

OMAR ASDRÚBAL VILCA MEJÍA

**Predição de mortalidade em cirurgia de coronária e/ou
valva no InCor: Validação de dois modelos externos e
comparação com o modelo desenvolvido localmente
(InsCor)**

Tese apresentada à Faculdade de Medicina da
Universidade de São Paulo para obtenção do título
de Doutor em Ciências

Programa de Cirurgia Torácica e Cardiovascular
Orientador: Prof. Dr. Noedir Antonio Groppo
Stolf.

Co-orientador: Prof. Dr. Luiz Augusto Ferreira
Lisboa.

São Paulo

2012

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Mejía, Omar Asdrúbal Vilca
Predição de mortalidade em cirurgia de coronária e/ou valva no InCor: Validação de dois modelos externos e comparação com o modelo desenvolvido localmente (InsCor)/ Omar Asdrúbal Vilca Mejía. -- São Paulo, 2012.

Tese(doutorado)-Faculdade de Medicina da Universidade de São Paulo.
Programa de Cirurgia Torácica e Cardiovascular.
Orientador: Noedir Antônio Groppo Stolf.
Co-orientador: Luiz Augusto Ferreira Lisboa.

Descritores: 1. Fatores de risco; 2. Procedimentos cirúrgicos cardiovasculares; 3. Ponte de artéria coronária; 4. Modelos probabilísticos; 5. Medição de risco; 6. Mortalidade hospitalar; 7. Estudos de validação.

USP/FM/DBD-091/12

Dedicatória

A Nelly Luz Mejía Fernandez, a quem dedico todo meu esforço.

À minha mãe, por toda sua inspiração concretizada na minha pessoa como fonte inesgotável de amor, estímulo e perseverança. Tua presença pode ser sentida a cada momento, grato por vir e saiba que sempre estaremos juntos.

Ao meu pai, por sua humildade, amizade e companhia, por me dar a oportunidade inicial de aprender e discernir.

À minha esposa, pela paciência nos dias de solidão, por saber me esperar chegar em casa e por sua entrega total, obrigado.

Aos meus filhos: Sofia, pedacinho da minha mãe; e Eduardo, o presente da mamãe, ambos me dão a força para poder continuar. Ao meu fiel Enzo, amigo de longas jornadas, por sua companhia, amor e aulas de humanidade.

Às minhas irmãs: Maria Luiza, Natalia, Andrea e Mariana, obrigado por sua consanguinidade e seu eterno amor incondicional.

Aos meus avós Alejandro e Clotilde, por seu carinho e incentivo de luta ao longo da minha vida.

Aos meus sogros e amigos: Pedro e Neusa, pela sua presença nos momentos difíceis e por sua confiança.

A Carlos Velândia e Eduardo Gazzoto, por compartilhar há décadas momentos inesquecíveis.

Agradecimientos

Ao Professor Noedir Stolf, a nossa grande inspiração, pela ajuda extrema na elaboração da tese.

Ao Dr. Luiz Lisboa, amigo, mestre, sempre exigindo mais, mas sempre de nosso lado.

Ao Professor Luiz Puig, sempre genial e aqui um exemplo das suas ideias.

Ao Dr Ricardo Dias, grande médico, a quem devo o aprendizado e principalmente a amizade.

Aos amigos feitos no tempo que durou o estudo: Professor Luiz Dallan, Professor Pablo Pomerantzeff, Professor Luiz Felipe, Professor Marcelo Jatene, Dr Fabio Gaiotto, Dr Anderson Benicio, Dr Lauro, aos anestesistas Dr Dagoberto, Dra Katia, Dra Raquel, Dr Celso. A todos eles expresse meu total agradecimento.

Aos professores Samir Nashef de Cambridge e Dr. Edward Hannan de Nova York, por seu apoio e colaboração na definição das variáveis do banco de dados.

A Adriana Quadros e Sandra Vitória dos Santos, por conseguir entender e traduzir os meus pensamentos.

A Neusa, Juliana e Eva da Pós-graduação; a Sandra, Raquel e Eurides da Unidade Cirúrgica de Cardiopatias Gerais e a Malú da Unidade Cirúrgica de Coronariopatia, a todas elas meu eterno agradecimento.

A Fátima e Marisa, por sua participação direta na coleta dos dados e no seguimento dos pacientes.

Aos novos colegas: Dr Vicente de Ávila Neto e André Martins, obrigado por ter acolhido os meus defeitos e virtudes.

A Creusa Dal Bó, pelo acompanhamento da análise estatística.

Ao CNPq, por seu apoio inesgotável na realização da tese.

*“O verdadeiro ato da descoberta
não consiste em encontrar novas terras,
mas sim em vê-las com novos olhos”*

Marcel Proust

Normatização adotada

Esta tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de *International Committee of Medical Journals Editors* (Vancouver)

Universidade de São Paulo. Faculdade de Medicina. Serviço de Biblioteca e Documentação. *Guia de apresentação de dissertações, teses e monografias*. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 2ª ed. São Paulo: Serviço de Biblioteca e Documentação. 2005.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

Normas do Acordo Ortográfico da Língua Portuguesa, de janeiro de 2009.

Sumário

LISTA DE ABREVIATURAS**LISTA DE SÍMBOLOS****LISTA DE TABELAS****LISTA DE FÍGURAS****RESUMO****SUMMARY**

1.	INTRODUÇÃO.....	1
2.	OBJETIVOS.....	10
3.	REVISÃO DA LITERATURA.....	11
3.1	Origem dos escores de risco.....	12
3.2	Formulação dos escores de risco.....	14
3.3	Validação dos escores de risco.....	21
3.4	Recalibração e remodelamento dos escores de risco.....	28
4.	MÉTODOS.....	37
4.1	Amostra.....	38
4.2	Critérios de inclusão e exclusão.....	38
4.3	Coleta, definição e organização dos dados.....	39
4.4	Validação do 2000 Bernstein-Parsonnet e do EuroSCORE.....	40
4.5	Formulação do modelo local (InsCor).....	41
4.6	Ética e Termo de Consentimento.....	43
5	RESULTADOS.....	45
5.1	Casuística.....	46
5.2	Resultados da validação do 2000 Bernstein Parsonnet e do EuroSCORE.....	49
5.2.1	Resultados da Calibração.....	49
5.2.2	Resultados da Discriminação.....	52
5.3	O modelo desenvolvido: O InsCor.....	54

5.4	Calibração do InsCor.....	60
5.5	Calibração do EuroSCORE.....	61
5.6	Calibração do 2000BP.....	61
6	DISCUSSÃO.....	64
6.1	Caracterização da amostra utilizada.....	65
6.1.1	Limitação do estudo.....	79
6.2	Considerações finais.....	80
7.	CONCLUSÕES.....	84
8.	ANEXOS.....	86
8.1	Planilha unificada para coleta de dados.....	87
8.2	Variáveis utilizadas no estudo definidas pelos escores 2000 Bernstein-Parsonnet e EuroSCORE.....	88
8.3	Ética e Termo de consentimento.....	101
9.	REFERÊNCIAS BIBLIOGRÁFICAS.....	102

Listas

LISTA DE ABREVIATURAS

ES:	EuroSCORE
2000BP:	2000 Bernstein Parsonnet
InCor-HCFMUSP:	Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
CRM:	Cirurgia de revascularização miocárdica
UTI:	Unidade de terapia intensiva
ROC:	<i>Receiver operating characteristic</i>
NYHA:	<i>New York Heart Association</i>
CCS:	<i>Canadian Cardiovascular Society</i>
col.	colaboradores
DATASUS:	Banco de dados do sistema único de saúde
CABG:	<i>Coronary artery bypass graft</i>
OR:	<i>Odds ratio</i>
dF:	<i>Degree free</i>
ep:	erro padrão
LI:	limite inferior
LS:	limite superior
pt:	pontos.
P:	significância estatística
ex.:	exemplo
Inc.:	<i>Incorporated</i>
h:	horas

LISTA DE SIMBOLOS

mg/dl:	miligrama por decilitro
mmHg:	milímetro de mercúrio
μmol/L:	micromol por litro
mm:	milímetros
ml/hr:	mililitro por hora
ml/Kg:	mililitros por quilo
Kg/m² :	quilo por metro quadrado
SI3:	Sistema Integrado de Informações Clínicas do InCor

LISTA DE TABELAS

	Pag.
Tabela 1. Mortalidade por tipo de procedimento cirúrgico.....	46
Tabela 2A. Prevalência dos fatores de risco no Grupo Estudo, comparado com os fatores de risco da população do EuroSCORE.....	47
Tabela 2B. Prevalência dos fatores de risco no Grupo Estudo, comparado com os fatores de risco da população do 2000 Bernstein Parsonnet.....	48
Tabela 3A. Mortalidade observada e esperada mediante uso do 2000BP como variável preditora nos grupos definidos pelo teste Hosmer – Lemeshow.....	50
Tabela 3B. Porcentagens da mortalidade observada e estimada por grupos de risco do 2000BP.....	50
Tabela 4A. Mortalidade observada e esperada mediante uso do ES como variável preditora nos grupos definidos pelo teste <i>Hosmer – Lemeshow</i>	51
Tabela 4B. Porcentagens da mortalidade observada e estimada por grupos de risco do ES.....	52
Tabela 5. Área abaixo da curva ROC para o 2000 BP e o ES realizada em 3000 pacientes.	53
Tabela 6. Análise do <i>bootstrap</i> : ranking das variáveis para desenvolvimento do InsCor.....	55
Tabela 7. O InsCor.....	56
Tabela 8. Valores de probabilidade estimada de óbito para o InsCor.....	58
Tabela 9. Mortalidade observada e esperada mediante o uso do InsCor como variável preditora nos grupos definidos pelo teste Hosmer – Lemeshow.....	60
Tabela 10. Calibração do InsCor - Análise por grupos de risco.....	60
Tabela 11. Calibração do EuroSCORE - Análise por grupos de risco.....	61
Tabela 12. Calibração do 2000 BP - Análise por grupos de risco.....	61
Tabela 13. Área abaixo da curva ROC para o InsCor, o 2000 BP e o ES realizada em 1000 pacientes.....	62

LISTA DE FIGURAS

	Pag.
Figura 1. Curva ROC para o 2000BP e o ES na avaliação do poder de discriminação realizada em 3.000 pacientes.....	53
Figura 2. Equação do modelo aditivo do InsCor.....	57
Figura 3. Curva da probabilidade estimada de óbito segundo o InsCor.....	58
Figura 4. Regra mnemotécnica para cálculo do InsCor.....	58
Figura 5. Curva ROC para o InsCor,o 2000BP e o ES na avaliação do poder de discriminação realizada em 1.000 pacientes.....	62

Resumo

Mejia, OAV. **Predição de mortalidade em cirurgia de coronária e/ou valva no InCor: Validação de dois modelos externos e comparação com o modelo desenvolvido localmente** [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2012. 118p.

Objetivo: Novas tendências na avaliação de risco trazem evidências de que modelos externos recalibrados ou remodelados funcionam melhor localmente. O objetivo deste estudo foi validar dois modelos externos e formular um modelo local, comparando-os na predição de mortalidade nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP.

Método: Entre 2007 e 2009, 3.000 pacientes foram sequencialmente operados de coronária e/ou valva no InCor-HCFMUSP. No banco de dados, foi realizada a validação dos modelos 2000 Bernstein-Parsonnet (2000BP) e EuroSCORE (ES), mediante testes de calibração e discriminação. O InCor de 2.000 pacientes foi elaborado mediante a utilização de técnicas de *bootstrap*. Nos próximos 1.000 pacientes foi realizada a validação interna do modelo, e seu desempenho medido frente ao 2000BP e ES.

Resultados: Houve uma diferença significativa na prevalência dos fatores de risco entre as populações do estudo, ES e 2000BP. Na validação externa dos modelos, o ES apresentou uma boa calibração ($P=0,596$); no entanto, o 2000BP revelou-se inadequado ($P=0,047$). Na discriminação, a área abaixo da curva ROC revelou-se boa para ambos os modelos, ES (0,79) e 2000BP (0,80). Utilizando a técnica de *bootstrap*, 10 variáveis: idade >70 anos, sexo feminino, cirurgia associada, infarto do miocárdio <90 dias, reoperação, cirurgia da valva aórtica, cirurgia da valva

tricúspide, creatinina <2mg/dl, fração de ejeção <30% e estado pré-operatório crítico (eventos), foram selecionadas para formulação do InsCor. Na validação interna do InsCor, a calibração foi adequada, com $P=0,184$. Na discriminação, a área abaixo da curva ROC foi boa (0,79). Neste grupo, a área abaixo da curva ROC foi de 0,81 e 0,82 para o ES e 2000BP, respectivamente, mostrando-se apropriada para ambos os modelos.

Conclusões: O InsCor e o ES tiveram melhor desempenho que o 2000BP em todas as fases da validação; porém o novo modelo, além de se identificar com os fatores de risco locais, é mais simples e objetivo para a predição de mortalidade nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP.

Summary

Mejia, OAV. **Mortality prediction in coronary bypass surgery and/or heart valve surgery at InCor: Validation of two external risk models and comparison to the locally developed model (InsCor)**[thesis]. São Paulo: “Faculdade de Medicina, Universidade de São Paulo”; 2012. 118p.

Background: New trends in risk assessment bring evidence that recalibrated or remodeled external models work best locally. The aim of this study was to validate two external models and formulate a local model, comparing them to predict mortality in patients who underwent coronary bypass surgery and/or heart valve surgery at InCor-HCFMUSP.

Method

Between 2007 and 2009, 3.000 patients were sequentially operated to coronary bypass surgery and/or heart valve surgery at InCor-HCFMUSP. The database was assessment to validate the models 2000 Bernstein-Parsonnet (2000BP) and EuroSCORE (ES) through calibration and discrimination tests. The InsCor of 2,000 patients (2/3 of database) was elaborated using bootstrap techniques. Over the next 1000 patients (1/3 of database) the internal validation of the InsCor was performed and its performance compared against the 2000BP and ES.

Results: Significant difference in the prevalence of risk factors was found among the external and study populations ($P < 0,001$). In the external validation of these models, the ES showed good calibration ($P = 0.596$); however, 2000BP was inadequate ($P = 0.047$). In discrimination, the area under the ROC curve was good for both models, ES (0.79) and 2000BP (0.80). With the bootstrap technique, 10 variables: age > 70 years, female, CABG + valve surgery, myocardial infarction <90 days, reoperation, aortic valve surgery, tricuspid valve surgery, creatinine <2mg/dl, ejection fraction <30% and critical preoperative state (events) were chosen to formulate the InsCor. In the validation of InsCor, the calibration was appropriate with $P = 0.184$. In discrimination, the area under the ROC curve was good (0.79). In this group, the area under the ROC curve was 0.81 and 0.82 for ES and 2000BP, respectively, being suitable for both models.

Conclusions: The InsCor and ES outperformed the 2000BP at all stages of validation, but the new model, besides identifying itself with the local risk factors, is more simple and objective for the prediction of mortality in patients who underwent coronary bypass surgery and/or heart valve surgery at InCor-HCFMUSP.

1. Introdução

A utilização de indicadores de resultados vem se configurando como uma forma objetiva de mensurar a qualidade dos serviços de saúde. Até recentemente, a avaliação da qualidade da assistência se restringia à verificação da estrutura e do processo existentes e a conformidade com padrões pré-estabelecidos. Uma das justificativas para a introdução de indicadores de resultados é o fato de estes minimizarem eventuais vícios relacionados aos princípios eleitos como padrão para a assistência a ser oferecida.

Entretanto, mesmo indicadores globais de efetividade como a taxa de mortalidade hospitalar, podem não ser diretamente comparáveis, uma vez que expressam também diferenças não devidas ao atendimento como, por exemplo, o perfil de risco dos pacientes atendidos [1,2]. Portanto, para a utilização de indicadores de resultados, deve-se assegurar que estes tenham sido ajustados por estas diferenças [3].

Estas comparações são importantes para subsidiar decisões técnico-administrativas e também para avaliar o impacto de mudanças introduzidas nas práticas assistenciais [4].

Uma das áreas em que estes modelos de ajuste de risco mais se desenvolveram foi a cirurgia cardíaca, particularmente a cirurgia de revascularização do miocárdio [5,6], tendo em vista a alta frequência destes procedimentos e o custo elevado dos mesmos.

A existência de sistemas de predição e ajuste de risco em cirurgia cardíaca remonta ao *Collaborative Study in Coronary Artery Surgery* (CASS) [7], que,

mediante análises de discriminação multivariável, elaboraram um modelo de risco com a finalidade de prever a mortalidade cirúrgica.

Mas o que realmente desencadeou a criação e/ou utilização dos escores de risco foi a publicação, pela *Health Care Financing Administration* (HCFA), em março de 1986 [8], dos resultados brutos, sem ajuste ao risco, da mortalidade nos hospitais que operavam doentes do *Medicare*. Tratava-se de documentos de interesse governamental que apareceram em vários jornais a nível nacional, causando efeitos adversos na maioria dos programas de cirurgia cardiovascular dos Estados Unidos. Sem estratificação de risco, cirurgiões e hospitais que tratam pacientes de alto risco aparentam possuir piores resultados que outros. Este fato pode prejudicar os sistemas de referência, a alocação de recursos e desencorajar o tratamento de pacientes de alto-risco, o que é especialmente indesejável em cirurgia cardíaca, tendo em vista que é precisamente este grupo de pacientes que mais se beneficia com o tratamento cirúrgico, apesar do risco aumentado [9,10]. Nenhum sistema de predição será suficientemente abrangente para estimar resultado específico para cada paciente. A estratificação de risco, porém, informa a pacientes e profissionais sobre o provável risco de complicações ou óbito para o grupo de indivíduos com perfil de risco similar, submetidos ao procedimento proposto [11]. A estruturação e o desenvolvimento de um escore de risco adequado para avaliar a contribuição dos fatores intrínsecos do paciente nos resultados requer a consideração de vários aspectos: evento resposta de interesse, tempo de observação para mensuração dos eventos selecionados, unidade de análise, fontes de dados, recursos humanos e financeiros necessários para coleta dos dados, perfil do profissional responsável pela coleta, definições padronizadas dos dados, coleta e consolidação dos dados,

checagem, técnicas analíticas, método de validação do modelo, tratamento dos dados ausentes [6,12].

Argumenta-se que a transição de um banco de dados para um sistema de escore sacrifica a precisão em busca da simplicidade [11]. Há dois extremos na seleção de um sistema de estratificação de risco. Pode-se obter precisão incluindo na avaliação do paciente um grande número de fatores de risco e comparando resultados em um grande banco de dados. Porém, este método requer grandes populações de pacientes e análises estatísticas complexas, não sendo de fácil aplicação no dia a dia e sem tecnologia de informação disponível. Por outro lado, modelos simples, contendo poucas variáveis, são de maior aplicabilidade, desde que contenham os fatores de risco significativos e que tenham um grau de predição adequado, por serem mais fáceis de serem incorporados na assistência ao paciente [3,13,14]

Atualmente a utilização de escores de risco na tomada de decisões em cirurgia de revascularização miocárdica é considerada recomendação IIa com grau de evidencia B [15]. No entanto, em razão do aumento da proporção de procedimentos sobre a valva incluindo a terapia transcater e as cirurgias minimamente invasivas, a aplicabilidade e adaptação destes modelos tornam-se necessárias. Portanto, existem evidências científicas que nos obrigam a utilizar estes modelos na cirurgia cardíaca moderna.

No entanto, para serem utilizados os modelos de risco precisam ser validados, o que estatisticamente significaria avaliar o desempenho em dois aspectos: calibração e discriminação [16]. Estas duas fases são obrigatoriamente sequenciais para uma adequada análise dos modelos. A calibração visa avaliar a capacidade para

a predição da mortalidade global e por grupos de risco. A discriminação é uma medida da capacidade de predizer bem um resultado específico [17].

Populações de pacientes e os fatores de risco associados a resultados adversos mudam com o tempo e diferem entre centros. Por tanto, modelos de risco derivados e validados em um local usualmente têm menor desempenho quando aplicados em outro local e inclusive no mesmo local ao longo do tempo. Médicos e gestores de saúde concordam que, para escolher um sistema de avaliação de risco, temos três opções básicas [18]:

- 1º. Utilizar um modelo externo, sabendo que os fatores de risco identificados ou a quantificação dada a eles, pode não ser ideal para a população em estudo (*ready-made*).
- 2º. Aceitar os fatores de risco utilizados pelo escore, mas ajustando o peso das variáveis aos da população em estudo através da recalibração das variáveis do modelo (recalibração).
- 3º. Desenvolver um modelo local a partir dos próprios dados (remodelamento).

No Brasil, nenhum escore preditor de mortalidade em cirurgia cardíaca foi adequadamente validado [13,16], embora vários já sejam utilizados. O desconhecimento do perfil de risco dos pacientes submetidos à cirurgia cardíaca no Brasil torna necessária a validação preditiva de escores externos ou a derivação de modelo local fortemente estruturado para analisar nossos pacientes e estratificá-los segundo a morbidade e complexidade do procedimento.

Sabemos que existem realidades diferentes entre as populações de origem dos modelos e os locais onde eles são aplicados. No Brasil, por exemplo, existem

diferenças nas características dos pacientes, na apresentação clínica devido a razões socio-econômicas, culturais e geográficas, na distribuição desigual de instalações médicas e na alta endemicidade de inflamação subclínica, infecção e doença reumática. Diferenças estas, que poderiam dificultar a aplicabilidade de um escore de risco. Por isso, é necessário ter muita cautela na escolha do modelo ideal, ou reforçar a ideia da criação de um escore de risco local. Assim, na escolha do modelo para analisar os resultados dos pacientes operados na instituição, foram exigidos os seguintes requisitos [19-21]:

1. Deverá permitir a comparação de resultados do atendimento em cenário nacional e internacional. As discussões do modelo em publicações científicas foram estipuladas como critério.
2. Deverá ser relevante para a população brasileira. Isso significa que o modelo já deve ter sido utilizado como preditor de risco em estudos clínicos nacionais.
3. Deverá ser relevante para os subgrupos analisados. Assim, o modelo deve predizer adequadamente por tipo de procedimento realizado, por sexo, por idade e outros.
4. A população de origem deverá ser similar à população de estudo com relação à faixa etária (adultos a partir de 18 anos) e ao motivo de intervenção (doença adquirida). Modelos pediátricos e escores específicos (escores congênitos) devem ser excluídos.
5. Deverá calcular a mortalidade esperada (mortalidade ajustada ao risco).

6. A predição do risco individual será uma característica secundária na escolha do modelo. A supra ou infraestimação do modelo em populações de estudo ou em casos isolados não será motivo de rejeição.
7. O número de variáveis a serem estudadas deverá ser o suficiente para não piorar a qualidade preditiva do modelo.
8. Deverá haver uma definição exata no que diz respeito ao parâmetro se ele é positivo ou negativo, ou à classificação da variável. Definições subjetivas e complexas provocam registros difíceis e incrementam o custo.
9. As variáveis escolhidas deverão estar disponíveis para todos os pacientes, cirurgias e laboratórios.
10. O cálculo da mortalidade esperada e a modelagem do escore deverão ser públicas e isentas de licenças.

Após a análise discriminativa na escolha dos modelos, os escores selecionados que mais se aproximaram à realidade brasileira foram: EuroSCORE (ES) [11], *Parsonnet initial* [22], 2000 Bernstein Parsonnet (2000BP) [23], *French score* [24], *Ontario Province Risk* (OPR) [25], *New York State Health Department* (NYSHD) [26] e *Society of Thoracic Surgeons* (STS) [27]. O STS foi excluído em razão de a licença não ter sido liberada. O NYSHD foi excluído por causa da variabilidade cronológica do sistema. O OPR e o *French score*, embora tenham apresentado resultados satisfatórios, foram excluídos por apresentarem escassas publicações e participações em estudos clínicos. O EuroSCORE, o primeiro escolhido, alcançou mais destaque porque validou satisfatoriamente o banco de

dados do STS [28], tornando-se o mais utilizado em ambos os lados do Atlântico. Ainda assim, Roque e cols. [29] já haviam demonstrado a aplicabilidade do EuroSCORE em diferentes países da Europa, concluindo que, embora haja diferenças epidemiológicas entre os países europeus, o poder de discriminação do EuroSCORE foi adequado na Espanha e excelente nos outros países. Ademais, muitas publicações favoráveis ao modelo do Parsonnet, inclusive na predição de morbidade, optaram pelo modelo atualizado. Assim, devido à sua abrangência dado o número de variáveis (41 variáveis), o 2000BP seria o modelo ideal (específico e discriminativo) na predição do risco individual.

