

Validation of EuroSCORE II in Chinese Patients Undergoing Coronary Artery Bypass Surgery

Li Shen, MD, Xi Chen, MD, Jianming Gu, MD, Song Xue, MD

Department of Cardiovascular Surgery, Renji Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China

ABSTRACT

Background: Comparisons between the EuroSCORE and EuroSCORE II in the patient populations for coronary artery bypass grafting are limited. The aim of the study was to compare the use of the EuroSCORE and EuroSCORE II as risk model for predicting in-hospital mortality in Chinese patients undergoing coronary artery bypass grafting (CABG).

Methods: Patients (n = 1598) with complete records of baseline and operative data were retrospectively collected from computerized records. The expected mortality rate for logistic EuroSCORE and EuroSCORE II was determined. Performance of the logistic EuroSCORE and EuroSCORE II model was assessed by comparing the observed and expected in-hospital mortality. The area under the receiver operating characteristics curve (AUC) values were calculated for these models to compare predictive power.

Results: Observed in-hospital overall mortality rate was 3.19%. The logistic EuroSCORE model (Hosmer-Lemeshow: $P < .05$, O/E = 0.73) over-predicted mortality (4.39%) and the EuroSCORE II model showed good calibration and discriminative capacity (area 0.762) in predicting in-hospital mortality (Hosmer-Lemeshow: $P = .191$, O/E = 1.24).

Conclusion: EuroSCORE II model reduces the overestimation of the calculated risk by logistic EuroSCORE in this population. EuroSCORE II risk model may be suitable in patients undergoing coronary artery bypass surgery in China.

INTRODUCTION

Risk prediction models have been developed in the last few decades that are an essential part of current cardiac surgical practice in estimating operative mortality and morbidity. Aimed at assessing a 30-day mortality using an additive (standard) model constructed from a multinational European adult cardiac surgery database, the European System for Cardiac Operative Risk Evaluation (EuroSCORE) was created in 1999 [Nashef 1999; Roques 1999], and validated in Europe, North America and elsewhere. Four years later, the 2003 logistic EuroSCORE model was developed to improve the predictive performance in high-risk patient groups [Roques 2003; Michel 2003; Shanmugam 2005].

Received November 6, 2016; received in revised form July 16, 2017; accepted July 27, 2017.

Correspondence: Song Xue, No. 160 Pujian Road, Pudong New Area, Shanghai P.R. China; (e-mail: xuesong64@163.com).

However, developments and achievements in perioperative treatment, surgical, and anesthesia techniques have decreased the mortality and morbidity of cardiac surgery over the years. A number of publications from different countries have identified that the initial EuroSCORE overestimates the risk of recent cardiac surgery procedures [Parolari 2010; Massoudy 2011; Zheng 2009; Parolari 2009]. On October 3, 2011, the EuroSCORE investigators launched EuroSCORE II model at the 2011 EACTS (European Association for Cardiothoracic Surgery) meeting in Lisbon as an update [Nashef 2012]. The EuroSCORE II was developed from an international cohort of 22,381 patients undergoing cardiac surgery during 2010 and published in 2012. Recent studies which have assessed the validity of this new score have reported mixed results, with EuroSCORE II performing better or similar to EuroSCORE [Biancari 2012; Chalmers 2013; Barili 2013]. Similar comparisons between the EuroSCORE and EuroSCORE II in the patient populations for coronary artery bypass grafting are limited. Therefore, we aimed to compare the use of the EuroSCORE and EuroSCORE II as risk model for predicting in-hospital mortality in Chinese patients undergoing coronary artery bypass grafting (CABG), whether or not in combination with other major cardiac surgery at our center.

PATIENTS AND METHODS

Study Population and Data Collection

Patients with incomplete data according to EuroSCORE and EuroSCORE II criteria were excluded; patients (n = 1598) with complete records of baseline and operative data at Renji Hospital in China from January 2005 to December 2014 were included. Relevant clinical characteristics were retrospectively collected from computerized records obtained from our patient database. Patients who underwent coronary artery bypass surgery, whether or not in combination with other major cardiac surgery, were included in the study. Risk factors recorded in patients' history were categorized according to the EuroSCORE and EuroSCORE

II criteria. Definitions proposed by the EuroSCORE were fully adopted. Of the new risk variables used in the EuroSCORE II, CCS (Canadian Cardiovascular Society) angina class, elective, urgent, emergency, or salvage surgery, preoperative dialysis and insulin-dependent diabetes were registered routinely as a patient's history in the database. The expected mortality rate for logistic EuroSCORE and EuroSCORE II was determined with the online interactive calculator (www.euroscore.org/calc.html). Creatinine clearance (CC) as an estimate of glomerular filtration rate was calculated using the

