Scholars Journal of Applied Medical Sciences
Abbreviated Key Title: Sch J App Med Sci
ISSN 2347-954X (Print) | ISSN 2320-6691 (Online)
Journal homepage: https://saspublishers.com/sjams/

OPEN ACCESS

Medicine

Performance of EuroSCORE and EuroSCORE II in Institut Jantung Negara (IJN), Kuala Lumpur, Malaysia

M. Azizan Petra*, M. Ezani Taib, M. F. Ramli, N. Jaafar, I. F. Gaafar

Institut Jantung Negara, Kuala Lumpur, Malaysia

DOI: 10.36347/sjams.2020.v08i06.001 | **Received:** 05.05.2020 | **Accepted:** 23.05.2020 | **Published:** 08.06.2020

***Corresponding author:** M. Azizan Petra

Abstract **Original Research Article**

Objective: Assessing the performance of European System for Cardiac Operative Evaluation (EuroSCORE) and EuroSCORE II. ***Method:*** 4145 patients who underwent cardiac surgery between 1st January 2015 to 31st December 2016 in Institut Jantung Negara (IJN) were included. The entire cohort and isolated coronary bypass graft (CABG) patients were analyzed by measuring the area under the receiver operating characteristic (ROC) curve for model discrimination and Hosmer-Lemeshow Chi-squared test for model calibration. Performance of both models was compared. ***Result:*** For the entire cohort, ROC curve for EuroSCORE was 0.679; EuroSCORE II was 0.615. For isolated CABG patients, ROC curve for EuroSCORE was 0.670; EuroSCORE II was 0.609. For the entire cohort, Hosmer-Lemeshow test showed no significant difference between expected and observed mortality according to EuroSCORE model (Chi-square = 5.284, P = 0.508) and EuroSCORE II model (Chi-square = 15.828, P = 0.050). For the isolated CABG patients, Hosmer-Lemeshow test showed no significant difference between expected and observed mortality according to EuroSCORE model (Chi-square = 5.365, P = 0.498) and EuroSCORE II model (Chi-square = 9.839, P = 0.276). For the entire cohort (Table 7), the observed and predicted mortality were 4.56% and 3.7% respectively for EuroSCORE; observed and predicted mortality were similar at 4.56% for EuroSCORE II. For isolated CABG patients (Table 8), the observed and predicted mortality were 3.62% and 3.36% respectively for EuroSCORE; the observed and predicted mortality were 3.62% and 3.97% respectively for EuroSCORE II. ***Conclusion:*** Despite poor discrimination under the ROC, the calibration of both models was good and acceptable to be used for risk prediction tools in our centre.

Keywords: EuroSCORE, receiver operating characteristic (ROC), predicted mortality.

INTRODUCTION

Risk stratification and prediction models has become paramount tools in current cardiac surgical practice. These tools have been consistently used to help in clinical decision making, informed consent, quality assurance and healthcare management [1, 6]. These models are also useful to allow surgeons to compare mortality and outcomes with international data.

There are several risk predictions models that are used internationally; EuroSCORE, EuroSCORE II, Parsonnet, et cetera. These models have been tested and widely used internationally. It has been validated with good results in European and North American population [1, 5]. These models are useful to predict patient outcome endpoints including long term mortality risk, intensive care unit stay, surgical complications and costs of cardiac surgery [1]. These models are based on Caucasian population and hence are most applicable when the demographics of pre-operative patients are similar to the population which were used in the risk scoring models. Therefore, the risk scoring models may not be useful in some parts of the world. Because of this, risk scoring system can only be reliably used when its validity and performance has been tested in the centre of application.

This study is aimed at assessing the performance of European System for Cardiac Operative Evaluation (EuroSCORE) and EuroSCORE II in IJN.

