

# EuroSCORE II and the importance of a local model, InsCor and the future SP-SCORE

*EuroSCORE II e a importância de um modelo local, InsCor e o futuro SP-SCORE*

Luiz Augusto Ferreira Lisboa<sup>1</sup>, MD, PhD; Omar Asdrubal Vilca Mejia<sup>1</sup>, MD, PhD; Luiz Felipe Pinho Moreira<sup>1</sup>, MD, PhD; Luís Alberto Oliveira Dallan<sup>1</sup>, MD, PhD; Pablo Maria Alberto Pomerantzeff<sup>1</sup>, MD, PhD; Luís Roberto Palma Dallan<sup>1</sup>, MD; Maria Raquel B. Massoti<sup>1</sup>, MD; Fabio B. Jatene<sup>1</sup>, MD, PhD

DOI: 10.5935/1678-9741.20140004

RBCCV 44205-1514

## Abstract

**Introduction:** The most widely used model for predicting mortality in cardiac surgery was recently remodeled, but the doubts regarding its methodology and development have been reported.

**Objective:** The aim of this study was to assess the performance of the EuroSCORE II to predict mortality in patients undergoing coronary artery bypass grafts or valve surgery at our institution.

**Methods:** One thousand consecutive patients operated on coronary artery bypass grafts or valve surgery, between October 2008 and July 2009, were analyzed. The outcome of interest was in-hospital mortality. Calibration was performed by correlation between observed and expected mortality by Hosmer Lemeshow. Discrimination was calculated by the area under the ROC curve. The performance of the EuroSCORE II was compared with the EuroSCORE and InsCor (local model).

**Results:** In calibration, the Hosmer Lemeshow test was inappropriate for the EuroSCORE II ( $P=0.0003$ ) and good for the EuroSCORE ( $P=0.593$ ) and InsCor ( $P=0.184$ ). However, the discrimination, the area under the ROC curve for EuroSCORE II was 0.81 [95% CI (0.76 to 0.85),  $P<0.001$ ], for the EuroSCORE was 0.81 [95% CI (0.77 to 0.86),  $P<0.001$ ] and for InsCor was 0.79 [95% CI (0.74-0.83),  $P<0.001$ ] showing up properly for all.

**Conclusion:** The EuroSCORE II became more complex and resemblance to the international literature poorly calibrated to predict mortality in patients undergoing coronary artery bypass grafts or valve surgery at our institution. These data emphasize the importance of the local model.

**Descriptors:** Risk Factors. Cardiovascular Surgical Procedures. Coronary Artery Bypass. Myocardial Revascularization. Coronary Disease. Heart Valve Diseases.

1. Heart Institute of the Clinics Hospital at the Faculty of Medicine, University of São Paulo (InCor-HCFMUSP), São Paulo, SP, Brazil.

Correspondence address:

Luiz Augusto Ferreira Lisboa  
Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo  
Av. Dr. Enéas de Carvalho Aguiar, 44 – 2º andar – sala 11 – Cerqueira César – São Paulo, SP, Brazil – Zip code: 05403-000  
E-mail: luiz.lisboa@incor.usp.br

This study was carried out at the Clinics Hospital at the Faculty of Medicine, University of São Paulo (InCor-HCFMUSP), São Paulo, SP, Brazil.

No financial support.

Article received on October 12<sup>th</sup>, 2013  
Article accepted on November 17<sup>th</sup>, 2013

#### Abbreviations, acronyms and symbols

EuroSCORE	European System for Cardiac Operative Risk Evaluation
InCor-HCFMUSP	Clinics Hospital at the Faculty of Medicine, University of São Paulo
ROC	Receiver Operating Characteristic
SPSS	Statistical Package for the Social Sciences
STS score	Society of Thoracic Surgeons score

#### Resumo

**Introdução:** O modelo mais utilizado para predição de mortalidade em cirurgia cardíaca foi recentemente remodelado, mas dúvidas referentes à sua metodologia e desenvolvimento têm sido relatadas.

**Objetivo:** O objetivo deste estudo foi avaliar o desempenho do EuroSCORE II na predição de mortalidade em pacientes submetidos a cirurgia de coronária e/ou valva na instituição.

