

Predicting Risk in Transcatheter Aortic Valve Implantation: Comparative Analysis of EuroSCORE II and Established Risk Stratification Tools

Miriam Silaschi¹ Lenard Conradi¹ Moritz Seiffert² Renate Schnabel² Gerhard Schön³
Stefan Blankenberg² Hermann Reichenspurner¹ Patrick Diemert² Hendrik Treede¹

¹Department of Cardiovascular Surgery, University Heart Center Hamburg, Hamburg, Germany

²Department of Cardiology, University Heart Center Hamburg, Hamburg, Germany

³Department of Medical Biometry and Epidemiology, University Heart Center Hamburg, Hamburg, Germany

Address for correspondence Miriam Silaschi, MD, Department of Cardiovascular Surgery, University Heart Center Hamburg, Martinistraße 52, Hamburg 20246, Germany (e-mail: m.silaschi@uke.de).

Thorac Cardiovasc Surg 2015;63:472–478.

Abstract

Background The logistic European System for Cardiac Operative Risk Evaluation (logEuroSCORE) II was developed to improve prediction of mortality in cardiac surgery. However, no specific tools are available for risk prediction in transcatheter aortic valve implantation (TAVI). The recently introduced EuroSCORE II was compared with established risk scores.

Patients and Methods We assessed 457 consecutive patients (80.5 ± 7.1 years, 52.3% female) undergoing TAVI. Preoperative risk evaluation included logEuroSCORE I, EuroSCORE II, Society of Thoracic Surgeons (STS), Ambler, and Parsonnet scores. Validity was assessed by receiver-operating characteristic (ROC) and area under the curve (AUC).

Results A 30-day mortality was 9.6% (44/457). Calculated scores were logEuroSCORE I 22.0%, confidence interval (CI) 21.0 to 24.6; EuroSCORE II 7.0%, CI 6.4 to 8.1; STS 7.9%, CI 7.7 to 9.5; Ambler score 6.9%, CI 5.7 to 7.0; and Parsonnet score 23.8%, CI 20.9 to 24.1. ROC analyses demonstrated no predictive value: logEuroSCORE I AUC 0.56, CI 0.47 to 0.65; EuroSCORE II AUC 0.54, CI 0.46 to 0.63; STS AUC 0.57, CI 0.49 to 0.66; Ambler AUC 0.52, CI 0.43 to 0.60; and Parsonnet AUC 0.51, CI 0.43 to 0.60. Accuracy and thresholds were measured on behalf of Youden index. Accuracy ranged between 44.2% (Parsonnet) and 66.3% (logEuroSCORE I). Thresholds were logEuroSCORE I 26%, EuroSCORE II 7%, STS 6%, Ambler 3%, and Parsonnet 19%.

Conclusions No risk evaluation system provided acceptable predictive ability. Scores derived from conventional cardiac surgery failed in risk prediction for TAVI. Specific risk tools are required. Until available, estimation of risk has to rely on judgment of an interdisciplinary heart team regarding individual patient factors.

Keywords

- ▶ transcatheter aortic valve implantation
- ▶ aortic valve disease
- ▶ aortic stenosis

received
February 26, 2014
accepted after revision
July 9, 2014
published online
September 5, 2014

© 2015 Georg Thieme Verlag KG
Stuttgart · New York

DOI <http://dx.doi.org/10.1055/s-0034-1389107>.
ISSN 0171-6425.

Introduction

The logistic European System for Cardiac Operative Risk Evaluation (logEuroSCORE) I has been developed to predict operative mortality in patients undergoing cardiac surgery.¹ Its accuracy in predicting mortality among surgical patients of advanced age and/or high risk is imprecise.² It is impaired by an overestimation of patients' individual risk of mortality and has been shown not to correlate with acute outcome.³ Consequently, the EuroSCORE II, a modified version of the former logistic EuroSCORE I, was introduced in 2011 for patients undergoing cardiac surgery.⁴ While in logistic EuroSCORE I, 17 binary patient-related factors are considered for risk prediction, its follow-up model, the EuroSCORE II includes two additional clinical factors (New York Heart Association classification, presence of insulin-dependent diabetes) and more specific response options (e.g., renal function according to creatinine clearance, more detailed description of left ventricular function and pulmonary hypertension) in risk analyses.

