Heart, Lung and Circulation (2019) **28,** 1670–1676
1443-9506/04/$36.00
https://doi.org/10.1016/j.hlc.2018.09.004

ORIGINAL ARTICLE

Assessment of the EuroSCORE II in a New Zealand Tertiary Centre

Navneet Singh [a,b*], Damian Gimpel, MBBS [b,c], Grant Parkinson, FRACS [b], Paul Conaglen, MS, FRACS [b], Felicity Meikle, MMedSc, FRACS [b], Zaw Lin, FRACS [b], Nand Kejriwal, MS, MCh, FRACS [b,c], Nicholas Odom, MD, FRCS, FRACS [b], David J. McCormack, FRCSEd (CTh), FRCS (CTh) [b,c], Adam El-Gamel, MD, FRCSEd (CTh), FRCS, FRACS [b,d]

[a]School of Medicine, The University of Auckland, Auckland, New Zealand
[b]Waikato Cardiothoracic Unit, Waikato Hospital, Hamilton, New Zealand
[c]Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand
[d]The University of Waikato, Hamilton, New Zealand

Received 4 August 2018; accepted 15 September 2018; online published-ahead-of-print 27 September 2018

Background The updated European System for Cardiac Operative Risk Evaluation (EuroSCORE II) is a well-established cardiac surgery risk scoring tool for estimating operative mortality. This risk stratification system was derived from a predominantly European patient cohort. No validation analysis of this risk model has been undertaken for the New Zealand population across all major cardiac surgery procedures. We aim to assess the efficacy (discrimination and calibration) of the EuroSCORE II for predicting mortality in cardiac surgical patients at a large New Zealand tertiary centre.

Methods Data was prospectively collected on patients undergoing cardiac surgery from September 2014 to September 2017 at Waikato Hospital, New Zealand. Patient demographic information, preoperative clinical risk factors and outcome data were entered into a national database. Included patients received either isolated coronary artery bypass grafting (CABG), isolated valve surgery, isolated thoracic aortic surgery, or a combination of these procedures. The primary outcome was the discrimination and calibration of predicted EuroSCORE II risk scores compared with observed 30-day mortality events.

Results 1,666 cardiac surgery patients were included during the study period, with an average EuroSCORE II of 2.97% (95% confidence interval (CI): 2.76–3.18). Nine hundred thirty-three (933) patients underwent isolated CABG, 384 underwent isolated valve surgery, 48 received isolated thoracic aortic surgery and 301 received combination procedures. Thirty-day (30-day) mortality events in each of these groups was 7, 4, 2 and 13 deaths respectively. There were 26 deaths across the total cohort at 30-days (observed mortality rate 1.56%). Discrimination analysis using receiver operating characteristic curves demonstrated the area under the curve (AUC) of the EuroSCORE II in each of these groups as 93.4% (95% CI: 91.6–94.9, $p < 0.0001$), 66.3% (95% CI: 61.3–71.0, $p = 0.37$), 37.0% (95% CI: 15.7–58.2, $p = 0.23$) and 74.8% (95% CI: 69.5–79.6, $p < 0.0001$) respectively. The total cohort AUC was 83.1% (95% CI: 81.2–84.9, $p < 0.0001$). Calibration analysis using Hosmer-Lemeshow tests for the subgroups revealed p-values of 0.848, 0.114, 0.638 and 0.2 respectively. The total cohort Hosmer-Lemeshow p-value was 0.317.

*Corresponding author at: Waikato Cardiothoracic Unit, Waikato Hospital, Pembroke Street, Hamilton West, Hamilton 3204, New Zealand., Email: nsin825@aucklanduni.ac.nz

Conclusions EuroSCORE II showed a strong discriminative ability for isolated CABG 30-day mortality in a New Zealand patient cohort. However, the scoring system discriminated poorly across valvular, thoracic aortic or complex combination cardiac surgical procedures. Good calibration of the EuroSCORE II was achieved across both the total cohort and subgroups. It is important to consider the performance of other cardiac surgery risk stratification models for the New Zealand population.

Keywords Cardiac surgery • Risk assessment • Validation studies • New Zealand

Introduction

One of the most widely used cardiac surgical risk scoring systems is the updated European System for Cardiac Operative Risk Evaluation (EuroSCORE II) [1]. The EuroSCORE II was created in 2011 to replace the logistic EuroSCORE model, which was developed to predict mortality after cardiac surgery [2]. The current model was designed by retrospectively studying a large patient registry of over 22,000 patients from 154 hospitals across 43 predominantly European countries.