**Métodos:** Mil pacientes, operados consecutivamente de coronária e/ou valva, entre outubro de 2008 e julho de 2009, foram analisados. O desfecho de interesse foi mortalidade in-

tra-hospitalar. A calibração foi realizada pela correlação entre mortalidade esperada e observada por meio do teste de Hosmer Lemeshow. A discriminação foi calculada pela área abaixo da curva ROC. O desempenho do EuroSCORE II foi comparado com os modelos EuroSCORE e InsCor (modelo local).

**Resultados:** Na calibração, o teste de Hosmer Lemeshow foi inadequado para o EuroSCORE II ( $P=0,0003$ ) e bom para os modelos EuroSCORE ( $P=0,593$ ) e InsCor ( $P=0,184$ ). No entanto, na discriminação, a área abaixo da curva ROC para o EuroSCORE II foi de 0,81 [IC 95% (0,76-0,85),  $P<0,001$ ]; para o EuroSCORE foi de 0,81 [IC 95% (0,77-0,86),  $P<0,001$ ] e para o InsCor foi de 0,79 [IC 95% (0,74-0,83),  $P<0,001$ ], revelando-se adequada para todos.

**Conclusão:** O EuroSCORE II se tornou mais complexo e, à semelhança com a literatura internacional, mal calibrado para prever mortalidade nos pacientes operados de coronária e/ou valva em nosso meio. Esses dados reforçam a importância do modelo local.

**Descritores:** Fatores de Risco. Procedimentos Cirúrgicos Cardiovasculares. Ponte de Artéria Coronária. Revascularização Miocárdica. Doença das Coronárias. Doenças das Valvas Cardíacas.

## INTRODUCTION

In modern medicine, the use of risk scores as predictors of cardiovascular events is well established [1]. Efficient models should be derived from prospective, compulsory and complete records, be built upon bootstrap statistical techniques and demonstrate adequate internal validation, strictly following the scientific principles [2,3]. Clearly risk models derived and validated on a local, usually have lower performance when applied elsewhere and even in the same location over time [4]. However, the first EuroSCORE created in 1999 [5] in the European population, was suitable in a contemporary Brazilian population [6-8].

Undoubtedly, the incorporation of the EuroSCORE on key services in Europe brought to mind the “Hawthorne” effect, explaining that nothing much has improved outcomes in cardiac surgery at the beginning of the century, as monitoring by EuroSCORE [9]. Over time, the remodeling of the EuroSCORE for countries that joined its mandatory use would be justifiable. Thus, the EuroSCORE II has aroused [10], from a record with 22,381 consecutive patients undergoing cardiac surgery in 154 hospitals in 43 countries (inside and outside Europe), over a 12-week period (May to July 2010).

This updated model has more variables than the first EuroSCORE, so in addition to the risk of having high discrimination power, it carries the risk of overfitting [11]. Thus, smaller models have good accuracy but unfortunately decrease the

power of discrimination. Still, we must not forget that “few variables as possible” prevails in a model in order to have a greater acceptance [12,13]. At the Heart Institute, Clinics Hospital of the Faculty of Medicine, University of São Paulo (Incor-HCFMUSP), the remodeling of EuroSCORE models and 2000 Bernstein-Parsonnet [8] together, using the bootstrap technique, gave rise to InsCor [14]. This model was similar to the first EuroSCORE and its performance was simpler than this and that the 2000 Bernstein-Parsonnet score to predict mortality in patients undergoing coronary and/or valve at Incor-HCFMUSP. This fact becomes more important when there is a need to assess the experience of treatment against the “casemix” location at a given time, as it has been done by several groups. The aim of this study was to validate the EuroSCORE II and compare it to InsCor and EuroSCORE models in patients undergoing coronary and/or valve on Incor-HCFMUSP.

## METHODS

### Sample size, inclusion and exclusion criteria

A retrospective analysis of prospectively collected data was performed at the Division of Cardiovascular Surgery, Incor - HCFMUSP. For validation of risk scores in a sample of at least 100 deaths, the study by Lisboa et al. [15] on the results of cardiovascular surgery at Incor-HCFMUSP of the past 23 years, was the basis for the study. For this, 1000 pa-

tients operated sequentially for coronary bypass or associated and/or isolated or combined valve surgery, including reoperations and in elective, urgent or emergency procedures, from October 2008 to July 2009, were selected. Of these, all filled the variables contained in the InsCor EuroSCORE models, however, only 900 patients included all variables required by Euroscore II. Patients younger than 18 years or undergoing other types of surgery other than CABG and/or valve surgery were excluded from the study.

