

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 Lisboa¹, MD, PhD; Omar Asdrubal Vilca Mejia¹, MD, PhD; Luiz Felipe Pinho Moreira¹, MD, PhD; Luís Alberto Oliveira Dallan¹, MD, PhD; Pablo Maria Alberto Pomerantzeff¹, MD, PhD; Luís Roberto Palma Dallan¹, MD; Maria Raquel B. Massoti¹, MD; Fabio B. Jatene¹, 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 evaluate 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. Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (InCor-HCFMUSP), São Paulo, SP, Brasil.

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, Brasil – CEP: 05403-000
E-mail: luiz.lisboa@incor.usp.br

Trabalho realizado no Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (InCor-HCFMUSP), São Paulo, SP, Brasil.

Não houve suporte financeiro.

Endereço para correspondência:
Luiz Augusto Ferreira Lisboa

Artigo recebido em 12 de outubro de 2013
Artigo aprovado em 17 de novembro de 2013

Abreviaturas, acrônimos & símbolos	
EuroSCORE	European System for Cardiac Operative Risk Evaluation
InCor-HCFMUSP	Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de 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.

INTRODUÇÃO

Na medicina moderna, a utilização de escores de risco como preditores de eventos está bem estabelecida [1]. Modelos eficientes devem ser originados de registros prospectivos, obrigatórios e completos, serem construídos mediante técnicas estatísticas de *bootstrap* e demonstrar adequada validação interna, seguindo estritamente os princípios científicos [2,3]. É evidente que modelos de risco derivados e validados em um local, usualmente, tenham menor desempenho quando aplicados em outro local e, inclusive, no mesmo local ao longo do tempo [4]. Contudo, o primeiro EuroSCORE criado em 1999 [5], em população europeia, mostrou-se adequado em uma população brasileira contemporânea [6-8].

Naturalmente, a incorporação do EuroSCORE nos principais serviços da Europa trouxe à vista o efeito “*Hawthorne*”, explicando que nada melhorou tanto os resultados em cirurgia cardíaca, no início do século, como a monitorização pelo EuroSCORE [9]. Com o tempo, o remodelamento do EuroSCORE para os países que aderiram seu uso obrigatório seria justificável. Dessa forma, deu-se origem ao EuroSCORE II [10], a partir de um registro com 22.381 pacientes consecutivos submetidos a cirurgia cardíaca, em 154 hospitais de 43 países (dentro e fora da Europa), no período de 12 semanas (maio a julho 2010).

Esse modelo atualizado apresenta mais variáveis que o primeiro EuroSCORE, portanto, além de ter o risco de alto poder de discriminação, traz o risco de *overfitting* [11]. As-

sim, modelos menores apresentam boa calibração, mas infelizmente diminuição do poder de discriminação. Mesmo assim, não devemos esquecer que “poucas variáveis quanto possíveis” é o que prevalece num modelo para ter maior aceitação [12,13]. No Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (InCor-HCFMUSP), o remodelamento conjunto dos modelos EuroSCORE e 2000 Bernstein-Parsonnet [8], com o uso da técnica de *bootstrap*, deu origem ao InsCor [14]. Esse modelo teve desempenho semelhante ao primeiro EuroSCORE e foi mais simples do que este e que o 2000 Bernstein-Parsonnet para prever mortalidade nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP. Esse fato se faz mais importante quando existe a necessidade de avaliar a experiência de tratamento frente ao “*casemix*” local em um tempo determinado, como vem sendo realizado por diversos grupos. Assim, o objetivo deste estudo foi validar o EuroSCORE II e compará-lo aos modelos InsCor e EuroSCORE, em pacientes operados de coronária e/ou valva no InCor-HCFMUSP.

MÉTODOS

Tamanho da amostra, critérios de inclusão e exclusão

Uma análise retrospectiva de dados coletados prospectivamente foi realizada na Divisão de Cirurgia Cardiovascular do InCor-HCFMUSP. Para validação dos escores de risco em uma amostra com pelo menos 100 óbitos, a publicação de

Lisboa et al. [15], sobre os resultados da cirurgia cardiovascular do InCor-HCFMUSP, nos últimos 23 anos, foi a base para o estudo. Para tanto, 1000 pacientes operados sequencialmente de revascularização miocárdica isolada ou associada e/ou a cirurgia valvar isolada ou associada, incluindo reoperações e nas modalidades eletiva, urgência ou emergência, no período de outubro de 2008 a julho de 2009, foram selecionados. Destes, todos preenchiam as variáveis contidas nos modelos InsCor e EuroSCORE, no entanto, somente 900 pacientes incluíam todas as variáveis requeridas pelo EuroSCORE II. Pacientes menores de 18 anos ou submetidos a outros tipos de cirurgia que não a revascularização miocárdica e/ou a cirurgia valvar foram excluídos do estudo.

