External Validation of European System for Cardiac Operative Risk Evaluation II (EuroSCORE II) for Risk Prioritization in an Iranian Population

Alireza Atashi1,2; Shahram Amini3, Mohammad Abbasi Tashnizi4, Ali Asghar Moeinipour4, Mathias Hossain Aazami5, Fariba Tohidnezhad1, Erfan Ghasemi6, Saeid Eslami1,7,8

Abstract

Introduction: The European System for Cardiac Operative Risk Evaluation II (EuroSCORE II) is a prediction model which maps 18 predictors to a 30-day post-operative risk of death concentrating on accurate stratification of candidate patients for cardiac surgery.

Objective: The objective of this study was to determine the performance of the EuroSCORE II risk-analysis predictions among patients who underwent heart surgeries in one area of Iran.

Methods: A retrospective cohort study was conducted to collect the required variables for all consecutive patients who underwent heart surgeries at Emam Reza hospital, Northeast Iran between 2014 and 2015. Univariate and multivariate analysis were performed to identify covariates which significantly contribute to higher EuroSCORE II in our population. External validation was performed by comparing the real and expected mortality using area under the receiver operating characteristic curve (AUC)

for discrimination assessment. Also, Brier Score and Hosmer-Lemeshow goodness-of-fit test were used to show the overall performance and calibration level, respectively.

Results: Two thousand five hundred eight one (59.6% males) were included. The observed mortality rate was 3.3%, but EuroSCORE II had a prediction of 4.7%. Although the overall performance was acceptable (Brier score=0.047), the model showed poor discriminatory power by AUC=0.667 (sensitivity=61.90, and specificity=66.24) and calibration (Hosmer-Lemeshow test, P<0.01).

Conclusion: Our study showed that the EuroSCORE II discrimination power is less than optimal for outcome prediction and less accurate for resource allocation programs. It highlights the need for recalibration of this risk stratification tool aiming to improve post cardiac surgery outcome predictions in Iran.

Keywords: Mortality. Decision Support Techniques. Risk Assessment. Cardiac Surgical Procedures.

INTRODUCTION

A growing literature shows the pervasiveness and importance of the need for reliable information on the cost-effectiveness of adult cardiac surgeries. Moreover, potential post-operative adverse events highlight the significance of perioperative clinical decision making. Various prediction models have been developed aiming to estimate risk-adjusted mortality, morbidity and length of intensive care unit stay following cardiac surgeries1,2. European System for Cardiac Operative Risk Evaluation (EuroSCORE) is a risk stratification tool which incorporates 18 variables describing patient, heart and proposed surgery to predict 30-day post-
operative risk of death\cite{6}. Predictive power of EuroSCORE II has been evaluated on different samples of target population in European countries. Vast majority of these studies have reported acceptable calibration (How many patients with a risk prediction of x% have experienced the outcome?) and discrimination (Who are the patients who have experienced the outcome associated with higher risk predictions and who are those that do not?) measures in comparison to Society of Thoracic Surgeons (STS) (esp. for patients undergoing coronary artery bypass grafting (CABG) procedure\cite{2}).

An international evaluation study was performed by Roques et al.\cite{7}, in 2000, to assess the predictive ability of EuroSCORE II summarized considering the demographic and clinical factors. The data were aggregated in a unique electronic dataset, (Available at: http://riskcalc.sts.org/stswebriskcalc/#). EuroSCORE II was calculated and inserted in dataset using online calculator. Similar studies in Iran reflect poor applicability of EuroSCORE II in patients undergoing different types of cardiac surgeries\cite{8,9}. Diverse surgical techniques and potential risk factors already have been stabilized in different communities may mislead prediction models and result in erroneous interpretations. Thus, mathematical localization studies are required in different geographical borders to assure its proper predictive function before routine clinical use\cite{10}. This study is conducted to investigate the accuracy of quantitative prioritization scores estimated by EuroSCORE II in an Iranian population.