**Collection, definition and organization of data**

Data were collected from electronic medical records system of the Incor (SI3) and stored in spreadsheets. Each worksheet has been adapted to take account of all the variables, respecting their definitions as described by EuroSCORE [9], EuroSCORE II [10] and InsCor [14] models. Patients were sorted according to the risk groups established by the scores and placed in the database made in Excel. The outcome of interest was in-hospital mortality, defined as death that occurred in the time interval between surgery and discharge.

**Validation of InsCor, EuroSCORE and EuroSCORE II**

To assess the performance of InsCor, EuroSCORE and EuroSCORE II in predicting mortality, the predictive validity of the models was performed. The analysis was performed using calibration and discrimination test. Calibration assesses the accuracy of the model to predict risk in a group of patients. The force calibration was assessed by testing the goodness of fit by the Hosmer-Lemeshow test. *P* value > 0.05 indicates that the model fits the data and predicts mortality properly. Discrimination measures the ability of the model to distinguish between patients at low and high risk. Discrimination was measured by use of the statistical technique

called area under the ROC (Receiver Operating Characteristic, sometimes called c-statistic or c-index).

**Statistical Analysis**

Statistical analysis was performed using the Statistical Package for Social Sciences software (SPSS) version 16.0 for Windows (IBM Corporation Armonk, New York). Continuous variables were expressed as mean±standard deviation and categorical variables as percentages. The logistic regression analysis for the outcome of in-hospital mortality was performed by using the value given to each patient by the InsCor, EuroSCORE and EuroSCORE II scores. Calibration and discrimination were measured for each score value in the patient population. The performance of the models was also measured by comparing mortality between observed and expected mortality in risk groups established by the models. The Fisher exact test was used for contingency tables. The *P* value <0.05 was considered significant.

**Ethics and written informed consent**

This study was approved by the Research Ethics Committee for Projects Analysis (CAPPesq) at Clinics Hospital of the University of São Paulo, with the number 1575.

**RESULTS**

**Calibration**

*InsCor*

Calibration of InsCor was adequate, with *P*=0.184 in the Hosmer-Lemeshow test. The average value of InsCor for survivors was significantly lower than for deaths (3.64 ± 3.5 and 7.96 ± 4.6, *P*<0.001). In Table 1, the InsCor calibration by risk group is presented.

Table 1. InsCor calibration - Analysis by risk group.

Risk	Number of cases	Mortality	
		% observed (95% CI)	% predicted (95% CI)
Low (0-3)	437	2.97 (1.38; 4.57)	4.35 (2.44; 6.26)
Mean (4-7)	317	10.09 (6.78; 13.41)	8.83 (5.71; 11.96)
High (≥8)	246	26.83 (21.29; 32.37)	26.02 (20.53; 31.50)

*Hosmer-Lemeshow test (P=0.184)*

**EuroSCORE**

The calibration of the EuroSCORE was also adequate, with  $P=0.593$  in the Hosmer-Lemeshow test. In Table 2, the calibration of the EuroSCORE by risk groups is presented.

**EuroSCORE II**

The calibration of the EuroSCORE II was not appropriate, with  $P=0.0003$  in the Hosmer-Lemeshow test. In Table 3, the calibration of the EuroSCORE II by risk group is presented.

**Discrimination**

***InsCor and EuroSCORE***

On discrimination, the area under the ROC curve of the EuroSCORE was 0.81 [95% CI (0.77 to 0.86),  $P<0.001$ ] and the InsCor was 0.79 [95% CI (0.74 to 0.83),  $P<0.001$ ] (Figure 1).

**EuroSCORE II**

On discrimination, the area under the ROC curve was 0.81 [95% CI (0.77 to 0.85)  $P<0.001$ ] for the EuroSCORE II (Figure 2).