Coleta, definição e organização dos dados

Os dados foram coletados do sistema eletrônico de prontuários do InCor (SI3) e armazenados em planilhas. Cada planilha foi adaptada de forma a contemplar todas as variáveis, respeitando-se suas definições como descritas pelos modelos EuroSCORE [9], EuroSCORE II [10] e InsCor [14]. Os pacientes foram ordenados segundo os grupos de risco estabelecidos pelos escores e colocados no banco de dados confeccionado no programa Excel. O desfecho de interesse foi mortalidade intra-hospitalar, definida como a morte ocorrida no intervalo de tempo entre a cirurgia e a alta hospitalar.

Validação do InsCor, EuroSCORE e EuroSCORE II

Para avaliar o desempenho do InsCor, EuroSCORE e EuroSCORE II na predição de mortalidade, foi realizada a validação preditiva dos modelos. A análise foi feita mediante teste de calibração e discriminação. A calibração avalia a acurácia do modelo para predizer risco em um grupo de pacientes. A força da calibração foi avaliada testando a qualidade do ajuste mediante o teste de Hosmer-Lemeshow. O valor de $P > 0,05$ indica que o modelo se ajusta aos dados e prediz a mortalidade adequadamente. A discriminação mede

a habilidade do modelo para distinguir entre pacientes de baixo e alto risco. A discriminação foi medida por meio da utilização da técnica estatística chamada área abaixo da curva ROC (*Receiver Operating Characteristic*, algumas vezes chamada *c-statistic* ou *c-index*).

Análise estatística

A análise estatística foi realizada com auxílio do *software Statistical Package for the Social Sciences* (SPSS) versão 16.0 para Windows (IBM Corporation Armonk, New York). As variáveis contínuas foram expressas como a média \pm desvio-padrão e as variáveis categóricas, como porcentagens. A análise de regressão logística para o desfecho mortalidade intra-hospitalar foi realizada por meio da utilização do valor dado para cada paciente pelos escores InsCor, EuroSCORE e EuroSCORE II. A calibração e a discriminação foram medidas para cada valor do escore na população de pacientes. O desempenho dos modelos foi também medido por meio da comparação entre a mortalidade observada e a mortalidade esperada nos grupos de risco estabelecidos pelos modelos. O teste exato de Fisher foi utilizado para as tabelas de contingência. O valor $P < 0,05$ foi considerado significativo.

Ética e termo de consentimento

Este trabalho foi aprovado pela Comissão de Ética para Análises de Projetos de Pesquisa (CAPPesq) do Hospital das Clínicas da Universidade de São Paulo, com o número 1575.

RESULTADOS

Calibração

InsCor

A calibração do InsCor foi adequada, com $P = 0,184$ no teste de Hosmer-Lemeshow. O valor médio do InsCor para os sobreviventes foi significativamente menor do que para os óbitos ($3,64 \pm 3,5$ e $7,96 \pm 4,6$, $P < 0,001$). Na Tabela 1, é apresentada a calibração do InsCor por grupos de risco.

Tabela 1. Calibração do InsCor - Análise por grupos de risco.

Risco	Número de casos	Mortalidade	
		% observada (IC a 95%)	% predita (IC a 95%)
Baixo (0-3)	437	2,97 (1,38; 4,57)	4,35 (2,44; 6,26)
Médio (4-7)	317	10,09 (6,78; 13,41)	8,83 (5,71; 11,96)
Alto (≥ 8)	246	26,83 (21,29; 32,37)	26,02 (20,53; 31,50)

Teste de Hosmer-Lemeshow ($P = 0,184$)

EuroSCORE

A calibração do EuroSCORE também foi adequada, com $P=0,593$ no teste de Hosmer-Lemeshow. Na Tabela 2, é apresentada a calibração do EuroSCORE por grupos de risco.

EuroSCORE II

A calibração do EuroSCORE II não foi adequada, com $P=0,0003$ no teste de Hosmer-Lemeshow. Na Tabela 3, é apresentada a calibração do EuroSCORE II por grupos de risco.

Discriminação

InsCor e EuroSCORE

Na discriminação, a área abaixo da curva ROC do 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), $P<0,001$] (Figura 1).

EuroSCORE II

Na discriminação, a área abaixo da curva ROC foi 0,81 [IC 95% (0,77-0,85) $P<0,001$] para o EUROSCORE II (Figura 2).