Surgical risk stratification systems are widely accepted as tools to identify candidates for transcatheter aortic valve implantation (TAVI) among high-risk patients with severe aortic stenosis. In general, the logistic EuroSCORE I and the Society of Thoracic Surgeons predicted risk of mortality (STS-PROM) score^{1,5,6} are used with accepted cutoff values of > 20 and > 10%, respectively, indicating high-risk of preoperative mortality. As TAVI is increasingly being used for treatment of comorbid high-risk patients of advanced age and indication is extended to patients with concomitant coronary artery disease⁷ or predominant aortic regurgitation,⁸ availability of a uniform and objective tool for decision making regarding the adequate choice of treatment is desirable. Since interventional techniques for treatment of severe aortic stenosis became available for high-risk patients, the definition of "high risk" has been critical among these patients. A reliable preoperative risk evaluation system designed specifically for patients undergoing TAVI has not been thoroughly validated to date. The objective of this study was to study the predictive reliability of five different surgical risk scores for patients undergoing TAVI, including the new EuroSCORE II.

Patients and Methods

Study Population and Data Collection

From January 2008 through August 2012, 457 consecutive patients underwent transfemoral or transapical TAVI (TA-AVI) using different commercially available devices at our center. Patients were allocated to TAVI when deemed unsuitable for conventional surgery due to contraindications or high risk by the local interdisciplinary heart team consisting of cardiologists and cardiac surgeons. Valve-in-valve or combined procedures were excluded from analysis. For all patients, logistic EuroSCORE I and STS-PROM were prospectively calculated. Furthermore, EuroSCORE II, Ambler, and Parsonnet scores were retrospectively calculated using web-based calculators.^{4,9,10} Logistic EuroSCORE I and Parsonnet

score are available as additive and logistic regression models. We calculated only the logistic regression model for both. With the Ambler score, risk factors are entered in an additive manner and can be transferred into a related probability of death by use of a lookup table or the following relationships: $\log \text{ odds} = 1.36 - 1.75 \times \exp(1.45 - 0.0716 \times S)$ and risk of in-hospital death (%) = $100/[1 + \exp(-\log \text{ odds})]$, where S is the sum of the risk scores for an individual patient.⁹

Statistical Analysis

Data are presented as absolute numbers and percentages for categorical variables and mean values and standard deviations for continuous variables. Dichotomous variables were compared using Fisher exact test and continuous variables by t tests. Correlations between scores are described by Spearman correlation coefficient r . A correlation coefficient of 0.5 to 1.0 suggests a strong positive correlation between two variables, whereas coefficients between 0.3 to 0.5 and 0.1 to 0.3 indicate medium and small correlation. The discriminatory power of the risk stratification systems was evaluated using a c -statistic, the area under the receiver-operating characteristic curve (ROC curve), with 95% confidence interval (CI). A c -statistic of 0.5 indicates no predictive ability, whereas a c -statistic of 1.0 would signify perfect discrimination. If area under the curve (AUC) reaches 0.7, the test is considered to be of acceptable predictive ability. We calculated the Youden index (maximum of sensitivity and specificity) to estimate an associated threshold to classify scores as positive/negative. Accuracy of the respective scores was defined as the sum of correct positive and correct negative predictions. Hosmer–Lemeshow goodness-of-fit test was applied to assess calibration of the respective tests, whereas a high p value and a low statistic indicate a well-calibrated test, demonstrating that model prediction is not significantly different from observed values. All statistical analysis was performed using SPSS (Armonk, New York, United States) 19.0 or the statistical package R version 2.12.2.

Results

Patients

A total of 457 patients underwent TAVI from March, 2008, to February, 2012, at our center. All patients were deemed high risk for surgery or inoperable by mutual agreement of a dedicated heart team consisting of interventional cardiologists, cardiac surgeons, and imaging specialists. Transfemoral and transapical access was used in 42.9% (196) and 57.1% (261) patients. Different devices were used, with the majority of patients receiving Edwards Sapien (Irvine, California, United States) (47.9%, $n = 219$) or Sapien XT (Irvine, California, United States) (23.4%, $n = 107$) devices. The Medtronic CoreValve (Minneapolis, Minnesota, United States) was implanted in 62 patients (13.5%), JenaValve (München, Germany) in 37 patients (8.1%), and Symetis Acurate (Ecublens, Switzerland) in 18 patients (3.9%); 52.3% of patients were female ($n = 239$), mean age was 80.5 ± 7.1 years. Detailed summary of patient baseline demographics and risk factors is provided in ► **Table 1**.

