



The usefulness of the EuroSCORE II model for predicting surgery mortality in a high specialty hospital in Mexico

Utilidad del modelo EuroSCORE II para predecir la mortalidad quirúrgica en un hospital de alta especialidad de México

Raúl Teniente-Valente,* Humberto Martínez-Bautista,† Miguel Ángel Chagolla-Santillán,* Iliana Acevedo-Bañuelos,* Ricardo Romo-Escamilla,* Iván García-Muñoz,* Mercedes Guadalupe Gutierrez-García,§ José Carlos Rodríguez-Jiménez,§ Benjamin Valente-Acosta,¶ Victor Hugo Vázquez-Martínez^{||}

Keywords:

EuroSCORE II, cardiac surgery, heart surgery risk calculator, validation.

Palabras clave:

EuroSCORE II, cirugía cardíaca, índice de riesgo en cirugía cardíaca, validación.

* Department of Cardiology and Cardiovascular Surgery of the HRAEB. León, Gto.

† Mathematics Research Center, Aguascalientes Unit. Ags.

§ Internal Medicine Resident. HRAEB. León, Gto.

¶ Internal Medicine and Infectious Diseases Service. Hospital ABC, Mexico City.

^{||} Research coordinator. Instituto Mexicano del Seguro Social. Reynosa, Tamaulipas.

ABSTRACT

Introduction: the EuroSCORE model has been used in various countries, including México, to estimate the probability of surgery-associated mortality. Several studies have shown deficiencies in its calibration while retaining good discrimination. The model was updated in 2012 and called EuroSCORE II. **Objective:** to evaluate the calibration and discrimination of the EuroSCORE II model in patients undergoing cardiac surgery in a high specialty hospital in México. **Material and methods:** an observational, cross-sectional, and retrospective study was performed. Patients ≥ 16 years old years who underwent cardiac surgery between the years 2008-2013 were included. The hospital mortality rate was obtained, and the EuroSCORE II was calculated online. Discrimination of the EuroSCORE II model was evaluated with the area under the curve of a receiver operating characteristics curve (AUC-ROC), and the calibration was assessed using χ^2 of Hosmer-Lemeshow (H-L) goodness of fit test and risk-adjusted ratio (RAMR). **Results:** three hundred thirty-eight patients were included. The mean age of participants was 49.8 ± 16.61 years; 162 were women (47.9%), and 176 (52.1%) were men. Surgery types were valvular 108 (31.9%), coronary revascularization 101 (29.8%), congenital 51 (15.08%), and other 78 (23.07%). The average EuroSCORE II was 4.1 (95% CI, 3.53-4.68). Mortality observed was 10.9%. The AUC-ROC was 0.806 (95% CI, 0.739-0.872), consistent with good discrimination. The χ^2 of H-L of 14.2, $p = 0.08$,

RESUMEN

Introducción: el modelo EuroSCORE ha sido utilizado en diversas latitudes, incluido México, para estimar la probabilidad de mortalidad operatoria. Varios estudios han mostrado deficiencias en su calibración, conservando buena discriminación. El modelo fue actualizado en el 2012, llamado EuroSCORE II. **Objetivo:** evaluar la calibración y discriminación del modelo EuroSCORE II en pacientes sometidos a cirugía cardíaca en un hospital de tercer nivel en México. **Material y métodos:** estudio observacional, transversal y retrospectivo. Se incluyeron pacientes ≥ 16 años operados entre 01/01/2008 y el 31/12/2013. Se registró la mortalidad hospitalaria y se calculó el puntaje de EuroSCORE II en línea. Se valoró la discriminación con el área bajo la curva (ABC) característica operativa del receptor (ROC). La calibración fue evaluada mediante prueba de χ^2 de bondad de ajustes de Hosmer-Lemeshow (H-L) y la razón ajustada al riesgo (RAMR). **Resultados:** 338 pacientes, edad media de los participantes 49.8 ± 16.61 años (16-80), 162 mujeres (47.9%). Tipos de cirugía: valvular 108 (31.9%), revascularización coronaria 101 (29.8%), congénitos 51 (15.08%) y otros 78 (23.07%). El puntaje promedio EuroSCORE II fue 4.1 (IC 95%, 3.53-4.68). Mortalidad observada 10.9%. El ABC ROC fue 0.806 (IC 95%, 0.739-0.872) compatible con buena discriminación. La χ^2 de H-L de 14.2, $p = 0.08$, compatible de calibración adecuada. La RAMR fue de 2.65, que indica infraestimación del modelo. **Conclusión:** el modelo EuroSCORE II mostró buena discriminación. La calibración fue adecuada de acuerdo

How to cite: Teniente-Valente R, Martínez-Bautista H, Chagolla-Santillán MÁ, Acevedo-Bañuelos I, Romo-Escamilla R, García-Muñoz I et al. The usefulness of the EuroSCORE II model for predicting surgery mortality in a high specialty hospital in Mexico. Cardiovasc Metab Sci. 2023; 34 (4): 150-158. <https://dx.doi.org/10.35366/113863>

Received:
07/02/2023
Accepted:
12/12/2023

compatible with adequate calibration. The RAMR was 2.65, indicating an underestimation of the model. **Conclusion:** the EuroSCORE II model showed good discrimination. The calibration was adequate according to the χ^2 of H-L, but the value of the RAMR suggests that the model underestimates the risk of mortality.

do con la χ^2 de H-L pero el valor de la RAMR sugiere que el modelo infraestima el riesgo de mortalidad.