Tabela 2. Calibração do EuroSCORE - Análise por grupos de risco

Risco	Número de casos	Mortalidade	
		% observada (IC a 95%)	% predita (IC a 95%)
Baixo (0-2)	333	2,10 (0,56; 3,64)	2,40 (0,76; 4,05)
Médio (3-5)	328	5,79 (3,26; 8,32)	5,79 (3,26; 8,32)
Alto (≥ 6)	339	25,07 (20,45; 29,69)	24,79 (20,18; 29,37)

Teste de Hosmer-Lemeshow ($P=0,593$)

Tabela 3. Calibração do EuroSCORE II - Análise por grupos de risco

Risco	Número de casos	Mortalidade	
		% observada (IC a 95%)	% predita (IC a 95%)
Baixo (0,17-0,80)	180	1,11 (-0,42; 2,64)	6,67 (3,02; 10,31)
Médio (0,81-1,22)	182	2,75 (0,37; 5,12)	6,59 (2,99; 10,20)
Médio Alto (1,23-2,02)	181	6,08 (2,60; 9,56)	7,73 (3,84; 11,63)
Alto (2,03-4,11)	182	14,84 (9,67; 20,00)	9,34 (5,11; 13,57)
Muito Alto (4,14-47,60)	175	31,43 (24,55; 38,31)	25,71 (19,24; 32,19)

Teste de Hosmer-Lemeshow ($P=0,0003$)

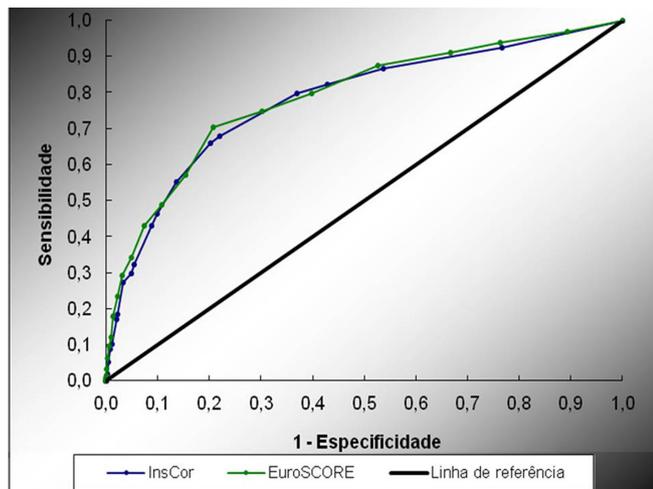


Fig. 1 – Curva ROC para o InsCor de 0,79 [IC 95% (0,74 – 0,83), $P<0,001$] e para o EuroSCORE de 0,81 [IC 95% (0,77 – 0,86), $P<0,001$] na avaliação do poder de discriminação realizada em 1.000 pacientes

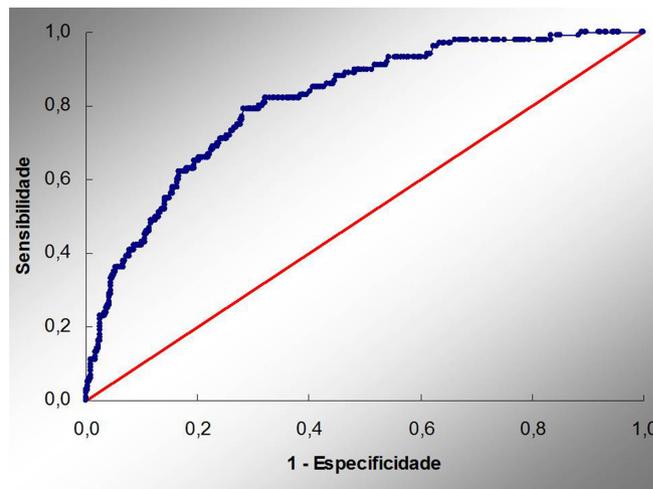


Fig. 2 – Curva ROC para o EuroSCORE II de 0,81 [IC 95% (0,77 – 0,85), $P<0,001$] na avaliação do poder de discriminação realizada em 900 pacientes

DISCUSSÃO

Os escores de risco deveriam ser fórmulas simplificadas, sem necessidade de assistentes pessoais digitais ou calculadoras para prever a mortalidade ou outros efeitos adversos à beira do leito. Eles são valiosos no auxílio das decisões terapêuticas e para o consentimento informado [16].

