enrollment, y (range) | 2010        | 2005 to 2013           | 1999 to 2015 |
| C-statistic      | 0.8095       | 0.82 (0.76–0.85)       | 0.72 (0.67–0.78) |
| C-statistic (range) | NA          | 0.737 to 0.861         | 0.50 to 0.95  |
| Any calibration reported, n (%) | 0 (0%)   | 4 (80%)                | 65 (89%)  |
| Change in discrimination, % | NA     | 2.6 (−16.0–13.1)      | −28.9 (−45.3–−9.5) |

EuroSCORE indicates European System for Cardiac Operative Risk Evaluation.
*All values are reported as median (interquartile range) unless otherwise specified.
†Validation data is reported at the population level only; patient-level data was not available.
‡Validation population are “related” if it meets all of the following criteria: (1) same type of intervention (eg, both surgical populations), (2) \(\pm 10\%\) absolute difference in the proportion of isolated valve procedure (eg, derivation population was 100% isolated valve and validation population was 95% isolated valve), and (3) overlapping years of enrollment. A validation population that does not meet all 3 criteria is “distantly related.”
§Change in discrimination is calculated as \([\text{Validation AUC}_0.5−\text{Derivation AUC}_0.5]/(\text{Derivation AUC}_0.5−0.5)\) × 100.
Table 5. STS (2009) Population Compared With External Validation Populations, Stratified by Relatedness

| Statistic* | STS Models (n=9) | Validation Populations†‡§ | Related | Distantly Related |
|------------|-----------------|---------------------------|---------|-------------------|
| Total patients, n | 109 759 | 37 395 | 49 530 |
| Total validations, n | NA | 33 | 37 |
| Age, y | Not Reported | 64.7 (56.6–73) | 81.6 (74.5–83) |
| Number of events, n | 9164 (3706–12 892) | 29 (12–82) | 38 (18–57) |
| Event rate, % | 8.3 (3.4–11.7) | 4.9 (2.7–12.6) | 9.1 (3.7–11.7) |
| Men, % | 55.4 | 56 (56.0–74.9) | 47.8 (43.6–55.3) |
| Type of intervention, n (%) | | | |
| Surgery | 9 (100%) | 33 (100%) | 8 (21.6%) |
| Percutaneous | 0 (0%) | 0 (0%) | 28 (75.7%) |
| Both | 0 (0%) | 0 (0%) | 1 (2.7%) |
| Valve-related, % | 100 | 100 (100–100) | 100 (100–100) |
| Enrollment, y (range) | 2002–2006 | 1999–2014 | 1999–2015 |
| C-statistic, median, IQR | 0.74 (0.70–0.77) | 0.72 (0.67–0.79) | 0.65 (0.6–0.71) |
| Calibation reported, n (%) | 0.643 to 0.805 | 0.612 to 0.86 | 0.5 to 0.81 |
| Change in discrimination, % | 9 (100%) | 18 (54.5%) | 28 (75.7%) |
| # | NA | –21.3 (–34.4–2.3) | –50.8 (–67.2–25.1) |

IQR indicates interquartile range; STS, Society of Thoracic Surgeons.

*All values are reported as median (interquartile range) unless otherwise specified.

†Validation data is reported at the population level only; patient-level data was not available.

‡Validation population is “related” if it meets all of the following criteria: (1) same type of intervention (eg, both surgical populations), (2) ±10% absolute difference in the proportion of isolated valve procedure (eg, derivation population was 100% isolated valve and validation population was 95% isolated valve), and (3) overlapping years of enrollment. A validation population that does not meet all 3 criteria is “distantly related.”

§Change in discrimination is calculated as [(Validation AUC–0.5)–(Derivation AUC–0.5)]/(Derivation AUC–0.5)×100.

validated, models often perform substantially worse than expected based on performance in derivation data sets. Notably, isolated external validations of VHD CPMs appear insufficient for broadly understanding CPM performance in the context of specific clinical decisions as predictive models may have highly variable performance across various databases. For patients under consideration for surgical VHD interventions, there are CPMs that have been extensively validated. The fidelity of TAVR CPM predictions is largely unknown, as these models have not been widely tested in external validations.

Predicted risk is central to procedural decision making for patients with VHD, however. individual risk estimates using published CPMs for VHD appear more uncertain than originally thought, especially when prediction models are derived on patients who are not closely related to the patients being treated. CPM performance (specifically discrimination) substantially degrades from the derivation population to the validation population, particularly when populations are “distantly related” with respect to procedure type (percutaneous versus open surgical), therapeutic era, and the need for concurrent revascularization. Without attention to these patient-level specifics, it is likely that there is widespread inappropriate use of CPMs that are informing treatment decisions for patients with VHD. While it is encouraging that newer models have been developed for TAVR patients, these CPMs have not been widely validated or integrated into contemporary guidelines, and have risk estimates that may become inaccurate as devices continue to improve and procedural techniques mature. The attenuated performance of these TAVR CPMs may also be related to the magnitude and significance of comorbid illnesses that are common for older treated adults and are rarely included as part of parsimonious modeling efforts. More work is needed to understand these risk factors.

The decrease in discrimination that is observed in this study may be attributable to model overfitting, differences in case mix (ie, narrower populations in the validation data set), and phenotypic heterogeneity. Ultimately, the relevant performance metrics for clinicians relate to the patients they are treating (with a specific intervention), not to performance measured at the time of CPM development. Rarely, discrimination appears to improve during validations. This is likely the result of differences between the derivation population and the validation population where some models are developed on more highly selected (narrow case-mix) cohorts than they are testing on. The data presented here demonstrates that CPMs externally validated multiple times show substantial variation in performance. This strongly suggests that adequate
performance demonstrated in a single external validation may be insufficient to assess the quality (and utility) of VHD CPMs and that a more tailored approach is needed to understand the trustworthiness of CPM predictions in specific settings.

There is increasing recognition of the central importance of CPM calibration. Surprisingly, calibration was reported in only 78% of the external validations of VHD CPMs. There is no agreed-upon standard for reporting model calibration and no consensus on interpreting this metric. Moreover, there are well-recognized limitations to the most commonly reported measure, the Hosmer-Lemeshow statistic (eg, sample-size dependence). Reporting of model calibration represents a poorly-recognized limitation to the most commonly reported metric. The optimal number of validations required to adequately assess CPM performance remains unknown. Ideally, CPMs are serially validated and recalibrated (if necessary) to optimize performance for specific, local clinical decision making. Without addressing these limitations, clinical decisions that leverage CPM outputs may be inaccurate and lead to harmful decisions.

This analysis offers a structure to consider which CPMs are most accurate (discrimination and calibration) and trustworthy (consistent performance in multiple external validations). For patients being considered for surgical valve interventions, EuroSCORE II (median validation c-statistic 0.82 [0.76, 0.85]), North West Quality Improvement Programme in Cardiac Interventions; NY, New York; OBSERVANT, Observational Study of Appropriateness, Efficacy and Effectiveness of AVR-TAVR Procedures for the Treatment of Severe Symptomatic Aortic Stenosis; STS, Society of Thoracic Surgeons; STT, Survival post-TAVI; TAVI, transcatheter aortic valve implantation.

*Change in discrimination is calculated as [(Validation AUC – Derivation AUC) / (Derivation AUC – 0.5)] × 100.

### Table 6. CPMs that Have Been Validated ≥2 Times in Related Populations

| De Novo CPM | Pub., Y | Model Name | External Validations in Related Populations (n) | Validation C-statistic, median (IQR) | % Change in Discrimination,* | Any Calibration Reported (%) |
|-------------|--------|------------|-----------------------------------------------|-------------------------------------|-----------------------------|-------------------------------|
| 2001        | STS (original): Isolated Valve | 2 | 0.77 (0.77, 0.77) | 2.6 (2.6–2.6) | 100 |
| 2004        | NNE Aortic | 2 | 0.76 (0.76, 0.77) | 4.0 (2.0–6.0) | 100 |
| 2007        | NWQIP | 2 | 0.78 (0.77, 0.78) | –1.8 (–2.7–2.9) | 100 |
| 2005        | Ambler | 4 | 0.73 (0.72, 0.76) | –15.2 (–18.9–2.2) | 100 |
| 2009        | STS: Mortality | 19 | 0.74 (0.71, 0.79) | –21.3 (–31.5–4.8) | 95 |
| 2009        | STS: Stroke | 2 | 0.65 (0.65, 0.66) | –20.9 (–23.8–17.9) | 0 |
| 2009        | STS: Prolonged Ventilation | 2 | 0.72 (0.68, 0.75) | –20.2 (–33.8–6.6) | 0 |
| 2009        | STS: Prolonged LOS | 2 | 0.67 (0.65, 0.68) | –38.1 (–43.5–32.8) | 0 |
| 2009        | STS: Renal Failure | 2 | 0.76 (0.72, 0.79) | –9.6 (–22.5–3.4) | 0 |
| 2009        | STS: DSWI | 2 | 0.68 (0.65, 0.70) | –13.7 (–24.8–2.7) | 0 |
| 2009        | STS: Composite AEs | 2 | 0.68 (0.65, 0.71) | –18.8 (–30.7–6.9) | 0 |
| 2009        | STS: Reoperation | 2 | 0.64 (0.63, 0.65) | –21.1 (–11.9–7.7) | 0 |
| 2011        | Aus-AVR Score | 3 | 0.72 (0.67, 0.72) | –22.9 (–40.4–20.4) | 100 |
| 2012        | EuroSCORE II | 5 | 0.82 (0.76, 0.85) | 2.6 (16.0–13.1) | 80 |
| 2013        | NY Operative Mortality Risk Score | 3 | 0.73 (0.71, 0.75) | –18.1 (–26.5–9.8) | 66.7 |
| 2014        | OBSERVANT Score | 4 | 0.60 (0.58, 0.61) | –57.8 (–63.7–50.7) | 50 |
| 2014        | STT: 30 d | 2 | 0.66 | 0 | 50 |
Figure 2. Percentage change in discrimination in external validations of valvular heart disease clinical prediction models, stratified by relatedness. Each bar represents a unique external validation that reports a c-statistic (n=205). Society of Thoracic Surgeons (2009) Models. Percentage change in discrimination is calculated as \( \frac{[\text{validation c-statistic} - 0.5] - [\text{derivation c-statistic} - 0.5]}{[\text{derivation c-statistic} - 0.5]} \times 100 \). STS indicates Society of Thoracic Surgeons.
Interventions Model (median validation c-statistic 0.78 [0.77, 0.78]), Northern New England Aortic Model (median validation c-statistic 0.76 [0.76, 0.77]), Ambler (median validation c-statistic 0.73 [0.72, 0.76]), and STS (2009) Mortality (median validation c-statistic 0.74 [0.71, 0.79]) have reasonable discrimination and multiple assessments of discrimination and calibration in external data sets. There are no CPMs for patients treated with TAVR that demonstrate good performance across multiple related validation databases. The trustworthiness of these newer risk estimates for TAVR remains under-studied.

There are several limitations to this work. Our review was limited to CPMs that provide enough information in the published report to calculate a risk prediction for a patient. Logistic regression models that did not report a full equation or intercept were not included. Cox regression models that did not report a point score or baseline hazard were excluded. The search for de novo VHD CPMs was last run in January 2017. While newer CPMs have been developed, there is often substantial delay before the publication of subsequent external validations. Notably, we present relative changes in discrimination to more accurately document changes on a clinically relevant scale, where small decreases in the C-statistic can result in large changes in clinically relevant performance. Lastly, this study was limited in its examination of CPM calibration, which is an important measure of model performance, but often poorly reported and without a widely-accepted summary measure.

While there are numerous available CPMs for patients with VHD, many have never been externally validated, and for those that have, discriminatory performance is often much worse than originally reported. We note that CPM performance is highly dependent on the cohort selected for study, suggesting that one-off external validations may inadequately assess performance. Instead of new CPM development, robust external validations of established TAVR CPMs and without a widely-accepted summary measure.

**Acknowledgments**

The statements in this publication are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute, its Board of Governors, or Methodology Committee.

**Sources of Funding**

This work was partially supported through a Patient-Centered Outcomes Research Institute Methods Award (ME-1606-35555), as well as by the National Institutes of Health (R03 AG056447 GEMSSTAR Grant and K23 AG055667) and the Bellows Foundation Grant from the American College of Cardiology.

**Disclosures**

None.

**References**

1. Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG, Tuzcu EM, Webb JG, Fontana GP, Makkar RR, Brown DL, Block PC, Guyton RA, Pichard AD, Bavaria JE, Herrmann HC, Douglas PS, Petersen JL, Akin JJ, Anderson WN, Wang D, Pocock S; PARTNER Trial Investigators. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. N Engl J Med. 2010;363:1597–1607.

2. Muller DWM, Farivar RS, Jansz P, Bae R, Waiters D, Clarke A, Grayburn PA, Stoler RC, Dahlie G, Rein KA, Shaw M, Scalisi GM, Guerrero M, Pearson P, Kapadia S, Gillinov M, Pichard A, Corso P, Popma J, Chiang M, Blanke P, Leipsic J, Soriaja P; Tendyne Global Feasibility Trial Investigators. Transcatheter mitral valve replacement for patients with symptomatic mitral regurgitation: a global feasibility trial. J Am Coll Cardiol. 2017;69:381–391.

3. Aboulhosn J, Cabalka AK, Levi DS, Himbert D, Testa L, Latib A, Makkar RR, Boudjemline Y, Kim DW, Kefer J, Bleiziffer S, Kerst G, Dvir D, McElhinney DB. Transcatheter valve-in-ring implantation for the treatment of residual or recurrent tricuspid valve dysfunction after prior surgical repair. JACC Cardiovasc Interv. 2017;10:53–63.