Study Design

A retrospective single centre study was performed. Patients included are those who underwent cardiac surgery in Insititut Jantung Negara, Kuala Lumpur, Malaysia between 1st January 2015 to 31 December 2016. Patients under 18 years old, closed heart surgery and transplant procedures are excluded from the study. Baseline characteristics of all included

patients are measured. The baseline characteristics used as risk factors for Euroscore and Euroscore II. This included the entire cohort and a subgroup of patients undergoing coronary artery bypass surgery (CABG) only were analysed.

Statistical analysis was performed by chi-square test. Data analysis was performed using SPSS 11.5 statistical software package. Observed and predicted mortality was compared in order to assess the performance of both models.

Data Analysis

The performances of the Euroscore models were analysed focusing on discrimination and calibration. The discrimination performance indicates the extent to which each model differentiates which patient will die or survive in the perioperative period. It was measured using receiver operating characteristic curve. Area under the curve were calculated with 95% confidence interval. An area of 0.5 reflects no discrimination and an area of 1.0 reflects perfect predictor. Areas of greater than 0.7 are generally thought to be useful [10].

Calibration refers to the agreement between observed outcomes and prediction. For example, 20 in - hospital deaths should be observed in a 100 patients group with 20% predicted mortality. We used the Hosmer – Lemeshow Chi Square statistic to measure these values over deciles of risks. P value greater than 0.05 shows a well calibrated model [1].

RESULTS

Total of 6118 patients underwent open heart surgery during the study year period. 1973 patients were excluded leaving 4145 patients for analysis and out of this, 2820 patients underwent isolated CABG.

Table-1: Total patient breakdown before and after exclusion

		2015	2016	Total
Before exclusion	All	3032	3086	6118
	Isolated CABG	1383	1549	2932
After exclusion	All	1984	2161	4145
	Isolated CABG	1314	1506	2820

The baseline variables and characteristics are measured for both groups for EuroScore and Euroscore II which are shown in Table-2. Most patient have normal EF in the overall and isolated CABG group; 50.1% and 45.5% respectively. Also, most patients underwent elective procedures in both groups; 70.2% and 69% respectively. It is important to note that there is a group of patients in the isolated CABG, who were planned to have multiple procedures but ended up to have isolated CABG.

Table-2: Baseline variables and characteristics are measured for both groups for EuroScore and Euroscore II.

	Overall		Isolated cabg	
Risk factors	**EuroSCORE**	**EuroSCORE II**	**EuroSCORE**	**EuroSCORE II**
age – mean ± sd (years)	59.3 ± 11.8		61.7 ± 8.6	
median (iqr)	61.2 (53.7, 67.4)		62.2 (56.1, 67.8)	
gender (female)	1016 (24.5)		455 (16.1)	
creatinine				
CC < 50	835 (20.1)		575 (20.4)	
CC > 50-85	1535 (37.0)		1126 (39.9)	
cc > 85	1728 (41.7)		1086 (38.5)	
extracardiac arteriopathy	80 (1.9)		71 (2.5)	
poor mobility	NIM	94 (2.3)	NIM	58 (2.1)
previous cardiac surgery	NIM	263 (6.3)	NIM	75 (2.7)
pulmonary disease	312 (7.5)		218 (7.7)	
active endocarditis	NIM	49 (1.2)	NIM	11 (0.4)
critical pre-operative state	NIM	107 (2.6)	NIM	89 (3.2)
diabetes on insulin	NIM	655 (15.8)		592 (21.0)
neurological disease	87 (2.1)	NIM	56 (2.0)	NIM
reoperation	1229 (29.7)	NIM	811 (28.8)	NIM
vt/vf	15 (0.4)	NIM	12 (0.4)	NIM
cpr	8 (0.2)	NIM	6 (0.2)	NIM
iabp	NIM	NIM		
ventilation	23 (0.6)	NIM	16 (0.6)	NIM
nyha i	NIM	1821 (43.9)	NIM	1333 (47.3)
nyha iI	NIM	1871 (45.1)	NIM	1235 (43.8)

