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

Jean L. Chao García[ab]✉, MD; Francisco J. Vázquez Roque[a], MD PhD; Roger Mirabal Rodríguez[a], MD; Gustavo J. Bermúdez Yera[a], MD; Ernesto Chaljub Bravo[a], MD; Yury Medrano Plana[a], MD; Noel L. Castillo García[a], MD; Vielka González Ferrer[a], MD; Yolepsis Quintero Fleites[a], MD; José L. Valdés Cantero[a], MD; Yoandy López de la Cruz[a], MD; Luis A. González Borges[a], MD; Álvaro L. Lagomasino Hidalgo[a], MD; and Deysi Rivero Valerón[c], 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

On-Line Versions:
Spanish - English

✉ JL Chao García
CIMEQ - Calle 216 y 11B
Rpto Siboney, Playa CP 12100
La Habana, Cuba.
E-mail address:
jeanluis.ssp@infomed.sld.cu

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 diference 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 mayor 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 surgery[1].

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 today[1].

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 cases[2].

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[3].

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[3].

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[4-7].

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[5,8].

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 1. Patients who died or had MACE after mitral valve replacement Cardiocentro "Ernesto Che Guevara", 2007-2010.

Total of patients	Deceased		MACE	
	Nº	%	Nº	%
158	7	4,4	29	18,3

Source: Department of Statistics. Cardiocentro "Ernesto Che Guevara". Santa Clara, Villa Clara, Cuba.

Table 2. General behavior of EuroSCORE.

n	Mean	Standard deviation	Minimum	Maximum
158	2,1	1,45	1	8

Table 3. Relationship of EuroSCORE with expected mortality and observed mortality.

n	Mean	Mortality (%)	
		Expected	Observed
158	2,1	3,9	4,4

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.

Discharge status	n	Mean	Standard deviation
Deceased	7	5,9	1,41
Alive	151	1,95	1,17

P = 0.00

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

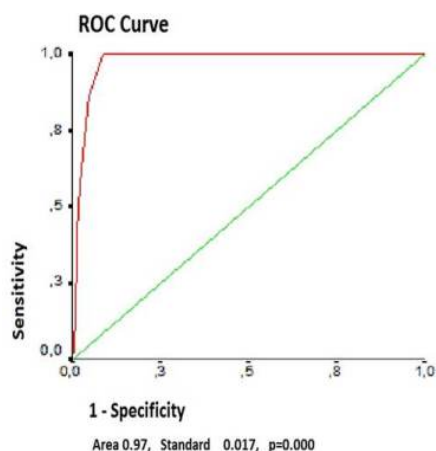| MACE	n	Mean	Standard deviation
Yes	29	3,45	2,37
No	129	1,84	0,93

P = 0.000

Area 0.97, Standard 0.017, p=0.000

Figure 1. ROC curve for predicting death, according to EuroSCORE. Cardiocentro "Ernesto Che Guevara", 2007-2010.
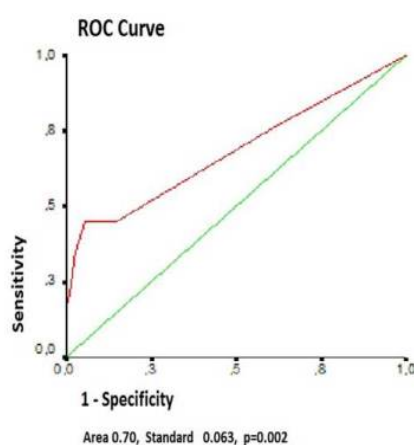

Area 0.70, Standard 0.063, p=0.002

Figure 2. ROC curve for predicting major adverse cardiac events, according to EuroSCORE.

The EuroSCORE predicted the mortality in our series with an area under the curve of 0.97 and a statistical significance of p = 0.000 (Figure 1), different from the presence of MACE (Figure 2) with an area under the curve of 0.70 and significance of p = 0.002.

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 studies[9,11,12], 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 0.92[5,8,13].

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.

Scaro et al.[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 Heikkinen[15], Syed[16], Collart[17] and Toumpoulis[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 *et al.*[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[19], Parsonnet[20] and Roques[21]. Other risk models[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[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 *et al.*[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