4. Holmes DR Jr, Mack MJ, Kaul S, Agnihotri A, Alexander KP, Bailey SR, Calhoun JH, Carabello BA, Desai MY, Edwards FH, Francis GS, Gardner TJ, Kappetein AP, Lindenbaum JA, Mukherjee C, Mukherji A, Otto CM, Ruiz CE, Sacco RI, Smith D, Thomas JD. 2012 ACCF/AATS/SCAI/STS expert consensus document on transcatheter aortic valve replacement. J Am Coll Cardiol. 2012;59:1200–1254.

5. Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP III, Fleisher LA, Jneid H, Mack MJ, McLeod CJ, O’Gara PT, Rogolin VH, Sundt TM III, Thompson A. 2017 AHA/ACC focused update of the 2014 AHA/ACC guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. Circulation. 2017;135:e1159–e1195.

6. Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP III, Guyton RA, O’Gara PT, Ruiz CE, Skubas NJ, Sorajja P, Sundt TM III, Thomas JD; ACC/AHA Task Force Members. 2014 AHA/ACC guideline for the management of patients with valvular heart disease: executive summary: a report of the American College Of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation. 2014;129:2490–2537.

7. Edwards FH, Cohen DJ, O’Brien SM, Peterson ED, Mack MJ, Shahian DM, Grover FL, Tuzcu EM, Thourani VH, Carroll J, Brennan JM, Brindis RG, Rumsfeld J, Holmes DR Jr; Steering Committee of the Society of Thoracic Surgeons/AmERICAN College of Cardiology Transcatheter Valve Therapy R. Development and validation of a risk prediction model for in-hospital mortality after transcatheter aortic valve replacement. JAMA Cardiol. 2016;1:46–52.

8. Duke University. Optimizing health outcomes in patients with symptomatic aortic valve disease (pcori-avr). National Institutes of Health. NCT number NCT02266251; 2014.

9. O’Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP; Society of Thoracic Surgeons Quality Measurement Task F. The society of thoracic surgeons 2008 cardiac surgery risk models: part 2–isolated valve surgery. Ann Thorac Surg. 2009;88:523–542.

10. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockwood AD, Euroscore ii. Eur J Cardiothorac Surg. 2012;41:734–744; discussion 744–753.

11. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg. 2008;135:180–187.

12. Wessler BS, Lai YN, Kramer W, Cangelosi M, Raman G, Lutz JS, Kent DM. Clinical prediction models for cardiovascular disease: tufts predictive analytics and comparative effectiveness clinical prediction model database. Circ Cardiovasc Qual Outcomes. 2015;8:368–375.

13. Van Calster B, Vickers AJ. Calibration of risk prediction models: impact on decision-analytic performance. Med Decis Making. 2015;35:162–169.

14. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol. 2015;68:25–34.
Clinical Prediction Models for VHD

15. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York, NY: Springer-Verlag; 2001.

16. Edwards FH, Peterson ED, Coombs LP, DeLong ER, Jamieson WR, Shroyer AL, Grover FL. Prediction of operative mortality after valve replacement surgery. J Am Coll Cardiol. 2001;37:885–892.

17. Nowicki ER, Birkmeyer NJ, Weintraub RW, Leavitt BJ, Sanders JH, Dacey LJ, Clough RA, Quinn RD, Charlesworth DC, Sisto DA, Uhlig PN, Olmedo EM, O’Connor GT. Northern New England Cardiovascular Disease Study Group and the Center for Evaluative Clinical Sciences, Dartmouth Medical School. Multivariable prediction of in-hospital mortality associated with aortic and mitral valve surgery in Northern New England. Ann Thorac Surg. 2004;77:1966–1977.

18. Kuduvalli M, Grayson AD, Au J, Grotte G, Bridgewater B, Fabri BM, Fabri BM. DOI: 10.1161/JAHA.119.011972

19. Harrell FE.

20. Monin JL, Lancellotti P, Monchi M, Lim P, Weiss E, Pierard L, Gueret P. Risk of perioperative mortality and long-term outcomes of percutaneous mitral valvuloplasty: a multifactorial score. Am J Med. 2009;122:581.e511–581.e519.

21. Monin JL, Lancellotti P, Monchi M, Lim P, Weiss E, Pierard L, Gueret P. Risk score for predicting outcome in patients with asymptomatic aortic stenosis. Circulation. 2009;120:69–75.

22. Guaragna JC, Bodanese LC, Bueno FL, Goldani MA. Proposed preoperative risk score for patients candidate to cardiac valve surgery. Arq Bras Cardiol. 2010;94:541–548.

23. Guo LX, Meng X, Zhang ZG, Bai T. Analysis of risk factors for valve replacements in 5,128 cases from a single heart center in China. Chin Med J (Engl). 2010;123:3509–3514.

24. Elmariah S, Lubitz SA, Shah AM, Miller MA, Kaplish D, Kothari S, Moreno PR, Gabrielli A, Elmariah S. Risk model and long-term outcomes of percutaneous mitral valve replacement: preliminary results from a series of 912 patients. Circulation. 2012;125:2119–2127.

25. Cioffi G, Cramarucic D, Dalsgaard M, Davidsen ES, Egstrup K, Rossebo AB, de Simone G, Gerds T. Left atrial systolic force and outcome in asymptomatic mitral stenosis. Echocardiography. 2012;29:1038–1044.

26. Holme I, Pedersen TR, Boman K, Egstrup K, Rossebo AB, de Simone G, Gerds T. An effective survival post TAVI function model for prolonged ventilation after adult heart valve surgery in a Chinese single center. Heart Lung. 2013;42:13–18.

27. Zheng Z, Fan H, Gao H, Li X, Yuan X, Meng J, Xu J, Song Y, Sun H, Hu S. Mortality risk model for heart valve surgery in China. J Heart Valve Dis. 2013;22:93–101.

28. Rankin JS, He X, O’Brien SM, Jacobs JP, Welke KF, Filardo G, Shahian DM. The society of thoracic surgeons risk model for operative mortality after multiple valve surgery. Ann Thorac Surg. 2013;95:1484–1490.

29. van der Kley F, Katsanos S, Joyce E, Tamborini G, Muratori M, Gripari P, Bax JJ. Score for patients candidate to cardiac valve surgery. Heart Lung. 2014;43:97–101.

30. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594.

31. van der Windt DA, Hayden JA, Perel P, Schroter S, Steyerberg EW, Moons KG, van der Wijden P, Thourani V, Green P, Rodes-Cabau J, Beohar N, Mack MJ, Leon MB, Cohen DJ; OBSERVANT Research Group. A simple risk tool (the observant score) for prediction of 30-day mortality after transcatheter aortic valve replacement. Eur J Cardiothorac Surg. 2011;39:815–821.

32. Rosenhek R, Iung B, Tornos P, Antunes MJ, Prendergast BD, Otto CM, De Bacquer D, Zuberbichler M, et al. Score for predicting in-hospital/30-day mortality for patients undergoing valve and valve/coronary artery bypass graft surgery. Ann Thorac Surg. 2013;95:1282–1290.

33. Wang C, Zhang GX, Lu FL, Li BL, Zou LJ, Han L, Xu Z. A local risk prediction model for prolonged ventilation after adult heart valve surgery in a Chinese single center. Eur Heart J. 2012;33:822–828,828a,828b.

34. Streyberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for developing and evaluating novel measures. Epidemiology. 2010;21:128–138.

35. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594.

36. Streyberg EW, Moons KG, van der Wijden P, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG; Group P. Prognosis research strategy (progress) 3: prognostic model research. PLoS Med. 2013;10:e1001381.

37. Streyberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York, NY: Springer; 2009.

38. Celermajer DS, Chow CK, Marijon E, Anstey NM, Woo KS. Cardiovascular disease in the developing world: prevalences, patterns, and the potential of early disease detection. J Am Coll Cardiol. 2012;60:1207–1216.

39. Baumgartner H, Falk V, Bax JJ, De Bonis M, Hamm C, Holm PJ, Iung B, Lancellotti P, Lansac E, Munoz DR, Rosenhek R, Sjogren J, Mas PT, Vahanian A, Wolter T, Wendler O, Windecker S, Zamorano JL. 2017 ESC/EACTS guidelines for the management of valvular heart disease. Rev Esp Cardiol (Engl Ed). 2018;71:110.

40. Aggarwal SK, Kasula S, Edupuganti MM, Raina S, Shahiels F, Almomani A, Payne JJ, Pothineni NV, Uretsky BF, Hakeem A. Clinical decision making for the hemodynamic “Gray Zone” (PFR 0.75-0.80) and long-term outcomes. J Invasive Cardiol. 2017;29:371–376.
SUPPLEMENTAL MATERIAL
Table S1. Search strategy used to identify all CVD/Cerebrovascular and VHD-specific CPMs.

| All CPM Search Terms                                                                                       | VHD CPM Search Terms                                                                 |
|-----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| ((predict$ adj1 model$) or (predict$ adj1 instrument$) or (predict$ adj1 score$) or (predict$ adj1 index)).mp. | heart valve prosthe$.mp. or exp cardiovascular surg$.mp. or exp valve replacement/ or exp valve intervention/ or exp aortic valve/ or exp mitral valve/ or exp tricuspid valve/ or exp pulmonic valve/ |
| ((prognos$ adj1 model$) or (prognos$ adj1 instrument$) or (prognos$ adj1 score$) or (prognos$ adj1 index)).mp. |                                                                                     |
| ((risk adj1 model$) or (risk adj1 instrument$) or (risk adj1 score$) or (risk adj1 index) or (risk assessment model or risk assessment instrument or risk assessment score)).mp. |                                                                                     |
| atrial fib$.mp. or exp Atrial Fibrillation/ or exp coronary artery disease/ or exp coronary disease/ or exp myocardial infarction/ or Myocardial infarct$.mp. or exp Heart Failure, Congestive/ or exp myocardial ischemia/ or exp cardiovascular diseases/ or exp Cerebrovascular Accident/ or *heart failure/ or *stroke/ or *acute coronary syndrome/ |                                                                                     |
| limit 6 to yr="1990 -Current" **Where current = May 15, 2012 publicationsstärkt**                                      |                                                                                     |
| (201205$ or 201206$ or 201207$ or 201208$ or 201209$ or 201210$ or 201211$ or 201212$ or 2013$ or 2014$ or 201501$ or 201502$ or 201503$).ed. |                                                                                     |
Table S2. De Novo Model Unabridged Overview.

