Surgical and transcatheter aortic valve procedures. The limits of risk scores

Marco Ranucci a, *, Fabio Guarracino b, Serenella Castelvecchio a, Rubia Baldassarri a, Remo Daniel Covello c, Giovanni Landoni d, for the ACEF Score Research Group

a Department of Cardiothoracic Anesthesia and Intensive Care, IRCCS Policlinico S. Donato, Via Morandi 30, 20097 San Donato Milanese, Milan, Italy
b Department of Cardiothoracic Anesthesia and Intensive Care Medicine, Azienda Ospedaliera Universitaria Pisana, Pisa, Italy
c Università Vita-Salute San Raffaele, Milan, Italy

Received 22 February 2010; received in revised form 26 April 2010; accepted 27 April 2010

Abstract

Transcatheter aortic valve implantation (TAVI) is an alternative to surgical aortic valve replacement in very high-risk patients with severe aortic stenosis. The present multicenter, retrospective study investigates the accuracy and calibration properties of the EuroSCORE and the age, serum creatinine, and ejection fraction (ACEF) score in selecting a population of patients to be referred to TAVI. The study includes 1053 surgical and 211 transcatheter procedures. The operative mortality rate within the surgical group was 2%. The EuroSCORE and the ACEF score had similar levels of accuracy; the ACEF score was well calibrated and the EuroSCORE overestimated the mortality risk. The observed mortality rate within the transcatheter group was 10.4%. Cut-off values corresponding to a mortality rate of 10% were 26 for the ACEF score and 2.5 for the EuroSCORE: both the EuroSCORE and the ACEF score may be used to refer patients to TAVI. However, they do not consider a number of ‘extreme’ risk conditions that may justify a transcatheter procedure even in absence of an overall elevated risk score. These risk conditions should be included in a specific risk model for referring patients for TAVI.

© 2010 Published by European Association for Cardio-Thoracic Surgery. All rights reserved.

Keywords: Risk-adjustment; EuroSCORE; Transcatheter aortic valve implantation

1. Introduction

Transcatheter aortic valve implantation (TAVI) was first used in humans in 2002 [1], and has to date been used in thousands of patients with severe symptomatic aortic stenosis at high-risk for surgical aortic valve replacement (AVR). In 2008, the European Association of Cardio-Thoracic Surgery and the European Society of Cardiology, in association with the European Association of Percutaneous Cardiovascular Interventions released a position statement on this matter [2]. This document reported that risk evaluation is of paramount importance in patient selection for TAVI, and that multidisciplinary clinical judgment is needed in the decision-making process, together with operative mortality risk prediction using the European System for Cardiac Operative Risk Evaluation (EuroSCORE) [3], the Society of Thoracic Surgeons Predicted Risk of Mortality (STS-PROM) [4], and the Amber score [5].

Different articles [6–11] reported a number of clinical conditions which may justify TAVI (previous chest X-rays irradiation; previous coronary operation with patent grafts; porcelain aorta; liver cirrhosis; hypoalbuminemia; poor nutritional status; frailty; severe respiratory failure; morbid obesity; concomitant cerebral aneurysm; anatomical abnormalities of chest wall; mediastinitis). Some authors [6–8, 11] investigated different risk scores in order to check their usefulness in identifying patients to refer for TAVI. The risk scores investigated were the additive [6] and logistic [6–8, 11] EuroSCORE, the STS-PROM [6–8, 11], and the Amber score [6]. The suggested values for the definition of patients suitable for TAVI are a predicted operative mortality rate >20% for the logistic EuroSCORE and >10% for the STS-PROM [2, 7, 8, 11]. The general assumption is that with values above these limits the observed operative mortality rate should exceed the one observed for TAVI, reported as about 12–15% (in-hospital) [9–11] or 20–30% (at one-year follow-up) [9, 11].

Recently, we demonstrated that a simple risk score based on age, serum creatinine, and ejection fraction (EF) (ACEF score) [12] can perform equivalently to the more complex EuroSCORE, still considered the gold standard in many European countries, in assessing mortality risk in elective cardiac surgical procedures.

