

CENTRO UNIVERSITÁRIO FEI
NAYARA MOREIRA ROSA

**ANÁLISE COMPARATIVA DE MÉTODOS DE PREVISÃO PARA
CLASSIFICAÇÃO E PREVISÃO DA PROBABILIDADE DE HOSPITALIZAÇÃO
PROLONGADA NO PERÍODO PÓS-OPERATÓRIO EM PACIENTES
SUBMETIDOS À CIRURGIA DE REVASCULARIZAÇÃO DO MIOCÁRDIO**

São Bernardo do Campo

2017

NAYARA MOREIRA ROSA

**ANÁLISE COMPARATIVA DE MÉTODOS DE PREVISÃO PARA
CLASSIFICAÇÃO E PREVISÃO DA PROBABILIDADE DE HOSPITALIZAÇÃO
PROLONGADA NO PERÍODO PÓS-OPERATÓRIO EM PACIENTES
SUBMETIDOS À CIRURGIA DE REVASCULARIZAÇÃO DO MIOCÁRDIO**

Dissertação apresentada ao Centro
Universitário FEI, como parte do requisito
necessário para obtenção do título de Mestre em
Engenharia Mecânica. Orientada pelo Prof. Dr.
João Chang Junior.

São Bernardo do Campo

2017

Moreira Rosa, Nayara .

Análise comparativa de métodos de previsão para classificação e previsão da probabilidade de hospitalização prolongada no período pós-operatório em pacientes submetidos à cirurgia de revascularização do miocárdio / Nayara Moreira Rosa. São Bernardo do Campo, 2017.

102 p. : il.

Dissertação - Centro Universitário FEI.

Orientador: Prof. Dr. João Chang Júnior.

1. Tempo de Permanência Pós-Operatório. 2. Fatores de Risco Cardiovasculares. 3. Regressão Logística. 4. Redes Neurais Artificiais. I. Chang Júnior, João, orient. II. Título.

Aluno: Nayara Moreira Rosa

Matrícula: 215124-9

Título do Trabalho: Análise comparativa de métodos de previsão para classificação e previsão da probabilidade de hospitalização prolongada no período pós-operatório em pacientes submetidos à cirurgia de revascularização do miocárdio.

Área de Concentração: Produção

Orientador: Prof. Dr. João Chang Junior

Data da realização da defesa: 13/02/2017

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

São Bernardo do Campo, 13 / 02 / 2017.

MEMBROS DA BANCA EXAMINADORA

Prof. Dr. João Chang Junior

Ass.: _____

Prof. Dr. Omar Asdrúbal Vilca Mejía

Ass.: _____

Prof.^a Dr.^a Claudia Aparecida de Mattos

Ass.: _____

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO

REPROVADO

VERSÃO FINAL DA DISSERTAÇÃO

**APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE
FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA
EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Rodrigo Magnabosco

A minha estimada mãe Cláudia Aparecida
Moreira, a quem dedico todo o meu esforço,
carinho e perseverança.

AGRADECIMENTOS

Agradeço primeiramente a Deus por fortalecer minhas condições físicas e emocionais nos momentos difíceis.

Agradeço também à minha mãe, por seu exemplo de vida e fonte inesgotável de amor que sempre me inspira ser uma cidadã e profissional cada vez melhor.

À minha criança e irmão que tanto amo, Luís Felipe, a quem espero um dia também ser fonte de inspiração.

Ao meu companheiro de vida, Alex Sandro, pelo companheirismo e paciência mesmo nos momentos difíceis que vivemos.

Ao prof. Chang, pela sua orientação na condução deste trabalho e sua sincera amizade.

Aos professores do Centro Universitário FEI, em São Bernardo do Campo – SP que, direta ou indiretamente, influenciaram no meu desenvolvimento pessoal e intelectual.

Aos profissionais que atuam na Secretaria da Pós-Graduação Stricto Sensu do Centro Universitário FEI, pelo atendimento sempre vigilante às necessidades do aluno.

À CAPES, pelo suporte financeiro que tanto tem me auxiliado durante este período de estudo.

À Prof. Dr^a. Suzana, pela atenção, carinho, auxílio pessoal e profissional.

Ao Dr. Omar, Dr. Alfredo e Enf. Elisandra, pelo suporte fundamental no desenvolvimento da pesquisa e também pela inspiração como profissionais que vão além do simples cumprimento da carga horária. São exemplos de profissionais que estão sempre em busca de garantir o bem-estar da equipe e dos pacientes no ambiente hospitalar.

À toda a equipe de trabalho do Sr. José Antônio, pela colaboração na validação dos dados.

*"A natureza é probabilística
A informação, incompleta
Os resultados, essenciais
Os recursos, limitados
E as decisões, inevitáveis"*

FRAZIER (1980 apud ROCHA; CARVALHO, 2001)

RESUMO

A capacidade de prever o tempo de permanência pós-operatório é uma importante ferramenta de apoio às intervenções médicas e na gestão organizacional da instituição de saúde, visto que uma hospitalização prolongada pode caracterizar o agravamento do estado clínico do paciente e gerar o aumento no consumo de recursos. A classificação dos fatores de risco apurados ainda no período pré-operatório, podem contribuir na capacidade dos métodos de previsão em estimar o risco de o paciente manter-se hospitalizado por um período estendido. Dessa forma, com base nos fatores de risco cardiovasculares este estudo objetiva avaliar a capacidade dos métodos de Regressão Logística Binária e Redes Neurais Artificiais Feedforward Multicamadas em prever a probabilidade de se estender o período de permanência do paciente no hospital, após o ato operatório. O presente estudo se configura como coorte prospectivo de pacientes submetidos à cirurgia de revascularização do miocárdio, coletado na base de dados de um hospital público brasileiro especializado em cardiologia, pneumologia e cirurgias cardíaca e torácica, delimitados por pacientes operados e que receberam alta hospitalar entre o período de novembro/2013 a janeiro/2016. De um total de nove fatores de risco investigados, cinco destes foram considerados estatisticamente significantes na predição, sendo aferido à clearance de creatinina como principal variável de impacto, seguida pelas variáveis tipo de admissão hospitalar, gênero, idade e fração de ejeção, respectivamente. As Redes Neurais Artificiais apresentou uma breve vantagem nos resultados quanto a modelagem com base na Regressão Logística (Curva ROC 67 e 66,5, respectivamente). Contudo, tais valores conferem aos métodos baixa capacidade discriminatória, de modo que o uso dos fatores de risco pré-operatórios selecionados pouco explicam a variabilidade do tempo de permanência pós-operatório. Há, portanto, a necessidade de ampliar o campo de pesquisa quanto aos fatores de risco discriminantes na estimação da probabilidade de hospitalização prolongada pós-operatória.

Palavras-chave: Tempo de Permanência Pós-Operatório. Fatores de Risco Cardiovasculares. Regressão Logística. Redes Neurais Artificiais.

ABSTRACT

The ability to predict the time of postoperative stay is an important tool to support the medical interventions and the organizational management of the health institution, since a prolonged hospitalization can characterize the worsening of the clinical condition of the patient and increase the consumption of resources. The risk factors classification established at the preoperative time may contribute in prediction methods to estimate the risk of the patient staying hospitalized for an extended period. Thus, based on cardiovascular risk factors, this study aims to evaluate the Binary Logistic Regression and Feedforward Multilayer Artificial Neural Networks methods to predict the probability of extending the period of patient stay in the hospital, after the surgery. For this, the present study is a prospective cohort of patients submitted to coronary artery bypass grafting, collected in the database of a Brazilian public hospital specialized in cardiology, pulmonology, cardiac and thoracic surgeries. The sample set is delimited by adult patients operated and discharged from November / 2013 to January / 2016. Among the nine risk factors analyzed as predictors of the prolonged hospitalization period, five of these were classified as predictors of the prolonged hospitalization period, measuring creatinine clearance as the main impact variable, followed by variables hospital admission, gender, age and ejection fraction, respectively. The Artificial Neural Networks presented a compact greater advantage than modeling based on Logistic Regression (Curve ROC 67 and 66.5, respectively). However, these results giving its low discriminatory capacity. So that, to use just these selected preoperative risk factors are not good to explain the variability of the postoperative permanence time. There is, therefore, the need to expand the field of research regarding the discriminant risk factors to estimate the probability of prolonged postoperative hospitalization.

Keywords: Postoperative length of stay. Preoperative Risk Factors. Logistic Regression. Artificial Neural Networks.

LISTA DE ILUSTRAÇÕES

Figura 1 – Publicações anuais: tempo de permanência hospitalar	15
Figura 2 – Estudos comparativos dos métodos de predição selecionados neste estudo ²	16
Quadro 1 – Fatores de risco preditivos do tempo de permanência pós-operatória.....	22
Figura 3 – Classificação dos modelos de previsão	33
Figura 4 – Curva da regressão logística	41
Figura 5 – Arquitetura básica de uma Rede Neural Artificial.....	45
Figura 6 – Rede feedforward de uma única camada.....	46
Figura 7 – Rede feedforward com múltiplas camadas.....	47
Figura 8 – Rede recorrente	49
Figura 9 – Relação das funções de ativação para redes neurais artificiais	52
Figura 10 – Procedimentos metodológico.....	53
Figura 11 – Fluxograma do processo de seleção dos artigos pesquisados.....	55
Figura 12 – Critérios de exclusão de dados.....	63
Figura 13 – Tempo de permanência pós-operatório estratificado	64
Figura 14 – Distribuição de frequência do TP pós-operatório	64
Figura 15 – Intervalo de confiança Bootstrapping	73
Figura 16 – Percentual de acerto da predição para cada variável inserida ao modelo	74
Figura 17 – Arquitetura da Rede Neural Artificial.....	76
Figura 18 – Curva ROC para os modelos de previsão.....	77
Figura 19 – Prevalência do tempo de permanência pós-operatório por fator de risc	100

LISTA DE TABELAS

Tabela 1 – Fatores de riscos pré-operatórios - variáveis contínuas	65
Tabela 2 – Variáveis categorizadas por grupo de risco	65
Tabela 3 – Fatores de riscos pré-operatórios - variáveis nominais e ordinais.....	66
Tabela 4 – Tabela de contingência variáveis explicativas versus TP pós-operatório.....	67
Tabela 5 – Correlação das variáveis contínuas ao TP pós-operatório nominal.....	68
Tabela 6 – Variáveis não presentes na equação	70
Tabela 7 – Interpretação preliminar dos coeficientes para o modelo final.	70
Tabela 8 – Diagnóstico -2 log likelihood value.....	71
Tabela 9 – Diagnóstico R ² de Nagelkerke.....	72
Tabela 10 – Diagnóstico teste Wald	73
Tabela 11 – Resultados da previsão – Regressão Logística.....	74
Tabela 12 – Resultados da previsão – Redes Neurais Artificiais	77
Tabela 13 – Área abaixo da curva ROC	77

LISTA DE ABREVIACOES

ANN	<i>Artificial Neural Networks</i>
CABG:	<i>Coronaty Artery Bypass Graft</i>
ClCr:	Clearance de creatinina
CRM:	Cirurgia de revascularizao do miocrdio
DATASUS:	Banco de dados do Sistema nico de Sade
dF	Graus de liberdade
DM:	Diabetes mellitus
DRC:	Doena renal crnica
ed.:	Editora
EuroSCORE	<i>European System for Cardiac Operative Risk Evaluation</i>
FE:	Frao de Ejeo
FEVD:	Frao de Ejeo do Ventrculo Esquerdo
HAP:	Hipertenso arterial pulmonar
HbA1C:	Hemoglobina glicosilada ou Hemoglicada
ID:	Idade
IMC:	ndice de Massa Corporal
Inc.:	Incorporado
LI	Limite inferior
LS:	Limite superior
-2LL:	<i>Log-likelihood Value</i>
NYHA:	<i>New York Heart Association</i>
OR:	<i>Odds ratio</i>
p.:	Pgina
PAP:	Presso Arterial Pulmonar
PMAP:	Presso mdia da artria pulmonar
PSAP:	Presso Sistlica da Artria Pulmonar
REOP:	Reoperao
RL:	Regresso Logstica
RNA:	Redes Neurais Artificiais
ROC	Receiver operating characteristic
SPSS	<i>Statistical Package for the Social Sciences</i>
TA:	Tipo de admisso (urgncia, emergncia e eletiva)
TP:	Tempo de permanncia
TPN:	Tempo de Permanncia Normal (Padro)
TPP:	Tempo de Permanncia Prolongado
UTI:	Unidade de Terapia Intensiva
v.:	Volume

SUMÁRIO

1	INTRODUÇÃO	13
1.1	JUSTIFICATIVA.....	15
1.2	DELIMITAÇÃO DA PESQUISA	17
1.3	ESTRUTURA DA DISSERTAÇÃO	17
2	OBJETIVOS	19
2.1	OBJETIVO GERAL.....	19
2.2	OBJETIVOS ESPECÍFICOS.....	19
3	REVISÃO TEÓRICA	20
3.1	FATORES DE RISCOS PARA DOENÇAS CARDIOVASCULARES	20
3.1.1	Idade	23
3.1.2	Gênero	24
3.1.3	Índice de Massa Corporal	25
3.1.4	Fração de ejeção	26
3.1.5	Disfunção renal	27
3.1.6	Diabetes mellitus	28
3.1.7	Hipertensão arterial pulmonar	29
3.1.8	Reoperações	30
3.1.9	Admissões eletivas, urgência ou emergência	31
3.2	MÉTODOS DE PREVISÃO PARA O TEMPO DE PERMANÊNCIA HOSPITALAR	31
3.2.1	Métodos aritméticos	33
3.2.2	Métodos de regressão	34
3.2.2.1	<i>Análise de sobrevivência</i>	34
3.2.2.2	<i>Métodos tradicionais de regressão</i>	35
3.2.3	Métodos orientados a dados	36
3.2.3.1	<i>Modelos de agrupamento</i>	37
3.2.3.2	<i>Métodos com técnicas de aprendizagem</i>	37
3.2.4	Modelos de multi-estágio	38
3.3	REGRESSÃO LOGÍSTICA	39
3.3.1	Métodos de diagnóstico	41
3.3.1.1	<i>Teste Razão de Verossimilhança</i>	41
3.3.1.2	<i>R² de Nagelkerke</i>	42

3.3.1.3	<i>Teste Bootstrapping</i>	43
3.3.1.4	<i>Teste Wald</i>	44
3.4	REDES NEURAIAS ARTIFICIAIS.....	44
3.4.1	Tipos de arquitetura para redes neurais artificiais.....	46
3.4.1.1	<i>Redes feedforward com uma única camada</i>	46
3.4.1.2	<i>Redes feedforward de múltiplas camadas</i>	47
3.4.1.3	<i>Redes recorrentes simples e de multicamadas</i>	48
3.4.2	Definição do número de camadas ocultas e neurônios.....	49
3.4.3	Processos de aprendizagem.....	50
3.4.4	Funções de ativação.....	51
4	MÉTODOS.....	53
4.1	BASE DE DADOS: CRITÉRIOS DE SELEÇÃO DOS ARTIGOS.....	53
4.2	AMOSTRA: TRATAMENTO DOS DADOS.....	55
4.2.1	Critério de exclusão de dados	57
4.3	PROPOSIÇÃO DOS MODELOS DE PREVISÃO.....	57
4.3.1	Caracterização da variável dependente	58
4.3.2	Construção do modelo de regressão logístico	58
4.3.2.1	<i>Ajuste e validação do modelo</i>	60
4.3.3	Construção do modelo de redes neurais artificiais	60
4.3.4	Análise de resíduos – Curva ROC	61
5	RESULTADOS.....	62
5.1	ANÁLISE DESCRITIVA DOS DADOS.....	62
5.2	PERFIL BIVARIADO: EXAME DA RELAÇÃO ENTRE VARIÁVEIS.....	66
5.3	ANÁLISE DA REGRESÃO LOGÍSTICA COMO MÉTODO DE MODELAGEM.....	69
5.3.1	Seleção das variáveis preditoras	69
5.3.2	Diagnósticos do modelo de regressão logística	71
5.4	MODELAGEM DAS REDES NEURAIAS ARTIFICIAIS.....	75
5.5	AVALIAÇÃO DO PODER DE DISCRIMINAÇÃO DOS MODELOS.....	77
6	DISCUSSÃO.....	79
7	CONSIDERAÇÕES FINAIS.....	83
7.1	LIMITAÇÕES DO ESTUDO.....	84
7.2	SUGESTÕES PARA TRABALOS FUTUROS.....	85
	REFERÊNCIAS	86
	APÊNDICE A – Validação cruzada das variáveis investigadas	99

1 INTRODUÇÃO

Como resposta à redução de custos e ganho de eficiência operacional, as instituições de saúde são frequentemente desafiadas pelo conflito entre minimizar o consumo de recursos sem reduzir a qualidade na assistência hospitalar. Para isso, métricas de avaliação do desempenho se configuram como ferramentas de apoio à tomada de decisão. Previsões que avaliam o risco de permanência prolongada vêm ganhando espaço, nesse sentido, como uma medida legitimada no âmbito das intervenções médicas e gerenciamento organizacional.

O tempo de permanência (TP) é então caracterizado pelo apoio às tomadas de decisões dos profissionais da saúde, ao mesmo tempo em que sinaliza o nível de desempenho organizacional para os gestores da instituição. Na primeira situação, o médico, por exemplo, pode estabelecer um tratamento paralelo ao principal que garanta menor risco de complicações pós-operatório, até então, traduzidas em um período estendido de hospitalização. No segundo aspecto, essa métrica se comporta como um estimador do consumo de recursos e custeio das despesas operacionais.

Na tangente à produção científica, o TP é concretizado em estudos direcionados a previsão da alta hospitalar (KHAIRUDIN; MOHD; HAMID, 2012; ALMASHRAFI et al., 2016b; KRELL; GIROTTI; DIMICK, 2014; AWAD; BADER-EL-DEN; MCNICHOLAS, 2016; CASTILLO, 2012; KHAIRUDIN, 2011) no estudo do tempo e movimento (ORTIGA et al., 2012), na gestão de custos (OSNABRUGGE et al., 2014), no monitoramento do consumo de recursos (SCHMIDT et al., 2013) e na satisfação do paciente e familiares aos serviços prestados (LeMOND; ALLEN, 2011).

Em termos gerais, o TP hospitalar corresponde ao período em que o paciente permanece hospitalizado para tratamento e recuperação da enfermidade, sem a intercorrência de readmissões no período (BORGHANS et al., 2012). Os autores ainda destacam que o conceito pode ser empregado nas diferentes fases do cuidado assistencial, cujos esforços neste estudo estão voltados ao período de recuperação do paciente após o ato cirúrgico, justificado pela alta taxa de óbito, readmissões e complicações incorridas no período pós-operatório.

O hospital estabelece um período mínimo padrão de permanência hospitalar, necessárias para a recuperação de pacientes em condições normais. Enquanto que, longos períodos de hospitalização podem aumentar o risco de morte do paciente e prejudicar o desempenho operacional da organização (KHAIRUDIN; MOHD; HAMID, 2012). Tal fato, expõe a necessidade do uso de técnicas de previsão como ferramenta de apoio ao plano de intervenção médico e na gestão dos recursos organizacionais que possam impactar no real período de

recuperação do paciente.

Uma das áreas médicas em que estudos do TP pós-operatório vem se desenvolvendo nos últimos anos é a insuficiência coronariana com necessidade de intervenção cirúrgica, particularmente na cirurgia de revascularização do miocárdio (CRM), tendo em vista o alto índice de cirurgias anuais e os custos incorridos com o procedimento (ALMASHRAFI et al., 2016a; CASTILLO, 2012).

Almashrafi et al. (2016a), Shinjo e Fuchigami (2015) e Oliveria et al. (2013) ainda destacam o uso dos fatores de riscos pré-operatórios cardiovasculares, comumente preconizados nos escores de risco de morte, como variáveis preditoras ao TP pós-operatório. Considerados determinantes ao quadro clínico do paciente, esses fatores podem apresentar diferentes graus de impacto na estimativa do TP e com diferentes classificações para cada população investigada. Tais fatos exigem avaliações estatísticas de precisão que contribuam na investigação do real impacto desses fatores ao prolongamento do período de hospitalização.

Dentre os métodos estatísticos de previsão que lidam com a relação dos fatores de risco ao TP pós-operatório, destacam-se publicações que apresentam Regressão Logística (RL) e Redes Neurais Artificiais (RNAs) como técnica proposta, dada a validação e disseminação dessas técnicas nos diversos campos do conhecimento. No uso da RL sobressaem-se os estudos de Azarfarin et al. (2014), Khairudin; Mohd; Hamid (2012), Pagowska-Klimek et al (2011), cujos resultados mostraram-se satisfatórios na estimativa do tempo de permanência prolongado. Enquanto que, Tsai et al. (2016) e Houthoof et al. (2015) também garantiram resultados satisfatório na aplicação das RNAs.

Na literatura brasileira poucos são os registros de pesquisas que abordam o TP pós-operatório. Em sua maioria, se limitam apenas a identificar os fatores de riscos associados ao período de hospitalização, tal como abordado por Oliveira et al. (2013), Silva et al. (2013) e Machado (2011). Carece, portanto no contexto brasileiro, estudos voltados à exploração dos métodos estatísticos de previsão no âmbito do tempo de permanência hospitalar.

Sendo assim, com base na relação dos fatores de riscos cardiovasculares ao tempo de permanência pós-operatória é proposto neste estudo, o uso destes métodos de previsão no contexto brasileiro. Onde, busca-se avaliar a capacidade dos respectivos métodos em prever a probabilidade de estender o período de permanência hospitalar após a intervenção cirúrgica. Para tanto, é definido a seguinte problemática de pesquisa:

Com base nos fatores de riscos, os métodos de previsão “Regressão Logística” e “Redes Neurais Artificiais” apresentam bom desempenho na previsão de probabilidade do tempo de permanência pós-operatório?

Cabe, portanto, nas próximas seções oferecer o embasamento teórico e aplicado necessários para o delineamento da problemática apresentada.

1.1 JUSTIFICATIVA

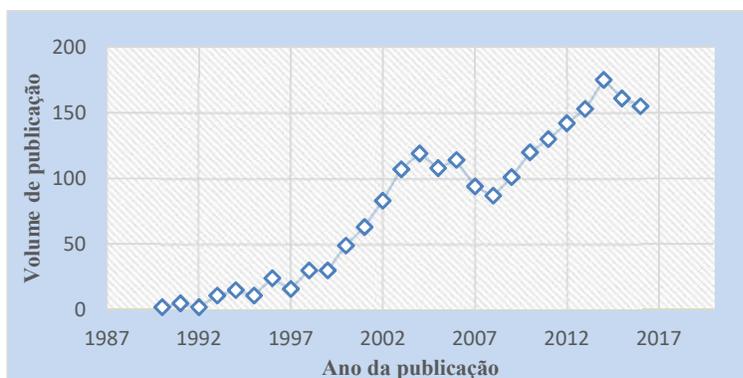
No Brasil, quase a totalidade da demanda de pacientes com indicação cirúrgica é coberto pelo Sistema Único de Saúde (SUS). Somente na cirurgia de revascularização do miocárdio (CRM) o sistema é responsável por 80% das intervenções praticadas em todo país, sejam estas realizadas por hospitais públicos, filantrópicos ou privados (PIEGAS; BITTAR; HADDAD, 2010).

São quase 290 milhões de reais repassados para hospitais conveniados ao SUS, apenas no ano de 2015. (MINISTÉRIO DA SAÚDE, 2016). Para tanto, tal cômputo reduz os repasses financeiros e, conseqüentemente, a capacidade operacional de atendimento da organização.

Nesse sentido, como suporte aos tomadores de decisões, a investigação do tempo de permanência hospitalar atua na proposição de melhores planos de intervenção médica frente a limitação de recursos; na maioria dos casos estimulado pelo uso de técnicas estatísticas de previsão (ALMASHRAFI et al., 2016a).

Em datações anteriores, por sua vez, as publicações eram dedicadas, em sua maioria, apenas na determinação dos fatores de risco associados ao TP pós-operatório. No âmbito da pesquisa científica, o TP apresenta avanços significativos na última década, justificado pela disseminação das técnicas de previsão no campo da ciência médica. (Figura 1).