Table 1 Patient baseline demographics and risk factors

	All patients (n = 457)	LES I	ES II	STS	Ambler	Parsonnet
Age (y)	80.5 ± 7.1	√	√	√	√ ^a	√
Gender (female)	52.3% (239)	√	√	√	√	√
NYHA functional class			√	√		
NYHA I	2.4% (11)					
NYHA II	12.3% (56)					
NYHA III	70.2% (321)					
NYHA IV	14.9% (68)					
LVEF 30–50%	14.4% (66)	√	√	√	√	√
LVEF < 30%	6.8% (31)	√	√	√	√	√
EOA (cm ²)	0.9 ± 3.2					
Mean/peak gradient (mm Hg)	36.1 ± 18.4/61.9 ± 28.4					
Prior TIA/stroke	20.1% (92)	√ ^b	√ ^b	√		
Coronary artery disease	45.5% (208)			√		
Extra-cardiac arthropathy	14.4% (166)	√	√	√		
Creatinine (mg/dL)	1.5 ± 1.3	√	√	√	√	
Pulmonary hypertension > 60 mm Hg	16.8% (77)	√	√			
COPD	25.4% (116)	√	√	√		
Previous cardiac surgery	22.0% (88)	√	√	√	√	√

Abbreviations: LES I, logistic EuroSCORE I; ES II, EuroSCORE II; STS, Society of Thoracic Surgeons Predicted Risk of Mortality; NYHA, New York Heart Association; LVEF, left ventricular ejection fraction; EOA, effective orifice area; TIA, transient ischemic attack; COPD, chronic obstructive pulmonary disease; continuous variables as mean ± standard deviation, categorical variables as percentages (n).

Note: Corresponding variables used for risk calculation in the respective scores are highlighted. Not all risk factors which are part of the presented scores are displayed.

^aAmbler score is the only score that categorizes age.

^bIn LES I and ES II, prior cerebrovascular events are comprised in the category “neurological dysfunction/poor mobility.”

Acute Mortality and Mean Scores

All-cause mortality was 9.6% (n = 44) at 30 days in this high-risk patient population. Mean predicted mortality by EuroSCORE II was 7.0 ± 5.9% and therefore lower than the observed mortality. Mean calculated scores are illustrated in ►Table 2. The PROM by EuroSCORE II in patients who deceased during 30-day follow-up was not significantly higher than in those who survived 30 days after the procedure with 6.8 and 7.1%, respectively (p = 0.826). Similar results were observed for the other tested scores, as the deceased patients did not have higher scores compared with survivors (p = 0.3–0.8). For detail, see ►Table 2 and ►Fig. 1.

Table 2 A 30-day mortality and mean risk scores

Mortality to 30 days	9.6% (44)
Logistic EuroSCORE I	22 ± 13.5%
EuroSCORE II	7.0 ± 5.9%
STS-PROM	7.9 ± 5.7%
Ambler score	6.9 ± 5.4%
Parsonnet score	23.8 ± 13.0%

Abbreviations: EuroSCORE, European System for Cardiac Operative Risk Evaluation; STS-PROM, Society of Thoracic Surgeons predicted risk of mortality.

Correlation of Scores

There was a strong linear correlation between logEuroSCORE I and EuroSCORE II in TAVI, as Spearman correlation coefficient *r* was 0.708, 95% CI 0.657 to 0.752 (p < 0.01). The logEuroSCORE I and STS-PROM correlated moderately (r = 0.558, 95% CI 0.489–0.619), as well as the EuroSCORE II and STS-PROM (r = 0.501, 95% CI 0.427–0.569). Correlation matrix of all other tested scores revealed medium correlation with r = 0.3 to 0.5.

Predictive Ability According to ROC Analysis

ROC analysis was performed for our large TAVI patient cohort. It revealed none of the tested risk stratification systems

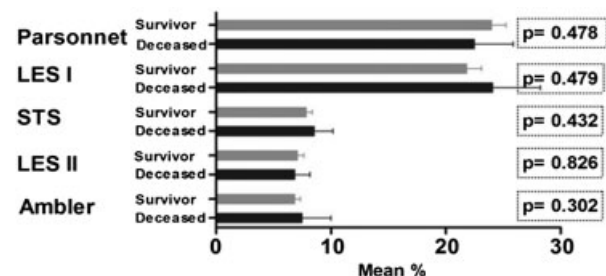


Fig. 1 Mean risk scores of survivors and deceased patients. LES I, logEuroSCORE I; LES II, EuroSCORE II.

to possess adequate predictive ability for acute mortality after TAVI: EuroSCORE II AUC 0.54 (95% CI 0.46–0.63), logEuroSCORE I AUC 0.56 (95% CI 0.47–0.65), STS-PROM AUC 0.57 (95% CI 0.49–0.66), Ambler score AUC 0.52 (95% CI 0.43–0.60), and Parsonnet score AUC 0.51 (95% CI 0.43–0.60). Comparison of EuroSCORE II to the other tested scores regarding AUC revealed no significant advantage ($p > 0.5$ in all cases). Derived accuracy of the respective scores was low: EuroSCORE II 63.8%, logEuroSCORE I 66.3%, STS-PROM 48.7%, Ambler score 17.8%, and Parsonnet score 44.2%. For detail, see ► **Table 3**.