| Author, Model Name | Year | Valve                          | Specific Intervention | Standardized Type of Intervention | Outcom e | Cente rs | Samp le Size | Numb er of Event s | EP V | Age | Continent | Model Method | C- statistic | Calibrati on Measure | Externally Validated? |
|-------------------|------|-------------------------------|-----------------------|-----------------------------------|----------|---------|-------------|-------------------|------|-----|-----------|---------------|---------------|----------------------|-----------------------|
| Edwards¹, STS (original) -- Isolated Valve | 2001 | Aortic/Mitral Surgery | AVR or MVR | 30 Day Operative Mortality | Multi-center | 4907 | 2291 | 121 | 63.7 | North America | Logistic regression | 0.766 | HL statistic, Calibration plot | Yes |
| Nowicki², NNE Aortic and Mitral Models | 2004 | Aortic Surgery | AVR +/- CABG | In-Hospital Mortality | Multi-center | 5793 | 360 | 40 | 69.5 | North America | Logistic regression | 0.75 | HL statistic | Yes |
|                     |      | Mitral Surgery | MVR or MVRRepair | In-Hospital Mortality | Multi-center | 3150 | 296 | 30 | 66.7 | North America | Logistic regression | 0.79 | HL statistic | Yes |
| Kuduvalli³, NWQIP  | 2007 | Aortic Surgery | AVR +/- CABG | In-Hospital Mortality | Multi-center | 4550 | 207 | 21 | 69 | Europe | Logistic regression | 0.78 | HL statistic | Yes |
| Cruz-Gonzalez⁴, PMV Score | 2009 | Mitral Percutaneous | Percutaneous Mitral Valvuoplast y | Procedural Success | Single center | 800 | 544 | 91 | 54.9 | North America | Logistic regression | NR | HL statistic | Yes |
| Monin⁵           | 2009 | Aortic Stenosis Natural History | None | Composi te (Non-MACE) | Single center | 107 | 62 | 9 | 72 | Europe | Logistic regression | 0.90 | HL statistic | Yes |
| O’Brien⁶, STS (new) -- Composite AEs | 2009 | Aortic/Mitral Surgery | AVR or MVR or MVRRepair | Composi te (Non-MACE) | Multi-center | 1097 | 20074 | 574 | NR | North America | Logistic regression | 0.721 | None | Yes |
| O’Brien⁶, STS (new) -- DSWI | 2009 | Aortic/Mitral Surgery | AVR or MVR or MVRRepair | DSWI | Multi-center | 1097 | 307 | 26 | NR | North America | Logistic regression | 0.704 | None | Yes |
| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centers | Sample Size | Number of Events | EPV | Age | Continent | Model Method | C-statistic | Calibration Measure | Externally Validated? |
|-------------------|------|-------|-----------------------------------|-----------------------|---------|---------|-------------|-----------------|-----|-----|------------|----------------|-------------|---------------------|----------------------|
| O’Brien⁶, STS (new) -- Mortality | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | 30 Day Mortality | Multi-center | 1097 59 | 3706 | 100 | NR | North America | Logistic regression | 0.805 | None | Yes |
| O’Brien⁶, STS (new) -- Prolonged LOS | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | Prolonged LOS | Multi-center | 1097 59 | 9718 | 270 | NR | North America | Logistic regression | 0.77 | None | Yes |
| O’Brien⁶, STS (new) -- Prolonged Ventilation | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | Prolonged Ventilation | Multi-center | 1097 59 | 12892 | 331 | NR | North America | Logistic regression | 0.77 | None | Yes |
| O’Brien⁶, STS (new) -- Renal Failure | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | Renal Failure | Multi-center | 1070 60 | 4673 | 173 | NR | North America | Logistic regression | 0.782 | None | Yes |
| O’Brien⁶, STS (new) -- Reoperation | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | Reoperation | Multi-center | 1097 59 | 9164 | 305 | NR | North America | Logistic regression | 0.643 | None | Yes |
| O’Brien⁶, STS (new) -- Short LOS | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | Prolonged LOS | Multi-center | 1097 59 | 41214 | 1178 | NR | North America | Logistic regression | 0.738 | None | No |
| O’Brien⁶, STS (new) -- Stroke | 2009 | Aortic/Mitral | Surgery | AVR or MVR or MVRRepair | Stroke | Multi-center | 1097 59 | 1751 | 92 | NR | North America | Logistic regression | 0.694 | None | Yes |
| Guaragna⁷, GuaragnaSCORE | 2010 | Aortic/Mitral | Surgery | Isolated Cardiac Valve Surgery +/- CABG | In-Hospital Mortality | Single center | 699 | 128 | 16 | 55.5 | South America | Logistic regression, Score | 0.82 | HL statistic, Calibration plot | Yes |
| Guo⁸ | 2010 | Aortic | Surgery | AVR | In-Hospital Mortality | Single center | 1087 | 45 | 6 | 60.24 | Asia | Logistic regression | NR | HL statistic | No |
| Mitral | 2010 | Surgery | MVR | In-Hospital Mortality | Single center | 1752 | 79 | 16 | 50.89 | Asia | Logistic regression | NR | HL statistic | No |
| Elmariah⁹, CRRAC the AV Score | 2011 | Aortic | Percutaneous | Balloon Aortic Valvuloplasty | 30 Day Mortality | Single center | 281 | 36 | 9 | 83 | North America | Cox regression, Score | 0.754 | HL statistic | No |
| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centers | Sample Size | Number of Events | EP V | Age | Continen t | Model Method | C-statistic | Calibration Measure | Externally Validated? |
|-------------------|------|-------|----------------------------------|-----------------------|---------|---------|-------------|-----------------|------|-----|------------|--------------|-------------|---------------------|---------------------|
| Bouleti10         | 2012 | Mitral | Percutaneous Mitral Commissurotomy | Percutaneous Mitral Commissurotomy | Composi (MACE) | Single center | 609 | 309 | 77 | 49 | Europe | Cox regression, Score | 0.74 | Calibration plot | No |
| Cioffi11          | 2012 | Aortic Stenosis | Natural History | None | Composi (MACE) | Multi-center | 1566 | 550 | 79 | 67 | Europe | Cox regression, Score | NR | None | No |
| Holme12, SEAS Score | 2013 | Aortic Stenosis | Natural History | None | 5 Year Mortality | Multi-center | 772 | 78 | 11 | 67.7 | Europe | Cox regression | 0.722 | HL statistic, Calibration plot, Brier score | No |
| Köttting13, German Aortic Valve Score | 2013 | Aortic | Percutaneous | AVR or TAVR | In-Hospital Mortality | Multi-center | 1114 | 416 | 28 | NR | Europe | Logistic regression, Score | 0.808 | HL statistic | Yes |
| Arnold14, 6 Month and 1 Year Models | 2014 | Aortic Stenosis | Percutaneous | TAVR | Composi (Non-MACE) | Multi-center | 2137 | 704 | 70 | 84 | North America, Europe | Logistic regression | 0.66 | HL statistics, Calibration plot | Yes |
|                       | 2014 | Aortic Stenosis | Percutaneous | TAVR | Composi (Non-MACE) | Multi-center | 2130 | 1073 | 134 | 84 | North America, Europe | Logistic regression | 0.66 | HL statistics, Calibration plot | Yes |
| Capodanno15, OBSERVANT T Score | 2014 | Aortic Stenosis | Percutaneous | AVR or TAVR | 30 Day Mortality | Multi-center | 1256 | 77 | 11 | 81.9 | Europe | Logistic regression, Score | 0.73 | HL statistic, Calibration plot, Brier score | Yes |
| D'Ascenzo16, Survival Post-TAVI (STT) - 30 Days and 1 Year Models | 2014 | Aortic | Percutaneous | TAVR | 30 Day Mortality | Multi-center | 1064 | 60 | 20 | 81.6 | Europe | Logistic regression | 0.66 | HL statistic | Yes |
|                       | 2014 | Aortic | Percutaneous | TAVR | 1 Year Mortality | Multi-center | 1064 | 165 | 55 | 81.6 | Europe | Logistic regression | 0.68 | HL statistic | Yes |
| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centers | Sample Size | Number of Events | EP V | Age | Continent | Model Method | C-statistic | Calibrated Measure | Externally Validated? |
|--------------------|------|-------|-----------------------------------|-----------------------|---------|---------|-------------|-----------------|------|-----|-----------|--------------|-------------|--------------------|----------------------|
| Jung17             | 2014 | Aortic | Percutaneous                      | TAVR                  | 30 Day Mortality | Multi-center | 2552        | 253             | 28   | 82.9| Europe    | Logistic regression, Score | 0.67        | HL statistic, Calibration in the large, Calibration plot | No                  |
| Debonnaire18, TAVI2-SCORE | 2015 | Aortic | Percutaneous                      | TAVR                  | 1 Year Mortality | Multi-center | 509         | 80              | 10   | 82  | Europe    | Cox regression, Score | 0.715       | HL statistic, Calibration in the large | Yes                |
| Edwards19          | 2016 | Aortic | Percutaneous                      | TAVR                  | In-Hospital Mortality | Multi-center | 1367        | 730             | 104  | 82.1| North America | Logistic regression | 0.67        | HL statistic, Calibration in the large, Calibration plot | Yes                |

**Isolated or Multiple Valve (± CABG)**

| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centers | Sample Size | Number of Events | EP V | Age | Continent | Model Method | C-statistic | Calibrated Measure | Externally Validated? |
|--------------------|------|-------|-----------------------------------|-----------------------|---------|---------|-------------|-----------------|------|-----|-----------|--------------|-------------|--------------------|----------------------|
| Koplan20           | 2003 | All   | Surgery                           | Cardiac Valve Surgery | Pacemaker Placement | Single center | 3116        | 168             | 24   | 65  | North America | Logistic regression, Score | NR         | None               | No                   |
| Ambler21, Ambler   | 2005 | Aortic, Mitral | Surgery                       | AVR and/or MVR       | In-Hospital Mortality | Multi-center | 1667         | 1067            | 76   | 64.8| Europe     | Logistic regression, Score | 0.77        | HL statistic, Calibration in plot | Yes                |
| Xu22               | 2006 | All   | Surgery                           | Valve Surgery        | Prolonged LOS       | Single center | 507          | 75              | 11   | 55.5| Asia       | Logistic regression | 0.81        | Calibration in table form | No                   |
| Hannan23           | 2007 | Aortic, Mitral | Surgery                        | Isolated Valve Surgery | In-Hospital Mortality | Multi-center | 1070         | 472             | 43   | NR  | North America | Logistic regression, Score | 0.794      | HL statistic, Calibration in plot | Yes                |
| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centres | Sample Size | Number of Events | EP | Age | Continent | Model Method | C-statistic | Calibration Measure | Externally Validated? |
|-------------------|------|-------|-----------------------------------|-----------------------|---------|---------|------------|----------------|-----|-----|------------|--------------|------------|---------------------|----------------------|
| Xu\(^{24}\), FUWAI Score | 2007 | All | Surgery | Valve Surgery +/- CABG | Prolonged LOS | Single center | 2193 | 345 | 27 | 53.29 | Asia | Logistic regression, Score | 0.76 | HL statistic, Calibration plot | Yes |
| Shi\(^{25}\) | 2010 | Aortic, Mitral | Surgery | AVR and/or MVR | In-Hospital Mortality | Single center | 158 | 8 | 1 | NR | Asia | Logistic regression | 0.7358 | None | No |
| Ariyaratne\(^{26}\), Aus-AVR Score | 2011 | Aortic, Mitral | Surgery | AVR +/- CABG +/- MVR | 30 Day Mortality | Multi-center | 3544 | 147 | 16 | NR | Australia | Logistic regression, Score | 0.78 | HL statistic, Calibration plot | Yes |
| Nashef\(^{27}\), EuroSCORE II | 2012 | All | Surgery | Major Cardiac Surgery | In-Hospital Mortality | Multi-center | 1682 | 8 | 656 | 36 | 64.6 | International | Logistic regression | 0.8095 | None | Yes |
| Hannan\(^{28}\), NY Operative Mortality Risk Score | 2013 | Aortic, Mitral | Surgery | Isolated Valve Surgery | 30 Day Mortality | Multi-center | 1345 | 5 | 542 | 49 | NR | North America | Logistic regression, Score | 0.781 | HL statistic | Yes |
| Wang\(^{29}\) | 2013 | All | Surgery | Valve Surgery | Prolonged Ventilation | Single center | 2400 | 303 | 25 | NR | Asia | Logistic regression | 0.789 | HL statistic | No |
| Zheng\(^{30}\) | 2013 | Aortic, Mitral | Surgery | AV and/or MV Surgery | In-Hospital Mortality | Multi-center | 6677 | 130 | 26 | 48 | Asia | Logistic regression, Score | 0.76 | HL statistic, Chi-square statistic, Calibration plot | No |

**Multiple Valve**

| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centres | Sample Size | Number of Events | EP | Age | Continent | Model Method | C-statistic | Calibration Measure | Externally Validated? |
|-------------------|------|-------|-----------------------------------|-----------------------|---------|---------|------------|----------------|-----|-----|------------|--------------|------------|---------------------|----------------------|
| Guo\(^{8}\) | 2010 | Aortic, Mitral | Surgery | AVR + MVR | In-Hospital Mortality | Single center | 818 | 55 | 14 | 58.4 | Asia | Logistic regression | NR | HL statistic | No |
| Rankin\(^{31}\), AM Preop | 2013 | Aortic, Mitral | Surgery | AVR + MVR | 30 Day Mortality | Multi-center | 2703 | 5 | 2541 | 116 | 70 | North America | Logistic regression | NR | Calibration plot | Yes |
| Author, Model Name | Year | Valve | Standardized Type of Intervention | Specific Intervention | Outcome | Centers | Sample Size | Number of Events | EPV | Age | Continent | Model Method | C-statistic Measure | Categorically Validated? |
|--------------------|------|-------|-----------------------------------|-----------------------|---------|---------|------------|----------------|------|-----|-----------|----------------|-----------------------|-------------------------|
| Rankin³¹, MT Preop | 2013 | Mitral, Tricuspid | Surgery | MVR + TVR | 30 Day Mortality | Multi-center | 1868 6 | 1420 | 71 | 70 | North America | Logistic regression | NR | Calibration plot | Yes |
| Rankin³¹, AMT Preop | 2013 | Aortic, Mitral, Tricuspid | Surgery | AVR + MVR + TVR | 30 Day Mortality | Multi-center | 4510 | 591 | 74 | 71 | North America | Logistic regression | NR | Calibration plot | Yes |
| Rankin³¹, AM Preop + Intraop | 2013 | Aortic, Mitral | Surgery | AVR + MVR | 30 Day Mortality | Multi-center | 2703 5 | 2541 | 110 | 70 | North America | Logistic regression | NR | Calibration plot | Yes |
| Rankin³¹, MT Preop + Intraop | 2013 | Mitral, Tricuspid | Surgery | MVR + TVR | 30 Day Mortality | Multi-center | 1868 6 | 1420 | 71 | 70 | North America | Logistic regression | NR | Calibration plot | Yes |
| Rankin³¹, AMT Preop + Intraop | 2013 | Aortic, Mitral, Tricuspid | Surgery | AVR + MVR + TVR | 30 Day Mortality | Multi-center | 4510 | 591 | 26 | 71 | North America | Logistic regression | NR | Calibration plot | Yes |

* ‘Isolated Valve’ indicates a single valve procedure; ‘Multiple Valve’ indicates intervention to > 1 valve; AVR indicates aortic valve replacement; MVR, mitral valve replacement; HL, Hosmer-Lemeshow; CABG, coronary artery bypass grafting; NR, not reported; MACE, major adverse cardiovascular events; AEs, adverse events; DSWI, deep sternal wound infections; LOS, length of stay; TAVR, transcatheter aortic valve replacement; TVR, tricuspid valve replacement.
Table S3. External Validations Overview.