Figura 1 – Publicações anuais: tempo de permanência hospitalar



Fonte: Autora

Sob a alusão desta temática, a disseminação é ainda garantida com o uso dos métodos de Regressão Logística e Redes Neurais Artificiais como métodos de previsão, dada a

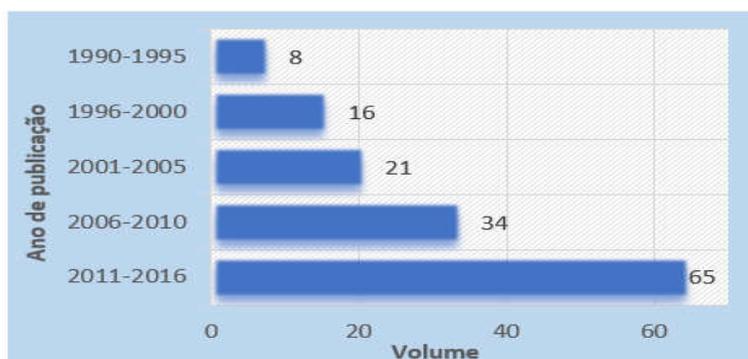
habilidade destas ferramentas em relacionar variáveis de impacto ao desfecho de TP prolongado.

Com propriedades dicotômicas a regressão logística emergiu como técnica estatística de previsão de alto rendimento, nas mais diversas áreas do conhecimento (OSBORNE, 2013). Em contrapartida, as RNA emergiram como uma alternativa à RL, justificado pelo bom desempenho em dados com comportamentos não lineares e a versatilidade do modelo em lidar tanto com predições baseado em resultados dicotômicos como também em métodos de classificação e previsão séries temporais (HAYKIN, 2003).

No Brasil, as pesquisas que abordam esta temática não seguem esta tendência de crescimento. Em sua maioria, os estudos se concentram na determinação dos fatores de risco que impactam no TP prolongado, carecendo, portanto, de abordagens que interagem esses fatores aos métodos de previsão.

Dado os resultados promissores desses modelos nos diversos campos da pesquisa e prática profissional. Vários estudos são direcionados na comparação da capacidade preditiva desses modelos no âmbito clínico, epidemiológico e no campo da engenharia (figura 2).

Figura 2 – Estudos comparativos dos métodos de predição selecionados neste estudo²



Fonte: Autora

²: Resultados obtidos na Base de dados Science Direct.

Portanto, dada a contemporaneidade de estudos relacionados ao tempo de permanência hospitalar pós-operatório, o impacto da cirurgia de revascularização do miocárdio na sociedade e nos gastos aos hospitais e governantes públicos, bem como a abrangência de estudos direcionados a aplicação de métodos preditivos no contexto do tempo de permanência é, então, proposto realizar uma análise comparativa de métodos de previsão para classificação e previsão da probabilidade de hospitalização prolongada no período pós-operatório, em pacientes submetidos à cirurgia de revascularização do miocárdio.

1.2 DELIMITAÇÃO DA PESQUISA

Essencialmente este trabalho apresenta certas delimitações que reflete o posicionamento da pesquisadora no desenvolvimento da pesquisa. Entre elas, destacam-se:

- a) abordagem do tempo de permanência hospitalar direcionada ao período pós-operatório, sendo, portanto, desconsiderados os períodos pré- e intra-operatório na análise preditiva;
- b) investigação do tempo de permanência no âmbito geral do período pós-operatório, ou seja, não foi definido para este estudo a investigação do tempo de permanência na UTI e na enfermaria, separadamente.
- c) Abordagem exclusiva à cirurgia de revascularização do miocárdio, não abrangendo, portanto, as demais cirurgias cardíacas;
- d) definição dos fatores de riscos cardiovasculares pré-operatório na relação ao tempo de permanência pós-cirúrgico. Foram desconsiderados fatores de riscos externos, bem como os fatores internos que estejam fora do contexto de análise pré-operatória;
- e) para mérito de comparação, considerou-se apenas os métodos de Regressão Logística e Redes Neurais Artificiais, garantindo para pesquisas futuras a exploração da capacidade preditiva dos demais métodos;
- f) por restrição de recursos materiais, ambas as modelagens limitou-se ao uso do *software* SPSS 23.0; e,
- g) coleta de dados unicêntrico, ou seja, em uma única instituição de saúde, fato que dificulta generalização e validação do modelo em outras instituições.

1.3 ESTRUTURA DA DISSERTAÇÃO

Este trabalho foi desenvolvido ao longo de 7 capítulos. Neste primeiro é introduzido a temática da pesquisa, os motivos que levaram a escolha da problemática do estudo, bem como as delimitações que norteiam a sua estrutura.

No segundo capítulo é dado destaque ao objetivo geral, bem como os objetivos específicos que deverão orientar o desenvolvimento da pesquisa, no atendimento à problemática abordada.

No capítulo três é apresentada a revisão teórica ao qual esta pesquisa foi embasada, iniciando-se com a investigação aos fatores de risco cardiovasculares pré-operatórios, em seguida é realizado um levantamento bibliográfico dos métodos de previsão aplicados ao tempo

de permanência pós-operatório e, para meios de comparação, os métodos de previsão Regressão Logística e Redes Neurais Artificiais são caracterizados quanto a sua estrutura analítica e os métodos estatísticos envolvidos na etapa de diagnóstico da modelagem.

No capítulo quatro é apresentada a metodologia de pesquisa, onde é sistematizado os procedimentos de pesquisa e análise de dados que fomentam o estudo.

No capítulo cinco são apresentados os resultados da pesquisa que vão desde a análise exploratória dos dados até a validação final dos modelos de previsão.

No capítulo seis são discutidos os resultados alcançados de acordo com pesquisas anteriormente realizadas.

No capítulo sete são apresentadas as considerações finais deste estudo, bem como as limitações e o desenvolvimento de propostas para continuidade da discussão do tema em pesquisas futuras.

Ao final, é dado destaque às referências bibliográficas que embasaram o estudo.

2 OBJETIVOS

Em consideração a importância de prever o risco de permanência prolongada no período pós-operatório, bem como os problemas enfrentados pelos tomadores de decisões em todos os níveis do setor hospitalar, há uma necessidade do maior envolvimento de ferramentas estatísticas que visam proporcionar resultados mais precisos e, conseqüentemente, contribuir no melhor entendimento da duração da estadia e os fatores de risco discriminantes.

Nesse sentido, este estudo é voltado para a avaliação dos métodos estatísticos Regressão Logística e Redes Neurais Artificiais como ferramentas de previsão da probabilidade do tempo de permanência hospitalar pós-operatório. Segue, então, a apresentação do objetivo geral e objetivos específicos que fundamentaram esta pesquisa.

2.1 OBJETIVO GERAL

Baseado nos fatores de riscos cardiovasculares para pacientes submetidos à cirurgia de revascularização do miocárdio, este estudo tem como objetivo geral avaliar a capacidade dos métodos de Regressão Logística e de Redes Neurais Artificiais na previsão de probabilidade do tempo de permanência prolongado pós-operatório.

2.2 OBJETIVOS ESPECÍFICOS

Como meio de fornecer o subsídio necessário para o alcance do objetivo geral e conseqüentemente o esclarecimento e desenvolvimento da pesquisa, foram definidos os seguintes objetivos específicos:

- a) identificar os fatores de risco cardiovasculares que se correlacionam ao tempo de permanência pós-operatório;
- b) aplicar o modelo de “Regressão Logística” para estimar o tempo de permanência pós-operatório;
- c) aplicar o modelo de “Redes Neurais Artificiais” para estimar o tempo de permanência pós-operatório;
- d) validar os modelos de previsão para o tempo de permanência pós-operatória;
- e) comparar o desempenho dos resultados obtidos por meio de cada modelo.

3 REVISÃO TEÓRICA

A revisão teórica é estrategicamente segmentada em três diferentes abordagens complementares a temática proposta. Na primeira parte (seção 3.1), é dado um embasamento teórico dos principais fatores de riscos pré-operatórios associados a doenças cardiovasculares, com ênfase nos casos necessários para a cirurgia de revascularização do miocárdio. Na segunda parte (seção 3.2), é realizado uma síntese dos principais modelos de previsão aplicados no contexto do tempo de permanência hospitalar, em geral. Na terceira parte (seção 3.3 e 3.4), portanto, é aprofundado o conhecimento nos métodos de previsão selecionados para a condução desta pesquisa: regressão logística e redes neurais artificiais; cujo foco está na melhor compreensão dos conceitos envolvidos por trás dos cálculos.

3.1 FATORES DE RISCOS PARA DOENÇAS CARDIOVASCULARES

Normalmente, as doenças cardiovasculares são produtos de um conjunto de fatores de riscos causais, sejam estes relacionados ao condicionamento físico e biológico, condições hereditárias ou desenvolvidas a partir de outras doenças. Vários fatores de risco, aparentemente modestos podem, em conjunto, resultar em um risco muito maior do que um único fator impressionantemente elevado. Tal fato pode apresentar impacto imediato no período necessário para o tratamento e recuperação do paciente (COONEY; DUDINA; GRAHAM, 2009).

Nos contornos da doença cardíaca, os primeiros registros de pesquisas que investigam a relação dos fatores de riscos ao tempo de permanência pós-operatório, pertencem a Silberbach et al. (1993) e Kalman et al. (1994). No primeiro caso, dados de 322 operações consecutivas de cardiopatas congênitos adultos foram coletados durante o período de 1985 e 1989. Na condição pré-operatória, o diagnóstico de anomalias como malformação extra cardíaca, falha de crescimento e o grau de oxigenação no sangue obteve impacto superior quando comparados aos fatores de riscos demográficos do paciente. No estudo seguinte, Kalman e os demais pesquisadores analisaram os fatores de riscos pré-operatórios associados a arteriopatia distal, onde foi observado a predominância da idade, histórico de doença cardiovascular e a indicação operatória como significantes ao prolongamento do TP pós-operatório de 275 pacientes analisados.

O tema ainda se mantém contemporâneo com a valorização crescente na gestão sistêmica hospitalar, de modo a reduzir custos, aumentar a capacidades de ocupação de leitos, introdução de novas tecnologias, desenvolvimento de novos métodos de tratamentos e

procedimentos cirúrgicos (WAGNER et al., 2013).

Para o TP pós-operatório os fatores de riscos são definidos em relação a condição orgânica e fisiológica do paciente no período pré-operatório, aos procedimentos clínicos e cirúrgicos, ao tipo de admissão cirúrgica (eletiva, urgência ou emergência), a ocorrência de reoperações e a precedência de doenças cardiopulmonares (LaPAR et al., 2013).

Estudos recentes fornecem uma vasta revisão dirigida a análise dos pacientes internados na Unidade de Tratamento Intensivo (UTI), cujo propósito central está em investigar os fatores de risco associados ao prolongamento desse período de estadia. Pagowska-Klimek et al. (2011) e Shinjo e Fushimi (2015) retratam como evidências em comum o gênero, idade, índice de massa corporal, reoperações, arritmia, fração de ejeção, hipertensão, diabetes mellitus, insuficiência renal e cirurgias não-eletivas como fatores de risco significativos à estimação do TP.

Em uma pesquisa de meta-análise conduzido por Almashrafi et al. (2016a) foram analisados 195 artigos que, de modo geral, reforçam a análise dos pesquisadores anteriores quanto aos fatores de riscos pré-operatórios. Neste estudo, os esforços de melhorar processos após cirurgia cardíaca também incidi nas UTIs. Os autores justificam que o critério é devido a percepção da UTI como principal ‘gargalo’ da prestação de serviços hospitalares, devido ao alto consumo de recursos durante a estadia do paciente e o aumento do risco de mortalidade nesse período de permanência.

Em contrapartida, alguns pesquisadores argumentam a necessidade de avaliar o processo pós-operatório em sua totalidade. Para Najafi e Goodarzynejad (2012) e Oliveira et al. (2013) o departamento de enfermaria também deve ser avaliado após o período cirúrgico, visto que uma estadia prolongada na enfermaria pode provocar lotação dos leitos e este se propagar nos limites de leitos da UTI.

No estudo proposto por Silva et al. (2013), além dos fatores de riscos pré-operatório comumente associados à cirurgia de revascularização do miocárdio, os autores ainda avaliam essas variáveis preditoras a partir de uma fonte financiadora da intervenção cirúrgica (Sistema Único de Saúde (SUS), planos de saúde ou recursos próprios). Como resultado, os pacientes agrupados em Não SUS tiveram TP superior aos pacientes do grupo SUS. No entanto, os autores sugerem a realização de novos estudos para identificar o motivo destas diferenças ainda não investigadas.

De modo geral, é apresentado no quadro 1 os principais estudos que abordam a identificação dos fatores de risco de impacto no TP pós-operatório. É também dado destaque a região onde foi realizada a pesquisa e o número de instituições que caracterizam o estudo como

unicêntrico ou multicêntricos.

Além disso, verifica a predominância dos fatores de riscos no tempo de permanência em UTIs como estratégias de análise e validação para a maioria das pesquisas analisadas. A idade e o gênero se mantiveram presentes nos consecutivos estudos, seguida pela insuficiência cardíaca, disfunção renal, hipertensão pulmonar e diabetes.

Quadro 1 – Fatores de risco preditivos do tempo de permanência pós-operatória

Referência	País	Número de Instituições	Número de Pacientes	Especialização cirúrgica	Abrangência do TP pós-operatório	Fatores de risco pós-operatório
Shinjo e Fushimi (2015)	Japão	268	2,491	revascularização do miocárdio sem circulação extracorpórea	pós-operatório completo	GR, IDADE, CA, TCA, FE, HP, IC, IMC, ANG, COAG, REOP, FU/NFU.
Almashrafi <i>et al</i> (2016)	Omã	1	600	cirurgia cardíaca (em geral)	UTI e pós-operatório completo	IDADE, TCA, IC, DR, HP, CA, ASI, IDADE, GR, IMC, DIAB, IMC, ANGIN, ACFA.
Da Silva <i>et al</i> (2013)	Brasil	1	2,840	cirurgia de revascularização do miocárdio	pós-operatório completo	IDADE, GR, SUS/NSUS, CA, ADM, FE, HP, IC.
Pagowska-Klimek <i>et al</i> (2011)	Polônia	1	693	cirurgia de cardiopatia congênita	UTI	IDADE, GR, IR, CI, REOP, DR, HP, NYHA (III-IV), SIE.
Najafi e Goodarzynjad (2012)	Iran	1	570	cirurgias coronarianas	UTI e enfermaria, separadamente	<u>Enfermaria:</u> IDADE, GR, IC, DB; <u>UTI:</u> IDADE, GR, BIA, ACFA, REOP.
Oliveira <i>et al</i> (2013)	Brasil	1	104	cirurgia de revascularização do miocárdio	UTI e enfermaria, separadamente	<u>Enfermaria:</u> FE, DB; <u>UTI:</u> TAB, DB.
Peterson <i>et al</i> (2002)	Estados Unidos	587	496,797	cirurgia de revascularização do miocárdio	pós-operatório completo	IDADE, BSA, EF, TC, DR, DPOC, DB.
De Cocker <i>et al</i> (2011)	Bélgica	1	395	cirurgia cardíaca (em geral)	UTI	IDADE, GR, NYHA (III-IV), DR, IC, ACFA, RM, SIE, BIA, TC, CA
Cacciatore <i>et al</i> (2012)	Itália	1	250	cirurgia cardíaca (em geral)	UTI	IDADE, GR, NYHA (III-IV), DR, AVC.
Azarfarin <i>et al</i> (2014)	Iran	1	280	cirurgia cardíaca (em geral)	UTI	IDADE, NYHA, FE, SIE, IC, REOP, HP, DB, TC.
Eltheni <i>et al</i> (2012)	Grécia	1	150	cirurgia cardíaca (em geral)	UTI	IDADE, GR, CR, ACFA, DR, DB.
Wang <i>et al</i> (2012)	China	1	4003	cirurgia valvar	UTI	IDADE, GR, IR, FE, ACFA, NYHA (III-IV), REOP, IC, CA

Fonte: Autora

Legenda: ASI: aneurisma do septo inter-atrial; SUS / NSUS: fonte financiadora dos cuidados hospitalares; IC: insuficiência cardíaca; DR: disfunção renal; ANG: angina; GR: gênero; CA: cirurgias associadas (cirurgia de revascularização do miocárdio + valvar ou outras relacionadas); FE: fração de ejeção; DB: diabetes; HP: hipertensão pulmonar; TC: tipo de cirurgia por admissão (eletiva, emergência e urgência); ACFA: arritmia cardíaca por fibrilação auricular; HEP: doença hepática; COAG: coagulopatias; REOP: reoperação; TAB: tabagismo; IR: insuficiência respiratória; HbA1c: hemoglobina glicosilada; BIA: balão intra-aórtico; NYHA (III-IV)¹: classe funcional da Associação do Coração de Nova York; DPOC: doença pulmonar obstrutiva crônica; BSA: área de superfície corporal; RM: regurgitação ou insuficiência mitral; SIE: suporte inotrópico endovenoso; AVC: acidente vascular cerebral; RC: ressuscitação cardiopulmonar; CR: creatinina.

Em conformidade aos estudos de Mihajlović *et al.* (2011), Pagowska-Klimek *et al.*

(2011) e Shinjo e Fushimi (2015), observa-se no quadro 1 a predominância dos fatores de riscos anteriormente mencionados, principalmente nas cirurgias de revascularização do miocárdio.

Modelos estratificados dos fatores de riscos, denominados de escores de risco, também são amplamente propostos na literatura com a finalidade de estimar o tempo de permanência pós-operatório (SULLIVAN; WALLACH; IOANNIDIS, 2016; NASHEF et al., 1999, 2012; BERNSTEIN; PARSONNET, 2000; D'AGOSTINO et al., 2001).

Para o tempo de permanência hospitalar, Toumpoulis et al. (2005) e Mihajlović et al. (2011), destacam que nos últimos anos estudos vem se intensificando no uso de modelos tradicionais – escores de risco de morte – para estimar esse período de hospitalização. Entre os escores de riscos adaptados é então proposto o uso do modelo European System for Cardiac Operative Risk Evaluation – EuroSCORE (NILSSON et al., 2004; TOUMPOULIS et al., 2005; MIHAJLOVIĆ et al., 2011; NIKOLIC, 2015), Parsonnet (LAWRENCE et al., 2000) e Ambler Score (PENA et al., 2010) e Cleveland Score (KURKI et al., 2001).

Contudo, assim como observado na prática tradicional destes modelos, a transição do banco de dados em um sistema estratificado de risco sacrificada a precisão em busca da simplicidade (NASHEF et al., 1999; COONEY; DUDINA; GRAHAM, 2009). Dessa forma, o uso do modelo estratificado não será investigado neste estudo, sendo apenas sugerido a investigação em estudos posteriores, com base na possibilidade de atender ao critério de parcimônia e ao requisito precisão.

Sendo assim, para maior compreensão do impacto dos fatores de risco, é apresentado na próxima subseção uma síntese dos principais fatores associados à CRM e alguns dos trabalhos desenvolvidos nesse campo de estudo.

Contudo, não se deve limitar o tema apenas a estes fatores de riscos, visto que a seleção e o impacto destas variáveis ao TP variam em diferentes populações analisadas. A identificação dos fatores de risco associados não deve ser tratada de forma isolada e sim em conjunto aos dados coletados, de modo a validar os fatores estatisticamente significantes ao modelo proposto.

3.1.1 Idade

Com o avanço de técnicas e tratamentos médicos que permitem o aumento da expectativa de vida da população, a idade é alvo frequente de estudos. Oliveira et al., (2013) e Carvalho Júnior (2015) enfatizam o aumento da expectativa de vida da população brasileira em mais de 10 anos na última década, passando de 63 para 73 anos.

Almashrafi et al. (2016a) argumentam que a idade é o fator mais comumente relatado

como estatisticamente significativa na predição do TP pós-operatório. A idade avançada está associada a maior morbimortalidade, devido a doenças crônicas pré-existentes que incidem em complicações pré, intra e pós-operatória. Normalmente, esses pacientes prolongam em até duas vezes mais o tempo de internação em UTIs do que a pacientes com idade inferior a 60 anos.

Shinjo e Fushimi (2015) ainda destacam a relação da idade ao prolongamento do tempo de permanência pós-operatório e como esta relação impacta no aumento dos custos hospitalares. Aos pacientes com idade entre 65-74 anos foi observado que os custos eram 3,0 % maior do que aqueles com idade inferior a 65 anos e até 5% maior em pacientes com idade superior a 74 anos.

Nesse contexto ainda é analisado o fator fragilidade conexo ao avanço da idade. Em tal aspecto, o processo de envelhecimento aumenta o número de mudanças fisiológicas e de doenças crônicas que declinam a reserva fisiológica e a capacidade funcional do paciente.

Contudo, há controvérsias quanto a real relação da fragilidade e o processo de envelhecimento. Lacasia-Melero (2016) chama a atenção a essa contestação e também na falta de consenso quanto ao significado do termo ‘frágil’ para a definição do estado clínico do paciente. A autora argumenta que fragilidade deve apresentar uma abordagem holística, de modo que o paciente deve ser avaliado quanto a sua capacidade funcional e hábitos diários ao invés de avaliar apenas o aspecto de envelhecimento. Ela ainda destaca que estudos clínicos que abordam a fragilidade ao invés da idade podem oferecer melhor desempenho dos resultados. Mas que, por outro lado, a mensuração da fragilidade pode representar um obstáculo à pesquisa, visto a dificuldade de avaliar a condição física do paciente, dado os hábitos por ele praticados no cotidiano.

Dessa forma, apesar das críticas relacionadas à pratica da idade nos estudos clínicos, esse fator de risco ainda continua sendo comumente relatado nas pesquisas. A idade é um fator de risco frequentemente empregado em estudos do escores de risco de morte e do tempo de permanência, mesmo para modelos já consolidados na prática médica e validadas em diferentes populações. Por esses motivos, é garantido neste estudo o impacto da variável Idade ao contexto do TP pós-operatório.

3.1.2 Gênero

Conceito associado ao aspecto biológico, o gênero é definido pela identidade dicotômica do sexo: masculino ou feminino.

A diferença de gêneros em doenças coronarianas é amplamente debatida na literatura e,

notavelmente mata mais mulheres do que homens, argumentado principalmente por diferenças físicas do músculo e órgão cardíaco (BARRETT-CONNOR, 1997; LAWTON, 2011 e PAPAKONSTANTINO et al., 2013).

Em consonância aos estudos que aborda a prevalência do sexo feminino, Butterworth et al. (2000) compararam o TP pós-operatório (exclusivamente UTI) de homens e mulheres submetidos à revascularização do miocárdio, onde se verificou uma maior associação do sexo feminino a gravidade do estado físico do paciente e conseqüentemente no aumento do período de internação.

Na década seguinte, Cocker et al. (2011) investigaram quais fatores de riscos são associadas a prorrogação da permanência do paciente na UTI após cirurgia cardíaca. Dentre os fatores de impacto e estatisticamente validados, o sexo feminino teve grande representatividade. Resultados similares também foram obtidos por Cacciatore et al. (2012), onde identificaram o sexo feminino como influenciador da permanência na UTI superior a três dias, independentemente da idade do paciente.

Além destes autores, Perterson et al. (2002), Shinjo e Fushimi (2015) e Almashrafi et al. (2016b) também relataram o predomínio do sexo feminino ao aumento do tempo de permanência após cirurgias cardiovasculares. De modo a jugar relevante a investigação quanto a associação deste fator ao TP pós-operatório.

3.1.3 Índice de Massa Corporal

O índice de massa corporal (IMC) é um fator de risco de avaliação frequente em doenças cardiovasculares (KAPLAN et al., 2014; COSTA et al., 2015). Conceitualmente, American Heart Association define IMC como uma medida de estimativa da gordura corporal de uma pessoa. Sua fórmula matemática avalia o peso corporal (quilogramas) em relação à altura (metros quadrados).

$$IMC = \frac{kg}{m^2} \quad (1)$$

O IMC é uma medida útil para população acima de 18 anos de idade, sendo recomendado uma abordagem diferente para crianças, menores de 18 anos e grávidas. Em geral, o cálculo é avaliado em cinco categorias que vão nos dois extremos da composição física corporal, são elas: baixo peso (IMC < 18,5), peso normal (IMC entre 18,5 e 24), sobrepeso

(IMC entre 25 e 29), obeso (IMC ≥ 30); para este último há ainda uma subclassificação para os diferentes níveis de obesidade (classe I para IMC entre 30 e 34, classe II para IMC entre 35 e 39, classe III ≥ 40) (WORLD HEALTH ORGANIZATION, 2002).

