Temporary validation of EuroSCORE model for assessing the results of mitral valve replacement surgery

Jean L. Chao Garcíaab, MD; Francisco J. Vázquez Roquea, MD PhD; Roger Mirabal Rodríguezb, MD; Gustavo J. Bermúdez Yeraa, MD; Ernesto Chaljub Bravoa, MD; Yury Medrano Planaa, MD; Noel L. Castillo Garcíaa, MD; Vielka González Ferrera, MD; Yolepsis Quintero Fleitesa, MD; José L. Valdés Canteroa, MD; Yoandy López de la Cruzb, MD; Luis A. González Borgesa, MD; Álvaro L. Lagomasino Hidalgoa, MD; and Deysi Rivero Valerónc, MD

a Department of Cardiac Electrophysiology and Pacing. Cardiocentro "Ernesto Che Guevara". Villa Clara, Cuba.
b Dr. Serafin Ruiz de Zárate Ruiz University of Medical Sciences. Villa Clara, Cuba.
c José Luis Miranda Pediatric Hospital. Villa Clara, Cuba.

Este artículo también está disponible en español

ARTICLE INFORMATION

Received: December 15, 2012
Accepted: January 20, 2012
Authors have no competing interests

Acronyms
MACE: major adverse cardiac events

On-Line Versions:
Spanish - English

ABSTRACT

Introduction y Objetive: The preoperative surgical risk estimate can be made through mathematical models or clinical prediction scales. The objective was to validate the EuroSCORE in order to assess the immediate results of mitral valve replacement surgery.

Method: Analytical, longitudinal study with a prospective follow-up of 158 patients who underwent surgery for mitral valve disease at the Cardiocentro Ernesto Che Guevara, in Villa Clara, from 2007 to 2010. EuroSCORE was calculated in all patients. Model calibration was assessed using the Hosmer-Lemeshow goodness of fit test. The discrimination capacity was analyzed by calculating the value of the area under the ROC curve.

Results: The mean EuroSCORE was 2.1 with a standard deviation of 1.45. Expected mortality was 3.9%, lower than the actual mortality that was 4.4%. The deceased had a mean score of 5.9, and those who survived 1.95 (p=0.00). Those who had serious complications showed a higher mean score than those who did not have them, 3.45 vs. 1.84; the difference is highly statistically significant (p=0.000). The area under the ROC curve was 0.97, with a significance level of 0.000 for those who died, and the presence of major cardiac adverse events with an area of 0.70 under the curve and a significance of 0.002.

Conclusions: The EuroSCORE is reliable to predict early mortality in patients who have undergone mitral valve replacement surgery in our hospital. EuroSCORE did not predict early major morbidity.

Key words: Heart Surgery, Mitral valve, Euroscore

Validación temporal del modelo EuroSCORE para la evaluación de los resultados de la cirugía de sustitución valvular mitral
ABSTRACT

**Introduction and Objective:** La estimación del riesgo quirúrgico preoperatorio puede hacerse a través de los modelos matemáticos o escalas de predicción clínica. El objetivo fue validar el Euroscore para evaluar los resultados inmediatos de la cirugía de sustitución valvular mitral.

**Método:** Estudio analítico, longitudinal, de seguimiento prospectivo en 158 pacientes intervenidos quirúrgicamente por enfermedad valvular mitral en el Cardiocentro Ernesto Che Guevara, de Villa Clara, durante los años 2007-2010. Se calculó el valor del Euroscore a todos los pacientes. La calibración del modelo se evaluó mediante la prueba de bondad de ajuste de Hosmer-Lemeshow. La capacidad de discriminación se analizó a través del cálculo del valor del área bajo la curva COR.

**Resultados:** El valor medio del Euroscore fue de 2,1 con una desviación típica de 1,45. La mortalidad esperada fue del 3,9 %, inferior a la observada que fue de 4,4 %. Los fallecidos presentaron medias del Euroscore de 5,9 vs. 1,95 de los que no fallecieron (p=0,00). Los que tuvieron complicaciones graves mostraron una puntuación media mayor que ante la ausencia de estas, 3,45 vs. 1,84, diferencias estadísticamente muy significativas (p=0,000). El área bajo la curva COR fue de 0,97, con significación de 0,000 para los que fallecieron, y la presencia de complicaciones graves con un área bajo la curva de 0,70 y significación de 0,002.

**Conclusiones:** El Euroscore es confiable para predecir la mortalidad precoz en los pacientes a los que se les ha realizado cirugía de sustitución valvular mitral en nuestro Hospital. El Euroscore no predijo morbilidad grave precoz.

