ORIGINAL ARTICLE

Aortic valve replacement for severe aortic stenosis in octogenarians: Patient outcomes and comparison of operative risk scores

António Tralhão, Rui Campante Teles, Manuel Sousa Almeida, Sérgio Madeira, Miguel Borges Santos, Maria João Andrade, Miguel Mendes, José Pedro Neves

Hospital de Santa Cruz, Western Lisbon Hospital Centre, Department of Cardiology, Carnaxide, Portugal

Hospital de Santa Cruz, Western Lisbon Hospital Centre, Department of Cardiothoracic Surgery, Carnaxide, Portugal

Received 14 September 2014; accepted 2 January 2015
Available online 8 July 2015

KEYWORDS
Aortic valve surgery; Aortic stenosis; Octogenarians; Risk assessment; EuroSCORE; STS score

Abstract

Introduction and Aim: Isolated aortic valve replacement (AVR) in octogenarians is associated with increased operative risk, due to higher prevalence of associated risk factors and other comorbidities, making outcome prediction essential. We sought to analyze operative mortality and morbidity and to compare the predictive accuracy of the logistic European System for Cardiac Operative Risk Evaluation score (EuroSCORE) I, EuroSCORE II and Society of Thoracic Surgeons (STS) score in this population.

Methods: We retrospectively enrolled 105 consecutive octogenarians with symptomatic severe aortic stenosis undergoing isolated AVR in a large-volume single center between January 2003 and December 2010 and calculated surgical risk scores.

Results: Mean logistic EuroSCORE I, EuroSCORE II and STS score were 14.6±11, 4.4±3.1 and 4.0±2.4%, respectively. Mean operative mortality was 5.7% (six patients). Two (1.9%) patients suffered an ischemic stroke, three (2.8%) required temporary hemodialysis and five (4.7%) had a permanent pacemaker implanted. Five (4.7%) required rethoracotomy. No myocardial infarction or sternal wound infection was observed. Calibration-in-the-large showed overestimation of operative mortality with logistic EuroSCORE I (p=0.036), whereas EuroSCORE II (p=1.0) and STS (p=1.0) showed good calibration. C-statistic values were 0.877 (95% CI 0.800–0.933) for logistic EuroSCORE I, 0.792 (95% CI 0.702–0.864) for EuroSCORE II and 0.702 (95% CI 0.605–0.787) for STS, without statistically significant differences.

* Corresponding author.
E-mail address: atralhao@gmail.com (A. Tralhão).

http://dx.doi.org/10.1016/j.repc.2015.01.016
0870-2551/Published by Elsevier España, S.L.U. on behalf of Sociedade Portuguesa de Cardiologia.
Introduction

Degenerative calcific aortic valve stenosis (AS) is the second most common valvular lesion in western countries, accounting for almost a third of native valve disease. Due to its degenerative etiology, the number of affected individuals rises sharply with age, reaching around 8.1% at the age of 85. The burden of AS is further heightened by increased longevity, meaning more elderly patients are expected to require therapeutic intervention in the future. Once severe AS becomes symptomatic, the natural course of the disease leads to a poor short-term prognosis, with only 25% survival at three years if left untreated. In the light of these findings, aortic valve replacement (AVR) is currently the recommended treatment (class I, level of evidence B) for severe symptomatic AS, in both European and American guidelines.

Currently, surgical AVR and, more recently, transcatheter aortic valve implantation (TAVI) are the only treatment options that have an impact on survival. A larger body of experience supports AVR as the reference treatment, with TAVI being reserved for patients unsuitable for AVR due to a high risk profile or inoperability. However, the precise operative risk cutoff figure above which AVR surgery should be withheld in a particular patient (for which age is often used as a surrogate) remains a matter of ongoing debate. Hence, there is a need to describe contemporary outcomes of conventional AVR as the population ages and transcatheter options become available. In recent decades, cardiac surgery has been performed with increased safety. Several series have consistently reported progressively lower operative mortality, ranging from 3.9% to 9%, in octogenarians undergoing AVR. Medium- and long-term survival after AVR surgery has also improved, paralleling age-matched healthy individuals.

In this regard, operative mortality scores can play a role in improving the identification of the subset of elderly patients who could be candidates for conventional surgery. The 2003 logistic European System for Cardiac Operative Risk Evaluation (EuroSCORE) I has gained widespread popularity as a screening tool. However, it has shown a tendency to overestimate mortality in contemporary surgical series.
of elective AVR patients. To overcome this limitation, an updated version (EuroSCORE II) was built that further refined the model variables, which could ultimately translate into better calibration and discriminative power. Overseas, the American counterpart of EuroSCORE II, the Society of Thoracic Surgeons (STS) score, has also been validated in a large cohort of patients and may be used as an alternative. However, controversy remains about the predictive accuracy of these scores in the elderly and their relative performance. We sought to assess the operative mortality of octogenarians undergoing isolated AVR and to compare the performance of the new EuroSCORE II against logistic EuroSCORE I and the STS score in predicting mortality outcomes in this population.