Para os pacientes classificados nos dois extremos das classes IMC, há uma maior predisposição a complicações pós-operatória e, este por sua vez, afeta o tempo de permanência no hospital.

Sob a análise do IMC superior a 30, dada natureza endêmica da obesidade na sociedade, inúmeros pacientes que necessitarão de algum tipo de cirurgia cardíaca, provavelmente estarão em excesso de peso (FLEGAL; CARROLL; OGDEN, 2012; VAN DER BOON et al., 2013).

A obesidade é muitas vezes associada ao surgimento de doenças secundárias que pode afetar significativamente a recuperação do paciente após a cirurgia. Terada et al. (2016) examinaram a associação da obesidade ao prolongamento do TP após cirurgia de revascularização do miocárdio, depois de confirmado sua relação ao surgimento de diabetes e colesterol LDL precoces.

Em contrapartida, uma vasta gama de estudos vem contrapondo a ideia de que a obesidade seja realmente um fator associativo. Ao et al. (2014) e Costa et al. (2015) frisam a não identificação da associação do obeso e sobrepeso ao aumento do tempo de permanência pós-operatório.

Autores que apoiem essa ideia defendem a existência do ‘paradoxo de obesidade’, onde pacientes obesos apresentam melhor recuperação pós-operatória que os pacientes considerados de massa corporal normal ou subnutrido.

Em seus estudos, Engel; Mcdonough e Smith (2009) destacam que pacientes abaixo do peso normal tiveram maior incidência de ventilação prolongada, reoperação por sangramento e falha renal, resultando em um tempo superior de internação pós-cirúrgico, do que pacientes considerados ‘gordos’. Ranucci et al. (2014) ainda destacam que o sobrepeso e obesidade reduz os efeitos nocivos da hemodiluição e transfusão em pacientes com maior risco de hipóxia pós-operatória.

Portanto, considerando as controvérsias encontrados na literatura, o impacto deste fator requer maiores investigações para que se confirme ou não sua associação ao TP, assim como proposto neste estudo.

3.1.4 Fração de ejeção

A fração de ejeção (FE) é uma medida percentual de quanto sangue é bombeado para

fora do ventrículo a cada vez que o coração se contrai. É uma importante maneira de determinar o quão bem o coração está bombeando sangue ao organismo, além de auxiliar no diagnóstico e acompanhamento de insuficiência cardíaca. Em outras palavras, corresponde a razão entre o débito sistólico e o volume diastólico final, valendo cerca de mais de 55% num indivíduo normal (AMERICAN HEART ASSOCIATION, 2015; WORLD HEALTH ORGANIZATION, 2002).

A FE pode ainda ser medida em exames clínicos com técnicas de imagens, no qual é predominante o uso de ecocardiograma, seguido por demais técnicas, como: cateterismo cardíaco, ressonância magnética cardiovascular, tomografia cardíaca computadorizada e teste de estresse nuclear (AMERICAN HEART ASSOCIATION, 2015).

O FE é um fator de risco relevante na gravidade de uma doença coronariana, sendo associada a insuficiência cardíaca que, por sua vez, eleva o período de permanência hospitalar ao comprometer a recuperação pós-operatória do paciente ou até mesmo elevar o risco de morte.

Em geral, estabelecem como valores de referência no estudo da fração de ejeção, os seguintes valores: fração de ejeção igual ou superior a 55% a função sistólica é adequada aos padrões normais de bombeamento do sangue, entre 45% e 54% percebe-se o início de leve disfunção sistólica, para 30% a 44% essa disfunção torna-se moderada, enquanto que valores inferiores a 30% são casos críticos de disfunção e deve ser acompanhado mais de perto pelo profissional da saúde (LaPAR et al., 2013).

Nesse sentido, Bootsma et al. (2015) estabeleceram implicações prognósticas para o efeito da fração de ejeção no ventrículo direito (FEVD) em pacientes submetidos à cirurgia cardíaca. Nos primeiros resultados foi identificado a correlação da FEVD média às primeiras 24 horas na UTI. Em seguida, confirmaram que FEVD inferior a 20% amplia o período do uso de ventilação mecânica após a cirurgia e também no uso de inotrópicos; impactando diretamente na maior taxa de mortalidade e no tempo de internação. Para Azarfarin et al. (2014) a fração de ejeção no ventrículo esquerdo inferior a 40% apresentou instabilidade hemodinâmica, sendo por isso necessário maior tempo de internação na UTI.

Além disso, Almashrafi et al. (2016a) na condução de uma revisão sistemática aos fatores que influenciam o TP pós-operatório, identificaram a baixa fração de ejeção como determinante em grande parte dos estudos analisados. Logo, verifica-se a elegibilidade da investigação deste fator ao estudo proposto.

3.1.5 Disfunção renal

A disfunção renal, também denominada de insuficiência renal, é caracterizada pela

incapacidade dos rins em remover resíduos e equilibrar os fluídos. O aparecimento do problema em pacientes com doenças cardíacas crônicas são frequentes, devido a coexistência de hipertensão arterial e diabetes que afetam a boa funcionalidade do coração e dos rins.

Taniguchi; Souza e Martins (2007) ainda destacam que pacientes diagnosticados com insuficiência renal permanecem internados por um período superior na unidade de terapia intensiva (UTI), os custos hospitalares são elevados e a diálise se torna necessária de 1% a 3% dos pacientes submetidos à cirurgia.

Em outro estudo, Barbosa et al. (2011) analisaram 3.890 pacientes que realizaram a cirurgia de revascularização do miocárdio em um hospital público brasileiro, onde 9,3% representam o grupo de pacientes com insuficiência renal crônica. A este grupo observou-se uma maior incidência de fibrilação atrial, síndrome de baixo débito cardíaco, maior tempo internado na UTI e na enfermaria pós-cirúrgica; além do aumento na taxa de mortalidade intra-hospitalar.

A identificação da disfunção renal é feita com base na taxa diária de creatinina liberada nos fluídos corporais. Comumente, essa taxa é usada como indicador da função renal através do cálculo da depuração (*clearance*) de creatinina (CARMONA et al., 2005). Em suma, refere-se a taxa ou razão da quantidade de sangue ou plasma depurado de creatinina por unidade de tempo – ml/min.

Para base de avaliação, esse indicador é estratificado em estágios da evolução da doença renal. O Ministério da Saúde (2008) define como padrões normais de depuração da creatinina, taxas com valor superior a 90 ml/min, enquanto que a depuração para valores entre 60 e 89 ml/min contribui com o leve decréscimo da função renal, entre 30-59 ml/min esse decréscimo já se torna moderado, lesão renal acentuada pode ser percebida para valores entre 15-29 ml/min e abaixo de 15 ml/min ocorre a falência da função renal.

A insuficiência renal desencadeia diversas complicações que impulsiona a necessidade de investigações mais detalhadas do seu real impacto direto e indireto no período de permanência hospitalar.

3.1.6 Diabetes mellitus

Diabetes mellitus (DM) é uma doença crônica que necessita de cuidados médicos contínuos, estratégias de redução de riscos multifatoriais e controle glicêmico. O uso de métodos que avaliam a frequência e a magnitude da hiperglicemia são essenciais no acompanhamento e controle da DM (AMERICAN DIABETES ASSOCIATION, 2013).

Sua prevalência na sociedade figura entre a quarta e a oitava posição como uma das principais causas de morte no mundo. O número crescente de diabéticos é acometido em virtude do crescimento e envelhecimento populacional, maior urbanização, prevalência de obesidade e sedentarismo (SOCIEDADE BRASILEIRA DE DIABETES, 2016).

Diretrizes nacionais e internacionais classificam os portadores de DM na categoria de alto risco em doenças cardiovasculares, capaz de avançar o aparecimento em até 15 anos. E para os portadores de diabetes dos tipos 1 e 2 com idade superior a 40 anos, geralmente apresentam risco de eventos coronarianos aumentados em até 2% (FOX et al., 2004; SOCIEDADE BRASILEIRA DE DIABETES, 2016).

Em termos gerais, possíveis pacientes com DM confirmam o diagnóstico a partir de testes como a hemoglobina glicolisada (Hb1Ac), onde valores de referência estabelecidos na literatura e vigentes na prática são essenciais na avaliação do nível médio glicêmico. Dessa forma, os pacientes que exibem níveis de Hb1Ac inferior a 5,7% proporcionam níveis normais de glicemia, para valores entre 5,7% e 6,4% há um aumento no risco de desenvolver diabetes (pré-diabetes), enquanto que valores acima de 6,5% confirmam o diagnóstico de diabetes.

Para a relação da DM ao tempo de permanência pós-operatório, pacientes com o diagnóstico confirmado aumenta a incidência de morte pós-cirúrgica, desenvolvimento de doença renal, acidente vascular cerebral, infecção de feridas esternal e a necessidade de suporte via intróptico. Nos estudos de Medhi et al. (2001), por sua vez, o nível glicêmico e a hemoglobina glicolisada (HbA1c) foram avaliados em 135 pacientes que aguardavam a realização da CRM, no centro médico de Westchester, Nova York, Estados Unidos. A hiperglicemia descontrolada de 7% ou mais foi utilizada como ponto de corte para TP de referência no pós-operatório de 6 dias. Os dados indicaram que o HbA1c causa maiores complicações e prolonga a estadia do paciente no hospital durante o período pós-operatório.

Najafi e Goodarzynejad (2012) ainda destacam a relevância do acompanhamento periódico dos pacientes com diabetes, visto que nestes casos, tanto o período de internação na UTI como na enfermaria se estende em até 48 horas e de 3 dia, respectivamente. Logo, o estudo revelou a necessidade do controle glicêmico anterior à cirurgia cardíaca para reduzir risco de morte e o tempo de recuperação do paciente. Torna-se, portanto, relevante avaliar o impacto da DM ao estudo aqui proposto.

3.1.7 Hipertensão arterial pulmonar

A hipertensão arterial pulmonar (HAP) é caracterizada pelo aumento da resistência

vascular ao evento circulatório. De forma que o aumento progressivo da resistência leva à insuficiência ventricular direita e até mesmo a morte precoce (CALLOU; RAMOS, 2009).

Ainda segundo as diretrizes da Sociedade Brasileira de Cardiologia apresentadas por Andrade et al. (2013) e Simão et al. (2013), os critérios de diagnósticos são: pressão sistólica da artéria pulmonar (PSAP) ≥ 30 mmHg e diastólica (PDAP) ≤ 15 mmHg, pressão média da artéria pulmonar (PMAP) ≥ 25 mmHg em repouso ou ≥ 30 mmHg durante a realização de exercício físico.

Para Kim et al. (2014) a HAP, além de agravar o risco de morte, favorece o aumento do período de internação pós-operatório, justificado por complicações decorrentes da cirurgia e o desenvolvimento de outras doenças relacionadas que exigem maior tempo para tratamento e recuperação do paciente.

Aos pacientes com doença arterial coronariana com quadro de hipertensão, há uma maior dificuldade de controle hemodinâmico e respiratório intraoperatório, fazendo-se necessário o uso de vasodilatadores como intervenção terapêuticas e uma maior permanência hospitalar para controle da pressão (CARMONA et al., 2005). Logo, a investigação desse fator como influenciador do tempo de permanência pós-operatória torna-se relevante.

3.1.8 Reoperações

Intervenção cirúrgica em reoperações exige maior acurácia técnica e cuidados especiais para a obtenção de resultados satisfatórios. Carmona et al. (2005) destacam que as reoperações são tecnicamente mais complexas, devido a existência de fibrose e distorção da anatomia que torna o procedimento mais delicado e prolongado.

De acordo com as Diretrizes Brasileira de Cardiologia apresentadas por Andrade et al. (2013) e Simão et al. (2013), as reoperações em cirurgias de revascularização do miocárdio, pode aumentar o sangramento pós-operatório com a necessidade de transfusões sanguíneas e maior período na unidade de tratamento intensivo. Além disso, os pacientes podem apresentar complicações respiratórias, neurológicas e até mesmo infarto do miocárdio, cujo risco de morte é de 3 a 5 vezes maior.

Sob o aspecto de período de permanência pós-operatório, Vivacqua et al. (2011) em um estudo do Institutional Review Board em Cleveland – Estados Unidos, abstraíram 18.891 informações de pacientes submetidos a reintervenção cirúrgica de revascularização do miocárdio ou combinada com outras cirurgias cardíacas, no período de 2000 a 2010. Entre os resultados identificados, foi observado que pacientes reoperados mantêm uma estadia superior

para amenizar o impacto de um possível sangramento pós-cirúrgico.

3.1.9 Admissões eletivas, urgência ou emergência

O paciente a ser submetido à cirurgia cardíaca pode dar entrada no hospital de modo eletivo (via de entrada pelo ambulatório) e não eletiva (urgência e emergência). No entanto, cirurgias não eletivas carrega um risco aumentado, sendo este aspecto considerado na maioria em modelos de estratificação do risco, como é o caso do EuroSCORE I e II (NASHEF et al., 1999, 2012).

Karthik et al. (2003) definem os conceitos de cirurgias eletivas e não eletivas. Sendo então a admissão eletiva definida como cirurgias pré-programadas e agendadas pelo profissional da saúde; no caso de cirurgias urgentes os pacientes não passaram pela lista de espera, mas requerem pronta atenção por razões de riscos à saúde; já as cirurgias de emergência são referenciadas por pacientes não programados e em situação crítica cuja intervenção deve ocorrer em até 24 horas.

Para esses pacientes não programados, há um aumento nas complicações cirúrgicas que elevam o período de internação pós-operatório. No caso da cirurgia de revascularização do miocárdio, Vegni et al. (2008) avaliaram o caráter eletivo *versus* urgente na cirurgia de revascularização do miocárdio, onde verificaram maior tempo de permanência na unidade de terapia intensiva, número de complicações e letalidade no período pós-operatório.

Nos casos de admissão de emergência, o estudo proposto por Schumer et al. (2016) avalia retrospectivamente informações de 5.940 pacientes submetidos à cirurgia cardíaca entre o período de janeiro de 2003 a dezembro de 2013, dos quais 212 pacientes foram admitidos como cirurgia de emergência. Em sua maioria, os pacientes experimentaram o uso do balão intra-aórtico, sangramento, diálise, morte intra-hospitalar e TP pós-operatório prolongado. Portanto, torna-se relevante investigar o impacto deste fator de risco ao caso investigado.

3.2 MÉTODOS DE PREVISÃO PARA O TEMPO DE PERMANÊNCIA HOSPITALAR

Avaliações clínicas são frequentemente suportadas pelo uso de ferramentas estatísticas. Diferentes estratégias de previsão podem então ser definidas com o interesse de estimar o tempo de permanência hospitalar, cuja decisão deve incidir em consonância às características que impactam no evento resposta do modelo.

A razão para se realizar previsão está na necessidade de antever e planejar o futuro, com

base em ações e comportamentos observados no passado e que se repetem ao longo do tempo, cujo mérito está na escolha de um modelo de boa exatidão e que esteja de acordo aos objetivos do estudo (WILKS, 2011).

Tais modelos fazem uso desde técnicas de aritmética simples até modelos com aplicações computacionais e estrutura matemática altamente complexa. De modo que, o uso destas ferramentas estatísticas, permite aos prestadores de serviços hospitalares o eficiente gerenciamento de custos, recursos e no atendimento satisfatório do paciente.

Nesse sentido, o TP hospitalar é uma métrica relativamente complexa, uma vez que sofre influência por diversos fatores externos e internos à organização, o qual muitos destes são de difícil mensuração e controle.

A escolha do modelo mais apropriado dependerá do tipo do delineamento e objetivo da pesquisa, das variáveis envolvidas no estudo e da maneira pela qual os dados foram coletados e organizados.

Diante disso, Awad; Bader-El-Den e McNicholas (2016), em continuação aos estudos de Castillo (2012), propuseram uma pesquisa designada de ‘estado da arte’, onde é realizado o mapeamento amplo das métricas de previsão, preconizadas ao longo das últimas três décadas. Para a triagem dos artigos, os autores avaliaram as práticas de estimação do TP em unidades de terapia intensiva e em instituições de saúde para o tratamento de doenças agudas.

Para tanto, foi realizado uma ‘varredura’ na literatura e, em seguida, a separação dos modelos de previsão em quatro grandes modalidades, que abrange: (1) métodos aritméticos, (2) métodos de regressão, (3) métodos orientados a dados e (4) métodos de multi-estágio (figura 3).

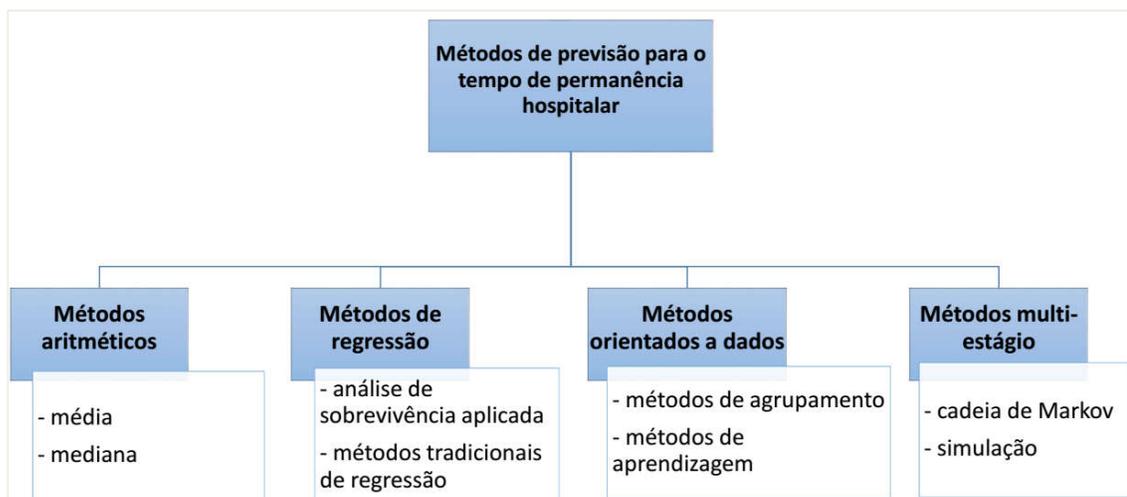
Dessa maneira, em cada uma das categorias são agrupados métodos de previsão que apresentam características em comum.

Logo, considerando a abordagem ampla e atual da pesquisa de Awad; Bader-El-Den e McNicholas (2016) e Castillo (2012), este estudo se propõe apresentar os métodos de previsão conforme configuração proposta pelos autores, juntamente com uma breve contextualização em cada uma das categorias.

São, ainda, delineados outros modelos de regressão não tratados nos estudos dos autores, mas que se atribui relevância ao tema proposto.

Sendo assim, a seguir é apresentado uma síntese de cada uma das categorias, juntamente com a exposição de alguns estudos predominantes no âmbito da estimação do TP pós-operatório.

Figura 3 – Classificação dos modelos de previsão



Fonte: Adaptado de Awad; Bader-El-Den e Mcnicholas (2016)

3.2.1 Métodos aritméticos

Para modelos preditivos de base aritmética simples, é usualmente computado a média e a mediana do evento investigado (AWAD; BADER-EL-DEN; MCNICHOLAS, 2016). Como uma maneira simples de medir o tempo de permanência, a validação desses modelos é comumente associada ao uso de estatística descritiva, onde é assumido que o evento estudado é normalmente distribuído e de baixa variabilidade.

Porém, na medida que a assimetria se configura no comportamento dos dados, uma alta variabilidade é então percebida e o uso dessas técnicas de previsão se tornam menos precisas e viáveis (WANKE; JULLIANELLI, 2011).

Ao tempo de permanência hospitalar, é comum o uso desse método de previsão em casos onde os resultados obtidos constituem uma avaliação secundária ao estudo ou, ainda, quando as condições acima são observadas.

Nesse sentido, a prática da estimação do TP pós-operatório, por meio do uso de média é apresentada nos estudos de Carvalho et al. (2011). Os autores avaliam a associação do TP nas complicações pós-operatórias e óbitos intra-hospitalar. As médias do tempo entre a cirurgia e a evolução para alta ou óbito foram estimadas de acordo com o *status* da presença das complicações e sua evolução. Para isso, fizeram uso de testes estatísticos na comparação de amostras dependentes para analisar as diferenças entre médias observadas nos seguintes pares de amostras: sobreviventes e óbitos sem complicação; sobreviventes e óbitos com complicação;

pacientes sobreviventes com e sem a complicação; e, pacientes não sobreviventes com e sem a complicação. Como resultado, os autores observaram que o TP era superior na presença de complicações em todos os subconjuntos definidos.

3.2.2 Métodos de regressão

Uma das preocupações dos métodos de regressão é criar modelos que especifiquem a estrutura de correlação entre os fenômenos observados, de modo a tornar conhecido seus atributos, bem como fazer afirmações de seus possíveis comportamentos.

A esta modalidade é concebida o título de um dos mais tradicionais métodos de previsão, tendo em vista sua aplicabilidade em diversos campos do conhecimento. No contexto do TP hospitalar, inclui-se a análise de covariáveis definidas com base em características do paciente e possíveis fatores de risco (CASTILLO, 2012).

Dentro desse grupo, os métodos de regressão são ainda fracionados em dois enfáticos subgrupos: análise de sobrevivência e modelos tradicionais de regressão, o qual serão discutidos a seguir.

3.2.2.1 Análise de sobrevivência

A análise de sobrevivência, também denominada de análise de sobrevida, é um conjunto de processos estatísticos utilizados na análise de dados, do qual a variável de interesse é definida como o tempo decorrente até que um acontecimento se verifique. Originalmente, o desfecho de interesse foi morte, cujo tempo decorrente até o seu acontecimento foi o foco da análise, fato que deu origem ao termo ‘análise de sobrevivência’ (SINGH; MUKHOPADHYAY, 2011).

Na área médica, a análise de sobrevida é predominante em estudos epidemiológicos sendo, no entanto, expandido para demais área da saúde, emergindo do interesse de conhecer e controlar eficientemente os processos hospitalares.

A este grupo destaca-se a aplicação do modelo Cox na estimação do tempo de permanência hospitalar, assim como observado nos estudos de Silber et al. (2003), Ravangard et al. (2011) e Cocker et al. (2011).

Para o primeiro estudo, é feito uso do modelo de regressão Cox na estimação do TP absoluto, do TP prolongado e na estimação da probabilidade de readmissão. Para isso, foram analisadas as variáveis comorbidades, renda, raça, localização da instituição de saúde (Nova York ou Pensilvânia) e o tipo de seguro como covariantes ao TP em pacientes com crise de

asma. Como resultados, os autores identificaram a prevalência da localização e probabilidade de readmissão como fatores associados ao TP pós-operatório.

Em outros estudos como o de Ravangard et al. (2011) é comparado o modelo de regressão Cox aos modelos preditivos paramétricos, o qual inclui o modelo exponencial, de Weibull, de Gompertz, log-normal, log-logística e gama. O estudo foi aplicado na determinação dos fatores que influenciam o tempo de permanência em pacientes das unidades de internação do Hospital da Mulher em Teerã, Irã. Em geral, o modelo Cox apresentou resultados eficientes na estimação, mas foi superado pelo modelo Gamma em relação a capacidade de discriminação das variáveis explicativas.

Já no trabalho do Cocker et al. (2011), o modelo Cox foi utilizado para previsão do TP em UTI pós-operatória, sendo o modelo validado em 395 pacientes submetidos à cirurgia. O modelo mostrou ser de boa calibração e excelente capacidade discriminativa em prever permanência na UTI em até 2, 5 e 7 dias, respectivamente.

3.2.2.2 Métodos tradicionais de regressão

A análise de regressão é a técnica de dependência mais amplamente usada e versátil, aplicável em diferentes facetas da tomada de decisão. É uma ferramenta analítica poderosa, planejada para explorar a relação de dependências entre as variáveis (HAIR JÚNIOR. et al, 2009).

Esses modelos são definidos pelo uso de técnicas estatísticas capaz de estabelecer relações entre variáveis que inter-relacionam em um determinado processo. Logo, ao estudo do TP hospitalar, enquadram-se ao grupo os modelos de regressão linear múltipla, regressão logística e árvore de regressão (AWAD; BADER-EL-DEN; MCNICHOLAS, 2016).