Table 1. Prevalence of Risk Factors in EuroSCORE II Population and the Study Sample

	EuroSCORE II Prevalence	Study Sample Prevalence	P
n	22381	1598	
Age, y	64.6 ± 12.5	65.8 ± 10.6	NS
Female, n (%)	6919 (30.9)	448 (28.0)	NS
BMI, kg/m ²	27.4 ± 4.8	24.57 ± 4.16	<.001
Diabetes mellitus, n (%)	5598 (25.0)	412 (25.8)	NS
IDDM, n (%)	1705 (7.6)	137 (8.6)	NS
Pulmonary disease, n (%)	2384 (10.7)	125 (7.8)	<.001
Neurological dysfunction, n (%)	713 (3.2)	68 (4.3)	<.05
Serum creatinine, μmol/L	96.4 ± 57.1	77.2 ± 23.1	<.001
CC, mL/min	83.6 ± 50.9	95.2 ± 18.3	<.001
On dialysis, n (%)	244 (1.1)	4 (0.3)	.001
Active endocarditis, n (%)	497 (2.2)	6 (0.4)	<.001
Critical preoperative state, n (%)	924 (4.1)	88 (5.5)	<.01
Isolated CABG, n (%)	10448 (46.7)	1340 (83.9)	<.001
Urgent, n (%)	4135 (18.5)	98 (6.1)	<.001
Emergency, n (%)	972 (4.3)	32 (2.0)	<.001
Salvage, n (%)	109 (0.5)	2 (0.1)	<.05

CC indicates creatinine clearance; IDDM, insulin-dependent diabetes mellitus; NS, not significant.

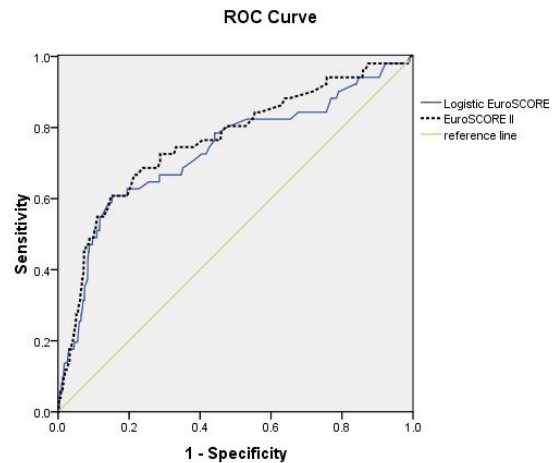
Cockcroft-Gault formula performed on the website as well. Operative mortality was defined as in-hospital death before discharge. Our Ethics Committee relieved the requirement of review for its retrospective manner.

Statistical Analysis

Data were analyzed using SPSS version 19.0 (SPSS, Chicago, IL, USA) for Windows.

Continuous variables were shown as mean ± SD and comparisons between them were performed by Student t test. Categorical variables were presented as percentages and differences were assessed using Fisher exact test or Pearson chi-square test. *P* values less than .05 were deemed statistically significant.

Performance of the logistic EuroSCORE and EuroSCORE II model was assessed by comparing the observed and expected in-hospital mortality. Calibration (statistical precision) was analyzed by Hosmer-Lemeshow goodness-of-fit test. A well-calibrated model gives a *P* value greater than .05. Discriminative powers were assessed using the area under the receiver operating characteristic (ROC) curves (AUC/c-statistic), which were used to assess how well the model could discriminate between survivors and non-survivors. Areas of greater than 0.7 are generally considered to be useful. All statistical tests were two-tailed.



ROC curve for logistic EuroSCORE (AUC = 0.733) and EuroSCORE II model (AUC = 0.762) when applied to the entire group.

RESULTS

Patients' Characteristics

The prevalence of risk factors in the EuroSCORE II population and our study sample are shown in Table 1. Some significant differences were seen between the two populations, although with the same age and sex group. The study population had a smaller calculated BMI and was more likely to have neurological dysfunction. More study patients were in a critical preoperative state and the EuroSCORE II patients were more likely to have pulmonary disease and active endocarditis. EuroSCORE II patients had worse renal function with a higher serum creatinine level and lower creatinine clearance rate. More EuroSCORE II patients underwent urgent or emergency operation. There were no significant differences in the prevalence of diabetes mellitus or insulin-dependent diabetes.

Calibration

For the entire cohort, observed hospital mortality was 3.19% (51 deaths). Table 2 and Table 3 show the predictive ability of the logistic EuroSCORE and EuroSCORE II for the entire group. Patients were also equally divided into three risk tertiles in each model for stratified analysis.