| De novo          | Validation                        | Author, Year | Index Condition | Outcome                        | Timeframe         | Continent             | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|------------------|-----------------------------------|--------------|----------------|--------------------------------|-------------------|-----------------------|-------------|-------------------|-------------|-----------------------|-------------------------------|
| STS (original) – | Isolated Valve                    | Edwards, 2001¹ | AVR or MVR     | operative mortality           | 30 days           | North America         | 25640       | 1231              | 0.773       | Yes                  | 2.631578947                   |
| Isolated Valve   |                                    | Brown, 2009² | AVR            | mortality                     | hospitalization   | North America         | 108,687     | 3197              | NA          | Yes                  | NA                           |
| NNE Aortic²      |                                   | Jin, 2005³   | AVR            | operative mortality           | until discharge   | North America         | 3324        | 142               | 0.75        | Yes                  | 0                             |
|                  |                                   | Ariyaratne, 2011² | AVR       | early postop mortality        | 30 days           | Australia             | 3544        | 147               | 0.77        | Yes                  | 8                             |
|                  |                                   | Wang, 2013³⁴ | valve surgery   | in-hospital mortality         | until discharge   | Asia                  | 12412       | 260               | NA          | Yes                  | -17.6                        |
| NNE Mitral²      |                                   | Jin, 2005³   | MVR             | operative mortality           | until discharge   | North America         | 1596        | 95                | 0.81        | Yes                  | 6.896551724                   |
| NWQIP³           |                                   | Kuduvalli, 2007³ | AVR +/-  | in-hospital mortality         | hospitalization   | Europe                | 816         | 33                | 0.78        | Yes                  | 0                             |
|                  |                                   | Ariyaratne, 2011² | AVR +/- |
|                  |                                   |              | CABG            | in-hospital mortality or mortality within 30 days of surgery | 30 days           | Australia             | 3306        | 120               | 0.77        | Yes                  | 3.571428571                   |
| PMV Score⁴       |                                   | Cruz-Gonzalez, 2009⁴ | percutaneous mitral valvuloplasty | PMV success       | 1 year               | North America         | 285         | 213               | NR          | No                   | NA                           |
|                  | Monin, 2009⁵                      | Monin, 2009³ | valve surgery   | morbidity/mortality           | mean 21 months    | Europe                | 107         | 56                | 0.89        | Yes                  | -2.5                         |
| STS (new) –      | Composite AEs³                   | Watanabe, 2013³⁵ | TAVR         | composite safety endpoint     | 30 days           | Europe                | 453         | 94                | 0.59        | No                   | -59.2760181                   |
| STS (new) –      | DSWI⁶                            | Wang, 2016³⁶ | AVR +/-  | composite morbidity           | 30 days           | Australia             | 450         | 152               | 0.627       | No                   | 42.53393665                   |
|                  |                                   | Wang, 2017³⁷ | MVR or MVRepair | composite morbidity          | 30 days           | Australia             | 407         | 77                | 0.732       | No                   | 4.977375566                   |
|                  |                                   | Wang, 2016³⁶ | AVR +/-  | DSWI                          | 30 days           | Australia             | 450         | 6                 | 0.631       | No                   | 35.78431373                   |
|                  |                                   | Wang, 2017³⁷ | MVR or MVRepair | mediastinitis            | 30 days           | Australia             | 407         | 4                 | 0.721       | No                   | 8.333333333                   |
| De novo | Validation |
|---------|------------|
| **Model** | **Author, Year** | **Index Condition** | **Outcome** | **Timeframe** | **Continent** | **Sample Size** | **Number of Events** | **C-statistic** | **Calibration Reported?** | **Change in Discrimination (%)** |
| | Piazza, 2010 | TAVR | periprocedural mortality | 30 days | Europe | 168 | 19 | 0.69 | Yes | 37.70491803 |
| | Basraon, 2011 | AVR | perioperative mortality | 30 days | North America | 537 | 32 | 0.73 | Yes | 24.59016393 |
| | Zhang, 2011 | valve surgery | prolonged postop ICU stay | mean LOS 79.44 +/- 59.76 hrs | Asia | 1333 | 187 | 0.70 | Yes | 34.42622951 |
| | Durand, 2013 | TAVR | mortality | 30 days | Europe | 250 | 19 | 0.58 | Yes | -73.7704918 |
| | Durand, 2013 | TAVR (transapical access) | mortality | 30 days | Europe | 60 | 7 | 0.55 | Yes | 83.60655738 |
| | Durand, 2013 | TAVR (transfemoral access) | mortality | 30 days | Europe | 190 | 12 | 0.66 | Yes | 47.54098361 |
| | Haensig, 2013 | TA-AVI | mortality | 30 days | Europe | 360 | 38 | 0.64 | Yes | 54.09836066 |
| | Haensig, 2013 | TA-AVI | in-hospital mortality | until discharge | Europe | 360 | 41 | 0.65 | Yes | 50.81967213 |
| | Laurent, 2013 | AVR | operative mortality | 30 days | Europe | 314 | 18 | 0.77 | Yes | 11.47540984 |
| | Wang, 2013 | valve surgery | in-hospital mortality | until discharge | Asia | 12412 | 260 | 0.735 | Yes | 22.95081967 |
| | Watanabe, 2013 | TAVR | mortality | 30 days | Europe | 453 | 57 | 0.6 | Yes | 67.21311475 |
| | Watanabe, 2013 | TAVR (transfemoral) | mortality | 30 days | Europe | 249 | 28 | 0.6 | Yes | 67.21311475 |
| | Watanabe, 2013 | TAVR (transfemoral approach, without early experience) | mortality | 30 days | Europe | NR | NR | 0.65 | Yes | 50.81967213 |
| | Watanabe, 2013 | TAVR (transapical/transaortic) | mortality | 30 days | Europe | 330 | 27 | 0.61 | Yes | 63.93442623 |
| | Barili, 2014 | MV surgery | in-hospital mortality | until discharge | Europe | 1239 | NR | 0.82 | Yes | 4.918032787 |
| | Barili, 2014 | MV surgery +/- CABG | in-hospital mortality | until discharge | Europe | 2202 | NR | 0.76 | Yes | 14.75409836 |
| De novo | Validation |
|---------|------------|
| Author, Year | Model | Condition | Outcome | Timeframe | Continent | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
| Beohar, 2014 | TAVR | mortality | 30 days | North America | 2552 | 165 | 0.6 | Yes | 67.21311475 |
| Chan, 2014 | MVR or MVRepair | perioperative mortality | 30 days | North America | 1154 | 11 | 0.74 | Yes | 21.31147541 |
| Osnabrugge, 2014 | AVR | in-hospital mortality | until discharge | North America | 4107 | 119 | 0.74 | Yes | 21.31147541 |
| Osnabrugge, 2014 | AVR +/- CABG | in-hospital mortality | until discharge | North America | 3480 | 143 | 0.74 | Yes | 21.31147541 |
| Osnabrugge, 2014 | MVRepair | in-hospital mortality | until discharge | North America | 1059 | 13 | 0.86 | Yes | 18.03278689 |
| Osnabrugge, 2014 | MVR | in-hospital mortality | until discharge | North America | 1071 | 59 | 0.79 | Yes | 4.918032787 |
| Rabbani, 2014 | valve replacement surgery | mortality | 30 days | Asia | 576 | 28 | 0.812 | Yes | 2.295081967 |
| Wendt, 2014 | AVR or TAVR | mortality | 30 days | Europe | 1512 | 95 | 0.708 | Yes | 31.80327869 |
| Wang, 2014 | isolated or multiple valve surgery | in-hospital mortality | until discharge | Asia | 9846 | 176 | 0.712 | Yes | 30.49180328 |
| Adamo, 2015 | percutaneous MVRepair | mortality | 30 days | Europe | 304 | 10 | 0.62 | Yes | -60.6557377 |
| Debonnaire, 2015 | TAVR | all-cause mortality | 1 year | Europe | 471 | 80 | 0.5 | Yes | -100 |
| Holinski, 2015 | repeat AVR | mortality | 30 days | Europe | 78 | 8 | 0.64 | Yes | 54.09836066 |
| Silaschi, 2015 | TAVR (transfemoral or transapical) | mortality | 30 days | Europe | 457 | 44 | 0.57 | Yes | 77.04918033 |
| Silva, 2015 | TAVR | mortality | 30 days | South America | 418 | 38 | 0.54 | Yes | -86.8852459 |
| Sinning, 2015 | TAVR | all-cause mortality | 1 year | Europe | 310 | 80 | 0.685 | No | -39.3442623 |
| Tralhao, 2015 | AVR | operative mortality | 30 days | Europe | 106 | 6 | 0.702 | Yes | -33.7704918 |
| Vassileva, 2015 | repeat AVR after prior CABG | operative mortality | median LOS 6 days (IQR 5-9) | North America | 6534 | 236 | NR | Yes | NA |
| De novo | Validation |
|---------|------------|
| **Model** | **Author, Year** | **Index Condition** | **Outcome** | **Timeframe** | **Continent** | **Sample Size** | **Number of Events** | **C-statistic** | **Calibration Reported?** | **Change in Discrimination (%)** |
| Wang, 2015 | AVR | operative mortality | 30 days | Australia | 620 | 18 | 0.716 | Yes | - | 29.18032787 |
| Wang, 2015 | AVR | post-operative complications | 14 days | Australia | 620 | 115 | 0.666 | Yes | - | 45.57377049 |
| Barili, 2016 | AVR | mortality | 30 days | Europe | 1444 | NR | 0.79 | Yes | - | 4.918032787 |
| Collas, 2016 | TAVR | mortality | 1 year | Europe | 225 | 38 | NR | No | NA |
| Halkin, 2016 | TAVR | all-cause mortality | 30 days | Asia | 1327 | 45 | 0.68 | Yes | - | 40.98360656 |
| Kortlandt, 2016 | percutaneous MVRepair | periprocedural mortality | 30 days | Europe | 136 | 5 | 0.65 | Yes | - | 50.81967213 |
| Peguero, 2016 | cardiac surgery | operative mortality | 30 days | North America | 2263 | 48 | 0.77 | Yes | - | 11.47540984 |
| Rosa, 2016 | TAVR | in-hospital mortality | until discharge | South America | 59 | 6 | NR | Yes | NA |
| Rosa, 2016 | TAVR | 30-day mortality | 30 days | South America | 59 | 8 | 0.81 | Yes | - | 1.639344262 |
| Wang, 2016 | AVR +/- CABG | operative mortality | 30 days | Australia | 450 | 29 | 0.699 | Yes | - | 34.75409836 |
| Wang, 2016 | valve surgery | in-hospital mortality | until discharge | Asia | 12412 | 260 | 0.735 | Yes | - | 22.95081967 |
| Yamaoka, 2016 | AVR +/- CABG | operative mortality | 30 days | Asia | 406 | 14 | 0.781 | Yes | - | 7.868852459 |
| Zbroński, 2016 | TAVR | mortality | 30 days | Europe | 156 | 15 | 0.55 | Yes | - | 83.60655738 |
| Balan, 2017 | TAVR | mortality | 30 days | North America | 426 | 18 | 0.674 | No | - | 42.95081967 |
| Balan, 2017 | SAVR | mortality | 30 days | North America | 297 | 14 | 0.791 | No | - | 4.590163934 |
| Balan, 2017 | TAVR (transfemoral) | mortality | 30 days | North America | NR | NR | 0.789 | No | - | 5.245901639 |
| Balan, 2017 | TAVR (transapical) | mortality | 30 days | North America | NR | NR | 0.583 | No | - | 72.78688525 |
| Schmid, 2017 | TAVR | all-cause mortality | 1 year | Europe | 74 | 10 | 0.734 | No | - | 23.27868852 |
| Schmid, 2017 | TAVR | all-cause mortality | 2 years | Europe | 74 | 18 | 0.646 | No | - | 52.13114754 |
| Wang, 2017 | MVR or MVRepair | operative mortality | 30 days | Australia | 407 | 10 | 0.850 | Yes | - | 14.75409836 |
| Model                          | Author, Year | Index Condition | Outcome | Timeframe | Continent | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|-------------------------------|--------------|----------------|----------|-----------|-----------|-------------|-----------------|-------------|----------------------|-------------------------------|
| STS (new)– Prolonged LOS⁶     | Wang, 2016³⁶ | AVR +/- CABG   | LOS > 14 days | 14 days   | Australia | 450         | 86              | 0.638       | No                   | -27.40740741                  |
|                              | Wang, 2017³⁷ | MVR or MVRepair| prolonged LOS | 14 days   | Australia | 407         | 56              | 0.696       | No                   | -27.40740741                  |
| STS (new)– Prolonged Ventilation⁶ | Wang, 2016³⁶ | AVR/CABG      | ventilation > 24 hours | 30 days   | Australia | 450         | 124             | 0.642       | No                   | -47.40740741                  |
|                              | Wang, 2017³⁷ | MVR or MVRepair| ventilation > 24 hours | 30 days   | Australia | 407         | 54              | 0.789       | No                   | 7.037037037                   |
| STS (new)– Renal Failure⁶    | Peguero, 2016⁶⁴ | valve surgery +/- CABG | operative mortality | 30 days   | North America | 2263     | 48              | 0.76        | Yes                  | -7.80141844                   |
|                              | Wang, 2016³⁶ | AVR +/- CABG   | renal failure | 30 days   | Australia | 450         | 6               | 0.682       | No                   | 35.46999291                   |
|                              | Wang, 2017³⁷ | MVR or MVRepair| renal failure | 30 days   | Australia | 407         | 12              | 0.828       | No                   | 16.31205674                   |
| STS (new)– Reoperation⁶      | Wang, 2016³⁶ | AVR/CABG      | reoperation   | 30 days   | Australia | 450         | 54              | 0.612       | No                   | 21.6783168                    |
|                              | Wang, 2017³⁷ | MVR or MVRepair| return to theater | 30 days   | Australia | 407         | 33              | 0.668       | No                   | 17.48251748                   |
| STS (new)– Stroke⁶           | Peguero, 2016⁶⁴ | valve surgery +/- CABG | operative mortality | 30 days   | North America | 2263     | 48              | 0.69        | Yes                  | -2.06185567                   |
|                              | Wang, 2016³⁶ | AVR/CABG      | stroke        | 30 days   | Australia | 450         | 15              | 0.642       | No                   | -26.80412371                  |
|                              | Wang, 2017³⁷ | MVR or MVRepair| stroke        | 30 days   | Australia | 407         | 7               | 0.665       | No                   | -14.9485361                   |
| GuaragnaSCORE⁷               | Sa, 2012⁷¹  | valve surgery +/- CABG | perioperative mortality | until discharge | South America | 491      | 74              | 0.781       | Yes                  | -12.1875                     |
|                              | Silva, 2015⁵⁴ | TAVR         | mortality     | 30 days   | South America | 418     | 38              | 0.52        | Yes                  | -93.75                       |
| German Aortic Valve Score¹³  | Sinning, 2015⁵⁵ | TAVR       | all-cause mortality | 1 year   | Europe    | 310       | 80              | 0.661       | No                   | 47.72727273                   |
|                              | Sinning, 2015⁵⁵ | TAVR       | all-cause mortality or rehospitalization | 1 year   | Europe    | 310       | 132             | 0.618       | No                   | 61.68831169                   |
|                              | Halkin, 2016⁶² | TAVR       | all-cause mortality | 30 days   | Asia      | 1327     | 45              | 0.52        | No                   | 93.50649351                   |
|                              | Kalendar, 2017⁷² | AVR     | mortality     | until discharge | Asia    | 35     | 6               | 0.647       | Yes                  | 52.27272727                   |
|                              | Martin, 2017⁷³ | TAVR       | mortality     | 30 days   | Europe    | 6676     | 360             | 0.59        | Yes                  | 70.77922078                   |
| De novo | Validation |
|---------|------------|
| **Model** | **Author, Year** | **Index Condition** | **Outcome** | **Timeframe** | **Continent** | **Sample Size** | **Number of Events** | **C-statistic** | **Calibration Reported?** | **Change in Discrimination (%)** |
| De novo | Schmid, 2017<sup>70</sup> | TAVR | all-cause mortality | 1 year | Europe | 74 | 10 | 0.703 | No | -34.09090909 |
| De novo | Schmid, 2017<sup>70</sup> | TAVR | all-cause mortality | 2 years | Europe | 74 | 18 | 0.554 | No | -82.46753247 |
| Arnold – 6 Month Model<sup>14</sup> | Arnold, 2016<sup>74</sup> | TAVR | poor outcome | 6 months | North America | 2830 | 882 | 0.646 | Yes | -8.75 |
| Arnold – 1 Year Model<sup>14</sup> | Arnold, 2016<sup>74</sup> | TAVR | poor outcome | 1 year | North America | 2325 | 1181 | 0.653 | Yes | -4.375 |
| **OBSERVANT Score<sup>15</sup>** | Collas, 2016<sup>61</sup> | TAVR | mortality | 1 year | Europe | 225 | 38 | NR | No | NA |
| | Halkin, 2016<sup>62</sup> | TAVR | all-cause mortality | 30 days | Asia | 1327 | 45 | 0.63 | No | -43.47826087 |
| | Zbróński, 2016<sup>68</sup> | TAVR | mortality | 30 days | Europe | 156 | 15 | 0.597 | Yes | -57.82608696 |
| | Martin, 2017<sup>73</sup> | TAVR | mortality | 30 days | Europe | 6676 | 360 | 0.57 | Yes | 69.56521739 |
| **Survival Post-TAVI (STT) – 30 days<sup>16</sup>** | D’Ascenzo, 2014<sup>16</sup> | TAVR | mortality | 30 days | Europe | 180 | 13 | 0.66 | Yes | 0 |
| | Collas, 2016<sup>61</sup> | TAVR | mortality | 1 year | Europe | 225 | 38 | NR | No | NA |
| **Survival Post-TAVI (STT) – 1 year<sup>16</sup>** | D’Ascenzo, 2014<sup>16</sup> | TAVR | mortality | 1 year | Europe | 180 | 63 | 0.67 | Yes | 5.555555556 |
| **TAVI2-SCORE<sup>18</sup>** | Collas, 2016<sup>61</sup> | TAVR | mortality | 1 year | Europe | 225 | 38 | NR | No | NA |
| **Edwards, 2016<sup>19</sup>** | Edwards, 2016<sup>19</sup> | TAVR | in-hospital mortality | until discharge | North America | 6868 | 300 | 0.66 | Yes | -5.882352941 |