Em relação aos modelos citados, Huang; Hooper e Marrie (2006) comparam a performance dos três modelos na estimação de TP pós-operatório, a partir da associação dos fatores de riscos em pacientes com diagnóstico de pneumonia. A perda de peso, *status* funcional, doenças do coração pré-diagnosticadas e disfunção renal foram elementos decisivos no TP superior a sete dias. Além disso, os autores verificaram uma leve superioridade da árvore de regressão na determinação dos fatores de risco de impacto ao tempo de permanência.

No estudo de pacientes em tratamento do câncer gastrointestinal, Pourhoseingholi et al. (2010) compara a regressão múltipla linear e regressão quantílica (estimativa aproximada da mediana na variável preditora) na estimação do tempo de permanência hospitalar e na inferência dos fatores de risco associados.

A abordagem de regressão logística – sendo sua capacidade preditiva avaliada neste estudo – é predominante em grande parte da literatura relacionada, sendo comumente aplicado na investigação do TP prolongado. Para isso, é avaliado os fatores de risco relacionados ao maior tempo de permanência do paciente no hospital, onde estima-se a probabilidade de determinado paciente permanecer por um período de internação superior a uma meta estabelecida.

A essa abordagem incluem-se os estudos de Azarfarin et al. (2014) que estimam o TP prolongado na UTI, com base no impacto dos fatores de risco associados. Pagowska-Klimek et al. (2011) discursam o impacto dos fatores de risco pré-operatórios ao prolongamento do tempo de permanência nas UTIs infantis de cirurgias cardíacas e, em seguida, são estimados a probabilidade desses pacientes permanecerem por mais de três dias na UTI. Almashrafi et al. (2016b) faz uso do método RL para identificar as variáveis preditoras e modelar o escore de risco para TP prolongado na UTI de um hospital de Oman. Esse escore apresenta três grupos distintos e que define o grau de risco do paciente quanto complicações pós-operatório; logo, para o primeiro grupo e de menor risco é definido um período padrão de permanência de até 5 dias, entre 6 e 10 dias o risco se torna moderado e acima de 11 dias o risco de ocorrência de complicações se torna elevado.

Shinjo e Fushigami (2015), também faz uso do método para avaliar o custo final com o tempo de permanência prolongado a partir dos fatores de riscos cardiovasculares, demográficos e regionais. Já Najafi e Goodarzynejad (2012) faz o uso tradicional da técnica para prever o a probabilidade de o paciente permanecer internado por um período prolongado na UTI (acima de 1 dia) e na enfermaria (acima de 3 dias). Por sua vez, Khairudin; Mohd e Hamid (2016) compararam o modelo de regressão logística e árvore de decisão na previsão da permanência prolongada em pacientes submetidos à CRM; a precisão preditiva para a árvore de decisão e o modelo de regressão logística foram de 65,86% e 75,87%, respectivamente, revelando a superioridade da RL frente ao outro modelo.

3.2.3 Métodos orientados a dados

A abordagem de métodos orientados a dados para previsão refere-se a técnicas capazes de classificar e agrupar informações. Tais técnicas são usadas para descobrir padrões que proporcionem a segmentação dos dados em grupos, proporcionando visibilidade de novas e interessantes relações entre os dados (AWAD; BADER-EL-DEN; MCNICHOLAS, 2016; CASTILLO, 2012). Na proposição de modelos de previsão ao TP hospitalar, inclui a esta

categoria os métodos de previsão por agrupamento e por aprendizagem.

Em modelos de agrupamento é proposto os métodos de *Cluster*, análise de correspondência e de Mineração de Dados por Awad; Bader-El-Den; Mcnicholas, (2016); Banu e Gomathy, (2014); além do método classificador Naive Bayes, proposto por Barbini et al. (2014).

Castillo (2012) ainda propõe métodos aprendizagem como técnicas eficiente de abordagens relacionadas as Redes Neurais Artificiais (RNA) – do inglês: *Artificial Neural Network* (ANN), Máquinas de Vetor Suporte (MVS) – do inglês: *Support Vector Machine* (SVM) e método Fuzzy.

3.2.3.1 Modelos de agrupamento

Os métodos baseados em técnicas de agrupamento são propostos em situações onde há pouco conhecimento a priori das leis que regem o sistema de onde foi coletado a amostra, de modo que, os dados são então agregados conforme padrões identificados.

Nos métodos de *Cluster*, Análise de Correspondência e Mineração dos Dados são descobertos padrões muitas vezes inexplorados, de modo a agrupar os fatores de riscos de acordo com sua maior ou menor influência (BANU; GOMATHY, 2014). Além disso, quando estes são aplicados em conjunto aos métodos de regressão, a acurácia deste último é melhorada substancialmente para a previsão do TP pós-operatório (ROUZBAHMAN; JOVICIC; CHIGNELL, 2016). No caso do método classificador de Naive Bayes, Barbini et al. (2014) propõem o uso desta ferramenta na seleção de variáveis associadas ao TP e também na modelagem de um sistema de pontuação que estime a probabilidade de o paciente permanecer internado por um período superior na UTI.

3.2.3.2 Métodos com técnicas de aprendizagem

A aplicação destes modelos leva em consideração que o evento resposta passa inicialmente por uma fase de aprendizagem ou treinamento, aonde são identificados diferentes comportamentos associadas com as variáveis de impacto e, que deverão se propagar em eventos futuros (CASTILLO, 2012).

Esses modelos apresentam ainda como vantagem, a capacidade de capturar efeitos não lineares no comportamento dos dados que, nos modelos tradicionais de previsão são comumente avaliados dentro do parâmetro de erro (HAYKIN, 2003).

Ao uso dessas técnicas no TP hospitalar, LaFaro et al. (2015) comparam um modelo de RNA aos modelos tradicionais de regressão para prever o tempo de permanência na UTI após cirurgia cardíaca. Com base nas características do paciente no período de pré-incisão cirúrgica, porém, não foi identificado a superioridade do RNA em relação aos demais modelos. Houthoof et al. (2015) avaliaram a capacidade preditiva do SVN e RNA para escore de risco de morte e tempo de permanência em UTI pós-operatória, apresentando resultados satisfatórios para ambos os modelos, com superioridade as RNA. Tsai et al. (2016), por sua vez, obteve resultados satisfatórios ao fazer uso das RNA para prever o TP em pacientes internados com um dos três diagnósticos: aterosclerose coronariana, insuficiência cardíaca e infarto agudo do miocárdio, em um hospital cristão de Tapei, Taiwan. A previsão foi satisfatória em 88% dos casos aplicados na validação do modelo.

Hachesu et al. (2013) avaliaram o uso dos métodos de aprendizagem RNA e MVS para prever o TP hospitalar, sendo identificado alta precisão na performance destes modelos. Já Coutinho et al. (2014) aplicam o modelo Fuzzy para prever o tempo médio de internação, com base no impacto de variáveis externas ao paciente, tais como: exposição dos pacientes aos poluentes do ar, temperatura e vento. O modelo mostrou-se satisfatório considerando o dióxido de enxofre (poluente do ar) como variável de impacto.

3.2.4 Modelos de multi-estágio

Os modelos de previsão multi-estágios são modelos baseados em processos estocásticos, permitindo que os indivíduos se movam entre um número finito de estados, cuja transição entre eles é definida por técnicas probabilísticas. A complexidade depende do número de estados do modelo e também das possíveis transições (MACHADO, 2011).

Em síntese, são definidos como modelos de estudos longitudinais, cuja finalidade é estimar a distribuição de probabilidade para o período de transição entre os estágios (definidos como procedimentos sequenciais) do evento de interesse.

Machado (2011) e Awad; Bader-El-Den e Mcnicholas (2016) compartilham da ideia de que no TP hospitalar as questões de interesse incluem a estimativa da taxa de progressão na recuperação do paciente e os efeitos dos fatores de risco ao período de internação.

Para isso, é modelado os diferentes processos envolvidos desde a admissão do paciente ou a sua inserção na sala de operação até a ocorrência da alta ou óbito e, por conseguinte, são determinadas as distribuições das transições que regem o comportamento do sistema em estudo.

De modo geral, neste grupo o TP pós-operatório estima a distribuição de probabilidades

do tempo de permanência hospitalar desde que o paciente foi liberado da sala cirúrgica – procedimento 1, até a sua alta ou óbito – procedimento 2.

É dado destaque ao modelo da Cadeia de Markov e demais modelos de simulação computacional. A principal vantagem desses modelos está na capacidade de capturar aspectos estocásticos do evento em questão (AWAD; BADER-EL-DEN; MCNICHOLAS, 2016). Para o tempo de permanência hospitalar, reforçam essa ideia os estudos de DeRienzo et al. (2016), Jeremic e Tan (2008) e Albornoz et al. (2006).

Sendo assim, para o próximo tópico pretende-se aproximar o pesquisador e o leitor aos métodos de previsão aplicados neste estudo. Contudo, não é o intuito detalhar o desenvolvimento do modelo de regressão logística e redes neurais artificiais, visto que os cálculos serão realizados com o auxílio de *softwares* especializados (vide seção 4), de modo que, não se faz necessário traçar detalhes de cunho operacional. Porém, considera-se relevante compreender quais as hipóteses e definições que fundamentam o desenvolvimento e validação dos modelos propostos, justificando assim, a tratativa apresentada.

3.3 REGRESSÃO LOGÍSTICA

Em um breve relato histórico, a regressão logística surgiu por volta da década de 40, sendo conferido notoriedade a partir do famoso estudo *Framingham Heart Study* com colaboração da Universidade de Boston – EUA. O principal objetivo foi identificar os fatores que concorrem no desencadeamento de doenças cardiovasculares e consecutivamente, a probabilidade de o paciente desenvolver doenças do coração nos 10 anos seguintes (DIAS FILHO; CORRAR, 2007). O modelo se expandiu da medicina, sendo hoje empregado nos diversos campos do conhecimento.

O que diferencia este método das demais técnicas de regressão refere-se: a capacidade de lidar com variável dependente não-métrica binária (dicotômica) ou múltipla; a acomodação de todos os tipos de variáveis independentes, sejam elas métrica ou não métricas e a não exigência de suposição de normalidade dos dados (HAIR JÚNIOR et al., 2009; CASELLA; BEAGER, 2014).

Para o evento resposta de interesse, o método de regressão logística pressupõe a definição de uma regra de decisão, cujos valores se encontram no intervalo de zero e um. Para isso, o método converte a variável dependente em razão de chance (também denominada *odds ratio*) com base logarítmica e assim evitar a predição de valores fora do intervalo (DIAS FILHO; CORRAR, 2007; HAIR JÚNIOR et al., 2009).

Por mera convenção, a probabilidade associada a cada *odds ratio* é então expressa na razão da probabilidade de sucesso em relação a probabilidade de fracasso que, por sua vez, deve ser convertido ao logaritmo natural e expresso em função das variáveis independentes, tal como segue a expressão matemática:

$$\ln \left(\frac{p(\text{sucesso})}{1-p(\text{fracasso})} \right) = b_0 + b_1 x_{1i} + \dots + b_k x_{ki} \quad (2)$$

Essa expressão pode ainda ser expressa na base e , descrita da seguinte forma:

$$P(\text{evento}) = \left(\frac{1}{1 + e^{-(b_0 + b_1 x_{1i} + \dots + b_k x_{ki})}} \right) \quad (3)$$

Onde, $P(\text{evento})$ expressa a probabilidade de ocorrência de determinado evento e (b_0, b_1, \dots, b_k) são os coeficientes estimados das variáveis independentes x_{ki} .

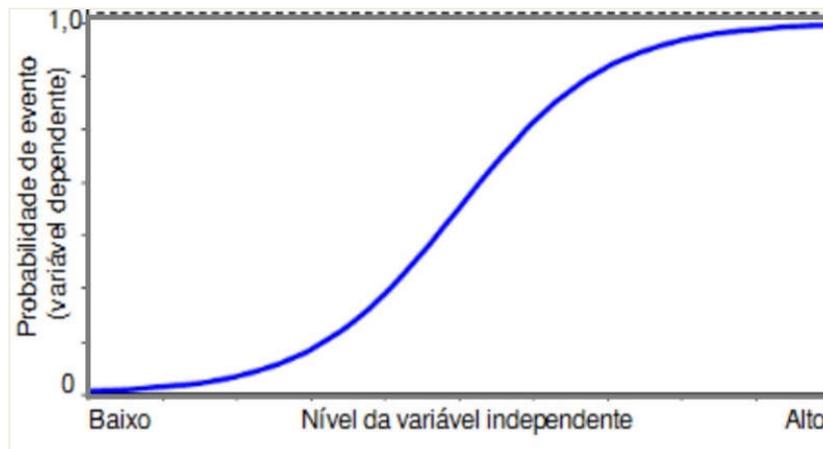
A probabilidade de ocorrência de determinado evento é então definida por $E(y/x_k) = P(\text{evento})$, de modo que os resultados são ainda interpretados como “zero” ou “um” para variável dicotômica ou ainda descritas em categorias no método múltiplo.

Para a estimação dos parâmetros das variáveis independentes é realizado o método de Máxima Verossimilhança – do inglês, *Maximum-Likelihood Estimation* (MLE). Tal método é uma forma de estimar os parâmetros de distribuição de probabilidades que maximize a função de verossimilhança, ou seja, o método busca-se por coeficientes que permita estimar a maior probabilidade possível de um evento ocorrer ou de certa característica estar presente.

Sob o ponto de vista conceitual, a probabilidade de ocorrência de determinado evento pode ainda ser representado graficamente, onde é avaliada a relação da estimação de probabilidade ao conjunto de variáveis predictoras, denominado de Curva da Regressão Logística ou Curva S, variando entre zero e um (figura 4). É uma ferramenta muito utilizada em estudos econômicos e em medicina para modelar a relação existente entre as variáveis.

As variáveis predictoras são então avaliadas quanto à disposição da amostra na curva, a fim de que quanto mais próximo das extremidades melhor será a capacidade preditiva do modelo. Outro fato a ser observado, é que quanto mais próximo das extremidades os valores da variável independente se encontram, maior deve ser o incremento dos valores para que a probabilidade se altere, fato que define a inclinação gráfica suave nas extremidades e o quão representativo é a variável ao modelo.

Figura 4 – Curva da regressão logística



Fonte: Dias Filho e Corrar (2007)

Em particular aos estudos do TP pós-operatório é comumente empregado o modelo de ordem binária, cujo intuito é estimar a probabilidade de o paciente permanecer internado em até t dias ou prosseguir internado por um período prolongado (igual ou maior $t+1$ dias) e vice-versa.

3.3.1 Métodos de diagnóstico

Diagnosticar o desempenho de um modelo de previsão é essencial para verificar se o comportamento da realidade pode ser aplicado a eventos futuros, contribuindo com informações acuradas e na qualidade das decisões. Para a regressão logística, se faz necessário o teste Razão de Verossimilhança, R^2 de Nagelkerke, Bootstrapping e Wald.

3.3.1.1 Teste Razão de Verossimilhança

O teste de Razão de Verossimilhança, também denominado de teste Qui-quadrado da Razão de Verossimilhança, trata-se de um indicador que visa avaliar o ajuste do modelo para cada variável explicativa inserida (DIAS FILHO; CORRAR, 2007).

O objetivo é avaliar a significância estatística na inclusão das variáveis preditoras, ou seja, o quanto essas variáveis podem melhorar o ajuste do modelo. Para isso, o teste avalia a diferença da razão entre o modelo composto apenas pela constante e o modelo acrescido da variável preditora (nulo *versus* composto). Se o teste estatístico suporta diferenças significativas entre os modelos, então pode-se afirmar que as variáveis são significantes na melhora do ajuste do modelo (HAIR JÚNIOR et al., 2009).

O teste é comumente empregado na literatura pela abreviação -2LL, cujo termo advém do inglês: *-2 log likelihood value*, onde a formulação matemática é definida da seguinte maneira:

$$-2LL = -2\ln \left[\frac{\text{verossimilhança SEM a variável}}{\text{verossimilhança COM a variável}} \right] \quad (4)$$

Ou ainda:

$$-2LL = -2\ln(L_S) + 2\ln(L_C) \quad (5)$$

Onde,

L_S é o cálculo da máxima verossimilhança sem a variável e L_C com a variável.

Esse cálculo da diferença permite inferir que testar a hipótese nula de que o modelo final apresenta um melhor ajuste do que o modelo base (com a inserção apenas da constante e probabilidade de que todos os pacientes permaneçam internados dentro do período padrão estimado) e por isso, a diferença do -2LL deve se aproximar de zero. Logo, quanto mais elevada a diferença no ajuste entre o modelo base e o modelo final, maior será o potencial das variáveis explicativas em prever os resultados.

$$\begin{aligned} H_{nulo}: -2\ln L_S - (-2\ln L_C) &= 0 \\ H_{alternativo}: -2\ln L_S - (-2\ln L_C) &\neq 0 \end{aligned} \quad (6)$$

Para este teste ainda é possível realizar a análise individual dos modelos a cada passo da inserção de novas variáveis, de modo que para -2LL com valor superior a 100 indicam que o modelo investigado pouco prediz sobre a realidade observada e baixos valores manifestam um bom ajuste entre o modelo e os dados (FIELD, 2013).

Porém, neste último caso, a análise individual pode ser um problema, visto que valores abaixo de 100 nem sempre expressa um bom ajuste se a diferença em relação ao modelo base não for estatisticamente significativa.

3.3.1.2 R^2 de Nagelkerke

Como uma extensão ao coeficiente de determinação – R^2 , a definição do R^2 de

Nagelkerke se faz importante na análise da proporção de variância explicada pelo modelo de regressão com base na abordagem da máxima verossimilhança. Maiores valores dessas medidas indicam melhor ajuste do modelo. (CASELLA; BERGER, 2014).

Outras medidas disponíveis e comparáveis ao R^2 de Nagelkerke é o R^2 de Cox-Snell e o Pseudo R^2 (HAIR JÚNIOR et al., 2009). Contudo, a escolha pelo R^2 de Nagelkerke é dada pela maior aproximação da medida ao clássico R^2 , independe do tamanho da amostra e seus parâmetros estão disponíveis na maioria dos *softwares* estatísticos.

3.3.1.3 Teste *Bootstrapping*

Introduzida por Efron em 1979, mas popularizada pelo mesmo autor e Tibshirani em 1993, a ideia principal é gerar diferentes e aleatórios conjuntos de dados de mesmo tamanho pelo processo de reamostragem. Novas estatísticas para cada amostra são então geradas como se esta primeira representasse exatamente toda a população.

Geralmente é definido 1000 reamostragem como satisfatório para diferentes número de amostras, sendo este valor estimado pelo autor após a realização de diversos estudos empíricos.

Em termos gerais, o teste *bootstrapping* é descrito sucintamente como um método de reamostragem aplicado para avaliar a variabilidade de estimadores, definir probabilidades ou os intervalos de confiança para os parâmetros do modelo, comparar resultados relatados em outros testes estatísticos, explorar a forma de distribuição da amostra e desenvolver distribuições preditivas com base no comportamento assintótico das variáveis (ADJEI; KARIM, 2016).

Na prática dos métodos de previsão, a técnica pode ser aplicada nas diferentes etapas da modelagem. Nos estudos de Billah et al. (2010) e Mejía et al. (2013), por exemplo, o método é aplicado junto ao procedimento de seleção das variáveis, a fim de produzir diferentes combinações aos modelos candidatos. No final, um modelo é então desenvolvido com base em uma pré-seleção das variáveis ranqueadas a partir dos modelos aspirantes.

Já em outros estudos, tais como proposto por Adjei e Karim (2016) e Altman et al. (2009), a medida é aplicada na validação dos parâmetros do modelo a partir da correção do viés ao longo do processo de reamostragem. Nesse processo, novos intervalos de confiança são estimados com base na melhoria incorporada durante o procedimento *bootstrapping*. Os coeficientes do modelo são então reavaliados quanto ao nível de confiança em permanecer no novo intervalo estimado.

3.3.1.4 Teste Wald

Para o teste estatístico Wald, a finalidade é aferir o grau de significância de cada coeficiente da equação logística, inclusive a constante. Mais precisamente, busca-se aferir se cada parâmetro estimado é significativamente diferente de zero, ou seja, testa a hipótese de que um determinado coeficiente não é nulo (HAIR JÚNIOR et al., 2009; DIAS FILHO; CORRAR, 2007).

$$\begin{aligned} H_{nulo}: b_k &= 0 \\ H_{alternativo}: b_k &\neq 0 \end{aligned} \quad (7)$$

Há, contudo, uma particularidade que deve ser considerada no uso dessa estatística: quando o valor absoluto do coeficiente é expressivo o erro-padrão a ele associado pode ficar distorcido e erros do tipo I no teste de hipótese podem vir a ocorrer.

Para contornar o problema, os autores recomendam o cálculo do -2LL com e sem a variável a que se refere o coeficiente investigado e, em seguida, compara-se os dois valores para verificar se o coeficiente em questão realmente exerce impacto significativo no modelo. Isso reforça a ideia de que é recomendável o uso de mais de um tipo de indicador como apoio na validação do modelo.

3.4 REDES NEURAIAS ARTIFICIAIS

A Regressão Logística e Redes Neurais Artificiais compartilham raízes comuns no reconhecimento de padrões estatísticos presentes nos dados. Contudo, a escolha deste último método pode ser justificada pela sua capacidade de prever os resultados em princípios estatísticos não-lineares e com dados de natureza não paramétrica (TSAI et al., 2016).

Em termos gerais, RNAs são métodos elaborados para solucionar problemas na ciência da inteligência artificial, a partir da construção de sistemas organizados em circuitos que simulam o funcionamento do cérebro humano, inclusive no seu processo de aprendizagem e erro.

Um dos trabalhos pioneiros registrados na literatura a respeito das RNA ocorreu em 1943, a partir das premissas apresentadas por McCullock e Pitts à revista de Neurocomputação - *Bulletin of Mathematical Biophysics*, no trabalho de publicação denominada '*A logical calculus of the ideas immanent in nervous activity*'. Os autores realizaram um estudo sobre o

comportamento do neurônio biológico com o objetivo de criar um modelo matemático capaz de representar esse comportamento com base no processo de aprendizagem.

Em relação aos métodos de previsão, nas RNAs o comportamento dos dados é transmitido aos eventos futuros por intermédio do processo de treinamento (aprendizagem) realizado na rede. Para isso, as conexões entre os neurônios, denominados de pesos sinápticos são introduzidos para armazenar o conhecimento adquirido durante o treino da rede (HAYKIN, 2003).

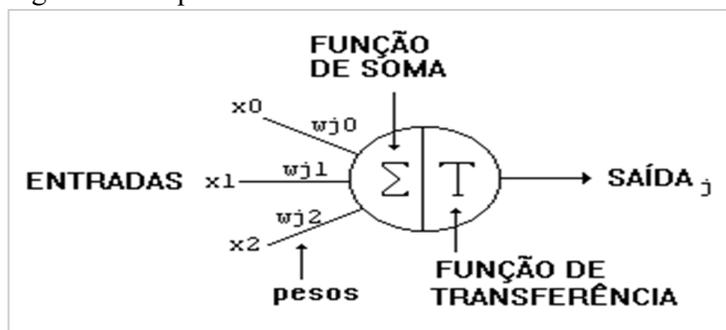
Em sua arquitetura básica, a RNA é composta pela estrutura tipicamente organizada em camadas, cujos neurônios (nós) são conectados àqueles da camada posterior e, assim sucessivamente, conforme destacado a seguir:

- a) camada de entrada: onde os padrões são apresentados à rede;
- b) camadas intermediárias ou ocultas: onde é feita a maior parte do processamento, através das conexões ponderadas; podem ser consideradas como extratoras de características;
- c) camadas de saída: onde o resultado final é então apresentado.

Nesse sentido, a figura 5 auxilia na interpretação dessa estrutura base, onde a camada de entrada é composta por x_j ($j = 0, 1, 2$) neurônios, designadas pelas variáveis predictoras do modelo de previsão. Esses nós são então conectados aos nós da camada seguinte, por um conjunto de linhas direcionadoras sinápticas que, no processo de treinamento adquirem pesos w_j ($j = 0, 1, 2$) para os ajustes aos estímulos recebidos.