INTRODUCTION

The European System for Cardiac Operative Risk Evaluation (EuroSCORE) is a probabilistic model designed in 1999 to estimate the risk of death in patients undergoing major cardiac surgery.^{1,2} The model has shown good calibration, with chi-square (χ^2) of Hosmer-Lemeshow (H-L) of 7.5, p2. It was initially published in an additive version and later in a logistic version, both of which have been evaluated in multiple latitudes and countries, showing generally good performance.³⁻⁷

Nevertheless, in recent years, some authors pointed out that the additive version of the EuroSCORE model overestimated mortality risk in low-risk patients and underestimated it in high-risk patients.⁸ On the other hand, several authors have shown that both the additive and logistic versions of the model overestimated risk, especially in high-risk patients.⁹⁻¹⁵ As a result of this situation, the Score underwent a redesign and was published in 2012, adopting the name EuroSCORE II.¹⁶

Some variables of the old model were redefined or modified; for example, renal function was evaluated by creatinine clearance with cut-off points of 85 mL/min, 50-84 mL/min, < 50 mL/min, and on dialysis, replacing the previous one based on serum creatinine levels. The variable neurological dysfunction was replaced by poor mobility due to neurological diseases and musculoskeletal diseases; pulmonary arterial hypertension was redefined as < 30 mm Hg, 30-55 mm Hg, and > 56 mm Hg. In total, 18 variables constitute the EuroSCORE II model. The new model has a predicted mortality of 3.95% and observed mortality of 4.18%, as well as discrimination evaluated through the AUC-ROC of 0.8095 and calibration through the χ^2 of H-L

of 15.48 with $p < 0.0505$. The calibration of the model was also assessed using the risk-adjusted risk ratio (RAMR), with a score of 1.058, which shows good calibration of the score.^{17,18}

The EuroSCORE II model has been subjected to external validation in different latitudes, with contradictory results.¹⁹⁻²⁵ It has nevertheless been used to evaluate the risk of mortality in patients undergoing cardiac surgery in our institution without confirmatory studies of its relevance.

Therefore, the present study aimed to evaluate the calibration and discrimination of the EuroSCORE II model in patients undergoing cardiac surgery at the *Hospital Regional de Alta Especialidad del Bajío* (HRAEB).

MATERIAL AND METHODS

Inclusion and exclusion criteria

An observational, cross-sectional, retrospective study was conducted at the *Hospital Regional de Alta Especialidad del Bajío* (HRAEB) in León, Guanajuato, Mexico. The study population consisted of all records (343) of patients who underwent cardiac surgery between 2008 and 2013. The inclusion criteria included having undergone cardiac surgery with and without a heart-lung machine and having all the information requested by the EuroSCORE II model. Four cases were eliminated since they did not contain the information requested by the model, and one case was triplicated, resulting in a total of 338 cases that were included.

Data collection

Four Cardiology and Cardiac Surgery Service physicians collected data from clinical records that met the inclusion criteria. A structured, purpose-built instrument was used to gather the information. The following variables were

Table 1: General characteristics and comorbidities of the study population (N = 338).

Characteristic	n (%)
Age [years], mean \pm SD (range)	49.87 \pm 16.6 (16-80)
Age groups [years]	
\leq 50	139 (41.12)
51-60	101 (29.88)
61-70	72 (21.30)
\geq 70	26 (7.69)
Gender	
Male	176 (52.07)
Female	162 (47.93)
Weight [kg], mean \pm SD	67.21 \pm 14.49
Height [cm], mean \pm SD	160.06 \pm 0.93
Body mass index [kg/m ²], mean \pm SD	26.07 \pm 5.05
Underweight [$<$ 18.5]	17 (5.03)
Normal weight [18.5-24.9]	135 (39.94)
Overweight [25.0-29.9]	111 (32.84)
Obesity I [30.0-34.9]	61 (18.05)
Obesity II [35.0-39.9]	11 (3.25)
Obesity III [\geq 40.0]	3 (0.89)
Diabetes mellitus type II	91 (26.92)
Insulin-dependent diabetes	42 (12.4)
Non-insulin dependent diabetes	49 (14.49)
Systemic arterial hypertension	153 (45.26)
Hypercholesterolemia	78 (23.08)
Hypertriglyceridemia	105 (31.07)
Smoking	104 (30.77)

