Comparison of the Additive, Logistic European System for Cardiac Operative Risk (EuroSCORE) with the EuroSCORE 2 to Predict Mortality in High-Risk Cardiac Surgery

Abstract

Background: The aim of this study was to compare the new EuroSCORE (ES) 2 prediction model in high-risk patients with the 2 other oldest additive ES (aES) and logistic ES (IES). Methods: Consecutive adult patients undergoing all cardiac surgery except heart transplantation and left ventricular assist device were included. The 3 risk scores were collected before surgery. We defined 4 high-risk groups of patients, patients ≥80 years, combined cardiac surgery, surgery of the thoracic aorta, and emergency cardiac surgery, and 2 low-risk groups, valve surgery and coronary artery bypass surgery. The predicted value of each score has been assessed by the area under the receiver operating characteristics curve (AUC). Results: The study had included 3301 patients. Thirty-day mortality was 3.9% (95% confidence interval (CI), 3.3 - 4.6%). The AUC of ES2 was 0.81 (0.77 - 0.84), 0.82 (0.78 - 0.85), 0.70 (0.64 - 0.76), 0.79 (0.74 - 0.83), 0.85 (0.83 - 0.87), and 0.88 (0.86 - 0.90) for octogenarians, thoracic aortic surgery, combined surgery, emergency surgery, coronary surgery, and valve surgery, respectively. These ES2 AUC values were higher than those obtained with the aES for octogenarians, and with the lES for octogenarians and valve surgery. The ES2 calibration was better than the aES and lES calibration for the whole population, and low-risk groups. The ES2 calibration was superior to aES and lES in high-risk groups, except for octogenarians and thoracic aortic surgery compared to IES. Conclusion: In high-risk cardiac surgery patients, ES2 only marginally improve the predicted 30-day mortality in comparison to other ES.

Keywords: Additive EuroSCORE, EuroSCORE 2, high-risk patients, logistic EuroSCORE, mortality

Introduction

Predicting perioperative mortality risk is essential before cardiac surgery. Scores are used not only to evaluate the results of a cardiac team but also for risk prediction and to modify the operative strategy.^[1] The most used scores in Europe were the additive (aES) first published in 1999^[2] and logistic (IES)^[3] European System for Cardiac Operative Risk (EuroSCORE, ES). However, the performance of these scores has gradually declined;^[4-6] this could be explained by the progressive mortality decline after cardiac surgery due to the technical progress in surgery, anaesthesiology, and perfusion. This decline is observed despite the changes in the patient's profile; nowadays patients undergoing cardiac surgery are older and have more morbidities than the population included in the original ES cohort. Consequently, IES overestimates mortality in high-risk patients.^[7] Since 2012,

the aES^[2] and IES scores have been replaced by EuroSCORE 2 (ES2)^[8] based on logistic regression analyses of 23 000 patients from 150 hospitals; the ES2 offers a better predictive value in unselected cardiac surgical population^[9-11] using only 18 different widely available preoperative clinical- and operation-related factors' variables

Predicting mortality in high-risk group is a major goal of risk scores. Only few studies showed whether ES2 is superior to the other scores to predict mortality in high-risk group. The aim of the present study was to assess the predictive performance of ES2 versus the aES and IES in predefine high-risk patients.

Patients and Methods

Study design

Consecutive adult patients undergoing cardiac surgery were included in this

How to cite this article: Guillet L, Moury PH, Bedague D, Durand M, Martin C, Payen JF, *et al.* Comparison of the additive, logistic European system for cardiac operative risk (EuroSCORE) with the EuroSCORE 2 to predict mortality in high-risk cardiac surgery. Ann Card Anaesth 2020;23:277-82.