Estimated Thresholds by Youden Indices

In accordance to Youden indices, a calculated cutoff value of 7% in EuroSCORE II corresponds to an estimated value of 26% in logEuroSCORE I for our cohort (see ► **Fig. 2**). Highest specificity was reached by logEuroSCORE I (68.6%), at an estimated threshold of 26%. Threshold in STS-PROM was 6%. Highest sensitivity of all scores was reached by the Ambler score (97.7%), but only provided a maximum specificity of 8.9%, at an estimated threshold for classification as high risk of 3%. Threshold was 19% in Parsonnet score. According to the estimated thresholds, deaths were not predicted by EuroSCORE II in 25 cases (56.8%), in logEuroSCORE I in 27 cases (61.4% of deaths), in STS-PROM in 13 cases (29.5%), Ambler score in 1 case (2.3%), and Parsonnet score in 21 cases (47.7%). For detail of ROC analyses and Youden indices in all tested scores, see ► **Figs. 2 and 3**.

Performance of EuroSCORE II in Different Risk Levels

Patients were furthermore divided into three levels of risk classification by EuroSCORE II, to evaluate the ability of discrimination between low-, intermediate-, and high-risk patients. The defined levels were low risk (EuroSCORE II < 3%), intermediate risk (EuroSCORE II 3–7%) and high risk (EuroSCORE II > 7%). Thirty-day survival decreased from low-risk (92.9%) to intermediate-risk (90.6%) and high-risk patients (88.3%), although this was not significant ($p = 0.490$). A division into subgroups of “non-high risk” (EuroSCORE II < 7%) and “high risk” (EuroSCORE II > 7%) by threshold estimation of Youden index revealed lower survival

Table 3 Summary of predictive performance of tested scores in our TAVI cohort

Score	AUC (95% CI)	Threshold (%)	Accuracy
Logistic EuroSCORE I	0.56 (0.47–0.65)	26	66.3
EuroSCORE II	0.54 (0.46–0.63)	7	63.8
STS-PROM	0.57 (0.49–0.66)	6	48.7
Ambler	0.52 (0.43–0.60)	3	17.8
Parsonnet	0.51 (0.43–0.60)	19	44.2

Abbreviations: AUC, area under the curve; CI, confidence interval; EuroSCORE, European System for Cardiac Operative Risk Evaluation; STS-PROM, Society of Thoracic Surgeons predicted risk of mortality; TAVI, transcatheter aortic valve implantation.

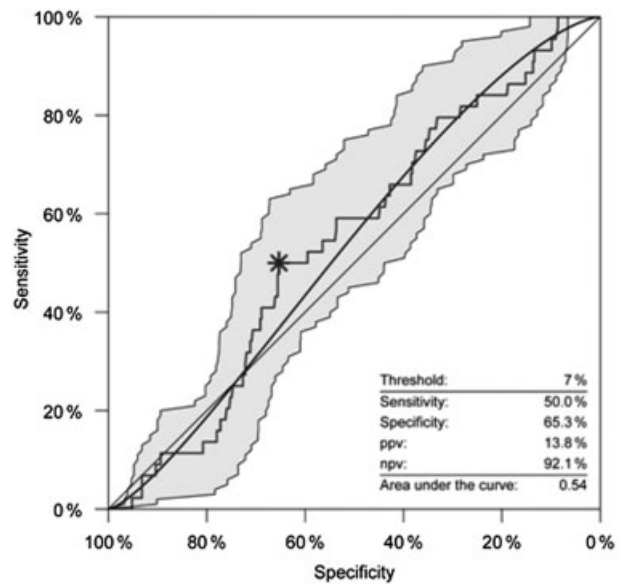


Fig. 2 c-statistic of logistic EuroSCORE II. Ppv, positive predictive value; npv, negative predictive value.

rates in the high-risk cohort, but this was not significant either (91.1 vs. 87.8%, $p = 0.247$).

Results from Hosmer–Lemeshow Test

Regression analyses made by Hosmer–Lemeshow goodness-of-fit test revealed that risk prediction of all tests was not significantly different from the observed events because p value was more than 0.05 in all tested models (EuroSCORE II $p = 0.133$; logEuroSCORE I $p = 0.389$; STS $p = 0.707$; Ambler $p = 0.969$; Parsonnet $p = 0.782$) and therefore the estimates fit to our data at an acceptable level.

