ORIGINAL ARTICLE

EMC

Rev Bras Cir Cardiovasc 2011; 26.1: 1-6

The 2000 Bernstein-Parsonnet score and EuroSCORE were similar in predicting mortality at the Heart Institute, USP

Os escores 2000 Bernstein-Parsonnet e EuroSCORE foram similares na predição da mortalidade no Instituto do Coração-USP

Omar Asdrúbal Vilca MEJÍA[1], Luiz A Ferreira LISBOA[2], Luiz Boro PUIG[3], Ricardo Ribeiro DIAS[4], Luís A. DALLAN[5], Pablo M. POMERANTZEFF[6], Noedir A.G. STOLF[7]

RBCCV 44205-1238

Abstract

Objective: **To evaluate the performance of 2000 Bernstein-Parsonnet (2000 BP) and additive EuroSCORE (ES) to predict surgical mortality at the Heart Institute from the University of São Paulo.**

Methods: **This is a prospective observational study. Seven hundred and seventy four patients underwent coronary artery bypass graft, valve, or a combined procedure between May and October, 2007. The mortality rate was estimated with the 2000 BP and ES. The correlation between both expected and observed mortality rates was validated through the calibration and discrimination test.**

Results: **The patients were stratified into five groups for the 2000 BP and three groups for the ES. The Hosmer-Lemeshow test for 2000 BP (*P* = 0.70) and for ES (*P* = 0.39) indicates a proper calibration. The ROC curve for the 2000 BP = 0.84 and for the ES = 0.81 confirms that the models are good predictors (*P*<0.001).**

Conclusion: **Both models are similar and adequate in predicting surgical mortality at the Heart Institute from the University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP).**

Descriptors: **Risk Factors. Cardiac Surgical Procedures. Hospital Mortality.**

Resumo

Objetivo: **Avaliar o desempenho dos escores 2000 Bernstein-Parsonnet (2000 BP) e EuroSCORE aditivo (ES) na predição de mortalidade cirúrgica no Instituto do Coração da Universidade de São Paulo (InCor-USP).**

Métodos: **Desenho prospectivo e observacional. Setecentos e quarenta e quatro pacientes consecutivos submetidos à cirurgia de revascularização miocárdica, valvar ou associada entre maio e outubro de 2007. A mortalidade foi calculada com os escores 2000 BP e ES. A correlação entre mortalidade estimada e mortalidade observada foi validada mediante testes de calibração e discriminação.**

Resultados: **Os pacientes foram estratificados em cinco grupos para o 2000 BP e três para o ES. O teste de Hosmer Lemeshow para o 2000 BP (*P* = 0,70) e para o ES (*P* = 0,39) indica uma boa calibração. A curva ROC para o 2000 BP = 0,84 e para o ES = 0,81 confirma que os modelos são bons preditores (*P* < 0,001).**

Conclusão: **Ambos os modelos são similares e adequados na predição de mortalidade cirúrgica no InCor-USP.**

Descritores: **Fatores de Risco. Procedimentos Cirúrgicos Cardíacos. Mortalidade Hospitalar.**

1. Cardiovascular Surgeon (Sciences Doctoral Student; Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP.)
2. Collaborator Professor – USP - (Assistant Physician; Surgical Unit of Heart Disease; Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP.)
3. Associate Professor - USP (Assitant Physician; Surgical Unit of Coronary heart diseases; Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP.)
4. Collaborator Professor - USP. (Assistant Physician; Surgical Unit of General Heart Diseases; Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP.)
5. Associate Professor - USP (Head Physician; Surgical Unit of Coronary Heart Diseases; Heart Institute; University of São Paulo School of Medicine Clinics Hospital(InCor-HCFMUSP.)
6. Associate Professor - USP (Head Physician; Surgical Unit of Heart Valve Diseases; Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP.)
7. Full Professor; Discipline of Cardiopneumology - USP (Division Director of Cardiovascular and Thoracic Surgery; Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP.)

This study was carried out at the Heart Institute; University of São Paulo School of Medicine Clinics Hospital (InCor-HCFMUSP).