Laura Guillet, Pierre H Moury, Damien Bedague, Michel Durand, Cécile Martin¹, Jean F Payen, Olivier Chavanon¹, Pierre Albaladejo

Departments of Anaesthesia and Intensive Care and 'Cardiac Surgery, Hopital Michallon, CHU Grenoble Alpes, F-38000, France

Submitted: 08-Nov-2018 Accepted: 10-Mar-2019 Published: 17-Jul-2020

Address for correspondence: Dr. Michel Durand, Pole Anesthésie-reanimation, Hopital Michallon, CHU Grenoble Alpes, F-38043, Grenoble, France. E-mail: MDurand@ chu-grenoble.fr



This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

retrospective monocentric study conducted from September 2012 to January 2018. Patients who underwent heart transplantation or left ventricular assistance were excluded from the analysis; no patient underwent VAD therapy or was transplanted as a direct consequence of a failed cardiac surgery during the study period. For each patient, the 3 scores were calculated before the surgical procedure. The 3 scores were determined using the online calculator (http:// www.euroscore.org/calc.html) provided by euroscore. org. Data were entered into a prospective database. Four groups of high-risk patients were defined from clinical criteria:^[3] patients aged 80 and over, combined cardiac surgery (defined as valve plus coronary surgery), thoracic aortic surgery (including aortic dissection), and emergency surgery defined as a surgical procedure needed prior to the next working day. In comparison, 2 low-risk groups were also studied: patients undergoing isolated coronary artery bypass grafting (CABG) or valve surgery.

The main goal was to compare the ES2 discriminatory power estimated by the area under the curve (AUC) receiver operating characteristic (ROC) in the 4 high-risk groups and to compare ES2 versus the aES and IES. Mortality was defined as the 30-day mortality.

Statistical analysis

The performance of each score to predict mortality was analysed according to 2 approaches, discrimination and calibration. The discrimination (accuracy of separating nonsurvivors and survivors) of each score was evaluated using the area under the ROC curve (AUC) with the calculation of the exact binomial confidence intervals (CI) of the AUC. The AUC of ES2 was compared to the AUC of the other scores using the Delong method.^[12] The discriminative power of the model is considered reasonable when the AUC is more than 0.7 and strong when the AUC is above 0.8. Calibration is the agreement between the predicted versus observed outcome: the observed - expected (O - E) mortality was calculated for each score in each sub-group of patients and was evaluated by comparison of the 95% CI of each percentage of the O - E. If the 95% CI of the observed mortality excluded the values of the expected mortality, observed mortality was considered statistically significant from the expected mortality. The score underestimated mortality if the value of O - E was greater than 0, and overestimated mortality if the value was less than 0. The calibration was also assessed by logistic regression using the Hosmer - Lemeshow test; the predicted mortality was compared with the observed mortality in the different groups of patients; and a nonsignificant Hosmer - Lemeshow test probability (P > 0.05) was considered clinically acceptable. A P value lower than 0.05 was considered statistically significant. Continuous variables were expressed as mean \pm standard deviation, and categorical variables were expressed as percentages and 95% CIs. Statistical analysis was performed using MedCalc® software for Windows, version 15.0 (Medcalc, Ostend, Belgium).

Results

Whole population

During the study period, 3301 patients were included. The demographic characteristics are shown in Table 1, and surgery and postoperative characteristics are shown in Table 2. The 30-day mortality was 3.9% (95% CI, 3.3 - 4.6%) corresponding to 129 patients. The performance of ES2 was significantly better than aES but not than the IES [Table 3]. The mean difference between predicted and observed mortality was -0.6 for ES2, 5.2 for IES, and 2.0 for aES [Table 4]. The ES2 had the best calibration in the whole population.

High-risk patients

Four hundred and forty-four (13.5%) patients were of age 80 and over. In this group, the distribution of surgical procedures was: valve replacement 46% (n = 195), isolated coronary surgery 23% (n = 112), combined surgeries 18% (n = 82), and thoracic aorta surgery 6% (n = 38). ES2 AUC was significantly higher than the aES and IES AUC [Table 3]. The discrimination of ES2 was significantly better than the discrimination of IES. In this group, the predicted and observed mortality was -1.1 for ES2, 7.9 for IES, and 2.3 for aES [Table 4].