**METHODS**

**Participants and Setting**

A retrospective single-center cohort study was conducted to include all consecutive patients undergoing cardiac surgeries at Emam Reza hospital, Northeast Iran from January 1, 2014 to December 31, 2015. Once the patient was hospitalized a cardiologist or a general physician evaluated pre- peri- and post-operative state to fill out the pre-designed structural paper form.

A total of 2907 patients were included and 30-day outcome was discovered using hospital information system or direct contact with patients’ family. About 11.2% (N=326) of records were excluded for regression analysis. The univariate and multivariate analysis are presented in Table 3. As these patients had no mortality, these factors were involved with a type of chronic kidney disease, 15.1% (N=24) underwent dialysis regularly, 10.6% (N=274) were current or past smokers and 2.2% (N=56) of patients were diagnosed with COPD. Table 1 summarizes some comparable information of our patients with the original EuroSCORE II population. As all procedures were elective operations, there were no urgent surgeries. Also, 23 patients undergoing valve surgery were suffering from active endocarditis, extra cardiac arteriopathy. Poor mobility was observed in 48 patients. No patient with COPD. Table 1 summarizes some comparable information of our patients with the original EuroSCORE II population.

**Statistical Analysis**

First, univariate and multivariate analysis of relevant EuroSCORE II prognostic factors were performed aiming to identify significant covariates which contributed to higher risk. EuroSCORE II was calculated and inserted in dataset using online calculator (Available at: http://riskcalc.sts.org/stswebriskcalc/#). The data were aggregated in a unique electronic dataset, summarized considering the demographic and clinical characteristics and were used for statistical analysis. The relation of each variable was addressed and the number of patients due to different values were compared to the original EuroSCORE II population. Then, the overall model performance was reported using Brier Score (A score function which measures the closeness of predictions to actual outcomes and result in a value from 0 for a perfect model to 0.25 for a non-informative model)\cite{11}. The area under the receiver operating characteristic curve (AUC) statistic was used to indicate the discriminative ability of model (while 1 refers to perfect discrimination, a value of 0.5 shows random classification). The Hosmer-Lemeshow goodness-of-fit test was employed to test the fitness of model to data by comparing observed to predicted mortality by decile of predicted probability\cite{12}. Analysis were performed using Medcalc-13.3.3.0 and R-3.3.1 (Resource Selection package).

**RESULTS**

**Patients’ Baseline Characteristics**

The mean age among the total of 2581 patients was 56.3±13.88 years (minimum=17 and maximum=93). The mortality rate was 3.3% (N=84). The mean height and weight of patients were 1.64±0.1 meters and 68.4±13.4 kilograms, respectively. About 7.8% (N=201) of patients aged 75 years and older and 22.2% (N=572) were diabetic. While 6.1% (N=158) were involved with a type of chronic kidney disease, 15.1% (N=24) underwent dialysis regularly, 10.6% (N=274) were current or past smokers and 2.2% (N=56) of patients were diagnosed with COPD. Table 1 summarizes some comparable information of our patients with the original EuroSCORE II population.

As all procedures were elective operations, there were no urgent surgeries. Also, 23 patients undergoing valve surgery were suffering from active endocarditis, extra cardiac arteriopathy. Poor mobility was observed in 48 patients. No patient with Canadian Cardiovascular Society (CCS) class 4 or with critical preoperative state was observed. Also, none of surgeries were on thoracic aorta. Some other details are presented in Tables 1 and 2. As these patients had no mortality, these factors were excluded for regression analysis. The univariate and multivariate analysis are presented in Table 3.

**Patients’ Heart Status**

Using New York Heart Association (NYHA), 37.1% (N=957) were classified as stage III cardiac failure patients, 6.2% (N=161) patients had a previous congestive heart failure during three months before surgery, 1% (N=26) of patients had atrial fibrillation. While 61.4% of surgeries were on-pump, the rest of procedures were performed off-pump. Table 2 shows more information about biological and clinical characteristics of patients.