Sabemos que o número de variáveis contidas no modelo escolhido é controverso. A prática clínica habitual leva à ideia intuitiva de que quanto mais variável, melhor o ajuste à realidade. Isso pode ser verdade, quando são utilizados modelos mais voltados para estimativas individuais (2000BP). Para predições sobre séries globais, um número relativamente pequeno de variáveis essenciais ou nucleares pode construir modelos estatisticamente eficazes, sendo que novas variáveis incrementariam de maneira marginal a capacidade preditiva (ES) [30]. Os parâmetros básicos para a escolha de um sistema de avaliação de risco foram analisados levando em consideração as características e compatibilidade da população local. Por tanto, os modelos 2000 Bernstein Parsonnet [23] e EuroSCORE [11] foram escolhidos para esta análise.

2. Objetivos

- Validar os modelos externos 2000 Bernstein-Parsonnet e EuroSCORE na predição de mortalidade em cirurgia de coronária e/ou valva no InCor.

- Desenvolver o modelo local (InsCor), validá-lo e compará-lo ao 2000 Bernstein Parsonnet e EuroSCORE na predição de mortalidade em cirurgia de coronária e/ou valva no InCor.

3. Revisão da Literatura

3.1 Origem dos escores de risco

Uma das primeiras publicações na área da predição de risco em cirurgia cardíaca talvez tenha sido realizada em 1975, quando Loop e col. [31] publicaram nos Estados Unidos uma análise feita em 1.188 sobreviventes e 60 pacientes que morreram após CRM. Foram identificadas quatro variáveis (cardiomegalia importante, insuficiência cardíaca congestiva descompensada, doença de três vasos e/ou obstrução do tronco da coronária esquerda e pressão diastólica final do ventrículo esquerdo) como fatores de risco independentes. Coeficientes, erro standard e exemplo dos cálculos realizados não foram relatados.

De fato, a existência de sistemas de predição e ajuste de risco em cirurgia cardíaca remonta ao *Collaborative Study in Coronary Artery Surgery* (CASS) [7], que mediante análises de discriminação multivariável (dados clínicos e angiográficos) formularam um modelo de risco para predizer mortalidade cirúrgica na CRM. As variáveis foram idade, estenose do tronco da coronária esquerda $\geq 90\%$, sexo feminino, escore de movimento da parede do ventrículo esquerdo e presença de rales. A calibração foi realizada mediante a comparação da mortalidade observada e a mortalidade estimada.

Em 1983, Paiement e col. [32] publicaram no Instituto do Coração de Montreal o que seria o primeiro índice de predição em cirurgia cardíaca (incluindo cirurgia valvar). Foram identificados oito fatores de risco (fração de ejeção baixa, insuficiência cardíaca congestiva, angina instável ou infarto de miocárdio recente, idade >65 anos, obesidade severa, reoperação, cirurgia de emergência e distúrbios

sistêmicos descompensados). O modelo foi revisado em 1992 por Tremblay e col. [33] com a inclusão de um novo fator de risco: cirurgia complexa.

Em 1986, um relatório da *Health Care Financing Administration* (HCFA) publicou, em março de 1986 [8], os resultados brutos, sem ajuste ao risco, da mortalidade nos hospitais que operavam doentes do *Medicare*. Segundo o relato, somente três hospitais (8,3%) com mortalidade observada menor que a mortalidade esperada ($MOME < 1$) deveriam continuar realizando cirurgia de revascularização miocárdica.

Em março de 1988, a Sociedade Americana de Cirurgiões Torácicos (STS) publicou um relatório, o *Ad Hoc Committee on Risk Factors for Coronary Bypass Surgery* [34], que considerou inapropriado e equívoco, o uso da mortalidade sem o adequado ajuste aos fatores de risco. Caso não houvesse, tal ajuste, isso poderia provocar alterações na seleção de pacientes com indicação cirúrgica, afetando principalmente os pacientes de alto risco.

A partir disso, múltiplos escores de risco foram desenvolvidos e orientados a ponderar os resultados em função da gravidade da enfermidade e das comorbidades.

Em 1988, Edwards e col. [35] desenvolveram nos Estados Unidos um modelo de predição através de análises bayesianas em 300 pacientes operados de CRM isolada. Na fase de validação, houve uma associação positiva entre mortalidade observada e mortalidade esperada. Os algoritmos bayesianos requeridos para o cálculo não foram revelados. Em 1989, o mesmo grupo publicou um modelo bayesiano derivado do registro do CASS [36]. As variáveis selecionadas foram idade, sexo, função ventricular, infarto de miocárdio prévio, extensão da doença

coronariana, angina instável e prioridade da cirurgia. O trabalho não descreveu a equação utilizada para calcular a mortalidade peri-operatória.

3.2 Formulação dos escores de risco

Em 1989, Parsonnet e Bernstein [22] analisaram dados retrospectivos de 3.500 pacientes operados consecutivamente de CRM, reparo ou substituição da valva mitral e/ou reparo ou substituição da valva aórtica. Foram desenvolvidos dois modelos de predição: um aditivo utilizando análises de regressão múltipla e outro utilizando regressão logística. O modelo de regressão logística foi constituído de 17 variáveis: idade, doença valvar aórtica, CRM isolada, CRM + outro procedimento, colesterol elevado, diabetes, estado catastrófico (ex. falha estrutural aguda), história familiar de doença coronariana, sexo feminino, hipertensão, aneurisma de ventrículo esquerdo, fração de ejeção, doença de valva mitral, obesidade, balão intra-aórtico pre-operatório, reoperação e tabagismo. O modelo aditivo constituído de 14 variáveis foi sexo feminino, obesidade mórbida, diabetes (sem especificação do tipo), hipertensão, fração de ejeção, reoperação, balão intra-aórtico pre-operatório, aneurisma ventricular esquerdo, CRM de emergência após complicações da angioplastia ou cateterização, dependência de diálise, outras raras circunstâncias (ex., paraplegia, dependência de marcapasso, doença congênita no adulto, asma severa), doença valvar mitral, pressão sistólica da artéria pulmonar $\geq 60\text{mmHg}$ e gradiente valvar aórtico $> 120\text{mmHg}$. Estes modelos de predição foram validados com sucesso em registros independentes.

Em 1990, Hannan e col. [26], utilizando o *New York State's new Cardiac Surgery Reporting System*, analisaram 7.596 pacientes operados de cirurgia cardíaca no estado de Nova York. Um modelo com variáveis pré-operatórias foi desenvolvido para avaliar a mortalidade intra-hospitalar. Este modelo logístico, similar ao modelo do Parsonnet, incluiu: idade, sexo feminino, reoperação, fração de ejeção, infarto de miocárdio prévio (<7dias), diabetes, dependência de diálise, estado catastrófico (defeito estrutural agudo, falha renal, choque cardiogênico), angina instável, insuficiência cardíaca congestiva, lesão de tronco da coronária esquerda > 90%, cirurgia valvar e outro tipo de cirurgia cardíaca (não CRM, não cirurgia valvar). Nenhum teste estatístico de validação foi mostrado na publicação.

Em 1992, Higgins e col. [37] publicaram nos Estados Unidos o que seria chamado de *Cleveland Clinic Score*, um modelo para avaliar a mortalidade e a morbidade. O modelo foi desenvolvido a partir de 5.051 pacientes operados consecutivamente de CRM. Um modelo aditivo único constituído por 13 variáveis para avaliar a morbidade e a mortalidade foi elaborado. Um modelo logístico com 9 variáveis para avaliar a mortalidade (emergência, creatinina > 168 $\mu\text{mol/L}$, fração de ejeção < 30%, hematócrito pre-operatório > 34%, idade, doença pulmonar obstrutiva crônica, cirurgia vascular prévia, reoperação e insuficiência da valva mitral) e outro de 13, para avaliar a morbidade também foram testados. No grupo de desenvolvimento, para morbidade, a área abaixo da curva ROC foi de 0,74 com ambos os modelos, e na mortalidade foi de 0,84 com o modelo logístico e de 0,83 com o modelo aditivo. No grupo da validação foram realizados unicamente testes de calibração, com os seguintes resultados: inadequado para morbidade ($P < 0,001$) e bom para mortalidade ($P = 0,3$).

Em 1992, O'Connor e col. [38] (*Northern New England Cardiovascular Disease Study Group*) publicaram nos Estados Unidos um estudo regional e prospectivo, um modelo logístico com oito variáveis (idade, sexo, área de superfície corpórea, presença de co-morbidades, CRM prévia, pressão diastólica final do ventrículo esquerdo, fracção de ejeção e prioridade da cirurgia) para predizer a mortalidade intra-hospitalar. No grupo de desenvolvimento, foram coletados dados de 3.055 pacientes operados de CRM isolada. No grupo de validação, 1.516 pacientes foram avaliados. Na calibração, o teste de Hosmer-Lemeshow revelou-se adequado, com $P=0,689$. Na discriminação, a área abaixo da curva ROC foi de 0,76.

Uns dois primeiros trabalhos sobre validação foi realizado por Nashef e col. [39] em 1992. Utilizando o Parsonnet inicial, 1.071 pacientes operados no hospital de Wythenshawe (Cambridge) foram avaliados. Na calibração, a comparação entre a mortalidade observada e a mortalidade esperada foi adequada. Os testes de discriminação não foram realizados, mesmo assim, o modelo foi generalizado para toda Inglaterra.

Tuman e col. Em 1992 [40] identificaram preditores de complicações, mortalidade pós-cirurgia cardíaca e tempo de permanência em terapia intensiva. Com base na “*Odds ratio*” e no grau de significância no modelo logístico para complicações, cada fator de risco recebeu um escore de 1 a 4, produzindo um escore de risco clínico simplificado, com valor máximo de 22. Foram estabelecidas três categorias de risco: baixa (escore de 0 a 5), moderada (escore de 6 a 9) e alta (escore ≥ 10).

Grover e col. em 1994 [41] avaliaram 43 instituições (*Veterans Affairs hospitals*) e desenvolveram modelos de estratificação de risco específicos para CRM,

com base em uma população de 8.552 pacientes. Para cirurgia valvar e outros procedimentos cardíacos, foram avaliados 2.104 pacientes.

A “*Society of Thoracic Surgeons*” [42] desenvolveu um modelo de predição de risco com base em mais de 80.000 pacientes submetidos à CRM. Treze fatores de risco foram identificados como preditores de mortalidade e foram definidos cinco grupos de taxas de mortalidade esperadas: 0 a 5%; 6 a 10%; 10 a 20%; 20 a 30% e 30 a 50%.

Em 1995, Tu e col. [25] publicaram no Canadá o *Ontario Province Score* para prever mortalidade, tempo de permanência na UTI e tempo de permanência hospitalar após cirurgia cardíaca. No grupo de desenvolvimento, 6213 pacientes foram operados de cirurgia cardíaca em nove hospitais de Ontário. Análises de regressão logística dos três desfechos deram origem a um modelo de seis variáveis (idade, sexo, fração de ejeção, tipo de cirurgia, urgência da cirurgia e reoperação). A validação temporal foi realizada em 6.885 pacientes. A área abaixo da curva ROC foi de 0,75 para mortalidade intra-hospitalar, 0,66 para tempo ≥ 6 dias de permanência na UTI e 0,69 para tempo ≥ 17 dias de permanência hospitalar.

Turner e col. em 1995 [43], em um estudo prospectivo envolvendo apenas uma instituição e 1.008 pacientes submetidos à cirurgia cardíaca, analisaram a combinação de variáveis pré-operatórias, intra-operatórias e pós-operatórias (incluindo as variáveis avaliadas por sistemas gerais de gravidade: APACHE II e III) das primeiras 24 horas de tratamento intensivo pós-operatório, para predição de mortalidade e permanência em unidade de terapia intensiva para um tempo > 24 horas.

Em 1996, o modelo Parsonnet sofreu a primeira modificação [44], o número de suas variáveis aumentou, chegando a 44, o modelo ganhou subjetividade e todas as variáveis foram recalibradas, embora ele continuasse sendo o mais utilizado na época.

Nesse mesmo ano, quando a cirurgia de revascularização miocárdica (CRM) era a cirurgia de grande porte mais realizada nos Estados Unidos e a mortalidade cirúrgica, o primeiro critério para julgar a qualidade dos resultados cirúrgicos, Magovern e col. [45], criaram um modelo para predição de morbidade e mortalidade a partir da análise de 1.567 pacientes operados consecutivamente de CRM. As variáveis listadas em ordem decrescente de importância foram: choque cardiogênico, operação de emergência, cateterismo de resgate, disfunção ventricular esquerda grave, aumento da idade, cardiomegalia, doença vascular periférica, insuficiência renal crônica, diabetes mellitus, baixo índice de massa corporal, sexo feminino, reoperação, anemia, doença cerebrovascular, doença pulmonar obstrutiva crônica, disfunção renal, albumina baixa, ureia elevada no sangue, insuficiência cardíaca congestiva e arritmias atriais. A validação foi adequada, na calibração, a proporção entre MOME para morbidade foi de 0,98, e para mortalidade de 0,97. A área abaixo da curva ROC foi 0,82 para morbidade e 0,86 para mortalidade. No entanto, a maior dificuldade, na época, foi uniformizar as definições, dadas as variáveis preditoras de mortalidade cirúrgica na CRM. Assim, foi criado o projeto chamado *Working Group Panel on The Cooperative CABG Database Project* [3], cujo propósito era desenvolver um sistema prático e padronizado para avaliar mortalidade na CRM.

Em 1996, Pons e col. [46] publicaram em Barcelona um modelo derivado de uma análise prospectiva e pré-operatória de 1.309 pacientes operados

consecutivamente em sete hospitais. A randomização da amostra fez com que 70% dos pacientes fizessem parte do grupo do desenvolvimento do modelo constituído de 14 variáveis: idade, infarto de miocárdio recente, classe funcional, doença hepática, presença de aneurisma do ventrículo esquerdo, creatinina $\geq 1,5$ mg/dl, presença de choque cardiogênico, ventilação mecânica pré-operatória, prioridade da cirurgia, reoperação, cirurgia da valva mitral, cirurgia da valva tricúspide, cirurgia da aorta torácica e associação de CRM + valva. No grupo da validação, apenas a calibração (P=0,034) foi realizada, embora tenha apresentado bons resultados em todos os subgrupos.

Uns dos primeiros trabalhos de validação externa do escore Parsonnet foi o realizado por Gabrielle e col. na França [47], onde 6.649 pacientes consecutivos foram analisados pelo modelo inicial e pelo modificado. Ambos se revelaram inadequados, com área abaixo da curva ROC de 0,64 para o Parsonnet inicial, e 0,70 para o Parsonnet modificado.

Até 1996, a Sociedade de Cirurgiões Torácicos dos Estados Unidos fornecia, para a predição de mortalidade na CRM, um software que continha um modelo bayesiano. A partir do *STS National Cardiac Surgery Database*, Edwards e col. [27] conseguiram criar um modelo de regressão logística, que é atualizado anualmente, utilizando dados de 332.604 pacientes operados de CRM.

Um trabalho realizado em 1999, por Martinez-Alario e col. [48], comparou o modelo do Parsonnet inicial com sistemas de avaliação de risco em pacientes críticos, tais como: *Acute Physiology and Chronic Health Evaluation* (APACHE I, II e III), *Simplified Acute Physiology Score* (SAPS I e II) and *Mortality Probability Models* (MPM I e II). A calibração foi adequada para todos os modelos, e mostrou

que mesmo que o modelo inicial do Parsonnet seja coletado somente no pré-operatório, ele consegue ter uma área abaixo da curva ROC de 0,857, em comparação com as seguintes áreas: 0,783 de MPM I; 0,796 de MPM II; 0,771 de SAPS II e 0,803 do APACHE II.

Em 1999, Nashef e col. [11], usaram o EuroSCORE *Project Database*, constituído de 14.781 pacientes, para elaborar um sistema de estratificação de risco que auxilie no controle da qualidade da cirurgia cardíaca, o que se tornaria uma das publicações mais importantes para monitorização dos resultados na cirurgia cardíaca moderna. Nasce assim o EuroSCORE, constituído de 17 variáveis (idade, sexo, doença pulmonar obstrutiva crônica, arteriopatía extracardíaca, disfunção neurológica, reoperação, creatinina sérica > 200 umol/L, endocardite ativa, estado pré-operatório crítico, angina instável, disfunção ventricular, infarto de miocárdio recente, hipertensão pulmonar, emergência, CRM associada, cirurgia na aorta torácica e rotura septal posinfarto) e com uma área abaixo da curva ROC de 0,76 no grupo da validação. Continha um sistema aditivo e classificado em três grupos: baixo (0-2), médio (3-5) e alto (≥ 6). Este score foi concebido para ser usado, em princípio, pela população europeia; devido à facilidade de seu uso e aplicabilidade no papel ou através de softwares, foi rapidamente disseminado.

Um dos primeiros trabalhos que compararam vários modelos de risco foi publicado na Alemanha em 2000, quando Geissler e col. [49] analisaram 504 pacientes operados de CRM utilizando seis escores: Parsonnet, Cleveland Clinic, French, EuroSCORE, Pons e Ontario Province Score. Os melhores resultados foram para o EuroSCORE e Parsonnet, com áreas abaixo das curvas ROC de 0,79 e 0,76, respectivamente.

Finalmente, no ano 2000 o escore Parsonnet foi modificado pela última vez, Bernstein e Parsonnet [23], atualizaram o modelo incluindo pacientes operados de coronária, valva e cirurgia associada. Tratava-se de um modelo com 47 variáveis derivadas da análise de 10.703 pacientes operados em 10 hospitais de Nova Jersey, o 2000 Bernstein Parsonnet. Os resultados deste estudo mostraram que o uso da regressão logística, seguido da translação para um formato simples, utilizando papel e caneta, proporcionaria um instrumento potencial na tomada de decisões em cirurgia cardíaca.

3.3 Validação dos escores de risco

Assim, pela simplicidade e popularidade, o modelo inicial do Parsonnet foi escolhido para comparar seu desempenho frente ao EuroSCORE. Desta forma, Kawachi e col. [50] em 2001, publicaram no Japão aquele que viria a ser uma das primeiras comparações dos modelos mais populares. Foram analisados 803 pacientes consecutivos operados de coronária, valva e aorta. Na validação, as áreas abaixo das curvas ROC foram de 0,72 para o Parsonnet e 0,82 para o EuroSCORE.

Uma das análises mais desafiadoras para o modelo EuroSCORE foi a realizada por Nashef e col., em 2002 [28], no banco de dados do STS nos Estados Unidos. Um total de 188.913 pacientes operados em 1995 e de 401.684 no período 1998-1999 de CRM e cirurgia valvar foram submetidos à análise para a validação do modelo. Na validação, os testes de calibração foram adequados e a área abaixo da curva ROC foi de 0,78. Houve diferença nas prevalências dos fatores de risco em

ambas as populações ($P < 0,001$). O EuroSCORE foi objeto de vários estudos. Vários observadores notaram uma tendência à subestimação do risco operatório nos pacientes de altíssimo risco com o EuroSCORE aditivo [51]. Dessa forma, Nashef e col. [52] relataram, em 2003, a equação da regressão logística do EuroSCORE. Utilizando as mesmas variáveis do modelo aditivo, o cálculo do escore seria feito com uma calculadora na página do EuroSCORE na internet (<http://www.euroscore.org>).

No início, a reação da comunidade foi favorável a começar utilizar o EuroSCORE logístico, mas houve grande controvérsia; todos sabiam que se trocava um modelo simples por um mais complexo, mantendo as mesmas variáveis.

Em 2003, Nilsson e col. [53] analisaram, na Suécia, com o EuroSCORE aditivo e o *STS score*, 4.497 pacientes operados de CRM isolada entre 1996 e 2001. As áreas abaixo da curva ROC foram de 0,84 para o EuroSCORE e 0,71 para o *STS score* ($p < 0,001$).

Um trabalho publicado na Arábia Saudita em 2004, por Syed [54], comparou o EuroSCORE aditivo com o modelo inicial do Parsonnet em 194 pacientes operados de cirurgia cardíaca. As áreas abaixo da curva ROC foram de 0,77 para o EuroSCORE e 0,69 para o modelo inicial do Parsonnet. Mas o tamanho da amostra, com apenas 13 óbitos, perde em credibilidade e em força estatística.

Em 2004, uma avaliação feita em Taiwan por Chen e col. [55], utilizou o EuroSCORE aditivo e logístico em 801 pacientes consecutivos operados de CRM. Com pouco mais de 80 óbitos, as áreas abaixo da curva ROC foram de 0,75 para o EuroSCORE aditivo e 0,74 para o EuroSCORE logístico.

Neste mesmo ano, mediante o registro prospectivo do *Northern New England Cardiovascular Disease Study Group*, Nowicki e col. [56] construíram um modelo aditivo para pacientes operados de valva mitral, e outro para operados de valva aórtica. No modelo aórtico, a correlação entre mortalidade observada e mortalidade esperada foi de 0,98, o teste de Hosmer Lemeshow obteve um $P=0,157$ e a área abaixo da curva ROC foi de 0,75. No modelo mitral, a correlação entre mortalidade observada e mortalidade esperada foi de 0,99, o teste de Hosmer Lemeshow obteve um $P=0,704$ e a área abaixo da curva ROC foi de 0,79. Assim, após a aceitação dos escores na CRM, começaram a surgir novos modelos criados especificamente para pacientes valvares.

Em 2005, Ambler e col. [57] publicaram em Londres um escore específico para pacientes operados de valva aórtica e /ou mitral com ou sem CRM associada. Foi feita uma análise, no período de 1995 a 2003, com 32.839 pacientes do banco de dados voluntário da Sociedade Britânica e Irlandesa de Cirurgia Cardiorácica. Na fase de validação, a área abaixo da curva ROC de 0,77 demonstrou um razoável poder discriminativo.

As comparações entre o EuroSCORE aditivo e EuroSCORE logístico continuaram sendo feitas. Em 2005, trabalhos feitos por Shanmugam e col. [58] e Jin e col. [59] em grandes populações confirmaram que o EuroSCORE logístico não apresentou vantagem significativa sobre o EuroSCORE aditivo, exceto nos pacientes de alto risco.

Em 2005, Seccareccia e col. [60] publicaram, na Itália, um modelo derivado do registro nacional voluntário para estudo da CRM, que estimava a mortalidade hospitalar com a finalidade de comparar resultados entre populações e hospitais. Este

modelo derivado de 17.231 pacientes operados de CRM isolada em 64 centros inclui como variáveis preditoras: idade, sexo feminino, choque cardiogênico, diabetes, diálise, hipertensão pulmonar, arritmia ventricular maligna, doença pulmonar obstrutiva crônica, creatinina sérica > 2mg/dl, arteriopatia extracardíaca, angina instável, CRM prévia, emergência e fração de ejeção. A fase de validação foi realizada em um grupo independente de 17.079 pacientes, com área abaixo da curva ROC de 0,80 e o teste de Hosmer-Lemeshow com P=0,354.

Por outro lado, em 2006, Hannan e col. [61] apresentaram um modelo logístico derivado de 16.120 pacientes operados de CRM isolada procedentes do *New York State Cardiac Surgery Reporting System*. Este novo modelo com 10 variáveis (idade, sexo feminino, estado hemodinâmico, fração de ejeção, infarto de miocárdio prévio, doença pulmonar obstrutiva crônica, aorta ascendente calcificada, doença arterial periférica, falha renal e reoperação) serviria para predizer a mortalidade intra-hospitalar. A validação temporal foi realizada em 14.692 pacientes operados de CRM isolada. Na discriminação, a área abaixo da curva ROC foi de 0,823. Na calibração, o teste de Hosmer-Lemeshow apresentou um P=0,47.

No mesmo ano, Nilsson e col. [62], utilizando o banco de dados da Universidade de Lund na Suécia, analisaram 6.222 pacientes submetidos à cirurgia cardíaca. Um total de 19 escores de risco foram analisados nos pacientes submetidos à cirurgia cardíaca. A maior área abaixo da curva ROC em 30 dias e 1 ano foi do EuroSCORE logístico (0,84 e 0,77) e do EuroSCORE aditivo (0,84 e 0,77), seguidos pelo *Cleveland Clinic* (0,82 e 0,76) e o *Magovern* (0,82 e 0,76). No subgrupo da CRM isolada, a maior área abaixo da curva ROC em 30 dias e 1 ano foi do EuroSCORE logístico (0,86 e 0,75), EuroSCORE aditivo (0,85 e 0,75), *New York*

State (0,84 e 0,75), *Cleveland Clinic* (0,84 e 0,75) e Parsonnet modificado (0,84 e 0,73). Nesta análise, ficou demonstrado que, em registros conhecidamente organizados e sistemáticos, nos quais são computados todos os pacientes e todas as variáveis, há um melhor desempenho dos modelos mais bem desenhados.

Um dos trabalhos mais importantes na avaliação e comparação do 2000 Bernstein-Parsonnet e EuroSCORE, foi o publicado por Berman e col. [63], em 2006. Estes modelos foram aplicados em 1.639 pacientes operados consecutivamente de coronária, valva e cirurgia associada, no *Rabin Medical Center* em Israel. A área abaixo da curva ROC foi de 0,83 para o 2000 Bernstein-Parsonnet e de 0,73 para o EuroSCORE.