Corresponding Author: Omar A. Vilca Mejía
Avenida Dr. Enéas de Carvalho Aguiar, 44, São Paulo, SP, Brasil. CEP 05403-000.
E-mail: omarvilca@incor.usp.br

Article received on July 12nd, 2010
Article accepted on December 20th, 2010

INTRODUCTION

Risk scores can make a real comparison of surgical outcomes, which allows a more objective evaluation of surgical indication.

One of the first studies to analyze the risk factors was performed by Parsonnet [1] in 1989, but included only patients undergoing coronary artery bypass grafting (CABG). Eleven years later, an analysis of 10,703 patients from 10 centers in New Jersey (USA) gave rise to the 2000 Bernstein-Parsonnet [2]. This new model, consisting of 47 variables could be used for patients undergoing coronary, valve, and associated surgeries.

Nevertheless, the most internationally used score is the EuroSCORE [3], which includes 17 risk factors, from 19,030 patients in 128 centers in Europe. The EuroSCORE has been shown to be effective, even when applied to non-European populations [4-7].

As the risk profile of patients undergoing cardiac surgery in Brazil remains unknown, it is necessary to apply these scores in order to analyze our patients and to stratify them according to the morbidity and complexity of the procedure.

The aim of this study was to analyze the performance of the 2000 Bernstein-Parsonnet (2000 BP) and the additive EuroSCORE (ES) to predict the operative mortality rate in the Heart Institute of the University of Sao Paulo Clinics Hospital (InCor-USP.)

METHODS

This study was approved by the Ethics Committee for Analysis of Research Projects (Comissão de Ética para Análises de Projetos de Pesquisa-CAPPesq) of the University of São Paulo Clinics Hospital under the number 1575, and the need for informed consent was waived in view of the observational and anonymous nature of the study.

This prospective study was conducted at the Division of Thoracic and Cardiovascular Surgery of the Heart Institute - University of Sao Paulo (InCor-USP). Due to the sample size, logistic regression studies made by Long [8], by bias simulation studies, conclude that to use samples < 100 patients are hardly unreliable, once the advisable is to use samples > 500 patients. The study group consisted of 744 patients operated on consecutively either in urgency or emergency elective modalities from May to October 2007. Inclusion criteria were: a) Heart valve surgery (replacement or repair), b) Coronary artery bypass grafting with and without cardiopulmonary bypass (including cases of left ventricular aneurysmectomy), c) Associated surgery (including one of the above-mentioned). Exclusion criteria included those patients undergoing surgical correction of congenital anomalies, heart transplantation, aortic diseases, cardiac tumors, and diseases of the pericardium.

No patient was excluded due to lack of information. Patient data were collected preoperatively from their own electronic medical records, and stored in a single spreadsheet. This spreadsheet has been adapted in such a way that all the variables in the 2000 BP and ES were included. All the definitions given to the variables by both scores along with their respective values were strictly respected, keeping the specific values (beta coefficient) for each risk factor according to their relevance with the death event.

Thus, after calculating the value of the 2000 BP and ES for each study patient, these patients were systematically arranged according to the risk groups established by the scores. All these data were stored in a database developed in Excel™.

The primary outcome was in-hospital mortality, defined as death occurring in the time frame between both patient surgery and discharge.

Statistical Analysis

To evaluate the performance of the 2000 Bernstein-Parsonnet (2000 BP) and additive EuroSCORE (ES) in predicting surgical mortality in the study group, the validation of the models was performed by calibration and discrimination tests.

Calibration evaluates the model in its ability to predict both the overall mortality and the different risk groups using the Hosmer-Lemeshow test. We used the logistic regression model where the dependent variable is mortality yes/no and the explanatory variable is the two scores (one model for each score). This way, a logistic regression model was set for each score, and this model was tested regarding its suitability.

On the other hand, the discrimination assesses the score effectiveness to predict which patients live and which patients die. This is represented by the area under the ROC curve. Thus, to a larger area under the ROC curve there is a better discriminatory strength of the model.

Statistical analysis was performed using SPSS version 15.0 for Windows where the continuous variables were expressed as mean ± standard deviation and categorical variables as percentages. The *P* value <0.05 was considered significant.