Cada conexão de entrada, já filtrada por seu peso, é então admitida à camada oculta por onde participará da somatória com as demais entradas e depois submetido a uma função de ativação. Nesse momento é determinado o limiar de disparo para a saída do neurônio que percorrerá a próxima camada ou exibirá a resposta final, conforme esquema similar ao neurônio biológico.

Figura 5 – Arquitetura básica de uma Rede Neural Artificial



Fonte: Tafner (2008)

3.4.1 Tipos de arquitetura para redes neurais artificiais

Para a descrição da arquitetura de uma rede neural, Haykin (2003) observa a existência de três tipos principais de arquitetura, são elas: redes *feedforward* com uma única camada, redes *feedforward* com múltiplas camadas e redes recorrentes.

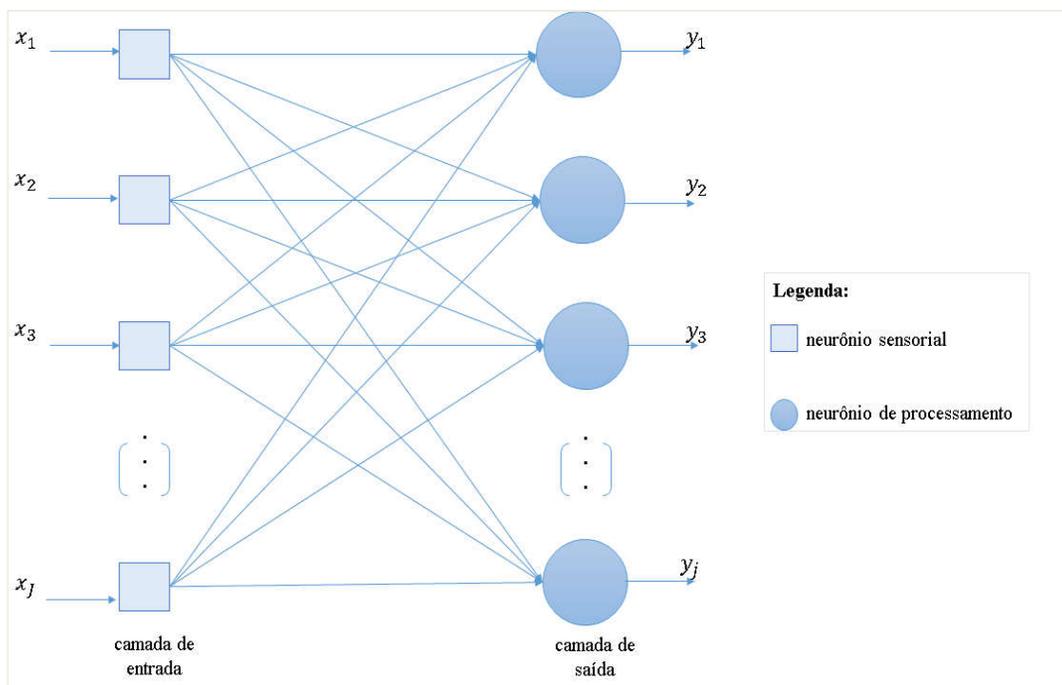
Dessa forma, para auxiliar no critério de escolha e na compreensão dos fenômenos envolvidos, cada uma das arquiteturas é assim sintetizada no tópico a seguir, conforme a literatura proposta por Haykin (2003) e Braga; Carvalho e Ludermir (2011).

3.4.1.1 Redes *feedforward* com uma única camada

Proeminentemente, uma rede definida como *feedforward* é caracterizada por uma estrutura onde inexistem ciclos, sendo por isso denominada de rede direta. E, para sua forma mais simples, a arquitetura é composta apenas por uma camada de entrada e uma de saída.

Geralmente, os neurônios de entrada são considerados propagadores puros, de modo que simplesmente repetem o sinal de entrada em sua saída distribuída. Já as unidades de saída costumam ser unidades processadoras da função lógica somatória, como apresentado na figura 6 a seguir. É uma arquitetura comumente aplicada à Rede Hebb, Perceptron Simples e Adaline.

Figura 6 – Rede *feedforward* de uma única camada



Fonte: Adaptado de Silva; Spatti e Flauzino (2016)

3.4.1.2 Redes feedforward de múltiplas camadas

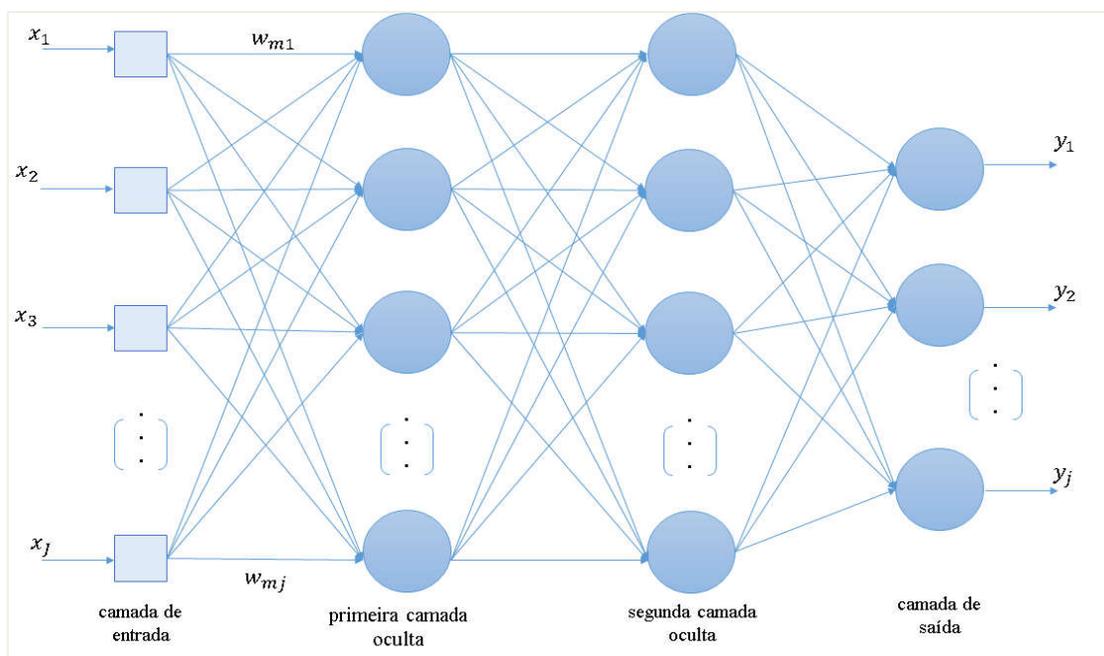
Essa arquitetura se distingue da anterior pela presença de uma ou mais camadas intermediárias ou ocultas, cujos neurônios essencialmente são unidades processadoras no exercício da função lógica.

A quantidade de camadas ocultas selecionadas aumenta o poder computacional de processamento e armazenamento de informação. Contudo aumenta também sua complexidade e o período gasto com o processamento de toda a rede e que, para o uso diário pode se tornar inviável. Fato que deve ser considerado durante o processo de definição do número de camada ocultas, normalmente definidas de maneira empírica.

Assim como a arquitetura anterior, redes com esta estrutura também se caracterizam pela alimentação das informações à frente, tal como segue na figura 7. As principais redes que faz uso dessa arquitetura referem-se a rede *Perceptron* Multicamadas, do inglês: *Multilayer Perceptron* (MLP), Madaline e as Redes de Função Base Radial.

Esse tipo de arquitetura vem ganhado espaço em pesquisas no campo da medicina, principalmente em estudos epidemiológicos que valer-se-ão de técnicas de previsão na investigação da doença, seus fatores de risco e o impacto deste último ao paciente, instituição de saúde e comunidade em geral (TSAI et al., 2016; PENG; PENG, 2008).

Figura 7 – Rede *feedforward* com múltiplas camadas



Fonte: Adaptado de Silva; Spatti e Flauzino (2016)

Em particular ao estudo do tempo de permanência pós-operatório, LaFaro et al. (2015) argumentam o uso desta arquitetura com base no desempenho satisfatório do método em outros estudos no campo da medicina, o número maior de recursos disponíveis para a modelagem e a simplicidade na utilização, quando comparado às demais arquiteturas de multicamadas.

No processo de treinamento da rede ainda é possível otimizar os resultados a partir da aplicação do algoritmo *backpropagation*, conhecido também como regra Delta generalizada ou gradiente de descida. Para melhor entendimento do princípio de funcionamento, Silva; Spatti e Flauzino (2016) esclarece que esse processo é dividido em duas etapas, a primeira é destinada ao ajuste dos pesos sinápticos conforme erro estimado da diferença entre o valor desejado e o obtido; denominada de procedimento *forward*, esta etapa segue sequencia padrão exposto na camada de saída.

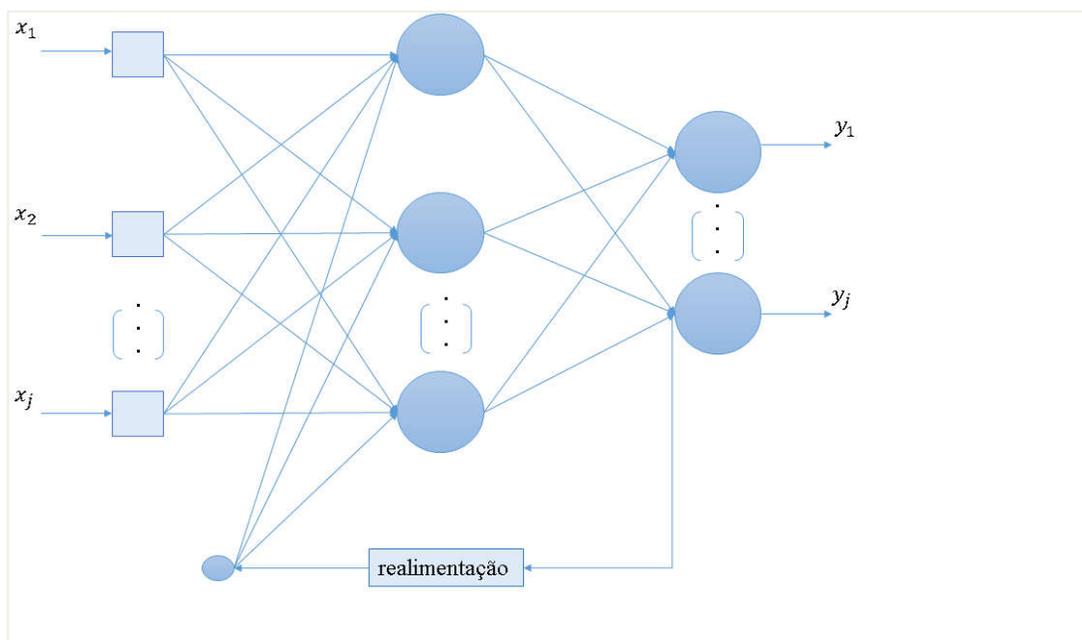
Já na segunda etapa, refere-se ao ajuste das demais camadas, seguindo direção contrária à saída de resultados, ou seja, o ajuste ocorre da camada de saída para a última camada oculta, esta por sua vez, realiza o ajuste aos pesos que a conectam à camada anterior, e assim sucessivamente até a camada de entrada da rede. Em termos gerais, este procedimento atua na minimização de erros em toda a rede.

3.4.1.3 Redes recorrentes simples e de multicamadas

Esse modelo de arquitetura se distingue das demais pela existência de *loop* de recorrência, de modo que na camada de saída possui ao menos uma ligação que realimenta a rede, assim como observado na figura 8. Redes *Bidirectional Associative Memory* (BAM), *Adaptative Resonance Theory* (ART 1 e ART2) e Redes de Elman são exemplos de redes recorrentes.

Dessa forma, embora RNAs diferem da maneira ao qual os neurônios estão conectados e como as entradas são processadas, os esforços neste estudo estão concentrados nas redes do tipo *feedforward*, conforme apresentado na próxima da seção.

Figura 8 – Rede recorrente



Fonte: Adaptado de Silva; Spatti e Flauzino (2016)

3.4.2 Definição do número de camadas ocultas e neurônios

A determinação do número camadas ocultas e o total de neurônios em cada uma delas ainda é questionado em muitos campos da pesquisa científica, visto que não há um consenso de soluções matemáticas que auxiliem na proposição de uma arquitetura ótima de rede. A experiência tem mostrado que a determinação do número de neurônios na camada intermediária é em essência um processo de tentativa e erro.

Além disso, subestimar esse número pode resultar no chamado *underfitting*: onde a estrutura é insuficiente para manipular as entradas, dificultando de sobremaneira a convergência dos algoritmos de treinamento. Por outro lado, número excessivo de neurônios na(s) camada(s) intermediária usualmente gera *overfitting*, ou seja, a diminuição da capacidade de generalização (HAYKIN, 2003).

Há, contudo, pesquisadores que, após realizarem diversas pesquisas empíricas defendem o uso de não mais que duas camadas ocultas para modelo *feedforward* multicamadas, pois além da complexidade na estrutura, os resultados destas pesquisas têm demonstrado que não há melhora significativa com números elevados de camadas ocultas (SILVA; SPATTI; FLAUZINO, 2016). Logo, este estudo se limita a modelagem com no máximo 2 camadas ocultas.

Quanto ao número de neurônio, mesmo não havendo consensos, são comumente

sugeridos para pontos de partida as seguintes condições:

- a) número intermediário entre o número de neurônios na camada de entrada e o número de neurônios na camada de saída;
- b) dois terços do número de neurônios na entrada mais o número de neurônios de saída;
- c) um número menor que o dobro de neurônios na camada de entrada (por exemplo, $2n - 1$, onde n é o número de neurônios na entrada); e,
- d) dez por cento do número amostras.

3.4.3 Processos de aprendizagem

Uma das propriedades mais marcantes dos métodos de aprendizagem e que difere ela aos demais métodos de previsão refere-se a habilidade da rede de ‘aprender’ conforme o comportamento dos dados ao qual a rede foi alimentada. No caso das RNA, o grande diferencial está em sua arquitetura que permite o mapeamento e classificação de diferentes características dos dados

Haykins (2003) argumenta que a aprendizagem, também denominado de treinamento, é o processo pelo qual os parâmetros de uma rede neural são adaptados através de um mecanismo de estímulos fornecidos pelo ambiente no qual a rede está inserida. O tipo de treinamento é definido pela forma na qual os parâmetros são modificados.

Nas palavras de Braga; Carvalho e Ludermir (2011) esse procedimento é realizando ajustando-se os pesos e limiares (contornos / *bias*) em processos iterativos, cuja finalidade é a obtenção de um algoritmo de aprendizagem que melhor represente o comportamento observado e, conseqüentemente, permita uma solução generalizada à certa classe do problema.

O fator importante no processo de treinamento é a maneira pela qual uma rede neural se relaciona com o ambiente. Nesse contexto existem alguns paradigmas de aprendizado que deve ser avaliado no processo de escolha quanto a melhor configuração da rede. Para isso, Silva; Spatti e Flauzino (2016) apresentam algumas estratégias de treinamento que devem ser consideradas e até mesmo combinadas no processo de modelagem, visto que influencia diretamente no modo de como é realizada a aprendizagem da rede:

- a) aprendizado supervisionado: é aplicado quando um agente externo a rede sugere a resposta desejada para o padrão de entrada – é o caso aplicado neste estudo, cuja variável resposta (TP estendido ou não) da amostra de treinamento é aplicada na definição das variáveis preditoras discriminantes;
- b) aprendizado não supervisionado: quando não existe um agente externo indicando a

resposta desejada para os padrões de entrada; e,

- c) aprendizagem por reforço: a fonte de aprendizado é a própria experiência do agente, cujo objetivo é adquirir uma política de ações que maximize seu desempenho geral.

Além disso, para os N pares (entrada e saída) do conjunto de treinamento no processo de aprendizado, a correção dos pesos pode ser executada de dois modos:

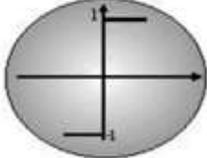
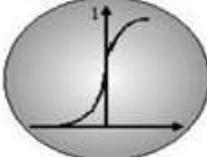
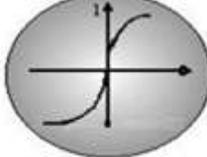
- a) aprendizagem por lote de padrões (*off-line* ou *batch*): os ajustes dos pesos que conectam cada camada só serão efetivados após a realização de todo o treinamento para uma época, pois a cada passo de ajuste leva em consideração o total de desvio observados nas amostras de treinamento frente aos valores desejados na camada de saída; e,
- b) aprendizagem padrão-por-padrão (*online*): ao contrário da estratégia anterior, neste caso o ajuste é realizado após a apresentação de cada amostra ao treinamento. Porém, para este processo de aprendizagem o ajuste é mais pontual e o comportamento a ser adquirido no treino e que serve de base à predição só se tornará representativa após a inserção de um número significativo de amostras.

Silva; Spatti e Flauzino (2016) ainda chama a atenção a necessidade de determinar a taxa de aprendizado adequadamente durante o processo de treinamento da rede neural. De modo que uma taxa de aprendizado muito baixa torna o processamento da rede muito lenta, ao passo que uma taxa de aprendizado muito alta provoca oscilações no treinamento e impede a convergência do processo de aprendizado ao resultado ótimo. Em geral, pesquisadores optam por valores entre 0,1 e 1,0.

3.4.4 Funções de ativação

Conforme já mencionado anteriormente, cada neurônio contribui para a estimação da saída. Nesse sentido, a função de ativação desempenha o papel de restringir a limites de amplitude de saída de um neurônio, ou seja, definir os intervalos aos quais a variável resposta poderá variar. Haykin (2003) expõe os seguintes intervalos $[0, 1]$ ou $[-1, 1]$ que variam de acordo com a função de ativação utilizada. As principais funções são apresentadas na figura 9 a seguir:

Figura 9 – Relação das funções de ativação para redes neurais artificiais

<p><i>Função limiar:</i> $f(x) = \begin{cases} 1 & \text{Se } x < k \\ 0 & \text{Se } x \geq k \end{cases}$ (10)</p>	
<p><i>Função logística:</i> $f(x) = \frac{1}{1+e^{(-g)}}$ (11) Onde, $g = \sum_{i=1}^p w_i \cdot x_i$</p>	
<p><i>Função hiperbólica:</i> $f(x) = \tanh(x)$ (12)</p>	

Fonte: Adaptado de Silva; Spatti e Flauzino (2016)

Dessa forma, como meio de comparação ao método de regressão logística, a função de ativação selecionada neste estudo segue a mesma configuração da regressão logística.

4 MÉTODOS

O desenvolvimento de um estudo de caráter científico deve ser conduzido mediante o conjunto de atividades sequenciais necessárias ao esboço da pesquisa. Para tanto, este estudo é arquitetado no decorrer de cinco etapas metodologicamente sistematizadas na figura 10.

Figura 10 – Procedimentos metodológico



Fonte: Autora

Tal como exposto, esta seção se compromete em subsidiar as atividades desenvolvidas no estudo, de modo que os procedimentos metodológicos de cada etapa vão sendo detalhadas no decorrer desta seção.

Na primeira etapa, a proposta da estimação do tempo de permanência hospitalar pós-operatório foi desenvolvida na delimitação do problema de pesquisa. Para tal, é exposto na próxima seção a estratégia de busca utilizadas nas bases de dados científicas; o qual proporcionou embasamento teórico necessário para a continuidade do estudo. A partir das três etapas seguintes (tratamento dos dados, formulação dos modelos e validação) são apresentados os resultados alcançados com a pesquisa, na etapa seguinte são então apresentados, com base nos resultados alcançados, a discussão e as considerações finais alcançadas com o desenvolvimento da pesquisa.

4.1 BASE DE DADOS: CRITÉRIOS DE SELEÇÃO DOS ARTIGOS

Para a abordagem desenvolvida nesta etapa do estudo, as pesquisas na literatura com a finalidade de embasamento teórico e conceitual foram realizadas em nove bases de dados bibliográficas de caráter pago e gratuito, disponibilizadas pela faculdade, são elas: Pub Med, Science Direct, SCOPUS, Proquest, IEEE Explore, EBSCO host, SAGE Journals, Scielo e Google Acadêmico.

O critério de seleção levou em consideração o idioma das respectivas pesquisas bibliográficas, o ano da publicação e o uso de palavras-chaves estratégicas no campo de busca.

No caso do idioma, por restrição do próprio pesquisador, foi limitado a publicações em

português, inglês e espanhol. Para a data de publicação, foram considerados periódicos desenvolvidos e divulgados na última década, com prioridade aos artigos de 2010 até os dias atuais; com ressalva de algumas publicações de caráter introdutório ao assunto pesquisado e que são considerados essenciais para descrever a evolução do tema no decorrer dos anos. Já no processo estratégico de busca, optou-se por uma etapa inicial desenvolvida com o auxílio do *software VOSviewer version 1.6.4* e a base de dados SCOPUS, onde foram identificadas as palavras-chaves relacionadas ao tema em estudo; em um segundo momento, estas palavras foram então aplicadas nas demais base de dados, e os artigos foram então selecionados conforme a manifestação destas palavras no título, resumo e palavras-chaves do próprio periódico.

Para tanto, as seguintes palavras-chaves foram então combinadas no mecanismo de busca: *length of stay, discharge, postoperative length of stay, hospitalization stay, hospital stay, length of hospitalization, prolonged length of stay, duration of stay, cardiac surgery, coronary surgery, postoperative cardiac length of stay, inpatient, coronary artery bypass grafting, cardiovascular risk factors, prediction models, length of stay risk score, EUROScore, tempo de permanência / tempo de permanência pós-operatório, doenças coronarianas e tempo de permanência prolongado.*

Em um segundo momento, também foram realizadas pesquisas no âmbito dos fatores de riscos pré-operatório cardiovascular, a partir das seguintes palavras-chaves: idade ou *age*, gênero ou *genre / sex*, índice de massa corporal ou *body mass index*, paradox weight, fração de ejeção ou *ejection fraction*, insuficiência / disfunção renal ou *failure / renal dysfunction*, hipertensão arterial pulmonar ou *pulmonary arterial hypertension*, diabetes melitus / melittus ou *diabetes mellitus*, hemoglicada / *glycated hemoglobina*, *fragility*, reoperação ou *reoperation*, cirurgia eletiva ou *elective surgery*, cirurgia de urgência ou *urgency surgery*, cirurgia de emergência ou *emergency surgery, clearance of creatinine.*

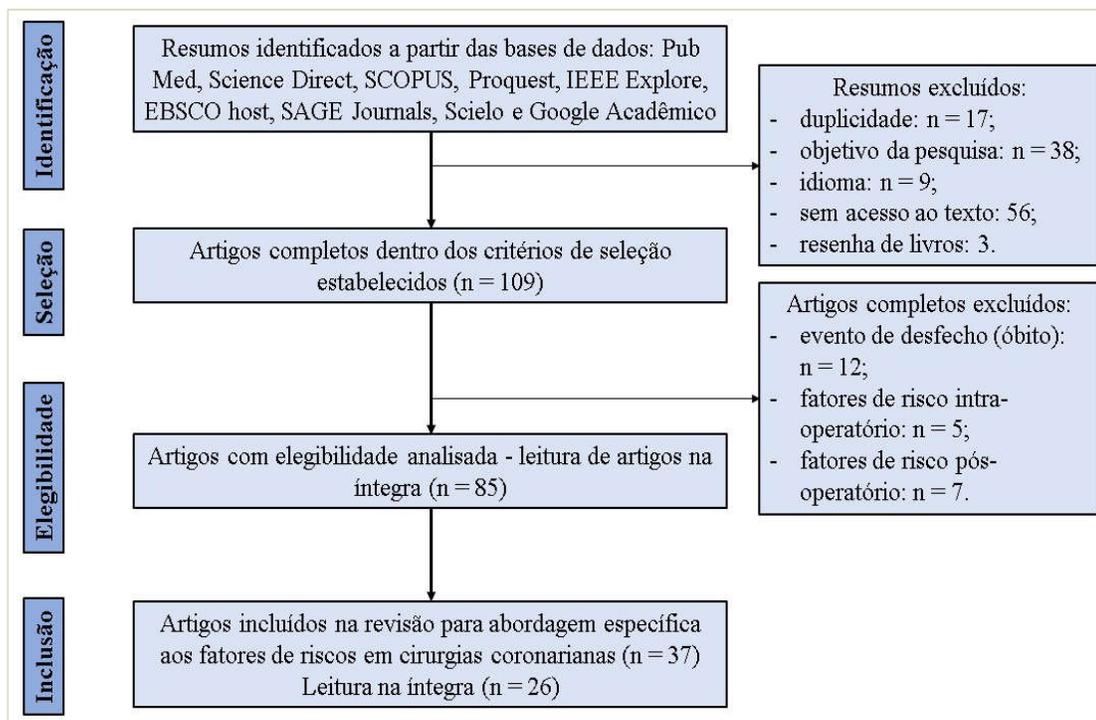
Sendo assim, a seguir são descritas as etapas da pesquisa bibliométrica, onde é dado destaque ao número de artigos selecionados em cada etapa e o número de artigos excluídos durante análise mais criteriosa, onde identificou-se o não atendimento ao tema proposto. Além disso, inicialmente a pesquisa foi restrita aos casos de cirurgia cardíaca, sendo posteriormente abrangido para pesquisas nas demais áreas cirúrgicas que tratam do assunto.