Um dos primeiros trabalhos realizados no Brasil, em Pernambuco, é de autoria de Moraes, que em 2006 [64] avaliou retrospectivamente a aplicabilidade do EuroSCORE aditivo em 752 pacientes operados de CRM. Com apenas 13 óbitos, a amostra teve uma área abaixo da curva ROC de 0,70.

Neste mesmo ano, Yap e col. [65] publicaram a validação do EuroSCORE logístico e aditivo em 8.331 pacientes consecutivos submetidos à cirurgia cardíaca em seis hospitais da Austrália. O banco de dados da Sociedade Australiana de Cirurgiões Torácicos e Cardíacos foi analisado retrospectivamente, e as definições de seis variáveis não foram idênticas às do EuroSCORE. A calibração para ambos os modelos foi inadequada, superestimando a mortalidade observada. Mesmo assim, a área abaixo da curva ROC foi de 0,83 para ambos os modelos e de 0,82 no subgrupo de CRM isolada também para ambos os modelos. No entanto, a discriminação foi adequada.

Um estudo sobre um banco de dados obrigatório, mas administrativo (DataSUS), que analisou a mortalidade global ajustada a quatro variáveis (idade, sexo, tipo de cirurgia e volume cirúrgico), foi realizado no Brasil em 2006, por Ribeiro [66]. Um total de 115.021 cirurgias cardíacas foram realizadas, com 9.211 óbitos (8%). Fato semelhante aconteceu quando o Medicare publicou os dados de mortalidade não ajustada ao perfil de risco dos pacientes nos Estados Unidos, em 1986.

Em 2006, Ivanov e col. [67] publicaram, no Canadá, o *Toronto Risk Score*, com a finalidade de prever eventos adversos (morte, infarto de miocárdio, síndrome de baixo débito, falha renal pós-operatória, acidente cerebrovascular e infecção profunda da ferida operatória) após a cirurgia cardíaca. A análise de regressão logística foi realizada em 12.683 pacientes operados de cirurgia cardíaca e um modelo aditivo foi construído. A fase de validação foi realizada em 4.378 pacientes operados em 2001. A área abaixo da curva ROC foi de 0,734 e o teste de Hosmer-Lemeshow teve um $P=0,9$.

Em 2006, Zhou e col. [68] relataram, na China, que 2.251 pacientes operados de CRM foram prospectivamente avaliados pelos modelos Parsonnet modificado, EuroSCORE, *Cleveland Clinic* e *Ontario Providence Registre*. As áreas abaixo da curva ROC para o Parsonnet modificado, *Cleveland Clinic*, *OPR* e EuroSCORE foram de 0,71; 0,75; 0,76 e 0,81, respectivamente.

Em 2007, De Varennes e col. [69] aplicaram, no Quebec, o modelo 2000 Bernstein Parsonnet em 4.883 pacientes consecutivos (eletivos ou emergência) submetidos à cirurgia cardíaca no período 2000-2005. A calibração foi realizada mediante a técnica de razão de mortalidade padronizada (RMP), calculada como a

razão entre a mortalidade observada (MO) e a mortalidade esperada (ME). Assim, $RMP=1$ indica que não existem diferenças entre MO e ME; $RMP<1$ indica que a qualidade do atendimento é superior ao estabelecido, e $RMP>1$ que o atendimento é inferior à qualidade do atendimento. Na análise, a RMP geral foi de 0,52 ($P=0,01$) e a mortalidade observada foi significativamente menor que a mortalidade esperada em todas as categorias. Os testes de discriminação não foram realizados.

Alguns especialistas na área dos escores de risco recomendam o uso de modelos mais específicos, por tipo de doença e tipo de procedimento, embora esta não seja a finalidade da origem e popularização dos modelos de risco.

Hannan e col. [70] publicaram, em 2007, um índice preditor de mortalidade intra-hospitalar orientado para pacientes operados de cirurgia valvar. Foi elaborado um modelo com 10.702 pacientes submetidos à cirurgia valvar isolada que incluía: tipo de cirurgia valvar, idade, sexo feminino, estado hemodinâmico, infarto de miocárdio < 14 dias, doença pulmonar obstrutiva crônica, endocardites, falha hepática, doença arterial periférica, falha renal e reoperação; e outro com 8.823 pacientes submetidos à cirurgia valvar e CRM composto de: tipo de cirurgia valvar/CRM, idade, sexo feminino, fração de ejeção, infarto de miocárdio < 7 dias, estado hemodinâmico, doença cerebrovascular, diabetes, endocardites, aorta ascendente calcificada, falha renal e reoperação. A fase de validação foi realizada no banco de dados de Nova York; a área abaixo da curva ROC foi de 0,794 para o modelo valvar e 0,750 para o modelo valvar + CRM. Na calibração, o teste de Hosmer-Lemeshow foi adequado para o modelo valvar ($P=0,52$) e inadequado para o modelo valvar + CRM ($P=0,04$).

3.4 Recalibração e remodelamento dos escores de risco

Antunes e col. [71] fizeram, em 2007, a análise de três escores: EuroSCORE logístico, o modelo inicial do Parsonnet e *Ontario Province Risk (OPR)* comparando-o ao modelo local (*Coimbra Score*). Um total de 4.567 pacientes operados de cirurgia de coronária isolada no Hospital Universitário de Coimbra foram analisados retrospectivamente. As áreas abaixo da curva ROC foram de 0,754 para o EuroSCORE logístico, 0,664 para o modelo inicial do Parsonnet e 0,683 para o *OPR*. O modelo local (idade, doença vascular periférica, FE<40%, cirurgia não eletiva, e reoperação) apresentou uma área abaixo da curva ROC de 0,752.

Outra análise que comparou o EuroSCORE com o modelo inicial do Parsonnet foi a publicada em 2007 por Au e col. [72] em Hong Kong. Um total de 1.247 pacientes operados de CRM e 1.406 pacientes operados de cirurgia valvar foram analisados. No grupo da CRM, a área abaixo da curva ROC foi de 0,76 para o EuroSCORE e 0,74 para o modelo inicial do Parsonnet. No grupo da cirurgia valvar, a área abaixo da curva ROC foi de 0,77 para o EuroSCORE e 0,79 para o modelo inicial do Parsonnet.

Em 2007, Gomes e col. [73] publicaram, no Brasil, um modelo composto de 46 variáveis: pré-operatórias, intra-operatórias e do primeiro dia do pós-operatório. Este modelo foi derivado da análise de 1.458 pacientes operados consecutivamente de cirurgia cardíaca no período de 2000 a 2003. Não houve outra amostra para realizar a validação do modelo, esta foi realizada no mesmo grupo de desenvolvimento do escore. Desta forma, a área abaixo da curva ROC foi de 0,84.

Em 2008, Campagnucci e col. [74] publicaram, no Brasil, uma análise retrospectiva com o EuroSCORE aditivo em 100 pacientes consecutivos operados de CRM. O tamanho da amostra limitou a análise estatística, prejudicando uma conclusão adequada do estudo.

Uma análise feita em 2008 por D'Errigo e col. [20] comparou o modelo derivado da *Italian CABG Outcome Project* e o EuroSCORE (aditivo e logístico), com o objetivo de prever a mortalidade hospitalar e de poder elaborar o ranking dos hospitais. Na calibração, o EuroSCORE logístico mostrou-se insatisfatório ($P < 0,001$) e superestimou a mortalidade em todos os grupos de risco. Mas, uma vez feita a recalibração das variáveis, o modelo alcançou níveis aceitáveis. O EuroSCORE aditivo ($P = 0,228$) e o modelo italiano ($P = 0,170$) mostraram calibração adequada. Ambos os modelos superestimaram discretamente a mortalidade nos pacientes de baixo e alto risco e o modelo italiano calibrou melhor os pacientes de médio risco. A área abaixo da curva ROC foi de 0,796 para o modelo italiano, 0,777 para o EuroSCORE logístico e 0,773 para o EuroSCORE aditivo.

Neste mesmo ano, Lafuente e col. [75] publicaram, na Espanha, uma análise retrospectiva para validação do EuroSCORE (aditivo + logístico) em 498 pacientes submetidos a CRM no Hospital Clínico de Barcelona entre 2004 e 2006. O desfecho foi mortalidade intra-hospitalar. A área abaixo da curva ROC foi de 0,830 para o EuroSCORE logístico e 0,839 para o EuroSCORE aditivo. Para o EuroSCORE logístico o teste de Hosmer-Lemeshow teve um $P = 0,32$, mostrando boa calibração.

Em 2009, Zheng e col. [21] publicaram, na China, a análise do EuroSCORE logístico no Registro Chinês de Cirurgia de Revascularização Miocárdica. Foram analisados 9.248 pacientes de 35 centros na China submetidos a CRM. Na calibração

que avaliou a mortalidade intra-hospitalar observada em comparação com a esperada, o EuroSCORE logístico superestimou a mortalidade observada. A área abaixo da curva ROC foi de 0,72. A análise revelou diferenças significativas na prevalência dos fatores de risco entre ambas as populações; bem como dificuldades na definição das variáveis e no registro obrigatório dos pacientes.

Em 2009, Ranucci e col. [14] publicaram, na Itália, uma análise em 11.150 pacientes submetidos à cirurgia cardíaca, mostrando que a limitação do número de variáveis usadas pelo EuroSCORE diminuiria o risco de sobreajuste, multicolinearidade e o erro humano, quando o modelo fosse aplicado a um número limitado de pacientes. A melhor acurácia foi obtida com cinco variáveis (idade, fração de ejeção, creatinina, cirurgia de emergência e cirurgia de revascularização miocárdica associada), com uma área abaixo da curva ROC de 0,76 em comparação a 0,75 do EuroSCORE logístico sem diferença significativa. Embora este novo modelo tenha resultado numa calibração melhor, na comparação com o EuroSCORE logístico ($P=0,678$ versus $P=0,185$).

Após o EuroSCORE não ter validado na população da Austrália, Reid e col. [76] publicaram, em 2009, um modelo pré-operatório aditivo de predição de mortalidade em 30 dias (AusSCORE), derivado da análise de 7.709 pacientes submetidos a CRM isolada. Após uma análise estatística, o modelo escolhido (idade, classe da NYHA, fração de ejeção, urgência da cirurgia, reoperação, hipercolesterolemia, doença vascular periférica e choque cardiogênico) teve uma calibração de 0,242 e uma área abaixo da curva ROC de 0,834. Quando aplicado aos pacientes do banco de dados que foram operados de cirurgia cardíaca, apresentou uma calibração de $P=0,591$ e uma área abaixo da curva ROC de 0,818. Uma

adequada sistematização no registro e análise dos dados colaborou para uma excelente aplicação das técnicas de validação.

A capacidade de discriminação e precisão do EuroSCORE na predição em cirurgia valvar tem sido ampliada recentemente, por dois motivos. O primeiro é que na maioria dos centros os procedimentos valvares (isolada ou combinada) atualmente representam mais de 50% do total de cirurgias cardíacas; por tanto, a estimativa de risco nestes pacientes tem adquirido importância. A segunda razão é estritamente relacionada à recente evolução nas opções de tratamento sobre a valva aórtica, que tem influenciado na adoção de procedimentos via transcateter nos pacientes de altíssimo risco ou muito idosos. Portanto, a adequação dos modelos de risco é essencial para selecionar os pacientes que se beneficiariam desta nova técnica, antes que decisões erradas sejam tomadas.

Sem dúvida, a maior força estatística para conhecer o desempenho dos modelos de risco são as metanálises. Havia certa polêmica sobre a aplicabilidade do EuroSCORE nos pacientes valvares, considerando que < 30% da população originária do EuroSCORE foram cirurgias valvares. Em 2010, Parolari e col. [77] publicaram, na Itália, uma metanálise para confirmar a validação do EuroSCORE (aditivo + logístico) neste tipo de pacientes. Foram analisados 12 estudos com 26.621 pacientes e 1.250 óbitos. A área abaixo da curva ROC foi de 0,730 (IC 95%, 0,717-0,743). Foi documentada uma constante tendência a superestimar a mortalidade para ambos os modelos do EuroSCORE, especialmente para o logístico. De fato, foi observado um caso paradoxal, no qual o mais antigo e menos sofisticado modelo (EuroSCORE aditivo) tem melhor desempenho que o modelo mais recente (EuroSCORE logístico) na predição de risco nos casos complexos.

Em 2010, o EuroSCORE logístico foi aplicado por Carvalho e col. [78] retrospectivamente em uma amostra de 546 pacientes operados de CRM em quatro hospitais públicos do Rio de Janeiro. A área abaixo da curva ROC foi de 0,62 (IC95%, 0,56-0,65). O EuroSCORE logístico subestimou notavelmente a mortalidade observada.

Na América Latina, com exceção do Brasil, existe um estudo na Colômbia, publicado em 2010, por Contreras e col. [79], avaliando a mortalidade intra-hospitalar com o EuroSCORE (aditivo + logístico). Um total de 153 pacientes consecutivos foram operados de cirurgia cardíaca. A área abaixo da curva ROC foi de 0,87 para o EuroSCORE logístico e de 0,86 para o EuroSCORE aditivo. Na calibração o teste de Hosmer- Lemeshow foi de 0,84 para o EuroSCORE logístico e de 0,89 para o EuroSCORE *aditivo*. O modelo bem desenhado é limitado devido ao tamanho da amostra e, conseqüentemente, pelo número de óbitos.

Em 2010, Billah e cols. [80] publicaram, na Austrália, um modelo pré-operatório preditor de mortalidade em 30 dias nos pacientes operados de cirurgia cardíaca. O modelo foi desenvolvido em 13.810 pacientes (60%) dentre os pacientes registrados no banco de dados da Sociedade Australiana de Cirurgiões torácicos e cardíacos. A fase de validação, escolhida randomicamente, foi realizada em 9.206 pacientes (40%). O método *bootstrap* na seleção de variáveis corrobora a força deste modelo em comparação com outros modelos da literatura. Na análise multivariável, doze variáveis foram escolhidas (idade, sexo, urgência do procedimento, tipo de procedimento, reoperação, NYHA classe funcional, medicação inotrópica, fração de ejeção, diálise pré-operatória, hipercolesterolemia, doença vascular prévia e índice de massa corpórea $>25 \text{ kg/m}^2$). A área abaixo da curva ROC do modelo foi de 0,813. Na

calibração, o teste de Hosmer-Lemeshow apresentou um $P=0,81$. Neste estudo foi iniciada a utilização do método de *bootstrap* como um método de maior evidência e estabilidade na escolha das variáveis preditoras.

Em 2010, Malik e col. [81] publicaram, na Índia, a validação do EuroSCORE aditivo em 1.000 pacientes consecutivos operados de cirurgia cardíaca (cirurgia valvar e/ou CRM). Todos os pacientes foram computados e as definições estritamente respeitadas. A área abaixo da curva ROC foi de 0,827. Na calibração, o teste de Hosmer-Lemeshow apresentou $P=0,73$. A diferença do perfil clínico dos pacientes entre ambas as populações foi marcada por uma alta prevalência de variáveis relacionadas com a apresentação tardia da doença na Índia. Do mesmo modo, uma inversão dos procedimentos realizados (30% pacientes valvares na população do EuroSCORE e 60% na população da Índia) poderia explicar a subestimação da mortalidade nos pacientes de alto risco.

Em março de 2011, Shih e col. [82] publicaram, no Taiwan, o desempenho do EuroSCORE (aditivo + logístico) em 1.240 pacientes submetidos à cirurgia cardíaca. Durante a validação, foram respeitadas todas as fases de uma adequada análise estatística. No entanto, ao igual que nos outros estudos, houve diferença estatística na prevalência dos fatores de risco entre o grupo de estudo e a população do EuroSCORE. As áreas abaixo da curva ROC foram de 0,839 para o EuroSCORE aditivo e de 0,845 para o EuroSCORE logístico. No subgrupo da CRM isolada a área abaixo da curva ROC foi de 0,873. Na calibração, todos os subgrupos, exceto para CRM, demonstraram uma boa aplicação do modelo. A análise multivariável revelou sete fatores de risco preditores de mortalidade intra-hospitalar (idade, condição pré-

operatória crítica, cirurgia de emergência, cirurgia na aorta torácica, CRM associada, fração de ejeção < 30 e creatinina sérica $> 2,26\text{MG/dl}$).

No mesmo mês, no InCor-HCFMUSP [83] publicamos uma análise prospectiva utilizando o modelo do 2000 Bernstein Parsonnet e o EuroSCORE aditivo. Foram analisados 774 pacientes consecutivos submetidos à cirurgia de coronária, valva e associada na modalidade eletiva ou urgência/emergência. A área abaixo da curva ROC foi de 0,835 para o 2000BP e de 0,807 para o EuroSCORE aditivo ($P < 0,001$). Na calibração, o 2000BP ($P=0,70$) e o EuroSCORE ($P=0,39$) foram adequados. As definições das variáveis foram idênticas à dos modelos utilizados, mas o tamanho da amostra foi a maior limitação do estudo.

Em março de 2011, Berg e col. [84] publicaram, na Noruega, uma análise utilizando o EuroSCORE (aditivo + logístico) em 5.029 pacientes operados de cirurgia cardíaca. Foram realizadas análises de regressão logística utilizando métodos de *bootstrap* para mortalidade intra-hospitalar. Um modelo pré-operatório com oito variáveis (idade, sexo, grau de urgência, tipo de cirurgia, cirurgia cardíaca prévia, disfunção renal, cardíaca e pulmonar) foi elaborado. A área abaixo da curva ROC foi de 0,857 (0,823-0,891) para o novo modelo, e de 0,821 (0,785-0,857) para o EuroSCORE, ($P=0,02$) Na calibração, o novo modelo superestimou a mortalidade nos pacientes de altíssimo risco.

Um trabalho publicado no Paquistão, em abril de 2011, por Qadir e col. [85] analisou retrospectivamente o EuroSCORE (aditivo + logístico) em 2.004 pacientes operados de CRM isolada. A área abaixo da curva ROC foi de 0,859 para o EuroSCORE logístico e de 0,866 para o EuroSCORE aditivo. A calibração foi boa para o EuroSCORE aditivo ($P=0,424$) e inadequada para o EuroSCORE logístico

($P < 0,001$). Ambos os escores superestimam discretamente a mortalidade no grupo de baixo e médio risco, sendo o EuroSCORE logístico mais acurado. A análise multivariável selecionou cinco variáveis como preditoras de mortalidade na população estudada (idade, sexo, falha renal, angina instável e infarto de miocárdio). O EuroSCORE aditivo teve melhor desempenho que o EuroSCORE logístico.

Por conseguinte, fica demonstrada a necessidade de encontrar um instrumento simples, abrangente e ao mesmo tempo eficiente para auxílio na tomada de decisões, principalmente, na aplicação de novas tecnologias.

Em outubro de 2011, Nashef e col. [86] apresentaram em Lisboa, no *25th European Association for Cardio-thoracic Surgery Annual Meeting*, o EuroSCORE remodelado que passou a ser denominado de EUROSCORE II. No estudo, 23.000 pacientes foram operados de cirurgia cardíaca em mais de 150 hospitais de 43 países entre maio e julho de 2010. Tratava-se de um sistema de registro voluntário e multinacional, embora somente logístico. No grupo de desenvolvimento do modelo foi utilizada a rede neural para a escolha das variáveis. Passaram a ser requeridas novas variáveis, como: cálculo da depuração de creatinina, diabete insulino dependente, classe da NYHA e classe 4 da CCS. Do mesmo modo, foram reclassificadas: fração de ejeção, hipertensão pulmonar, urgência do procedimento e tipo de procedimento realizado. 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 inicial. Os autores referam que a discriminação do novo modelo foi muito boa, embora a modelagem não tenha sido relatada.

É importante deduzir da extensa publicação de modelos preditivos de risco e suas diversas aplicações geográficas e por grupos de risco que, embora os escores

sejam ferramentas estatisticamente bastante sofisticadas, eles ainda não atingiram um grau de infalibilidade. Estes escores avaliam somente uma parte ínfima das múltiplas variáveis conhecidas e desconhecidas do paciente e da estrutura assistencial, que influenciam diretamente no resultado final do processo. Por esta razão, as conclusões derivadas da sua aplicação devem ser cuidadosamente analisadas.

4 Métodos

4.1 Amostra

Este estudo prospectivo e observacional foi conduzido na Divisão de Cirurgia Cardiovascular do Departamento de Cardiopneumologia do Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (InCor-HCFMUSP).

O tamanho da amostra para fazer a validação preditiva dos escores provenientes de populações diferentes deverá observar como princípio geral pelo menos 100 óbitos [17]. Para isto, a publicação de Lisboa e col. [87], sobre os resultados do InCor-HCFMUSP, nos últimos 23 anos, relata uma mortalidade global de 6,9%, sendo 4,8% para cirurgia coronária eletiva e 8,4% para cirurgia valvar eletiva. Como no nosso estudo incluímos pacientes operados de coronária e/ou valva, consideramos razoável utilizar a menor mortalidade como parâmetro para o cálculo da amostra, neste caso a cirurgia coronária eletiva, resultando em 2.084 pacientes.

4.2 Critérios de inclusão e exclusão

Critérios de inclusão: Todos os pacientes operados consecutivamente entre maio de 2007 e julho de 2009, na modalidade eletiva, urgência ou emergência de:

- Cirurgia valvar (troca ou plastia);
- Cirurgia de coronárias (com ou sem a utilização de circulação extracorpórea);

- Cirurgia associada (cirurgia de revascularização miocárdica e cirurgia valvar).

Critérios de exclusão: Pacientes submetidos à cirurgia de cardiopatia congênita, transplante cardíaco, afecções da aorta, tumores cardíacos, doenças do endocárdio ou do pericárdio.

4.3 Coleta, definição e organização dos dados

Os dados foram coletados pré-operatoriamente à avaliação clínica e do sistema eletrônico de prontuários do InCor (SI3) e armazenados numa planilha única. Esta planilha (Anexo I) foi adaptada, de forma a contemplar todas as variáveis descritas pelo modelo do 2000 Bernstein Parsonnet [23] e do EuroSCORE [11]. Um total de 60 variáveis pré-operatórias (demográficas, clínicas e laboratoriais) por paciente foram coletadas. Todas as definições atribuídas às variáveis por ambos os escores foram respeitadas junto a seus respectivos valores segundo sua relevância com o evento morte (Anexo II). Assim, após o cálculo do valor do 2000BP e do ES para cada paciente, eles foram ordenados segundo os grupos de risco estabelecidos pelos escores e colocados no banco de dados confeccionado no programa Excel com esta finalidade. Todos os pacientes foram acompanhados até a alta hospitalar. Nenhum paciente foi excluído da análise por falta de dados. O desfecho clínico foi a mortalidade intrahospitalar, definida como a morte ocorrida no intervalo de tempo entre a cirurgia e a alta hospitalar.

4.4 Validação do 2000 Bernstein Parsonnet e do EuroSCORE

Para avaliar o desempenho do 2000BP e do ES na predição de mortalidade, foi realizada a validação preditiva dos modelos em 3.000 pacientes. A análise foi feita mediante teste de calibração e discriminação.

Calibração. - A calibração avalia a acurácia do modelo para predizer risco em um grupo de pacientes. Em outras palavras, se o modelo propõe que a mortalidade em 1.000 pacientes seria 5% e a mortalidade observada é de 5% ou perto de isso, diremos que o modelo está bem calibrado. A força da calibração foi avaliada testando a qualidade do ajuste mediante o teste de Hosmer-Lemeshow [88, 89]. O valor de $P > 0,05$ indica que o modelo se ajusta aos dados e prediz a mortalidade adequadamente.

Discriminação. - A discriminação mede a habilidade do modelo para distinguir entre pacientes de baixo e alto risco. Em outras palavras, se a maioria dos óbitos ocorre em pacientes que o modelo identifica como de alto risco, diremos que o modelo tem uma boa discriminação. Ao contrário, se a maioria dos óbitos ocorre em pacientes que o modelo identifica como de baixo risco, diremos que o modelo tem pobre discriminação. A discriminação é medida por meio da utilização da técnica estatística chamada área abaixo da curva ROC (algumas vezes chamada *c-statistic* ou *c-index*). Assim, a excelente discriminação refere-se a valores acima de 0,97; a muito boa discriminação está no intervalo entre 0,93 e 0,96; boa discriminação entre 0,75 e 0,92; abaixo de 0,75 corresponde a modelos deficientes na capacidade de discriminação [90]. Na prática, os modelos raramente excedem 0,85.

Análise estatística. - A análise estatística foi realizada através do *software* SPSS versão 16.0 para Windows (SPSS Inc, Chicago, Illinois). 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 intrahospitalar foi realizada por meio da utilização do valor dado para cada paciente pelos escores 2000BP e ES. 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.