RESULTS

Table I show that the hospital mortality was 6.7% (50/744 patients). Of all procedures, 53.5% were coronary artery bypass surgery, 27.3% heart valve surgery, and 19.22% were associated surgeries. The patients' age ranged from 13 to 89 years (mean 59.3 ± 14.4), and 36% were female.

Table 2 shows the prevalence of risk factors in the study group, in both 2000 BP and ES scores. We can observe that there is a higher prevalence of neurological dysfunction, reoperation, creatinine > 2.3, active endocarditis, critical

preoperative status, recent myocardial infarction, and pulmonary hypertension in our sample. However, these variables were not statistically significant.

Table 3 presents the observed and estimated mortality using the 2000 BP score as the predictor variable in the groups defined by the Hosmer-Lemeshow test. Likewise, the ES as the predictor variable is presented in Table 4.

In Figure 1, evaluating the discrimination strength of the models, we observed that the area under the ROC curve for the BP 2000 score was 0.84 and for the ES score it was of 0.81 (P <0.001). Therefore, the cutoff point given by the ROC curve for these scores would be:

* For the 2000 BP e" 27, the chance of death is 11.6 times higher than when the 2000 BP score is < 27 (95% CI = 5.7 to 23.7; S = 80%; E = 74.4%; PPV = 18.4%; NPV = 98.1%; A = 74.7%; P <0.001).

PPV = Positive Predictive Value

NPV = Negative Predictive value

* For the ES ≥ 6.5, the chance of death is 11.8 times higher than when is S < 6.5 (95% CI = 5.9 to 23.6; S = 78.0%; E = 77.0 %; PPV = 19.6%; NPV = 98.0%; A = 77%; *P* <0.001).

Table 1. Mortality due to the type of surgical procedure.

Procedure	Nº of Death	Nº of Death	Percentage
CABG			
CABG	398	15	3.8%
CABG (OFF-pump)	136	3	2.2%
CABG (On-pump)	262	12	4.6%
CABG (without Reop)	373	11	2.9%
Reop CABG	34	4	11.8%
Valve			
AoV	96	13	13.5%
MiV	107	9	8.4%
Doble valve	47	4	8.5%
Triple valve	5	2 (Reop)	40%
AoV (without Reop)	62	5	8.1%
MiV (without Reop)	59	1	1.7%
Valve reop	113	19	16.8%
Associated			
CABG + AoV	33	1	3%
CABG + MiV	17	0	0%
CABG + Double valve	2	0	0%
IVC post-AMI	2	0	0%
Others	39	6	15.4%

CABG= Coronary artery bypass grafting; Reop= Reoperation; Off-pump CABG = without cardiopulmonar bypass (extracoporeal cirulation) support; On-pump CABG – with cardiopulmonar bypass (extracorporeal circulation) support; AoV= Aortic Valve; MiV= Mitral valve; IVC post-AMI = interventricular communication post-acute myocardial infarction.

Table 2. Risk Factors Prevalence on the Study Group in comparison with the population risk factor from EuroScore and 2000 Bernstein Parsonnet.

Risk factor	Study (N=744)	EuroScore (N=19,030)	2000 BP (N=10,703)
Age	59.53	62.5	NC
Mean	44.8%	33.2%	NC
< 60	15.2%	17.8%	NC
60-64	15.3%	20.7%	NC
65-69	12.1%	17.9%	18.5%
70-74	12.6%	9.6%	13.7%
≥75	30.0%	27.8%	31.3%
Female	2.6%	3.9%	10.8%
Chronic Pulmonary Disease	6.6%	11.3%	9.1%
Extracardiac Arteriopathy	8.3%	1.4%	8.4%
Neurological disorder	19.8%	7.3%	7.6%
Prior heart surgery	4.3%	1.8%	4.5%
Creatinine > 2,3 mg/dl	3.0%	1.0%	NC
Active endocarditis	16.1%	4.1%	NC
Critical preoperative status (condition)	4.6%	8.0%	NC
Unstable Angina	29.8%	25.6%	38.6%
EF < 30	5.5%	5.8%	8.4%
Recent AMI	17.2%	9.7%	NC
Pulmonary hypertension	9.8%	2.0%	10.7%
Emergency	3.6%	4.9%	NC
Associated surgery	18.6%	36.4%	NC
Thoracic aorta surgery	1.9%	2.4%	NC
Post-infarction IVC	0.3%	0.2%	NC

EE: Ejection fraction; AMI: Acute myocardial infarction; IVV: Interventricular communication; NC: not clarified by 2000 BP score

Table 3 Estimate and observed mortality percentages per 2000 BP risk group.