Methods

Between June 2003 and November 2010, 106 consecutive octogenarians with symptomatic severe AS underwent isolated AVR as their first surgical cardiac intervention in a high-volume center and were included in this study. Aortic valve disease severity was graded according to the European Society of Cardiology (ESC) guidelines on valvular heart disease. All patients underwent comprehensive transthoracic Doppler echocardiographic assessment at the institution’s core laboratory. Significant obstructive coronary artery disease requiring concomitant coronary artery bypass grafting (CABG) was excluded by coronary angiography. Patients with associated valvular lesions requiring intervention were also excluded from the study. All operations were performed using a standard approach, with a median sternotomy, the application of extracorporeal circulation and cold blood cardioplegia. Choice of type and size of valve prosthesis was left to the surgeon’s discretion.

Operative mortality was defined as death occurring within 30 days of surgery or during the index admission. Logistic EuroSCORE I, II and STS score were computed using the respective online calculators, from data retrospectively retrieved from patients’ medical records. Since data on ‘poor mobility’ were missing for a substantial proportion of patients, EuroSCORE II was calculated without this variable; an analysis by Barili et al. showed that omitting this variable did not reduce the discriminative accuracy of EuroSCORE II. Post-operative complications were defined as myocardial infarction (MI), stroke, need for hemodialysis, atrioventricular block requiring permanent pacemaker implantation, sternal wound infection and redo thoracotomy for bleeding. Variable and outcome data were collected retrospectively from patients’ medical records.

Continuous variables are reported as mean ± standard deviation or median (interquartile range [IQR]). Categorical variables are reported as proportions (%). Due to the small population size, calibration-in-the-large was performed using the chi-square test (two-sided), with a significant p-value indicating reduced calibration. The discriminative power of each score was assessed using receiver operating characteristic (ROC) curve analysis and C-statistics for a 95% confidence level. A value of 0.5 signifies absence of discriminative power, while a value of 1 signifies a perfect result. ROC curves were compared using the method proposed by DeLong et al., in which a non-significant p means equal discriminative power (95% confidence level). Statistical Package for the Social Sciences version 21.0 (IBM SPSS Modeler, Chicago, IL) and Medcalc version 6.0 (MedCalc Software, Ostend, Belgium) were used for data processing and statistical analysis.

Results

Preoperative variables

Patients’ mean age was 83.0±2.1 years and 39 (36.8%) were male (Table 1). The surgical procedure was performed electively in 96 patients (90.1%) and as an urgent intervention in the other 10 (9.9%). Twenty-four patients (22.6%) had left ventricular ejection fraction (LVEF) <50%, 37 (34.9%) were in NYHA class III or IV, 39 (36.8%) had angina, 16 (15.1%) had extracardiac arteriopathy and 18 (17.0%) presented creatinine clearance <30 ml/min/m².

Perioperative variables

A bioprosthetic valve was implanted in all patients (Table 2); Mitroflow™ was the most widely used model (66%). The most used ring size was 21 mm, with a median of 21 (19 mm: 14 cases; 21 mm: 49 cases; 23 mm: 40 cases; 25 mm: 3 cases).

Median mechanical ventilation time was 10 h (IQR 8–14 h), median intensive care unit stay was 24 h (IQR 14–60 h) and median hospital stay was 9 days (IQR 9–12 days).

Operative mortality, surgical risk scores and post-operative complications

Operative mortality occurred in six patients (5.7%). Causes of death were heart failure in three cases (pump failure in two patients, patient-prosthesis mismatch in one), suspected aortic rupture in one, sepsis in one and pneumonia in one.

Mean logistic EuroSCORE I, EuroSCORE II and STS score were 14.6±11%, 4.4±3.1% and 4.0±2.4%, respectively (Table 3 and Figure 1). Calibration-in-the-large showed

![Figure 1](image-url)
Non-fatal complications occurred in fifteen patients (14.2%). Two (1.9%) suffered an ischemic stroke (one with full recovery, one with mild residual hemiparesis), three (2.8%) required temporary hemodialysis and five (4.7%) had a permanent pacemaker implanted due to complete atrioventricular block. No MI or sternal wound infection was observed. Five (4.7%) patients required repeat thoracotomy due to hemodynamic instability or suspected ongoing blood loss.