SD = standard deviation.

obtained: demographics, comorbidities, unadjusted mortality until hospital discharge, defined as death occurring during the index hospitalization; type of surgery, defined as the procedure or procedures performed during the index surgery whether it was a) valve surgery, b) coronary revascularization surgery, c) surgery to correct congenital malformation(s) and d) surgery of a different type (valve surgery plus coronary revascularization surgery, aortic surgery, closure of postinfarction ventricular septal defect, traumatic heart injury, pericardial resection). In addition, the variables required for the calculation of the EuroSCORE II mortality risk score were collected using the online calculator on the EuroSCORE website: <https://www.euroscore.org/index.php?id=1>

Data of the participants, obtained from the clinical file, were subjected to an anonymization procedure to dissociate the personal data from the holder, not allowing the participant to be identified due to the structure, content or degree of disaggregation.

Statistical analysis

Quantitative variables are presented as means and standard deviations if normally distributed or as median and interquartile ranges when not normally distributed. Qualitative variables are presented as frequency percentages and compared with χ^2 , or Fisher's exact test. Quantitative variables of two groups were compared with Student's t-test when normally distributed and comparison of three or more averages with analysis of variance (ANOVA). A significance level of $p \leq 0.05$ was accepted.

Discrimination is the ability of a mathematical model to identify patients who will survive from those who will die (accuracy), and it was evaluated by the area under the receiver operating characteristic curve (AUC-ROC). Values ≤ 0.5 indicate that the model does not discriminate better than chance, and values of one indicate perfect discrimination. Values greater than 0.75 identify systems with good model discrimination capability.

The calibration compares the expected episodes with the observed ones across the risk range. It was evaluated using the χ^2 of the H-L goodness-of-fit test, which calculates a C-statistic which measures the difference between the model's expected mortality values and the mortality values observed in risk decile groups of the population studied. The lower the value of this statistic and the p -value > 0.05 , the better the calibration of the model (expected and observed mortality are close, and there is no statistical difference between them). A p -value greater than 0.05 suggests that the model has a good calibration and consequently predicts the probability of dying for patients across the risk range well.

We also calculated the RAMR obtained by the coefficient of observed mortality to expected mortality (RAMR = O/E), which has also been proposed to evaluate calibration.¹⁸ A ratio of 1.0 means that the score or test model predicts mortality in a perfect way (the same number

of [observed] patients die as the number of expected [predicted] patients). A RAMR > 1.0 means that the model underestimates mortality, while a RAMR < 1.0 implies that the model overpredicts (overestimates) mortality.¹⁸

Finally, to obtain final results, a 1000-sample bootstrap procedure was conducted, using robust errors, taking care of any possible errors

derived from the distribution of the data and the reduced sample size.

Microsoft Excel spreadsheet was utilized as the database, and descriptive and inferential statistics were performed in Stata version 16.

RESULTS

The present study was a retrospective review of 338 consecutive patient records of patients who underwent cardiac surgery with or without a heart-lung machine at the HRAEB. The mean age and standard deviation (SD) of the evaluated population were 49.9 ± 16.6 years with a range of 16-80 years. 47.9% were women. The average weight was 67.2 ± 14.4 kg; 32.8% were overweight, and 18% had obesity. 26.9% had diabetes mellitus, a higher figure than that observed in our general population; 45.2% had systemic arterial hypertension, also higher than that observed in our general population, and 30.7% were smokers (Table 1).

The surgical procedures performed were valve surgery 108 (31.95%), coronary artery bypass surgery 101 (29.88%), congenital surgery 51 (15.08%), and miscellaneous surgery 78 (23.07%). The aortic clamping time and extracorporeal circulation time were 88.7 ± 43.4 and 118.88 ± 54.4 minutes, respectively. Thirty-seven patients died during the index hospitalization (10.9%) (Table 2). Table 3 shows the number and percentage of each of the variables of the EuroSCORE II model found in our study population. Table 4 compares the variables found in the population from which the EuroSCORE II was obtained and the population evaluated at the HRAEB. The population from which the EuroSCORE II model was derived was almost 15 years older than our study population (64.6 vs 49.9 years). Body weight (77.9 vs 67.2 kg) and height (168.5 vs 160 cm) were also higher. The prevalence of chronic obstructive pulmonary disease (10.7 vs 6.2), the percentage of emergent surgery (4.3 vs 0.9), isolated coronary revascularization (46.7 vs 29.9), and valve surgery (45.5 vs 31.9) were also more frequent in the EuroSCORE II population than in our population.

In contrast, the percentage of women (47.9% vs 30.9%), insulin-dependent diabetes mellitus (12.4% vs 7.6%), poor mobility (7.4% vs 3.2%), infective endocarditis (7.4% vs 2.2%),

Table 2: Surgical procedures.