Risk	OM/EM	EM	OM	N	%
0-8.5	0.60	1.17	0.70	143	19.2
9-15	0.65	1.97	1.28	156	21.0
15.5-22	0.60	3.36	2.01	149	20.0
22.5-31.5	1.28	6.42	8.22	146	19.6
≥32	1.04	20.58	21.33	150	20.2

OM/EM= Observed Mortality/Expected Mortality; EM= Expected Mortality; OM= Observed Mortality; N= Number of patients; %= Percentage

Table 4. Estimate and Observed mortality percentages per ES risk group.

Risk	OM/EM	EM	OM	N	%
0-2	0.87	1.57	1.36	221	29.7
3-5	0.77	3.41	2.64	265	35.6
≥6	1.06	14.52	15.50	258	34.6

OM/EM= Observed Mortality/Expected Mortality; EM= Expected Mortality; OM= Observed Mortality; N= Number of patients; %= Percentage

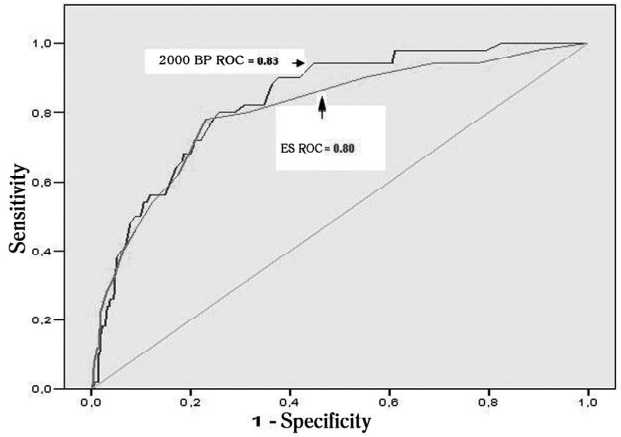

*Fig. 1 – ROC Curve - *2000 BP ROC = ROC Curve given by the 2000 Bernstein Parsonnet; ES ROC= ROC Curve given by the Euroscore*

DISCUSSION

As a part of evidence-based medicine, risk scores should be widely used in daily practice, although they need to be validated.

In this study, we have chosen the 2000 BP score because it is a simplified model, it is easy to use at the bedside, impartial and objective in predicting postoperative complications and remains at the intensive care unit [9,10]. The choice for the EuroSCORE was due to its widespread use in the literature and the good correlation in different populations in Europe, North America, and Japan [11,12].

The publications with the greatest impact on the mortality regarding cardiac surgery in Brazil have shown increased mortality rate. However, it does not characterize the clinical profile of patients, thus limiting their interpretations [13,14]. The EuroSCORE was used successfully. It was retrospectively and posteriorly used in both CABG [4,15] and heart valve surgeries [16], respectively, at the Heart Institute of Pernambuco. There are studies showing a better performance of the EuroSCORE model in relation to the original Parsonnet model score [5-7,12]. These studies describing 20 years of existence may have weakened the predictive strength of the model. This way, the article by Dr. Berman [17] is perhaps the most important article concerning the proper statistical analysis to compare the 2000 BP with the EuroSCORE. Thus, similar results can be obtained with good accuracy.

Comparing the characteristics of the study population in relation to the populations of origin of both the EuroSCORE and 2000 BPscores, we can observe (Table 1) that there is a higher prevalence of neurological dysfunction, reoperation, creatinine > 2.3, active endocarditis, critical preoperative state, recent myocardial infarction, and pulmonary hypertension, which would explain the differences. Nevertheless, these differences were not significant between both expected and observed mortality in high-risk patients. Comparing the characteristics of the sample population compared to populations of origin of the EuroSCORE and 2000 BP, we can observe (Table 1), there is a higher prevalence of neurological dysfunction, reoperation, creatinine> 2.3, active endocarditis, preoperative status critical, recent infarction and pulmonary hypertension, which would explain the differences, but no significant mortality between expected and observed mortality in high risk patients. However, significant differences found in patients undergoing CABG in the state of Rio de Janeiro hindered the discriminative strength of EuroSCORE [18].