Table 1: Demographical characteristics			
Patient related factors	% or mean (SD)		
Age (y)	67.6 (11.9)		
Female/Male	28%/72%		
Height (cm)	169 (12)		
Weight (kg)	77 (17)		
Diabetes			
Type I	9%		
Type II	18%		
Pulmonary disease	7%		
Peripheral arteriopathy	14%		
Hypertension	63%		
Ischemic attack history			
Transient	2%		
Complete	15%		
Renal impairment			
Creatinine (µmol/L)	98 (66)		
Creatinine >200 µmol/L	2%		
On dialysis	1%		
Left ventricular ejection fraction			
Good >50%	86%		
Moderate 30-50%	12%		
Poor <30%	2%		

Four hundred and eighty-three patients had a thoracic aortic surgery (15%): 219 had Bentall surgery (44%), 224 had

Table 2: Characteristics of surgery and postoperative	
in-hospital stay; CPB: cardio pulmonary bypass	

Perioperative factors	% or mean	
	(SD)	
Surgery		
Elective	73%	
Urgent	17%	
Emergency	10%	
Surgeries without CPB (ECC)	3%	
OPCABG	93%	
Pericardiectomy	4%	
Cardiac wound	3%	
Surgery with CPB	97%	
Duration of ECC (min)	106 (45)	
Duration of aortic clamping (min)	69 (31)	
Length of Surgery (h)	4 (1,3)	
Postoperative period		
Duration of mechanical ventilation (h)	9 (46)	
Duration of intensive care unit stay (d)	4 (5)	
Duration of hospital stay (d)	12 (7)	
Postoperative bleeding within the 24 h (mL)	505 (350)	
Transfusion during a hospital stay	27%	

aortic aneurysm surgery (50%) including 50 surgery for aortic dissection (6%). The ES2 discrimination was better in this group than one of the 2 other scores [Table 3]. The predicted and observed mortality was -2.6 for ES2, 7.3 for IES, and 0.9 for aES [Table 4].

A combined surgery was performed in 334 patients (9%). There was a significant difference between ES2 and aES AUC [Table 3]. The predicted and observed mortality was -1.1 for ES2, 2.1 for IES, and 0.8 for aES [Table 4].

emergency cardiac surgery for An was done 323 patients (11%). The distribution of surgeries was valve surgery 31% (n = 100), thoracic aortic surgery 29% (n = 93), coronary surgery 22% (n = 71), aortic dissection 15% (n = 48), combined surgeries 3% (n = 11), and others 4.2% (n = 14). There was no significant difference between the AUC of the 3 scores [Table 3]. The predicted and observed mortality was -6.4 for ES2, 6.7 for IES, and -6.3 for aES [Table 4]. Figure 1 shows the ROC analyses for the four different high-risk groups.

Low-risk patients

A CABG was performed in 1196 patients (36%). There was no significant difference between the AUC of the 3

Table 3: I	ROC curves AUC values [exact b	inomial confi dence interval (95%	6 CI)]
	EuroSCORE		
	2	Logistic	Additive
Whole population	0.86 [0.85-0.87]	0.84 [0.83-0.86]	0.83 [0.82-0.84]**
HL	< 0.0001	< 0.0001	0.59
Age >80	0.81 [0.77-0.84]	0.73 [0.69-0.78]*	0.72 [0.68-0.76]**
HL	0.09	0.10	0.39
Thoracic aortic surgery	0.82 [0.78-0.85]	0.79 [0.74-0.83]	0.78 [0.74-0.82]
HL	0.03	0.35	0.002
Combined surgery	0.71 [0.66-0.76]	0.72 [0.67-0.77]	0.65 [0.60-0.70]
HL	0.21	0.054	0.72
Emergency surgery	0.79 [0.74-0.83]	0.82 [0.77-0.86]	0.79 [0.74-0.83]
HL	0.02	0.02	0.48
Coronary surgery	0.85 [0.83-0.87]	0.87 [0.85-0.89]	0.85 [0.83-0.87]
HL	0.06	0.31	0.41
Valve surgery	0.88 [0.88-0.93]	0.85 [0.83-0.87]*	0.85 [0.83-0.87]*
HL	0.03	0.14	0.80
* D -0.05 ** D -0.00 1.		D	