**Discussion**

The present study aimed to assess up-to-date figures concerning operative mortality for isolated AVR in the geriatric population and to compare the performance of three commonly used risk calculation tools to predict operative outcomes.

**Operative results**

In this series of octogenarians undergoing isolated AVR for severe symptomatic AS in a single center, operative mortality was 5.6%. In a population generally regarded as high-risk, we consider this figure acceptable.

Elayda et al. published one of the first surgical series, reporting a mortality of 5.3% in 77 patients aged 80 or over undergoing AVR, between 1975 and 1991. The authors found LVEF <45% to be an independent predictor of mortality, together with hypertension and associated CABG. Since then, the number of published surgical series and observational studies has increased modestly, but not without some discrepancies (Table 4). Gehlot et al. described an operative mortality of 13% in 322 octogenarians, of whom 43% had concomitant CABG and 56% were urgent or emergent procedures, both of them independent predictors of operative mortality. The presence of coronary artery disease requiring CABG and clinical instability are two recognized factors that can contribute to increased operative mortality, by 2- and 4-13-fold, respectively. Taken together, these two studies seem to suggest a lower risk associated with isolated elective AVR surgery in octogenarians. However, two other surgical series show higher than average mortality rates. Bloomstein et al. reported mortality of 16.7% in 180 septuagenarians and octogenarians undergoing isolated AVR surgery. In this series, a valve size of 19 mm

---

**Table 1** Pre-operative variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>83.1 ± 2.2</td>
</tr>
<tr>
<td>Male (%)</td>
<td>39</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.59 ± 0.1</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>67.4 ± 13</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.5 ± 4.4</td>
</tr>
<tr>
<td>Smoking (current or past) (%)</td>
<td>44</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>81</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>20</td>
</tr>
<tr>
<td>Dyslipidemia (%)</td>
<td>68</td>
</tr>
<tr>
<td>Previous MI</td>
<td>10</td>
</tr>
<tr>
<td>Previous PCI</td>
<td>8</td>
</tr>
<tr>
<td>Previous cardiac surgery</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 2** Operative variables.

<table>
<thead>
<tr>
<th>Type of prosthetic valve implanted</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biologic</td>
<td>106</td>
</tr>
<tr>
<td>Carpentier-Edwards Perimount™ (%)</td>
<td>9</td>
</tr>
<tr>
<td>Carpentier-Edwards Perimount Myna™ (%)</td>
<td>2</td>
</tr>
<tr>
<td>Mitroflow™ (%)</td>
<td>70</td>
</tr>
<tr>
<td>Mosaic Ultra™ (%)</td>
<td>9</td>
</tr>
<tr>
<td>St. Jude Epic™ (%)</td>
<td>14</td>
</tr>
<tr>
<td>St. Jude Trifecta™ (%)</td>
<td>2</td>
</tr>
<tr>
<td>Prosthesis ring size (median, mm)</td>
<td>21</td>
</tr>
<tr>
<td>Duration of ICU stay (median, h)</td>
<td>24</td>
</tr>
<tr>
<td>Hospital stay (median, days)</td>
<td>9</td>
</tr>
</tbody>
</table>

ICU: intensive care unit.

---

overestimation of operative mortality with logistic EuroSCORE I (14.6 vs. 5.6%, p=0.036), whereas EuroSCORE II (4.4 vs. 5.6%, p=1.0) and STS score (4.0 vs. 5.7%, p=1.0) exhibited similar mean values, closer to the observed mortality.

ROC curve analysis (Figure 2) yielded significant C-statistic values of 0.877 (95% CI 0.800–0.933) for logistic EuroSCORE I, 0.792 (95% CI 0.702–0.864) for EuroSCORE II and 0.702 (95% CI 0.605–0.787) for STS score. Comparison of ROC curves using the DeLong method was not significant for differences in discriminative power (Table 3).
Table 3  Operative risk score analysis.

<table>
<thead>
<tr>
<th>Score Type</th>
<th>Mean ± SD</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Calibration-in-the-large</th>
<th>C-statistic (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic EuroSCORE I</td>
<td>14.6 ± 11</td>
<td>5.8</td>
<td>64.3</td>
<td>p=0.036</td>
<td>0.877 (0.800–0.933)</td>
<td>0.001</td>
</tr>
<tr>
<td>EuroSCORE II</td>
<td>4.4 ± 3.1</td>
<td>1.5</td>
<td>17.0</td>
<td>p=1.0</td>
<td>0.792 (0.702–0.864)</td>
<td>0.009</td>
</tr>
<tr>
<td>STS score</td>
<td>4.0 ± 2.4</td>
<td>2.3</td>
<td>17.6</td>
<td>p=1.0</td>
<td>0.702 (0.605–0.787)</td>
<td>0.090</td>
</tr>
</tbody>
</table>

CI: confidence interval; SD: standard deviation; STS: Society of Thoracic Surgeons.