Predictive Ability According to Access

Subgroup analysis of transfemoral and TA-AVI patients revealed no statistical difference compared with our findings in the overall TAVI cohort, as AUC was below 0.7 for all tested scores in the separate groups.

Discussion

The EuroSCORE II has recently been introduced to improve preoperative risk stratification in cardiac surgery.⁴ Its discriminative capacity in patients undergoing cardiac surgery with a low- to medium-risk profile has recently been demonstrated.¹¹ As the logEuroSCORE I is widely being used to categorize patients eligible for TAVI, it seems possible that the EuroSCORE II will equally serve as a benchmark for inoperable candidates.

We present an evaluation of the EuroSCORE II in a real world, clinical experience of 457 TAVI patients. Furthermore, we describe the predictive ability of the EuroSCORE II in comparison to four other systems of cardiac preoperative risk evaluation.

Overall mortality at 30 days was 9.6% ($n = 44$) and corresponds to mortality rates reported in the literature.^{12,13}

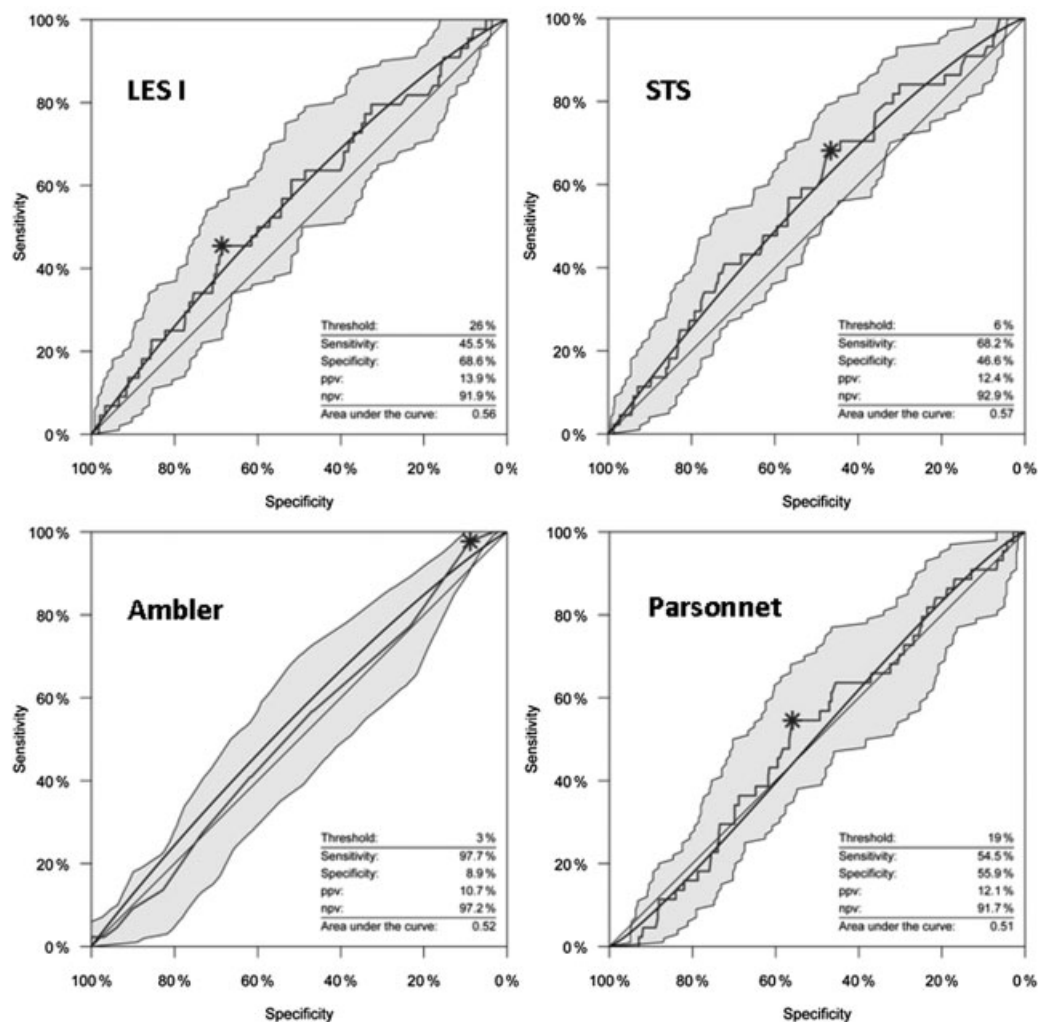


Fig. 3 c-statistics of tested scores.