*P<0.05; **P<0.02 compared to ES2; HL=Hosmer-Lemeshow statistics P

 Table 4: Comparison of the observed and predicted mortality with the 3 scores in the whole population and the

 different subgroups

different subgroups						
	Observed		EuroSCORE			
	mortality (%)	2	Logistic	Additive		
Whole population	3.9 [3.3-4.6]	3.3 [2.7-3.9]	8.1 [7.2-9.1]*	5.9 [5.1-6.8]*		
Age >80	6.3 [4.2-9.1]	5.2 [3.3-7.8]	14.2 [10.9-18.2]*	8.6 [6.1-11.7]		
Thoracic aortic surgery	7.3 [5.1-10.1]	4.6 [2.9-6.9]*	14.6 [11.5-18.5]*	8.1 [5.7-11.0]		
Combined surgery	6.0 [3.7-9.3]	4.9 [2.7-7.8]	8.1 [5.3-11.8]	6.8 [4.4-10.3]		
Emergency surgery	16.1 [12.1-22.1]	9.7 [6.5-13.6]	22.8 [18.0-28.8]	9.8 [6.8-14.0]		
Coronary surgery	1.8 [0.6-2.3]	2.1 [1.4-3.1]	4.5 [3.4-5.9]*	4.2 [3.1-5.5]*		
Valve surgery	3.4 [1.7-4.2]	3.4 [2.4-4.7]	8.7 [7.1-10.6]*	6.4 [5.0-8.0]*		

*P<0.05 vs. observed mortality

Annals of Cardiac Anaesthesia | Volume 23 | Issue 3 | July-September 2020



Figure 1: ROC curves for the different high-risk groups

scores [Table 3]. In this group, the discrimination of ES2 was better than the 2 other ES [Table 4].

The study included 1144 patients after a valve surgery (35%). There was a significant difference between ES2 AUC values and both aES and IES [Table 3]. The ES2 has the best calibration in this group.

Discussion

In this study, ES2 had a better performance than the other ES in the whole population. Mortality prediction was better with the ES2 in the 2 low-risk group, CABG or valve replacement, with a good calibration. The performance of the ES2 was only marginally different from the aES and IES in the 4 high-risk groups. In our centre, ES2 constantly underestimates mortality in high-risk groups.

These results are similar to those found in previous studies comparing the ES2 score's performance to predict mortality, specifically in high-risk groups. In a previous study comparing the predictive value of the ES2 with the original ES in high-risk patients defined as a $IES \ge 10$,^[13] the ES2 calibration was poor. Similarly, Ranucci^[14] compared

the clinical performance of the ES2 and ACEF scores in high-risk patients, showing poor clinical relevance for these patients. On the other side, several studies^[15-20] that had not specifically included high-risk patients showed better discrimination for ES2 than previous scores with AUC values (around 0.80) close to the values of the original study;^[8] however, in these studies, the ES2 underestimates mortality for high-risk patients and overestimates mortality for low-risk patients. One limitation is that none of these studies had defined the high-risk patients' groups with the same criteria or with the same ES cut-off value. Previously, aES and IES had shown low calibration in high-risk groups^[15,21] and the whole population^[22] even if they performed better than older scores.^[23,24] Weak calibration of aES and lES models was previously found in octogenarians,^[25] in the combined surgery group^[26] or the isolated valve surgery group.^[27]

The poor calibration of these scores in high-risk patients might be partly explained by the difference between current population undergoing cardiac surgery and the population used to develop these scores 20 years ago. Indeed, the patients in our study seemed to be at higher risk than in Nashef's initial study.^[8] In our study, patients were older with a mean age of 67.9 years vs 64.6 years in the ES2 study and 62.5 years in the original ES study. Surgery risk profiles in this study were also different from those observed in ES2 study, 15% vs 7% for thoracic aorta surgery and 11% vs 5% for emergency surgery, respectively. On the contrary, low-risk patients were less numerous in our study than in the ES2 publication, 34% vs 47% for CABG and 33% vs 46% for valve surgery. In the original ES study, about 60% of the patients were CABG with a mean aES value 4.8,^[2] the mean aES value was 6.0 in this study. We chose to define high-risk groups of patients according to clinically relevant criteria^[3] and not to scores values, since 1) any cut-off value is controversial^[15-17,25,26] and 2) the predictive value of a score cannot be calculated in a high-risk group defined by a high value of the same score. We did not evaluate the ES2 calibration in the acute aortic dissections group because their number was too limited (n = 50, 1.5%). A previous study showed the limited performance of the ES2 in this specific group.^[28] Finally, the best way to improve the risk- adjusted result of a cardiac surgeon is probably to operate mainly low-risk patients.