Procedure	n (%)
Valvular	108 (31.95)
Aortic valve replacement	33 (9.76)
Mitral valve replacement + tricuspid repair	23 (6.80)
Mitral valve replacement	20 (5.19)
Mitral and aortic valve replacement	12 (3.55)
Tricuspid valve replacement	7 (2.07)
Mitral valve replacement + coronary revascularization	6 (1.77)
Mitral, aortic, and tricuspid valve replacement	4 (1.18)
Valve conduit + coronary reimplantation	3 (0.88)
Isolated coronary artery bypass grafts	101 (29.88)
Off-pump coronary artery bypass	77 (76.25)
Single-vessel coronary artery bypass	8 (2.36)
Double-vessel coronary artery bypass	43 (12.72)
Triple-vessel coronary artery bypass	46 (13.60)
More than triple-vessel coronary artery bypass	4 (1.18)
Congenital anomalies	51 (15.08)
Atrial septal defect	24 (7.10)
Patent ductus arteriosus	7 (2.07)
Coarctation of the aorta	5 (1.47)
Ventricular septal defect	4 (1.18)
Tetralogy of Fallot	3 (0.88)
Ebstein's anomaly	1 (0.29)
Other congenital anomalies	7 (2.07)
Miscellaneous	78 (23.07)
Pericardial window/resection	34 (10.05)
Thoracic aorta	23 (6.80)
Myxomas	6 (1.77)
Epicardial permanent pacemaker	5 (1.47)
Post-infarction ventricular septal rupture	3 (0.88)
Metastatic tumors	3 (0.88)
Stab wounds	2 (0.59)
Gunshot wounds	1 (0.29)
IV septum rupture + free wall rupture due to MI	1 (0.29)
Aortic clamp time, mean \pm SD	88.7 ± 43.4
Extracorporeal circulation time, mean \pm SD	118.88 ± 54.4
Patients who died	37/338 (10.95)

Table 3: Distribution, number, and percentage of EuroSCORE II variables in the study.

Characteristic	n (%)
Patient-related factors	
Age [years], mean ± SD	49.87 ± 16.61
Women	162 (47.93)
Extracardiac arteriopathy	19 (5.62)
Creatinine clearance (Cockcroft-Gault) [mL/min]	
85	201 (59.46)
50-85	101 (29.88)
< 50	16 (4.73)
On dialysis	20 (5.91)
Poor mobility	25 (7.39)
Previous cardiac surgery	47 (13.90)
Chronic obstructive pulmonary disease	21 (6.21)
Active endocarditis	24 (7.10)
Critical preoperative condition	49 (14.49)
Diabetes under insulin control	42 (12.42)
Cardiac-related factors	
NYHA functional class	
Class I	46 (13.6)
Class II	172 (50.9)
Class III	95 (28.1)
Class IV	25 (7.4)
SCC class IV angina	39 (11.5)
Left ventricular ejection fraction	
More than 50%	223 (65.97)
31-50	113 (33.43)
21-30	2 (0.59)
≤ 20	0 (0.00)
Recent myocardial infarction	46 (13.64)
Pulmonary artery systolic pressure [mmHg]	
No	115 (30.02)
Moderate [31-54]	171 (50.59)
Severe [≥ 55]	53 (15.68)
Surgery-related factors	
Type of surgery	
Elective	179 (53.0)
Urgent	155 (45.9)
Emergency	3 (0.9)
Salvage	1 (0.3)
Extent of surgery	
Isolated CABG	90 (26.6)
One non-CABG procedure	189 (55.9)
Two procedures	38 (11.2)
Three procedures	21 (6.2)
Surgery on the thoracic aorta	23 (6.8)
NYHA = New York Heart Association. SCC = Canadian Society of Cardiology. CABG = coronary artery bypass grafts.	

and critical preoperative status (14.5% vs 1.7%) were higher in the present series. The predicted risk in the EuroSCORE II model was 3.95%, and in this study group, it was 4.10%.

The EuroSCORE II model in the present study had an AUC-ROC of 0.806 (95% CI, 0.739-0.872), consistent with good discrimination (Figure 1). The χ^2 of H-L was 14.2, with $p = 0.08$, which is compatible with good calibration. However, the other tool proposed to evaluate calibration, the RAMR,¹⁸ was 2.65, consistent with the EuroSCORE II model, generally underestimating perioperative mortality in our series (Figure 2).

DISCUSSION

Implementing a mathematical model to predict operative mortality in cardiac surgery requires evaluating its performance in the hospital where it is to be used. The present study included a retrospective series of 338 adult patients who underwent cardiac surgery and aimed to evaluate the discrimination and calibration of the EuroSCORE II model in the HRAEB.

The EuroSCORE model for predicting perioperative mortality in general cardiac surgery in adult patients was widely used in the first decade and part of the second decade of the present century in various latitudes, showing generally good predictive performance. Nevertheless, it has shown deficiencies in its calibration while retaining good discrimination in recent years. For this reason, the model was updated in 2012 and called EuroSCORE II.