*Isolated or Multiple Valve*

| Ambler<sup>21</sup> | De Bacco, 2008<sup>75</sup> | Implantation of bovine pericardial bioprosthesis | in-hospital mortality | until discharge | South America | 703 | 101 | 0.729 | Yes | 15.18518519 |
| | Dewey, 2008<sup>76</sup> | AVR | mortality | mean 4.2 +/- 2.7 years | North America | 97 | 39 | NR | Yes | NA |
| | Tran, 2010<sup>77</sup> | AVR | mortality | 1 year | North America | 394 | 23 | 0.799 | Yes | 10.74074074 |
| | Laurent, 2013<sup>83</sup> | AVR | operative mortality | 30 days | Europe | 314 | 18 | 0.70 | Yes | 25.92592593 |
| | Wang, 2013<sup>34</sup> | Valve surgery | in-hospital mortality | until discharge | Asia | 3479 | 112 | 0.677 | Yes | 34.44444444 |
| Model | Author, Year | Index Condition | Outcome | Timeframe | Continent | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|-------|--------------|-----------------|---------|-----------|-----------|-------------|-----------------|-------------|-----------------------|-----------------------------|
|       | Silaschi, 201553 | TAVR (transfemoral or transapical) | mortality | 30 days | Europe | 457 | 44 | 0.52 | Yes | 92.59259259 |
|       | Silva, 201554 | TAVR | mortality | 30 days | South America | 418 | 38 | 0.57 | Yes | 74.07407407 |
|       | Wang, 201656 | valve surgery | in-hospital mortality | until discharge | Asia | 12412 | 260 | 0.674 | Yes | 35.55555556 |
|       | Yamaoka, 201657 | AVR +/- CABG | operative mortality | 30 days | Asia | 406 | 14 | 0.709 | Yes | 22.59259259 |
|       | Zbroński, 201658 | TAVR | mortality | 30 days | Europe | 156 | 15 | 0.54 | Yes | 85.18518519 |
|       | Hannan, 200723 | isolated valve surgery | in-hospital mortality | North America | 9662 | 504 | NR | No | NA |
|       | van Gameren, 200878 | isolated valve surgery | hospital mortality | hospitalization period | Europe | 904 | 25 | 0.86 | Yes | 22.44897959 |
|       | Wang, 201334 | valve surgery +/- CABG | in-hospital mortality | hospitalization period | Asia | 3479 | 112 | 0.682 | Yes | -38.0952381 |
|       | FUWAI Score24 | Zhang, 201140 | valve surgery | prolonged postop ICU stay | mean LOS 79.44 +/- 59.76 hrs | Asia | 1333 | 187 | 0.81 | Yes | 19.23076923 |
|       | Aus-AVR Score26 | Ariyaratne, 201126 | AVR | early postoperative mortality | 30 days | Australia | 3544 | 147 | 0.73 | Yes | 17.85714286 |
|       | Wang, 201558 | AVR | operative mortality | 30 days | Australia | 620 | 18 | 0.716 | Yes | 22.85714286 |
|       | Wang, 201559 | AVR | post-operative complications | 14 days | Australia | 620 | 115 | 0.618 | Yes | 57.85714286 |
|       | EuroSCORE II27 | Barili, 201379 | MV surgery | in-hospital mortality | until discharge | Europe | NR | NR | 0.79 | No | -6.300484653 |
|       | Barili, 201379 | AV, MV, or TV surgery | in-hospital mortality | until discharge | Europe | NR | NR | 0.8 | No | -3.069466882 |
|       | Carnero-Alcazar, 201380 | cardiac surgery | post-operative complications | 30 days | Europe | 3798 | 215 | 0.85 | Yes | 13.08562197 |
|       | Chalmers, 201381 | AVR | in-hospital mortality | until discharge | Europe | 814 | 19 | 0.69 | Yes | 38.61066236 |
|       | Chalmers, 201381 | AVR +/- CABG | in-hospital mortality | until discharge | Europe | 517 | 23 | 0.74 | Yes | 22.4557351 |
|       | Chalmers, 201381 | MV surgery | in-hospital mortality | until discharge | Europe | 340 | 5 | 0.87 | Yes | 19.54765751 |
| Model | Author, Year | Index Condition | Outcome | Timeframe | Continent | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|-------|--------------|----------------|---------|-----------|-----------|-------------|-----------------|-------------|----------------------|-------------------------------|
| Durand, 2013 | TAVR (all) | mortality | 30 days | Europe | 250 | 19 | 0.66 | Yes | 48.30371567 |
| Durand, 2013 | TAVR (transapical) | mortality | 30 days | Europe | 60 | 7 | 0.52 | Yes | 93.53796446 |
| Durand, 2013 | TAVR (transfemoral) | mortality | 30 days | Europe | 190 | 12 | 0.71 | Yes | 32.14862682 |
| Haensig, 2013 | TA-AVI in-hospital mortality | until discharge | Europe | 360 | 38 | 0.51 | Yes | 96.76898223 |
| Haensig, 2013 | TA-AVI mortality | 30 days | Europe | 360 | 41 | 0.50 | Yes | -100 |
| Howell, 2013 | valve surgery +/- CABG in-hospital mortality | until discharge | Europe | 933 | 90 | 0.67 | Yes | -45.0726979 |
| Sedaghat, 2013 | TAVR (transfemoral) | mortality | 30 days | Europe | 206 | 14 | 0.71 | Yes | 32.14862682 |
| Sedaghat, 2013 | TAVR (transfemoral) | mortality | 1 year | Europe | 206 | 56 | 0.70 | Yes | 35.37964459 |
| Wang, 2013 | valve surgery in-hospital mortality | until discharge | Asia | 12412 | 260 | 0.693 | Yes | 37.64135703 |
| Watanabe, 2013 | TAVR mortality | 30 days | Europe | 453 | 57 | 0.68 | Yes | 41.84168013 |
| Watanabe, 2013 | TAVR (transfemoral) | mortality | 30 days | Europe | 249 | 28 | 0.74 | Yes | 22.45557351 |
| Watanabe, 2013 | TAVR (transfemoral approach, without early experience) | mortality | 30 days | Europe | NR | NR | 0.75 | Yes | 19.22455574 |
| Watanabe, 2013 | TAVR (transapical/transaortic) | mortality | 30 days | Europe | 330 | 27 | 0.61 | Yes | 64.45880452 |
| Zhang, 2013 | isolated or multiple valve surgery in-hospital mortality | until discharge | Asia | 3479 | 112 | 0.69 | Yes | 40.22617124 |
| Zhang, 2013 | isolated valve surgery in-hospital mortality | until discharge | Asia | 1106 | 26 | 0.792 | Yes | 5.654281099 |
| Zhang, 2013 | multiple valve surgery in-hospital mortality | until discharge | Asia | 2373 | 86 | 0.605 | Yes | 66.07431341 |
| Barili, 2014 | isolated MV surgery in-hospital mortality | until discharge | Europe | 1239 | NR | 0.81 | Yes | 0.161550889 |
| Model | Author, Year | Index Condition | Outcome | Timeframe | Continent | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|-------|--------------|----------------|---------|-----------|-----------|-------------|-----------------|-------------|------------------------|-----------------------------|
| Barili, 2014<sup>44</sup> | associated MV surgery | in-hospital mortality | until discharge | Europe | NR | NR | 0.75 | Yes | 19.22455574 |
| Barili, 2014<sup>44</sup> | MV surgery +/- CABG | in-hospital mortality | until discharge | Europe | NR | NR | 0.74 | Yes | 22.45557351 |
| Barili, 2014<sup>85</sup> | elective major cardiac surgery | in-hospital mortality | until discharge | Europe | 12201 | 210 | 0.80 | Yes | 3.069466882 |
| Koszta, 2014<sup>86</sup> | major cardiac surgery | mortality | 30 days | Europe | 2287 | 123 | 0.8177 | Yes | 2.649434572 |
| Osnabrugge, 2014<sup>87</sup> | AVR | in-hospital mortality | until discharge | North America | 4107 | 119 | 0.71 | Yes | 32.14862682 |
| Osnabrugge, 2014<sup>87</sup> | AVR +/- CABG | in-hospital mortality | until discharge | North America | 3480 | 143 | 0.72 | Yes | 28.91760905 |
| Osnabrugge, 2014<sup>87</sup> | MVR | in-hospital mortality | until discharge | North America | 1059 | 13 | 0.82 | Yes | 3.392568659 |
| Osnabrugge, 2014<sup>87</sup> | MVR | in-hospital mortality | until discharge | North America | 1071 | 59 | 0.78 | Yes | 9.531502423 |
| Rabbani, 2014<sup>48</sup> | valve replacement surgery | mortality | 30 days | Asia | 576 | 28 | 0.816 | Yes | 2.100161551 |
| Rabbani, 2014<sup>48</sup> | MVR | mortality | 30 days | Asia | 247 | 7 | 0.898 | Yes | 28.59450727 |
| Rabbani, 2014<sup>48</sup> | AVR | mortality | 30 days | Asia | 137 | 4 | 0.747 | Yes | 20.19386107 |
| Rabbani, 2014<sup>48</sup> | DVR | mortality | 30 days | Asia | 86 | 2 | 0.637 | Yes | 55.73505654 |
| Rabbani, 2014<sup>48</sup> | MVR +/- CABG | mortality | 30 days | Asia | 57 | 11 | 0.773 | Yes | 11.79321486 |
| Rabbani, 2014<sup>48</sup> | AVR +/- CABG | mortality | 30 days | Asia | 49 | 4 | 0.521 | Yes | 93.21486268 |
| Spiliopoulos, 2014<sup>87</sup> | AVR +/- CABG | perioperative mortality | 30 days | Europe | 222 | 14 | 0.77 | Yes | 12.76252019 |
| Spiliopoulos, 2014<sup>87</sup> | AVR +/- CABG | late mortality | beyond 30 days | Europe | 202 | 21 | 0.718 | Yes | -29.5638126 |
| Wang, 2014<sup>88</sup> | valve surgery | in-hospital mortality | until discharge | Asia | 11170 | 226 | 0.72 | Yes | 28.91760905 |
| Wang, 2014<sup>88</sup> | isolated non-CABG surgery | in-hospital mortality | until discharge | Asia | 3696 | NR | 0.76 | Yes | 15.99353796 |
| Wang, 2014<sup>88</sup> | 2 procedures | in-hospital mortality | until discharge | Asia | 5006 | NR | 0.67 | Yes | -45.0726979 |
| Model | Author, Year | Index Condition | Outcome | Timeframe | Continent | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|-------|--------------|----------------|---------|-----------|-----------|-------------|----------------|-------------|----------------------|-------------------------------|
| Wang, 2014 | 3+ procedures | in-hospital mortality | until discharge | Asia | 2468 | NR | 0.73 | Yes | 25.68659128 |
| Wendt, 2014 | AVR or TAVR | mortality | 30 days | Europe | 1512 | 95 | 0.712 | Yes | 31.50242326 |
| Wendt, 2014 | TAVR (transfemoral) | mortality | 30 days | Europe | 291 | 34 | 0.554 | Yes | 82.55250404 |
| Wendt, 2014 | TAVR (transapical) | mortality | 30 days | Europe | 155 | 12 | 0.837 | Yes | 8.885298869 |
| Debonnaire, 2015 | AVR or TAVR | mortality | 30 days | Europe | 1512 | 95 | 0.712 | Yes | 8.885298869 |
| Holinski, 2015 | repeat AVR | mortality | 30 days | Europe | 1512 | 95 | 0.712 | Yes | 8.885298869 |
| Moscarelli, 2015 | minimally invasive MV surgery +/- TVR | in-hospital mortality | until discharge | Europe | 1609 | 28 | 0.846 | Yes | 11.79321486 |
| Poullis, 2015 | AVR | in-hospital mortality | until discharge | Europe | 814 | NR | NR | Yes | NA |
| Poullis, 2015 | MVR | in-hospital mortality | until discharge | Europe | 340 | NR | NR | Yes | NA |
| Poullis, 2015 | AVR +/- CABG | in-hospital mortality | until discharge | Europe | 517 | NR | NR | Yes | NA |
| Silaschi, 2015 | TAVR (transfemoral or transapical) | mortality | 30 days | Europe | 457 | 44 | 0.54 | Yes | 87.07592892 |
| Silva, 2015 | TAVR | mortality | 30 days | South America | 418 | 38 | 0.54 | Yes | 87.07592892 |
| Tralhao, 2015 | AVR | operative mortality | 30 days | Europe | 106 | 6 | 0.792 | Yes | 5.654281099 |
| Wang, 2015 | AVR | operative mortality | 30 days | Australia | 620 | 18 | 0.711 | Yes | 31.82552504 |
| Wang, 2015 | AVR | morbidity/mortality | 30 days | Australia | 620 | 115 | 0.649 | Yes | 51.85783522 |
| Halkin, 2016 | TAVR | all-cause mortality | 30 days | Asia | 1327 | 45 | 0.70 | No | 35.37964459 |
| Kortlandt, 2016 | MVR | periprocedural mortality | 30 days | Europe | 136 | 5 | 0.54 | Yes | 87.07592892 |
| Patrat-Delon, 2016 | cardiac surgery for acute | in-hospital mortality | until discharge | Europe | 149 | 32 | 0.78 | Yes | 9.531502423 |
| De novo | Validation |
|---------|-------------|
| **Model** | **Author, Year** | **Index Condition** | **Outcome** | **Timeframe** | **Continent** | **Sample Size** | **Number of Events** | **C-statistic** | **Calibration Reported?** | **Change in Discrimination (%)** |
| Rosa, 2016$^{65}$ | infective endocarditis | TAVR | mortality | 30 days | South America | 59 | 8 | 0.77 | Yes | -12.76252019 |
| Wang, 2016$^{36}$ | | AVR/CABG | operative mortality | 30 days | Australia | 450 | 29 | 0.669 | Yes | -45.39579968 |
| Wang, 2016$^{66}$ | | valve surgery | in-hospital mortality | until discharge | Asia | 12412 | 260 | 0.704 | Yes | -34.08723748 |
| Yamaoka, 2016$^{67}$ | AVR +/- CABG | operative mortality | 30 days | Asia | 406 | 14 | 0.704 | Yes | -34.08723748 |
| Allyn, 2017$^{92}$ | elective cardiac surgery with CPB | post-operative mortality | until discharge | Europe | 6520 | 411 | 0.737 | No | -23.42487884 |
| Bomberg, 2017$^{93}$ | cardiac surgery | mortality | 30 days | Europe | 856 | 27 | 0.74 | No | -22.45557351 |
| Bomberg, 2017$^{93}$ | cardiac surgery | mortality | 6 months | Europe | 809 | 49 | 0.76 | No | -15.99353796 |
| Bomberg, 2017$^{93}$ | cardiac surgery | mortality | 2 years | Europe | 809 | 84 | 0.74 | No | -22.45557351 |
| Kalender, 2017$^{72}$ | AVR | mortality | until discharge | Asia | 35 | 6 | 0.603 | Yes | -66.72051696 |
| Kar, 2017$^{94}$ | valve surgery +/- CABG | mortality | until discharge | Asia | 911 | 52 | 0.76 | Yes | -15.99353796 |
| Kar, 2017$^{94}$ | valve surgery +/- CABG | mortality | until discharge | Asia | 427 | 18 | 0.83 | Yes | -6.62358643 |
| Mateos-Pañero, 2017$^{95}$ | valve surgery +/- CABG | mortality | until discharge | Europe | 866 | 53 | 0.861 | Yes | 16.63974152 |
| Mateos-Pañero, 2017$^{95}$ | valve surgery +/- CABG | mortality | until discharge | Europe | 427 | NR | 0.767 | Yes | -13.73182553 |
| Mateos-Pañero, 2017$^{95}$ | valve surgery +/- CABG | mortality | until discharge | Europe | 119 | NR | 0.954 | Yes | 46.68820679 |
| Schmid, 2017$^{70}$ | TAVR | all-cause mortality | 1 year | Europe | 74 | 10 | 0.669 | No | -45.39579968 |
| Schmid, 2017$^{70}$ | TAVR | all-cause mortality | 2 years | Europe | 74 | 18 | 0.552 | No | -83.19870759 |
| Wang, 2017$^{37}$ | MVR or MVRepair | operative mortality | 30 days | Australia | 407 | 10 | 0.817 | Yes | 2.423263328 |
| Model                        | Author, Year | Index Condition | Outcome                          | Timeframe | Continent     | Sample Size | Number of Events | C-statistic | Calibration Reported? | Change in Discrimination (%) |
|------------------------------|--------------|-----------------|----------------------------------|-----------|---------------|-------------|------------------|-------------|------------------------|-------------------------------|
| NY Operative Mortality Risk Score²⁸ | Hannan, 2013²⁸ | isolated valve surgery | in-hospital/30-day mortality | 30 days   | North America | 12354       | NR               | NR          | No                     | NA                            |
|                              | Jin, 2013⁶⁶  | isolated valve surgery | in-hospital/30-day mortality | 30 days   | Europe        | 4021        | 105              | 0.777       | Yes                    | 1.423487544                      |
|                              | Wang, 2016⁶⁶ | isolated valve surgery | in-hospital mortality | hospitalization period | Asia       | 5152        | 84               | 0.683       | Yes                    | 34.87544484                      |
| Multiple Valve               |              |                 |                                  |           |               |             |                  |             |                        |                               |
| AM Preop³¹                    | Rankin, 2013³¹ | AV + MV surgery | operative mortality | 30 days   | North America | NR         | NR               | 0.71        | Yes                    | NA                            |
| MT Preop³¹                    | Rankin, 2013³¹ | AV + MV surgery | operative mortality | 30 days   | North America | NR         | NR               | 0.722       | Yes                    | NA                            |
| AMT Preop³¹                   | Rankin, 2013³¹ | AV + MV surgery | operative mortality | 30 days   | North America | NR         | NR               | 0.702       | Yes                    | NA                            |
| AM Preop + Intraop³¹          | Rankin, 2013³¹ | AV + MV surgery | operative mortality | 30 days   | North America | NR         | NR               | 0.714       | Yes                    | NA                            |
| MT Preop + Intraop³¹          | Rankin, 2013³¹ | AV + MV surgery | operative mortality | 30 days   | North America | NR         | NR               | 0.727       | Yes                    | NA                            |
| AMT Preop + Intraop³¹         | Rankin, 2013³¹ | AV + MV surgery | operative mortality | 30 days   | North America | NR         | NR               | 0.706       | Yes                    | NA                            |