The EuroSCORE II utilises a variety of patient-specific factors and procedural aspects to estimate the incidence of perioperative mortality in the context of cardiac surgery. The model is designed to predict operative mortality with high levels of discrimination and calibration across various cardiac surgical procedures. In this context, discrimination is the ability of the model to distinguish between patients who will die following cardiac surgery versus those who will not [3]. Calibration is an assessment of how closely the predicted and observed mortality rates align with one another.

Contemporary studies have considered the efficacy of risk stratification models in accurately predicting operative mortality following major cardiac surgical procedures. For example, Dedda et al. (2013) revealed that, for an Italian population, the EuroSCORE II had a differential level of accuracy in predicting operative mortality depending on the type of cardiac surgical procedure carried out [4]. Similar results were found by Wang et al. (2013) for a Chinese population, Rabbani et al. (2014) for a Pakistan population and Ad et al. (2016) for a North American population [5,6,7].

However, no validation analysis of the EuroSCORE II operative risk model has been undertaken for New Zealand patients across the full breadth of major cardiac surgical procedures. We, therefore, aim to assess the efficacy of this internationally-derived risk stratification tool for the New Zealand population. We hypothesise that the EuroSCORE II has poor discrimination and calibration for a New Zealand population group.

Material and Methods

All adult patients undergoing cardiac surgery at Waikato Hospital, New Zealand between September 2014 and September 2017 were prospectively enrolled into a national database. Hospital ethics approval was gained prior to study conduct.

Patients undergoing isolated coronary artery bypass surgery, isolated valve surgery, isolated thoracic aortic surgery, or a combination of these procedures were analysed. A description of the types of procedures included in these subgroups is provided in Table 1.

Patient demographic, clinical, operative and mortality outcome data was extracted from the national database. This database contained information on all risk factors and input variables in the EuroSCORE II calculator [8]. Any missing data was acknowledged as 'unknown' in data tables. Available data was used to calculate a patient's EuroSCORE II mortality risk as per the online calculator. For the purposes of the calculator, creatinine clearance was calculated using the Cockcroft-Gault equation [9].

The primary outcome was to assess how well the predicted EuroSCORE II mortality risk correlated with observed 30-day mortality events for each patient group (isolated coronary surgery, valve surgery, thoracic aortic surgery, or

Table 1 Description of subgroups.

Cohort	Included procedures
Isolated CABG	Single CABG
	Multiple CABG
Isolated valve surgery	Aortic valve repair/replacement
	Mitral valve repair/replacement
	Tricuspid valve repair/replacement
	Multiple valve repair/replacement
Isolated thoracic aortic surgery	Aortic root surgery (valve-sparing)
	Ascending aortic surgery
	Aortic arch surgery
	Descending thoracic aortic surgery
	Aortic dissection repair
	PEARS
Combination procedures	CABG + valve repair/replacement
	CABG + thoracic aortic surgery
	Valve repair/replacement + thoracic aortic surgery[a]
	CABG + Other
	Valve surgery + Other
	Thoracic aortic surgery + Other

N.B. 'Other' includes procedures such as atrial ablation for atrial fibrillation or atrial septal defect closure.

Abbreviations: CABG, coronary artery bypass graft; PEARS, personalised external aortic root support.

[a]Includes Bentall procedure.

Table 2 Baseline demographic and procedural characteristics.

Characteristic		N = 1666 (%)
Gender	Males	1258 (76)
	Females	408 (24)
Age (mean +/− SD, years)		65 +/− 11
EuroSCORE II (mean +/− SD, %)		2.97 +/− 4.46
Ethnicity	NZ Europeans	1,220 (73)
	Maori	287 (17)
	Other	159 (10)
Procedure	Isolated CABG	933 (56)
	Isolated valve	384 (23)
	Isolated thoracic aortic	48 (3)
	Combination	301 (18)

Abbreviations: NZ, New Zealand; CABG, coronary artery bypass graft; SD, standard deviation.

combination procedures) using discrimination and calibration statistics.

Statistical Analyses

Data was analysed using Biosoft easyROC software and Medcalc for Windows, version 18.5 (Medcalc Software, Ostend, Belgium) [10,11].