The new model improved its discrimination: AUC-ROC 0.8095 (95% CI, 0.7820-0.8360) and its calibration, χ^2 of H-L of 15.48, with $p < 0.0505$.¹⁶ In addition to being evaluated by the classical χ^2 method of H-L goodness-of-fit, the calibration was also assessed by the RAMR, which was 1.058.

Like the EuroSCORE II, our study included patients who underwent general cardiac surgery (ischaemic, valvular, congenital, and mixed) with and without a heart-lung machine. Compared to the population used by Nashef et al., from which the EuroSCORE II model was derived,¹⁶ our population appears to have advantages concerning risk: younger

population, lower prevalence of chronic obstructive pulmonary disease, and emerging surgeries. However, it was more prevalent in other relevant prognostic variables: more women, insulin-dependent diabetes mellitus, poor mobility, infective endocarditis, and preoperative critical condition (Table 4). Therefore, it is unsurprising that the EuroSCORE II in our population was higher but only slightly compared to the original EuroSCORE II model population, 4.10% vs 3.95%.

On the other hand, several studies have evaluated the discrimination and calibration of the EuroSCORE II model with contradictory results. Some studies corroborate good to very good discrimination and good calibration, while other studies, although corroborating good discrimination, question its calibration. Di Dedda et al., in a retrospective series of 1,090 patients, report discrimination, with

AUC-ROC of 0.81 (95% CI, 0.78-0.83) and good calibration, with an observed mortality of 3.75%, expected of 3.10 and conclude that «the EuroSCORE II represents a useful update of the previous version of the EuroSCORE, with much better clinical performance and the same good level of accuracy».²⁶ Barili et al., in a retrospective validation study of the new model involving 12,325 general cardiac surgery patients report an AUC-ROC of 0.82 (95% CI, 0.80-0.85) consistent with good discrimination and «optimal calibration but only up to 30% of predicted mortality».²³ Gao et al., in a series of 1,628 Chinese patients, reported good discrimination with AUC-ROC of 0.90 and good calibration, with the χ^2 of H-L of 0.071 ($p > 0.05$). However, discrimination and calibration decreased efficiency up to five years after patient follow-up.²⁷ Borracci et al., in a prospective series of 2,000 Argentinean

Table 4: Variables from the present study and the EuroSCORE cohort.

Variable	HRAEB	EuroSCORE II
Number, n	338	22,381
Age [years], mean	49.9	64.6
Women, %	47.9	30.9
Weight (kg), mean	67.2	77.9
Height (cm), mean	160	168.5
Diabetes mellitus (total), %	26.9	25.0
Insulin-dependent diabetes, %	12.4	7.6
Chronic obstructive pulmonary disease, %	6.2	10.7
Poor mobility, %	7.4	3.2
Extracardiac arteriopathy, %	5.6	
Infective endocarditis, %	7.4	2.2
Serum creatinine [mg/dL]	1.2	1.3
Ejection fraction, %		
> 50	65.9	
31-50	33.4	
21-30	0.6	
< 20	0.0	
Critical preoperative status, %	14.5	1.7
Emergency surgery, %	0.9	4.3
Isolated CABG, %	29.9	46.7
Valve surgery, %	31.9	45.5
EuroSCORE II, %	4.1	3.9

HRAEB = *Hospital Regional de Alta Especialidad del Bajío*. CABG = coronary artery bypass grafts.

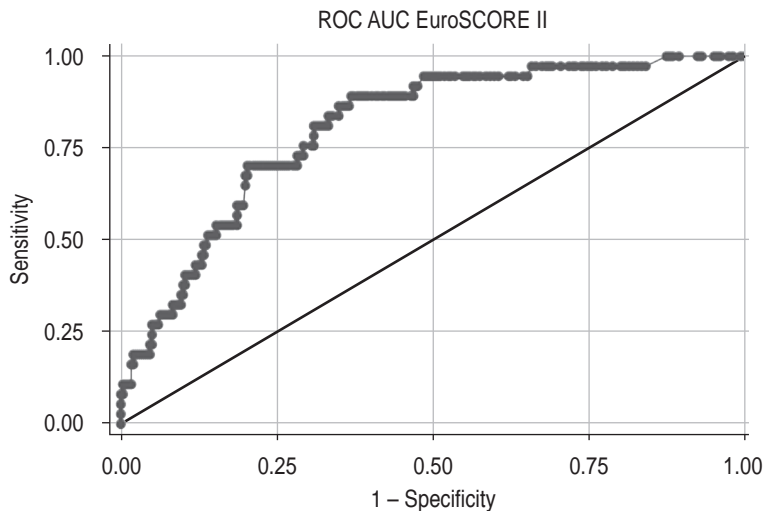


Figure 1: Area under ROC curve = 0.8057 (95% CI, 0.738-0.873).