**Performance Measures**

As mentioned before, the overall mortality was 3.3%. When applied to the current data set, the EuroSCORE II predicted a mortality of 4.7%. This means that the current risk-adjusted mortality ratio (RAMR=observed/predicted) for the previous additive model is about 0.67 and not adequately enough for outcome prediction or resource allocation programs.
Atashi A, et al. - External validation of European System for Cardiac Operative Risk Evaluation II

Brazilian Journal of Cardiovascular Surgery

**Table 1.** Comparison of demographic and comorbidity characteristics between the original EuroSCORE II population and an Iranian sample[2].

| Variable                          | Frequencies (% or mean (SD) [range] of original EuroSCORE II Population (N=22381) | Frequencies (% or mean (SD) [range] of our Population (N=2581) |
|-----------------------------------|-------------------------------------------------------------------------------------|----------------------------------------------------------------|
| Age                               | 64.6 (12.5) [18–95]                                                               | 56.3 (13.88) [17-94]                                              |
| Gender                            | Female 6919 (30.9%) Male 15462 (69.1%)                                             | 1044 (40.4%) 1537 (59.6%)                                       |
| Height (cm)                       | 168.5 (9.6) [100–213]                                                            | 164.1 (10.0) [104-195]                                           |
| Weight (kg)                       | 77.9 (15.9) [30–182]                                                             | 68.4 (13.4) [28-132]                                            |
| BMI (kg/m²)                       | 27.4 (4.8) [9.6–82.6]                                                            | 25.4 (4.8) [10.1-62.6]                                          |
| Diabetes on insulin               | 5643 (25.2%)                                                                      | 572 (22.2%)                                                      |
| NYHA                              | Class II NA                                                                        | 1008 (37.0%)                                                    |
|                                  | Class III NA                                                                       | 957 (37.1%)                                                     |
|                                  | Class IV NA                                                                        | 96 (3.6%)                                                       |
| Chronic pulmonary disease         | 2384 (10.7%)                                                                      | 56 (2.2%)                                                       |
| Serum creatinine (mg/dl)          | 1.13 (0.92)                                                                        | 1.09 (0.98)                                                     |
| Renal failure                     | Dialysis 108 (0.5%)                                                               | 158 (6.1%)                                                      |
|                                  | Dialysis NA                                                                       | 23 (0.9%)                                                       |
| LV function (ejection fraction)   | EF≤50 NA                                                                           | 1150 (44.6%)                                                    |
|                                  | S1-70 NA                                                                           | 782 (30.3%)                                                     |
|                                  | EF≤70 NA                                                                           | 20 (0.8%)                                                       |
| Recent MI                         | NA                                                                                 | 161 (6.2%)                                                      |
| Pulmonary hypertension            | NA                                                                                 | 190 (7.3%)                                                      |
| Previous cardiac surgery          | NA                                                                                 | 8 (0.3%)                                                        |
| Urgency                           | Urgent operation 4135 (18.5%)                                                     | None                                                            |
|                                  | Emergency 972 (4.3%)                                                              | None                                                            |
|                                  | Elective 17 165 (76.7%)                                                           | 2581 (100%)                                                    |
|                                  | Salvage 109 (0.5%)                                                                | None                                                            |

NA=not available; BMI=body mass index; NYHA=New York Heart Association functional classification; LV=left ventricle; MI=myocardial infarction

The Brier Score lower than 0.05 indicates acceptable overall performance. However, poor discrimination may be revealed by AUC=0.667 (cut off=3.0, sensitivity=61.90, and specificity=66.24). Also, the Hosmer-Lemeshow test showed unacceptable matching of predicted probabilities to observed events (P-value<0.01) (Table 4). Performance measures of EuroSCORE II are presented in Figure 1 and Table 4.