LogEuroSCORE I and Parsonnet score overestimated observed mortality with 22.04 and 23.84%, respectively, whereas mean EuroSCORE II, Ambler score, and STS-PROM tended to underestimate mortality. These results give the impression that mortality would be lower if surgical aortic valve replacement (AVR) had been performed. It should be mentioned that there are some risk factors existent in the study patient population which are not reflected by the analyzed scores, the most important one among them being frailty. These risk factors are evidently more frequent among TAVI patients compared with surgical candidates and should be taken into consideration when judging mortality rates.

For all tested scores, survivors had similar mean scores compared with deceased patients, without significant differences (–Fig. 1). ROC curve analyses revealed that none of the five tested scores was a reliable predictor of mortality. AUC was highest in STS-PROM (0.57) but did not reach a value of 0.7 and does therefore not serve as a reliable predictor of perioperative mortality.

Even though the EuroSCORE II is an adequate predictor of mortality in a standard patient cohort undergoing cardiac surgery,¹¹ with an AUC of approximately 0.85 (95% CI 0.83–

0.87), this observation can obviously not be transferred to a high-risk population undergoing TAVI. Discrimination of a risk model depends on the heterogeneity of the case mix. It will by definition be poorer in a population consisting only of high-risk patients. The EuroSCORE II was developed by means of large patient databases predominantly undergoing coronary surgery, and no TAVI patient has been in the database for score development. Therefore, it may not be valid in specific subgroups, particularly high-risk patients receiving novel valve intervention. This observation does not seem unexpected; nevertheless, in the absence of specific risk prediction tools for TAVI, the logEuroSCORE I was introduced into clinical routine for TAVI evaluation, leading to misinterpretation. Our observations imply that conventional risk scores may only be valid for surgical candidates.

Concerning the different technical approaches of a TAVI procedure with transfemoral and transapical being the most widely used, subgroup analysis did not reveal significant differences in predictive capability either, as the maximum AUC was 0.63 for logEuroSCORE I in transapical (TA) TAVI and 0.61 for STS-PROM in transfemoral (TF) TAVI. A similar observation was made by Watanabe et al in an analysis of

453 TAVI patients.¹⁴ They observed an AUC of 0.68 for EuroSCORE II in the overall patient cohort and emphasized its inadequateness in predicting mortality after TAVI. In contrast to our study, their series indicated the EuroSCORE II had more precise performance in transfemoral patients, reaching an AUC of 0.75 (95% CI 0.63–0.88, $p = 0.01$) and it was therefore judged as an acceptable predictor of 30-day mortality. Similar to our observations, other tested systems (logEuroSCORE I, STS-PROM) also failed ROC analysis and were not adequate predictors. Stähli et al stated that the EuroSCORE II is a superior predictor compared with logEuroSCORE I and STS-PROM (AUC = 0.7).¹⁵ A similar observation was made by Sedaghat et al, who state that the logEuroSCORE I reflects 1-year mortality accurately, while EuroSCORE II is a better predictor of 30-day mortality.¹⁶ AUC in EuroSCORE II reported by these authors appears higher than in our cohort; however, AUC gets imprecise and 95% CI broader with smaller patient populations. The presented patient populations by Sedaghat et al and Stähli et al were relatively small (206 and 350 patients vs. 457 patients in our cohort), and confidence intervals for AUC in EuroSCORE II overlapped between ours and the reported studies (e.g., 0.70 [95% CI 0.61–0.78] in Stähli et al vs. 0.54 [0.46–0.63] in our group), suggesting that differences in AUC are not significant.

Furthermore, cohorts reported by Stähli et al as well as Sedaghat et al consisted mainly of patients treated via transfemoral access (83% in Stähli et al and 100% in Sedaghat et al), whereas in our cohort, 57.1% of patients received TA-AVI. Possibly, different access routes may have contributed to the observed difference in AUC. In addition, mortality rates differed (9.1% in our cohort vs. 6.8% in Sedaghat et al) as well as some baseline characteristics (mean age: 82.2 vs. 80.5) and mean score values.

In a recent publication, the EuroSCORE II was evaluated in a small group of 76 TA-AVI patients and its performance compared with conventional AVR.¹⁷ The authors found that correlation between logEuroSCORE I and EuroSCORE II was poor in TAVI, with correlation coefficient of only 0.382. A stronger correlation between logEuroSCORE I and EuroSCORE II was present in our cohort (0.708). While in our cohort sensitivity and specificity of all tested scores was poor, Goetzenich et al stated that the STS-PROM was high in specificity but very low in sensitivity, the logEuroSCORE I high in sensitivity and fair in specificity. In contrast to our findings, the STS-PROM had the lowest predictive ability of all scores in their TA-AVI cohort (AUC 0.561). Youden index calculated a cutoff value of 2.8% for EuroSCORE II in their entire population (TA-AVI + AVR), with a peak sensitivity of 100% and a fair specificity of 67%. This observation may be due to the small proportion of TAVI patients in their cohort, as the majority underwent surgical aortic-valve replacement (SAVR). In contrast, we calculated the maximum Youden index at a threshold of 7% in EuroSCORE II, within acceptably low sensitivity (50.0%) and specificity (65.3%). With this cutoff value, the EuroSCORE II consequently defined 163 patients as “high risk,” whereas 25 patients were defined as “non-high risk” below the EuroSCORE II of 7% but died within 30 days of follow-up. In our experience, the poorest test was