4.5 Formulação do modelo local (InsCor)

A técnica de *bootstrap* junto a procedimentos de seleção automatizada de variáveis “stepwise” foi utilizada para desenvolver um modelo parcimonioso através da regressão logística múltipla [91]. Esta técnica foi introduzida por Efron [92] e tenta o que seria desejável realizar na prática: “repetir a experiência”. As observações são escolhidas de forma aleatória e as estimativas, recalculadas. Trata a amostra observada como se esta representasse toda a população e, desta forma, retira amostras de mesmo tamanho com repetição, gerando assim um grande número de amostras. Aplica-se a técnica de análise estatística em cada amostra e obtêm-se as estimativas desejadas. Espera-se que estas estimativas converjam para uma estimativa

única. São fortes as evidências que indicam que os modelos preditivos produzidos com técnicas de *bootstrap* são os mais estáveis, têm melhor acurácia e são constituídos por variáveis que são verdadeiros preditores reproduzíveis e independentes [93]. Para esta análise o banco de dados constituído por 3.000 pacientes foi dividido em dois grupos: o grupo de desenvolvimento e o grupo de validação do modelo.

Grupo de desenvolvimento do modelo (InsCor). – Dos primeiros 2.000 pacientes, foram selecionadas 1.000 amostras *bootstrap* (cada uma contendo 2.000 pacientes) para o desenvolvimento do modelo. Por tanto, foram construídas 1.000 amostras com reposição contendo o mesmo número de casos de óbito e de não óbito da amostra original. Para cada amostra, foi rodado o modelo de regressão logística multivariado com processo de seleção “*stepwise*” e anotadas as variáveis que foram selecionadas nos 1.000 modelos obtidos (ranking das variáveis). As variáveis utilizadas nestes modelos foram as que apresentaram $p < 0,10$ na análise univariada. As variáveis contínuas foram analisadas com a partição já escolhida durante a análise univariada.

Grupo da validação do modelo (InsCor). – Como todo modelo de predição, o InsCor precisa ser validado. A validação “otimista” denominada “*goodness of fit*” é a forma mais fraca de validação (dados não apresentados). A avaliação do desempenho do modelo em dados não pertencentes ao grupo de desenvolvimento do modelo é mais importante. Isso pode ser alcançado utilizando a validação interna. Para a **validação interna**, uma abordagem comum é dividir a população total de pacientes em duas partes (frequente 2:1); desenvolver o modelo utilizando o primeiro grupo e avaliar seu desempenho preditivo no segundo grupo.

Para isso, o InsCor teria que passar pelos testes de calibração e discriminação como mencionados anteriormente. Para melhor discernimento do estudo, decidimos também aplicar o 2000BP e o ES no mesmo grupo de validação para poder compará-los ao InsCor. No entanto, a validação interna não dá capacidade para o modelo ser utilizado fora do local onde foi desenvolvido [13, 16]

Análise estatística. – No grupo de desenvolvimento do modelo foi utilizada a técnica de *bootstrap* com a construção de 1.000 amostras de reposição do conjunto de 2.000 pacientes através do *software SAS System version 9.2 for Microsoft Windows* [94]. O *SAS System* foi escolhido para realizar a técnica de *bootstrap* por existir uma maior afinidade com ele, neste tipo de análise. As variáveis com $P < 0,10$ foram identificadas como plausíveis preditores de mortalidade após cirurgia cardíaca. O risco aditivo foi obtido a partir dos coeficientes beta estimados do modelo de predição de risco proposto. As categorias de risco foram definidas de tal forma que fossem de tamanho semelhante [11]. No grupo de validação (1.000 pacientes), o desempenho do InsCor foi comparado ao do 2000BP e ES por meio da utilização do *software SPSS versão 16.0 para Windows* (SPSS Inc, Chicago, Illinois). A calibração e a discriminação foram medidas para cada valor do escore na população de pacientes. O valor $P < 0,05$ foi considerado significativo.

4.6 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 (Anexo III), sendo isentada a necessidade do termo de consentimento livre e esclarecido pelo tipo de desenho aplicado.

5 Resultados

5.1 Casuística

Todos os pacientes operados de coronária e/ou valva entre maio de 2007 e julho de 2009 no InCor-HCFMUSP foram incluídos no estudo. Dos 3.000 pacientes operados, 268 (8,9%) morreram. Do total de procedimentos, 57,7% (1.731) foram operados de coronária; 36,8% (1104) de valva e 5,5% (165) de coronária e valva. A mortalidade por tipo de procedimento cirúrgico é apresentada em subgrupos para uma análise na **tabela 1**.

Tabela 1. Mortalidade por tipo de procedimento cirúrgico.

Procedimento	Nº Pacientes	Nº Óbitos	%
RM			
CRM S/CEC (Isolada, primária)	429	14	3,26%
CRM C/CEC (Isolada, primária)	1122	64	5,70%
CRM (Reoperação)	90	13	14,44%
CRM + Aneurismectomia	81	7	8,64%
CIV(PósIAM)	11	7	63,63%
Valvar			
VAo (Isolada, primária)	288	28	9,72%
VAo (Reoperação)	106	16	15,09%
VMi (Isolada, primária)	258	15	5,81%
VMi (Reoperação)	212	29	13,68%
VTri	5	2 (Reop.)	40%
Dupla valva	218	34	15,60%
Triple valva	17	3	17,65%
Associada			
CRM + VAo	108	21	19,44%
CRM + VMi	80	14	17,50%
*Outras	18	5	27,78%

CRM= cirurgia de revascularização miocárdica; Reop. = Reoperação; S/CEC= Sem circulação extracorpórea; C/CEC = com circulação extracorpórea; VAo= valva aórtica; VMi= valva mitral; CIV(PósIAM)= Comunicação interventricular pós-infarto agudo do miocárdio. *Outras: cirurgias associadas (exceção de CRM + cirurgia valvar).

Para efeito descritivo, mostramos na **tabela 2A** a prevalência dos fatores de risco na população de estudo e na população do ES. Do mesmo modo, na **tabela 2B**, é apresentada a prevalência dos fatores de risco na população de estudo e na população do 2000BP. Pelo fato destas populações serem potencialmente comparáveis, foi analisada a diferença estatística das prevalências dos fatores de risco do ES e do 2000BP, em relação à população de estudo.

Tabela 2A. Prevalência dos fatores de risco no Grupo Estudo, comparado com os fatores de risco da população do EuroSCORE.

VARIÁVEIS	ESTUDO	EUROSCORE	P
	(N=3000)	(N=19030)	
Idade			
<60	44,27%	33,20%	< 0,001
60-64	15,80%	17,80%	0,007
65-69	13,87%	20,70%	< 0,001
70-74	12,20%	17,90%	< 0,001
>75	11,50%	9,60%	0,001
Feminino	35,90%	27,80%	< 0,001
Doença pulmonar crônica	2,60%	3,90%	< 0,001
Arteriopatia extracardíaca	4,80%	11,30%	< 0,001
Disfunção neurológica	6,90%	1,40%	< 0,001
Cirurgia cardíaca previa	17,80%	7,30%	< 0,001
Creatinina > 2,3mg/dl	4,40%	1,80%	< 0,001
Endocardite ativa	4,10%	1,00%	< 0,001
Estado pré-operatório crítico	10,30%	4,10%	< 0,001
Angina instável	7,00%	8,00%	0,059
FE 30 – 50	26,10%	25,60%	0,569
FE <30	5,80%	5,80%	0,998
IAM recente	16,80%	9,70%	< 0,001
Hipertensão pulmonar	8,10%	2,00%	< 0,001
Emergência	3,10%	4,90%	< 0,001
Cirurgia associada	6,90%	36,40%	< 0,001
Cirurgia na aorta torácica	0,70%	2,40%	< 0,001
CIV pós-infarto	0,50%	0,20%	0,002

FE: Fração de ejeção; IAM: Infarto agudo do miocárdio; CIV: Comunicação interventricular.

Tabela 2B. Prevalência dos fatores de risco no Grupo Estudo, comparado com os fatores de risco da população do 2000 Bernstein Parsonnet.

VARIÁVEIS	ESTUDO	2000BP	P
<i>Idade</i>	(N=3000)	(N=10703)	
70-74	12,20%	18,50%	< 0,001
>75	11,50%	13,70%	0,002
Feminino	35,90%	31,30%	< 0,001
Doença pulmonar crônica	2,60%	10,80%	< 0,001
Arteriopatia extracardíaca	4,80%	9,10%	< 0,001
Disfunção neurológica	6,90%	8,40%	0,008
Cirurgia cardíaca previa	17,80%	7,60%	< 0,001
Creatinina > 2,3mg/dl	4,40%	4,50%	0,809
FE 30 – 50	26,10%	38,60%	< 0,001
FE <30	5,80%	8,40%	< 0,001
Hipertensão pulmonar	8,10%	10,70%	< 0,001

FE: Fração de ejeção.

Alem do que foi apresentado na **tabela 2A e 2B**, nossos dados mostram que dos pacientes que foram operados: 11,8% tiveram lesão de tronco da coronária esquerda > 50%; 9,1% foram considerados obesos mórbidos; 2,7% apresentaram balão intra-aórtico; 21,9% tiveram indicação de tratamento cirúrgico da valva aórtica; 26,2% tiveram indicação de tratamento cirúrgico da valva mitral; 1,2% tinham critérios de choque cardiogênico; 5,6% apresentaram insuficiência tricúspide importante; 2,2% já eram dependentes de marca-passo; 2,1% foram operados durante as primeiras 48 horas do infarto agudo do miocárdio; 1,7% foram entubados antes da cirurgia; 0,4% tinham diagnóstico de cirrose hepática; 6% tinham creatinina > 2mg/dl; 1,5% apresentaram ao menos uma carótida com lesão > 70%; 6,9% tiveram

diagnóstico prévio de acidente cerebrovascular ou de limitação motora importante; 11,2% tiveram diagnóstico de fibrilação atrial; 18,1% tiveram átrio esquerdo > 50 mm e 37,3% apresentaram diagnóstico de dislipidemia.

5.2 Resultados da Validação do 2000 Bernstein Parsonnet e EuroSCORE *(3.000 pacientes)*

5.2.1 Resultados da Calibração

-2000 Bernstein Parsonnet

Foi encontrada associação entre o modelo 2000BP e óbito com OR: 1,079 ($P < 0,001$). O teste de Hosmer-Lemeshow mostrou uma *goodness-of-fit statistic* = 15,678 com 8 graus de liberdade, $P = 0,0472$ (**tabela 3A**). Para uma melhor adequação da análise, o 2000BP foi dividido em cinco categorias, **tabela 3B**. O 2000BP apresenta um ajuste inadequado nos subgrupos estabelecidos.

Tabela 3A. Mortalidade observada e esperada mediante uso do 2000BP como variável preditora nos grupos definidos pelo teste Hosmer – Lemeshow.

Grupo	Total	ÓBITO = 1		ÓBITO = 0	
		Observada	Esperada	Observada	Esperada
1	293	6	5,42	287	287,58
2	301	4	7,43	297	293,57
3	300	3	9,17	297	290,83
4	292	12	10,82	280	281,18
5	272	6	12,15	266	259,85
6	286	12	15,52	274	270,48
7	299	25	20,97	274	278,03
8	296	33	27,90	263	268,10
9	299	55	43,60	244	255,40
10	362	113	116,03	249	245,97

* Goodness of fit statistic = 15.678 com 8 DF (P=0.0472).

Tabela 3B. Percentagens da mortalidade observada e estimada por grupos de risco do 2000BP.

Risco	MO/ME	ME	MO	N	%
< 9	0,77	2,19	1,68	594	19,8
9 -14	0,75	3,38	2,53	592	19,7
14,1 – 19,9	0,67	4,84	3,23	558	18,6
20 – 28,9	1,18	8,24	9,75	595	19,8
≥ 29	1,05	24,21	25,42	661	22

MO/ME= Mortalidade observada/Mortalidade esperada; ME= Mortalidade esperada; MO= Mortalidade observada; N= Número de pacientes; %= Porcentagem.

-EuroSCORE

Foi encontrada a associação entre o modelo ES e óbito com OR: 1,337 ($P < 0,001$). O teste de Hosmer-Lemeshow mostrou uma *goodness of fit statistic* = 5.5301 com 7 graus de liberdade, $P = 0,5956$ (**tabela 4A**). Para uma melhor adequação da análise, o ES foi dividido em três categorias, **tabela 4B**. O ES apresenta um ajuste adequado nos subgrupos estabelecidos.

Tabela 4A. Mortalidade observada e esperada mediante uso do ES como variável preditora nos grupos definidos pelo teste *Hosmer – Lemeshow*.

Grupo	Total	ÓBITO = 1		ÓBITO = 0	
		Observada	Esperada	Observada	Esperada
1	296	5	4,83	291	291,17
2	378	10	8,20	368	369,80
3	287	6	8,27	281	278,73
4	397	12	15,15	385	381,85
5	358	19	18,04	339	339,96
6	269	14	17,83	255	251,17
7	264	19	22,89	245	241,11
8	345	53	43,97	292	301,03
9	406	131	129,83	275	276,17

* *Goodness of fit statistic* = 5.5301 com 7 DF ($P = 0.5956$)

Tabela 4B. Porcentagens da mortalidade observada e estimada por grupos de risco do ES.

Risco	MO/ME	ME	MO	N	%
0 – 2	1	2,19	2,19	961	32,03
3 – 5	0,88	4,98	4,39	1024	34,13
≥ 6	1,03	19,41	20,00	1015	33,83

MO/ME= Mortalidade observada/Mortalidade esperada; ME= Mortalidade esperada; MO= Mortalidade observada; N= Número de pacientes; %= Porcentagem.

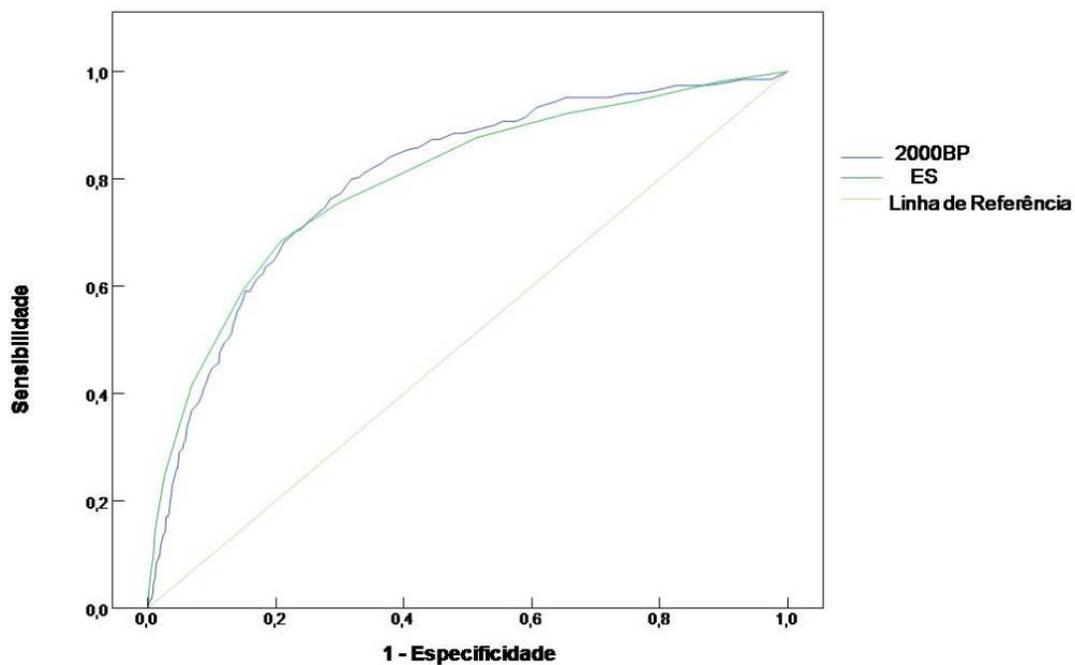
5.2.2 Resultados da Discriminação (Curvas ROC, figura 1)

-2000 Bernstein Parsonnet

Avaliando o poder de discriminação do 2000BP, observamos que a área abaixo da curva ROC foi de 0,800 (IC95%, 0,772 – 0,827, P=0,014) (**tabela 5**).

-EuroSCORE

Avaliando o poder de discriminação do ES, observamos que a área abaixo da curva ROC foi de 0,796 (IC95%, 0,766 – 0,826, P=0,015) (**tabela 5**).



*2000BP ROC= Curva ROC apresentada pelo escore 2000 Bernstein Parsonnet; ES ROC= Curva ROC resultante do EuroSCORE.

Figura 1. Curva ROC para o 2000BP e o ES na avaliação do poder de discriminação realizada em 3.000 pacientes.

Tabela 5. Área abaixo da curva ROC para o 2000 BP e o ES da análise realizada em 3.000 pacientes.

	Área	IC 95%	ep	P
2000BP	0,800	0,772 – 0,827	0,014	< 0,001
ES	0,796	0,766 – 0,826	0,015	< 0,001

5.3. O modelo desenvolvido: InsCor

Após a obtenção dos modelos fizemos uma ordenação das variáveis de acordo com o número de vezes em que elas foram selecionadas, e escolhemos as 10 primeiras, por acreditarmos que com um número reduzido de variáveis poderíamos obter um bom score. Assim, foram selecionadas para o modelo final as variáveis que foram selecionadas em mais de 40% das 1.000 amostras *bootstraps*. Os resultados são resumidos na **tabela 6**. No ranking, sexo feminino e cirurgia de revascularização miocárdica (CRM) + cirurgia valvar foram selecionados como preditores independentes em pelo menos 52% das amostras. Foi identificado o infarto do miocárdio recente (<90dias) em pelo menos 50% das amostras *bootstrap*. A reoperação foi identificada como preditora em pelo menos 47% das amostras *bootstrap*. Idade > 70% e tratamento cirúrgico da valva aórtica foram selecionadas em pelo menos 45% das amostras. Tratamento cirúrgico da valva tricúspide e creatinina sérica > 2mg/dl foram selecionados como preditores independentes em pelo menos 44% das amostras *bootstrap*. Fração de ejeção < 30% foi selecionada em pelo menos 43% das amostras. Evento, variável composta, presente quando uma das seguintes situações está incluída: balão intra-aórtico pré-operatório, choque cardiogênico, taquicardia ventricular, intubação oro-traqueal pré-operatória, insuficiência renal aguda, uso de inotrópicos no preoperatório e/ou massagem cardíaca prévia. Esta variável foi selecionada em pelo menos 42% das amostras *bootstrap*.

Tabela 6. Análise do *bootstrap*: ranking das variáveis para desenvolvimento do InsCor.

Variável	Número de vezes que foram escolhidas
Sexo feminino	52
CRM + cirurgia valvar	52
Infarto recente <90dias	50
Reoperação	47
Idade > 70 anos	45
Tratamento da valva aórtica	45
Tratamento da valva tricúspide	44
Creatinina > 2mg	44
Fração de ejeção < 30%	43
Evento	42
Distúrbios neurológicos	40
Fibrilação atrial	40
Pressão pulmonar sistólica >60mmHg	39
Endocardite aguda	38
Angina instável	38
Cirurgia de emergência	38
Insuficiência congestiva	36
Tratamento da valva mitral	32
Dependente de marcapasso	29
CIV pós-IAM	29
IAM 48h	28
Cirrose hepática	28

CRM: Cirurgia de revascularização miocárdica. CIV pós-IAM: Comunicação interventricular pós-infarto agudo do miocárdio; IAM 48h: Infarto agudo do miocárdio até 48 horas do diagnóstico.

O InsCor foi o melhor dentre os três modelos escolhidos do ranking das variáveis (dados não mostrados), quanto à simplicidade, objetividade e força estatística para a predição de mortalidade. As variáveis do modelo selecionado foram: sexo feminino, CRM + cirurgia valvar, infarto recente, reoperação, idade > 70anos, tratamento cirúrgico da valva aórtica, tratamento cirúrgico da valva tricúspide, creatinina > 2mg/dl, fração de ejeção < 30% e “eventos”. Não houve nenhuma evidência de interação de primeira ordem e multicolinearidade entre as variáveis do modelo final (**tabela 7**).

Tabela 7. O InsCor

Variáveis	“Odds ratio”	IC a 95%	
		LI	LS
Idade ≥ 70 anos	3,14	2,26	4,37
Sexo feminino	1,62	1,17	2,25
CRM + cirurgia valvar	2,28	1,39	3,75
Infarto recente < 90 dias	1,40	0,94	2,09
Reoperação	2,87	2,03	4,05
Tratamento da valva aórtica	2,29	1,62	3,23
Tratamento da valva tricúspide	2,58	1,55	4,30
Creatinina > 2mg/dl	4,70	3,01	7,33
Fração de ejeção < 30 %	3,24	2,02	5,20
*Eventos	5,48	3,84	7,82

*Eventos: Inclui pelo menos uma das seguintes situações prévias à cirurgia: balão intra-aórtico, choque cardiogênico, taquicardia ou fibrilação ventricular, intubação orotraqueal, insuficiência renal aguda, uso de drogas inotrópicas e massagem cardíaca.

Assim, a partir dos “*Odds ratios*” das variáveis selecionadas foi criado o modelo aditivo (**figura 2**):

$$\begin{aligned} \text{InsCor:} & (3 * \text{idade} \geq 70) + (2 * \text{sexofeminino}) + (2 * \text{CRM} + \text{valvar}) + (2 * \text{Infartorece} \\ & \text{nte}) + (3 * \text{Reoperação}) + (2 * \text{TTOVAo}) + (3 * \text{TTOVTric.}) + (5 * \text{Creatinina} > 2) \\ & + (3 * \text{FE} < 30) + (5 * \text{Eventos}). \end{aligned}$$

Figura 2. Equação do modelo aditivo do InsCor.

O modelo aditivo do InsCor pode variar de 0 a 30. A partir, dos valores do InsCor e fazendo uso da regressão logística pode-se estimar a probabilidade de óbito dos pacientes. Na **tabela 8**, apresentamos a probabilidade para alguns valores do InsCor e na **figura3** as probabilidades para todos os valores. O modelo foi dividido em três categorias: baixo risco (0-3); médio risco (4-7) e alto risco (≥ 8).

Tabela 8. Valores de probabilidade estimada de óbito para o InsCor.

Escore	Probabilidade estimada de óbito
0	0,02
5	0,07
10	0,18
15	0,40
20	0,67
25	0,86
30	0,95

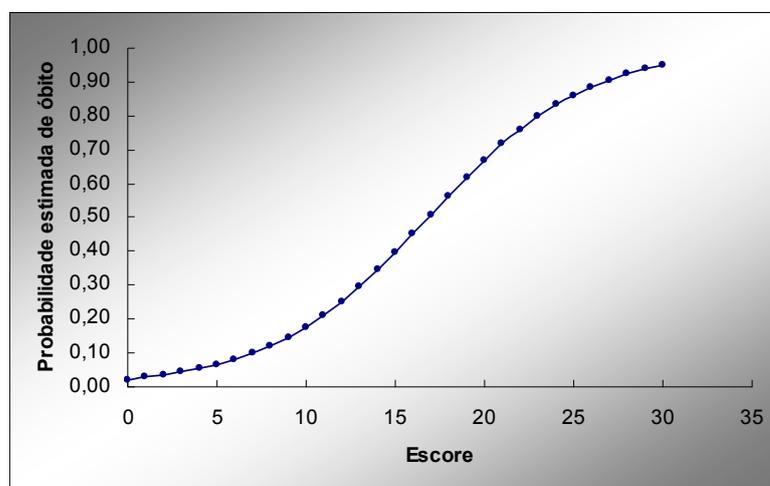


Figura 3. Curva da probabilidade estimada de óbito segundo o InsCor.

Para praticidade e melhor adoção do modelo, foi criada uma regra mnemotécnica com a palavra FICARTE, na qual às 3 primeiras letras (FIC x 2) contêm 2 variáveis cada uma e as outras (ARTE x 1) uma variável cada uma (**figura 4**).

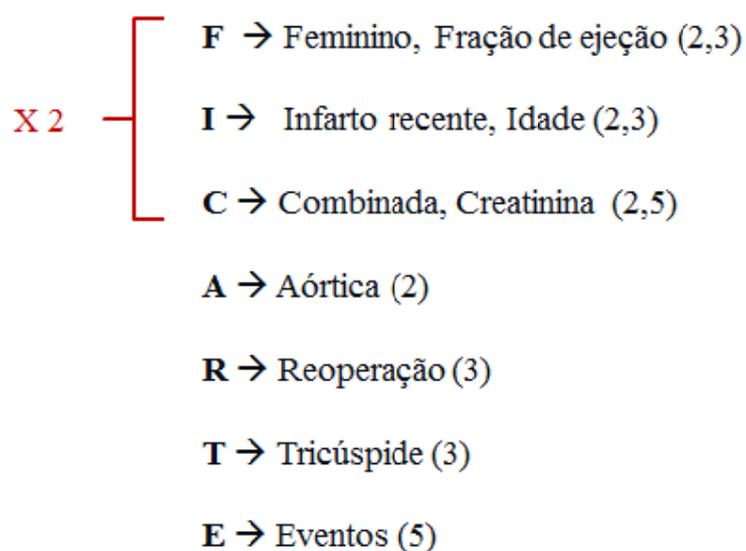


Figura 4. Regra mnemotécnica para cálculo do InsCor.