**DISCUSSION**

**Main Finding**

Our single-center study, based on consecutive patients who underwent cardiac surgery revealed that EuroSCORE II demonstrated a moderate statistical overall performance with poor discrimination and calibration measures remain as concerning issues regarding 30-day post-operative mortality prediction after adult cardiac surgery. The analysis of ROC curve showed that the EuroSCORE II discrimination power is less than optimal (AUC=0.667) for outcome prediction and less accurate for resource allocation programs, because, references consider an AUC value more than 0.7 as an acceptable value for least useful prediction models[5]. Although, the Brier score less than 0.05 indicates good overall performance for the model[12], the Hosmer-Lemeshow test showed unacceptable matching of predicted probabilities to observed events. In general, EuroSCORE II did not predict the outcome for our population as well as it did for the European populations. Thus, recalibration process seems to be essential for Iranian population prior to daily clinical use.

It is well known that risk assessment is central in the evaluation of the perioperative risk. The application of risk stratification tools gives an objective appraisal of risk for both physicians and patients and presents a good estimation for
Table 2. EuroSCORE II characteristics by patients’ demographic and clinical characteristics; number of patients and relation with mortality. The expected and observed mortality can also be compared by any variable[2].

| Variable                        | Number of Patients N (%) | Mortality N (%) | Mortality Predicted truly N (%) by EuroSCORE II |
|---------------------------------|--------------------------|-----------------|-----------------------------------------------|
| **Gender**                      |                          |                 |                                               |
| Male                            | 1537 (59.6%)             | 40 (47.6%)      | 27                                            |
| Female                          | 1044 (40.4%)             | 44 (52.4%)      | 29                                            |
| **Age**                         |                          |                 |                                               |
| 21-40                           | 301 (11.7%)              | 6 (7.1%)        | 3                                             |
| 41-60                           | 1215 (47.1%)             | 30 (35.7%)      | 21                                            |
| 61-80                           | 979 (37.9%)              | 45 (53.6%)      | 31                                            |
| >80                             | 46 (1.8%)                | 3 (3.6%)        | 1                                             |
| **BMI (kg/m²)**                 |                          |                 |                                               |
| ≤18.5                           | 120 (4.6%)               | 10 (11.9%)      | 8                                             |
| (18.5-23)                       | 608 (23.6%)              | 21 (25%)        | 14                                            |
| (23-25)                         | 437 (16.9%)              | 10 (11.9%)      | 8                                             |
| (25-30)                         | 832 (32.2%)              | 28 (33.4%)      | 21                                            |
| >30                             | 364 (14.1%)              | 7 (8.4%)        | 5                                             |
| **Valve Surgery (weight of the intervention)** | |                 |                                               |
| Isolated CABG                   | 2071 (80.2%)             | 54 (64.3%)      | 37                                            |
| AVR                             | 76 (2.9%)                | 2 (2.4%)        | 2                                             |
| MVR                             | 195 (7.6%)               | 14 (16.7%)      | 11                                            |
| TVR                             | 16 (0.6%)                | 2 (2.4%)        | 2                                             |
| MVR+TVR                         | 54 (2.1%)                | 4 (4.7%)        | 1                                             |
| AVR+MVR                         | 47 (1.8%)                | 3 (3.6%)        | 2                                             |
| ASD+TVR                         | 10 (0.4%)                |                |                                               |
| ASD                             | 7 (0.3%)                 |                |                                               |