nyha iII	NIM	307 (7.4)	NIM	156 (5.5)
nyha iV	NIM	17 (0.4)	NIM	8 (0.3)
ccs 4	NIM	42 (1.0)	NIM	37 (1.3)
unstable angina	408 (9.8)	NIM	369 (13.1)	NIM
lvef > 50%	2077 (50.1)		1284 (45.5)	
lvef 31-50%	1531 (36.9)		1172 (41.6)	
lvef 21-30%	274 (6.6)		202 (7.2)	
lvef <20%	29 (0.7)		22 (0.8)	
recent mi (<90 days)	606 (14.6)		531 (18.8)	
pulmonary hpt (31-55mmhg)	NIM	135 (3.3)	NIM	71 (2.5)
pulmonary hpt (>55mmhg)	NIM	56 (1.4)	NIM	10 (0.4)
systolic pressure >60mmhg	23 (0.6)	NIM	4 (0.1)	NIM
urgency				
elective	2911 (70.2)		1947 (69.0)	
emergency	112 (2.7)		83 (2.9)	
urgent	187 (4.5)		170 (6.0)	
salvage	-		-	
planned operation				
isolated cabg	2771 (66.9)		2719 (96.4)	
single non cabg	580 (14.0)		6 (0.2)	
2 procedures	585 (14.1)		39 (1.4)	
3 procedures	154 (3.7)		9 (0.3)	
planned surgery on thoracic aorta	103 (2.5)		15 (0.5)	
POSTINFARCT SEPTAL RUPTURE	4 (0.1)	NIM	3 (0.1)	NIM

For the entire cohort (Figure-1), the ROC curve for EuroSCORE was 0.679 (95% confidence interval (CI): 0.639 – 0.720] p<0.001); EuroSCORE II was 0.615 (95% CI: 0.576 – 0.655) p<0.001). For isolated CABG patients (Figure 2), the ROC curve for EuroSCORE was 0.670 [95% CI: 0.614 – 0.726] p<0.001); EuroSCORE II was 0.609 (95% CI: 0.556 – 0.661), p<0.001).

For the entire cohort, the Hosmer-Lemeshow (HL) goodness-of-fit test showed no significant difference between observed and predicted mortality according to EuroSCORE model (Table-3), (Chi-square = 5.284, P = 0.508) and EuroSCORE II model (Table-4), (Chi-square = 15.828, P = 0.050).

For the subgroup of isolated CABG patients (Table-5), the Hosmer-Lemeshow (HL) goodness-of-fit test showed no significant difference between observed and predicted mortality according to EuroSCORE model (Table-5), (Chi-square = 5.365, P = 0.498) and EuroSCORE II model (Table-6), (Chi-square = 9.839, P = 0.276).

For the entire cohort (Table-7), the observed and predicted mortality were 4.56% and 3.7% respectively for EuroSCORE; observed and predicted mortality were similar at 4.56% for EuroSCORE II. For isolated CABG patients (Table-8), the observed and predicted mortality were 3.62% and 3.36% respectively for EuroSCORE; the observed and predicted mortality were 3.62% and 3.97% respectively for EuroSCORE II.

Table-3: Goodness-of-fit test for EuroSCORE (Overall)

Risk group	ALIVE patients		DEAD patients		total
	Observed	Expected	Observed	Expected	
1	319	321.1	7	4.9	326
2	487	483.5	6	9.5	493
3	588	590.1	17	14.9	605
4	773	776.7	29	25.3	802
5	547	544.2	20	22.8	567
6	408	412.7	27	22.3	435
7	347	343.2	20	23.8	367
8	487	484.4	63	65.6	550

p value = 0.508, Chi-square = 5.284

**The logistic regression fits the data since H-L test > 0.05

Table-4: Goodness-of-fit test for EuroSCORE II (Overall)

Risk group	ALIVE patients		DEAD patients		total
	Observed	Expected	Observed	Expected	
1	407	400.0	8	15.0	415
2	406	399.6	9	15.4	415
3	405	399.3	10	15.7	415
4	395	399.0	20	16.0	415
5	396	398.5	19	16.5	415
6	394	397.8	21	17.2	415
7	395	400.9	24	18.1	419
8	400	398.7	18	19.3	418
9	386	393.5	29	21.5	415
10	372	368.7	31	34.3	403

p value = 0.050, Chi-square = 15.828

**The logistic regression fits the data since H-L test > 0.05

Table-5: Goodness-of-fit test for EuroSCORE (Isolated CABG)