A remark made regarding the analysis of this study and other studies applied in different populations [4-7,11,15-17] suggests that it is not the type of the patient as a result of his/her ethnicity that defined the surgical outcomes, but the patient as a result of the control of risk factors. For that reason, we should not continue insisting on the premise that a different ethnic group would not validate the applicability of these models [19,20].

There is a growing tendency to report that the EuroSCORE is overestimating the mortality in several of populations [12,20]. Most records are from spontaneous records without inclusion of all patients and often without the adhesion of the exact definitions of variables [21]. However, when we show the totality of the patients consecutively operated on, strictly respecting the definitions and limitations of the scores, we are able to reproduce its similar efficacy to that found by other authors [4-7,17].

The limitations of our study were as follows:

1) Although it is a prospective study with patients from different regions of Brazil, data were collected from a single tertiary center.

2) It has not been determined the cause of death, even knowing that death from cardiac causes would better reflect the surgical outcome.

3) In order to validate a score of a sample, it would be necessary at least a cohort of 100 deaths [22]. However, the initial analysis allows us an objective assessment of how the behavior of the scores is.

4) The additive EuroSCORE because of its nature tends to underestimate the risk in high-risk patients [23]. Although this has not been demonstrated in multicenter studies [24-26].

5) However, it is questionable to include the surgeries of the thoracic aorta in this analysis. We can justify saying that in the origin of the models used, the aorta surgeries were the minority. Thus, the extrapolation to the aorta would be unnatural. Furthermore, the validation of these scores in the surgeries of the thoracic aorta has demonstrated little effectiveness [5,27,28].

CONCLUSION

The similarity between both observed and expected mortality by the ES and 2000 BP scores allows us to confirm that the values given by the scores to the various risk factors can be applied to our patients. Therefore, we conclude that both models are similar and adequate in predicting the mortality of patients undergoing CABG, heart valve, and associated surgeries at the Heart Institute of University of São Paulo.

REFERENCES

1. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation. 1989;79(6 Pt 2):I3-12.

2. Bernstein AD, Parsonnet V. Bedside estimation of risk as an aid for decision-making in cardiac surgery. Ann Thorac Surg. 2000;69(3):823-8.

3. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salomon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg. 1999;16(1):9-13.

4. Moraes F, Duarte C, Cardoso E, Tenório E, Pereira V, Lampreia D, et al. Avaliação do EuroSCORE como preditor de mortalidade em cirurgia de revascularização miocárdica no Instituto do Coração de Pernambuco. Rev Bras Cir Cardiovasc. 2006;21(1):29-34.

5. Kawachi Y, Nakashima A, Toshima Y, Arinaga K, Kawano H. Risk stratification analysis of operative mortality in heart and thoracic aortic surgery: comparison between Parsonnet and EuroSCORE additive model. Eur J Cardiothorac Surg. 2001;20(5):961-6.

6. Au WK, Sun MP, Lam KT, Cheng LC, Chiu SW, Das SR. Mortality prediction in adult cardiac surgery patients: comparison of two risk stratification models. Hong Kong Med J. 2007;13(4):293-7.

7. Syed AU, Fawzy H, Farag A, Nemlander A. Predictive value of EuroSCORE and Parsonnet scoring in Saudi population. Heart Lung and Circulation. 2004;13(4):384-8.

8. Nemes S, Jonasson JM, Genell A, Steineck G. Bias en odds ratios by logistic regression modelling and sample size. BMC Med Res Methodol. 2009;9(1):56.

9. Lawrence DR, Valencia O, Smith EE, Murday A, Treasure T. Parsonnet score is a good predictor of the duration of intensive care unit stay following cardiac surgery. Heart. 2000;83(4):429-32.