| AVR+MVR+TVR                     | 18 (0.7%)                | 1 (1.2%)        |                                               |
| PVR                             | 10 (0.4%)                |                |                                               |
| 2 procedures                    | 122 (4.7%)               | 7 (8.3%)        | 6                                             |
| 3 procedures                    | 19 (0.7%)                | 1 (1.2%)        |                                               |
| **Other**                       | 77 (3%)                  | 4 (4.7%)        | 3                                             |
| **Ejection fraction**           |                          |                 |                                               |
| ≤50                             | 1150 (44.6%)             | 64 (76.1%)      | 31                                            |
| 51-70                           | 782 (30.3%)              | 19 (22.6%)      | 10                                            |
| >70                             | 20 (0.8%)                | 1 (1.2%)        | 1                                             |
| **Diabetes mellitus**           |                          |                 |                                               |
| Yes                             | 572 (22.2%)              | 32 (38.1%)      | 19                                            |
| No                              | 2009 (77.8%)             | 52 (61.9%)      | 37                                            |
| **COPD**                        |                          |                 |                                               |
| Yes                             | 56 (2.2%)                | 3 (3.6%)        | 3                                             |
| No                              | 2525 (97.8%)             | 81 (96.4%)      | 51                                            |
| **Mortality**                   |                          |                 |                                               |
| Alive                           | 2497 (96.7%)             | N/A             | N/A                                           |
| Dead                            | 84 (3.3%)                | N/A             | 56                                            |
| **Previous cardiac surgery**    |                          |                 |                                               |
| Yes                             | 8 (0.3%)                 | 3 (3.6%)        | 1                                             |
| No                              | 2573 (99.7%)             | 81 (96.4%)      | 55                                            |
| **Recent MI**                   |                          |                 |                                               |
| Yes                             | 161 (6.2%)               | 7 (8.3%)        | 4                                             |
| No                              | 2420 (93.8%)             | 77 (91.7%)      | 52                                            |
| **NYHA**                        |                          |                 |                                               |
| Class II                        | 1008 (37.0%)             | 7 (8.3%)        | 2                                             |
| Class III                       | 957 (37.1%)              | 24 (28.5%)      | 20                                            |
| Class IV                        | 963 (36.6%)              | 4 (4.7%)        | 4                                             |
| **Renal failure**               |                          |                 |                                               |
| Yes                             | 158 (6.1%)               | 14 (16.6%)      | 10                                            |
| No                              | 2423 (93.9%)             | 70 (83.3%)      | 46                                            |
| **Dialysis**                    |                          |                 |                                               |
| Yes                             | 23 (0.9%)                | 2 (2.4%)        | 2                                             |
| No                              | 2558 (99.1%)             | 82 (97.6%)      | 54                                            |
| **Pulmonary hypertension**      |                          |                 |                                               |
| Yes                             | 190 (7.3%)               | 15 (17.9%)      | 10                                            |
| No                              | 2391 (92.6%)             | 64 (82.1%)      | 46                                            |