Table 2. EuroSCORE calibration - Analysis by risk group.

Risk	Number of cases	Mortality	
		% observed (95% CI)	% predicted (95% CI)
Low (0-2)	333	2.10 (0.56; 3.64)	2.40 (0.76; 4.05)
Mean (3-5)	328	5.79 (3.26; 8.32)	5.79 (3.26; 8.32)
High ( $\geq 6$ )	339	25.07 (20.45; 29.69)	24.79 (20.18; 29.37)

*Hosmer-Lemeshow test (P=0.593)*

Table 3. EuroSCORE II calibration - Analysis by risk group.

Risk	Number of cases	Mortality	
		% observed (95% CI)	% predicted (95% CI)
Low (0.17-0.80)	180	1.11 (-0.42; 2.64)	6.67 (3.02; 10.31)
Mean (0.81-1.22)	182	2.75 (0.37; 5.12)	6.59 (2.99; 10.20)
Mean-High (1.23-2.02)	181	6.08 (2.60; 9.56)	7.73 (3.84; 11.63)
High (2.03-4.11)	182	14.84 (9.67; 20.00)	9.34 (5.11; 13.57)
Very High (4.14-47.60)	175	31.43 (24.55; 38.31)	25.71 (19.24; 32.19)

*Hosmer-Lemeshow test (P=0.0003)*

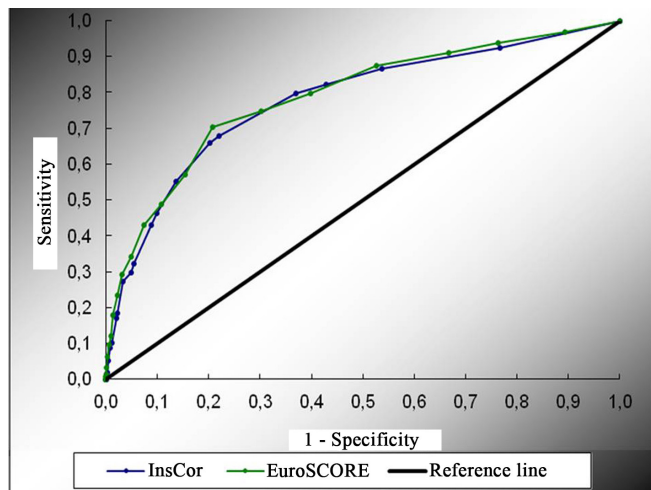


Fig. 1 - ROC curve for InsCor 0.79 [95% CI (0.74 to 0.83),  $P < 0.001$ ] and the EuroSCORE was 0.81 [95% CI (0.77 to 0.86),  $P < 0.001$ ] in the evaluation of the power of discrimination performed in 1,000 patients

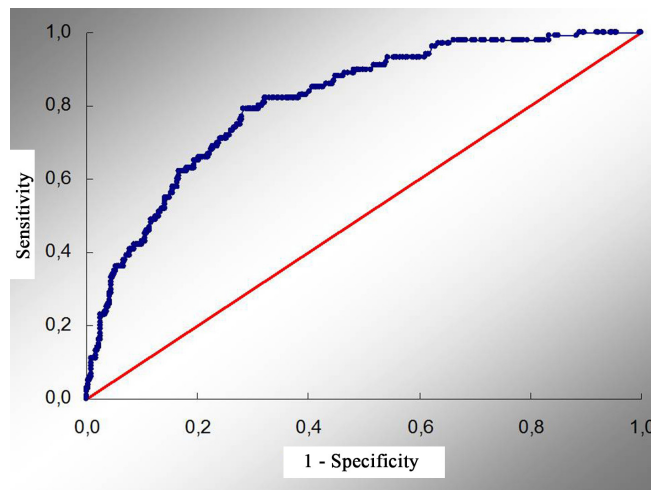


Fig. 2 - ROC curve for EuroSCORE II 0.81 [95% CI (0.77 to 0.85),  $P < 0.001$ ] in the evaluation of the power of discrimination performed in 900 patients

## DISCUSSION

Risk scores should be simplified formulas without the need for personal digital assistants or calculators to predict mortality or other adverse effects at the bedside. They are a valuable aid in therapeutic decisions and for informed consent [16].