1. Kennedy JW, Kaiser GC, Fisher LD, Maynard C, Fritz JK, Myers W, *et al*. Multivariate discriminant analysis of the clinical and angiographic predictors of operative mortality from the Collaborative Study in Coronary Artery Surgery (CASS). J Thorac Cardiovasc Surg. 1980;80(6):876-87.
2. Cortina Romero JM. Scores de gravedad y complejidad en cirugía cardíaca. Usos y limitaciones. Rev Esp Cardiol. 2005;58(5):473-6.
3. Investigadores ESMUCICA. Estudio multicéntrico de cirugía cardiaca. Pacientes valvulares.. Rev Argent Cardiol. 2001;69(1):1-79.
4. Lemeshow S, Hosmer D. A review of goodness of fit statistic for use in the development of logistic regresión models. Am J Epidemiol. 1982;115(1):92-106.
5. Hosmer DW, Lemeshow S. Applied logistic regression. 2da ed. New York: John Wiley & Sons; 2000.
6. Vázquez Roque FJ, Pita Fernández S, Dueñas Fernández R, Lastayo Casanova R, Martínez Espinosa C, Hidalgo Menéndez P. El euroscore como predictor de mortalidad precoz y de complicaciones mayores en la cirugía coronaria a corazón latiendo. Rev Arg Cir Cardiovas [Internet]. 2005 [citado 11 Dic 2012];3(3):[aprox. 5 p.]. Disponible en: http://www.caccv.org.ar/raccv-anterior/Vol03N03/euroscorecomopredictor.htm
7. Lemeshow S, Klar J, Teres D. Outcome prediction for individual intensive care patients: useful, misused, or abused? Intensive Care Med. 1995;21(9):770-6.
8. Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. Ann Thorac Surg. 2005;79(1):16-20.
9. Careaga G, Martínez G, Anza LM, Avila A. Euroscore para predecir morbimortalidad en cirugía cardiaca valvular. Cir Ciruj. 2008;76:497-505.
10. Grinberg M, Masutti Jonke V, Sampaio Orismar R, Sobreira Spina G, Tarasoutchi F. Validação de un novo escore de riesgo cirúrgico para cirugía valvular. Arq Bras Cardiol . 2009;92 (4):320-5.
11. Roques F, Nashef SA, Michel P; EuroSCORE study group. Risk factors for early mortality after valve surgery in Europe in the 1990s: lessons from the EuroSCORE pilot program. J Heart Valve Dis. 2001; 10(5):572-7.

12. Vieira JC, Bodanese LC, Bueno FL, Goldani MA. Propuesta de score de riesgo preoperatorio para pacientes candidatos a cirugía cardiaca valvular. Arq Bras Cardiol. 2010;94(4):523-30

13. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operaiting characteristic (ROC) curve. Radiology. 1982 Apr;143(1):29-36.

14. Scaro G, Shireishi P, Gerbaudo L, Gallo C, *et al*. Valor predictivo del Euroscore en el hospital público. 4to Congreso Virtual de Cardiología [Internet]. Argentina: FAC; 2005. Disponible en:
http://www.fac.org.ar/ccvc/llave/tl447/tl447.pdf

15. Heikkinen J, Biancari F, Satta J, Salmela E, Mosorin M, Juvonen T, *et al*. Predicting immediate and late outcome after surgery for mitral valve regurgitation with EuroSCORE. J Heart Valve Dis. 2007;16(2):116-21.

16. Syed AU, Fawzy H, Farag A, Nemlander A. Predictive value of EuroSCORE and Parsonnet scoring in Saudi population. Heart Lung Circ. 2004;13(4):384-8.

17. Collart F, Feier H, Kerbaul F, Mouly-Bandini A, Riberi A, Mesana TG, *et al*. Valvular surgery in octogenarians: operative risks factors, evaluation of Euroscore and long term results. Eur J Cardiothorac Surg. 2005;27(2):276-80.

18. Toumpoulis IK, Anagnostopoulos CE, Swistel DG, De Rose JJ. Does EuroSCORE predict length of stay and specific postoperative complications after cardiac surgery? Eur J Cardiothorac Surg. 2005;27(1):128-33.

19. Vázquez Roque FJ, Fernández Tarrío R, Pita S, Cuenca JJ, Herrera JM, Campos V, et al. Evaluación preoperatoria del riesgo en la cirugía coronaria sin circulación extracorpórea Rev Esp Cardiol. 2005; 58(11):1302-9.

20. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in adquired adult heart disease. Circulation. 1989;79(6 Pt 2):I3-12.

21. Nashef S, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R; The Euroscore study group. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg.1999; 16(1):9-13.

22. Wang C, Yao F, Han L, Zhu J, Xu ZY. Validation of the European system for cardiac operative risk evaluation (EuroSCORE) in Chinese heart valve surgery patients. J Heart Valve Dis. 2010;19(1):21-7.

23. Andrade IN, Moraes Neto FR, Oliveira JP, Silva IT, Andrade TG, Moraes CR. Assessment of the EuroSCORE as a predictor for mortality in valve cardiac surgery at the Heart Institute of Pernambuco. Rev Bras Cir Cardiovascular. 2010;25(1):11-8.