BMI=body mass index; CABG=coronary artery bypass grafting; AVR=aortic valve replacement; MVR=mitral valve replacement; TVR=tricuspid valve replacement; ASD=atrial septal defect; PVR=pulmonary valve replacement; COPD=chronic obstructive pulmonary disease; N/A=not applicable; MI=myocardial infarction

*Analysis by independent-samples t test.
*Analysis by one-way ANOVA.

Sum of percentages may not result in 100% due to missing data.
Table 3. Univariate and multivariate analysis of EuroSCORE II prognostic factors[3].

| Characteristic                  | Mean ± SD | Univariate Analysis | Multivariate Analysis<sup>a</sup> |
|---------------------------------|-----------|---------------------|-----------------------------------|
|                                 | β (95% CI) | P value             | β (95% CI) | P value               |
| Age (year)<sup>a</sup>          | 56.3±13.88| 0.063 (0.058 to 0.068) | <0.001   | 0.058 (0.051 to 0.065) | <0.001 |
| Gender                          |           |                     |          |                      |
| Female                          | 3.5±1.95  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| Male                            | 2.6±2.16  | -0.88 (-1.05 to -0.72) | <0.001   | -0.82 (-0.98 to -0.66) | <0.001 |
| Creatinine clearance            |           |                     |          |                      |
| <50                             | 4.4±2.16  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| 50-85                           | 2.9±1.99  | -1.49 (-1.70 to -1.28) | <0.001   | -0.92 (-1.13 to -0.71) | <0.001 |
| >85                             | 1.8±1.53  | -2.57 (-2.8 to -2.34) | <0.001   | -1.06 (-1.32 to -0.79) | <0.001 |
| Dialysis                        | 3.0±2.31  | -1.33 (-2.12 to -0.55) | <0.001   | -0.78 (-1.73 to -0.02) | 0.045 |
| Chronic lung disease            |           |                     |          |                      |
| No                              | 2.9±2.11  | 1 [References]      | 0.061    | 1 [Reference]        |      |
| Yes                             | 3.5±2.43  | 0.54 (-0.03 to 1.10) | <0.001   | 0.15 (-0.31 to 0.61) | 0.529 |
| Diabetes on insulin             |           |                     |          |                      |
| No                              | 3.0±2.12  | 1 [References]      | 0.859    | 1 [Reference]        |      |
| Yes                             | 2.9±2.14  | -0.02 (-0.22 to 0.18) | <0.001   | -0.19 (-0.37 to -0.01) | 0.043 |
| NYHA                            |           |                     |          |                      |
| Class I                         | 2.4±1.97  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| Class II                        | 2.9±2.18  | 0.47 (0.18 to 0.77)  | <0.001   | 0.2 (-0.002 to 0.40) | 0.052 |
| Class III                       | 3.3±2.09  | 0.86 (0.56 to 1.17)  | <0.001   | 0.14 (0.07 to 0.35)  | 0.199 |
| Class IV                        | 3.8±1.98  | 1.44 (0.81 to 2.07)  | <0.001   | 0.52 (0.08 to 0.95)  | 0.019 |
| Left ventricular function       |           |                     |          |                      |
| ≤50                             | 3.5±2.19  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| 51-70                           | 2.3±1.84  | -1.22 (-1.40 to -1.03) | <0.001   | -0.87 (-1.03 to -0.70) | <0.001 |
| >70                             | 2.0±1.76  | -1.52 (-2.43 to -0.61) | <0.001   | -1.37 (-2.06 to -0.69) | <0.001 |
| Recent myocardial infarction    |           |                     |          |                      |
| No                              | 2.8±2.07  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| Yes                             | 4.7±2.19  | 1.81 (1.48 to 2.14)  | <0.001   | 1.86 (1.56 to 2.15)  | <0.001 |
| Pulmonary hypertension          |           |                     |          |                      |
| No                              | 2.8±2.03  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| Yes                             | 5.2±1.99  | 2.41 (2.11 to 2.71)  | <0.001   | 2.04 (1.77 to 2.31)  | <0.001 |
| Weight of the intervention      |           |                     |          |                      |
| Isolated CABG                   | 2.7±2.06  | 1 [References]      | <0.001   | 1 [Reference]        |      |
| Single non-CABG                 | 4.1±2.02  | 1.38 (1.14 to 1.62)  | <0.001   | 1.48 (1.25 to 1.72)  | <0.001 |
| 2 Procedures                    | 4.2±1.93  | 1.55 (1.21 to 1.90)  | <0.001   | 1.89 (1.56 to 2.22)  | <0.001 |
| 3 Procedures                    | 4.1±1.6   | 1.45 (0.74 to 2.15)  | <0.001   | 1.58 (0.91 to 2.25)  | <0.001 |

<sup>a</sup>Using linear regression, all variables were found to be associated with EuroSCORE II, except for chronic lung disease, diabetes on insulin, and NYHA.