* ‘Isolated Valve’ indicates a single valve procedure; ‘Multiple Valve’ indicates intervention to > 1 valve; AVR indicates aortic valve surgery (repair or replacement); HL, Hosmer-Lemeshow; MVR, mitral valve surgery (repair or replacement); NR, not reported; NA, not applicable; CABG, coronary artery bypass grafting; TAVR, transcatheter aortic valve replacement; TA-AVI, transapical aortic valve implantation; AEs, adverse events; TVR, tricuspid valve surgery (repair or replacement); DVR, double valve surgery (repair or replacement); MACE, major adverse cardiovascular events.
Abstracts identified
(n = 1205)
All articles citing the VHD de novo models, excluding editorials, commentaries, book chapters, non-English text

Abstracts excluded
(n = 966)

Full-text articles assessed for eligibility
(n = 239)

Full-text articles excluded
(n = 171)
No validation performed; population <50% VHD

Articles included
(n = 68)
Supplemental References:

1. Edwards FH, Peterson ED, Coombs LP, DeLong ER, Jamieson WR, Shroyer ALW, Grover FL. Prediction of Operative Mortality after Valve Replacement Surgery. *J Am Coll Cardiol.* 2001;37:885-892.

2. Nowicki ER, Birkmeyer NJ, Weintraub RW, Leavitt BJ, Sanders JH, Dacey LJ, Clough RA, Quinn RD, Charlesworth DC, Sisto DA, Uhlig PN, Olmstead EM, O’Connor GT. Multiple Variable Prediction of in-Hospital Mortality Associated with Aortic and Mitral Valve Surgery in Northern New England. *Ann Thorac Surg.* 2004;77:1966-1977.

3. Kuduvalli M, Grayson AD, Au J, Grotte G, Bridgewater B, Fabri BM, North West Quality Improvement Programme in Cardiac Interventions. A Multi-Centre Additive and Logistic Risk Model for in-Hospital Mortality Following Aortic Valve Replacement. *Eur J Cardiothorac Surg.* 2007;31:607-613.

4. Cruz-Gonzalez I, Sanchez-Ledesma M, Sanchez PL, Martin-Moreiras J, Jneid H, Rengifo-Moreno P, Inglessis-Azuaje I, Maree AO, Palacios IF. Predicting Success and Long-Term Outcomes of Percutaneous Mitral Valvuloplasty: A Multifactorial Score. *Am J Med.* 2009;122:581.e511-589.

5. Monin JL, Lancellotti P, Monchi M, Lim P, Weiss E, Pierard L, Gueret P. Risk Score for Predicting Outcome in Patients with Asymptomatic Aortic Stenosis. *Circulation.* 2009;120:69-75.

6. O’Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, Shewan CM, Dokholyan RS, Peterson ED, Edwards FH, Anderson RP, Society of Thoracic Surgeons Quality Measurement Task Force. The Society of Thoracic Surgeons 2008 Cardiac Surgery Risk Models: Part 2—Isolated Valve Surgery. *Ann Thorac Surg.* 2009;88:S23-S42.

7. Guaragna JC, Bodanese LC, Bueno FL, Goldani MA. Proposed Preoperative Risk Score for Patients Candidate to Cardiac Valve Surgery. *Arq Bras Cardiol.* 2010;94:541-548.

8. Guo LX, Meng X, Zhang ZG, Bai T. Analysis of Risk Factors for Valve Replacements in 5,128 Cases from a Single Heart Center in China. *Chin Med J (Engl).* 2010;123:3509-3514.

9. Elmariah S, Lubitz SA, Shah AM, Miller MA, Kaplish D, Kothari S, Moreno PR, Kini AS, Sharma SK. A Novel Clinical Prediction Rule for 30-Day Mortality Following Balloon Aortic Valvuloplasty: The Cracc the Av Score. *Catheter Cardiovasc Interv.* 2011;78:112-118.

10. Bouleti C, Iung B, Laouenan C, Hambert D, Brochet E, Messika-Zeitoun D, Detaint D, Garbarz E, Cormier B, Michel PL, Mentre F, Vahanian A. Late Results of Percutaneous Mitral Commissurotomy up to 20 Years: Development and Validation of a Risk Score Predicting Late Functional Results from a Series of 912 Patients. *Circulation.* 2012;125:2119-2127.

11. Cioffi G, Cramariuc D, Dalsgaard M, Davidsen ES, Egstrup K, Rossebo AB, de Simone G, Gerdts E. Left Atrial Systolic Force and Outcome in Asymptomatic Mild to Moderate Aortic Stenosis. *Echocardiography.* 2012;29:1038-1044.

12. Holme I, Pedersen TR, Boman K, Egstrup K, Gerdts E, Kesaniemi YA, Malbecq W, Ray S, Rossebo AB, Wachtell K, Willenheimer R, Gohlke-Barwolf C. A Risk Score for Predicting Mortality in Patients with Asymptomatic Mild to Moderate Aortic Stenosis. *Heart.* 2012;98:377-383.
13. Kotting J, Schiller W, Beckmann A, Schafer E, Dobler K, Hamm C, Veit C, Welz A. German Aortic Valve Score: A New Scoring System for Prediction of Mortality Related to Aortic Valve Procedures in Adults. *Eur J Cardiothorac Surg.* 2013;43:971-977.