However, to be incorporated into the risk models they must be validated. Validating a model means to investigate its calibration and discrimination of a population under certain conditions. Proper calibration and especially good discrimination are the most important factors of a model. Thus, in a model with high discrimination power, many variables are needed in general. In this situation, there is the risk of overfitting. An important feature for adherence of the model is that it is simple and comprehensive, so that the methodology is important [17].

In the history of cardiac surgery, the risk prediction model with greater impact was the EuroSCORE and was published in 1999 by Nashef et al. [5], with more than 108,000 references on Google search and some 1,300 formal citations in the medical literature. This model includes 17 risk factors, from 19,030 patients from 128 centers in Europe. In 2012, in Brazil, the remodeling of EuroSCORE and 2000 Bernstein-Parsonnet models together through the bootstrap technique, gave rise to InsCor [14]. This parsimonious model

consists of 10 variables and can be used for predicting mortality in cardiovascular procedures of adults.

Over time, countries that have adopted strict monitoring by the EuroSCORE, in the past decade, had to adjust the model to their new “Hawthorne Effect” results. Thus, in October 2011, Nashef et al. [10] presented in Lisbon, in 25<sup>th</sup> European Association for Cardio-Thoracic Surgery Annual Meeting, the EuroSCORE remodeled, which came to be called EuroSCORE II. In this study, 23,000 patients underwent cardiac surgery in more than 150 hospitals in 43 countries between May and July 2010. In the internal model validation, on calibration, the observed mortality was 3.9% and the expected mortality by EuroSCORE II of 3.77%, compared to 4.6% of the EuroSCORE. The authors also reported that discrimination of the new model was very good, although the model was not described in the presentation.

In our study, the discrimination of three models proved to be adequate, which means that qualitatively the variables included in the models are the same that have strong relationship with mortality. However, the calibration with respect to the amount or intensity of each predictor variable was adequate for InsCor and EuroSCORE and bad for the EuroSCORE II. Faced with these results, we were waiting for the complete version of the EuroSCORE II, held in January 2012 [10].

After careful analysis of this publication, we point out



some problems in the internal validation of the EuroSCORE II, justifying inadequate external validation of the model. Our analysis is consistent and supported by several sequential international publications [18-20], being reinforced by editorial that demonstrated that, in fact, there are problems in the design of the EuroSCORE II [21,22]. In general, problems with randomly division into two groups for development and validation of the model and details such as the  $P=0.0505$  (ideal  $>0.05$ ) value in the Hosmer-Lemeshow test, stating a good calibration, are questionable [23]. It is doubtful, especially considering the association of this statistical value with some clinical significance.

The term EuroSCORE was also inappropriate, since several non-European countries participated in the remodeling of the model. With this in mind, it would be better to calculate the mortality rate itself or the local risk-adjusted hospital, since the model was built to predict death in a wide variety of groups, making it difficult to forecast specific clinical scenarios. Another reason for poor calibration would be the large number of highly correlated risk factors, including confounding variables and over-adjusted to a certain types of procedures or specific subgroups of patients.

Upon publication of the EuroSCORE II, it was not reported if analyzes of first order interaction and multicollinearity were performed, so many variables could overestimate the risk of certain categories of patients (e.g., intermediate risk or extreme risk). In the follow-up, there was inefficient management of patients with loss of data, where the bias arises due to significant differences between individuals with complete data and those with missing data. Thus, a regression coefficient calculated for a predictor may be influenced if missing data were associated with the outcome. In EuroSCORE II, the authors could have chosen otherwise imputation to preserve these cases.

In general, the performance of the participating centers, with major failures in the supply of data, especially in the follow-up, was poor [21]. Furthermore, there should be more careful in order to not to increase the number of variables at all times, since models with only a few variables are very stable and, if robust they may achieve good calibration. The inclusion of many variables increases the risk of errors that can be caused by differences in interpretation of definitions, types, or conflicting information. The reduced number of variables without affecting its accuracy ("few variables as possible") in comprehensive models is one of the most important aspects of the cost, popularity and applicability of risk scores [12,24].