NYHA=New York Heart Association; CABG=coronary artery bypass graft

Table 4. Performance measures calculated for EuroSCORE II scoring system.

| Scoring System          | Overall Performance | Discrimination | Calibration |
|-------------------------|---------------------|----------------|-------------|
|                         | Brier Score (min-max) [STD] | AUC | SE | 95% CI | Sensitivity | Specificity | H-L Test |
| EuroSCORE II            | 0.047 (0.0-1.0) [0.12] | 0.667 | 0.0307 | 0.648-0.685 | 0.619 | 0.662 | Chi²(8)=936.66, P<0.01 |

AUC=area under the ROC curve; SE=standard error; CI=confidence interval, H-L=Hosmer-Lemeshow
The observed 30-day mortality rate in our sample (3.3%) was similar to those published by Roques et al. [7] (3.4%), Nashef et al. [8] (3.9%), Geissler et al. [5] (4%), and Pitkänen et al. [1] (2%). While Mir Mohammad Sadeghi et al. [9] reported similar mortality rate in Isfahan (central Iran), four years later Jamaati et al. [10] evaluated EuroSCORE II on a sample containing 12.2 mortality rate. An AUC of 66.7% in our study is lower than all similar studies including 78% by Geissler et al. [5], 77% by Pitkänen et al. [1], and 75.4% by Antunes et al. [24]. This is while similar studies in Iran confirmed the poor discriminative ability of EuroSCORE II [9,10].

Currently, a great interest for prediction models as powerful tools for outcome prediction, cost-effectiveness strategies, reasonable resource allocation, and consequently quality control process have been growing [1,9,25].

Due to the results of our study, despite the little differences between two populations (Tables 1 to 3) the EuroSCORE II may not be completely reliable for risk periodization or resource allocation programs in Iran. Poor performance measures for EuroSCORE II highlights the need for reformulating this risk stratification tool aiming to improve post cardiac surgery outcome predictions in Iran. It may be done by calibrating mortality risk scoring model (e.g. EuroSCORE model) for the region or creating new models with accurate localized parameter sets [11,20].

**Limitation**

Although sampling was done in one of the largest hospitals performing various cardiac procedures and the study has adequate sample size, including just one center may affect the generalizability of results to the entire country.

**Future Studies**

Regarding the key prognostic role of prediction models, further investigation of clinical risk factors and recalibration process seems to be essential on large samples of target population from different centers around country aiming to improve outcome predictions.

**CONCLUSION**

Our study showed that the EuroSCORE II discrimination power is less than optimal for outcome prediction and less accurate for resource allocation programs. It highlights the need for recalibration this risk stratification tool aiming to improve post cardiac surgery outcome predictions in Iran.

**ACKNOWLEDGEMENT**

The authors would like to express their appreciation to the staff of cardiac surgery ICU of Emam Reza hospital, Mashhad, for their tremendous support in data collection. We thank Proff. Ameen Abu-Hanna for remarks on an earlier version of this paper. This study was a part of the first author’s PhD thesis which was supported by a grant from Mashhad University of Medical Sciences Research Councils.
REFERENCES

1. Pitkänen O, Niskanen M, Rehnberg S, Hippeläinen M, Hynynen M. Intra-institutional prediction of outcome after cardiac surgery: comparison between a locally derived model and the EuroSCORE. Eur J Cardiothorac Surg. 2000;18(6):703-10.

2. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. Eur J Cardiothorac Surg. 2012;41(4):734-44.

3. Braunwald E, Fauci AS, Kasper DL, Hauser SL, Longo DL, Jameson JL, eds. Harrison's principles of internal medicine. 15th ed. New York: McGraw-Hill Professional Publishing; 2001.

4. Roques F, Nashef SA, Michel P, Pinna Pintor P, David M, Baudet E; EuroSCORE Study Group. Does EuroSCORE work in individual European institutions? Eur J Cardiothorac Surg. 2000;18(6):27-30.

5. Geissler HJ, Holz P, Maroš S, Kuhn-Régnier F, Mehlihorn U, Südkamp M, et al. Risk stratification in heart surgery: comparison of six score systems. Eur J Cardiothorac Surg. 2000;17(4):400-6.

6. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon AA. Authors' roles & responsibilities

AA Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; drafting the work; final approval of the version to be published

SA Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; data gathering management; final approval of the version to be published

MHA Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; data gathering management; final approval of the version to be published

FT Statistical analysis and interpretation of data for the work; final approval of the version to be published

EG Statistical analysis and interpretation of data for the work; final approval of the version to be published

SE Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; data gathering management; final approval of the version to be published

This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Braz J Cardiovasc Surg 2018;33(1):40-6

Atashi A, et al. - External validation of European System for Cardiac Operative Risk Evaluation II

4. Roques F, Nashef SA, Michel P, Pinna Pintor P, David M, Baudet E; EuroSCORE Study Group. Does EuroSCORE work in individual European institutions? Eur J Cardiothorac Surg. 2000;18(6):703-10.

5. Geissler HJ, Holz P, Maroš S, Kuhn-Régnier F, Mehlihorn U, Südkamp M, et al. Risk stratification in heart surgery: comparison of six score systems. Eur J Cardiothorac Surg. 2000;17(4):400-6.

11. Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. Circulation. 1999;99(6):2098-104.

12. Lichtenstein S, Fischhoff B, Phillips LD. Calibration of probabilities: the state of the art to 1980. In: Kahneman D, Slovic P, Tversky A, eds. Judgement under uncertainty: heuristics and biases. New York: Cambridge University Press, 1982.

13. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol. 1982;115(1):92-106.

14. Madeira S, Rodrigues R, Trallhão A, Santos M, Almeida C, Marques M, et al. Assessment of perioperative mortality risk in patients with infective endocarditis undergoing cardiac surgery: performance of the EuroSCORE I and II logistic models. Interact Cardiovasc Thorac Surg. 2016;22(2):141-8.

15. Clark RE. The STS Cardiac Surgery National Database: an update. Ann Thorac Surg. 1995;59(6):1376-80.

16. Ferguson TB Jr, Dzubiyan SW Jr, Edwards FH, Eiken MC, Shoeyr AL, Patolero PC, et al. The STS National Database: current changes and challenges for the new millennium. Committee to Establish a National Database in Cardiothoracic Surgery, The Society of Thoracic Surgeons. Ann Thorac Surg. 2000;69(3):680-91.

17. Cochrane AL. Archie Cochrane in his own words. Selections arranged from his 1972 introduction to “Effectiveness and Efficiency: Random Reflections on the Health Services” 1972. Control Clin Trials. 1989;10(4):428-33.

18. Parolini A, Pesce LL, Trezzi M, Cavallotti L, Kassem S, Loardi C, et al. EuroSCORE performance in valve surgery: a meta-analysis. Ann Thorac Surg. 2010;89(3):787-93.

19. 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):22-9.

20. Lebret 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):1304-8.

21. Borracci RA, Rubino M, Celano L, Ingino CA, Allende NG, Ahuad Guerrero RA. Prospective validation of EuroSCORE II in patients undergoing cardiac surgery in Argentinean centres. Interact Cardiovasc Thorac Surg. 2014;18(5):539-43.

22. Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Medratta N, et al. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. Eur J Cardiothorac Surg. 2013;43(4):688-94.

23. Wang L, Han QQ, Qiao F, Wang C, Zhang XW, Han L, et al. Performance of EuroSCORE II in patients who have undergone heart valve surgery: a multicentre study in a Chinese population. Eur J Cardiothorac Surg. 2014;45(2):359-64.

24. Antunes PE, Eugénio L, Ferrão de Oliveira J, Antunes MJ. Mortality risk prediction in coronary surgery undergoing infective endocarditis undergoing cardiac surgery: performance of the EuroSCORE I and II logistic models. Interact Cardiovasc Thorac Surg. 2011;40(6):1304-8.

25. Hekmat K, Raabe A, Kroener A, Fischer U, Suedkamp M, Geissler HJ, et al. Risk stratification models fail to predict hospital costs of cardiac surgery patients. Z Kardiol. 2005;94(11):748-53.