This study has some limitations. The first limitation remains in the retrospective data collection, providing a possible bias in scores calculation. However, the data were obtained from a prospective database and there were only few missing data (less than 5%), which were easily retrieved from the medical charts and anesthesia sheets. Another limitation is that this is a single-center study. However, it included a high number of patients, which is larger than previous studies.^[13] The definition of mortality is different for each score which makes comparison difficult. In our study, mortality was defined as 30-day mortality. In the ES2 study,^[8] mortality was defined as mortality at discharge from the same hospital as the operation took place. In the original ES study, the mortality was defined as the operative mortality within the 30 days after surgery regardless of location. The 30-day mortality is probably more relevant and limits bias provided by potential missing data as for 90-day mortality but is usually slightly higher than the operative mortality.^[8] Finally, as it is recognise d by one of the main designers of the ES,^[29] score discrimination invariably is the lowest in subgroups

The ES2 is well calibrated in a low-risk group and the whole cardiac population, and ES2 can be used as a benchmark for quality control. ES2 can't be applied to individual patients but it may provide the patient some objective data to made decision. However, clinicians must be aware that the ES2 underestimate mortality in high-risk groups.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

- Arangalage D, Cimadevilla C, Alkhoder S, Chiampan A, Himbert D, Brochet E, et al. Agreement between the new EuroSCORE II, the logistic EuroSCORE and the society of thoracic surgeons score: Implications for transcatheter aortic valve implantation. Arch Cardiovasc Dis 2014;107:353-60.
- Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9-13.
- Roques F, Nashef SA, Michel P, Gauducheau E, de Vincentiis C, Baudet E, et al. Risk factors and outcome in European cardiac surgery: Analysis of the EuroSCORE multinational database of 19030 patients. Eur J Cardiothorac Surg 1999;15:816-22.
- Engebretsen KV, Friis C, Sandvik L, Tønnessen T. Survival after CABG-better than predicted by EuroSCORE and equal to the general population. Scand Cardiovasc J 2009;43:123-8.
- Gummert JF, Funkat A, Osswald B, Beckmann A, Schiller W, Krian A, et al. EuroSCORE overestimates the risk of cardiac surgery: Results from the national registry of the German society of thoracic and cardiovascular surgery. Clin Res Cardiol 2009;98:363-9.
- Nissinen J, Biancari F, Wistbacka JO, Loponen P, Teittinen K, Tarkiainen P, et al. Is it possible to improve the accuracy of EuroSCORE. Eur J Cardiothorac Surg 2009;36:799-804.
- Kalavrouziotis D, Li D, Buth KJ, Légaré JF. The European system for cardiac operative risk evaluation (EuroSCORE) is not appropriate for withholding surgery in high-risk patients with aortic stenosis: A retrospective cohort study. J Cardiothorac Surg 2009;4:32.
- Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. Eur J Cardiothorac Surg 2012;41:734-44.
- Michel P, Roques F, Nashef SA, EuroSCORE Project Group. Logistic or additive Euro SCORE for high-risk patients? Eur J Cardiothorac Surg 2003;23:684-7.
- Gogbashian A, Sedrakyan A, Treasure, T. EuroSCORE: A systematic review of international performance. Eur J Cardiothorac Surg 2004;25:695-700.
- Shanmugam G, West M, Berg G. Additive and logistic EuroSCORE performance in high-risk patients. Interact Cardiovasc Thorac Surg 2005;4:299-303.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 1988;44:837-45.
- 13. Howell NJ, Head SJ, Freemantle N, van der Meulen TA, Senanayake E, Menon A, *et al.* The new EuroSCORE II does not improve prediction of mortality in high-risk patients undergoing cardiac surgery: A collaborative analysis of two European centres. Eur J Cardiothorac Surg 2013;44:1006-11.
- Ranucci M, Di Dedda U, Castelvecchio S, La Rovere MT, Menicanti L. In search of the ideal risk-scoring system for very high-risk cardiac surgical patients: A two-stage approach. J Cardiothorac Surg 2016;11:13.
- 15. Barili F, Pacini D, Capo A, Ardemagni E, Pellicciari G, Zanobini M, et al. Reliability of new scores in predicting perioperative mortality after isolated aortic valve surgery: A comparison with the society of thoracic surgeons score and logistic EuroSCORE. Ann Thorac Surg 2013;95:1539-44.
- Di Dedda U, Pelissero G, Agnelli B, De Vincentiis C, Castelvecchio S, Ranucci M. Accuracy, calibration and clinical performance of the new Euro SCORE II risk stratification system. Eur J Cardiothorac Surg 2013;43:27-32.
- Paparella D, Guida P, Di Eusanio G, Caparrotti S, Gregorini R, Cassese M, *et al.* Risk stratification for in-hospital mortality after cardiac surgery: External validation of EuroSCORE II in a prospective regional registry. Eur J Cardiothorac Surg 2014;46:840-8.
- 18. Guida P, Mastro F, Scrascia G, Whitlock R, Paparella D. Performance