Risk group	ALIVE patients		DEAD patients		total
	Observed	Expected	Observed	Expected	
1	310	311.8	6	4.2	316
2	480	476.8	5	8.2	485
3	385	387.5	11	8.5	396
4	437	441.7	17	12.3	454
5	341	338.9	10	12.1	351
6	254	253.5	11	11.5	265
7	210	209.8	12	12.2	222
8	301	297.9	30	33.1	331

p value = 0.498 Chi-square = 5.365

**The logistic regression fits the data since H-L test > 0.05

Table-6: Goodness-of-fit test for EuroSCORE II (Isolated CABG)

Risk group	ALIVE patients		DEAD patients		total
	Observed	Expected	Observed	Expected	
1	293	290.4	6	8.6	299
2	278	273.6	4	8.4	282
3	277	273.5	5	8.5	282
4	273	273.3	9	8.7	282
5	271	273.1	11	8.9	282
6	270	273.7	13	9.3	283
7	268	272.3	14	9.8	282
8	269	271.6	13	10.4	282
9	273	270.5	9	11.5	282
10	246	246.0	18	18.0	264

p value = 0.276 Chi-square = 9.839

**The logistic regression fits the data since H-L test > 0.05

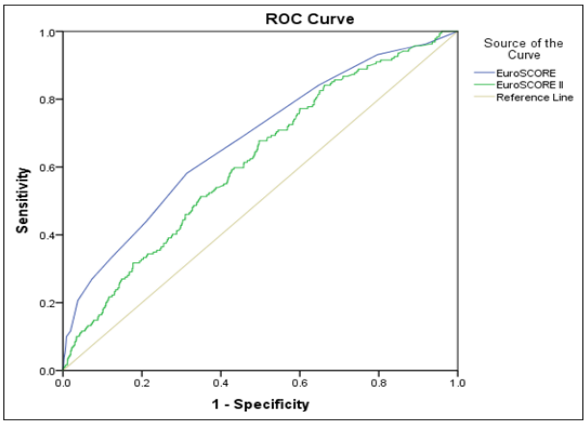

Fig-1: ROC curves for EuroSCORE and EuroSCORE II (Overall)

Area Under the Curve					
Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
EuroSCORE	.679	.021	.000	.639	.720
EuroSCORE II	.615	.020	.000	.576	.655
a. Under the nonparametric assumption					
b. Null hypothesis: true area = 0.5					

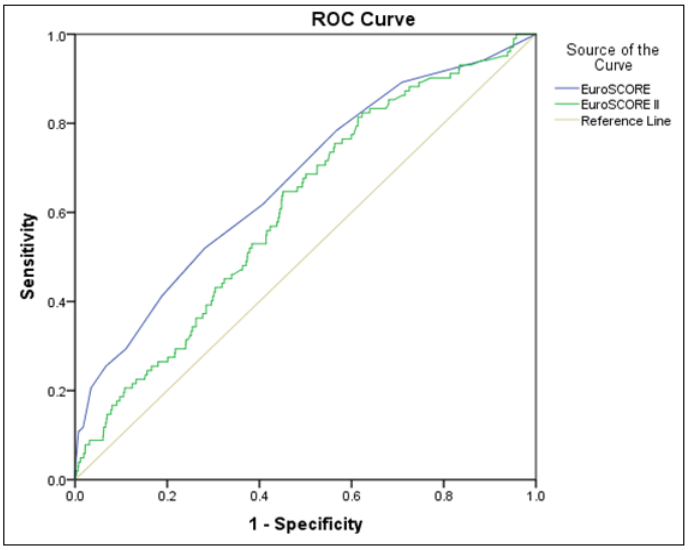

Fig-2: ROC curves for EuroSCORE and EuroSCORE II (Isolated CABG)