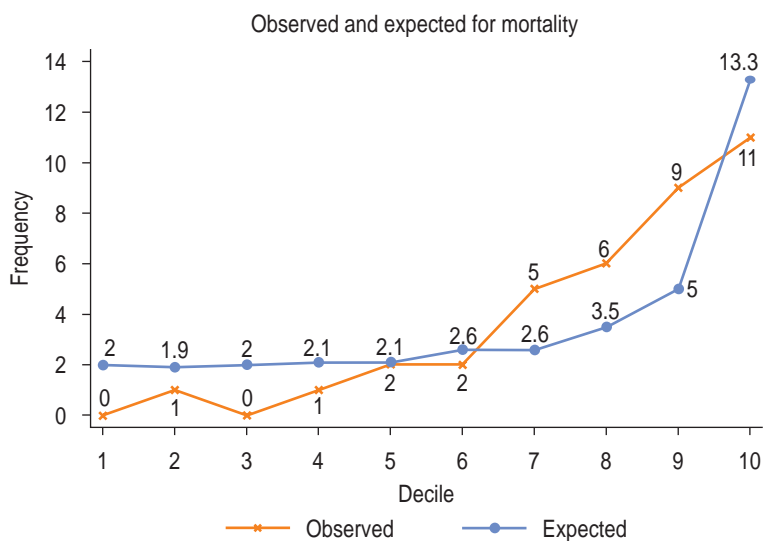


Figure 2: Expected mortality vs observed mortality.

patients undergoing general cardiac surgery, reported an AUC-ROC of 0.80 (95% CI, 0.75-0.85), compatible with good discrimination and χ^2 of H-L values of 11.4 ($p = 0.178$), consistent with good calibration.²⁸ Kinkel et al., in a retrospective series of 704 adult patients undergoing general cardiac surgery, found an AUC-ROC of 0.821 (95% CI, 0.772-0.871) and a χ^2 of H-L = 17.7, $p = 0.64$, consistent with good discrimination and calibration. In the present retrospective study of 338 patients who underwent general cardiac surgery, we found

AUC-ROC of 0.806 (95% CI, 0.739-0.872) compatible with good discrimination and a χ^2 of H-L of 14.2, $p = 0.08$, suggestive of adequate calibration.²⁴

García-Valentín et al. report the results of a prospective study conducted in Spain to validate the EuroSCORE II in which 20 Spanish hospitals participated, recruiting 4,034 adult patients who underwent general cardiac surgery. The AUC-ROC was 0.78 (95% CI, 0.76-0.82), compatible with good discrimination, and χ^2 of H-L of 38.98 ($p < 0.001$), consistent with poor calibration.²⁹ In the same sense, Kunt et al., in a retrospective series of 428 adult patients from Turkey who underwent coronary artery bypass surgery, found an AUC-ROC of 0.72 (95% CI, 0.62-0.81), compatible with acceptable discrimination. The observed mortality was 7.9%, the predicted mortality by EuroSCORE II was 1.7%, and they concluded that while the model showed good discrimination, it significantly underestimated the risk of perioperative death.²²

The RAMR has yet to be used in most studies evaluating the calibration of the EuroSCORE II model, with some exceptions. For example, Alvarez-Cabo, in an ambispective series of 206 adult Spanish patients undergoing coronary revascularization surgery, in addition to the χ^2 of H-L to evaluate calibration, also used the RAMR, with a point value of 0.83, suggesting a slight overestimation of the model, supported by the 95% confidence interval of the point value of the RAMR reported.²⁹ Similarly, Borracci et al., in their prospective series of 2,000 Argentinean adult patients undergoing general cardiac surgery, used in their calibration evaluation in addition to the χ^2 of H-L the RAMR, whose point value was 1.4 suggestive of slight underestimation of risk; «the clinical validation of the model, based on the ratio of observed/expected mortality, showed that the system performed better in the lowest and highest risk groups while underestimating the risk in the intermediate groups.»²⁸

We used the present work's H-L χ^2 and the RAMR to evaluate the calibration. According to χ^2 of H-L, the model has an adequate calibration. Nevertheless, according to the RAMR (point value 2.65), the model, in general, underestimates the risk of death in our study

population; in graph 2, we can see how, in the first deciles, the model overestimates the risk, but in the high-risk deciles, underestimation of the risk predominates.