No entanto, para serem incorporados, os modelos de risco precisam ser validados. Validar um modelo significa investigar sua calibração e discriminação em relação a uma população sob determinadas condições. Adequada calibração e, principalmente, boa discriminação são os fatores mais importantes de um modelo. Assim, em um modelo com alto poder de discriminação, são necessárias, em geral, muitas variáveis. Nessa situação, corre-se o risco de *overfitting*. Uma característica importante para a aderência do modelo é que seja simples e abrangente, para isso a metodologia é importante [17].

Na história da cirurgia cardíaca, o modelo de predição de risco com maior impacto foi o EuroSCORE, publicado em 1999, por Nashef et al. [5], com mais de 108.000 referências na pesquisa Google e cerca de 1.300 citações formais na literatura médica. Esse modelo inclui 17 fatores de risco, provenientes de 19.030 pacientes em 128 centros da Europa. Em 2012, no Brasil, o remodelamento conjunto dos modelos EuroSCORE e 2000 Bernstein-Parsonnet, por meio da técnica de *bootstrap*, deu origem ao InsCor [14]. Esse modelo parcimonioso é constituído por 10 variáveis e pode ser utilizado para predição de mortalidade em procedimentos cardiovasculares de adulto.

Com o tempo, os países que adotaram estritamente a monitorização pelo EuroSCORE, na década passada, precisaram ajustar o modelo para seus novos resultados “Efeito Hawthorne”. Assim, em outubro de 2011, Nashef et al. [10] apresentou em Lisboa, no 25th *European Association for Cardio-Thoracic Surgery Annual Meeting*, o EuroSCORE remodelado, que passou a ser denominado de EuroSCORE II. Nesse estudo, 23.000 pacientes foram operados de cirurgia cardíaca em mais de 150 hospitais de 43 países, entre maio e julho de 2010. Na validação interna do modelo, na calibração, a mortalidade observada foi de 3,9% e a mortalidade esperada pelo EUROSCORE II de 3,77%, em comparação a 4,6% do EuroSCORE. Os autores referiam, também, que a discriminação do novo modelo era muito boa, embora a modelagem não fosse descrita na apresentação.

Em nosso estudo, a discriminação dos três modelos revelou-se adequada, o que significa que qualitativamente as variáveis incluídas pelos modelos são as mesmas que apresentam forte relação com mortalidade. Entretanto, a calibração, que diz respeito à quantidade ou à intensidade preditora de cada variável, foi adequada para o InsCor e EuroSCORE e ruim para o EuroSCORE II. Ante esses resultados, ficamos aguardando a publicação completa do EuroSCORE II, realizada em janeiro de 2012 [10].

Após criteriosa análise dessa publicação, apontamos alguns problemas na validação interna do EuroSCORE II, que justificam inadequada validação externa do modelo. A nossa análise é concordante e respaldada por várias publicações internacionais sequenciais [18-20], ficando reforçada por editoriais que demonstraram, de fato, os problemas encontrados

no delineamento do EuroSCORE II [21,22]. De um modo geral, problemas com a divisão aleatória dos grupos de desenvolvimento e de validação do modelo e detalhes como o do valor $P=0,0505$ (ideal $>0,05$) no teste de Hosmer-Lemeshow, afirmando uma boa calibração, são questionáveis [23]. Isso é duvidoso, principalmente considerando a associação desse valor estatístico com algum significado clínico.

O termo EuroSCORE também ficou inadequado, já que vários países não-europeus participaram no remodelamento do modelo. Com isso em mente, seria melhor calcular a taxa de mortalidade própria do local ou hospital ajustada ao risco, já que o modelo foi construído para prever óbito em uma ampla variedade de grupos, o que dificultaria previsões para cenários clínicos específicos. Outro motivo para pobre calibração seria o grande número de fatores de risco altamente correlacionáveis, incluindo variáveis de confusão e sobreajustadas a certos tipos de procedimentos ou subgrupos específicos de pacientes.

Na publicação do EuroSCORE II, não foi relatado se foram realizadas análises de interação de primeira ordem e multicolinearidade, portanto, muitas variáveis poderiam sobrestimar o risco de certas categorias de pacientes (por exemplo, risco intermediário ou risco extremo). No seguimento, houve manuseio ineficiente dos pacientes, com perda de dados, onde o viés surge em decorrência de diferenças significativas entre os indivíduos com dados completos e indivíduos com dados faltantes. Assim, um coeficiente de regressão calculado para um preditor pode estar influenciado se dados faltantes tiveram associação com o resultado. No EuroSCORE II, os autores poderiam ter escolhido outra forma de imputação para preservar esses casos. Em geral, a atuação dos centros participantes, com importantes falhas no fornecimento de dados, principalmente no seguimento, foi precária [21]. Por outro lado, deve haver mais cuidado para não aumentar o número de variáveis o tempo todo, já que modelos com apenas algumas variáveis são bastante estáveis e, se robustos, podem conseguir boa calibração. A inclusão de muitas variáveis aumenta o risco de erros que podem ser causados por diferenças na interpretação das definições, erros de digitação ou informações conflitantes. O número reduzido de variáveis sem afetar sua acurácia (“poucas variáveis quanto possíveis”) em modelos abrangentes é uma das características mais importantes com relação a custo, aplicabilidade e popularidade dos escores de risco [12,24].