the Ambler score with the lowest accuracy of all tested scores of 17.8%. In daily clinical routine, it is rapidly obtained and variables are easily acquired. Nevertheless, it tested the majority of our cohort positive. This was because threshold was low (3%), and the majority of patients was characterized as high risk as their mean Ambler score was $6.9 \pm 5.4\%$. The highest specificity was provided by the logEuroSCORE I (68.6%), which to date has been the standard tool in daily clinical TAVI evaluation. For detail, see ►Figs. 2 and 3. In our experience, none of the tested scores was useful in prediction of mortality in TAVI patients, whereas in other publications, this observation was not as obvious as in ours. Watanabe et al reported an acceptable prediction of mortality by EuroSCORE II in transfemoral TAVI patients but a poor overall ability.¹⁴ Goetzenich et al described a good correlation of EuroSCORE II with acute outcome, but demanded an establishment of generalized cutoff values to identify suitable TAVI candidates.¹⁷ Whether a safe threshold for high-risk patients in TAVI can be provided by EuroSCORE II or whether different systems of preoperative risk stratification have to be established must be further evaluated in larger patient cohorts.

The poor performance of EuroSCORE II and other widely used scores in TAVI emphasizes the need for new scores, designed and calibrated for this particular subgroup of patients. As a consequence, novel risk stratification tools specifically designed for TAVI were published recently, one of the most important to mention being the SURTAVI model.¹⁸ The authors emphasized that routinely used scores do not correlate uniformly (e.g., STS and logEuroSCORE I) and therefore fail to provide standardized allocation of a patient to either TAVI or AVR. In our sample, logEuroSCORE I and STS-PROM correlated moderately. As the status “high risk” is often defined by the criterion “logEuroSCORE I > 20% or STS-PROM > 10%,” there should at least be a strong positive correlation between the respective scores, to provide a universally valid allocation to a high-risk patient cohort. The principles of the SURTAVI model are based on the division of TAVI candidates into three risk levels (low-, intermediate-, and high risk), dependent on the factor age and the number of additional risk factors present. Ten risk factors were defined by the authors, with several risk factors known from previous scores and supplemented by the factor *frailty* defined by the presence of at least one out of three variables: (1) Katz score “activities of daily living,” (2) ambulation “walking aid/assist,” and (3) diagnosis of (pre)dementia.¹⁸ Variables constituting a contraindication for TAVI were excluded from the model (e.g., coronary artery disease not amendable to percutaneous coronary intervention (PCI), infectious endocarditis). The SURTAVI model was tested in the Rotterdam database on patients with aortic stenosis undergoing either AVR or TAVI and verified by the fact that mortality differed significantly between the three levels of risk stratification ($p < 0.001$). Nevertheless, the model was retrospectively tested on a total amount of 1,082 cases and has not yet been validated and needs further evaluation in clinical practice. Besides the SURTAVI model, the German Aortic Valve Disease Score (AKL-Score) has been developed recently for isolated aortic valve procedures and is based on data from 11,794 patients

who received isolated AVR or TF-TAVI. Fifteen variables were identified to have significant impact on 30-day survival. The performance of the AKL-Score was assessed with ROC analysis and Hosmer–Lemeshow test and showed that predictive power was acceptable. AUC was 0.808 in ROC analysis and $p = 0.776$ in Hosmer–Lemeshow test, respectively.¹⁹ Hosmer–Lemeshow test revealed good calibration of all tested scores in our analysis. Moreover, this risk stratification model was calibrated on a mixed patient population that included surgical procedures and did not comprise transfemoral TAVI patients. A possible validity of the AKL-Score for TAVI candidates alone must further be evaluated in clinical practice.

Limitations

This is an observational, retrospective single-center study and as in any retrospective analysis may contain hidden bias. Therefore, conclusions drawn from results of our analyses have to be interpreted with caution. Furthermore, several risk scores such as the Northern New England and the New York State Score were omitted from the present article and not taken into account for conclusions.