**Palabras clave:** Cirugía cardíaca, Válvula mitral, Euroscore

INTRODUCTION

The analysis and prediction of the outcome of interventions in cardiovascular surgery is as old as cardiovascular surgery itself. They date back to a study in the eighties in the United States on clinical and angiographic predictors in myocardial bypass surgery¹.

The Health Care Financing Administration (HCFA) published in March 1986 the unadjusted results of mortality in hospitals operating patients with MEDICARE, which led to a change in the position of the Society of Thoracic Surgeons (STS) in the U.S. They believed that the use of mortality data without appropriate adjustment to the risk factors was inappropriate and misleading. From that moment on, there was an emergence of systems oriented to weight the results based on the severity of the disease and the presence of comorbidity, as conceived today¹.

The preoperative surgical risk assessment can be carried out in different ways, medical judgment is the oldest. However, this method is less efficient than the mathematical models or clinical prediction scales, which are just equations that are determined through statistical techniques after analyzing a sample in a given population.

The use of scales and scores developed at other institutions, which are based on populations and times different from those where they are used, is one of the limitations of mathematical methods. It can be said that virtually any model accurately coincides with another in the variables found. This is one reason why their estimates vary considerably when they are used in another universe.

The use of stratification models has become a requirement for surgical decision making in the specialty because they are used to weigh up the risks and benefits of the procedure. Their use in preoperative individual estimates may be the use most required by surgeons, anesthesiologists, intensivists and cardiologists, but the development of these systems was aimed at achieving global estimates on series of patients and not on individual cases².

The selection of a risk model appropriate to the performance of the surgical team is vital for the estimation of the results. For this reason, it is important to study the validity of stratification systems used to continuously assess surgical results; in this case, the surgical results of patients who have undergone mitral valve replacement surgery at the Cardiocentro “Ernesto Che Guevara”, in Villa Clara.
**METHOD**
An analytical, longitudinal, prospective follow-up study on the relationship between the preoperative outcome of EuroSCORE and the morbidity and mortality in the immediate postoperative period of mitral valve replacement surgery was conducted at the Cardiocentro “Ernesto Che Guevara” in Santa Clara between January 2007 and December 2010.

The variables included in the predictor EuroSCORE were identified in the medical records. For that purpose, a database was created in Excel. It was exported to SPSS® v.15.0, where the validation analysis was conducted. Values not explicitly reflected in these records were considered as absent risk factors at the time of testing.

The probability of death was calculated by adding the relative weights of each risk variable in each individual, according to the EuroSCORE.

The postoperative variables that were used in our study were: the fact of dying or not dying, and having major adverse cardiac events (MACE) or not having them.

Mortality: Death occurring during hospitalization as a result of surgical operation, regardless of the length of stay, or within 30 days after surgery.

Major adverse cardiac events: low cardiac output that needed intraaortic balloon pump counterpulsation or mechanical ventilation, severe arrhythmia (ventricular fibrillation and complete atrioventricular block), respiratory complications requiring mechanical ventilation for more than 48 hours; focal neurological lesions confirmed by clinical method, computed tomography, or both; diffuse encephalopathy for more than 24 hours requiring mechanical ventilation or showing a seriously disturbed mental state, kidney failure requiring dialysis or ultrafiltration, reoperation, mediastinitis, and generalized sepsis.

Model calibration was assessed using the Hosmer-Lemeshow goodness of fit test. Comparison of means was performed with the Mann Whitney test, once the normality of variables was determined using Kolmogorov-Smirnov test.

The lower the value of the statistic the better it calibrates the model. A value of p > 0.05 indicates that the model fits the data, and therefore predicts.

The discrimination capacity was analyzed by calculating the value of the area under the ROC curve. An excellent discrimination would be values greater than 0.97. A very good discrimination would be in the range from 0.93 to 0.96 and good discrimination between 0.75 and 0.92. The values below 0.75 indicate models that are deficient in their discrimination.

**RESULTS**
Of the 158 patients studied, 7 died, for a mortality of 4.4%, and 29 had MACE, accounting for 18.3% (Table 1).

The mean EuroSCORE was 2.1, with a standard deviation of 1.45 (Table 2), and had one as minimum and eight as maximum.

<table>
<thead>
<tr>
<th>Total of patients</th>
<th>Deceased</th>
<th>MACE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N°</td>
<td>%</td>
</tr>
<tr>
<td>158</td>
<td>7</td>
<td>4,4</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>158</td>
<td>2,1</td>
<td>1,45</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>Mean</th>
<th>Mortality (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expected</td>
</tr>
<tr>
<td>158</td>
<td>2,1</td>
<td>3,9</td>
</tr>
</tbody>
</table>

Table 3 shows that the expected mortality when applying the score was 3.9%, for a mean score of 2.1; which is lower than the mortality observed, that was 4.4%. The deceased showed higher mean scores than the living (5.9 vs. 1.95), with statistical significance of...
p=0.00 (Table 4), while those patients with major adverse cardiac events (Table 5) had a mean score higher than those who did not have them (3.45 vs. 1.84), hence these differences were highly statistically significant (p=0.000).