Construído o InsCor, este passou pelos processos de calibração e discriminação.

Validação do InsCor e comparação com o EuroSCORE e o 2000BP (1000 pacientes).

Respeitando os conceitos estatísticos, foi realizada a validação interna do InsCor no grupo da validação (1.000 pacientes). No mesmo grupo, foi realizada também a validação do EuroSCORE e do 2000 Bernstein Parsonnet para compará-los ao desempenho do InsCor.

5.4 Calibração do InsCor

A Calibração do InsCor foi adequada, com $P=0,184$ no teste de Hosmer-Lemeshow (**tabela 9**). 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 10** é apresentada a calibração do InsCor por grupos de risco.

Tabela 9. Mortalidade observada e esperada mediante o uso do InsCor como variável preditora nos grupos definidos pelo teste Hosmer – Lemeshow.

Grupo	Total	ÓBITO = 1		ÓBITO = 0	
		Observada	Esperada	Observada	Esperada
1	224	3	7,73	221	216,27
2	213	10	11,37	203	201,63
3	99	5	6,55	94	92,45
4	69	5	5,64	64	63,36
5	149	22	14,98	127	134,02
6	117	20	18,16	97	98,84
7	129	46	46,58	83	82,42

* *Goodness of fit statistic* = 7,5339 com 5 DF ($P=0,1839$)

Tabela 10. Calibração do InsCor - Análise por grupos de risco

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

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

5.5 Calibração do EuroSCORE

Tabela 11. Calibração do EuroSCORE - Análise por grupos de risco

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

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

5.6 Calibração do 2000BP

Tabela 12. Calibração do 2000BP - Análise por grupos de risco

Risco	Número de casos	Mortalidade	
		% observada (IC a 95%)	% predita (IC a 95%)
0 – 8,5	210	0,95 (-0,36; 2,27)	2,38 (0,32; 4,44)
9 - 14	198	3,03 (0,64; 5,42)	4,55 (1,64; 7,45)
14,5 - 20	217	5,53 (2,49; 8,57)	6,45 (3,18; 9,72)
20,5 – 30,5	197	14,21 (9,34; 19,09)	11,17 (6,77; 15,57)
>= 31	178	35,39 (28,37; 42,42)	33,71 (26,76; 40,65)

Teste de Hosmer-Lemeshow (P=0,157)

Na discriminação, a área abaixo da curva ROC foi 0,79 para o InsCor; 0,81 para o ES e 0,82 para o 2000BP, **figura 5**. Na **tabela 13**, observamos que não houve diferença significativa na discriminação com os modelos InsCor, ES e 2000BP.

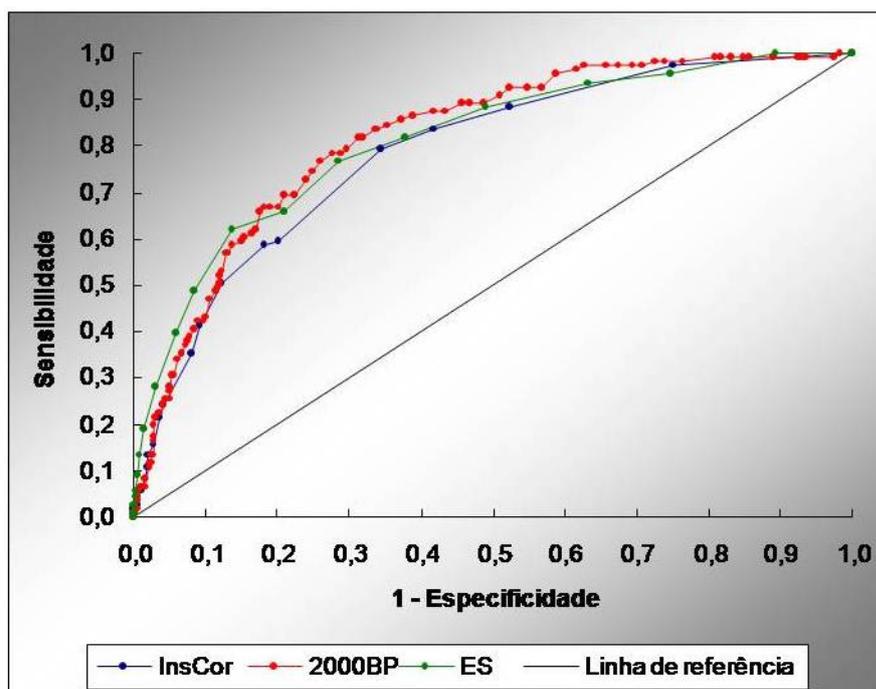


Figura 5. Curva ROC para o InsCor, o 2000BP e o ES na avaliação do poder de discriminação realizada em 1.000 pacientes.

Tabela 13. Área abaixo da curva ROC para o InsCor, o 2000 BP e o ES realizada em 1000 pacientes.

	Área	IC 95%	ep	P
InsCor	0,79	0,74 – 0,83	0,02	< 0,001
2000BP	0,82	0,78 – 0,86	0,02	< 0,001
ES	0,81	0,77 – 0,86	0,02	< 0,001

Desta forma, o valor dado aos fatores de risco pré-operatórios selecionados pelo modelo do InsCor pode ser utilizado para predizer o risco de mortalidade intrahospitalar nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP.

6.1 Caracterização da casuística utilizada

A comparação de modelos de predição de risco revela mais similaridades que diferenças entre os mesmos. A avaliação das variáveis relacionadas ao paciente mostra uma considerável sobreposição nas características que foram significativas e utilizadas por cada programa para a predição da mortalidade. Como pode ser notado, cada um dos sistemas avaliados utiliza outras variáveis além das referentes ao paciente. Variáveis como prioridade da cirurgia, bem como a informação de procedimentos médicos e de intervenções incorporam características tanto do paciente como do processo de assistência. Decisões das equipes médicas e/ou cirúrgicas antes do procedimento cirúrgico têm importante influência nos resultados.

Portanto, a inclusão destas variáveis nos modelos de risco torna estes modelos mais difíceis de serem interpretados, uma vez que eles misturam elementos relativos à assistência com aqueles intrínsecos à condição clínica pré-operatória do paciente.

Vários fatores de risco reconhecidamente relacionados com complicações e mortalidade, embora sejam valiosos na predição de risco, apresentam dificuldades práticas de avaliação ou não disponibilidade para todos os pacientes, tornando sua inclusão em modelos de predição bastante difícil [95]. Há também considerável variação entre os estudos quanto aos fatores que se mostraram estatisticamente associados ao risco aumentado de complicações ou mortalidade, refletindo diferenças nas populações avaliadas [30]. As complicações e a mortalidade são influenciadas pela técnica cirúrgica e anestésica, e particularmente por fatores como: tempo de permanência em circulação extracorpórea, adequação da proteção miocárdica e

controle hemodinâmico [15,96]. Eventos peri-operatórios podem ampliar ou neutralizar o risco estabelecido com base nos fatores pré-operatórios. Tendo em vista estes aspectos, alguns autores realizaram a avaliação do paciente no momento da admissão em unidade de terapia intensiva na tentativa de captar a influência nos resultados de fatores diretamente ligados ao procedimento cirúrgico [37, 43, 97, 98].

Com exceção de Turner e col. [43] os achados dos estudos avaliados confirmam a impressão de que a condição clínica pré-operatória do paciente é o principal determinante dos resultados cirúrgicos e que estes resultados podem ser estimados de forma adequada mesmo sem considerar outros fatores, como a ocorrência de erros técnicos ou de avaliação intra-operatória ou preservação inadequada do miocárdio no decorrer da circulação extracorpórea. A estabilidade clínica do paciente imediatamente antes da cirurgia é o componente de risco mais importante para os resultados em curto prazo, enquanto as disfunções crônicas têm maior impacto na evolução em longo prazo, envolvendo qualidade de vida e estado funcional pós-cirurgia [6].

A melhor demonstração é o trabalho de Martinez-Alario e col. [48] quando compararam o Parsonnet inicial com os sistemas gerais de avaliação de gravidade na predição de risco de mortalidade em cirurgia cardíaca (APACHE I, II, III; SAPS I, II; MPM I e II). O modelo de Parsonnet, que inclui apenas variáveis pré-operatórias, foi mais adequado para predição de mortalidade que os sistemas gerais, que consideram variáveis pós-operatórias.

Outro aspecto a ser considerado é que, restringindo-se as variáveis preditoras àquelas relacionadas às condições pré-operatórias, possibilita-se a utilização dos modelos de predição de risco como instrumento de tomada de decisão

previamente ao procedimento, o que se torna impossível com a inclusão de variáveis per e pós-operatórias.

Sugere-se que os modelos de predição de risco não contenham variáveis instituições específicas [99] e que sejam validados em populações diferentes da que originou o modelo [100]. Em geral, os modelos de predição de risco têm um melhor desempenho na população que os originou do que quando aplicado em populações diferentes. Charlson e col. [101] sugerem que parte desta variabilidade possa ser devida a diferenças na definição da variável resposta ou na coleta de dados. Assim por exemplo, uma explicação da melhora substancial dos resultados em Nova York, após implantar o registro obrigatório, foi a padronização das variáveis do banco de dados [67]. A grande vantagem dos modelos de predição de risco em cirurgia cardíaca baseados em dados pré-operatórios é permitir o ajuste dos indicadores institucionais pela gravidade dos pacientes, tornando comparáveis os resultados em diferentes momentos e entre instituições e profissionais distintos, e possibilitando o seu uso para o planejamento do atendimento prestado [102, 103, 104].

Sabemos que não é possível comparar as instituições e profissionais de forma justa sem antes considerar as variáveis preditoras a serem incluídas nos modelos de predição de risco. Os estudos analisados podem ser considerados reflexos deste posicionamento. Jones e col. [3] analisando especificamente a CRM, sugeriram que a maior parte das informações relacionadas ao prognóstico está contida em relativamente poucas variáveis clínicas. Este grupo de trabalho, denominado PANEL, identificou e propôs sete variáveis essenciais (idade, sexo, prioridade da cirurgia, função do ventrículo esquerdo, cirurgia cardíaca prévia, lesão de tronco esquerdo e número de coronárias acometidas) que, segundo eles, deveriam

estar presentes em qualquer banco de dados relacionados à CRM, uma vez que se relacionam fortemente com mortalidade operatória. O grupo PANEL também identificou 13 variáveis, categorizadas como nível 1, e sugerem sua inclusão nos bancos de dados de CRM, pela alta probabilidade de estarem relacionadas com a mortalidade precoce, e mais 24 variáveis opcionais, ou nível 2, cujo poder preditor não foi claramente demonstrado, mas que podem ser interessantes para pesquisa ou decisões administrativas. Tu e col. [46] testaram a proposta do grupo PANEL e concluíram que modelos simples contendo apenas as variáveis essenciais são tão efetivos quanto modelos complexos para a comparação de resultados entre as instituições “pouca variável quanto possíveis”. Para uma variável ter impacto significativo no ajuste de indicadores de complicações ou mortalidade, deve possuir pelo menos quatro características: deve ser um preditor da variável resposta estatisticamente significativo, a distribuição da variável deve ser significativamente diferente entre os grupos a serem comparados, deve ser altamente frequente e não possuir correlação com outros fatores no banco de dados incluídos no modelo de predição de risco.

Na avaliação dos trabalhos de validação, um dos melhores trabalhos feitos com o Parsonnet foi realizado por Berman e col. [63] em Israel, em 2006. A última atualização do Parsonnet, o 2000 Bernstein Parsonnet foi comparada, ao EuroSCORE aditivo. Foram analisados 1.639 pacientes operados consecutivamente de coronária, valva e associada. A área abaixo da curva ROC foi 0,83 para o 2000 Bernstein Parsonnet e 0,73 para o EuroSCORE aditivo. Este resultado foi similar à nossa análise inicial realizada no InCor-HCFMUSP [83] em 744 pacientes mediante utilização dos mesmos escores de risco e dos mesmos procedimentos. Nesta análise,

o teste de Hosmer-Lemeshow para o 2000BP ($P = 0,70$) e para o ES ($P = 0,39$) indicaram uma boa calibração. Assim mesmo, a curva ROC para o 2000BP = 0,84 e para o ES = 0,81 foi adequada para prever mortalidade e com melhor desempenho do que os resultados do estudo realizado em Israel. No entanto, a interpretação da validação dos modelos fica limitada pelo número de óbitos. Na análise final com 3.000 pacientes podemos observar que para este tamanho de amostra a calibração se inverte; o 2000 Bernstein Parsonnet já não calibra ($P=0,047$) e o ES aumenta sua calibração ($P=0,597$). Contudo, a boa discriminação persiste com uma área abaixo da curva ROC de 0,80 para o ES e de 0,81 para o 2000BP. Neste contexto, a literatura atesta que a origem, natureza e evolução destes modelos são favoráveis a um melhor desempenho do EuroSCORE em populações maiores, enquanto o 2000BP o faz com grupos menores. Uma explicação para isto é conhecida na estatística como sobreajuste ou “*overfitting*” dos modelos quando apresentam muitas variáveis [16, 30, 105, 106]. Para confirmar estes fatos, em uma análise feita em 1,000 pacientes da mesma amostra, o 2000BP apresentou uma calibração com $P = 0,157$ e o ES com $P=0,593$. Por tanto, na medida em que a amostra aumenta de tamanho a calibração do EuroSCORE melhora, e a do 2000BP piora. Mesmo assim, a persistência de uma adequada discriminação para ambos os modelos indica que as variáveis contidas nos modelos são verdadeiras preditoras de mortalidade [13, 107].

Em 2007, Antunes e col. [71], além de fazer a validação de modelos externos a partir de 4567 pacientes operados de CRM isolada formularam um modelo com as seguintes variáveis: idade, doença vascular periférica, fração de ejeção $< 40\%$, cirurgia não eletiva, e reoperação. O modelo local, chamado de *Coimbra score*, apresentou uma área abaixo da curva ROC de 0,752. O InsCor, um

modelo com 10 variáveis que serviria não somente para pacientes operados de CRM mas também de cirurgia valvar e associada apresentou as seguintes variáveis compartilhadas com o *Coimbra score*: idade ≥ 70 a, reoperação e fração de ejeção $< 30\%$. Mesmo que o InsCor tenha apresentado três variáveis específicas para doença valvar, este apresentou uma área abaixo da curva ROC de 0,79, comparada a 0,75 do *Coimbra score*.

Em 2008, D'Errigo e col. [20], após fazer a validação do *EuroSCORE*, elaboraram um modelo derivado da *Italian CABG Outcome Project*. Este modelo italiano era constituído de 14 variáveis (idade, sexo feminino, choque cardiogênico, diabetes, diálise, hipertensão pulmonar, arritmia ventricular maligna, doença pulmonar obstrutiva crônica, creatinina > 2 mg/dl, arteriopatía extracardiaca, angina instável, CRM prévia, emergência e fração de ejeção). Cinco variáveis são compartilhadas com o InsCor (idade, sexo feminino, creatinina > 2 mg/dl, reoperação e fração de ejeção). Além disso, choque cardiogênico, arritmia ventricular maligna e angina instável são variáveis que encaixam na variável “eventos” do InsCor. Nesta análise, o *EuroSCORE* aditivo ($P=0,228$) e o modelo italiano ($P=0,170$) mostraram adequada calibração. No entanto, na nossa análise, a calibração foi melhor com $P=0,595$ para o *EuroSCORE* e $P=0,184$ para o InsCor. Na discriminação, a área abaixo da curva ROC foi de 0,796 para o modelo italiano e 0,773 para o *EuroSCORE* aditivo, semelhante ao encontrado no nosso estudo com uma área abaixo da curva ROC de 0,79 para o InsCor e 0,81 para o *EuroSCORE* aditivo.

Em 2009, Ranucci e col. [14] demonstraram, na Itália, que a limitação do número de variáveis usadas pelo *EuroSCORE* diminuiria o risco de sobreajuste, multicolinearidade e erro humano. Assim, o melhor modelo foi obtido com cinco

variáveis (idade, fração de ejeção, creatinina, cirurgia de emergência e cirurgia de revascularização miocárdica associada). Destas, apenas cirurgia de emergência não está incluída no InsCor. A área abaixo da curva ROC para este modelo de cinco variáveis foi de 0,76 em comparação a 0,75 do EuroSCORE sem diferença significativa. Assim, a obtenção de um novo modelo com os próprios dados, recalibrando-lo periodicamente, pode proporcionar a mesma acurácia, melhor calibração e alto valor de informação, especialmente quando estes modelos são aplicados a séries institucionais de 1.000 a 1.500 pacientes anuais [18].

Considerando que o EuroSCORE e mesmo o modelo recalibrado não conseguiram ter validade na população da Austrália [65], Reid e col. [76] publicaram em 2009 um modelo aditivo de predição de mortalidade, chamado AusSCORE, dirigido a pacientes operados de CRM isolada. O modelo escolhido de oito variáveis (idade, classe da NYHA, fração de ejeção, urgência da cirurgia, reoperação, hipercolesterolemia, doença vascular periférica e choque cardiogênico) teve uma calibração de 0,242 e uma área abaixo da curva ROC de 0,834. As mesmas variáveis do InsCor compartilhadas com o Coimbra score (idade, fração de ejeção e reoperação) foram compartilhadas com o AusSCORE. Além disso, a variável choque cardiogênico se encaixa na variável “eventos” do InsCor. Uma adequada sistematização no registro e análise dos dados, além do tamanho da amostra, contribuiu para uma excelente aplicação das técnicas de validação.

Em 2010, Billah e cols. [80] publicaram, na Austrália, um modelo pré-operatório preditor de mortalidade, desta vez para pacientes operados de cirurgia cardíaca. O método *bootstrap* para seleção de variáveis foi novidade na elaboração de modelos preditivos em cirurgia cardíaca. Doze variáveis foram escolhidas (idade,

sexo, urgência do procedimento, tipo de procedimento, reoperação, NYHA classe funcional, medicação inotrópica, fração de ejeção, diálise pré-operatória, hipercolesterolemia, doença vascular prévia e índice de massa corpórea $>25 \text{ kg/m}^2$). Este modelo foi elaborado com a mesma metodologia que o InsCor, porém acabou ficando maior e menos objetivo, compartilhando cinco variáveis (idade, sexo, reoperação, fração de ejeção e tipo de procedimento), além da variável medicação inotrópica encaixada na variável “eventos” do InsCor. Na calibração do modelo australiano, o teste de Hosmer-Lemeshow apresentou um $P=0,81$, comparado a $0,18$ do InsCor. A área abaixo da curva ROC do modelo australiano foi de $0,813$ contra $0,79$ do InsCor. O tamanho da população para a validação do modelo australiano foi 10 vezes maior que o grupo de validação do InsCor podendo ter influenciado nos testes de validação.

Em 2010 Malik e col. [81] publicaram, na Índia, a validação do EuroSCORE aditivo em mais de 1.000 pacientes consecutivos operados de cirurgia cardíaca. Na calibração, o teste de Hosmer-Lemeshow apresentou $P=0,73$. A área abaixo da curva ROC foi de $0,827$. A similaridade com a validação do Euroscore aditivo na nossa população poderia estar correlacionada ao perfil clínico dos pacientes caracterizados por uma alta prevalência de doença reumática.

Em março de 2011, Shih e col. [82] publicaram, no Taiwan, o desempenho do EuroSCORE (aditivo + logístico) em 1.240 pacientes submetidos à cirurgia cardíaca.

Na calibração, todos os subgrupos exceto para CRM demonstraram uma boa aplicação do modelo. As áreas abaixo da curva ROC foram de $0,839$ para o EuroSCORE aditivo e de $0,845$ para o EuroSCORE logístico. A análise multivariável

revelou sete fatores de risco preditores de mortalidade intra-hospitalar (idade, condição pré-operatória crítica, cirurgia de emergência, cirurgia na aorta torácica, cirurgia associada, fração de ejeção < 30 e creatinina sérica $> 2,26\text{mg/dl}$). Quatro variáveis são compartilhadas com as do InsCor (idade, condição pré-operatória crítica, cirurgia associada, fração de ejeção < 30). A variável creatinina sérica $> 2,26\text{mg/dl}$ do modelo foi semelhante à do InsCor, no qual a creatinina foi $> 2\text{mg/dl}$.

Em março de 2011, Berg e col. [84] publicaram, na Noruega, uma análise utilizando o EuroSCORE (aditivo + logístico) em 5.029 pacientes operados de cirurgia cardíaca. Foram realizadas análises de regressão logística utilizando métodos de *bootstrap* para mortalidade. Um modelo pré-operatório com oito variáveis (idade, sexo, grau de urgência, tipo de cirurgia, cirurgia cardíaca prévia, disfunção renal, cardíaca e pulmonar) foi construído. As variáveis compartilhadas com as do modelo do InsCor foram três: idade, sexo feminino e cirurgia cardíaca prévia. A variável disfunção renal do modelo não foi detalhada para que ela pudesse ser comparada à do InsCor. A área abaixo da curva ROC foi de 0,857 para o novo modelo e de 0,821 para o EuroSCORE, ($P=0,02$). Na calibração o novo modelo superestimou mortalidade nos pacientes de altíssimo risco.

Em abril de 2011, um trabalho publicado no Paquistão por Qadir e col. [85] analisou retrospectivamente o EuroSCORE (aditivo + logístico) em 2.004 pacientes operados de CRM isolada. A calibração foi boa para o EuroSCORE aditivo ($P=0,424$) e inadequada para o EuroSCORE logístico ($P<0,001$). A área abaixo da curva ROC foi de 0,859 para o EuroSCORE logístico e de 0,866 para o EuroSCORE aditivo. A análise multivariável selecionou cinco variáveis como preditoras de mortalidade na população estudada (idade, sexo, falha renal, angina instável e infarto

de miocárdio recente). Das cinco variáveis, três foram compartilhadas com as do InsCor (idade, sexo e infarto de miocárdio recente). A variável falha renal não foi detalhada para que pudesse ser comparada com a variável creatinina $> 2\text{mg/dl}$ do InsCor. No entanto, como demonstrado em várias análises o EuroSCORE aditivo teve melhor desempenho que o EuroSCORE logístico.

Assim, podemos ver que na maioria dos modelos, uma vez determinadas as variáveis independentes e dependentes, o processo de desenvolvimento dos modelos segue, em geral, a sequência apresentada por Clark [108]: análise univariada, análise multivariada e avaliação do grau de adequação do modelo.

A regressão logística múltipla foi a análise estatística mais utilizada para determinar a associação independente entre cada fator de risco e o evento desfecho. Poucos estudos utilizaram o modelo Bayesiano [35, 109, 110], a justificativa apresentada foi a flexibilidade deste modelo para acomodar mudanças temporais na população de pacientes, além de lidar melhor com problemas associados a dados incompletos ou definições inconsistentes. Modelos derivados de redes neurais surgiram posteriormente, porém em decorrência da complexidade na aplicabilidade não se mostraram práticos [111-115]. Com o tempo, o EUROSCORE II [86] trouxe novamente a rede neural para o remodelamento do modelo mais utilizado no mundo. No entanto, a técnica de *bootstrap* junto com procedimentos de seleção automatizada de variáveis “*stepwise*” é a mais nova inovação em cirurgia cardíaca. As evidências descrevem que estes modelos são mais estáveis, têm melhor acurácia e incluem variáveis que são verdadeiros preditores reproduzíveis e independentes [91, 92].

Escores podem ser desenvolvidos com base na opinião de especialistas ou levar em consideração a magnitude do “*Odds ratios*” obtida pela regressão logística. A vantagem do sistema de escore é sua praticidade, pois exprime o risco global do paciente representado pela soma dos valores atribuídos a cada uma das variáveis independentes. A desvantagem é a menor precisão atribuída aos sistemas de escore quando comparados aos modelos de regressão, uma vez que são feitos arredondamentos e agrupamento de variáveis contínuas [2]. Porém, a diferença de “*performance*” parece não ser clinicamente importante [37].