Accuracy and discrimination of the EuroSCORE II model for our data was assessed using receiver operating characteristic (ROC) curves. Nonparametric curves were created, with p-values calculated using the exact binomial method with type 1 error set at 0.05. The area under the ROC curves (AUCs) was compared across patient groups. AUCs are presented as a percentage with a 95% confidence interval and associated p-value at a significance level of 0.05.

Calibration was assessed using the Hosmer-Lemeshow logistic regression test. A corresponding chi-squared statistic and p-value was obtained. A p-value <0.05 was taken to mean that the model was a poor fit for the data.

Continuous parameters are presented as mean (± standard deviation, SD) and categorical data as number (percentage).

Results

There was a total of 1,666 patients undergoing major cardiac surgery during the study period. The baseline characteristics and preoperative risk profile of these patients is presented in Tables 2 and 3. The majority of the patient sample were male (76%) and of New Zealand European (73%) ethnicity. The mean sample EuroSCORE II was 2.97% (95% confidence interval [CI]: 2.76–3.18). The most common procedure carried out was isolated coronary artery bypass grafting, which made up 56% of all procedures performed.

Table 3 Cohort risk profile preoperatively.

Characteristic		Value (%)
Renal impairment	Normal	821 (49.3)
	Moderate	693 (41.6)
	Severe	152 (9.1)
	Dialysis	0 (0)
Extracardiac arteriopathy		136 (8.2)
Poor mobility		72 (4.3)
Previous cardiac surgery		262 (15.7)
Chronic lung disease		190 (11.4)
Active endocarditis		22 (1.3)
Critical preoperative state		112 (7.3)
Diabetes on insulin		109 (6.5)
NYHA class	1	634 (38.1)
	2	615 (36.9)
	3	339 (20.3)
	4	75 (4.5)
	Unknown	3 (0.2)
CCS class 4 angina		133 (8.0)
Recent myocardial infarction		475 (28.5)
Pulmonary hypertension	Moderate	218 (13.1)
	Severe	43 (2.6)
Urgency	Elective	1,042 (62.5)
	Urgent	538 (32.3)
	Emergency	81 (4.9)
	Salvage	5 (0.3)
Surgery on thoracic aorta		156 (9.4)

Abbreviations: NYHA, New York Heart Association classification for heart failure; CCS, Canadian Cardiovascular Society.

Table 4 30-day mortality events.

Cohort	Number of 30-day deaths (% of cohort)
Isolated CABG (n = 933)	7 (0.8)
Isolated valve (n = 384)	4 (1.0)
Isolated thoracic aortic (n = 48)	2 (4.2)
Combination (n = 301)	13 (4.3)

Abbreviations: CABG, coronary artery bypass graft.

There was a total of 26 deaths (observed mortality rate of 1.56%) across the total patient sample at 30 days post-procedure, with the greatest proportion of deaths occurring in the combination procedures group at 4.3% (Table 4).

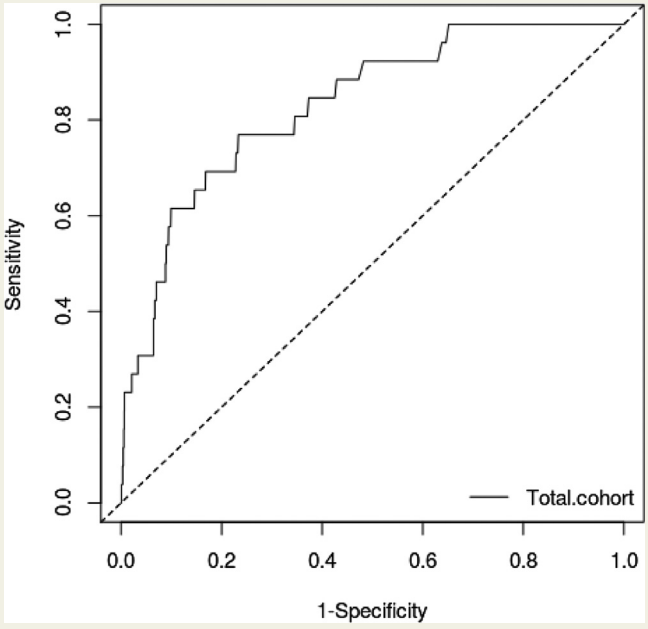

Figure 1 ROC curve for the total patient cohort.
Abbreviations: ROC, receiver operating characteristic.