Another concern with the EuroSCORE II is that the primary outcome was mortality at the base hospital, and we cannot forget that, in actual practice, it is common for patients to be transferred to other hospitals in accordance with clinical outcome.

Recently, Kunt et al. [20] compared the EuroSCORE,

STS score and EuroSCORE II in a population of 428 patients who underwent isolated coronary surgery, between 2004 and 2012 in Turkey. The mortality rate was 7.9% and the predicted mortality was 6.4% for the additive EuroSCORE, 7.9% for the logistic EuroSCORE, 1.7% for the EuroSCORE II and 5.8% for STS score. The area under the ROC curve for the additive EuroSCORE, logistic EuroSCORE, STS score and EuroSCORE II was 0.7, 0.7, 0.72 and 0.62, respectively.

In the modern evolution of risk assessment, it has been widespread the concept of applying external models and remold them to the characteristics of the region [25]. To apply a risk score, it must first be remolded (adaptation of the variables and their weights) or at least recalibrated (adjusting the weights of the variables) and never used form of ready-made (without adaptation of the variables and their weights) [24]. In Brazil, the adhesion of a model itself is of paramount importance, especially by differences in patient characteristics, clinical presentation due to socioeconomic, cultural and geographical reasons, the uneven distribution of medical facilities and the high endemicity of subclinical inflammation, infection and rheumatic disease [25].

Thus, the external validation InsCor is required. We are already in advanced work in collaboration with seven centers of large representation of the state of São Paulo, for the study and creation of the SP-SCORE [26].

Importantly, risk scores are based on the experience of the participating teams, patients with regional characteristics and certain infrastructure and time. A model cannot be transported to other locations or be included in the same location over time without performing preliminary validation tests, so it is important to know the limitations of these instruments.

### Limitations

Although data were collected prospectively, this is a retrospective analysis. However, the collection within the electronic database was "blind", or that is, we selected the first 1000 patients undergoing coronary and/or valve within the period studied without knowledge of clinical outcome. Another important factor is that, as the study was retrospective, only 900 patients had all the data to calculate the EuroSCORE II. To minimize this limitation, we performed an analysis with 100 unselected patients and observed that the mortality of these patients showed no statistical difference with the selected group to perform validation of the EuroSCORE II.

### CONCLUSION

The InsCor and EuroSCORE were adequate in all phases of the validation. However, the errors found in the design of the EuroSCORE II were also manifest in the calibration of patients undergoing coronary and/or valve surgery on InsCor-HCFMUSP. These data reinforce the importance of InsCor local model and future SP-SCORE.

Author' roles and responsibilities	
LAF	Study design, analysis of results and writing of the manuscript
OAVM	Study design, collection of data and writing of the manuscript
LFPM	Evaluation of results and statistics
LAOD	Evaluation of results and discussion
PMAP	Evaluation of results and discussion
LRPA	Medical records analysis and risk factors
MRBM	Medical records analysis and risk factors
FBJ	Study design and discussion