A idade é um determinante crítico de complicações e mortalidade pós-cirurgia cardíaca na maioria dos estudos, embora a importância atribuída à idade varie amplamente [7, 11, 22, 23, 25, 26, 32, 36, 37, 38, 40, 45, 46]. A média das idades na população submetida à cirurgia cardíaca vem aumentando. Edwards e col. [27] em estudo com duração de 10 anos mostraram que a idade média passou de 58,5 anos, na década de 1980, para 64,1 anos na década de 1990 ($p < 0.0001$). Quando o EuroSCORE considera risco acima de 60 anos, o modelo inicial do Parsonnet [22], assim como o Inscor considera aumento de risco apenas na faixa etária ≥ 70 . Hannan e col. [26] verificaram aumento de risco na faixa etária ≥ 75 anos. A maioria dos estudos inclui o sexo feminino como fator de risco para mortalidade em cirurgia cardíaca [7, 11, 22, 23, 25, 36, 38, 45]. O risco aumentado entre as mulheres é atribuído a vários fatores. Parsonnet e col. [22] e Higgins e col. [37] examinaram as variáveis sexo e superfície corporal na mesma população e mostraram que, sem o ajuste para superfície corporal, o sexo feminino apresenta um maior risco, uma vez que a população feminina caracteriza-se por ter uma superfície corporal reduzida. Neste contexto, o menor diâmetro das artérias no sexo feminino também interfere

nos resultados [116]. Atualmente, o diagnóstico e o tratamento das doenças cardíacas isquêmicas estão em estágio mais avançado na população feminina em comparação com a masculina, refletindo uma abordagem inadequada ou inapropriada da dor torácica na mulher e introduzindo um viés de seleção nas avaliações [117, 118]. Este aspecto, além das diferenças das características clínicas e angiográficas entre os sexos, também é destacado por Tuman e col. [40] uma vez que o sexo feminino permaneceu sendo fator de risco significativo nos modelos logísticos, apesar da consideração da superfície corporal. Outros fatores de risco para doença isquêmica do coração própria ao sexo feminino, como a redução de estrógeno com a idade, possivelmente interferem nas diferenças de risco encontradas entre os sexos. Nos estudos avaliados, a reoperação foi particularmente investigada, encontrando-se relação direta com o aumento de óbito [11, 20, 22, 23, 25, 26, 32, 38, 45]. No estudo de Tu e col. [25] a reoperação mostrou-se associada de forma significativa com mortalidade (OR= 3,22; $p < 0,001$) sendo compatível com a hipótese de que a reoperação aumenta o risco de óbito per e pós-operatório precoce; mas, os pacientes sobreviventes apresentam um tempo de permanência apenas um pouco mais prolongado que aqueles submetidos ao primeiro procedimento.

Marcadores de alteração de reserva fisiológica pré-operatória incluem o grau de disfunção do ventrículo esquerdo (presença de insuficiência cardíaca congestiva ou fração de ejeção), função renal (creatinina sérica elevada, dependência de diálise), condições pulmonares (doença pulmonar obstrutiva crônica) [119]. O valor da fração de ejeção pode variar, dependendo do tipo de procedimento em que é mensurado (cateterismo ou ecocardiograma) e das condições presentes no momento da medida (arritmia cardíaca, doença valvar etc), mas é uma forma objetiva de

avaliação da função do ventrículo esquerdo. A definição de fração de ejeção reduzida varia entre os estudos, tendo sido encontrados os valores de $< 30\%$, $< 35\%$ e $< 40\%$. Parsonnet e col. [22] verificaram um aumento de 4% na taxa de mortalidade quando a FE foi $< 30\%$. O InsCor considera FE $<30\%$ como preditor de mortalidade, da mesma forma que o EuroSCORE e o 2000BP. Hannan e col. [119] encontraram risco para mortalidade 4,06 vezes maior para os indivíduos avaliados com FE $< 20\%$ e de 1,63 vezes para os indivíduos com FE entre 30 e 39%, quando comparados àqueles com FE $\geq 40\%$. O'Connor e col. [38] encontraram um risco de óbito 2 vezes maior entre os indivíduos com FE $< 40\%$ quando comparados àqueles cuja FE foi $\geq 60\%$. Alguns estudos definem disfunção renal como creatinina sérica elevada; outros, como dependência de diálise. Higgins e col. [37] verificaram risco elevado de mortalidade em pacientes com creatinina sérica $\geq 1,9$ mg/dl (OR= 3,69; $p < 0.0001$) e moderado no grupo com creatinina sérica entre 1,6 e 1,8 mg/dl. O modelo do InsCor, assim como o modelo italiano [20], considera creatinina > 2 mg/dl como preditor de mortalidade. Este valor de creatinina está relacionado com a taxa de filtração glomerular < 60 ml/min, o qual indica o limite da função renal normal [120]

O InsCor considera a cirurgia de valva tricúspide como preditor de mortalidade, assim como os modelos Pons[46] e 2000BP[23]. No entanto em vários modelos [11, 20, 22, 23] a hipertensão pulmonar é considerada. Poderia tratar-se de variáveis colineares, no entanto o 2000BP considera as duas variáveis no modelo, e na análise de *bootstrap* (ranking das variáveis) para a formulação do InsCor a hipertensão pulmonar ocupou a 13ª posição. No estudo, em geral a maioria dos casos de hipertensão pulmonar é secundária a doença na valva mitral, pode-se deduzir que em geral quando há indicação de tratamento cirúrgico da valva tricúspide existem

níveis mais avançados de hipertensão pulmonar. Uma alta prevalência de hipertensão pulmonar na população do estudo tem criado uma rotina no controle da hipertensão pulmonar no pós-operatório imediato, fazendo com que somente os estágios mais avançados sejam considerados de risco.

Desde a década de 1970, procedimentos em pacientes com Infarto agudo de miocárdio recente aumentaram os índices de mortalidade, levando a retrazar o procedimento. Atualmente, quando indicado a cirurgia deve-se tentar estabilizar o paciente para diminuir os riscos. Esta variável preditora incluída pelo InsCor é considerada importante por outros modelos (11, 26, 32, 36, 46).

Na verdade os únicos modelos de risco que consideram a cirurgia da valva aórtica como uma variável preditora de mortalidade são o modelo inicial do Parsonnet [22] e o InsCor. À exceção do InsCor, nenhum modelo atual considera a cirurgia valvar aórtica como preditor de mortalidade. Variáveis ocultas, não esclarecidas no estudo e relacionadas à indicação mais tardia da doença podem ter influenciado nos resultados.

A cirurgia associada (CRM + cirurgia valvar) é um fator de risco importante de mortalidade em cirurgia cardíaca. O aumento no tempo de CEC, riscos na proteção miocárdica e o sangramento no posoperatorio são variáveis relacionadas com estes procedimentos. Esta variável é considerada preditora no InsCor e nos modelos Parsonnet inicial [22], Pons [46], 2000BP [23] e EuroSCORE [11].

As evidências confirmam que a condição clínica pré-operatória do paciente é o principal determinante dos resultados cirúrgicos. Variáveis isoladas prévias à cirurgia, tais como: balão intra-aórtico, choque cardiogênico, taquicardia ou fibrilação ventricular, intubação orotraqueal, insuficiência renal aguda, uso de drogas

inotrópicas e massagem cardíaca são importantes, mas perdem força na análise por serem infreqüentes. Portanto, a variável combinada estado pré-operatório crítico se faz necessária, sendo variável preditora no InsCor e em outros modelos de risco [11, 22, 26, 32].

Um aspecto a ser considerado é a possibilidade de aplicação dos modelos de predição de risco a nível individual. Estes modelos são obtidos a partir de grandes populações e devem ser tratados como tal. Os sistemas de predição são válidos quando aplicados a grandes grupos de pacientes, como na comparação entre instituições, mas apresentam limitações significativas quando utilizados na predição de risco individual e precisam ser utilizados com cautela [121, 122].

A predição de risco em cirurgia cardíaca permanece como uma questão complexa. O desenvolvimento de indicadores de desempenho clínico/cirúrgico e a expansão da base de conhecimentos médicos que permita a identificação de métodos mais efetivos de assistência médica são necessários, mas não suficientes para a melhoria da assistência.

Os modelos desenvolvidos, particularmente aqueles baseados em variáveis préoperatórias, representam um grande avanço neste sentido, mas precisam ser mais bem testados em outras populações e países para que se possa determinar a sua validade externa.

6.1.1 Limitações do estudo

As limitações do estudo foram: **Primeiro**, é um estudo coletado de um único centro, onde a limitação mais importante é a generalização dos resultados. Em

segundo lugar, em virtude de sua natureza, o EuroSCORE aditivo tende a subestimar o risco nos pacientes de alto risco; embora isso não tenha sido demonstrado em trabalhos multicêntricos [123], concordamos que a escolha do EuroSCORE logístico seria a mais adequada se comparássemos unicamente pacientes de alto risco, o que não foi o objetivo do trabalho. Por último, embora a mortalidade hospitalar (até 30 dias após cirurgia) pareça ser mais completa que a mortalidade intra-hospitalar (até a alta hospitalar), as definições atuais sugerem que ambas têm acurácia equivalente sendo a mortalidade intra-hospitalar mais prática e fácil de usar [124].

6.2 Considerações finais

O surgimento dos escores de risco em cirurgia cardíaca ocorreu como resposta à necessidade de ajustar diversos resultados a uma padronização estabelecida, baseando-se unicamente no perfil de risco do paciente. Trata-se de uma exigência da sociedade e das instituições frente a um número grande de procedimentos cardiovasculares considerados de alta complexidade. Posteriormente, seu uso foi disseminado na carreira por mostrar melhores resultados e aumentar a demanda nos serviços de atendimento. Talvez a utilização atual mais importante seja no pré-operatório para auxílio na tomada de decisões (novas tecnologias) e no pós-operatório na prevenção dos efeitos adversos e controle dos custos principalmente na unidade de cuidados intensivos. Não devemos esquecer que, no tempo e espaço, variações nos sistemas de prevenção, diagnóstico e tratamento dos fatores de risco e

das doenças cardiovasculares podem alterar a acurácia dos modelos. Por outro lado, são claramente definidos os fatores de risco relacionados com a morbimortalidade nos pacientes operados por doenças cardiovasculares. Assim, para poder utilizar estes modelos matemáticos em diferentes tempos e espaços, devemos primeiro validá-los respeitando os princípios estatísticos: amostra, coleta de dados, calibração e discriminação. Lamentavelmente, em quase todas as validações externas realizadas com os modelos mais populares, ocorrem erros na coleta dos dados, e principalmente na definição das variáveis. Confirmou-se que, quando existiu uma adequada aplicação dos escores, sobretudo em registros obrigatórios (todos os pacientes), houve um bom desempenho dos modelos [53, 62, 63, 82, 83, 84]

Assim, na procura por um modelo ideal ficou claro que a preferência foi dada a um modelo parcimonioso, abrangente e ao mesmo tempo simples, objetivo e eficaz. Sem dúvidas o EuroSCORE seria o escolhido.

Muito bem citado o efeito Hawthorne trazido à luz pelo Professor Nashef [125], para explicar que nada foi mais benéfico para os resultados em cirurgia cardíaca no início do século do que a monitorização dos resultados e a incorporação do EuroSCORE na maioria das unidades de cirurgia cardiovascular no mundo. Com o tempo, a recalibração das variáveis e, certamente o remodelamento do EuroSCORE se faria necessário.

Em nossa realidade, a falta de uma correta validação de modelos externos, obrigatória em uma população em desenvolvimento e com alta prevalência de doença reumática e de Chagas, dificultaram sua aplicabilidade. Neste trabalho, podemos ver que uma adequada aplicação do EuroSCORE e do 2000 Bernstein Parsonnet nos

permitiu conhecer o desempenho dos modelos e dos resultados cirúrgicos no InCor-HCFMUSP, citados anteriormente.

São questionáveis as opiniões de especialistas na tentativa de ajustar o EuroSCORE, como por exemplo, na indicação do implante transapical de prótese aórtica, multiplicando o EuroSCORE do paciente pela razão de mortalidade do hospital (mortalidade observada/mortalidade esperada).

Adaptações das diretrizes internacionais em função do momento da indicação cirúrgica variam em relação à experiência e aos resultados do hospital; assim, por exemplo, na indicação de tratamento cirúrgico do aneurisma da aorta ascendente pelo diâmetro e da insuficiência mitral importante assintomática.

Por outro lado, formulamos o conceito da conversão de risco, similar à conversão de moeda, na qual um coeficiente, variável segundo os resultados de cada país, é capaz de comparar resultados internacionalmente, como é a proposta do EUROSCORE II.

A pergunta que se poderia fazer é: por que ajustar um modelo, e não criar um modelo? O que é melhor?

Sem dúvida criar um modelo é mais complexo, mas é mais fidedigno para avaliações de risco feitas no local onde ele foi formulado (diferenças na prevalência dos fatores de risco).

Crescer acima de teus resultados ou crescer acima de outros resultados, o que é mais fácil para melhorar o desempenho, sabendo da existência de muitas variáveis ocultas?

Crescer acima de seus resultados é mais fácil porque você pode saber de outras variáveis que possam estar influenciando nos seus resultados além da chance de identificar variáveis ocultas muitas vezes aleias ao perfil clínico do paciente.

Diante da urgência de um modelo mais bem calibrado, surgiram novos conceitos, a partir de aplicações recentes dos escores de risco, nas quais modelos próprios foram desenhados, validados e comparados a estes escores [14, 20, 71, 80, 82, 84, 85].

Desta forma, a formulação e validação interna do InCor tornou o modelo um instrumento importantíssimo para os pacientes atendidos no InCor-HCFMUSP. O InCor permitiria a monitorização dos resultados ao longo do tempo, o controle e o seguimento dos fatores de risco conhecidos e desconhecidos podendo alterar sua prevalência, sabendo que variáveis conhecidas sempre arrastam variáveis ocultas que podem influenciar nos resultados. O InCor é um modelo parcimonioso derivado de técnicas de *bootstrap*, o que faz com que suas variáveis sejam as mais eficazes para a predição da mortalidade. Ainda assim, ele é abrangente e pode ser utilizado em 80-90% de todos os procedimentos cardiovasculares realizados em adultos no InCor.

Por fim, podemos dizer que o InCor tem desempenho semelhante ao EuroSCORE e é mais simples do que este e que o 2000BP para prever a mortalidade nos pacientes operados no InCor. Sendo assim, sua validação externa é obrigatória para que ele se consolide como o primeiro modelo de risco em cirurgia cardíaca no Brasil.

7. Conclusões

- O **2000 Bernstein Parsonnet** revelou-se ruim na calibração e bom na discriminação, sendo questionável na validação para predizer a mortalidade nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP.
- O **EuroSCORE** revelou-se bom, tanto na calibração como na discriminação, sendo adequado na validação para predizer a mortalidade nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP.
- O **InsCor** foi desenvolvido localmente através de técnicas de *bootstrap*. Este modelo mostrou-se adequado na validação e comparável ao EuroSCORE e o 2000 Bernstein Parsonnet na predição de mortalidade nos pacientes operados de coronária e/ou valva no InCor-HCFMUSP.

8. Anexos

Anexo II. Variáveis utilizadas no estudo definidas pelos escores 2000 Bernstein-Parsonnet e EuroSCORE.

-2000BP.

1. Sexo feminino(6pt)

Abramov D, Tamariz MG, Sever JY, Christakis GT, Bhatnagar G, Heenan AL, Goldman BS, Fremes SE. *The influence of gender on the outcome of coronary artery bypass surgery.* Ann Thorac Surg. 2000;70(3):800-5; discussion 806.

Becker RC, Corrao JM, Alpert JS. *Coronary artery bypass surgery in women.* Clin Cardiol. 1988;11(7):443-8.

2. Idade

70-75 (2,5pt)

76-79 (7pt)

≥80 (11pt)

- . Hirose H, Amano A, Yoshida S, Takahashi A, Nagano N, Kohmoto T. *Coronary artery bypass grafting in the elderly.* Chest. 2000; 117(5):1262-70.

- Hochberg MS, Levine FH, Daggett WM, Akins CW, Austen WG, Buckley MJ. *Isolated coronary artery bypass grafting in patients seventy years of age and older: early and late results.* J Thorac Cardiovasc Surg. 1982;84(2):219-23.

- Alexander KP, Anstrom KJ, Muhlbaier LH, Grosswald RD, Smith PK, Jones RH, Peterson ED. *Outcomes of cardiac surgery in patients > or = 80 years: results from the National Cardiovascular Network.* J Am Coll Cardiol. 2000;35(3):731-8.

-Akins CW, Daggett WM, Vlahakes GJ, Hilgenberg AD, Torchiana DF, Madsen JC, Buckley MJ. *Cardiac operations in patients 80 years old and older.* Ann Thorac Surg. 1997; 64(3):606-14; discussion 614-5.

3. Insuficiência congestiva(Cardiomegalia) (2,5pt)

Pacientes com quadro clínico de insuficiência cardíaca congestiva (edema de pernas e/ou tornozelo e/ou dificuldade para respirar) em uso crônico de Inibidores da Enzima Convertedora de Angiotensina e Diuréticos incluídos na prescrição do dia anterior à cirurgia

e apresentando cardiomegalia no ecocardiograma transtorácico (diâmetros maiores que os estabelecidos).

O **Eletrocardiograma** (ECG) e a **radiografia do tórax** podem também conferir cardiomegalia e congestão pulmonar quando não houver métodos mais acurados.

-Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine, 8th ed.,2007.

4. DPOC (doença pulmonar obstrutiva crônica) severo. (6pt)

Paciente em uso prolongado de corticosteroides e/ou broncodilatadores > 6 meses confirmados na prescrição do dia anterior à cirurgia e com diagnóstico de DPOC severo nos testes de Função Pulmonar, aqui no Instituto são usados os critérios do *National Heart, Lung, and Blood Institute and World Health Organization*.

1. The GOLD Workshop Panel. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease: NHLBI/ WHO Workshop Report. Bethesda, Md: National Heart, Lung, and Blood Institute; 2001. NIH Publication N°. 2701.

TABLE 1
GOLD CLASSIFICATION OF COPD

Stage	Characteristics
0: At risk	Normal spirometry
I: Mild COPD	Chronic symptoms (cough, sputum production)
	FEV1/FVC < 70% . FEV1 greater than or equal to 80% predicted With or without chronic symptoms (cough, sputum production)
II: Moderate COPD	FEV1/FVC < 70%
	FEV1 greater than or equal to 30% to < 80% predicted
	IIa: FEV1 greater than or equal to 50% to < 80% predicted
	IIb: FEV1 greater than or equal to 30% to < 50% predicted With or without chronic symptoms (cough, sputum production, dyspnea)
III: Severe COPD	FEV1/FVC < 70%
	FEV1 < 30% predicted or FEV1 < 50% predicted plus respiratory failure or clinical signs of right heart failure

GOLD, Global Initiative for Chronic Obstructive Lung Disease; COPD, chronic obstructive pulmonary disease; FEV1, forced expiratory volume in one second; FVC, forced vital capacity; respiratory failure, arterial partial pressure of oxygen < 60 mm Hg with or without arterial partial pressure of carbon dioxide greater than or equal to 50 mm Hg while breathing air at sea level.

Data extracted from The GOLD Workshop Panel.[1]

5. Diabetes (3pt)

Paciente em uso de antidiabéticos orais e/ou qualquer tipo de Insulina na prescrição do dia anterior à cirurgia e/ou níveis de glicemia em jejum $\geq 126\text{mg/dl}$ ou $>200\text{mg/dl}$ 2 horas após

o teste de tolerância oral de glicose. Não foram incluídos os pacientes com níveis em jejum entre 100 e 125mg/dl ou entre 140 e 199mg/dl no teste de tolerância oral de glicose, considerados como pré-diabéticos.

<http://www.diabetes.org/about-diabetes.jsp>

6. Lesão detronco da coronária esquerda (TCE) - Estenose > 50% (2,5pt)
 Paciente com lesão de TCE>50% no Cateterismo e confirmado no laudo do estudo coronariográfico. Pacientes com lesão de TCE= 50% foram excluídos.

- Göl MK, Ozsöyler I, Sener E, Göksel S, Saritaş A, Taşdemir O, Bayazit K. *Is left main coronary artery stenosis a risk factor for early mortality in coronary artery surgery?* J Card Surg. 2000; 15(3): 217-22.

- Hata M, Shiono M, Sezai A, Iida M, Yoshitake I, Wakui S, Taoka M, Negishi N, Sezai Y. *Outcome of emergency conventional coronary surgery for acute coronary syndrome due to left main coronary disease.* Ann Thorac Cardiovasc Surg. 2006; 12(1):28-31.

7. Fração de Ejeção (FE)

- 30-49% (6,5pt)

- <30% (8pt)

O método usado pelo 2000BP para medida da Fração de Ejeção (FE) foi o Ecocardiograma Transtorácico o qual ficou definido como exame preferencial. O Eco-Stress, o Ecocardiograma transesofágico e a ventriculografia invasiva também foram utilizados. Quando usada a ventriculografia invasiva para avaliação da fração de ejeção utilizou-se o método angiográfico de Simpson, através do *software* do aparelho Philips Integris3000 (Philips, Netherlands), disponível no InCor.

*Pacientes com FE: 50% foram excluídos.

8. Hipertensão. - Pressão arterial sistêmica (HAS) \geq 140/90mmhg ou antecedente de hipertensão ou tratamento com anti-hipertensivos. (3pt)

O uso de medicação anti-hipertensiva foi confirmado na prescrição do dia anterior à cirurgia. A classificação diagnóstica de HAS no InCor é a mesma que a estabelecida pela American Heart Associatione American College of Cardiology.

Table 1. Classification and Management of Blood Pressure for Adults Aged 18 Years or Older

BP Classification	Systolic BP, mm Hg*	and	Diastolic BP, mm Hg*	Lifestyle Modification	Management*	
					Initial Drug Therapy	
					Without Compelling Indication	With Compelling Indications†
Normal	<120		<80	Encourage	No antihypertensive drug indicated	Drug(s) for the compelling indications‡
Prehypertension	120-139	or	80-89	Yes	Thiazide-type diuretics for most; may consider ACE inhibitor, ARB, β-blocker, CCB, or combination	Other antihypertensive drugs (diuretics, ACE inhibitor, ARB, β-blocker, CCB) as needed
Stage 1 hypertension	140-159	or	90-99	Yes	2-Drug combination for most (usually thiazide-type diuretic and ACE inhibitor or ARB or β-blocker or CCB)§	Other antihypertensive drugs (diuretics, ACE inhibitor, ARB, β-blocker, CCB) as needed
Stage 2 hypertension	≥160	or	≥100	Yes		

Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin-receptor blocker; BP, blood pressure; CCB, calcium channel blocker.
* Treatment determined by highest BP category.
† See Table 6.
‡ Treat patients with chronic kidney disease or diabetes to BP goal of less than 130/80 mm Hg.
§ Initial combined therapy should be used cautiously in those at risk for orthostatic hypotension.

The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. The JNC 7 Report. JAMA. 2003;289:(doi:10.1001/jama.289.19.2560).

9. – Obesidade mórbida. Acima de 1.5 vezes o peso ideal (1pt).

Existe muita controvérsia no que diz respeito à definição do peso ideal. O método usado pelo 2000BP é o mesmo estabelecido pelo Parsonnet em 1989 e não é atualmente utilizado. Classificamos a obesidade mórbida segundo o Índice de Massa Corporal (IMC), por ser um método prático e dos mais usados na avaliação de risco de morbimortalidade nas práticas atuais da cardiologia.

No InCor, todos os pacientes submetidos à cirurgia tem o peso e altura medidos no dia anterior à cirurgia, podendo facilmente ser calculados. Valores de IMC ≥ 40 foram considerados como Obesidade Mórbida.

IMC = peso / (altura)².

IMC - Classificação

- <18,4 = Muito magro
- 18,5 - 24,9 = Saudável
- 25 - 29,9 = Sobrepeso (Pré-Obesidade)
- 30 - 34,9 = Obesidade Grau I
- 35 - 39,9 = Obesidade Grau II
- ≥ 40 = Obesidade Grau III (Mórbida)

-Syrakas CA, Neumaier-Prauser P, Angelis I, Kiask T, Kemkes BM, Gansera B. *Is extreme obesity a risk factor for increased in-hospital mortality and postoperative morbidity after cardiac surgery? Results of 2251 obese patients with BMI of 30 to 50.* Cardiovascular Surgery, City Hospital Munich Bogenhausen.

10. BIA pré-operatório.

Presente antes da cirurgia. (4pt)

11. Reoperação.

Antecedente de cirurgia cardíaca, em que houve necessariamente a aperturado pericárdio.

Primeira reoperação: (10pt)

Segunda ou subsequente reoperação: (20pt)

12. Tratamento da valva aórtica. - Procedimento proposto (0pt)

Quando é proposto no pré-operatório o tratamento cirúrgico ou a exploração da valva aórtica.

13. Tratamento da valva mitral. - Procedimento proposto (4,5pt)

Quando é proposto no pré-operatório o tratamento cirúrgico ou a exploração da valva mitral.

14. Cirurgia de revascularização miocárdica + cirurgia valvar. Combinação da CRM e cirurgia valvar. (6pt)

Quando é proposto no pré-operatório o tratamento cirúrgico de uma ou mais valvas cardíacas junto à revascularização miocárdica.