In this study, the observed mortality of 10.9% stands out, higher than that reported in the EuroSCORE II and multiple validation studies in European and Anglo-Saxon populations, but similar to that reported in our country: 9.68% in the Rodríguez-Chávez series of 1,188 valve surgery patients to validate the EuroSCORE³⁰ and 12.5% at 30 days reported in the Kinkel series of 704 patients undergoing general cardiac surgery.²⁴ The explanation for this high mortality is given by factors related to the patient: biological status, socioeconomic status, and low health education that leads to seeking medical attention late, when the condition has already produced advanced structural and functional cardiac and extracardiac damage,³⁰ as well as factors related to health care centers: resources, infrastructure, experience, previous results, among others.³¹

In summary, the performance of the updated EuroSCORE II model evaluated in our population showed good discrimination. As assessed by χ^2 of H and L, the calibration suggests adequate calibration. However, when calibration was evaluated using the RAMR, the result was consistent with underestimating the risk of death, in line with the observed mortality.

Limitations

The study's period and the presentation of the results are long. However, the model has retained relevance and contributes to re-evaluating the EuroSCORE II calibration measure. Future studies should include patients who intervened during recent years and increase the sample size or design a prospective study with an adequate sample size to have a fairer evaluation of the performance of the EuroSCORE II model in the Mexican context.

CONCLUSIONS

In the present study, the EuroSCORE II model showed good discrimination and adequate calibration based on the χ^2 data of Hosmer

and Lemeshow. However, the data obtained from the RAMR indicates that the model underestimates the risk of death in the medium- and high-risk groups of patients. Considering the limitations mentioned above, we considered continuing the EuroSCORE II instrument in the hospital where the study was conducted. Ultimately, it is necessary to design a model to measure the risk of operative mortality in cardiac surgery in Mexico that includes characteristics and variables specific to the Mexican population, not contemplated so far by traditional international instruments.

REFERENCES

1. Roques F, Nashef SAM, Gauducheau ME, de Vicentis C, Baudet E, Cortina MD 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: 816-823.
2. Nashef SAM, Roques F, Michel P, Gauducheau ME, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg.* 1999; 16: 9-13.
3. Roques F, Michel P, Goldstone AR, Nashef SA. The logistic EuroSCORE. *Eur Heart J.* 2003; 24 (9): 881-882.
4. Pitkanen O, Niskanen M, Rehnberg S, Hippelainen M, Hynynen M. Intra-institutional outcome prediction after cardiac surgery: comparison between a locally derived model and the EuroSCORE. *Eur J Cardiothorac Surg.* 2000; 18: 703-710.
5. Roques F, Nashef SAM, Michel P, Pintor P, David M, Baudet E, The EuroSCORE study group. Does EuroSCORE work in individual European countries? *Eur J Cardiothorac Surg.* 2000; 18: 27-30.
6. Teniente-Valente R, González-Bravo FE, Chagolla-Santillán MA, Acevedo-Bañuelos I, Bernal-Ruiz EA, López-Luis LR et al. Validation of the EuroSCORE model in patients undergoing heart surgery in Regional Hospital of High Speciality of Bajío. *Rev Mex Cardiol.* 2018; 29 (3): 134-143.
7. Nashef SAM, Roques F, Hammill BG, Peterson ED, Michel P, Grover F et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. *Eur J Cardiothorac Surg.* 2002; 22: 101-105.
8. Gogbashian A, Sedrakyan A, Treasure T. EuroSCORE: a systematic review of international performance. *Eur J Cardiothorac Surg.* 2004; 25: 695-700.
9. Bhatti F, Grayson AD, Grotte G, Fabri BM, Au J, Jones M et al. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? *Heart.* 2006; 92: 1817-1820.
10. Basraon J, Chandrashekhar YS, John R, Agnihotri A, Kelly R, Ward H et al. Comparison of risk scores to estimate perioperative mortality in aortic valve