Figure 2  Distribution and receiver operating characteristic curves for logistic EuroSCORE I (a), EuroSCORE II (b) and STS score (c).

Table 4  Examples of literature series of surgical AVR in elderly patients.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Mean age</th>
<th>No. of patients</th>
<th>Operative mortality (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elayda et al.</td>
<td>1993</td>
<td>82.6</td>
<td>77</td>
<td>5.7</td>
<td>27% mortality with additional surgical procedures</td>
</tr>
<tr>
<td>Gehlot et al.</td>
<td>1996</td>
<td>82.7</td>
<td>322</td>
<td>13.0</td>
<td>56% urgent or emergent procedures, 43% concomitant CABG</td>
</tr>
<tr>
<td>Bloomstein et al.</td>
<td>2001</td>
<td>76</td>
<td>180</td>
<td>16.7</td>
<td>23.8% mortality with body surface area &lt;1.8 m²</td>
</tr>
<tr>
<td>Mistlaen et al.</td>
<td>2001</td>
<td>73</td>
<td>400</td>
<td>3.8</td>
<td>Previous CABG did not influence mortality</td>
</tr>
<tr>
<td>Kohl et al.</td>
<td>2007</td>
<td>82.8</td>
<td>162</td>
<td>9.0</td>
<td>32% mortality for urgent procedures</td>
</tr>
<tr>
<td>Leontyev et al.</td>
<td>2009</td>
<td>82.7</td>
<td>282</td>
<td>10.8</td>
<td>EuroSCORE I imprecise for perioperative mortality discrimination</td>
</tr>
<tr>
<td>Ferrari et al.</td>
<td>2010</td>
<td>82.2</td>
<td>124</td>
<td>5.6</td>
<td>Very similar results to present study</td>
</tr>
<tr>
<td>Litmathe et al.</td>
<td>2011</td>
<td>82.7</td>
<td>304</td>
<td>5.8</td>
<td>50% mechanical valves</td>
</tr>
<tr>
<td>Dell’Amore et al.</td>
<td>2012</td>
<td>82.0</td>
<td>188</td>
<td>4.3</td>
<td>No mortality difference between AVR and AVR+CABG patients</td>
</tr>
<tr>
<td>Di Eusanio et al.</td>
<td>2012</td>
<td>82.1</td>
<td>638</td>
<td>4.5</td>
<td>One of the largest literature series</td>
</tr>
<tr>
<td>Saxena et al.</td>
<td>2012</td>
<td>83.4</td>
<td>531</td>
<td>4.0</td>
<td>One of the lowest mortality results</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>81.1 a</td>
<td>3208</td>
<td>6.8 a</td>
<td></td>
</tr>
</tbody>
</table>

AVR: aortic valve replacement; CABG: coronary artery bypass grafting.

a Mean value.
was associated with significantly increased mortality (26.1% in 66 patients with 19 mm valve vs. 20% in 55 patients with 21 mm valve vs. 23% in 64 patients with 23–29 mm valve, $p=0.017$). Moreover, there was a strong correlation between small body surface area, small valve size, and female gender. In a series by Leontyev et al., operative mortality was also distinctly higher (10.8%) than in other studies. These authors included patients with low (logistic EuroSCORE I $\leq$10%), intermediate (10–20%) and high ($>20\%$) operative risk. However, in the 102 low-risk patients, operative mortality was 7.2%. More recently, Di Eusano et al. reported 4.5% operative mortality in a large regional series of 638 octogenarians undergoing isolated AVR surgery, with a mean logistic EuroSCORE I of 13.0%±7.9, very similar to our results. In another large series, Saxena et al. reported a 4.0% 30-day mortality rate in 531 octogenarians undergoing isolated AVR surgery. Although no formal surgical risk score calculation was mentioned, precluding a more accurate comparison, these results are also in agreement with the present series.

Understandably, naturally diminished global physiological and organ reserve makes this population more vulnerable to operative complications. Nevertheless, in the present series we observed a low rate of complications in isolated AVR, similar to those reported in the recent series by Di Eusano et al. (1.3% for stroke, 0.3% for MI, 4.4% for complete atrioventricular block and 3.7% for redo thoracotomy).