Conclusion

None of the five tested risk stratification systems including the new EuroSCORE II provided adequate prediction of acute mortality in our large routine TAVI cohort. Likely, scoring systems derived from classic cardiac surgery databases are inadequate for risk prediction in TAVI patients. Therefore, specific risk models are needed for high-risk patients undergoing TAVI. Recently, the SURTAVI model and the AKL-Score were developed for clinical use. These new models need to be tested in clinical routine and proof of validity is pending at present. Until then, evaluation of perioperative risk has to rely on clinical judgment of individual patient factors by an interdisciplinary heart team consisting of cardiologists and cardiac surgeons.

Note

M.S. und L.C. contributed equally to this work.

Conflict of Interest

The authors state that there are no conflicts of interest with regard to the present study.

References

- 1 Roques F, Nashef SA, Michel P, et al. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. *Eur J Cardiothorac Surg* 1999;15(6): 816–822, discussion 822–823
- 2 Leontyev S, Walther T, Borger MA, et al. Aortic valve replacement in octogenarians: utility of risk stratification with EuroSCORE. *Ann Thorac Surg* 2009;87(5):1440–1445
- 3 Gummert JF, Funkat A, Osswald B, et al. EuroSCORE overestimates the risk of cardiac surgery: results from the national registry of the German Society of Thoracic and Cardiovascular Surgery. *Clin Res Cardiol* 2009;98(6):363–369
- 4 Nashef SA, Roques F, Sharples LD, et al. EuroSCORE II. *Eur J Cardiothorac Surg* 2012;41(4):734–744, discussion 744–745
- 5 STS Adult Cardiac Surgery Database Risk Model Variables – Data Version 2.73. Available at: <http://riskcalc.sts.org/STSWebRiskCalc273/de.aspx>. Accessed December 2012
- 6 Roques F, Michel P, Goldstone AR, Nashef SA. The logistic EuroSCORE. *Eur Heart J* 2003;24(9):881–882
- 7 Conradi L, Seiffert M, Franzen O, et al. First experience with transcatheter aortic valve implantation and concomitant percutaneous coronary intervention. *Clin Res Cardiol* 2011;100(4): 311–316
- 8 Seiffert M, Diemert P, Koschyk D, et al. Transapical implantation of a second-generation transcatheter heart valve in patients with noncalcified aortic regurgitation. *JACC Cardiovasc Interv* 2013; 6(6):590–597
- 9 Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. Generic, simple risk stratification model for heart valve surgery. *Circulation* 2005;112(2):224–231
- 10 Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. *Circulation* 1989;79(6 Pt 2):13–112
- 11 Carnéro-Alcázar M, Silva Guisasola JA, Requillo Lacruz FJ, et al. Validation of EuroSCORE II on a single-centre 3800 patient cohort. *Interact Cardiovasc Thorac Surg* 2013;16(3):293–300
- 12 Cao C, Ang SC, Indraratna P, et al. Systematic review and meta-analysis of transcatheter aortic valve implantation versus surgical aortic valve replacement for severe aortic stenosis. *Ann Cardiothorac Surg* 2013;2(1):10–23
- 13 Conradi L, Seiffert M, Treede H, et al. Transcatheter aortic valve implantation versus surgical aortic valve replacement: a propensity score analysis in patients at high surgical risk. *J Thorac Cardiovasc Surg* 2012;143(1):64–71
- 14 Watanabe Y, Hayashida K, Lefèvre T, et al. Is EuroSCORE II better than EuroSCORE in predicting mortality after transcatheter aortic valve implantation? *Catheter Cardiovasc Interv* 2013;81(6): 1053–1060
- 15 Stähli BE, Tasnady H, Lüscher TF, et al. Early and late mortality in patients undergoing transcatheter aortic valve implantation: comparison of the novel EuroScore II with established risk scores. *Cardiology* 2013;126(1):15–23
- 16 Sedaghat A, Sinning JM, Vasa-Nicotera M, et al. The revised EuroSCORE II for the prediction of mortality in patients undergoing transcatheter aortic valve implantation. *Clin Res Cardiol* 2013; 102(11):821–829
- 17 Goetzenich A, Deppe I, Schnöring H, et al. EuroScore 2 for identification of patients for transapical aortic valve replacement—a single center retrospective in 206 patients. *J Cardiothorac Surg* 2012;7:89
- 18 van Mieghem NM, Head SJ, van der Boon RM, et al. The SURTAVI model: proposal for a pragmatic risk stratification for patients with severe aortic stenosis. *EuroIntervention* 2012;8(2):258–266
- 19 Kötting J, Schiller W, Beckmann A, et al. 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(5):971–977