Na figura 11, é dado destaque inicialmente a base dada utilizada nesta pesquisa e nas etapas seguintes são revelados os artigos lidos, excluído e por qual motivo.

De modo geral, os periódicos são de delineamento experimental (ensaios clínicos, randomizados ou não) ou observacional (estudos de caso-controle e estudos de coorte), cujo

desfecho final é a alta com/sem acompanhamento até 90 dias posterior a liberação do paciente. Também foram realizadas algumas pesquisas em *sites* específicos para fins de desenvolvimento do tema, tais como: Ministério da Saúde – SUS, American Heart Association, World Health Organization, American Diabetes Association, Sociedade Brasileira de Diabetes, Sociedade Brasileira de Cardiologia.

Figura 11 – Fluxograma do processo de seleção dos artigos pesquisados



Fonte: Autora

4.2 AMOSTRA: TRATAMENTO DOS DADOS

Conjunto amostral terá como base um estudo de coorte prospectivo, coletado no banco de dados de um hospital público universitário brasileiro de alta complexidade, especializado em cardiologia, pneumologia e cirurgias cardíacas e torácicas.

Os dados foram coletados em registros do próprio hospital, cujas informações foram inicialmente coletadas pela própria equipe médica da instituição. O processo de seleção das variáveis explicativas levou em consideração os fatores de riscos mencionados na literatura, bem como as informações disponíveis no banco de dados do hospital em questão.

No geral, a amostra coletada apresenta 12 fatores de risco, no qual 9 deles foram contemplados neste estudo. Tal como expostos na seção 3.1, foram selecionados os seguintes

fatores de riscos pré-operatórios e as devidas abreviações são aplicadas neste estudo: idade (ID), gênero (GN), índice de massa corporal (IMC), fração de ejeção (FE), disfunção renal – *clearance* de *creatinina* (ClCr), diabetes mellitus – hemoglobina glicada / hemoglicada – (HbA1c), pressão arterial pulmonar (PAP), reoperações (REOP) e tipo de admissão – eletiva, urgência ou emergência – (TA).

Casella e Berger (2014), antes de iniciar o processo de modelagem, os autores argumentam a necessidade de transformar as variáveis de natureza qualitativa em quantitativas, a partir do processo de codificação ou categorização, denominadas *dummies*. É o caso das variáveis GN, TA, PAP e REOP que se tornaram variáveis quantitativas nominais neste estudo.

No âmbito de pesquisas da área médica, as variáveis ditas quantitativas contínuas são, em muitos casos, estratificadas em quantitativas nominais ou ordinais, a fim de mensurar e classificar os pacientes em diferentes grupos de risco. Tal categorização segue estudos e práticas consolidadas na área médica que podem contribuir na melhor adequação dos dados aos métodos de previsão. Incluem-se neste grupo as variáveis ID, IMC, FE, HbA1c e ClCr.

Em contrapartida, alguns pesquisadores como Babyak (2004), Sauerbrei e Royston (2010) e Harrel (2015) chamam a atenção para a perda de informações advindas do processo de categorização e reforçam a ideia de verificar se realmente há uma contribuição real desse procedimento na modelagem preditiva.

Logo, para validar o real impacto destas variáveis, é proposto neste estudo a investigação das variáveis tanto no seu formato nominal e ordinal, bem como em sua configuração de variável contínua, para os casos identificados.

A amostra é ainda analisada quanto a sua significância estatística, visto que o seu tamanho e as características observadas devem ser capazes de representar a população da qual a mesma foi selecionada. Logo, para a aplicação destes dados à regressão logística, Dias Filho e Corrar (2007) julgam necessário haver uma relação de ao menos 30 observações para cada variável analisada. Tal condição é então garantida no estudo de coorte realizado.

Ao julgamento do número de observações que deverá representar o aspecto de TP prolongado, é observado por Almashrafi et al. (2016b), a falta de uma definição clara na literatura quanto ao número aceitável de observações. Diante disso, os autores sugerem a predisposição de que menos de 75 percentis dos pacientes contenham o período de permanência superior a 10 dias, como garantia de observações com boa capacidade de avaliar as associações existentes entre as variáveis. Logo, esta condição deverá ser garantida pelo conjunto amostral selecionado neste estudo.

4.2.1 Critério de exclusão de dados

Para o conjunto amostral, a exclusão de dados é baseada nos seguintes critérios: óbito como evento resposta, dados duplicados, procedimento cirúrgico principal não corresponde à cirurgia de revascularização do miocárdio, data de cirurgia ou alta inserida incorretamente, mais de duas variáveis explicativas inseridas incorretamente para o mesmo paciente e a presença de dados caracterizados como *outliers*.

Para os cinco primeiros critérios, empregou-se o *software* PostgreSQL 9.6.1, cujo mapeamento e filtro dos dados permitiu aplicar os critérios de exclusão mencionados. Enquanto isso, para o último critério de exclusão: identificação de *outliers*, fez-se o uso da análise estatística do Box-Plot a partir ferramentas matemáticas disponíveis no Excel 2016 – versão 1611. Ao final, de 1.387 amostras coletadas, limitou-se o estudo a 1.215 pacientes.

Ao primeiro critério, Brock (2011) relatam que pacientes que vieram à óbito antes de atingir a estabilidade clínica ou receber a alta do hospital, impede o acompanhamento do real impacto de cada fator de risco ao respectivo período de permanência hospitalar, visto que o óbito não é o evento resposta de interesse. A esse comportamento é sugerido desconsiderar os dados de óbitos ou ainda é atribuído o aspecto de censura a estes dados, mais recomendado em casos onde são aplicados regressão de Cox.

No segundo critério são excluídas informações repetidas e, para o terceiro critério são desconsiderados pacientes, cujo procedimento cirúrgico principal não seja a CRM. Ao quarto e quinto critério é realizado a exclusão de variáveis incorretamente inseridas. Enquanto que no sexto critério, ferramentas estatísticas são aplicadas para a identificação e eliminação de *outliers* que, nas palavras de Freitas et al. (2012), refere-se a dados que não estão em conformidade com o comportamento geral da amostra.

4.3 PROPOSIÇÃO DOS MODELOS DE PREVISÃO

Nesta seção é apresentado as atividades desenvolvidas na modelagem matemática e no processo de diagnóstico dos modelos de previsão. Para tanto, é definido a caracterização aplicada neste estudo para a variável dependente, em seguida é formalizado o processo de modelagem da regressão logística e redes neurais artificiais, e finalmente é exposto os testes de diagnósticos a serem realizados para explorar o real impacto das variáveis explicativas ao TP pós-operatório, bem como a real capacidade preditiva dos modelos.

4.3.1 Caracterização da variável dependente

Na elaboração dos modelos de previsão é necessário, primeiramente, definir o tempo de hospitalização padrão para pacientes sem complicações pós-operatório, fornecendo subsídios para a caracterização do TP normal e prolongado.

Por conseguinte, a utilização dos modelos de regressão logística e RNA faz uso da variável dependente TP pós-operatória em instâncias dicotômicas: na primeira é definido o tempo de permanência normal (TPN) e na segunda instância o tempo de permanência prolongado (TPP). No caso da instituição de saúde de onde é extraído os dados, é definido a meta de alta hospitalar até 7 dias após a intervenção cirúrgica coronariana, com um *delay* de até 3 dias dentro do padrão de normalidade. Enquanto que, em períodos posteriores a 10 dias pode ser caracterizado por complicações no quadro clínico do paciente que o leva a permanecer por um período maior de internação, sendo definido neste estudo como o tempo de permanência prolongado.

Esse padrão de normalidade assumido ($TP \leq 10$ dias) age em consonância aos estudos desenvolvidos por Almashrafi et al. (2016a) e Oliveira et al. (2013) que definem esse período como limite essencial para a recuperação de pacientes submetidos à CRM sem elevar a taxa de readmissão hospitalar.

Aos modelos de previsão propostos neste estudo, é desenvolvida a seguinte premissa: tempo de permanência pós-operatória é estendido para hospitalização superior a 10 dias.

Para tanto, as seguintes nomenclaturas são usadas como um modo de testar esta proposição:

- a) 0: tempo de permanência normal ($TP \leq 10$);
- b) 1: tempo de permanência prolongado ($TP > 10$).

4.3.2 Construção do modelo de regressão logístico

Em extensão aos modelos de regressão linear, as etapas de desenvolvimento do modelo de regressão logística podem proceder pelo processo de eliminação *backward* ou seleção *forward*, sendo este último o mais amplamente utilizado (HAIR JÚNIOR et al., 2009). Os dois métodos fornecem uma maneira rápida e eficiente de definir o número de variáveis independentes; ao mesmo tempo que auxilia na escolha da equação mais conveniente, segundo critério da parcimônia a ser considerado.

Dessa forma, com a técnica bem difundida e considerando a facilidade da pesquisadora

é, então, selecionado o procedimento de regressão *stepwise* para o desenvolvimento deste estudo. Para isso, faz-se uso do *software* SPSS versão 23.0.0 no processo de modelagem.

Ao critério de decisão para inclusão ou exclusão de uma variável independente implica em testar o valor p da variável para um dado nível de significância previamente definido. Logo, quanto menor o valor p , mais estatisticamente significativa é a variável dentro do modelo.

Para esse procedimento, os seguintes passos são traçados:

PASSO 1: inicia-se a modelagem selecionando a variável explicativa de maior estatística de escore em relação à variável dependente. Esta é uma medida associativa usada na regressão logística para selecionar variáveis no procedimento *stepwise* a partir da redução do 2LL e, conseqüentemente, o aumento de sua medida associativa. Sob este aspecto, analisa-se o respectivo valor p para as seguintes situações:

- a) para valor $p < \alpha$ prossiga ao passo seguinte;
- b) para valor $p > \alpha$ procura-se a próxima variável de menor valor na função log-verosimilhança.

Em suma, a todos os modelos possíveis são verificados a combinação da variável independente às demais.

PASSO 2: são feitas comparações do módulo de variação da função log-verosimilhança sem e com o acréscimo de cada uma das variáveis explicativas aptas no passo anterior. No SPSS, esta avaliação leva ainda em consideração o valor do teste Qui-Quadrado ao nível de confiança e graus de liberdade definidos. Assim sendo, a decisão deve levar em consideração:

- a) para as variáveis explicativas com valor $p < \alpha$ e o módulo de variação inferior com hipótese inicial de que todos os pacientes permanecerão internados por no máximo de 10 dias, então segue-se para a variável seguinte e assim sucessivamente;
- b) o procedimento é encerrado se pelo menos uma das condições anteriores não forem satisfeitas.

Além disso, para validação do modelo costuma-se dividir a amostra em dois grupos: o primeiro é disponível para a modelagem e realização dos testes estatísticos de diagnósticos, enquanto que, no segundo grupo dispõe-se de dados para validar o modelo desenvolvido. Para tanto, é comumente empregada dois terços para treinamento e um terço na validação, ou ainda, três quartos para treinamento para um quarto na validação.

Neste estudo, após a realização de alguns testes preliminares, optou-se pela segunda opção, visto que no método Redes Neurais Artificiais há a necessidade de um número maior de dados para validar todas as possíveis combinações que o conjunto de variáveis selecionadas poderia proporcionar. De modo que, buscou-se realizar a modelagem com o mesmo número de

dados para ambos os modelos de previsão.

4.3.2.1 Ajuste e validação do modelo

Para o primeiro teste de ajuste do modelo, a performance deve ser percebida com a redução do -2LL no processo *stepwise forward*. Para isso, deve haver um decaimento do -2LL para cada variável incluída ao modelo, ao nível de significância de 5%.

Em contrapartida, a medida R^2 de Nagelkerke deve aumentar à medida que previsores são adicionados, garantindo que as variáveis inseridas no modelo sejam capazes de explicar grande variabilidade na ocorrência do evento resposta ou se será necessário incluir novas variáveis para garantir o melhor ajuste do modelo.

E, para validar o modelo e garantir que não haja vícios nos estimadores que possam comprometer os testes anteriores, é então aplicado método de reamostragem *bootstrapping*. Já o teste Wald é aplicado como complementariedade à medida anterior, com o intuito de avaliar o real impacto de cada variável explicativa no desempenho do modelo. Para isso, é avaliado o coeficiente destas variáveis, sob o intuito de rejeitar a hipótese nula de que estes coeficientes não são significativos ao nível de significância de 5%.

4.3.3 Construção do modelo de redes neurais artificiais

O modelo segue procedimento realizado na regressão logística – método *stepwise*, em que a cada interação uma nova variável é adicionada e seu impacto no modelo avaliado.

Para base de cálculos e treinamento da rede inicialmente era de intento fazer-se uso do *software* Matlab R2014. Contudo, por restrições de recursos materiais a modelagem da rede é então garantida pelos recursos oferecidos no SPSS 23.0.0 que, apesar de ser restrito aos tipos de arquiteturas disponíveis, apresenta uma cobertura satisfatória quanto a arquitetura *feedforward* multicamadas aplicadas neste estudo.

Em suma, para a modelagem da rede a seguinte configuração é definida: Rede *Feedforward Multicamadas* (Rede *Perceptron Multicamadas*), aprendizagem supervisionado com procedimento de ajuste dos pesos sinápticos por lote (*online*), algoritmo de otimização *backpropagation* e função de ativação logística.

Para avaliar a *performance* do modelo RNA, a amostra é dividida em duas partes: a primeira parte para treinamento da rede e a segunda para validação, numa proporção 75% - 25% respectivamente.

Diferentemente do método de regressão logística que realiza diferentes técnicas e testes estatísticas para validação do modelo, nas redes neurais artificiais esta validação é garantida principalmente pelo teste de especificidade e sensibilidade propostos na curva ROC e a mensuração da acurácia do modelo, cuja finalidade é avaliar a quantidade de acertos dos resultados preditos em relação aos resultados observados/desejados.

4.3.4 Análise de resíduos – Curva ROC

Etapa comum aos dois métodos de previsão, os resíduos indicam o quanto o modelo se ajusta a amostra de dados utilizada, dada a diferença entre o valor previsto e o valor desejado (real).

A Curva ROC (*Receiver Operating Characteristic*), nesse sentido, é uma ferramenta de avaliação da *performance* de modelos regressão de ordem binária (CABRAL, 2013), tal como ocorre na regressão logística e RNA para saídas binárias. É uma forma mais eficiente de demonstrar a relação antagônica entre a sensibilidade e a especificidade dos resultados.

Nesta situação, as medidas de sensibilidade e especificidade fornecem a probabilidade de que o teste classifique corretamente um indivíduo dentro do evento resposta esperado. A avaliação de eficiência do teste é estudar o efeito de diversos pontos de cortes em termos destas duas medidas para o desenvolvimento da curva ROC.

Seja R o valor correspondente a área abaixo da curva ROC, em geral os autores orientam-se:

- a) para $R = 0,5$ não há discriminação;
- b) para $0,5 < R < 0,7$ baixa discriminação;
- c) para $0,7 \leq R < 0,8$ a discriminação é aceitável;
- d) para $0,8 \leq R < 0,9$ há boa discriminação; e,
- e) para $R \geq 0,9$ discriminação excelente.

5 RESULTADOS

A partir da sistematização das etapas de pesquisas apresentadas na seção anterior, esta seção propõe inicialmente a apresentação de uma análise descritiva quanto a população investigada, bem como a investigação dos resultados obtidos na aplicação das ferramentas de modelagem, de modo a auxiliar na previsão da probabilidade de o paciente manter-se internado por um período prolongado após intervenção cirúrgica.

5.1 ANÁLISE DESCRITIVA DOS DADOS

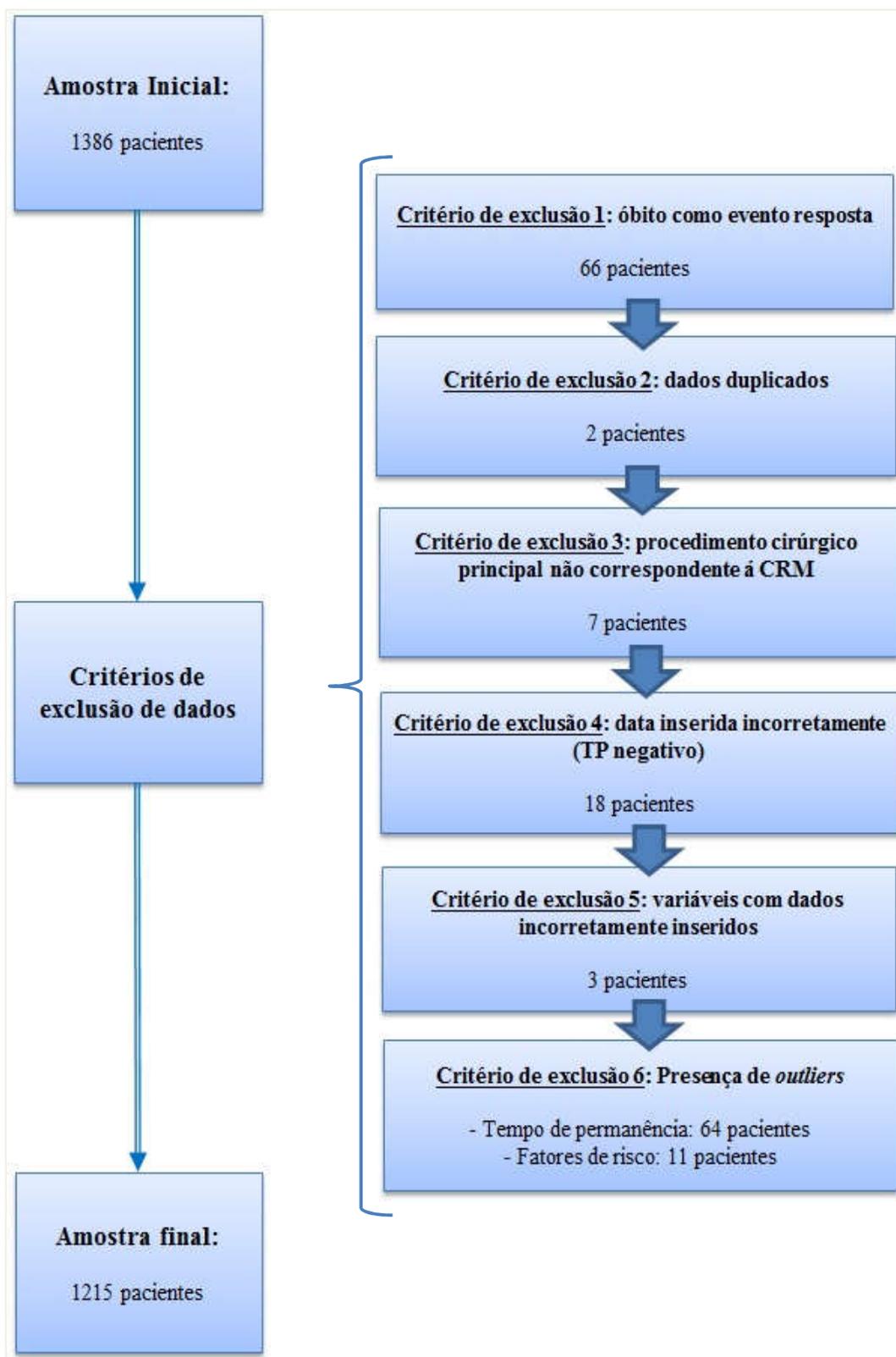
A análise descritiva inicia-se com o processo de eliminar vieses ou ruídos que possam comprometer a qualidade do banco de dados. No total, foram excluídas 171 amostras do coorte inicial de 1.386 pacientes.

No processo de tratamento dos dados, tal como segue na figura 12, seis critérios de exclusão foram definidos. Dentre eles, o critério com maior número de dados excludentes – Casos de *outliers*, seguida da exclusão de amostras tendo óbitos como desfecho.

Em geral, o conjunto amostral é caracterizado por cirurgias de revascularização do miocárdio ocorridas entre o período de novembro / 2013 a janeiro / 2016 e idade igual ou superior a 32 anos. Além disso, 993 pacientes (81,73%) realizaram o procedimento cirúrgico com circulação extracorpórea, 211 pacientes (17,37%) realizaram a colocação de *stents*, 142 pacientes (11,69%) sofreram enfarte ainda no período pré-operatório, 33 pacientes (2,72%) tiveram diagnóstico de fibrilação atrial, 22 pacientes (1,81%) garantiram o uso do balão intra-aórtico e 3 pacientes (0,25%) foram operados nas primeiras 48 horas – sob regime de admissão por emergência.

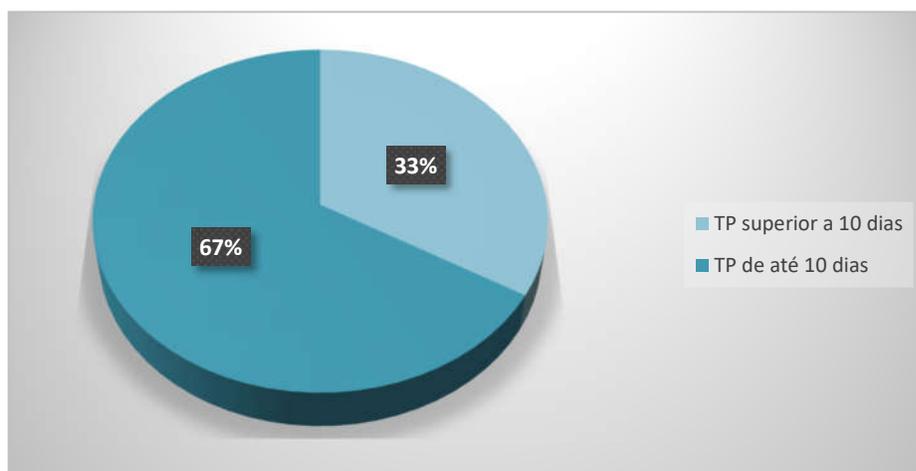
Para o particular estudo do TP pós-operatório, observa-se na figura 13 que 407 pacientes (33,50%) permaneceram internados por mais de 10 dias após a realização do procedimento cirúrgico e menos de dois terços dos pacientes permaneceram internados dentro do período padrão estimado pelo hospital.

Figura 12 – Critérios de exclusão de dados.



Fonte: Autora

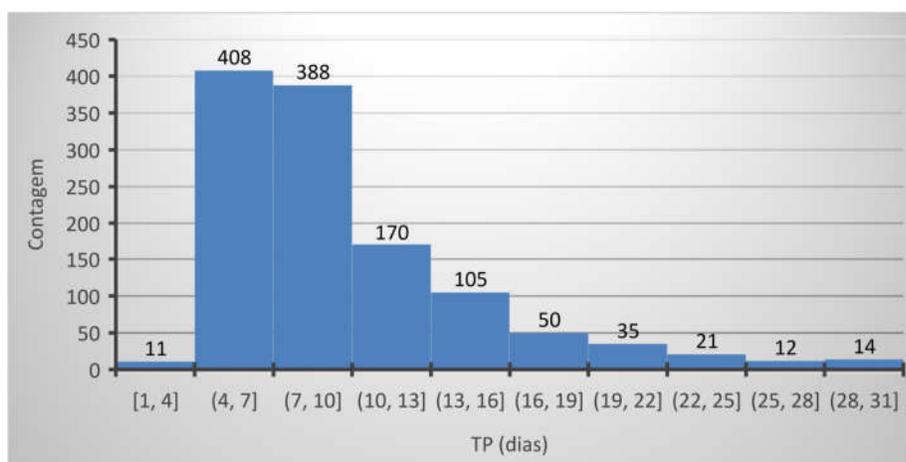
Figura 13 – Tempo de permanência pós-operatório estratificado



Fonte: Autora

No contexto da distribuição de frequência do TP pós-operatório (figura 14), confere a concentração de pacientes que se mantiveram internados entre 5 a 10 dias, cuja distribuição se configura assimétrica concentrada a esquerda.