15. Choque cardiogênico (diurese < 10 ml/hr) (12pt)

Paciente grave submetido à cirurgia em suporte hemodinâmico com drogas vasoativas e/ou BIA e que necessariamente apresente diurese < 10 ml/hr.

- Mehta RH, Grab JD, O'Brien SM, Glower DD, Haan CK, Gammie JS, Peterson ED; Society of Thoracic Surgeons National Cardiac Database Investigators. *Clinical characteristics and in-hospital outcomes of patients with cardiogenic shock undergoing coronary artery bypass surgery: insights from the Society of Thoracic Surgeons National Cardiac Database*. *Circulation*. 2008 Feb 19;117(7):876-85.

16. Endocardite Ativa (6,5pt)

Paciente submetido à cirurgia com quadro clínico sugestivo ou confirmado de endocardite ativa da valva nativa ou prótese valvar e que necessariamente faz uso de antibioticoterapia pré-operatória para endocardite ativa. No InCor são usados os critérios de Duke.

- Lukes AS, Bright DK, Durack DT. *Diagnosis of infective endocarditis*. *Infect Dis Clin North Am*. 1993 Mar;7(1):1-8.

-Haldar SM, O'Gara PT. *Infective endocarditis: diagnosis and management*. *Nat Clin Pract Cardiovasc Med*. 2006 Jun;3(6):310-7.

17. Endocardite Tratada (0pt)

Antecedente de diagnóstico de endocardite infecciosa com necessidade de internação hospitalar para tratamento clínico e/ou cirúrgico.

18. Aneurisma do ventrículo esquerdo ressecável (1,5pt)

Paciente com diagnóstico de aneurisma do ventrículo esquerdo (AVE) com indicação de tratamento cirúrgico no pré-operatório. Paciente com AVE sem indicação cirúrgica não foi considerado.

19. Tratamento cirúrgico da valva tricúspide proposto (5pt)

Quando é proposto, no pré-operatório, o tratamento cirúrgico ou exploração da valva tricúspide. Como em todos os casos foi considerado o princípio da “intenção de tratar”(O fato de tratar ou não cirurgicamente não exclui o risco da gravidade do procedimento).

20. Dependente de marcapasso (0pt)

Paciente submetido à cirurgia com marcapasso temporário ou definitivo por não apresentar automatismo cardíaco próprio.

21. IAM com 48h de evolução (4pt)

Paciente submetido à cirurgia com diagnóstico clínico e eletrocardiográfico de infarto agudo de miocárdio (IAM) com menos de 48 h de evolução e necessariamente confirmado pelos níveis seriados de enzimas cardíacas. Foram escolhidos níveis de Troponina T ou I e, na sequência, os níveis de CPK-MB para o diagnóstico. Para isto, foram usados no InCor os mesmos critérios da American Heart Association e da American College of Cardiology (Tab2). O ecocardiograma é útil na avaliação dos pacientes em UTI, com intubação oro traqueal e em ocasiões especiais pela dificuldade clínica, mas sempre confirmado por enzimas cardíacas.

- Pollack CV Jr, Antman EM, Hollander JE; American College of Cardiology; American Heart Association. *2007 focused update to the ACC/AHA guidelines for the management of patients with ST-segment elevation myocardial infarction: implications for emergency department practice*. Ann Emerg Med. 2008 Oct;52(4):344-355.e1. Epub 2008 Jun 2.

-Pollack CV Jr, Braunwald E. *2007 update to the ACC/AHA guidelines for the management of patients with unstable angina and non-ST-segment elevation myocardial infarction: implications for emergency department practice*. Ann Emerg Med. 2008 May;51(5):591-606. Epub 2007 Nov 26.

- Thielmann M, Massoudy P, Neuhäuser M, Tsagakis K, Marggraf G, Kamler M, Mann K, Erbel R, Jakob H. *Prognostic value of preoperative cardiac troponin I in patients undergoing emergency coronary artery bypass surgery with non-ST-elevation or ST-elevation acute coronary syndromes*. Circulation. 2006 Jul 4; 114(1 Suppl):I448-53.

22. Defeito agudo do septo ventricular (12pt)

Paciente submetido à cirurgia corretiva da comunicação interventricular aguda pós-infarto agudo de miocárdio diagnosticado no cateterismo cardíaco e/ou ecocardiograma.

23. Taquicardia ventricular (TV); fibrilação ventricular (FV) ou abortamento de morte súbita (PCR) (1pt)

Paciente com antecedente de TV, FV ou PCR com registro eletrocardiográfico submetido à cirurgia na mesma internação.

24. Asma (1pt)

Paciente submetido à cirurgia com diagnóstico de asma e uso crônico de inaladores adrenérgicos B2 agonistas e/ou corticoterapia conferidos na prescrição do dia anterior. Nas provas de função pulmonar, medidas de volume expiratório forçado no primeiro segundo (VEF1) < 80% ou VEF1/CVF < 70 % indicam quadro obstrutivo segundo a American Academy of Allergy, Asthma and Immunology e American Collegue of Allergy, Asthma and Immunology (ACAAI).

http://www.aaaai.org/professionals/resources/pdf/asthma_algorithm1998.pdf

Em que medida a DPOC difere da asma?

É importante distinguir a DPOC da asma, pois cada uma das doenças reage de forma mais eficaz a tratamentos diferentes.

As histórias naturais da DPOC e da asma são completamente diferentes, com etiologias e tratamentos distintos. Algumas das principais diferenças entre as duas doenças incluem:

- A idade com que os doentes começam a sofrer de asma é, geralmente, precoce, ao passo que os doentes com DPOC tendem a ser mais velhos - geralmente, com idade superior a 40 anos;
- Não existe ligação direta entre a asma e o tabagismo, ao passo que a DPOC está fortemente associada ao mesmo;
- A dispneia ou falta de ar quando há esforço físico é muito mais comum na DPOC do que na asma;
- Os sintomas da DPOC são progressivos e, geralmente, pioram com a idade, enquanto os sintomas da asma são em regra mais episódicos e repetitivos com o passar do tempo;
- Finalmente, os fenômenos inflamatórios são uma parte fundamental da asma, ao passo que há muito menos informações disponíveis da inflamação na DPOC.

25. IOT pré-operatória (4pt)

Paciente submetido à cirurgia em intubação endotraqueal por qualquer intercorrência e/ou etiologia.

26. Púrpura trombocitopênica idiopática (12pt)

Paciente submetido à cirurgia com diagnóstico de trombocitopenia associada à presença de megaplaquetas e tempo de sangramento aumentado, mas com tempo de protrombina e tempo de tromboplastina parcial ativada normal, sem causa aparente e muitas vezes confirmada pela presença de anticorpos antiplaquetários.

*Gernsheimer T, Stratton J, Ballem PJ, Slichter SJ. *Mechanisms of response to treatment in autoimmune thrombocytopenic purpura*. N Engl J Med 1989; 320:974-80.

27. Hipertensão pulmonar (Pressão média > 30mmHg) (11pt)

Ao igual que no 2000BP, usamos como método de avaliação da pressão pulmonar o ecocardiograma transtorácico devido a sua acessibilidade, praticidade e por não ser invasivo. No entanto, o valor de pressão arterial média requerido pelo modelo 2000BP não são fornecidos no laudo do ecocardiograma transtorácico do InCor. Este seria calculado a partir da pressão sistólica da artéria pulmonar utilizando métodos de correlação.

Por tanto, a pressão média da artéria pulmonar (PMAP) não deve ser confundida com a pressão sistólica da artéria pulmonar (PSAP), que é frequentemente relatado no ecocardiograma. A fórmula para o cálculo seria $\rightarrow PMAP = 0,61 \times (PSAP + 2)$.

Assim, uma PSAP de 40 mm Hg tipicamente significa uma PMAP > 25 mmHg.

-Simonneau G, Galiè N, Rubin LJ, et al (June 2004). "Clinical classification of pulmonary hypertension". J. Am Coll Cardiol. 43 (12 Suppl S): 5S–12S.

*Pacientes sem ecocardiograma transtorácico (nem externo), mas com medidas de pressão pulmonar no cateterismo cardíaco foram considerados.

28. Cirrose (12,5pt)

Paciente submetido à cirurgia com diagnóstico clínico e laboratorial de Cirrose hepática confirmada necessariamente com estudos de imagem (ultrassom ou tomografia) e/ou biópsia hepática. A avaliação dos exames pré-operatórios do paciente foi realizada seguindo as definições determinadas pela AASLD: *American Association for the Study of Liver Diseases*.

<http://www.aasld.org/practiceguidelines/Documents/Practice%20Guidelines/Hepatic.pdf>

<http://www.mayoclinic.com/print/cirrhosis/DS00373/DSECTION=all&METHOD=print>

29. Dependente de diálise. (13,5pt)

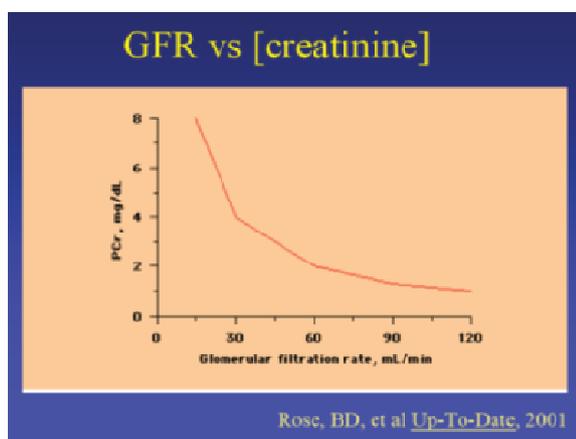
Paciente que realiza hemodiálise no pré-operatório de uma cirurgia cardíaca.

30. Insuficiência renal, aguda ou crônica. (3,5pt)

Um aspecto importante da doença renal é que, uma vez que um rim é danificado até certo grau, ele continua a deteriorar-se ainda que a doença renal subjacente possa ser curada. Isto é comumente conhecido como o ponto de não retorno (PNR). O ponto de não retorno é geralmente considerado como quando a creatinina sérica atinge 2,0 mg/dl nas medições dos EUA, ou cerca de 175 $\mu\text{mol/L}$ nas medições do sistema internacional.

Chronic Renal Failure

- Irreversible decline in renal function (gfr) due to a reduction in functioning renal mass
 - Stage 1 (gfr >90 ml/min)
 - Stage 2 (gfr 60-89 ml/min)
 - Stage 3 (gfr 30-59 ml/min)
 - Stage 4 (gfr 15-29 ml/min)
 - Stage 5 (gfr < 15 ml/min)
 - Serum creatinine: approximate indicator of renal function



Critérios de consenso para o diagnóstico de IRA são:

- Risco: creatinina sérica aumentada uma vez e meia o valor prévio ou a produção de débito urinário, em seis horas, de menos de 0.5 ml/kg de peso corpóreo;
- Injúria: creatinina sérica aumentada duas vezes o valor prévio ou a produção de débito urinário, em doze horas, de menos de 0.5 ml/kg de peso corpóreo;
- Falência: creatinina sérica aumentada três vezes ou débito urinário abaixo de 0.3 ml/kg, em vinte e quatro horas;
- Perda: IRA persistente ou mais de quatro semanas de perda completa da função renal;
- Insuficiência Renal estágio terminal: doença renal em estágio final, com mais de três meses (tratada como doença renal crônica).

- Hirose H, Amano A, Takahashi A, Nagano N. *Coronary artery bypass grafting for patients with non-dialysis-dependent renal dysfunction (serum creatinine > or =2.0 mg/dl)*. Eur J Cardiothorac Surg. 2001 Sep;20(3):565-72.

- Zakeri R, Freemantle N, Barnett V, Lipkin GW, Bonser RS, Graham TR, Rooney SJ, Wilson IC, Cramb R, Keogh BE, Pagano D. *Relation between mild renal dysfunction and outcomes after coronary artery bypass grafting*. Circulation. 2005 Aug 30; 112(9 Suppl):I270-5.

31. Aneurisma da aorta abdominal assintomático (0,5pt)
Paciente submetido à cirurgia cardíaca com diagnóstico (achado) de aneurisma da aorta abdominal assintomático.

32. Doença de carótida (bilateral ou 100% de oclusão unilateral) (2pt)

Paciente com obstrução de uma carótida (100%) ou ambas carótidas com lesão >70% diagnosticados com ultrassom Doppler.

33. Doença vascular periférica severa (3,5pt)

Paciente submetido à cirurgia com diagnóstico de doença vascular periférica severa mediante exames de ultrassonografia Doppler (Índice tornozelo-braquial < 0.5).

*Foram considerados os pacientes com ultrassonografia duplex, angiografia, tomografia ou ressonância magnética mostrando oclusão arterial severa.

http://www.vascularweb.org/patients/NorthPoint/Leg_Artery_Disease.html.

34. Reação a produtos sanguíneos (11pt)

Foram considerados pacientes submetidos à cirurgia com antecedente de reação prévia ou atual a produtos sanguíneos e/ou hemoderivados referidos pelo paciente e/ou familiar e/ou prontuário, caracterizados por prurido, avermelhamento, hipotensão, lesões dérmicas, edema de glotes, etc. diagnosticados pelo médico e excluindo outras etiologias,

35. Distúrbios neurológicos severos (acidente vascular cerebral (AVC) curado, paraplegia, distrofia muscular, hemiparesia)(5pt)

- AVC curado. - Paciente submetido à cirurgia com antecedente de acidente vascular cerebral (hemorrágico ou isquêmico) com ou sem sequela, diagnosticado e internado no hospital.

- Paraplegia. - Paciente submetido à cirurgia com perda de controle e sensibilidade nos membros inferiores, impedindo-o de caminhar e dificultando a posição de sentado. Pacientes comparalisia da metade do corpo (hemiplegia) também foram considerados.

- Distrofia muscular. - Paciente submetido à cirurgia com diagnóstico prévio de doença genética caracterizada por fraqueza progressiva e degeneração dos músculos esqueléticos usados no movimento voluntário.

- Hemiparesia. - Paciente submetido à cirurgia com déficit parcial da capacidade de executar movimentos voluntários (menos severo que hemiplegia) na metade do corpo.

http://www.nlm.nih.gov/cgi/mesh/2008/MB_cgi?field=uid&term=D010264

<http://www.mda.org/disease/>

http://www.nlm.nih.gov/cgi/mesh/2008/MB_cgi?field=uid&term=D006429

36. PTCAoufalhanacateterização. (5,5pt)

Paciente submetido à cirurgia com laudo médico de prévia realização de angioplastia, colocação de stent (artéria culpada ou não) ou de não ter conseguido fazer o estudo angiográfico por dificuldade técnica (cateterização frustra). Todos os cateterismos foram analisados previamente à cirurgia e a existência de stent foi registrada.

37. Abusodesubstâncias. (4,5pt)

Paciente submetido à cirurgia com diagnóstico de abuso de substâncias. No InCor são usados os critérios do DSM IV. Pacientes com diagnóstico de dependência de substâncias foram considerados por estarem numa fase mais avançada que o abuso de substâncias. Por tanto, foram considerados: tabagistas severos, usuários de drogas (cocaína, morfina e outros), usuários de psicotrópicos (antidepressivos, ansiolíticos e outros), portadores de doenças crônicas em uso de analgésicos, quimioterápicos e/ou corticoterapia de forma prolongada (>12meses) e pacientes com alcoolismo crônico.

- ABUSO DE SUBSTÂNCIAS - DSM.IV (Manual Diagnóstico e Estatístico de Transtornos Mentais 4º Ed.).

Tabagismo:

Foram considerados pacientes submetidos à cirurgia com diagnóstico de tabagismo severo
Ex-tabagistas não foram considerados.

O teste mais usado no InCor para medir a quantidade de cigarros que uma pessoa tenha fumado durante um longo período de tempo é multiplicando o número de maços de cigarros fumados por dia pelo número de anos que a pessoa fumou. Por exemplo, um maço ano é igual a fumar 20 cigarros por dia durante 1 ano, ou 40 cigarros por dia na metade de um ano assim por diante, Assim por exemplo, calcular “Maços/Ano”

Número de maços por ano = (o número de cigarros fumados por dia x o número de anos fumados) / 20 Por exemplo: um paciente que fumava 15 cigarros por dia durante 40 anos tem um $(15 \times 40) / 20 = 30$ maços/ano. (1 maço tem 20 cigarros). Um maço/ano, significa que está fumando 20 cigarros por dia durante um ano.

*A OMS classifica em: Leve (<10 cigarros/dia); Moderado (10-20 cigarros/dia); Severo (>20 cigarros/dia).

Ex-tabagista: > 1 ano sem fumar.

http://www.cancer.gov/Templates/db_alpha.aspx?CdrID=306510.

National Cancer Institute Definition of Pack Year Retrieved from "http://en.wikipedia.org/wiki/Pack_year"

-EuroSCORE.

1. Idade (paracada cinco anos a partir dos 60a). (1pt)
2. Sexo feminino (1pt)
3. DPOC (usocronicode broncodilatadores ou esteroides) (1pt)
Pacientes com DPOC diagnosticados nos testes de função pulmonar, em uso crônico de broncodilatadores ou esteroides.

4. Arteriopatia extracardiaca. - Uma ou mais das seguintes apresentações:
(2pt)
Claudicação intermitente; oclusão de carótida ou estenose > 50% diagnosticado no ultrassom Doppler ou ressonância magnética; intervenção prévia ou futura da aorta abdominal, das artérias dos membros ou das carótidas.
5. Disfunção neurológica. - Transtorno severo da deambulação ou das atividades diárias. (2pt)
6. Cirurgia cardíaca prévia. (qualquer procedimento com abertura prévia do pericárdio). (3pt)
7. Creatinina sérica >2,26mg/dl. (2pt)
8. Endocardite ativa. - Paciente operado com diagnóstico de endocardite ativa em vigência de antibioticoterapia. (3pt)
9. Estado pré-operatório crítico. - Um ou mais dos seguintes:
Taquicardia ventricular, fibrilação ventricular, abortamento de morte súbita, massagem cardíaca pré-operatória, IOT antes de chegar à sala de cirurgia, suporte inotrópico pré-operatório, balão intra-aórtico e/ou falha renal aguda (anúria ou oligúria <10 ml/hr) (3pt)
10. Angina instável. - Angina de repouso usando nitrato endovenoso até chegar à sala de cirurgia (2pt)
11. Disfunção do ventrículo esquerdo.
O ecocardiograma transtorácico ficou definido como exame preferencial. O eco stress ou Ecocardiograma transesofágico e a ventriculografia invasiva também foram utilizados. Quando usada a ventriculografia invasiva para avaliação da fração de ejeção utilizou-se o método angiográfico de Simpson, através do *software* do aparelho Philips Integris 3000 (Philips, Netherlands), disponível no InCor.
- | | | |
|-----------|----------|-------|
| Moderado: | 30 e 50% | (1pt) |
| Severo: | <30% | (3pt) |
12. Infarto miocárdico recente. - <90 dias (2pt)
Paciente com diagnóstico de infarto agudo de miocárdio confirmado com níveis de Troponina T ou I ou níveis de CPK-MB. Foram considerados os pacientes encaminhados com diagnóstico de infarto agudo de miocárdio sempre que estes apresentaram laudo dos exames confirmatórios.
13. Hipertensão pulmonar. - Pressão sistólica da artéria pulmonar >60mmhg (2pt)

Diagnóstico feito através do ecocardiograma. Foram consideradas as medidas feitas no cateterismo cardíaco.

14. Emergência. - Cirurgia antes do próximo dia de trabalho. (2pt)

15. Outro procedimento cardíaco importante além da cirurgia de revascularização miocárdica. (2pt)

16. Cirurgia na aorta torácica. - Cirurgias incluindo aorta ascendente, arco aórtico ou aorta descendente. (3pt)

17. Rotura septal pós-infarto. – Diagnosticado no Ecocardiograma transtorácico e/ou transesofágico e/ou ventriculografia invasiva. (4pt)

-Definições das variáveis não contempladas pelos escores analisados:

*Dislipidemia: Exames sanguíneos pré-operatórios confirmando um ou mais dos valores citados:

-Colesterol Total > 240mg/dl

-LDL >160 mg/dl

-VLDL > 40 mg/dl

-HDL < 35mg/dl(Homem) ou <45mg/dl (mulher)

-Triglicérides >200mg/dl

*Circulation.2004;110:e296-e297. **Measurement of Cholesterol**, 2004 American Heart Association, Inc.

*Fibrilação atrial: Diagnosticado no ECG ou no monitor cardíaco prévio ao começo da cirurgia.

*Átrio esquerdo: medida do átrio esquerdo >40 mm feito no ecocardiograma mais recente.

Anexo III.

HOSPITAL DAS CLÍNICAS
DA
FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO
DIRETORIA CLÍNICA
COMISSÃO DE ÉTICA PARA ANÁLISE DE PROJETOS DE PESQUISA - CAPPesq
CADASTRO DE PROTOCOLO DE PESQUISA

Registro (uso reservado à Secretaria da CAPPesq)	
Nº do Protocolo: 1063/07	Tipo: Humanos
Instituto: INCOR	
Registro on-line nº: 1575	Data de Entrada: 12/11/2007

Este projeto envolve:

Pacientes HC	Não
Médicos ou Funcionários HC (como sujeitos de pesquisa)	Não
Documentos HC(Prontuários e Outros)	Não
Materiais estocados no HC	Não

1. Título do Protocolo de Pesquisa

“Validação dos Escores 2000 Bernstein-Parsonnet e Euroscore: Análise pré-operatória dos fatores de risco relacionados à mortalidade hospitalar na população brasileira atendida no InCor (Escore-InCor?)”.
--

9. Referências Bibliográficas

1. Dubois RW, Rogers WH, Moxley JH, Draper D, Brook RH. *Hospital Inpatient Mortality*. N Engl J Med 1987; 317: 1674-80.
2. Green J, Wintfeld N, Sharkey P, Passman LJ. *The importance of severity of illness in assessing hospital mortality*. JAMA 1990; 263: 241-246.
3. Jones RH, Hannan EL, Hammermeister KE, DeLong ER, O'Connor GT, Luepker RV, Parsonnet V, Pryor DB. *For The Working Group Panel on the Cooperative CABG Database Project*. J Am Coll Cardiol 1996; 28: 1478-87.
4. Clark RE. *Calculating risk and outcomes: The Society of Thoracic Surgeons Database*. Ann Thorac Surg 1996; 62: S2-5.
5. DeLong ER, Peterson ED, DeLong DM, Muhlbaier LH, Hackett S, Mark DB. *Comparing risk-adjustment methods for provider profile*. Stat Methods 1997, 16 (23); 2645-64.
6. Daley J. *Criteria by which to evaluate risk-adjusted outcomes programs in cardiac surgery*. Ann Thorac Surg 1994; 58: 1827-35.
7. Kennedy JW, Kaiser GC, Fisher LD, Myers W, Mudd JG, Coggin J, Ryan TJ. *Multivariate discriminant analysis of the clinical and angiographic predictors of operative mortality from the Collaborative Study in Coronary Artery Surgery (CASS)*. J Thorac Cardiovasc Surg 1980; 80: 876-87.
8. Health Care Financing Administration. *Medicare Hospital Mortality Information, volume I, 1986* (HCFA publication No. 01-002). Washington, DC: US Government Printing Office; 1987.
9. Mark DB. *Implications of costs in treatment selection for patients with coronary disease*. Ann Thorac Surg 1996; 61: S12-S15.

10. Moustapha A, Anderson HV. *Revascularization interventions for ischemic heart disease*. *Curr Opin Cardiol* 2000; 15(6): 463-71.
11. Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, the EuroSCORE study group. *European system for cardiac operative risk evaluation (EuroSCORE)*. *Eur J Cardio-thorac Surg* 1999; 16: 9-13.
12. Chatfield C. *Model uncertainty, data mining and statistical inference (with discussion)*. *Journal of the Royal Statistical Society, Series A* 1995; 158: 419-466.
13. Omar RZ, Ambler G, Royston P, Eliahoo J, Taylor KM. *Cardiac Surgery Risk Modeling for Mortality: A Review of Current Practice and Suggestions for Improvement*. *Ann Thorac Surg* 2004; 77:2232-7.
14. Ranucci M, Castelvechio S, Menicanti L, Frigiola A, Pelissero G. *Accuracy, calibration and clinical performance of the EuroSCORE: can we reduce the number of variables?* *Eur J Cardio-thorac Surg* 2010; 37: 724-29.
15. Hillis LD, Smith PK, Anderson JL, Bittl JA, Bridges CR, Byrne JG, Cigarroa JE, DiSesa VJ, Hiratzka LF, Hutter AM Jr, Jessen ME, Keeley EC, Lahey SJ, Lange RA, London MJ, Mack MJ, Patel MR, Puskas JD, Sabik JF, Selnes O, Shahian DM, Trost JC, Winniford MD, Jacobs AK, Anderson JL, Albert N, Creager MA, Ettinger SM, Guyton RA, Halperin JL, Hochman JS, Kushner FG, Ohman EM, Stevenson W, Yancy CW. *2011 ACCF/AHA guideline for coronary artery bypass graft surgery: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines*. *J Thorac Cardiovasc Surg* 2012 Jan; 143 (1): 4-34.
16. Altman DG, Royston P. *What do we mean by validating a prognostic model?* *Statist. Med.* 2000; 19:453-73.
17. Harrell FE. *Regression modeling strategies*. Berlin: Springer-Verlag, 2001.