**Table 4.** Euroscore and its relationship with alive and deceased patients.

<table>
<thead>
<tr>
<th>Discharge status</th>
<th>n</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deceased</td>
<td>7</td>
<td>5.9</td>
<td>1.41</td>
</tr>
<tr>
<td>Alive</td>
<td>151</td>
<td>1.95</td>
<td>1.17</td>
</tr>
</tbody>
</table>

P = 0.00

**Table 5.** Euroscore relationship with the presence or absence of MACE.

<table>
<thead>
<tr>
<th>MACE</th>
<th>n</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>29</td>
<td>3.45</td>
<td>2.37</td>
</tr>
<tr>
<td>No</td>
<td>129</td>
<td>1.84</td>
<td>0.93</td>
</tr>
</tbody>
</table>

P = 0.000

**DISCUSSION**

The easiest way to validate a model is the direct comparison between expected mortality and observed mortality. Another way, more accurate and complex, is determining their discriminating power.

Careaga et al.⁹ considered the Euroscore to be a very effective tool for predicting mortality in patients undergoing mitral valve replacement, especially in those at high risk. On the other hand, Grinberg et al.¹⁰ found no relationship between the expected mortality and the observed mortality in patients who had undergone mitral valve surgery with the use of the EuroSCORE and Parsonnet 97.

Typically, patients who die have more risk factors; therefore, they have a greater weight on the score. This difference in the mean of this score between the living and the dead showed a high statistical significance.

This behavior is reasonable, since both, in our series and in other studies⁹,¹¹,¹², the deceased are related to higher scores in preoperative risk than those who survived. Among the authors who indicate this behavior are Careaga et al.⁹, Roques et al.¹¹ and Vieira et al.¹².

The area under the curve is a number between 0 and 1, which gives an idea of the discrimination capacity of the model, and is usually called C-statistic or C-index. Values which are equal or close to 0.5 make the model not applicable, because a positive result could be true or be a false positive. Values higher than 0.97 would mean there is an excellent discrimination. A very good discrimination would be in the range of 0.93 to 0.96, and good discrimination between 0.75 and 0.92⁵,⁸,¹³.

When it is said that the EuroSCORE has an area under the curve of 0.97 for predicting mortality, it means that if living and dead patients are randomly se-
lected, and the score is applied to them, 97% of those who died would have a value in the score higher than the living.

Scarò et al.\textsuperscript{14} conducted a study in Argentina with 123 patients to test the discriminating power of a risk score, and concluded that the EuroSCORE was not useful in predicting mortality in the population, especially in terms of medium and high risk, when comparing observed mortality with expected mortality.

In our work, areas under the curve greater than or equal to 0.80 were shown, demonstrating an excellent discriminating power, and, therefore, the effectiveness of the model as a predictor of mortality in our series. These results are superior to those of Heikki\textsuperscript{15}, Syed\textsuperscript{16}, Collart\textsuperscript{17} and Toumpoulis\textsuperscript{18} who used the same risk score and obtained areas under the curve indicating a good capacity of the model, but that are inferior in value to ours.

According to Careaga \textit{et al.}\textsuperscript{9}, the EuroSCORE could slightly overestimate expected morbidity; however, according to the standard deviation found, it is concluded that it can be used in our population.

Most risk scores in cardiac surgery are designed considering mortality as dependent variable; this is shown in the studies conducted by Pons\textsuperscript{19}, Parsonnet\textsuperscript{20} and Roques\textsuperscript{21}. Other risk models\textsuperscript{11} have used the combination of morbidity and mortality as a dependent variable.

In our study, when analyzing morbidity, the EuroSCORE had an area under the curve of less than 0.75. It is below the value observed in the mortality analysis, which confirms that the score was devised considering preoperative variables that are dependent on mortality and not on morbidity. This indicates that, in our hospital, this risk score must not be used to predict MACE. However, other authors\textsuperscript{22} did demonstrate the usefulness of this score to predict morbidity, length of hospital stay and costs, showing areas under the curve above 0.70. Andrade \textit{et al.}\textsuperscript{23}, in the Heart Institute of Pernambuco, found areas under the curve greater than 0.75 in all cases, to predict MACE and death, using as a basis the EuroSCORE.

CONCLUSIONS

The EuroSCORE was reliable to predict early mortality in patients who underwent mitral valve replacement surgery in our hospital, but did not predict the early onset of disease that would endanger the patient’s life.

REFERENCES