Figura 14 – Distribuição de frequência do TP pós-operatório



Fonte: Autora

Influenciada por essa configuração e dispersa por uma amplitude de até 30 dias, a média do TP pós-operatório é estimada em 10,35 dias com desvio padrão de ± 5 dias. Contudo, em tal perspectiva a média encontra-se no limite superior do TP pós-operatório padrão, influenciada por valores extremos. Logo, julga-se necessário considerar a mediana como medida de tendência central, cujo valor estimado em 8 dias apresenta menor influência ao nível de

dispersão dos dados.

A efeito descritivo dos fatores de risco cardiovasculares, na tabela 1 as variáveis contínuas são expressas em mediana, média e desvio-padrão, respectivamente. Ao passo que na tabela 2 e 3 é dado destaque as variáveis categóricas (ordinais e nominais) manifestadas em frequência e porcentagem.

Tabela 1 – Fatores de riscos pré-operatórios - variáveis contínuas

Fatores de risco cardiovascular (variáveis contínuas)	Mediana	Média (\pm DP) ¹
Idade (ID - anos)	63,0	63,2 (\pm 9,14)
Índice de Massa Corporal (IMC - kg/m ²)	26,6	27,16 (\pm 4,14)
Fração de ejeção (FE - %) – feminino	60,0	57,86 (\pm 9,51)
Fração de ejeção (FE - %) – masculino	60,0	56,06 (\pm 10,57)
Clearance de creatinina (CLCR ¹ - mL/min)	68,0	69,89 (\pm 24,93)
Hemoglicada (HbA1c)	5,6	4,25 (\pm 3,43)

Fonte: Autora

¹: Desvio padrão

Para as variáveis contínuas independentes, percebe-se comportamentos diferentes daquele apresentado pela variável dependente, de modo que estas primeiras garantem uma maior proximidade nos valores aferidos pela média em relação a mediana, revelando assim uma distribuição com comportamento mais simétrico dos dados.

Uma segunda configuração das variáveis contínuas é ainda apresentada na tabela 2, onde estas variáveis são então categorizadas em grupos de riscos, conforme relação dos diferentes níveis à gravidade da doença. Tal prática de estratificação, segundo Fedorov; Mannino; Zhang (2009) é comum em pesquisas no campo da medicina que visa analisar a população por grupos de risco, propensos a diferentes níveis de agravamento da doença.

Tabela 2 – Variáveis categorizadas por grupo de risco

Continua

Fatores de risco cardiovascular (variáveis categorizadas)	Frequência	Porcentagem (%)
Idade (ID)		
1 (32 \leq ID \leq 56 anos)	279	23,0
2 (57 \leq ID \leq 62 anos)	281	23,1
3 (63 \leq ID \leq 69 anos)	349	28,7
4 (70 \leq ID \leq 87 anos)	306	25,2
Índice de Massa Corporal (IMC)		
1 (IMC \leq 18,5 kg/m ²)	8	0,7
2 (18,5 kg/m ² < IMC \leq 24,9 kg/m ²)	391	32,2
3 (25,0 kg/m ² \leq IMC \leq 29,9 kg/m ²)	554	45,6
4 (30,0 kg/m ² \leq IMC \leq 34,9 kg/m ²)	202	16,6
5 (35,0 kg/m ² \leq IMC \leq 39,9 kg/m ²)	52	4,3
6 (IMC \geq 40 kg/m ²)	8	0,7

Conclusão

Fatores de risco cardiovascular (variáveis categorizadas)	Frequência	Porcentagem (%)
Fração de ejeção (FE)		
1 (FE ≥ 55%)	967	79,6
2 (45% ≤ FE ≤ 54%)	88	7,2
3 (30% ≤ FE ≤ 44%)	144	11,9
4 (FE ≤ 30%)	16	1,3
Clearance de creatinina (CLCR)		
1 (CLCr ≥ 90 mL/min)	244	20,1
2 (89 ≥ CLCr ≥ 60 mL/min)	543	44,7
3 (59 ≥ CLCr ≥ 30 mL/min)	383	31,5
4 (29 ≥ CLCr > 15 mL/min)	20	1,6
5 (CLCr ≤ 15 mL/min)	25	2,1
Hemoglicada (HbA1c)		
1 (HbA1c < 5,7%)	609	50,1
2 (5,7% ≤ HbA1c ≤ 6,4%)	300	24,7
3 (HbA1c ≥ 6,5%)	306	25,2

Fonte: Autora

A decisão de categorizar foi baseada nos grupos de riscos já consolidados na literatura e que servem de referências na prática hospitalar, conforme exposto na seção 2. É dada, porém, exceção apenas à idade, visto que não há um consenso na literatura que define a faixa etária dentro do contexto deste estudo.

Tabela 3 – Fatores de riscos pré-operatórios - variáveis nominais e ordinais

Fatores de risco cardiovascular (variáveis ordinais e nominais)	Frequência	Porcentagem (%)
Gênero (GN)		
0 (masculino)	905	74,5
1 (feminino)	310	25,5
Reoperação (REOP)		
0 (não)	1.196	98,4
1 (sim)	19	1,6
Tipo de Admissão (TA)		
1 (eletiva)	570	46,9
2 (urgência e/ou emergência)	645	53,1

Fonte: Autora

Para as variáveis ordinais e nominais expostas na tabela 3, observa-se neste primeiro momento o predomínio de pacientes do sexo masculino (905 pacientes, 74,5%), baixa representatividade de casos com reoperação e um número superior de admissões por urgência e emergência em relação as admissões eletivas.

5.2 PERFIL BIVARIADO: EXAME DA RELAÇÃO ENTRE VARIÁVEIS

Nos primeiros passos da análise bivariada, é confrontado os fatores de risco

cardiovasculares ao TP pós-operatório – tabela 4, onde é avaliado a existência de correlação entre a variável dependente e as variáveis independentes. Para isso, a tabela de contingência abaixo define a frequência absoluta, o percentual para cada grupo do TP pós-operatório – em parêntese – e o teste Qui-Quadrado para validar a existência ou não de correlação; sob a hipótese nula de não haver nenhuma relação de dependência entre as variáveis ao nível de significância de 5%.

Tabela 4 – Tabela de contingência variáveis explicativas *versus* TP pós-operatório

VARIÁVEIS ESTRATIFICADAS		FREQUÊNCIA		SIGNIFICÂNCIA Valor p (Teste Qui Quadrado)
		TP ≤ 10 dias	TP > 10 dias	
ID	1	213 (26,4%)	66 (16,2%)	0,000
	2	206 (25,5%)	75 (18,4%)	
	3	218 (27,0%)	131 (32,2%)	
	4	171 (21,1%)	135 (33,2%)	
GN	0	631 (78,1%)	274 (67,3%)	0,000
	1	177 (21,9%)	133 (32,7%)	
IMC	1	6 (0,7%)	2 (0,5%)	0,644*
	2	262 (32,5%)	129 (31,7%)	
	3	368 (45,5%)	186 (45,7%)	
	4	134 (16,6%)	68 (16,7%)	
	5	33 (4,1%)	19 (4,7%)	
	6	5 (0,6%)	3 (0,7%)	
REOP	0	791 (97,9%)	405 (99,5%)	0,052*
	1	17 (2,1%)	2 (0,5%)	
FE	1	661 (81,8%)	306 (75,2%)	0,007
	2	59 (7,3%)	29 (7,1%)	
	3	81 (10,0%)	63 (15,5%)	
	4	7 (0,9%)	9 (2,2%)	
PAP	1	793 (98,1%)	389 (96,6%)	0,059*
	2	15 (1,9%)	18 (3,4%)	
ClCr	1	188 (23,3%)	56 (13,8%)	0,000
	2	Conclusão 385 (47,6%)	158 (38,8%)	
	3	215 (26,6%)	168 (41,3%)	
	4	8 (1,0%)	12 (2,9%)	
	5	12 (1,5%)	13 (3,2%)	
TA	1	414 (51,2%)	156 (38,3%)	0,000
	2	394 (48,8%)	251 (61,7%)	
HbA1c	1	416 (51,5%)	193 (47,4%)	0,046
	2	206 (25,5%)	94 (23,1%)	
	3	186 (23,0%)	120 (29,5%)	

Fonte: Autora

* Hipótese nula aceita

Ainda segundo informações expostas na tabela 4 juntamente com a exposição gráfica disponível no Apêndice A as variáveis ID, GN, FE, ClCr, TA, HbA1c comprovaram certa relação de dependência com a variável TP pós-operatório, cujo valor p manteve-se inferior ao

nível de significância de 5%. No caso particular das variáveis IMC, REOP e PAP a associação não foi encontrada e a frequência absoluta em cada estrato não apresentam resultados expressivos que diferencie o comportamento dessas variáveis nas duas categorias do TP pós-operatório.

Para a variável ID verifica-se um aumento expressivo do número de pacientes das classes 3 e 4 (acima de 63 anos e 70 anos, respectivamente) quando comparado as demais faixas etárias e que permaneceram internados por um período superior a 10 dias após a cirurgia. Já na variável GN, o destaque é dado ao aumento da prevalência do sexo feminino ao grupo de pacientes que se manteve internados por períodos prolongados.

No caso da variável FE, há um aumento de pacientes nos estratos 3 e 4 que permanecem internados por mais de 10 dias. Enquanto que na variável ClCr, nota-se o crescimento de internações por mais de 10 dias nos estratos 3, 4, e 5. E por fim, para a variável TA as admissões não eletivas proporcionam percentual superior no TP prolongado quando comparado as admissões eletivas.

Impulsionada por críticas de Babyak (2004), Sauerbrei e Royston (2010) e Harrel (2015) quanto a perda de informações no exercício da categorização de variáveis contínuas, é então proposto verificar a existência de associação entre variáveis explicativas contínua e o TP pós-operatório dicotomizado. Para isso aplica-se a correlação de Spearman como medida de análise estatística, ao nível de significância de 5%, conforme apresentado na tabela 5.

Tabela 5 – Correlação das variáveis contínuas ao TP pós-operatório nominal

Fatores de risco cardiovascular (variáveis contínuas)	Correlação	Significância estatística
Idade (ID_ - anos)	0,178	0,000
Índice de Massa Corporal (IMC_kg/m ²)	0,019	0,515*
Fração de ejeção (FE - %) – feminino	-0,175	0,009
Fração de ejeção (FE - %) – masculino	-0,149	0,007
Clearance de creatinina (ClCr - mL/min)	-1,181	0,000
Hemoglicada (HbA1c)	0,038	0,189*

Fonte: Autora.

* A correlação não é significativa no nível 0,05 (inferência bilateral).

Em geral, observa-se resultados similares quanto a significância estatística das variáveis contínuas em relação a sua configuração categorizada. E, em casos da existência de variáveis não significativas, estas são então desconsideradas nas etapas seguintes da modelagem.

Quanto ao aspecto de independência entre as variáveis explicativas, não foi evidenciado multicolineariedade de primeira ordem na análise de correlações.

5.3 ANÁLISE DA REGRESSÃO LOGÍSTICA COMO MÉTODO DE MODELAGEM

A modelagem por meio da Regressão Logística, inicia-se o processo de seleção de variáveis, seguindo pela avaliação da capacidade discriminatória, definição do modelo final e o diagnóstico de acurácia desse modelo, subsequentemente. Conforme apresentado na subseção a seguir.

5.3.1 Seleção das variáveis preditoras

A seleção e inclusão das variáveis explicativas ao modelo é realizado pelo método de inclusão *stepwise forward*, cuja significância é assegurada pelo teste da razão de verossimilhança. Contudo, dada a rigorosidade observada para seleção das variáveis ao nível de significância de 5%, incluir-se-ão no modelo as variáveis com valor p inferior a 0,15 e, a cargo de remoção, variáveis com valor p superior a 0,20 (HOSMER e LEMESHOW, 2013).

No programa SPSS 23.0 é inicialmente definido um modelo de base para efeito comparativo (denominado no SPSS de “bloco 0”). Esse modelo baseia na inserção da constante e da probabilidade a priori de que todos os pacientes permanecerão internados dentro do período padrão estimado pelo hospital. Sob tais condições o modelo base obteve 65% de acertos na estimativa inicial da modelagem.

Para as variáveis explicativas ainda não incluídas no estudo, o procedimento de inserção inicia-se com a avaliação da estatística de escore para cada variável, sob a hipótese nula de não haver associação entre as variáveis explicativas e a variável dependente (HAIR JÚNIOR et al., 2009). Nesse sentido, a tabela 6 apresenta um resumo geral das variáveis candidatas a entrar no modelo.

Para a estatística de escore disponível na tabela, a pontuação máxima adquirida ao nível de significância aceitável pode ser atribuída a variável categórica ClCr, cujo escore no primeiro estrato corresponde a 45,527 pontos e, por isso, deverá ser a primeira variável a ser incluída no modelo. Em contrapartida, as variáveis HbA1c e REOP não garantiram grau de significância em todos os seus estratos, a mesma situação é ainda percebida na variável HbA1c de medida contínua. No caso das variáveis que não atende aos requisitos do teste de significância serão desconsideradas para as próximas etapas da modelagem.

Seguindo o mesmo procedimento para as demais variáveis explicativas previamente selecionadas, a modelagem alcança seu melhor resultado na quinta interação, a partir da inclusão das variáveis ClCr, TA, GN, FE e ID respectivamente; garantindo assim um percentual

de acerto de 68,6%. Dentre estas variáveis, apenas FE foi mantida na configuração contínua, devido ao nível de significância satisfatória e a maior pontuação na estatística de escore.

Tabela 6 – Variáveis não presentes na equação

Variáveis escalares (contínuas)	Variáveis		Estatística de escore	Significância
		ID		38,379
	FE		6,856	0,009
	CICr		39,647	0,000
	HbA1c		1,728	0,256*
Variáveis categóricas (nominal e ordinal)	ID	1	36,023	0,000
		2	7,604	0,006
		3	6,784	0,049
		4	20,706	0,000
	GN	0	6,584	0,049
		1	16,526	0,000
	REOP	0	3,986	0,158*
		1	4,572	0,032
	CICr	1	45,527	0,000
		2	8,533	0,003
		3	26,98	0,000
		4	6,411	0,011
		5	3,922	0,048
	FE	1	8,064	0,045
		2	0,22	0,639*
		3	3,579	0,059
		4	4,103	0,043
	TA	1	5,483	0,028
		2	18,109	0,000
	HbA1c	1	6,014	0,204*
2		0,838	0,368*	
3		6,002	0,081	

Fonte: Autora

* Hipótese nula aceita

Para as variáveis incluídas no modelo, as seguintes análises podem ser obtidas na interpretação dos coeficientes (β_i), em termos de *odds ratio* (razão de chances ou razão de possibilidades):

Tabela 7 – Interpretação preliminar dos coeficientes para o modelo final.

Continua

Variáveis na equação	β_i	OR	95% C.I. para EXP (B)		Conclusões preliminares	
			Inferior	Superior		
CICr	1	-	-	-	Nos casos de pacientes que apresentam clearance de creatinina entre 15 e 29 mL/min é acometido quase 2,5 vezes de mais chances (OR) do TP pós-operatório ser prolongado. E para pacientes com clearance de creatinina inferior a 15 mL/min este fator é aumentado próximo de 3 vezes mais chances de permanecer internado por um período prolongado.	
	2	-0,035	0,966	0,634		1,472
	3	0,497	1,644	1,033		2,617
	4	0,090	2,482	0,794		7,753
	5	1,137	3,117	1,241		7,830

Conclusão

Variáveis na equação		B _i	OR	95% C.I. para EXP (B)		Conclusões preliminares
TA	2	0,450	1,568	1,174	2,094	Para pacientes admitidos à cirurgia por meio da admissão de urgência ou emergência, a chance de haver prolongamento no TP pós-operatório é aumentada em 1,6 vezes.
GN	1	0,464	1,590	1,151	2,196	Aos pacientes do sexo feminino, a chance de prolongar o TP pós-operatório é aumentada em 1,5 vezes.
FE	-	-0,180	0,980	0,969	0,996	O aumento do fluxo sanguíneo observado por meio da fração de ejeção reduz em 2% [(1-0,98)*100%] as chances do paciente permanecer internado por um período prolongado após a CRM.
ID	1	-	-	-	-	Em pacientes cuja idade é superior a 63 anos, a chance de permanecer internado por um período prolongado é aumentado em até 1,7 vezes (estrato 3 e 4)
	2	0,156	1,169	0,742	1,842	
	3	0,553	1,738	0,742	1,842	
	4	0,562	1,754	1,103	2,789	

Fonte: Autora

Sendo assim, a partir da estimação do intervalo de confiança para *odds ratio* na tabela 7 expõe informações complementares que colaboram na investigação das variáveis de impacto ao TP pós-operatório.

5.3.2 Diagnósticos do modelo de regressão logística

A etapa de diagnóstico inicia-se com a investigação do ajuste do modelo na incorporação das variáveis preditoras. Esse processo é iniciado com o teste de razão de verossimilhança, comumente denominado de -2LL. No procedimento é avaliada a capacidade de ajuste do modelo com a variável em relação ao modelo inicial, cuja análise estatística é firmada pelo teste Qui-quadrado, ao nível de significância estatística de 5%. Dessa forma, conforme modelagem realizada no SPSS 23.0, os seguintes resultados para esse exame são então definidos:

Tabela 8 – Diagnóstico -2 log likelihood value

Variável incluída	Passo <i>stepwise forward</i>	Teste Qui-quadrado	Significância
CICr	Passo 1	37,825	0,000
TA	Passo 2	49,332	0,000
GN	Passo 3	55,558	0,000
FE	Passo 4	60,439	0,000
ID	Passo 5	69,339	0,000

Fonte: Autora

Nesse primeiro momento, verifica-se que a inclusão das variáveis preditivas ao modelo garante significância estatística, e consequentemente rejeita a hipótese nula de que a diferença

de ajuste entre os modelos seja igual a zero. Logo, a relação de dependência anteriormente observada na análise bivariada é então confirmada nesta etapa do estudo. Na etapa posterior de diagnóstico é mensurada o quanto da variabilidade do modelo pode ser explicado por cada variável preditora selecionada. Neste caso, o R^2 de Nagelkerke é então aplicada, como exposto na tabela 9.

Tabela 9 – Diagnóstico R^2 de Nagelkerke

Variável incluída	Passo <i>stepwise forward</i>	R^2 de Nagelkerke
CICr	Passo 1	0,056
TA	Passo 2	0,073
GN	Passo 3	0,089
FE	Passo 4	0,101
ID	Passo 5	0,116

Fonte: Autora

Em geral, mesmo que as variáveis explicativas tenham demonstrado melhora no ajuste do modelo, o teste R^2 de Nagelkerke ajuda a identificar o quanto deste ajuste realmente esclarece a variabilidade do TP pós-operatório. Tal fato expõe a necessidade de investigar demais variáveis que contribuam no entendimento do modelo.

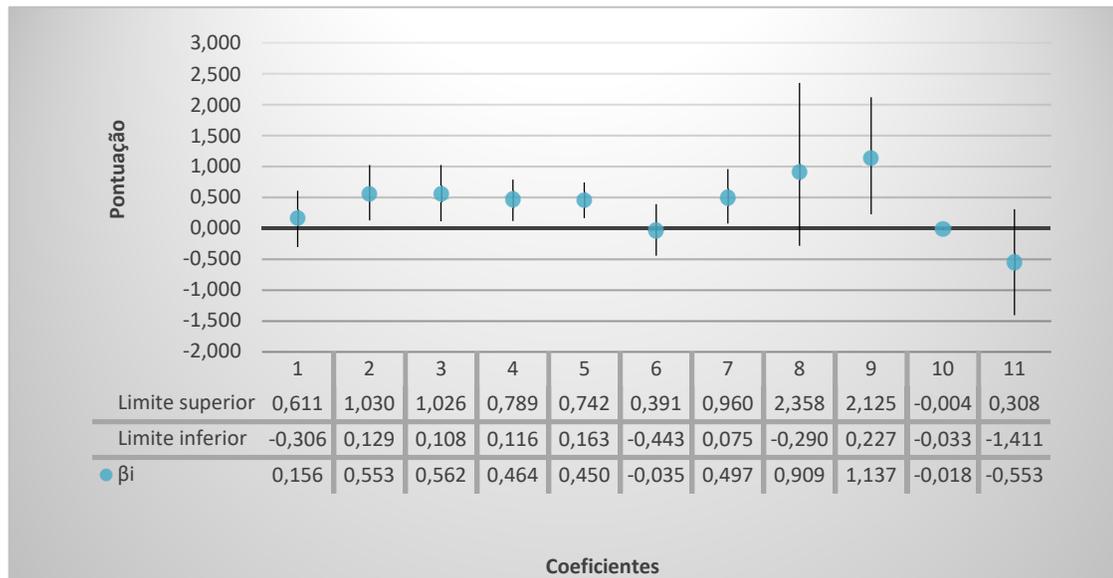
Para o teste de Bootstrapping, 1000 reamostragem aleatórias foram combinadas na tentativa de garantir coeficientes com boa capacidade preditiva e variabilidade próximo dos padrões de distribuição normal.

Logo, na figura 15 os resultados combinados do teste definem novos intervalos de confiança, no intuito de avaliar o quanto dos possíveis vícios na modelagem podem comprometer a qualidade dos coeficientes no ajuste geral do modelo.

No contexto geral, os coeficientes apresentaram significância estatística em relação aos intervalos de confiança estimados no método Bootstrapping, cujo vício no processo de modelagem não foi encontrado.

O teste Wald, por sua vez, finaliza a fase de diagnóstico da modelagem logística, cuja hipótese nula é rejeitada para a maior parte dos casos observados, dada a avaliação dos coeficientes como estatisticamente diferente de zero. É dado ressalvas, contudo, os estratos 2 e 4 da variável clearance de creatinina, estrato 2 da variável idade e a constante que não obtiveram a significância estatística necessária que justificasse manter os respectivos coeficientes no modelo final (tabela 10).

Figura 15 – Intervalo de confiança Bootstrapping



Fonte: Autora

Tabela 10 – Diagnóstico teste Wald

Variáveis na equação		B_i	Erro Padrão	Teste Wald	Significância	95% C.I. para EXP (B)	
						Inferior	Superior
CICr	1	-		16,265	0,002	-	-
	2	-0,035	0,215	0,026	0,871*	0,634	1,472
	3	0,497	0,237	4,392	0,036	1,033	2,617
	4	0,090	0,581	2,445	0,118*	0,794	7,753
	5	1,137	0,470	5,850	0,015	1,241	7,830
TA	2	0,450	0,148	9,279	0,002	1,174	2,094
GN	1	0,464	0,165	7,915	0,005	1,151	2,196
FE	-	-0,180	0,007	6,293	0,012	0,969	0,996
ID	1	-		8,824	0,032	-	-
	2	0,156	0,232	0,455	0,499*	0,742	1,842
	3	0,553	0,224	6,065	0,014	0,742	1,842
	4	0,562	0,237	5,643	0,018	1,103	2,789
Constante	-	-0,553	0,445	1,545	0,214*	-	-

Fonte: Autora

* Não rejeita a hipótese nula.

Após submetidos a todos os testes de diagnósticos e a exclusão das variáveis e estratos não significantes, o modelo logístico pode então ser expresso a partir da seguinte equação:

$$P_{TP} = \beta_{CICr3} * CICr_3 + \beta_5 * CICr_5 + \beta_{TA} * TA_1 + \beta_{GN} * GN_1 + \beta_{FE} * FE_1 + \beta_{ID3} * ID_3 + \beta_{ID4} * ID_4 \quad (8)$$

Onde, P_{TP} diz respeito ao evento de desfecho (tempo de permanência pós-operatório), em que β refere-se aos coeficientes β_i dispostos na tabela 10, enquanto x_i é expresso pela

nomenclatura das variáveis preditoras.