The logistic EuroSCORE and EuroSCORE II predicted a mortality rate of 4.39% (±3.84%) and 2.58% (±2.40%) for the entire cohort, respectively. The results show that the logistic EuroSCORE model (Hosmer-Lemeshow: *P* < .05, O/E = 0.73) overpredicted mortality and the EuroSCORE II model showed good calibration in predicting in-hospital mortality (Hosmer-Lemeshow: *P* = .191, O/E = 1.24)

Discriminative Power

The discriminative power of the logistic EuroSCORE and EuroSCORE II model for the entire group was acceptable with an area under the ROC curve of 0.733 for the logistic EuroSCORE model, 0.762 for the EuroSCORE II model (Figure).

Table 2. Predicted and Observed Mortality by Logistic EuroSCORE Risk Levels for All Patients

	Patients (deaths)	Observed mortality rate	Expected mortality rate	Recalibration (O/E)	H-L Test P Value
Logistic EuroSCORE					
Low risk*	533 (9)	1.69%	1.75% (0.88%-2.54%)	0.97	
Medium risk	533 (8)	1.50%	3.22% (2.54%-4.34%)	0.47	
High risk	532 (34)	6.39%	8.19% (4.34%-48.85%)	0.78	
Total	1598 (51)	3.19%	4.39%	0.73	<.05

*Patients were divided into three approximately equal risk tertiles for logistic EuroSCORE analysis.

Table 3. Predicted and Observed Mortality by Logistic EuroSCORE II Risk Levels for All Patients

	Patients (deaths)	Observed mortality rate	Expected mortality rate	Recalibration (O/E)	H-L Test P Value
EuroSCORE II					
Low risk*	533 (5)	0.94%	1.02% (0.56%-1.25%)	0.92	
Medium risk	533 (9)	1.69%	1.79% (1.25%-2.24%)	0.94	
High risk	532 (37)	6.95%	4.94% (2.24%-24.33%)	1.41	
Total	1598 (51)	3.19%	2.58%	1.24	.191

*Patients were divided into three approximately equal risk tertiles for EuroSCORE II analysis.

DISCUSSION

Risk models are useful in current cardiac surgical practice and enable comparison between centers. Local validation of updated operative risk models are necessary to polish the quality of clinical research and patients' care. The most widely used original additive and logistic EuroSCORE model is no longer suitable for predicting in-hospital mortality in contemporary cardiac surgery. It overestimated the individual patient risk which may severely affect any critical assessment of clinical outcomes and lead to a false sense of reassurance, and may compromise patients' welfare [Choong 2009]. The EuroSCORE investigators published a modified risk scoring method in order to enhance and optimize its usefulness in contemporary cardiac surgical practice in April 2012 [Nashef 2012]. It was reported [Biancari 2012] that the EuroSCORE II performs better than its original version in predicting operative mortality and morbidity after isolated CABG in Europe. The EuroSCORE was introduced to China in 2000 due to the lack of a local risk model. But, reconfirmed in this study, Zheng et al [Zheng 2009] in 2008 reported that the logistic EuroSCORE over-predicted CABG patients at all risk tertiles in China with possible reasons being advances in surgical, anesthetic, and intensive care as well as different epidemiology and comorbidities. One year after the update of EuroSCORE, it was reported that the EuroSCORE II model gives an accurate prediction for individual operative risk in patients undergoing single valve surgery but an imprecise prediction in patients undergoing multiple valve surgery in China [Zhang 2013].

In this study, therefore, we have analyzed the predictive value of the EuroSCORE II model for predicting operative mortality in patients undergoing coronary artery bypass surgery at our center. We validated the good risk prediction performance of EuroSCORE II in our CABG patients. Compared with the poor calibration of logistic EuroSCORE, the EuroSCORE II proved to be a successful advance with a better calibration and discriminatory ability for the entire CABG group. However, when the whole study group is analyzed, EuroSCORE II is good, but in high-risk tertile (expected mortality range from 2.24% to 24.33% individually) with an expected mortality rate of 4.94% and observed mortality rate of 6.95% (O/E = 1.41), the EuroSCORE II shows a trend of slightly underestimating the true mortality. Despite the insignificant H-L test value and acceptable AUC of the entire group, this result in the high-risk tertile could have negative consequences for surgeons who operate on patients with certain risk profiles.