Outra preocupação com o EuroSCORE II é que o desfecho primário foi mortalidade no hospital de base, e não podemos esquecer que, na prática atual, é comum que os pacientes sejam transferidos para outros hospitais de acordo com a evolução clínica.

Recentemente, Kunt et al. [20] compararam o EuroSCORE, EuroSCORE II e o STS score em uma população de 428 pacientes operados de coronária isolada, entre 2004 e 2012, na Turquia. A mortalidade observada foi de 7,9% e

a mortalidade predita foi de 6,4% para EuroSCORE aditivo, 7,9% para o EuroSCORE logístico, 1,7% para o EuroSCORE II e 5,8% para o STS score. A área abaixo da curva ROC para o EuroSCORE aditivo, EuroSCORE logístico, EuroSCORE II e STS score foi de 0,7, 0,7, 0,72 e 0,62, respectivamente.

Na evolução da moderna avaliação de risco, vem sendo difundido o conceito de aplicar modelos externos e remodelá-los às características próprias da região [25]. Para aplicar um escore de risco, ele deverá primeiro ser remodelado (adaptação das variáveis e seus pesos) ou pelo menos recalibrado (adaptação dos pesos das variáveis) e nunca utilizado da forma *ready-made* (sem adaptação das variáveis e seus pesos) [24].

No Brasil, a aderência de um modelo próprio é de suma importância, principalmente pelas diferenças nas características dos pacientes, na apresentação clínica devido a razões socioeconômicas, culturais e geográficas, na distribuição desigual de instalações médicas e na alta endemicidade de inflamação subclínica, infecção e doença reumática [25]. Dessa forma, a validação externa do InsCor se faz necessária. Já estamos em trabalho avançado, em colaboração com sete centros de grande representatividade do Estado de São Paulo, para o estudo e criação do SP-SCORE [26].

É importante salientar que escores de risco são baseados na experiência das equipes participantes, com pacientes de características regionais e com infraestrutura e tempo determinado. Um modelo não pode ser transportado para outros locais ou ser incluso no mesmo local ao longo do tempo sem a realização de testes prévios de validação, portanto, é importante conhecer as limitações desses instrumentos.

Limitações

Apesar dos dados serem coletados prospectivamente, trata-se de uma análise retrospectiva. Entretanto, a coleta dentro do banco de dados eletrônico foi “cega”, ou seja, foram selecionados os primeiros 1000 pacientes operados de coronária e/ou valva, dentro do período estudado, sem o conhecimento da evolução clínica. Outro fator de importância é que, como o estudo foi retrospectivo, apenas 900 pacientes apresentavam todos os dados para cálculo do EuroSCORE II. Para minimizar essa limitação, realizamos uma análise com os 100 pacientes não selecionados e observamos que a mortalidade desses pacientes não apresentou diferença estatística com o do grupo selecionado para realizar a validação do EuroSCORE II.

CONCLUSÃO

O InsCor e o EuroSCORE foram adequados em todas as fases da validação. No entanto, os erros encontrados no desenho do EuroSCORE II também foram manifestos na calibração dos pacientes operados de coronária e/ou valva no InCor-HCFMUSP. Esses dados reforçam a importância do modelo local InsCor e do futuro SP-SCORE.

Papéis & responsabilidades dos autores

LAF	Desenho do estudo, análise dos resultados e redação do manuscrito
OAVM	Desenho do estudo, levantamento dos dados e redação do manuscrito
LFPM	Avaliação dos resultados e estatística
LAOD	Avaliação dos resultados e discussão
PMAP	Avaliação dos resultados e discussão
LRPA	Análise de prontuários e dos fatores de risco
MRBM	Análise de prontuários e dos fatores de risco
FBJ	Desenho do estudo e discussão

REFERÊNCIAS

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 21st century cardiac surgery practice. *Eur J Cardiothorac Surg.* 2013;43(1):32-3.
3. Hannan EL, Cozzens K, King SB 3rd, 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, Castelvechchio 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, Kurtcepe 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.