14. Arnold SV, Reynolds MR, Lei Y, Magnuson EA, Kirtane AJ, Kodali SK, Zajarias A, Thourani VH, Green P, Rodes-Cabau J, Beohar N, Mack MJ, Leon MB, Cohen DJ, PARTNER Investigators. Predictors of Poor Outcomes after Transcatheter Aortic Valve Replacement: Results from the Partner (Placement of Aortic Transcatheter Valve) Trial. *Circulation.* 2014;129:2682-2690.

15. Capodanno D, Barbanti M, Tamburino C, D'Errigo P, Ranucci M, Santoro G, Santini F, Onorati F, Grossi C, Covello RD, Caprannizzo P, Rosato S, Seccareccia F, OBSERVANT Research Group. A Simple Risk Tool (the Observant Score) for Prediction of 30-Day Mortality after Transcatheter Aortic Valve Replacement. *Am J Cardiol.* 2014;113:1851-1858.

16. D'Ascenzo F, Capodanno D, Tarantini G, Nijhoff F, Ciua C, Rossi ML, Brambilla N, Barbanti M, Napodano M, Stella P, Saia F, Ferrante G, Tamburino C, Gasparetto V, Agostoni P, Marzocchi A, Presbitero P, Bedogni F, Cerrato E, Omede P, Conrotto F, Salizzoni S, Biondi Zoccai G, Marra S, Rinaldi M, Gaita F, D'Amico M, Moretti C. Usefulness and Validation of the Survival Post Tavi Score for Survival after Transcatheter Aortic Valve Implantation for Aortic Stenosis. *Am J Cardiol.* 2014;114:1867-1874.

17. Iung B, Laouenan C, Himbert D, Eltchaninoff H, Chevreul K, Donzeau-Gouge P, Fajadet J, Leprince J, Leguerrier A, Lievre M, Prat A, Teiger E, Laskar M, Vahanian A, Gilard M, FRANCE 2 Investigators. Predictive Factors of Early Mortality after Transcatheter Aortic Valve Implantation: Individual Risk Assessment Using a Simple score. *Heart.* 2014;100:1016-1023.

18. Debonnaire P, Fusini L, Wolterbeek R, Kamperidis V, van Rosendaal P, van der Kley F, Katsanos S, Joyce E, Tamborini G, Muratori M, Gripari P, Bax JJ, Marsan NA, Pepi M, Delgado V. Value of the "Tavi2-Score" Versus Surgical Risk Scores for Prediction of One Year Mortality in 511 Patients Who Underwent Transcatheter Aortic Valve Implantation. *Am J Cardiol.* 2015;115:234-242.

19. Edwards FH, Cohen DJ, O'Brien SM, Peterson ED, Mack MJ, Shahian DM, Grover FL, Tuzcu EM, Thourani VH, Carroll J, Brennan JM, Brindis RG, Rumsfeld J, Holmes DR, Jr., Steering Committee of the Society of Thoracic Surgeons/American College of Cardiology Transcatheter Valve Therapy Registry. Development and Validation of a Risk Prediction Model for in-Hospital Mortality after Transcatheter Aortic Valve Replacement. *JAMA Cardiol.* 2016;1:46-52.

20. Koplan BA, Stevenson WG, Epstein LM, Aranki SF, Maisel WH. Development and Validation of a Simple Risk Score to Predict the Need for Permanent Pacing after Cardiac Valve Surgery. *J Am Coll Cardiol.* 2003;41:795-801.

21. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, Simple Risk Stratification Model for Heart Valve Surgery. *Circulation.* 2005;112:224-231.

22. Xu J, Ge Y, Pan S, Liu F, Shi Y. A Preoperative and Intraoperative Predictive Model of Prolonged Intensive Care Unit Stay for Valvular Surgery. *J Heart Valve Dis.* 2006;15:219-224.

23. Hannan EL, Wu C, Bennett EV, Carlson RE, Culliford AT, Gold JP, Higgins RS, Smith CR, Jones RH. Risk Index for Predicting in-Hospital Mortality for Cardiac Valve Surgery. *Ann Thorac Surg.* 2007;83:921-929.

24. Xu J, Ge Y, Hu S, Song Y, Sun H, Liu P. A Simple Predictive Model of Prolonged Intensive Care Unit Stay after Surgery for Acquired Heart Valve Disease. *J Heart Valve Dis.* 2007;16:109-115.

25. Shi JH, Meng X, Han J, Li Y, Wang JG, Zhang HB, Jia YX, Gurbanov E, Zhuang XJ. A Mortality Risk Assessment Model for Cardiac Valve Replacement Surgery and Its Application in the Use of Prophylactic Extracorporeal Membrane Oxygenation. *Int Surg.* 2010;95:227-231.
26. Ariyaratne TV, Billah B, Yap CH, Dinh D, Smith JA, Shardey GC, Reid CM. An Australian Risk Prediction Model for Determining Early Mortality Following Aortic Valve Replacement. *Eur J Cardiothorac Surg.* 2011;39:815-821.

27. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, Lockowandt U. Euroscore II. *Eur J Cardiothorac Surg.* 2012;41:734-744; discussion 744-735.

28. Hannan EL, Racz M, Culliford AT, Lahey SJ, Wechsler A, Jordan D, Gold JP, Higgins RS, Smith CR. Risk Score for Predicting in-Hospital/30-Day Mortality for Patients Undergoing Valve and Valve/Coronary Artery Bypass Graft Surgery. *Ann Thorac Surg.* 2013;95:1282-1290.

29. Wang C, Zhang GX, Lu FL, Li BL, Zou LJ, Han L, Xu ZY. A Local Risk Prediction Model for Prolonged Ventilation after Adult Heart Valve Surgery in a Chinese Single Center. *Heart Lung.* 2013;42:13-18.

30. Zheng Z, Fan H, Gao H, Li X, Yuan X, Meng J, Xu J, Song Y, Sun H, Hu S. Mortality Risk Model for Heart Valve Surgery in China. *J Heart Valve Dis.* 2013;22:93-101.

31. Rankin JS, He X, O'Brien SM, Jacobs JP, Welke KF, Filardo G, Shahian DM. The Society of Thoracic Surgeons Risk Model for Operative Mortality after Multiple Valve Surgery. *Ann Thorac Surg.* 2013;95:1484-1490.

32. Brown JM, O'Brien SM, Wu C, Sikora JA, Griffith BP, Gammie JS. Isolated Aortic Valve Replacement in North America Comprising 108,687 Patients in 10 Years: Changes in Risks, Valve Types, and Outcomes in the Society of Thoracic Surgeons National Database. *J Thorac Cardiovasc Surg.* 2009;137:82-90.

33. Jin R, Grunkemeier GL, Starr A, Providence Health System Cardiovascular Study Group. Validation and Refinement of Mortality Risk Models for Heart Valve Surgery. *Ann Thorac Surg.* 2005;80:471-479.

34. Wang C, Li X, Lu FL, Xu JB, Tang H, Han L, Xu ZY. Comparison of Six Risk Scores for In-Hospital Mortality in Chinese Patients Undergoing Heart Valve Surgery. *Heart Lung Circ.* 2013;22:612-617.

35. Watanabe Y, Hayashida K, Lefevre T, Chevalier B, Hovasse T, Romano M, Garot P, Farge A, Donzeau-Gouge P, Bouvier E, Cormier B, Morice MC. Is Euroscore II Better Than Euroscore in Predicting Mortality after Transcatheter Aortic Valve Implantation? *Catheter Cardiovasc Interv.* 2013;81:1053-1060.

36. Wang TK, Choi DH, Ramanathan T, Ruygrok PN. Comparing Performance of Risk Scores for Combined Aortic Valve Replacement and Coronary Bypass Grafting Surgery. *Heart Lung Circ.* 2016;25:1118-1123.

37. Wang TK, Harmos S, Gamble GD, Ramanathan T, Ruygrok PN. Performance of Contemporary Surgical Risk Scores for Mitral Valve Surgery. *J Card Surg.* 2017;32:172-176.

38. Piazza N, Wenaweser P, van Gameren M, Pilgrim T, Tzikas A, Otten A, Nuis R, Onuma Y, Cheng JM, Kappetein AP, Boersma E, Juni P, de Jaegere P, Windecker S, Serruys PW. Relationship between the Logistic Euroscore and the Society of Thoracic Surgeons Predicted Risk of Mortality Score in Patients Implanted with the Corevalve Revalving System—a Bern-Rotterdam Study. *Am Heart J.* 2010;159:323-329.

39. Basraon J, Chandrashekar YS, John R, Agnihotri A, Kelly R, Ward H, Adabag S. Comparison of Risk Scores to Estimate Perioperative Mortality in Aortic Valve Replacement Surgery. *Ann Thorac Surg.* 2011;92:535-540.

40. Zhang CX, Xu JP, Ge YP, Wei Y, Yang Y, Liu F, Shi Y. Validation of Four Different Risk Stratification Models in Patients Undergoing Heart Valve Surgery in a Single Center in China. *Chin Med J (Engl).* 2011;124:2254-2259.
41. Durand E, Borz B, Godin M, Tron C, Litzler PY, Bessou JP, Dacher JN, Bauer F, Cribier A, Eltchaninoff H. Performance Analysis of Euroscore II Compared to the Original Logistic Euroscore and STS Scores for Predicting 30-Day Mortality after Transcatheter Aortic Valve Replacement. *Am J Cardiol.* 2013;111:891-897.

42. Haensig M, Holzhey DM, Borger MA, Schuler G, Shi W, Subramanian S, Rastan AJ, Mohr FW. Is the New Euroscore II a Better Predictor for Transapical Aortic Valve Implantation? *Eur J Cardiothorac Surg.* 2013;44:302-308; discussion 308.

43. Laurent M, Fournet M, Feit B, Oger E, Donal E, Thebault C, Biron Y, Beneux X, Sellin M, Le Reveille S, Flecher E, Leguerrier A. Simple Bedside Clinical Evaluation Versus Established Scores in the Estimation of Operative Risk in Valve Replacement for Severe Aortic Stenosis. *Arch Cardiovasc Dis.* 2013;106:651-660.

44. Barili F, Pacini D, Grossi C, Di Bartolomeo R, Alamanni F, Parolari A. Reliability of New Scores in Predicting Perioperative Mortality after Mitral Valve Surgery. *J Thorac Cardiovasc Surg.* 2014;147:1008-1012.

45. Beohar N, Whisenant B, Kirtane AJ, Leon MB, Tuzcu EM, Makkar R, Svensson LG, Miller DC, Smith CR, Pichard AD, Herrmann HC, Thourani VH, Szeto WY, Lim S, Fischbein M, Fearon WF, O’Neill W, Xu K, Dewey T, Mack M. The Relative Performance Characteristics of the Logistic European System for Cardiac Operative Risk Evaluation Score and the Society of Thoracic Surgeons Score in the Placement of Aortic Transcatheter Valves Trial. *J Thorac Cardiovasc Surg.* 2014;148:2830-2837.e2831.

46. Chan V, Ahraari A, Ruel M, Elmistekawy E, Hynes M, Mesana TG. Perioperative Deaths after Mitral Valve Operations May Be Overestimated by Contemporary Risk Models. *Ann Thorac Surg.* 2014;98:605-610; discussion 610.

47. Osnabrugge RL, Speir AM, Head SJ, Fonner CE, Fonner E, Kappetein AP, Rich JB. Performance of Euroscore II in a Large US Database: Implications for Transcatheter Aortic Valve Implantation. *Eur J Cardiothorac Surg.* 2014;46:400-408; discussion 408.

48. Rabbani MS, Qadir I, Ahmed Y, Gul M, Sharif H. Heart Valve Surgery: Euroscore Vs. Euroscore II Vs. Society of Thoracic Surgeons Score. *Heart Int.* 2014;9:53-58.

49. Wendt D, Thielmann M, Kahler P, Kastner S, Price V, Al-Rashid F, Patsalis P, Erbel R, Jakob H. Comparison between Different Risk Scoring Algorithms on Isolated Conventional or Transcatheter Aortic Valve Replacement. *Ann Thorac Surg.* 2014;97:796-802.

50. Wang L, Lu FL, Wang C, Tan MW, Xu ZY. Society of Thoracic Surgeons 2008 Cardiac Risk Models Predict in-Hospital Mortality of Heart Valve Surgery in a Chinese Population: A Multicenter Study. *J Thorac Cardiovasc Surg.* 2014;148:3036-3041.

51. Adami M, Capodanno D, Cannata S, Giannini C, Laudisa ML, Barbanti M, Curello S, Imme S, Maffeo D, Grasso C, Bedogni F, Petronio AS, Ettori F, Tamburino C, GRASP-IT Investigators. Comparison of Three Contemporary Surgical Scores for Predicting All-Cause Mortality of Patients Undergoing Percutaneous Mitral Valve Repair with the Mitraclip System (from the Multicenter Grasp-IT Registry). *Am J Cardiol.* 2015;115:107-112.

52. Holinski S, Jessen S, Neumann K, Konertz W. Predictive Power and Implication of Euroscore, Euroscore II and STS Score for Isolated Repeated Aortic Valve Replacement. *Ann Thorac Surg.* 2015;21:242-246.

53. Silaschi M, Conradi L, Seiffert M, Schnabel R, Schon G, Blankenberg S, Reichenspurner H, Diemert P, Treede H. Predicting Risk in Transcatheter Aortic Valve Implantation: Comparative Analysis of Euroscore II and Established Risk Stratification Tools. *Thorac Cardiovasc Surg.* 2015;63:472-478.
54. Silva LS, Caramori PR, Nunes Filho AC, Katz M, Guaragna JC, Lemos P, Lima V, Abizaid A, Tarasoutchi F, Brito FS, Jr. Performance of Surgical Risk Scores to Predict Mortality after Transcatheter Aortic Valve Implantation. *Arq Bras Cardiol.* 2015;105:241-247.