This multicenter study is aimed to determine (a) the accuracy and calibration of the logistic EuroSCORE and of the ACEF score in predicting operative mortality in patients undergoing a conventional aortic valve operation, and (b) the effectiveness of these scores for selecting a patient population to be referred for TAVI rather than for surgery.
2. Methods

Three Italian Institutions (IRCCS Policlinico S. Donato, Azienda Ospedaliera Universitaria Pisana, and Università Vita-Salute San Raffaele) participated in this retrospective study. The study was based on the existing institutional databases of each Institution. The Ethics Committee or Institutional Review Board approved the study and waived the need for an informed consent of the patients.

Two cohorts of patients were included in the study population: (a) surgical cohort: patients who underwent isolated surgical AVR due to moderate to severe aortic valve stenosis with surgical indication in the period 2005–2009; patients who underwent AVR associated with coronary operations or other valve procedures, and patients who underwent AVR due to aortic valve regurgitation were excluded; (b) TAVI cohort: patients with severe aortic stenosis who received a TAVI procedure in the period 2007–2009.

Operative techniques of surgical AVR included full sternotomy approach, upper ministernotomy or right anterior thoracotomy as determined by surgeon preference, arterial cannulation in the ascending aorta, direct aortic clamping, moderate systemic hypothermia, and multidose, cold crystalloid or blood cardioplegia.

TAVI procedures were performed without cardiopulmonary bypass, implanting the balloon-expandable Edwards–Sapien Bioprosthesis (Edwards Life-sciences Inc, Orange, CA, USA) or the self-expandable CoreValve ReValving System (CRSTM, CoreValve Inc, CA, USA) through retrograde transfemoral or transaxillary approach, under general anesthesia or local anesthesia plus sedation.

All the Institutions routinely and prospectively collected patient data and stratified the operative mortality risk using the logistic EuroSCORE. Among the fields used to calculate the EuroSCORE, age, preoperative serum creatinine value, and left ventricular EF were available to calculate the ACEF score. The ACEF score and the resulting operative mortality risk were calculated according to the equations published in the original article [12].

Model discrimination (accuracy) of the two predictive models was assessed using a receiver operating characteristics (ROC) analysis, producing an area under the curve (AUC) with 95% confidence intervals (Cls).

Calibration of the model was assessed by comparing the observed mortality rate (% with 95% CI) with the predicted mortality rate (% with 95% CI) for both the models, and with a Hosmer–Lemeshow statistics (test difference between expected and observed mortality at different risk deciles), producing a y^2-value and a P-value that, if significant, was indicative of a poor calibration.

Differences between continuous variables were explored with a non-parametric test (Mann–Whitney U-test for independent samples).

For all the statistical calculations, a P < 0.05 was considered statistically significant.

An SPSS 13.0 (SPSS Inc, Chicago, IL, USA) computerized statistical program was used for all the statistical calculations.

3. Results

Details of the surgical (n = 1053) and TAVI (n = 211) patient population are reported in Table 1. Patients referred to TAVI had a higher risk profile than surgical patients, with significantly older age, lower EF, higher preoperative serum creatinine value, and higher EuroSCORE and ACEF score. Within the surgical group, the risk profile was worse in non-survivors, with all factors excepted EF being significantly different between survivors and non-survivors. Conversely, within the TAVI group the only factor that was different between survivors and non-survivors was the ACEF predicted mortality, with a significantly higher predicted mortality in non-survivors.

3.1. Operative mortality prediction for conventional surgical operations

Twenty-one (2%, 95% CI 1.2–2.8%) patients out of the 1053 patients who underwent surgical AVR died during the hospital stay or after discharge within 30 days from the operation (Fig. 1). The predicted mortality rate according to the logistic EuroSCORE was 6.2% (95% CI 5.8–6.6%), significantly (P = 0.001) overestimated. The predicted mortality rate according to the ACEF score was 2.5% (95% CI 2.3–2.7%), not significantly different from the observed.