Table 5 ROC curve subgroup analysis.

Procedure	Area under curve (%)	95% CI	ROC P-value
Isolated CABG	93.4	91.6–94.9	<0.0001
Isolated valve	66.3	61.3–71.0	0.37
Isolated thoracic aortic	37.0	15.7–58.2	0.23
Combination	74.8	69.5–79.6	<0.0001

Abbreviations: CABG, coronary artery bypass graft; ROC, receiver operating characteristic.

Table 6 Hosmer-Lemeshow test analysis.

Procedure	Chi-squared statistic	p-value
Isolated CABG	4.10	0.848
Isolated valve	12.94	0.114
Isolated thoracic aortic	5.18	0.638
Combination	11.03	0.2

Abbreviations: CABG, coronary artery bypass graft.

The receiver operating characteristic (ROC) curve for the total patient sample (Figure 1) revealed an area under the curve (AUC) of 83.1% (95% CI: 81.2–84.9, $p < 0.0001$). The logistic regression Hosmer-Lemeshow p-value obtained was 0.317.

For subgroup analysis (Figures 2–5, Table 5), the highest AUC of 93.4% (95% CI: 91.6–94.9, $p < 0.0001$) was achieved in the isolated CABG group. Statistically significant AUC values were not achieved for the isolated valve surgery and isolated thoracic aortic surgery groups ($p = 0.37$ and 0.23 respectively).

Logistic regression Hosmer-Lemeshow test results for the subgroups are presented in Table 6. This calibration analysis revealed that the EuroSCORE II model was a good fit for all the subgroups.

Discussion

We found that the EuroSCORE II model achieves reasonable discrimination for our total cohort of New Zealand cardiac surgical patients. Subgroup analysis revealed that the risk scoring system achieves strong discrimination for patients undergoing isolated coronary surgery, but relatively poor discrimination for patients undergoing non-coronary or complex combination procedures. Calibration analysis for the total cohort and for all subgroups showed good fit of the data to the EuroSCORE II model. This is the first time that the performance of the EuroSCORE II model has been assessed across different types of major cardiac surgery in a New Zealand population.

A valid risk model should exhibit both robust discrimination and calibration in order to have good predictability for mortality outcomes [12]. The original EuroSCORE model was updated because it did not predict operative mortality well, particularly for valve surgery [13]. Our subgroup analysis has shown that the updated EuroSCORE II still has weak predictability with poor discrimination outside of isolated coronary surgery. The EuroSCORE II authors (Nashef et al. 2012) estimated the ability of their model to discriminate with a ROC curve AUC of 80.95% across all procedures. The

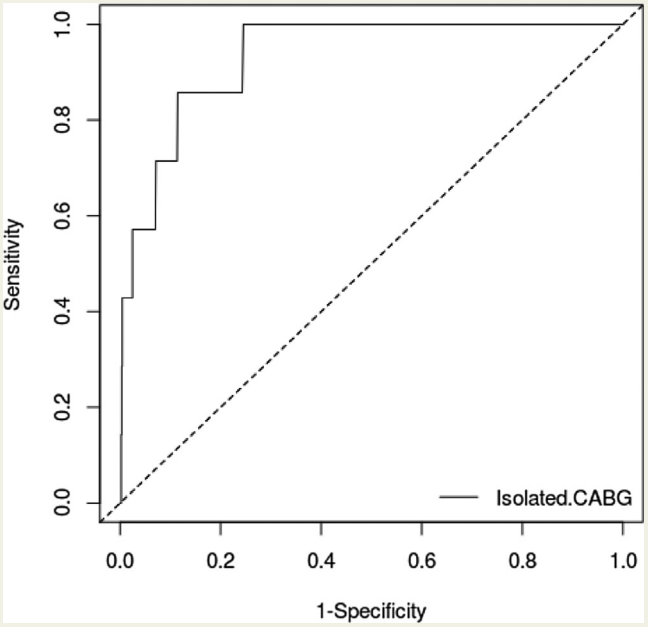

Figure 2 ROC curve for isolated CABG surgery group.
Abbreviations: CABG, coronary artery bypass graft; ROC, receiver operating characteristic.