55. Sinning JM, Wollert KC, Sedaghat A, Widera C, Radermacher MC, Descoups C, Hammerstingl C, Weber M, Stundl A, Ghanem A, Widder J, Vasa-Nicotera M, Mellert F, Schiller W, Bauersachs J, Zur B, Holdenrieder S, Welz A, Grube E, Pencina MJ, Nickenig G, Werner N, Kempf T. Risk Scores and Biomarkers for the Prediction of 1-Year Outcome after Transcatheter Aortic Valve Replacement. *Am Heart J.* 2015;170:821-829.

56. Tralhão A, Teles RC, Almeida MS, Madeira S, Santos MB, Andrade MJ, Mendes M, Neves JP. Aortic Valve Replacement for Severe Aortic Stenosis in Octogenarians: Patient Outcomes and Comparison of Operative Risk Scores. *Rev Port Cardiol.* 2015;34:439-446.

57. Vassileva CM, Aranki S, Brennan JM, Kaneko T, He M, Gammie JS, Suri RM, Thourani VH, Hazelrigg S, McCarthy P. Evaluation of the Society of Thoracic Surgeons Online Risk Calculator for Assessment of Risk in Patients Presenting for Aortic Valve Replacement after Prior Coronary Artery Bypass Graft: An Analysis Using the Sts Adult Cardiac Surgery Database. *Ann Thorac Surg.* 2015;100:2109-2115; discussion 2115-2106.

58. Wang TK, Choi DH, Stewart R, Gamble G, Haydock D, Ruygrok P. Comparison of Four Contemporary Risk Models at Predicting Mortality after Aortic Valve Replacement. *J Thorac Cardiovasc Surg.* 2015;149:443-448.

59. Wang TK, Choi DH, Haydock D, Gamble G, Stewart R, Ruygrok P. Comparison of Risk Scores for Prediction of Complications Following Aortic Valve Replacement. *Heart Lung Circ.* 2015;24:595-601.

60. Barili F, Pacini D, D’Ovidio M, Ventura M, Alamanni F, Di Bartolomeo R, Grossi C, Davoli M, Fusco D, Perucci C, Parolari A. Reliability of Modern Scores to Predict Long-Term Mortality after Isolated Aortic Valve Operations. *Ann Thorac Surg.* 2016;101:599-605.

61. Collas VM, Van De Heyning CM, Paelinck BP, Rodrigus IE, Vrints CJ, Bosmans JM. Validation of Transcatheter Aortic Valve Implantation Risk Scores in Relation to Early and Mid-Term Survival: A Single-Centre Study. *Interact Cardiovasc Thorac Surg.* 2016;22:273-279.

62. Halkin A, Steinvil A, Witberg G, Barsheteshet A, Barkagan M, Assali A, Segev A, Fefer P, Guetta V, Barbash IM, Kornowski R, Finkelstein A. Mortality Prediction Following Transcatheter Aortic Valve Replacement: A Quantitative Comparison of Risk Scores Derived from Populations Treated with Either Surgical or Percutaneous Aortic Valve Replacement. The Israeli Tavr Registry Risk Model Accuracy Assessment (Irmma) Study. *Int J Cardiol.* 2016;215:227-231.

63. Kortlandt FA, van ’t Klooster CC, Bakker AL, Swaans MJ, Kelder JC, de Kroon TL, Rensing BJ, Eefting FD, van der Heyden JA, Post MC. The Predictive Value of Conventional Surgical Risk Scores for Periprocedural Mortality in Percutaneous Mitral Valve Repair. *Neth Heart J.* 2016;24:475-480.

64. Peguero JG, Lo Presti S, Issa O, Podesta C, Parise H, Layka A, Brenes JC, Lamelas J, Lamas GA. Simplified Prediction of Postoperative Cardiac Surgery Outcomes with a Novel Score: R2chads2. *Am Heart J.* 2016;177:153-159.

65. Rosa VE, Lopes AS, Accorsi TA, Fernandes JR, Spina GS, Sampaio RO, Paixao MR, Pomerantz Eff PM, Lemos Neto PA, Tarasoutchi F. Euroscore Ii and Sts as Mortality Predictors in Patients Undergoing Tavi. *Rev Assoc Med Bras (1992).* 2016;62:32-37.

66. Wang C, Tang YF, Zhang JJ, Bai YF, Yu YC, Zhang GX, Han L. Comparison of Four Risk Scores for in-Hospital Mortality in Patients Undergoing Heart Valve Surgery: A Multicenter Study in a Chinese Population. *Heart Lung.* 2016;45:423-428.
67. Yamaoka H, Kuwaki K, Inaba H, Yamamoto T, Kato TS, Dohi S, Matsushita S, Amano A. Comparison of Modern Risk Scores in Predicting Operative Mortality for Patients Undergoing Aortic Valve Replacement for Aortic Stenosis. *J Cardiol*. 2016;68:135-140.

68. Zbronski K, Huczek Z, Puchta D, Paczwa K, Kochman J, Wilimski R, Scisio P, Rymuza B, Filipiak KJ, Opolski G. Outcome Prediction Following Transcatheter Aortic Valve Implantation: Multiple Risk Scores Comparison. *Cardiol J*. 2016;23:169-177.

69. Balan P, Zhao Y, Johnson S, Arain S, Dhoble A, Estrera A, Smalleg R, Nguyen TC. The Society of Thoracic Surgery Risk Score as a Predictor of 30-Day Mortality in Transcatheter Vs Surgical Aortic Valve Replacement: A Single-Center Experience and Its Implications for the Development of a Tavr Risk-Prediction Model. *J Invasive Cardiol*. 2017;29:109-114.

70. Schmid J, Stojakovic T, Zweiker D, Scharnagl H, Maderthaner RD, Scherr D, Maier R, Schmidt A, Marz W, Binder JS, Rainer PP. St2 Predicts Survival in Patients Undergoing Transcatheter Aortic Valve Implantation. *Int J Cardiol*. 2017;244:87-92.

71. Sa MP, Sa MV, Albuquerque AC, Silva BB, Siqueira JW, Brito PR, Vasconcelos FP, Lima Rde C. Guaragnascore Satisfactorily Predicts Outcomes in Heart Valve Surgery in a Brazilian Hospital. *Rev Bras Cir Cardiovasc*. 2012;27:1-6.

72. Kalender M, Baysal AN, Karaca OG, Boyacioglu K, Kayalar N. Validation of German Aortic Valve Score in a Multi-Surgeon Single Center. *Braz J Cardiovasc Surg*. 2017;32:77-82.

73. Martin GP, Sperrin M, Ludman PF, de Belder MA, Gale CP, Toff WD, Moat NE, Trivedi U, Buchan I, Mamas MA. Inadequacy of Existing Clinical Prediction Models for Predicting Mortality after Transcatheter Aortic Valve Implantation. *Am Heart J*. 2017;184:97-105.

74. Arnold SV, Afifalo J, Spertus JA, Tang Y, Baron SJ, Jones PG, Reardon MJ, Yakubov SJ, Adams DH, Cohen DJ. Prediction of Poor Outcome after Transcatheter Aortic Valve Replacement. *J Am Coll Cardiol*. 2016;68:1868-1877.

75. De Bacco G, De Bacco MW, Sant’Anna JR, Santos MF, Sant’Anna RT, Prates PR, Kalil RA, Nesralla IA. Applicability of Ambler’s Risk Score to Patients Who Have Undergone Valve Replacement with Bovine Pericardial Bioprosthesis. *Rev Bras Cir Cardiovasc*. 2008;23:336-343.

76. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of Risk Algorithms in Predicting Early and Late Operative Outcomes in High-Risk Patients Undergoing Aortic Valve Replacement. *J Thorac Cardiovasc Surg*. 2008;135:180-187.

77. Tran HA, Roy SK, Hebsur S, Barnett SD, Schlauch KA, Hunt SL, Holmes SD, Ad N. Performance of Four Risk Algorithms in Predicting Intermediate Survival in Patients Undergoing Aortic Valve Replacement. *Innovations (Phil).* 2010;5:407-412.

78. van Gameren M ea. Do We Need Separate Risk Stratification Models for Hospital Mortality after Heart Valve Surgery? - Pubmed - Ncbi. *Ann Thorac Surg*. 2008;85:921-930.

79. Barili F, Pacini D, Rosato F, Parolari A. The Role of Surgical Procedures on Discriminative Performance of the Updated Euroscore II. *J Thorac Cardiovasc Surg*. 2013;146:986-987.

80. Carnero-Alcazar M, Silva Guisasola JA, Renguillo Lacruz FJ, Maroto Castellanos LC, Cobiella Carnicer J, Villagran Medinilla E, Tejerina Sanchez T, Rodriguez Hernandez JE. Validation of Euroscore II on a Single-Centre 3800 Patient Cohort. *Interact Cardiovasc Thorac Surg*. 2013;16:293-300.

81. Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N, Poullis M. Validation of Euroscore II in a Modern Cohort of Patients Undergoing Cardiac Surgery. *Eur J Cardiothorac Surg*. 2013;43:688-694.
82. Howell NJ, Head SJ, Freemantle N, van der Meulen TA, Senanayake E, Menon A, Kappetein AP, Pagano D. The New Euroscore II Does Not Improve Prediction of Mortality in High-Risk Patients Undergoing Cardiac Surgery: A Collaborative Analysis of Two European Centres. *Eur J Cardiothorac Surg.* 2013;44:1006-1011; discussion 1011.

83. Sedaghat A, Sinning JM, Vasa-Nicotera M, Ghanem A, Hammerstingl C, Grube E, Nickenig G, Werner N. The Revised Euroscore II for the Prediction of Mortality in Patients Undergoing Transcatheter Aortic Valve Implantation. *Clin Res Cardiol.* 2013;102:821-829.

84. Zhang GX, Wang C, Wang L, Lu FL, Li BL, Han L, Xu ZY. Validation of Euroscore II in Chinese Patients Undergoing Heart Valve Surgery. *Heart Lung Circ.* 2013;22:606-611.

85. Barili F, Pacini D, Rosato F, Roberto M, Battisti A, Grossi C, Alamanni F, Di Bartolomeo R, Parolari A. In-Hospital Mortality Risk Assessment in Elective and Non-Elective Cardiac Surgery: A Comparison between Euroscore II and Age, Creatinine, Ejection Fraction Score. *Eur J Cardiothorac Surg.* 2014;46:44-48.

86. Koszta G, Sira G, Szatmari K, Farkas E, Szerafin T, Fulesdi B. Performance of Euroscore II in Hungary: A Single-Centre Validation Study. *Heart Lung Circ.* 2014;23:1041-1050.

87. Spiliopoulos K, Bagiatis V, Deutsch O, Kemkes BM, Antonopoulos N, Karangelis D, Haschemi A, Gansera B. Performance of Euroscore II Compared to Euroscore I in Predicting Operative and Mid-Term Mortality of Patients from a Single Center after Combined Coronary Artery Bypass Grafting and Aortic Valve Replacement. *Gen Thorac Cardiovasc Surg.* 2014;62:103-111.

88. Wang L, Han QQ, Qiao F, Wang C, Zhang XW, Han L, Xu ZY. Performance of Euroscore II in Patients Who Have Undergone Heart Valve Surgery: A Multicentre Study in a Chinese Population. *Eur J Cardiothorac Surg.* 2014;45:359-364.

89. Moscarelli M, Bianchi G, Margaryan R, Cerillo A, Farneti P, Murzi M, Solinas M. Accuracy of Euroscore II in Patients Undergoing Minimally Invasive Mitral Valve Surgery. *Interact Cardiovasc Thorac Surg.* 2015;21:748-753.

90. Poullis M, Pullan M, Chalmers J, Mediratta N. The Validity of the Original Euroscore and Euroscore II in Patients over the Age of Seventy. *Interact Cardiovasc Thorac Surg.* 2015;20:172-177.

91. Patrat-Delon S, Rouxel A, Gacouin A, Revest M, Flecher E, Fouquet O, Le Tulzo Y, Lerolle N, Tattevin P, Tadie JM. Euroscore II Underestimates Mortality after Cardiac Surgery for Infective Endocarditis. *Eur J Cardiothorac Surg.* 2016;49:944-951.

92. Allyn J, Allou N, Augustin P, Philip I, Martinet O, Belghiti M, Provenchere S, Montravers P, Ferydunos C. A Comparison of a Machine Learning Model with Euroscore II in Predicting Mortality after Elective Cardiac Surgery: A Decision Curve Analysis. *PLoS One.* 2017;12:e0169772.

93. Bomberg H, Klingele M, Wagenpfeil S, Spanuth E, Volk T, Sessler DI, Schafer HS, Groesdonk HV. Presepsin (Scd14-St) Is a Novel Marker for Risk Stratification in Cardiac Surgery Patients. *Anesthesiology.* 2017;126:631-642.

94. Kar P, Geeta K, Gopinath R, Durga P. Mortality Prediction in Indian Cardiac Surgery Patients: Validation of European System for Cardiac Operative Risk Evaluation II. *Indian J Anaesth.* 2017;61:157-162.

95. Mateos-Panero B, Sanchez-Casado M, Castano-Moreira B, Paredes-Astillero I, Lopez-Almodovar LF, Bustos-Molina F. Assessment of Euroscore and SAPS III as Hospital Mortality Predicted in Cardiac Surgery. *Rev Esp Anestesiol Reanim.* 2017;64:273-281.
96. Jin R, Grunkemeier GL, Providence Health & Services and Swedish Health Services Cardiovascular Disease Study Group. Validation of New York Operative Mortality Risk Score for Valve and Valve/Coronary Artery Bypass Grafting Operations. *Ann Thorac Surg.* 2013;95:1291-1296.