Accuracy of the two models (Fig. 2) was slightly better for the ACEF score (AUC 0.733, 95% CI 0.647–0.819, P = 0.001) than for the logistic EuroSCORE (AUC 0.718, 95% CI 0.624–0.812, P = 0.001). Calibration of the ACEF score was good (y^2 12.5, P = 0.013 at the Hosmer–Lemeshow test) and poor for the logistic EuroSCORE (y^2 20.6, P = 0.008 at the Hosmer–Lemeshow test).

Table 1. Risk profile of the surgical and TAVI patient population

<table>
<thead>
<tr>
<th>Factor</th>
<th>Surgical (n = 1053)</th>
<th>TAVI (n = 211)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Survivors n = 1032 (98%)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>65.1 ± 13</td>
<td>64.8 ± 13.4</td>
</tr>
<tr>
<td>Ejection fraction (%)</td>
<td>55.2 ± 9.9</td>
<td>55.2 ± 9.9</td>
</tr>
<tr>
<td>Serum creatinine (mg/dl)</td>
<td>1.1 ± 0.7</td>
<td>1.1 ± 0.7</td>
</tr>
<tr>
<td>Logistic EuroSCORE</td>
<td>6.2 ± 6.5</td>
<td>6.1 ± 6.4</td>
</tr>
<tr>
<td>ACEF score</td>
<td>1.3 ± 0.5</td>
<td>1.3 ± 0.5</td>
</tr>
<tr>
<td>ACEF predicted mortality (%)</td>
<td>2.6 ± 2.9</td>
<td>2.6 ± 2.9</td>
</tr>
</tbody>
</table>

Data are expressed as mean ± standard deviation. *P < 0.001 non-survivors vs. survivors, within groups; ^P < 0.05 non-survivors vs. survivors, within groups; "P < 0.001 TAVI group vs. surgical group.

TAVI, transcatheter aortic valve implantation; ACEF, age, serum creatinine and ejection fraction.
Fig. 1. Predicted and observed mortality rates for surgical aortic valve replacement, and observed mortality rate for transcatheter aortic valve procedures. ES, EuroSCORE; ACEF, age, serum creatinine and ejection fraction; TAVI, transcatheter aortic valve implantation; CI, confidence intervals.

Fig. 2. Receiver operating characteristics analysis for accuracy of the EuroSCORE and the ACEF score in predicting operative mortality in surgical aortic valve replacement.

<table>
<thead>
<tr>
<th>Cut-off value</th>
<th>Observed mortality % (95% CI)</th>
<th>Predicted mortality % (95% CI)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic EuroSCORE 20</td>
<td>7.5 (0–16)</td>
<td>31 (27–34)</td>
<td>14.3</td>
<td>96.4</td>
<td>7.5</td>
<td>96.4</td>
</tr>
<tr>
<td>Logistic EuroSCORE 26</td>
<td>10 (0–24)</td>
<td>38 (33–44)</td>
<td>9.5</td>
<td>98.2</td>
<td>10.0</td>
<td>98.2</td>
</tr>
<tr>
<td>ACEF score 2.5</td>
<td>8.3 (0–20)</td>
<td>16 (12–21)</td>
<td>9.5</td>
<td>97.8</td>
<td>8.3</td>
<td>98.1</td>
</tr>
</tbody>
</table>

CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; ACEF, age, serum creatinine, and ejection fraction.

3.2. Risk stratification for TAVI patients selection

Twenty-two patients (10.4%) out of the 211 who underwent TAVI suffered operative mortality. Details of this patient population are shown in Table 1.

Different cut-off values for the logistic EuroSCORE and the ACEF scores were explored to identify a population of conventional surgery patients having a similar (10%) operative mortality risk. The cut-off value of 20 for the logistic EuroSCORE was included in the analysis by default being the value suggested by the existing guidelines.

The cut-off values corresponding to a 10% mortality risk were 26 for the logistic EuroSCORE and 2.5 for the ACEF score. A cut-off value of 20 for the logistic EuroSCORE, as proposed by the existing guidelines [2], corresponds to an observed mortality risk of 7.5% for surgical operations (Table 2).

Twenty patients were identified as high-risk patients for surgery according to a logistic EuroSCORE cut-off value of 26, with two operative deaths (10%). Twenty-four patients were identified as high-risk patients according to an ACEF cut-off value of 2.5, with two operative deaths (8.3%).