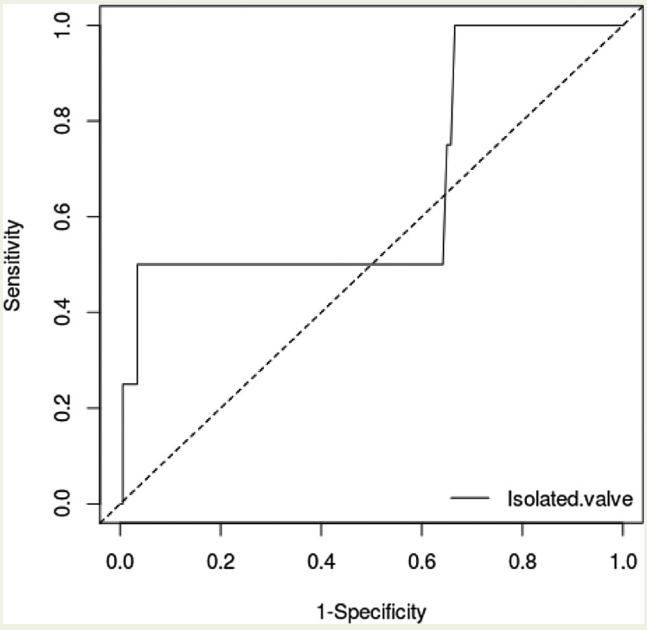

Figure 3 ROC curve for isolated valve surgery group.
Abbreviations: ROC, receiver operating characteristic.

present study did not find this level of discrimination was reached across all procedure subgroups. This result questions the validity of using the EuroSCORE II for New Zealand cardiac surgery patients outside of isolated CABG surgery. Therefore, a similar validation analysis of other international risk stratification tools (such as the AusSCORE II) to the New Zealand population is warranted [14]. If such an analysis also yielded unsatisfactory discrimination results, then this may suggest that the international data or the risk variables themselves used in current international risk models do not accurately estimate operative mortality for New Zealand patients. Hence the development of a New Zealand-specific cardiac surgical risk scoring system may be warranted.

The discrimination results reflect both the design of the EuroSCORE II model and biological variability. The majority of patients in the cohort used to originally develop the risk

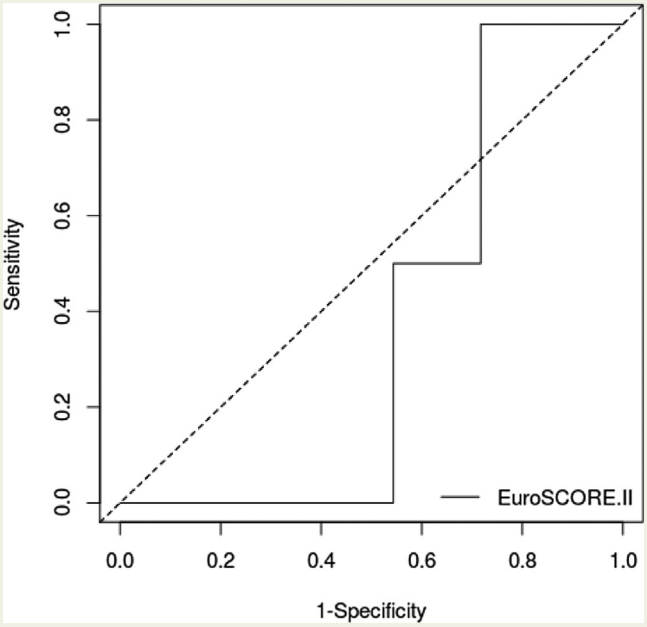

Figure 4 ROC curve for isolated thoracic aortic surgery group.
Abbreviations: ROC, receiver operating characteristic.