Area Under the Curve					
Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
EuroSCORE	.670	.028	.000	.614	.726
EuroSCORE II	.609	.027	.000	.556	.661
a. Under the nonparametric assumption					
b. Null hypothesis: true area = 0.5					

Table-7: Observed and predicted mortality by EuroSCORE and EuroSCORE II risk level for all cohort

Risk group	Patients (death)	Observe death (%)	Predicted death (%)
Low risk (0-2)	1424 (30)	2.11 (1.36-2.86)	1.20 (1.16-1.24)
Medium risk (3-5)	1804 (76)	4.21 (3.28-5.14)	3.80 (3.76-3.83)
High risk (>=6)	917 (83)	9.05 (7.19-10.91)	7.42 (7.31-7.53)
Total	4145 (189)	4.56 (3.92-5.20)	3.70 (3.63-3.78)

EuroSCORE

Risk group	Patients (death)	Observe death (%)	Predicted death (%)
Low risk (<=1.22)	412 (8)	1.94 (0.61-3.27)	0.89 (0.87-0.92)
Medium risk (1.23-2.02)	988 (25)	2.53 (1.55-3.51)	1.62 (1.60-1.63)
High risk (2.03-4.11)	1229 (59)	4.80 (3.60-6.00)	2.93 (2.90-3.00)
Very high risk (4.12-47.6)	1516 (97)	6.40 (5.17-7.63)	8.78 (8.48-9.07)
Total	4145 (189)	4.56 (3.92-5.20)	4.56 (4.41-4.70)

EuroSCOREII

Table-8: Observed and predicted mortality by EuroSCORE and EuroSCORE II risk level for isolated CABG

Risk group	Patients (death)	Observe death (%)	Predicted death (%)
Low risk (0-2)	1197 (22)	1.84 (1.08-2.60)	1.07 (1.02-1.11)
Medium risk (3-5)	1070 (38)	3.55 (2.44-4.66)	3.82 (3.78-3.87)
High risk (>=6)	553 (42)	7.59(5.38-9.80)	7.42 (7.27-7.56)
Total	2820 (102)	3.62 (2.93-4.31)	3.36 (3.26-3.45)

EuroSCORE

Risk group	Patients	Observe death (%)	Predicted death (%)
Low risk (<=1.22)	305 (6)	1.97 (0.41-3.53)	0.90 (0.87-0.92)
Medium risk (1.23-2.02)	757 (13)	1.72 (0.79-2.65)	1.62 (1.61-1.64)
High risk (2.03-4.11)	863 (39)	4.52 (3.13-5.91)	2.93 (2.89-2.97)
Very high risk (4.12-47.6)	895 (44)	4.92 (3.50-6.34)	8.00 (7.66-8.34)
Total	2820 (102)	3.62 (2.93-4.31)	3.97 (3.82-4.12)

EuroSCOREII

DISCUSSION

Risk stratification and risk scoring systems in adult cardiac surgery are becoming important as they provide reliable estimations of the risks associated with surgical procedure and they permit, in some cases to a certain extent, comparison of outcomes among institutions and surgeons by adjusting for a variable mixture of cases. In addition, these models may also provide a more accurate assessment of the indication for surgery in each individual patient which in turn can facilitate a more precise balance between the potential risks and benefits. This is an important role of risk prediction models in cardiac surgery, as not only it will help us to decide surgery suitability, it will also help family and patients to have improved informed consent [2].

However, risk prediction models may not be suitable to certain groups of our patient population. Ready-made models may not accurately predict local outcomes and will require validation prior to use [1].

We have shown that in this study that both models do not predict outcomes accurately in the entire cohort. Even though both risk prediction models represented under-prediction rates for the entire cohort, Euroscore II appears to perform better at risk prediction with an under-prediction rate of only 2% when compared to Euroscore which produced an under-prediction rate of 20%.