The two cut-off values identified for the logistic EuroSCORE and the ACEF score had a positive predictive value (PPV) and a sensitivity of about 10%, therefore, they can select a patient population with a mortality risk for conventional surgery of 10%. However, 19 patients who actually suffered operative mortality were below the identified cut-off values. These 19 patients had a mean (± standard deviation) logistic EuroSCORE of 8.1 ± 4.94, an ACEF score of 1.48 ± 0.35 and a predicted mortality rate (%) according to the ACEF score of 3 ± 1.6. Therefore, both the risk models did not identify them as high-risk patients.

4. Discussion

The results of our study confirm that the mortality risk in surgical stenotic AVR may be predicted with an acceptable level of accuracy with the logistic EuroSCORE. However, this risk model overestimates the actual mortality risk, with a predicted/observed mortality ratio of 3.1, therefore it is not well calibrated. Conversely, the ACEF score, which considers only three risk factors, has the same level of accuracy but is well calibrated, with a predicted/observed mortality ratio of 1.25. These results confirm, in the aortic valve surgery setting, the data already published on the ACEF score [12].

The observed mortality at 30 days in the TAVI group (about 10%) is in agreement with that reported in recently published articles on the transarterial approach [6–11, 13]. Both the risk models applied in the present study (Euro-
SCORE and ACEF) may identify a patient population with an operative mortality risk of 10%, and the adequate cut-off values are 2.5 for the ACEF score and 26 for the logistic EuroSCORE. The cut-off value of 20 for the logistic EuroSCORE, as suggested by the existing guidelines, appears too low in our series, and selects a patient population with an operative mortality risk of 7.5%. This confirms other observations highlighting that the EuroSCORE may not be the ideal risk score for selecting patients to be referred to TAVI (7, 8), or at least that its cut-off value should be higher than 20.

With the identified cut-off values of 26 (logistic EuroSCORE) and 2.5 (ACEF score) both tests have a specificity and a PPV which ranges around 10%, that means that they are correctly identifying a population with a mortality risk similar to the one identified for TAVI in the same institutions. Our results may be interpreted from different perspectives: actually, they demonstrate that both the risk scores could identify only two out of the 21 patients who died after the conventional operation, and who maybe could have benefited from a TAVI procedure. As a consequence, it is reasonable to refer patients for TAVI if they have a risk score above the identified cut-off values for the EuroSCORE and the ACEF score, but the main problem remains the selection of patients who may belong to a high-risk category even in presence of a low or medium risk profile as defined by these risk scores.

Risk models are defined by their discrimination accuracy, calibration, stratification capacity, and classification accuracy.

Discrimination accuracy means how well the model may discriminate between patients having or not having an event (mortality) and is usually assessed through a ROC curve analysis, with the AUC or c statistics as a measure of the accuracy. Calibration defines how well predicted events are reflected in observed events. Risk stratification capacity is defined as the proportion of patients stratified into relevant risk categories. For example, patients could be stratified according to the predicted operative mortality into low, medium, and high-risk classes. Classification accuracy is the ability to discriminate low from high-risk, and is related to the reliability of a model to make a binary discrimination at a predefined threshold. This last property is exactly what we would like to obtain from a risk model aimed at selecting patients for TAVI rather than for conventional surgery, and it makes little sense to test these model based on their accuracy and calibration as has been previously done [6–8].

Many ‘extreme’ and rare conditions (frailty, porcelain aorta, oxygen-dependent chronic obstructive pulmonary disease…) may justify a TAVI procedure even in absence of a very high preoperative risk score. It is our opinion that these conditions, rather than being incorporated into a new score, could be applied as ‘correcting tools’ for the existing scores. A similar approach was followed by the interventional cardiologists, who created the Clinical Syntax Score by simply multiplying the Syntax Score (coronary anatomy) by the ACEF Score, and demonstrated a higher accuracy of this score in predicting outcome after percutaneous coronary interventions [14]. The same approach could be followed for TAVI risk stratification, starting with the simple ACEF score and separately adding the possible multipliers represented by the ‘extreme’ risk conditions. Further studies are obviously needed to test this hypothesis.

References