18. Ivanov J, Tu JV, Naylor CD. *Ready-Made, Recalibrated, or Remodeled? Issues in the Use of Risk Indexes for Assessing Mortality after Coronary Artery Bypass Graft Surgery*. *Circulation*. 1999; 99: 2098-2104.
19. Staudinger O, Ostermann H, Laufer G, Schistek R, Staudinger B, Tilg B. *Evaluation of cardiac scoring models for an Austrian cardiac register*. *Interact CardioVasc Thorac Surg* 2007; 6: 470-3.
20. D'Errigo P, Seccareccia F, Rosato S, Manno V, Badoni G, Fusco D, Perucci CA, the 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 Cardio-thorac Surg* 2008; 33: 325-33.
21. Zheng Z, Li Y, Zhang S, Hu S, the 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 Cardio-thorac Surg* 2009; 35:54-61.
22. 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(suppl 1): 3-12.
23. Bernstein AD, Parsonnet V. *Bedside Estimation of Risk as an Aid for Decision-Making in Cardiac Surgery*. *Ann Thorac Surg* 2000; 69: 823-8.
24. Roques F, Gabrielle F, Michel P, De Vincentiis C, Baudet M. *Quality of care in adult heart surgery: proposal for a self-assessment approach based on a French multicenter study*. *Eur J Cardiothorac Surg* 1995; 9: 433-40.
25. Tu JV, Jaglal SB, Naylor CD, and the Steering Committee of the Provincial Adult Cardiac Care Network of Ontario. *Multicenter Validation of a Risk Index for Mortality, Intensive Care Unit Stay, and Overall Hospital Length of Stay after Cardiac Surgery*. *Circulation*. 1995; 91: 677-84.

26. Hannan EL, Kilburn H, O'Donnell JF, Lukacik G, Shields EP. *Adult Open Heart Surgery in New York State. An Analysis of Risk Factors and Hospital Mortality Rates.* JAMA. 1990; 264: 2768-74.
27. Edwards FH, Grover FL, Shroyer LW, Schwartz M, Bero J. *The Society of Thoracic Surgeons National Cardiac Surgery Database: Current Risk Assessment.* Ann Thoracic Surg 1997; 63: 903-8.
28. Nashef SAM, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, Wyse RKH, Ferguson TB. *Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery.* Eur J Cardio-thorac Surg 2002; 22: 101-5.
29. Roques F, Nashef SAM, Michel P, Pinna Pintor P, David M, Baudet E. *The EuroSCORE Study Group. Does EuroSCORE work in individual European countries?* Eur J Cardiothorac Surg 2000; 18: 27-30.
30. Cortina JM. *Scores de gravedad y complejidad em cirujía cardíaca. Usos y Limitaciones.* Rev Esp Cardiol. 2005; 58(5): 473-6.
31. Loop FD, Berrettoni JN, Pichard A, Siegel W, Razavi M, Effler DB. *Selection of the candidate for myocardial revascularization: A profile of high risk based on multivariate analysis.* J Thorac Cardiovasc Surg 1975; 69:40-51.
32. Paiement B, Pelletier C, Dyrda I, Maillé JG, Boulanger M, Taillefer J, Sahab P, Delorme M, Dupont E. *A simple classification of the risk in cardiac surgery.* Can J Anaesth. 1983; 30:61-8.
33. Tremblay NA, Hardy JF, Perrault J. *A simple classification of the risk in cardiac surgery: The first decade.* Soc Cardiovas Anesthesiol 1992 Abstracts, p167.

34. Kouchoukos NT, Ebert PA, Grover FL, Lindesmith GG. *Report of the Ad Hoc Committee on Risk Factors for Coronary Artery Bypass Surgery*. Ann Thorac Surg. 1988; 45: 348-9.
35. Edwards FH, Albus RA, Zajtchuk R, Graeber GM, Barry MJ, Rumisek JD, Arishita G. *Use of a Bayesian statistical model for risk assessment in coronary artery surgery*. Ann Thorac Surg 1988; 45: 437-440.
36. Edwards FH, Albus RA, Zajtchuk R, Graeber GM, Barry M. *A quality assurance model of operative mortality in coronary artery surgery*. Ann Thorac Surg 1989; 47:646-49.
37. Higgins TL, Estafanos FG, Loop FD, Beck GJ, Blum JM, Paranandi L. *Stratification of Morbidity and Mortality Outcome by Preoperative Risk Factors in Coronary Artery Bypass Patients. A Clinical Severity Score*. JAMA. 1992; 267: 2344-48.
38. O'Connor GT, Plume SK, Olmstead EM, Coffin LH, Morton JR, Maloney CT, Nowicki ER, Levy DG, Tryzelaar JF, Hernandez F, Adrian L, Casey KJ, Bundy D, Soule DN, Marrin CAS, Nugent WC, Charlesworth DC, Clough R, Katz S, Leavitt BJ, Wennberg JE. *Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group*. Circulation 1992; 85: 2110-18.
39. Nashef SAM, Carey F, Silcock MM, Oommen PK, Levy RD, Jones MT. *Risk stratification for open heart surgery: trial of the Parsonnet system in a British hospital*. BMJ 1992; 305: 1066-7.
40. Tuman KJ, McCarthy RJ, March RJ, Najafi H, Ivankovich AD. *Morbidity and duration of ICU stay after cardiac surgery. A model for preoperative risk assessment*. Chest. 1992 Jul; 102(1): 36-44.

41. Grover FL, Johnson RR, Shroyer LW, Marshall G, Hammermeister KE. *The Veterans Affairs Continuous Improvement in Cardiac Surgery Study*. Ann Thorac Surg 1994; 58: 1845-51.
42. Edwards FH, Clark RE, Schwartz M. Coronary artery bypass grafting: *The Society of Thoracic Surgeons National Database Experience*. Ann Thorac Surg 1994; 57: 12-9.
43. Turner JS, Morgan CJ, Thakrar B, Pepper JR. *Difficulties in predicting outcomes in cardiac surgery patients*. Crit Care Med 1995; 23: 1843-50.
44. Parsonnet V, Dean D, Bernstein AD. *Clinical Usefulness of Risk-Stratified Outcome Analysis in Cardiac Surgery in New Jersey*. Ann Thorac Surg 1996; 61: S8-11.
45. Magovern JA, Sakert T, Magovern GJ, Benckart DH, Burkholder JA, Liebler GA, Magovern GS. *A Model That Predicts Morbidity and Mortality After Coronary Artery Bypass Graft Surgery*. J Am Coll Cardiol 1996; 28: 1147-53.
46. Pons JMV, Granados A, Espinas JA, Borrás JM, Martín I, Moreno V. *Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model*. Eur J Cardio-thorac Surg 1996; 11: 415-23.
47. Gabrielle F, Roques F, Michel P, Bernard A, de Vicentis C, Roques B, Brenot R, Baudet E, David M. *Is the Parsonnet score a good predictive score of mortality in adult cardiac surgery?: assessment by a French multicentre study*. Eur J Cardio-thorac Surg 1997; 11: 406-14.
48. Martínez-Alario J, Tuesta ID, Plasencia E, Santana M, Mora ML. *Mortality Prediction in Cardiac Surgery Patients: Comparative Performance of Parsonnet and General Severity Systems*. Circulation. 1999; 99: 2378-2382.

49. Geissler HJ, Holzl P, Marohl S, Kuhn-Régnier, Mehlhorn U, Sudkamp M, Rainer de Vivie E. *Risk stratification in heart surgery: comparison of six score systems*. Eur J Cardio-thorac Surg 2000; 17: 400-6.
50. 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 Cardio-thorac Surg 2001; 20: 961-66.
51. Sergeant P, de Worm E, Meyns B. *Single centre, single domain validation of the Euroscore on a consecutive sample of primary and repeat CABG*. Eur J Cardio-thorac Surg 2001; 20(6): 1176-82.
52. Nashef SAM, Roques F, Michel P. *The logistic Euroscore*. European Heart Journal 2003; 24: 1-2.
53. Nilsson J, Algotsson L, Högglund P, Luhrs C, Brandt J. *Early Mortality in Coronary Bypass Surgery. The EuroSCORE versus the Society of Thoracic Surgeons Risk Algorithm*. Ann Thorac Surg 2004; 77: 1235-40.
54. Syed AU, Fawzy H, Farag A, Nemlander A. *Predictive Value of EuroSCORE and Parsonnet Scoring in Saudi Population*. Heart Lung and Circulation 2004; 13: 384-88.
55. Chen CC, Wang CC, Hsieh SR, Tsai HW, Wei HJ, Chang Y. *Application of European system for cardiac operative risk evaluation (EuroSCORE) in coronary bypass surgery for Taiwanese*. Interact CardioVasc Thorac Surg 2004; 3: 562-5.
56. Nowicki ER, Birkmeyer NJO, Weintraub RW, Leavitt BJ, Sanders JH, Dacey LJ, Clough RA, Quinn RD, Charlesworth DC, Sisto DA, Uhlig PN, Olmstead EM, O'Connor GT. *Multivariate prediction of in-hospital mortality associated with aortic and mitral surgery in Northern New England*. Ann Thorac Surg. 2004 Jun; 77(6): 1966-77.

57. Ambler G, Omar RZ, Royston P, Kinsman R, Keogh BE, Taylor KM. *Generic, simple risk stratification model for heart valve surgery*. *Circulation*. 2005; 112: 224-31.
58. Shanmugam G, West M, Berg G. *Additive and Logistic EuroSCORE performance in high risk patients*. *Interact CardioVasc Thorac Surg* 2005; 4: 299-303.
59. Jin R, Grunkemeier GL. *Additive vs. logistic risk models for cardiac surgery mortality*. *Eur J Cardiothorac Surg* 2005; 28: 240-43.
60. Seccareccia F, Perucci CA, D'Errigo P, Arcà M, Fusco D, Rosato S, Greco D. on behalf of the research group of the Italian CABG Outcome Study. *The Italian CABG Outcome Study: short-term outcomes in patients with coronary artery bypass graft surgery*. *Eur J Cardio-thorac Surg* 2006; 29: 56-64.
61. Hannan EL, Wu C, Bennett EV, Carlson RE, Culliford AT, Gold JP, Higgins RSD, Isom OW, Smith CR, Jones RH. *Risk stratification of in-hospital mortality for coronary artery bypass graft surgery*. *J Am Coll Cardiol* 2006; 47: 661-8.
62. Nilsson J, Algotsson L, Høglund P, Luhrs C, Brandt J. *Comparison of 19 pre-operative risk stratification models in open-heart surgery*. *European Heart Journal* 2006; 27: 867-74.
63. Berman M, Stamler A, Sahar G, Georghiou GP, Sharoni E, Brauner R, Medalion B, Vidne BA, Kogan A. *Validation of the 2000 Bernstein-Parsonnet Score Versus the EuroSCORE as a Prognostic Tool in Cardiac Surgery*. *Ann Thorac Surg* 2006; 81: 537-41.
64. Moraes F, Duarte C, Cardoso E, Tenório E, Pereira V, Lampreia D, Wanderley J, Moraes CR. *Assessment of the EuroSCORE as a predictor for*

- mortality in myocardial revascularization surgery at the Heart Institute of Pernambuco.* Braz J Cardiovasc Surg 2006; 21(1): 29-34.
65. Yap CH, Reid C, Yui M, Rowland MA, Mohajeri M, Skillington PD, Seevanayagam S, Smith JA. *Validation of the EuroSCORE model in Australia.* Eur J Cardio-thorac Surg 2006; 29: 441-6.
66. Ribeiro ALP, Gagliardi SPL, Nogueira JLS, Silveira LM, Colosimo EA, Nascimento CAL. *Mortality related to cardiac surgery in Brazil, 2000-2003.* J Thorac Cardiovasc Surg. 2006; 131: 907-9.
67. Ivanov J, Borger MA, Rao V, David TE. *The Toronto Risk Score for adverse events following cardiac surgery.* Can J Cardiol 2006; 22(3): 221-7.
68. Zheng Z, Zhou Y, Gao HW, Hu SS. *Prediction value for operative mortality of four different coronary artery bypass graft risk stratification models in Chinese patients.* Zhonghua Xin Xue Guan Bing Za Zhi. 2006 Jun; 34(6): 504-7.
69. De Varennes B, Lachapelle K, Cecere R, Ergina P, Shum-Tim D, Tchervenkov C, Sampalis J. *Application of the Parsonnet scoring system for a Canadian cardiac surgery program.* Can J Cardiol 2007; 23(13): 1061-65.
70. Hannan EL, Wu C, Bennett EV, Carlson RE, Culliford AT, Gold JP, Higgins RSD, Smith CR, Jones RH. *Risk Index for Predicting In-Hospital Mortality for Cardiac Valve Surgery.* Ann Thorac Surg 2007; 83: 921-30.
71. Antunes PE, Eugenio L, Ferrão de Oliveira J, Antunes MJ. *Mortality risk prediction in coronary surgery: a locally developed model outperforms external risk models.* Interact Cardiovasc Thorac Surg. 2007; 6(4):437-41.
72. 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: 293-7.

73. Gomes RV, Tura B, Mendonça Filho HT, Almeida Campos LA, Rouge A, Matos Nogueira PM, Oliveira Fernandes, Rocha Dohmann HF, Cunha BA. *A First Postoperative Day Predictive Score of Mortality for Cardiac Surgery*. Ann Thorac Cardiovasc Surg 2007; 13(3):159-64.
74. Campagnucci VP, Silva AM, 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.
75. Lafuente S, Trilla A, Bruni L, Gonzalez R, Bertrán MJ, Pomar JL, Asenjo MA. *Validation of the EuroSCORE Probabilistic Model in Patients Undergoing Coronary Bypass Grafting*. Rev Esp Cardiol. 2008; 61(6): 589-94.
76. Reid C, Billah B, Dinh D, Smith J, Skillington P, Yii M, Seevanayagam S, Mohajeri M, Shardey G. *An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: The AusSCORE*. J Thorac Cardiovasc Surg. 2009; 138(4):904-10.
77. Parolari A, Pesce LL, Trezzi M, Cavalotti L, Kassem S, Loardi C, Pacini D, Tremoli E, Alamanni F. *EuroSCORE Performance in Valve Surgery: A Meta-analysis*. Ann Thorac Surg. 2010; 89:787-93.
78. Carvalho MR, Souza e Silva NA, Klein CH, Oliveira GM. *Applications of the EuroSCORE in coronary artery bypass surgery in public hospitals in Rio de Janeiro, Brazil*. Rev Bras Cir Cardiovasc. 2010; 25(2):209-17.
79. Contreras AE, Ferrero Guadagnoli A, Brenna EJ, Pogonza P, Coppa LA, Paladini G, Salomone OA, Chiotti IG, Estrada CD. *EuroSCORE system validation for cardiac surgery in Cordoba city*. Rev Fac Cien Med Univ Nac Cordoba. 2010; 67(3): 104-7.

80. Billah B, Reid CM, Shardey GC, Smith JA. *A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort.* Eur J Cardio-thorac Surg 2010; 37: 1086-1092.
81. Malik M, Chauhan S, Malik V, Gharde P, Kiran U, Pandey RM. *Is EuroSCORE applicable to Indian patients undergoing cardiac surgery?* Ann Card Anaesth 2010; 13: 241-5.
82. Shih HH, Kang PL, Pan JY, Wu TH, Wu CT, Lin CY, Lin YH, Chou WT. *Performance of European system for cardiac operative risk evaluation in Veterans General Hospital Kaohsiung cardiac surgery.* J Chin Med Assoc. 2011; 74(3):115-20.
83. Mejía OA, Lisboa LA, Puig LB, Dias RR, Dallan LA, Pomerantzeff PM, Stolf NA. *The 2000 Bernstein-Parsonnet score and EuroSCORE are similar in predicting mortality at the Heart Institute, USP.* Rev Bras Cir Cardiovasc. 2011; 26(1):1-6.
84. Berg KS, Stenseth R, Pleym H, Wahba A, Videm V. *Mortality risk prediction in cardiac surgery: comparing a novel model with the EuroSCORE.* Acta Anaesthesiol Scand. 2011; 55(3):313-21.
85. Qadir I, Perveen S, Furnaz S, Shahabuddin S, Sharif H. *Risk stratification analysis of operative mortality in isolated coronary artery bypass graft patients in Pakistan: comparison between additive and logistic EuroSCORE models.* Interact Cardiovasc Thorac Surg. 2011; 13(2):137-41.
86. Nashef SAM, Roques F, Sharples L, Nilsson J, Smith C, Lockowandt U. *EUROSCORE II.* Interact Cardiovasc Thorac Surg 2011; 13(Suppl. 2): S73. (Presented at 25th EACTS Annual Meeting; 2011 October 1-5; Lisbon, Portugal. Abstracts)
87. Lisboa LAF, Moreira LFP, Dallan LAO, Pomerantzeff PMA, Costa R, Puig LB, Marcial MLB, Jatene FB, Stolf NAG. *Vinte e três anos de evolução da*

- cirurgia cardiovascular em serviço de atendimento terciário no Brasil: Análise de 67000 procedimentos registrados no banco de dados. *Arq Bras Cardiol.* 2007 Sep; 89(1): 205.
88. Lemeshow S, Hosmer DW Jr. *A review of goodness of fit statistic for use in the development of logistic regression models.* *Am J Epidemiol* 1982; 115:92-106.
89. Hosmer DW, Lemeshow S. *Applied logistic regression.* 2nded. New York: John Wiley and Sons; 2000.
90. Jones CM, Athanasiou T. *Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests.* *Ann Thorac Surg.* 2005; 79: 16-20.
91. Austin P, Tu J. *Bootstrap methods for developing predictive models.* *The American Statistician* 2004; 58(2):131—7.
92. Efron B. and Tibshirani R. (1993). *An Introduction to the Bootstrap.* London: Chapman and Hall.
93. Altman DG, Vergouwe Y, Royston P, Moons KGM. *Prognosis and prognostic research: validating a prognostic model.* *BMJ* 2009; 338: b605.
94. SAS System version 9.2 for Microsoft Windows [computer software]. North Carolina State University, USA, SAS Institute Inc.; 2008.
95. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bastos PG, Sirio CA, Murphy DJ, Lotring T. *APACHE III prognostic system.* *Chest* 1991; 100: 1619-1636.
96. Steg PG, Dabbous OH, Feldman LJ, Cohen-Solal A, Aumont MC, Lopez-Sendon J, Budaj A, Goldberg RJ, Klein W, Anderson FA Jr. *Determinants and prognostic impact of heart failure complicating acute coronary*

- syndromes: observations from the Global Registry of Acute Coronary Events (GRACE)*. *Circulation* 2004; 109: 494-9.
97. Becker RH, Zimmerman JE, Knaus WA. *The use of APACHE III to evaluate ICU length of stay, resource use and mortality after coronary artery surgery*. *J Cardiovasc Surgery* 1995; 36: 1-11.
98. Ryan TA, Rady MY, Bashour CA, Leventhal M, Lytle B, Starr NJ. *Predictors of outcome in cardiac surgery patients with prolonged intensive care stay*. *Chest* 1997; 112: 1035-42.
99. Poses RM, Cebul RD, Collins M, Fager SS. *The importance of disease prevalence in transporting clinical prediction rules*. *Ann Intern Med*. 1986; 105: 586-91.
100. Wasson JH, Sox HC, Neff RK, Goldman L. *Clinical prediction rules. Applications and methodological standards*. *N Engl J Med* 1985; 313: 793-99.
101. Charlson M, Ales KL, Simon R, Mackenzie CR. *Why predictive indexes perform less well in validation studies: is it magic or methods?* *Arch Intern Med*. 1987; 147: 2155-61.
102. Roper WL, Winkenwerder W, Hackbarth GM, Krakauer H. *Effectiveness in health care. An initiative to evaluate and improve medical practice*. *N Engl J Med* 1988; 319: 1197-1202.
103. The Society of Thoracic Surgeons. *Report of the Ad hoc Committee on physician-specific mortality rates for cardiac surgery*. *Ann Thorac Surg* 1993; 56: 1200-2.
104. Hammermeister KE, Johnson R, Marshall G. *Continuous assessment and improvement in quality of care. A model from the Department of Veterans Affairs Cardiac Surgery*. *Ann Thorac Surg* 1994; 219: 281-90.

105. Concato J, Feinstein AR, Holford TR. *The risk of determining risk with multivariable models*. Annals of Internal Medicine 1993; 118: 201-210.
106. Chatfield C. *Model uncertainty, data mining and statistical inference (with discussion)*. Journal of the Royal Statistical Society, Series A 1995; 158: 419-466.
107. Nashef SA. *Applying and evaluating risk models*. Eur J Cardiothorac Surg. 2012; 41(2): 314-5.
108. Clark RE. *Calculating risk and outcome: The Society of Thoracic Surgeons Database*. Ann Thorac Surg 1996; 62: S2-5.
109. Marshall G, Shroyer ALW, Grover FL, Hammermeister KE. *Bayesian-logit. Model for risk assessment in coronary artery bypass grafting*. Ann Thorac Surg 1994; 57: 1492-500.
110. Keogh BE, Kinsman R. *National Adult Cardiac Surgical Database Report. The Society of Cardiothoracic Surgeons of Great Britain and Ireland*. Play hatch, Berkshire, UK: Dendrite Clinical Systems Ltd., 1998.
111. Tu JV, Weinstein MC, McNeil BJ, Naylor CD. *Predicting mortality after coronary artery bypass surgery: what do artificial neural networks learn? The Steering Committee of the Cardiac Care Network of Ontario*. Med Decis Making. 1998; 18:229-35.
112. Orr RK. *Use of a probabilistic neural network to estimate the risk of mortality after cardiac surgery*. Med Decis Making. 1997; 17: 178-85.
113. Lippmann RP, Shahian DM. *Coronary artery bypass risk prediction using neural networks*. Ann Thorac Surg. 1997; 63: 1635-43.
114. Nilsson J, Ohlsson M, Thulin L, Hoglund P, Nashef SAM, Brandt J. *Risk Factor identification and mortality prediction in cardiac surgery using artificial neural networks*. J Thorac Cardiovasc Surg 2006; 132: 12-9.

115. Peng SY, Peng SK. *Predicting adverse outcomes of cardiac surgery with the application of artificial neural networks*. *Anaesthesia* 2008; 63: 705-13.
116. Loop FD, Golding LR, MacMillian JP, Cosgrove DM, Lytle BW, Sheldon WC. *Coronary artery surgery in women compared with men: analyses of risks and long-term results*. *J Am Coll Cardiol*. 1983; 1: 383-390.
117. Tobin JN, Wassertheil-Smoller S, Wexler JP, Steingart RM, Budner N, Lense L, Wachspress J. *Sex bias in considering coronary bypass surgery*. *Ann Intern Med*.1987; 107(1): 19-25.
118. Ayanian JZ, Epstein AM. *Differences in the use of procedures between women and men hospitalized for coronary heart disease*. *N Engl J Med*, 1991; 325:221.
119. Hannan EL, Kilburn H, Raez M. *Improving the outcomes of coronary artery bypass surgery in New York State*.*JAMA* 1994; 271: 761-767.
120. Hirose H, Amano A, Takahashi A, Nagano N. *Coronary artery bypass grafting for patients with non-dialysis-dependent renal dysfunction (serum creatinine > or =2.0 mg/dl)*. *Eur J Cardiothorac Surg*. 2001; 20(3):565-72.
121. Lee KL, Pryor DB, Harrell FE, Califf RM, Behar VS, Floyd WL, Morris JJ, Waugh RA, Whalen RE, Rosati RA. *Predicting outcome in Coronary Disease. Statistical models versus expert clinicians*. *Am J Med* 1986; 80: 553-560.
122. Diamond GA. *Future imperfect: the limitations of clinical prediction models and the limits of clinical prediction*. *J Am Coll Cardiol* 1989; 14:12A-22A.
123. Jin R, Grunkemeier GL. *Does the logistic Euroscore offer an advantage over the additive model?* *Interact Cardio Vasc Thorac Surg* 2006; 5:15-17.

124. Likosky DS, Nugent WC, Clough RA, Weldner PW, Quinton HB, Ross CS, O'Connor GT. *Comparison of three measurements of cardiac surgery mortality for the Northern New England Cardiovascular Disease Study Group*. Ann Thorac Surg. 2006 Apr; 81(4):1393-5.
125. Nashef SA, EuroSCORE Project team. *The New EuroSCORE Project. Nowa skala EuroSCORE*. Kardiol Pol. 2010; 68(1): 128-9.