EuroSCORE II investigators adjusted the grading strategy of the risk factors for potential improvement [Nashef 2012]. For instance, calculated by Cockcroft-Gault formula, with a better sensitivity in mild renal insufficiency patients, creatinine clearance instead of serum

creatinine level was used to evaluate the severity of renal failure. This might have significantly strengthened the predictive ability of EuroSCORE II in patients with potential renal injury. Due to the different baseline of the patients, our patients had better renal function with a lower serum creatinine level and higher creatinine clearance rate. This meant

that more patients with mild renal insufficiency had been considered when using the EuroSCORE II model. Moreover, a better grading of operation weight, and urgency and symptomatic status of cardiac-related factors such as NYHA (New York Heart Association) class & CCS angina class with the other advances may have also significantly contributed to improving this risk predicting method.

A limitation of this study is that the analysis was done in a single center and in a retrospective way. The assessment of performance of the EuroSCORE II model for predicting operative mortality in China could be biased, thus the results need to be further reexamined with a larger number of patients in a multi-center database. Moreover, the EuroSCORE II predicts hospital mortality, defined as all-cause (either cardiac or non-cardiac) mortality in the hospital where the operation took place, which means that we are unable to determine the major cause of death. In-hospital mortality, not 30-day status, may also underestimate the operative mortality.

In conclusion, the present study suggests that the EuroSCORE II model has succeeded in improving the performance of the EuroSCORE model in CABG patients in China. Additional validation of EuroSCORE II will be necessary in larger populations.

REFERENCES

- Barili F, Pacini D, Capo A, et al. 2013. Does EuroScore II perform better than its original versions? A multi centre validation study. *Eur Heart J* 34:22-9.
- Biancari F, Vasques F, Mikkola R, Martin M, Lahtinen J, Heikkinen J. 2012. Validation of EuroSCORE II in patients undergoing coronary artery bypass surgery. *Ann Thorac Surg* 93:1930-5.
- Chalmers J, Pullan M, Fabri B, et al. 2013. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. *Eur J Cardiothorac Surg* 43:688-94.
- Chen CC, Wang CC, Hsieh SR, Tsai HW, Wei HJ, Chang Y. 2004. Application of European system for cardiac operative risk evaluation (EuroSCORE) in coronary artery bypass surgery for Taiwanese. *Interact Cardiovasc Thorac Surg* 3:562-5.
- Choong CK, Sergeant P, Nashef SA, Smith JA, Bridgewater B. 2009. The EuroSCORE risk stratification system in the current era: how accurate is it and what should be done if it is inaccurate? *Eur J Cardiothorac Surg* 35:59-61.
- Kawachi Y, Nakashima A, Toshima Y, Arinaga K, Kawano H. 2001. Risk stratification analysis of operative mortality in heart and thoracic aorta surgery: comparison between Parsonnet and EuroSCORE additive model. *Eur J Cardiothorac Surg* 20:961-6.
- Massoudy P, Sander J, Wendt D, Thielmann M, Jakob H, Herget-Rosenthal S. 2011. Does the EuroSCORE equally well predict perioperative cardiac surgical risk for men and women? *Minim Invasive Ther Allied Technol* 20:67-71.
- Michel P, Roques F, Nashef SA. 2003. Logistic or additive EuroSCORE for high-risk patients? *Eur J Cardiothorac Surg* 23:684-7.
- Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. 1999. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 16:9-13.
- Nashef SA, Roques F, Hammill BG, et al. 2002. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. *Eur J Cardiothorac Surg* 22:101-5.
- Nashef SA, Roques F, Sharples LD, et al. 2012. EuroSCORE II. *Eur J Cardiothorac Surg* 41:734-744; discussion 44-5.
- Parolari A, Pesce LL, Trezzi M, et al. 2009. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: single institution experience and meta-analysis. *Eur Heart J* 30:297-304.
- Parolari A, Pesce LL, Trezzi M, et al. 2010. EuroSCORE performance in valve surgery: a meta-analysis. *Ann Thorac Surg* 89:787-93.
- Roques F, Nashef SA, Michel P, et al. 1999. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. *Eur J Cardiothorac Surg* 15:816-22; discussion 22-23.
- Roques F, Michel P, Goldstone AR, Nashef SA. 2003. The logistic EuroSCORE. *Eur Heart J* 24:881-2.
- Shanmugam G, West M, Berg G. 2005. Additive and logistic EuroSCORE performance in high risk patients. *Interact Cardio Vasc Thorac Surg* 4:299-303.
- Zhang GX, Wang C, Wang L, et al. 2013. Validation of EuroSCORE II in Chinese patients undergoing heart valve surgery. *Heart Lung Circ* 22:606-11.
- Zheng Z, Li Y, Zhang S, Hu S. 2009. The Chinese coronary artery bypass grafting registry study: how well does the EuroSCORE predict operative risk for Chinese population? *Eur J Cardiothorac Surg* 35:54-8.