## REFERENCES

1. Kolh P, Wijns W. Essential messages from the ESC/EACTS guidelines on myocardial revascularization. *Eur J Cardiothorac Surg.* 2012;41(5):983-5.
2. Takkenberg JJ, Kappetein AP, Steyerberg EW. The role of EuroSCORE II in 21<sup>st</sup> century cardiac surgery practice. *Eur J Cardiothorac Surg.* 2013;43(1):32-3.
3. Hannan EL, Cozzens K, King SB 3<sup>rd</sup>, Walford G, Shah NR. The New York State cardiac registries: history, contributions, limitations, and lessons for future efforts to assess and publicly report healthcare outcomes. *J Am Coll Cardiol.* 2012;59(25):2309-16.
4. Shahian DM, Normand SL. Comparison of "risk-adjusted" hospital outcomes. *Circulation.* 2008;117(15):1955-63.
5. 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(1):9-13.
6. 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.
7. Andrade IN, Moraes Neto FR, Oliveira JP, Silva IT, Andrade TG, Moraes CR. Assessment of the EuroSCORE as a predictor for mortality in valve cardiac surgery at the Heart Institute of Pernambuco. *Rev Bras Cir Cardiovasc.* 2010;25(1):11-8.
8. Mejia OA, Lisboa LA, Dallan LA, Pomerantzeff PM, Moreira LF, Jatene FB, et al. Validation of the 2000 Bernstein-Parsonnet and EuroSCORE at the Heart Institute - USP. *Rev Bras Cir Cardiovasc.* 2012;27(2):187-94.
9. Nashef SA; EuroSCORE Project team. The New EuroSCORE Project. Nowa skala EuroSCORE. *Kardiologia Pol.* 2010;68(1):128-9.
10. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg.* 2012;41(4):734-44.
11. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med.* 2000;19(4):453-73.
12. Tu JV, Sykora K, Naylor CD. Assessing the outcomes of coronary artery bypass graft surgery: how many risk factors are enough? Steering Committee of the Cardiac Care Network of Ontario. *J Am Coll Cardiol.* 1997;30(5):1317-23.
13. Ranucci M, Castelvécchio S, Conte M, Megliola G, Speziale G, Fiore F, et al. The easier, the better: age, creatinine, ejection fraction score for operative mortality risk stratification in a series of 29,659 patients undergoing elective cardiac surgery. *J Thorac Cardiovasc Surg.* 2011;142(3):581-6.
14. Mejia OA, Lisboa LA, Puig LB, Moreira LF, Dallan LA, Pomerantzeff PM, et al. InsCor: a simple and accurate method for risk assessment in heart surgery. *Arq Bras Cardiol.* 2013;100(3):246-54.
15. Lisboa LA, Moreira LF, Mejia OV, Dallan LA, Pomerantzeff PM, Costa R, et al. Evolution of cardiovascular surgery at the Instituto do Coração: analysis of 71,305 surgeries. *Arq Bras Cardiol.* 2010;94(2):162-8.
16. Hannan EL, Racz M, Culliford AT, Lahey SJ, Wechsler A, Jordan D, et al. Risk score for predicting in-hospital/30-day mortality for patients undergoing valve and valve/coronary artery bypass graft surgery. *Ann Thorac Surg.* 2013;95(4):1282-90.
17. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.
18. Carnero-Alcázar M, Silva Guisasaola JA, Reguillo Lacruz FJ, Maroto Castellanos LC, Cobiella Carnicer J, Villagrán Medinilla E, et al. Validation of EuroSCORE II on a single-centre 3800 patient cohort. *Interact Cardiovasc Thorac Surg.* 2013;16(3):293-300.
19. Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N, et al. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. *Eur J Cardiothorac Surg.* 2013;43(4):688-94.
20. Kunt AG, Kurtcephe M, Hidiroglu M, Cetin L, Kucuker A, Bakuy V, et al. Comparison of original EuroSCORE, EuroSCORE II and STS risk models in a Turkish cardiac surgical cohort. *Interact Cardiovasc Thorac Surg.* 2013;16(5):625-9.
21. Sergeant P, Meuris B, Pettinari M. EuroSCORE II, illum qui est gravitates magni observe. *Eur J Cardiothorac Surg.* 2012;41(4):729-31.

22. Collins GS, Altman DG. Design flaws in EuroSCORE II. *Eur J Cardiothorac Surg.* 2013;43(4):871.
23. Nezic D, Borzanovic M, Spasic T, Vukovic P. Calibration of the EuroSCORE II risk stratification model: is the Hosmer-Lemeshow test acceptable any more? *Eur J Cardiothorac Surg.* 2013;43(1):206.
24. Mejía OA, Lisboa LA. The risk of risk scores and the dream of BraSCORE. *Rev Bras Cir Cardiovasc.* 2012;27(2):xii-xiii.
25. Sá MP, Sá MV, Albuquerque AC, Silva BB, Siqueira JW, Brito PR, et al. GuaragnaSCORE satisfactorily predicts outcomes in heart valve surgery in a Brazilian hospital. *Rev Bras Cir Cardiovasc.* 2012;27(1):1-6.
26. Mejía OA, Lisboa LA, Dallan LA, Pomerantzeff PM, Trindade EM, Jatene FB, et al. Heart surgery programs innovation using surgical risk stratification at the São Paulo State Public Healthcare System: SP-SCORE-SUS study. *Rev Bras Cir Cardiovasc.* 2013;28(2):263-9.