- replacement surgery. *Ann Thorac Surg.* 2011; 92: 535-540.
11. Parolari A, Pesce LL, Trezzi M, Loardi C, Kassem S, Brambillasca C et al. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: single institution experience and meta-analysis. *Eur Heart J.* 2009; 30: 297-304.
 12. Qadir I, Perveen S, Furnaz S, Shahabuddin S, Sharif H. Risk stratification analysis of operative mortality in isolated coronary artery bypass graft patients in Pakistan: comparison between additive and logistic EuroSCORE models. *Interact Cardiovasc Thorac Surg.* 2011; 13 (2): 137-141.
 13. Lebreton G, Merle S, Inamo J, Hennequin JL, Sanchez B, Rilos Z et al. Limitations in the inter-observer reliability of EuroSCORE: what should change in EuroSCORE II? *Eur J Cardiothorac Surg.* 2011; 40 (6): 1304-1308.
 14. Shih HH, Kang PL, Pan JY, Wu TH, Wu CT, Lin CY et al. Performance of European System for cardiac operative risk evaluation in Veterans General Hospital Kaohsiung cardiac surgery. *J Chin Med Assoc.* 2011; 74 (3): 115-120.
 15. Akar AR, Kurtcepe M, Sener E, Alhan C, Durdu S, Kunt AG et al.; The working Group for the Turkish Society of Cardiovascular Surgery and Turkish Ministry of Health. Validation of the EuroSCORE risk models in Turkish adult cardiac surgical population. *Eur J Cardiothorac Surg.* 2011; 40 (3): 730-735.
 16. Nashef SAM, Roques F, Sharples L, Nilsson J, Smith C, Goldstone A et al. EuroSCORE II. *Eur J Cardiothorac Surg.* 2012; 41 (4): 734-745.
 17. Sergeant P, Meuris B, Pettinari M. EuroSCORE II, illum qui est gravitates magni observe. *Eur J Cardiothorac Surg.* 2012; 41 (4): 729-731.
 18. Nezic D, Borzanovic M, Spasic T, Vukovic P. Calibration of the EuroSCORE II risk stratification model: is the Hosmer-Lemeshow test acceptable any more? *Eur J Cardiothorac Surg.* 2013; 43 (1): 206.
 19. Carnero AM, Silva JA, Reguillo FJ, Maroto LC, Cobiella J, Villagrán E et al. Validation of the EuroSCORE II on a single-centre 3800 patients cohort. *Interact Cardiovasc Thorac Surg.* 2013; 16 (4): 293-300.
 20. Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N et al. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. *Eur J Cardiothorac Surg.* 2013; 43 (4): 688-694.
 21. Borde D, Gandhe U, Hargave N, Pandey K, Khullar V. The application of European system for cardiac operative risk evaluation II (EuroSCORE II) and Society of Thoracic Surgeons (STS) risk-score for risk stratification in Indian patients undergoing cardiac surgery. *Ann Card Anaesth.* 2013; 16 (3): 163-166.
 22. Kunt A, Kurtcepe M, Hidiröglu M, Cetin L, Kucuker A, Bacuy V et al. Comparison of original EuroSCORE II and STS risk models in a Turkish cardiac surgical cohort. *Interact Cardiovasc Thorac Surg.* 2013; 16 (5): 625-629.
 23. Barili F, Pacini D, Capo A, Rasovic O, Grossi C, Alamanni F et al. Does EuroSCORE II perform better than its original versions? A multicentre validation study. *Eur Heart J.* 2013; 34 (1): 22-29.
 24. Kinkel CR, Rendón EH, Zuñiga LM, Rosas M et al. EuroSCORE II validation as a method for cardiac surgery risk stratification in Mexico. *Cardiovasc Thorac Surg.* 2019; 4: 1-8.
 25. Teniente-Valente R, Martínez-Bautista H, Chagolla-Santillán MA, Acevedo-Bañuelos I, García-Muñoz I, Romo-Escamilla R. ¿Cuál es el rendimiento del modelo EuroSCORE II en nuestro medio? *Cardiovasc Metab Sci.* 2020; 31 (supl. 4): s348.
 26. Di Dedda H, Pelissero G, Agnelli B, De Vincentiis C, Castelvechchio S. Accuracy, calibration and clinical performance of the new EuroSCORE II risk stratification system. *Eur J Cardiothorac Surg.* 2013; 43 (1): 27-32.
 27. Gao F, Shan L, Wang C, Meng X, Chen J, Han L, et al. Predictive ability of european heart surgery risk assessment system II (EuroSCORE II) and the Society of Thoracic Surgeons (STS) score for in-hospital and medium-term mortality of patients undergoing coronary artery bypass grafting. *Int J Gen Med.* 2021; 14: 8509-8519.
 28. Borracci R, Rubio M, Baldi-Jr J, Giorgini J, Higa C. Validación prospectiva multicéntrica del EuroSCORE II en Argentina. *Arch Cardiol Mex.* 2019; 89 (4): 315-323.
 29. Álvarez-Cabo R, Meana B, Díaz R, Hernández-Vaquero D, Pizcoya C, Mencia P, et al Utilidad de EuroSCORE-II en pacientes con cardiopatía isquémica. *Cir Cardiov.* 2017; 24 (2): 56-62.
 30. Rodríguez-Chávez L, Figueroa-Solano J, Muñoz-Consuegra CE, Avila-Vanzzinic N, Kuri-Alfaro J. EuroSCORE subestima el riesgo de mortalidad en cirugía cardíaca valvular de población mexicana. *Arch Cardiol Mex.* 2017; 87 (1): 18-25.
 31. Carosella VC, Grancelli H, Stutzbach P, Sigal AR, Lerech E, Morcos L et al. Estimación del riesgo en cirugía cardíaca en el "mundo real": ArgenSCORE ajustado al centro. *Rev Argent Cardiol.* 2021; 89 (1): 3-12.
- Declaration of confidentiality and patient consent:** given the nature of the study design, the patient's informed consent was not required. The confidentiality of every patient was kept.
- Clinical trial registration and approval number:** given the study design, it does not apply.
- Declaration of interests:** the authors have no economic or commercial interest in this research.
- Funding:** without support from any institution.
- Correspondence:**
Dr. Raúl Teniente Valente
E-mail: dr.teniente@yahoo.com