**Comparison of operative risk scores**

All scores showed good discriminative power in our series, but the logistic EuroSCORE I overestimated mortality by almost three times the observed value. Previous studies have demonstrated a systematic exaggeration of operative mortality by a factor of two to three. As EuroSCORE was originally devised in 1995, a large part of this discrepancy can probably be attributed to natural improvements in surgical techniques and post-operative patient management, and changes in population case mix and outcomes in recent years. By contrast, both EuroSCORE II and STS score showed excellent calibration.

Nevertheless, the logistic EuroSCORE I showed good discriminative power, which was not statistically different from EuroSCORE II or STS score, so ultimately no particular score was favored. Barili et al. also found this lack of added value when using ROC curves to compare the two versions of EuroSCORE in general cardiac surgery. In another study by the same authors, the STS score showed a slightly higher C-statistic, but again without statistically significant difference.

Still, theoretical and methodological concerns have been voiced regarding calibration of EuroSCORE II, particularly regarding the validity of the Hosmer-Lemeshow test for this purpose. Barili et al. set out to answer such concerns in a recent study including more than 12,000 patients undergoing cardiac surgery. The authors employed four different methods to assess calibration (including the use of calibration curves) and demonstrated that EuroSCORE II has an optimal calibration performance until 30% actual mortality, beyond which it progressively overestimates mortality. On the basis of these findings, EuroSCORE II could be considered an appropriate tool to assess operative risk in the general octogenarian population, including our study cohort. However, in a truly high-risk population, such lack of calibration can be crucial, particularly when considering transcatheter therapy.

Our findings appear to suggest that EuroSCORE II should be the preferred tool for estimation of operative risk mortality in octogenarians undergoing isolated AVR surgery. The STS score, although equally well calibrated, has the potential disadvantages of being derived from an American cohort and being more cumbersome and time-consuming due to a larger set of variables.

More than a decade after the introduction of TAVI for the treatment of severe AS, should age be a decisive factor in the decision whether to deny AVR surgery? Naturally, age per se, or as a surrogate of an increased prevalence of operative risk factors, is always present in all risk scores. An analysis by Lung et al. showed that low LVEF and greater age were significantly associated with the decision to deny surgery to 72 out of 216 patients with severe AS aged ≥75. Neurological dysfunction was the only comorbidity significantly linked with the decision not to operate, which seems to suggest that age itself, rather than associated comorbidities, is used as a true criterion more than a surrogate marker to assess a patient’s pre-operative physiological status. Still, issues related to quality of life in the very old, which are inherently subjective and difficult to evaluate, should also be considered when deciding on the potential benefit of AVR.

Our patient cohort, although limited in size and inevitably suffering the limitations and biases of a single-center analysis, contributes to the growing evidence that isolated AVR surgery can indeed be performed with acceptable results in selected octogenarians. Furthermore, current age boundaries are progressively being challenged. Yerebakan et al. reported 7.6% overall operative mortality in a cohort of 119 nonagenarian patients, of whom 66 underwent isolated AVR.

Proper patient selection has a decisive role in achieving good operative results. A number of conditions such as LVEF <30%, severe renal failure, severe lung disease, emergent surgery and simultaneous CABG have been identified as markers of poor outcome. Each of these variables, when added to age as the sole variable contributing to increased operative risk, may advise against AVR, not only because of the expected high operative mortality but also due to the reduced long-term benefit of the intervention.

In an increasingly aging population, the number of patients who need a therapeutic intervention will continue to grow. From these results, and given the advent of TAVI and the data from the PARTNER trial, the question arises of what is the best treatment option. The answer can be found by an objective assessment of each case, and for the most difficult and doubtful cases, a multidisciplinary evaluation by the “heart team” would be desirable to find the best option for the patient. In this regard, the development of a new, integrated score, perhaps including variables such as frailty, could help to define the best treatment strategy for each individual case, similar to what the SYNTAX II score did for revascularization in coronary artery disease. Further studies are needed to confirm the feasibility of this strategy.
Conclusions

This study adds to the evidence gathered so far on the safety of isolated AVR in elderly patients, in both morbidity and mortality outcomes. However, further studies are needed to better assess the applicability of these results to patients at increased risk.

EuroSCORE II and the STS score should be the preferred operative risk assessment instruments, offering an important contribution to contemporary clinical decision-making in selected octogenarians with severe symptomatic AS.

Ethical disclosures

Protection of human and animal subjects. The authors declare that no experiments were performed on humans or animals for this study.

Confidentiality of data. The authors declare that no patient data appear in this article.

Right to privacy and informed consent. The authors declare that no patient data appear in this article.

Conflicts of interest

The authors have no conflicts of interest to declare.

References