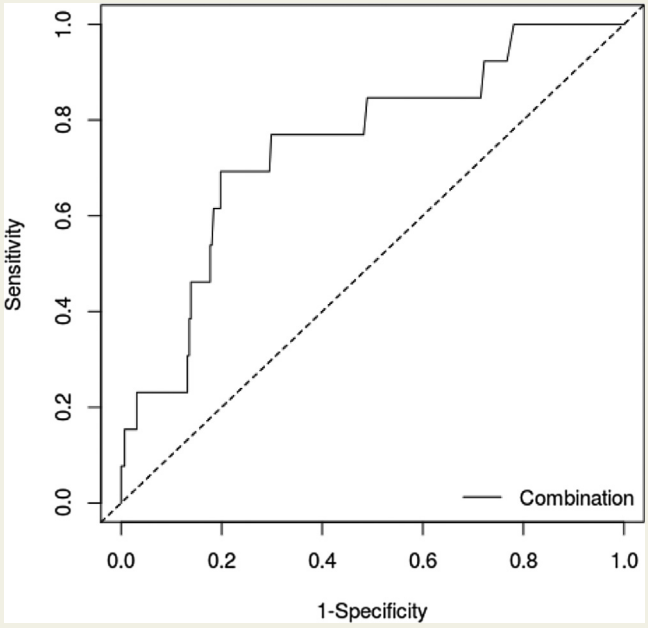

Figure 5 ROC curve for combination procedures group.
Abbreviations: ROC, receiver operating characteristic.

scoring system underwent isolated CABG, and so the model is inherently best powered for this particular subgroup of procedures. Furthermore, coronary surgery from a biological standpoint is more predictable in its outcomes based on the preoperative risk factors measured, compared to the highly variable outcomes of more risk-averse thoracic aortic surgery [15].

The primary limitation in our analysis was a relatively small sample size with a low number of mortality events occurring, particularly in subgroup analysis for isolated valve surgery and isolated thoracic aortic surgery. Because the study was underpowered in the valve surgery subgroup, we did not stratify and analyse this group across aortic, mitral and tricuspid surgery separately, or stratify single

valve versus multiple valve surgery. Given the importance of accurately assessing the efficacy of the EuroSCORE II in these subgroups, a larger repeat analysis is indicated.

Conclusions

The EuroSCORE II has limited ability to predict mortality outcomes for our cohort outside of isolated coronary surgery. Further work at a larger scale is advisable to assess the alignment of the EuroSCORE II risk scoring tool to the wider New Zealand population and ensure that the use of the model in the clinical management of patients is both safe and accurate. A similar validation analysis of other commonly used international cardiac surgical risk scoring systems to the New Zealand population would also be helpful. There is a potential future role in producing a cardiac surgery risk stratification tool specific to New Zealand.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of Interest

None.

References

[1] Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. Eur J Cardiothorac Surg 2012;4:734–44.
[2] Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothoracic Surg 1999;16(1):9–13.
[3] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology 2010;21(1):128–38.
[4] Dedda UD, Pelissero G, Agnelli B, Di Vincentiis C, Castelvecchio S, Ranucci M. Accuracy, calibration and clinical performance of the new EuroSCORE II risk stratification system. Eur J Cardiothoracic Surg 2013;43(1):27–32.
[5] Wang L, Han Q, 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 Cardiothoracic Surg 2013;45(2):359–64.
[6] Rabbani MS, Qadir I, Ahmed Y, Gul M, Sharif H. Heart valve surgery: EuroSCORE vs. EuroSCORE II vs. Society of Thoracic Surgeons score. Heart Int 2014;9(2):53–8.
[7] Ad N, Holmes SD, Patel J, Pritchard G, Shuman DJ, Halpin L. Comparison of EuroSCORE II, original EuroSCORE, and the Society of Thoracic Surgeons risk score in cardiac surgery patients. Ann Thorac Surg 2016;102:573–9.
[8] EuroSCORE Study Group. EuroSCORE interactive calculator; 2011, Available from: http://www.euroscore.org/calc.html. [Accessed 8th June 2018].
[9] Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16(1):31–41.
[10] Goksuluk D, Korkmaz S, Zararsiz G, Karaağaoğlu AE. easyROC: An interactive web-tool for ROC curve analysis using R language environment. R J 2016;8(2):213–30.
[11] MedCalc for Windows Software version 18.5, Ostend, Belgium. Available from: https://www.medcalc.org/index.php. [Accessed 12th June 2018].
[12] Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. JAMA 2017;318(14):1377–84.
[13] Parolari 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.
[14] Billah B, Huq MM, Smith JA, Sufi F, Tran L, Shardey GC, et al. AusSCORE II in predicting 30-day mortality after isolated coronary artery bypass grafting in Australia and New Zealand. J Thorac Cardiovasc Surg 2014;148(5):1850–5.
[15] Hernandez-Vaquero D, Diaz R, Pascual I, Alvarez R, Alperi A, Rozado J, et al. Predictive risk models for proximal aortic surgery. J Thorac Dis 2017;9(Suppl 6):S521–5.