However, it is important to note that the calibration which is calculated by the Hosmer-Lemeshow (HL) goodness-of-fit test show good results confirming statistical precision with P values of >0.05. However, despite still showing discriminatory ability, this study shows poor discriminatory abilities for both risk models where the ROC curve values were 0.609 - 0.679.

For both models, the under-prediction rates may be due to several factors. The first factor could be due to the difference in the underlying co-morbidities between Asian and European populations. For example, higher prevalence in diabetes and hypertension [10]. The other factor could be information bias; some information were not disclosed by patients for example family history and smoking.

The major limitation of this study is that it is done only as a single centre study. Although this study was done from a high-volume institution, it is still important to consider the practice of other institutions. This data has not been validated externally. Other than that, this study could have been done over a longer time period. The other limitation of the study is its retrospective nature which may introduce potential bias; particularly in regards to the accuracy of data which relied on individuals recalling the information.

CONCLUSION

Euroscore and Euroscore II is relatively a simple tool to be used for risk stratification. It is an easy scoring system including most of the usual risk factors [6]. However, despite poor discrimination under the ROC curve, the calibration of both models was good and acceptable to be used for risk prediction tools in our centre. However, we still should be cautious in the utilization of these risk stratification models. The differences in demographic characteristics should be respected among different population groups.

REFERENCES

1. Yap CH, Reid C, Yii M, Rowland MA, Mohajeri M, Skillington PD, Seevanayagam S, Smith JA. Validation of the EuroSCORE model in Australia. European journal of cardio-thoracic surgery. 2006 Apr 1;29(4):441-6.
2. Barili F, Pacini D, Capo A, Rasovic O, Grossi C, Alamanni F, Di Bartolomeo R, Parolari A. Does EuroSCORE II perform better than its original versions? A multicentre validation study. European heart journal. 2013 Jan 1;34(1):22-9.
3. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, EuroSCORE Study Group. European system for cardiac operative risk evaluation (Euro SCORE). European journal of cardio-thoracic surgery. 1999 Jul 1;16(1):9-13.
4. Michel P, Roques F, Nashef SA, EuroSCORE Project Group. Logistic or additive EuroSCORE for high-risk patients?. European Journal of Cardio-thoracic surgery. 2003 May 1;23(5):684-7.
5. Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, Wyse RK, Ferguson TB. Validation of European system for cardiac operative risk evaluation (EuroSCORE) in North American cardiac surgery. European journal of cardio-thoracic surgery. 2002 Jul 1;22(1):101-5.
6. Sergeant, P., de Worm, E., & Meyns, B. (2001). Single centre, single domain validation of the EuroSCORE on a consecutive sample of primary

and repeat CABG. *European journal of cardio-thoracic surgery*, *20*(6), 1176-1182.

7. Garcia-Valentin A, Mestres CA, Bernabeu E, Bahamonde JA, Martín I, Rueda C, Domenech A, Valencia J, Fletcher D, Machado F, Amores J. Validation and quality measurements for EuroSCORE and EuroSCORE II in the Spanish cardiac surgical population: a prospective, multicentre study. European Journal of Cardio-Thoracic Surgery. 2016 Feb 1;49(2):399-405.
8. Amr MA, El-shorbagy AA. Evaluation of accuracy of EuroSCORE II in prediction of in-hospital mortality in patients underwent mitral valve replacement in Egypt. Journal of the Egyptian Society of Cardio-Thoracic Surgery. 2016 Aug 1;24(2):135-42.
9. Akar AR, Kurtcephe M, Sener E, Alhan C, Durdu S, Kunt AG, Güvenir HA. Validation of the EuroSCORE risk models in Turkish adult cardiac surgical population. European journal of cardio-thoracic surgery. 2011 Sep 1;40(3):730-5.
10. Zheng Z, Li Y, Zhang S, Hu S. The Chinese coronary artery bypass grafting registry study: how well does the EuroSCORE predict operative risk for Chinese population?. European journal of cardio-thoracic surgery. 2009 Jan 1;35(1):54-8.