of the European system for cardiac operative risk evaluation II: A meta-analysis of 22 studies involving 145,592 cardiac surgery procedures. J Thorac Cardiovasc Surg 2014;148:3049-57.

- Garcia-Valentin A, Mestres CA, Bernabeu E, Bahamonde JA, Martín I, Rueda C, *et al.* Validation and quality measurements for EuroSCORE and EuroSCORE II in the Spanish cardiac surgical population: A prospective, multicentre study. Eur J Cardiothorac Surg 2016;49:399-405.
- Kieser TM, Rose MS, Head SJ. Comparison of logistic EuroSCORE and EuroSCORE II in predicting operative mortality of 1125 total arterial operations. Eur J Cardiothorac Surg 2016;50:509-18.
- Ranucci M, Castelvecchio S, Menicanti LA, Scolletta S, Biagioli B, Giomarelli P. An adjusted EuroSCORE model for high-risk cardiac patients. Eur J Cardiothorac Surg 2009;36:791-7.
- Ghazy T, Kappert U, Ouda A, Conen D, Matschke K. A question of clinical reliability: Observed versus EuroSCORE-predicted mortality after aortic valve replacement. J Heart Valve Dis 2010;19:16-20.
- 23. Kawachi Y, Nakashima A, Toshima Y, Arinaga K, Kawano H. Risk stratification analysis of operative mortality in heart and thoracic aorta surgery: Comparison between Parsonnet and EuroSCORE additive model. Eur J Cardiothorac Surg 2001;20:961-6.

- Zingone B, Pappalardo A, Dreas L. Logistic versus additive EuroSCORE. A comparative assessment of the two models in an independent population sample. Eur J Cardiothorac Surg 2004;26:1134-40.
- Leontyev S, Walther T, Borger MA, Lehmann S, Funkat AK, Rastan A, et al. Aortic valve replacement in octogenarians: Utility of risk stratification with EuroSCORE. Ann Thorac Surg 2009;87:1440-5.
- Karthik S, Srinivasan AK, Grayson AD, Jackson M, Sharpe DA, Keenan DJ, et al. Limitations of additive EuroSCORE for measuring risk stratified mortality in combined coronary and valve surgery. Eur J Cardiothorac Surg 2004;26:318-22.
- 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;89:787-93.
- 28. Ge Y, Sun L, Zhu J, Liu Y, Cheng L, Chen L, et al. Can EuroSCORE II predict the mortality and length of intensive care unit stay after total aortic arch replacement with stented elephant trunk implantation for DeBakey type I aortic dissection? J Thorac Cardiovasc Surg 2013;61:564-8.
- Nashef SA, Sharples LD. Editorial comment: Pride without prejudice: EuroSCORE II, the STS score and the high-risk patient subset. Eur J Cardiothorac Surg 2013;44:1012.