10. Doering LV, Esmailian F, Laks H. Perioperative predictors of ICU and hospital costs in coronary artery bypass graft surgery. Chest 2000;118(3):736-43.

11. Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. Eur J Cardiothorac Surg. 2002;22(1):101-5.

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

13. Ribeiro AL, Gagliardi SP, Nogueira JL, Silveira LM, Colosimo EA, Nascimento CAL. Mortality related to cardiac surgery in Brazil, 2000-2003. J Thorac Cardiovasc Surg. 2006;131(4):907-9.

14. David TE. Should cardiac surgery be performed in low-volume hospitals? J Thorac Cardiovasc Surg. 2006;131(4):773-4.

15. Campagnucci VP, Pinto e Silva AMR, Pereira WL, Chamlian EG, Gandra SM, Rivetti LA. EuroSCORE and the patients undergoing coronary bypass surgery at Santa Casa de São Paulo. Rev Bras Cir Cardiovasc. 2008;23(2):262-7.

16. Andrade ING, MoraesNeto FR, Oliveira JPSP, Silva ITC, Andrade TG, Moraes CRR, et al. Avaliação do EuroSCORE como preditor de mortalidade em cirurgia cardíaca valvar no Instituto do Coração de Pernambuco. Rev Bras Cir Cardiovasc.2010;25(1):11-8.

17. Berman M, Stamler A, Sahar G, Georghiou GP, Sharoni E, Brauner R, et al. Validation of the 2000 Bernstein-Parsonnet score versus the EuroSCORE as a prognostic tool in cardiac surgery. Ann Thorac Surg. 2006;81(2):537-40.

18. Carvalho MRM, Souza e Silva NA, Klein CH, Oliveira GMM. Aplicação do EuroSCORE na cirurgia de revascularização miocárdica em hospitais públicos do Rio de Janeiro. Rev Bras Cir Cardiovasc. 2010;25(2):209-17.

19. D'Errigo P, Seccareccia F, Rosato S, Manno V, Badoni G, Fusco D, et al; Research Group of the Italian CABG Outcome Project. Comparison between an empirically derived model and the EuroSCORE system in the evaluation of hospital performance: the example of the Italian CABG Outcome Project. Eur J Cardiothorac Surg. 2008;33(3):325-33.

20. Zheng Z, Li Y, Zhang S, Hu S; Chinese CABG Registry Study. The Chinese coronary artery bypass grafting registry study: how well does the EuroSCORE predict operative risk for Chinese population? Eur J Cardiothorac Surg. 2009;35(1):54-8.

21. Menicanti LA. Editorial: The surgeon, the statistics and the data. Eur J Cardiothorac Surg. 2008;33(3):323-4.

22. Harrel FE. Regression models strategies. Berlin: Springer-Verlag; 2001.

23. Roques F, Michel P, Goldstone AR, Nashef SA. The logistic Euroscore. Eur Heart J. 2003;24(9):881-2.

24. Jin R, Grunkemeier GL; Providence Health System Cardiovascular Study Group. Does the logistic EuroSCORE offer an advantage over the additive model? Interact Cardiovasc Thorac Surg. 2006;5(1):15-7.

25. Gogbashian A, Sedrakyan A, Treasure T. EuroSCORE: a systematic review of international performance. Eur J Cardiothorac Surg. 2004;25(5):695-700.

26. Nilsson J, Algotsson L, Höglund P, Lührs C, Brandt J. Comparison of 19 pre-operative risk stratification models in open-heart surgery. Eur Heart J. 2006;27(7):867-74.

27. Matsuuara K, Ogino H, Matsuda H, Minatoya K, Sasaki H, Yagihara T, Kitamura S. Limitations of EuroSCORE for measurement of risk-stratified mortality in aortic arch surgery using selective cerebral perfusion: is advanced age no longer a risk? Ann Thorac Surg. 2006;81(6):2084-7.

28. Barmettler H, Immer FF, Berdat PA, Eckstein FS, Kipfer B, Carrel TP. Risk-stratification in thoracic aortic surgery: should the EuroSCORE be modified? Eur J Cardiothorac Surg. 2004;25(5):691-4.