Para o objetivo final de prever o risco de o paciente permanecer internado por um período prolongado, a expressão (8) pode ainda ser escrita segundo a probabilidade de ocorrência do evento:

$$P(evento) = \frac{1}{1 + e^{-(PTP)}} \quad (9)$$

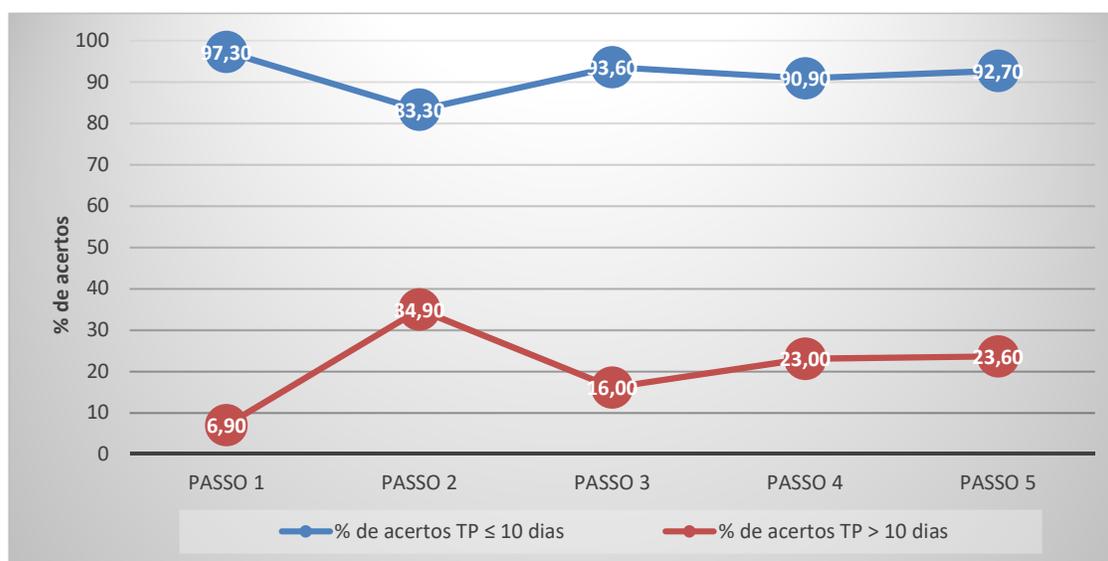
Logo, com base na equação (9), o nível de acurácia do modelo confirmado para a fase de treinamento é de 68,6% de acertos, enquanto que na etapa de validação houve uma melhora no desempenho do modelo.

Tabela 11 – Resultados da previsão – Regressão Logística

Amostra	Observado	TP ≤ 10 dias	TP >10 dias	Porcentagem correta
Treinamento	TP ≤ 10 dias	549	43	92,7%
	TP >10 dias	243	75	23,60%
	<u>Porcentagem global</u>	87,03%	12,97%	68,57%
Validação	TP ≤ 10 dias	582	10	91,10%
	TP >10 dias	259	59	15,40%
	<u>Porcentagem global</u>	92,42%	7,58%	70,44%

Fonte: Autora

Figura 16 – Percentual de acerto da predição para cada variável inserida ao modelo



Fonte: Autora

Os valores mostraram ser superior às estimativas do modelo base, porém o percentual

ainda ficou aquém do esperado para um modelo considerado de boa capacidade preditiva. Além disso, conforme observado na figura 16, a inclusão das variáveis a cada passo pouco melhorou nos resultados esperados, sendo também observado aspectos de superestimação do modelo para o TP pós-operatório de até 10 dias, visto que o percentual de acerto para esta situação é bem superior do que TP superior a 10 dias.

5.4 MODELAGEM DAS REDES NEURAIS ARTIFICIAIS

Para as RNA, a estrutura de desenvolvimento do método segue os conceitos desenvolvidos na seção 2, sendo então proposto a seguinte configuração de modelagem:

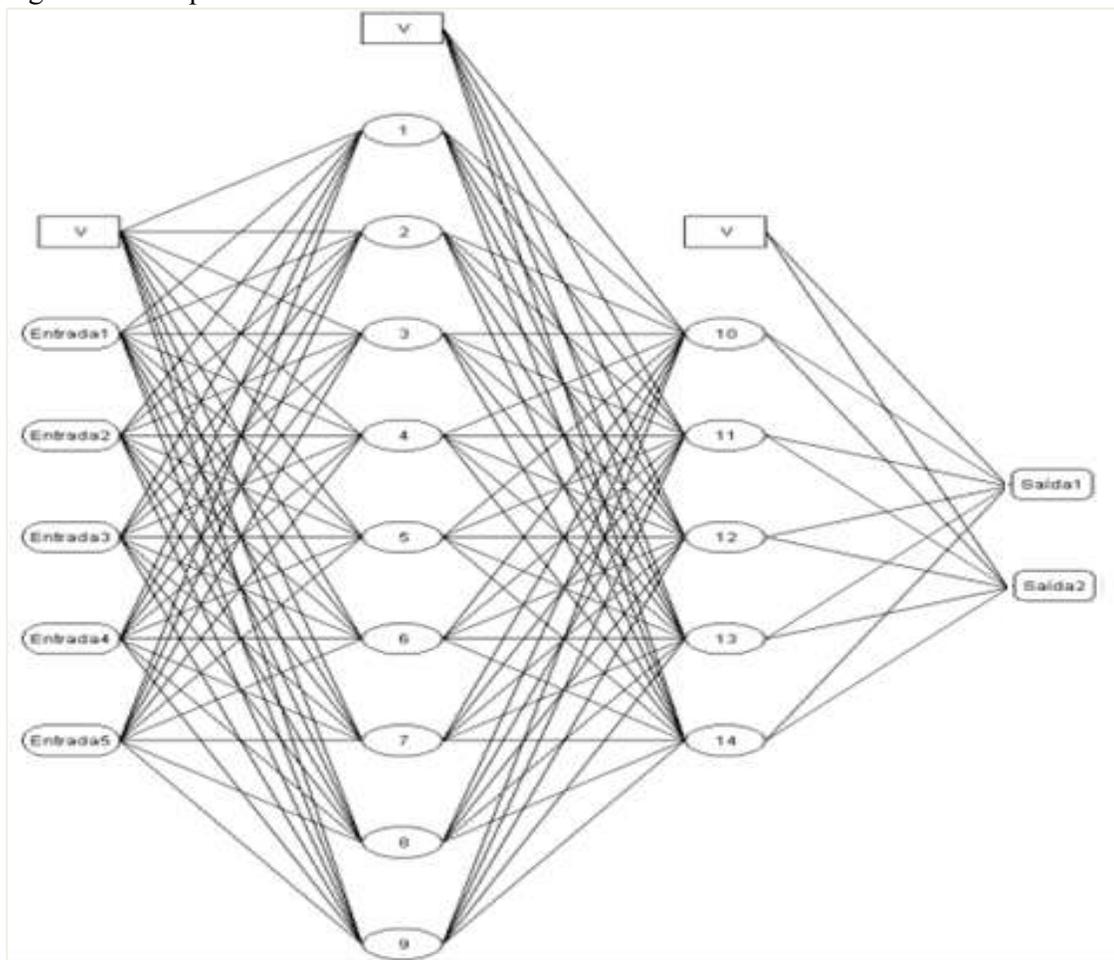
- c) arquitetura *feedforward* multicamadas, justificado pela ampla aplicação em diversos estudos clínicos e epidemiológicos;
- d) processo de aprendizagem supervisionado com método de correção dos pesos sinápticos em lotes, visto que o pesquisador definirá os parâmetros do modelo e o ajuste aos erros deverá acontecer apenas após a ocorrência de cada iteração (uma época);
- e) algoritmo de treinamento *backpropagation*;
- f) taxa de aprendizagem: 0,4;
- g) máximo de época: 10.000;
- h) número de camadas ocultas: 2;
- i) número de neurônios nas camada oculta: 10 e 6, respectivamente, incluindo os neurônio para o acúmulo do vieses;
- j) número de interações sem redução do erro;
- k) função de ativação: logística ou sigmoide; e,
- l) seleção das variáveis explicativas (input): método *stepwise forward*.

Para a definição tanto do número de neurônio da camada oculta como do número de camadas ocultas, primeiramente empregou-se os pressupostos da subseção 3.5.2, onde são apresentados alguns métodos de seleção. Dentre eles, o método que apresentou melhores resultados parciais refere-se ao pressuposto de que o número de neurônio seja igual ou inferior ao dobro de neurônios na camada de entrada. Porém para garantir a melhor configuração, outras redes com número de neurônios próximo da quantidade anteriormente definida são então testadas (método da ‘tentativa e erro’), valendo-se as medições tanto para uma como para duas camadas ocultas.

E por fim é definida a arquitetura da rede no SPSS 23.0, conforme figura 17, juntamente

com as respectivas variáveis explicativa selecionadas para a entrada do modelo. A arquitetura que apresentou uma melhor performance no SPSS apresenta nove neurônios na primeira camada oculta e 5 na segunda camada, um neurônio do viés para cada camada e dois neurônios na camada de saída, com o intuito de separar os valores preditos quanto aos dois possíveis valores desejados (dados reais).

Figura 17 – Arquitetura da Rede Neural Artificial



Fonte: Autora

As mesmas variáveis de entradas anteriormente selecionadas na regressão logística são agora avaliadas na performance das RNA. Para as demais variáveis pouco impacto foi identificado no processo de modelagem e, por isso, foram desconsideradas da amostra. Ao final, a rede previamente estabelecida e treinada alcança o total de 70,82 % de previsões corretamente estimadas.

Com base nesta configuração de rede, a seguir é apresentado na tabela 12 os resultados da previsão tanto para a etapa de treinamento como para a etapa de validação do modelo.

Tabela 12 – Resultados da previsão – Redes Neurais Artificiais

Amostra	Observado	TP ≤ 10 dias	TP >10 dias	Porcentagem correta
Treinamento	TP ≤ 10 dias	397	27	93,60%
	TP >10 dias	205	48	19,00%
	<u>Porcentagem global</u>	88,90%	11,10%	65,70%
Validação	TP ≤ 10 dias	155	15	91,10%
	TP >10 dias	53	10	15,40%
	<u>Porcentagem global</u>	89,27%	10,73%	70,82%

Fonte: Autora

5.5 AVALIAÇÃO DO PODER DE DISCRIMINAÇÃO DOS MODELOS

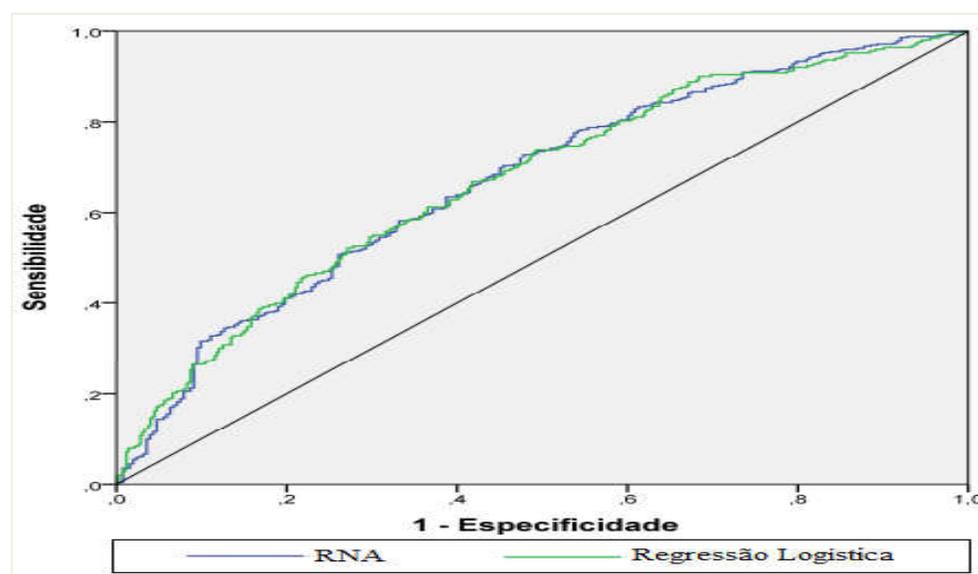
Após a modelagem, a etapa de validação apenas é finalizada com a análise de resíduos e como estes afetam o poder de discriminação dos modelos de previsão. Para isso, é sugerido o uso da curva ROC para avaliar a sensibilidade e a especificidade dos modelos, tal como apresentado na tabela 13 e figura 18 a seguir:

Tabela 13 – Área abaixo da curva ROC

MODELOS	ÁREA
Regressão Logística	66,5
Redes Neurais Artificiais	67,0

Fonte: Autora

Figura 18 – Curva ROC para os modelos de previsão



Fonte: Autora

Dessa forma, dado os fatores de risco pré-operatórios selecionados foi admitida baixa discriminação dos modelos para a previsão de probabilidades do tempo de permanência pós-operatório superior a 10 dias, em pacientes submetidos à cirurgia de revascularização do miocárdio.

6 DISCUSSÃO

Na composição descritiva geral dos dados, sejam estes configurados como contínuos ou categóricos, observa-se que pacientes acima de 63 anos e do gênero masculino predominam entre os pacientes submetidos a cirurgia de revascularização do miocárdio. Para os índices de massa corporal é registrado leve sobrepeso da população estudada; o que de fato, para esta situação em particular não representa um agravante ao estado clínico pós-cirúrgico (vide seção 3.1.3).

No que diz respeito a fração de ejeção, em média, é predominante pacientes dentro dos padrões normais ($FE \geq 55\%$) de referência. E, no caso da variável clearance de creatinina há prevalência de pacientes nos estratos 2 e 3 do grupo de risco; esses resultados mostram certa preocupação quanto a redução ao índice de filtração glomerular que cometera parte da população investigada, podendo causar complicações pós-operatório e o desenvolvimento de doença renal crônica. No que diz respeito ao fator de risco diabetes mellitus, analisada por meio da hemoglobina glicolisada, verificou-se que em média os pacientes mantêm níveis de glicemia no sangue dentro dos padrões normalidade.

Em termos gerais, as variáveis idade, gênero, fração de ejeção, clearance de creatinina e o tipo de admissão estiveram presentes na maioria dos estudos analisados, assim como observado por Almashrafi et al. (2016b), Shinjo e Fushimi (2015) Azarfarin et al. (2014), Khairudin; Mohd e Hamid (2012) e Oliveira et al (2013). Ainda vale ressaltar que nem todos os fatores de risco pré-operatórios foram abordados neste estudo e, por isso, os resultados alcançados são considerados reflexos deste posicionamento.

Outro posicionamento definido neste estudo, refere-se ao delineamento do tempo de permanência pós-operatório, assumindo o padrão de normalidade de até dez dias de hospitalização. Essa definição segue o comportamento observado por Almashrafi et al. (2016b), Azarfarin et al. (2014), Khairudin et al. (2012); bem como refere-se ao atual padrão definido pela instituição de saúde, de onde foram coletados os dados.

Contudo, em estudos como de Peterson et al (2002), Shahian et al (2009) e Khairudin (2011) o tempo de permanência pós-operatório é abordado sobre a premissa de que são necessários quatorze dias de internação para recuperação completa do paciente pós-cirúrgico. Os autores argumentam que altas precoces poderão elevar o indicador de readmissões hospitalares ou até mesmo elevar a taxa de óbito nos próximos 90 dias após a alta.

No que diz respeito ao processo de validação dos modelos de previsão, com base nos trabalhos dispostos na literatura, a regressão logística binária mostra-se predominante entre os

estudos, garantindo classificação dos fatores de risco de impacto e a estimação do tempo de permanência pós-operatório.

Os fatores de risco, intitulados de variáveis preditoras no processo de modelagem, em sua maioria foram selecionadas por meio do método *stepwise forward*. Enquanto que para os testes de diagnóstico prevalece o uso dos métodos estatísticos tradicionais: -2LL, Nagelkerke, Hosmer e Lemeshow e Wald. Apenas ALMASHRAFI et al. (2016b) garantiu o uso de métodos Bootstrapping para seleção das variáveis e diagnóstico da performance do modelo. Enquanto isso, para as redes neurais artificiais, a arquitetura *feedforward* multicamadas com algoritmo *backpropagation* foram dominantes, cujas técnicas estatísticas de diagnósticos se restringiram ao uso da curva ROC e métodos de mensuração da acurácia.

Ainda sob a perspectiva dos métodos de previsão, Azarfarin et al. (2014) validaram o uso do modelo de regressão logística binária para a seleção das variáveis relacionadas ao tempo de permanência na UTI pós-operatória, garantindo um total de acertos de 72% das probabilidades previstas. Já Pagowska-Klimek et al (2011) fizeram uso da regressão logística múltipla, garantindo 76% de acertos para TP padrão de onze dias, em um hospital especializado em cirurgias cardíacas infantil. Em outro caso, Almashrafi et al. (2016b) fizeram uso da regressão logística múltipla com sete variáveis de entrada, cuja performance mensurada a partir da curva ROC garantiu uma área abaixo da curva no valor de 70,5%.

Já na validação de modelos baseados nas RNA, LaFaro et al. (2015) aplicaram a arquitetura *feedforward* multicamadas para prever o tempo de permanência na UTI, mas o modelo não obteve boa capacidade preditiva, de modo que área abaixo da curva ROC foi inferior a 60%. Como trabalhos futuros os autores recomendaram a inclusão de novas variáveis que melhore o poder preditivo. Houthoof et al. (2015), por sua vez, modelaram uma rede neural Perceptron Multicamadas com 32 variáveis de entrada, cujos resultados garantiram boa acurácia e uma área abaixo da curva ROC de 70,3%.

No âmbito desta pesquisa, as RNAs apresentaram uma leve superioridade quando comparado a Regressão Logística, no que diz respeito a avaliação da acurácia dos modelos (70,82% *versus* 70,44% de acertos, respectivamente) e no poder de discriminação (67% *versus* 66,5% da área abaixo da curva ROC, respectivamente). Tal fato pode ser justificado pela capacidade das RNA em lidar com comportamentos não lineares dos dados.

Porém, em geral, ambos os métodos de previsão não desempenharam boa capacidade preditiva, visto que os resultados não garantiram a área abaixo da curva ROC igual ou superior a 70%. Além disso, conforme resultados obtidos com os testes de diagnósticos *Log-Likelihood* e R^2 de Nagelkerke, apesar das variáveis pré-operatórias selecionadas apresentarem associação

ao TP pós-operatório, a capacidade de discriminação destas variáveis pouco explicam da variabilidade presente no comportamento dos dados. Torna-se, portanto, desejável realizar novos estudos que avalie o impacto de outros fatores de risco na previsão do tempo de permanência pós-operatório.

Sob este aspecto, Almashrafi et al. (2016a) argumentam que além dos fatores cardiovasculares, previamente selecionados para este estudo é necessário avaliar aspectos clínicos como possíveis preditores; é o caso de pacientes com *stents*, balão intra-aórtico, marca-passo que podem apresentar complicações após a realização da cirurgia. Pagowska-Klimek et al (2011) ainda sugerem a inserção na investigação dos fatores associados exclusivamente ao estilo de vida do paciente, tais como o uso de cigarros, frequência das atividades físicas, consumo de bebidas alcóolicas, entre outros.

Shinjo e Fushimi (2015), por sua vez, vai um pouco além e sugere investigar o tamanho do hospital e sua localização como possíveis preditores, com base em estudos epidemiológicos que identificaram diferenças significativas entre as regiões do estudo. Em outro estudo, Lacasia-Melero (2016) chama a atenção ao conceito de fragilidade e sugere a prática deste fator em substituição a idade, visto que este último pode não referenciar corretamente o real estado físico e clínico do paciente.

É ainda sugerido por Najafi e Goodarzynejad (2012) e Oliveira et al. (2013) a investigação do TP pós-operatório separado entre UTI e enfermagem. Neste sentido, Azarfarin et al. (2014) avalia fatores de risco intra e pós-operatório para a análise do tempo de permanência na UTI pós-operatória, nas primeiras 96 horas. Os resultados mostraram-se satisfatórios quanto a força preditiva das variáveis selecionadas. Para isso, foram coletados dados dos fatores de risco intra-operatório, com destaque ao uso de inotrópicos, tempo de anestesia e prática de circulação extra-corpórea; além dos fatores de risco pós-operatório: infecções, hemodiálises e o período de intubação.

Além dos fatores de risco cardiovasculares, Azarfarin et al. (2014) ainda chamam a atenção ao impacto de fatores de risco ‘não médicos’, tal como foram denominados pelos autores; sendo estes, em muitos casos considerados fatores limitantes a correta previsão do TP pós-operatório. Enquadram-se a este grupo a dificuldade de autorização de transferência da UTI para a enfermagem ou mesmo de alta durante os finais de semana e feriados, a permanência de pacientes na UTI em casos de desequilíbrio da fibrilação atrial no pós-operatório – podendo ser tratado este paciente na enfermagem; ou ainda, a ausência de leitos e equipamentos de transporte dos pacientes.

De modo geral, mesmo não garantindo modelos de boa acuracidade e discriminação

para a previsão de probabilidade do tempo de permanência pós-operatório, há vários aspectos que vão além dos contornos definidos para este estudo e que podem gerar resultados promissores.

7 CONSIDERAÇÕES FINAIS

A proposição de métodos preditivos em cirurgias cardíacas, especialmente na cirurgia de revascularização do miocárdio, veio em resposta a necessidade de se antever às situações críticas e proporcionar aos profissionais melhores ferramentas de suporte às tomadas de decisões.

A utilização de recursos após procedimentos cirúrgicos invasivos, podem aumentar dramaticamente quando surgem complicações no período subsequente à cirurgia. Nesse sentido, o tempo de permanência é um índice de recuperação pós-operatória, atuando como medida do impacto cumulativo da cirurgia, morbidade pré-operatória e complicações pós-operatórias.

Dessa forma, manter o conhecimento do TP sob a pauta dos fatores de risco promove o desenvolvimento de vias clínicas eficientes e otimiza o gerenciamento de recursos (ROWAN et al., 2011). Em suma, os fatores de risco cardiovasculares são de relevância clínica na definição de prioridades cirúrgicas que neutralize o risco de complicações e comprometa a recuperação do paciente no período pós-operatório.

Vários métodos de previsão e com diferentes abordagens estatísticas foram propostos nas literaturas, sendo implementados em sua maioria nos ambientes de cuidados intensivos. Muitos desses métodos aplicam a informação pré-operatória como principal impacto ao desfecho clínico, sejam elas de cunho demográfico, clínico e fisiológicas. Além de fornecer uma medida geral da gravidade da doença e como esta poderá afetar a recuperação do paciente (ROWAN et al., 2011; ALMASHRAFI et al., 2016b).

Nesse sentido, a seguinte problemática foi levantada neste estudo: Com base nos fatores de riscos, os métodos de previsão “Regressão Logística” e “Redes Neurais Artificiais” apresentam bom desempenho na previsão de probabilidade do tempo de permanência pós-operatório?

Como resposta ao questionamento levantado, os fatores de risco selecionados para este estudo foram insuficientes para explicar grande parte da variabilidade do tempo de permanência pós-operatório. Logo, os modelos de previsão propostos não apresentaram boa performance preditiva.

Contudo, há uma ampla dimensão de fatores de risco associados a CRM e que não foram indagadas nos contornos deste estudo. Tal fato, demonstra a necessidade de dar continuidade ao estudo em trabalhos futuros, de modo a identificar as variáveis mensuráveis de grande impacto a temática abordada e como o comportamento destas variáveis são percebidas nos

métodos de previsão para garantir maior acurácia nos resultados.

Vale ainda destacar que o resultado obtido deve, em grande parte, aos contornos defendidos neste estudo, de modo que diversas formas de tratamento dos dados, proposição de outras técnicas de previsão ou mesmo novas definições de parâmetro para os atuais métodos analisados podem proporcionar diferentes resultados.

Por conseguinte, para um modelo de previsão selecionado que atenda as especificações estatísticas e profissionais, deve-se ainda levar em consideração a necessidade de se realizar atualizações periódicas, devido a alterações no ambiente interno e/ou externo ao paciente e que pode afetar terminantemente a acurácia dos modelos de previsão.

7.1 LIMITAÇÕES DO ESTUDO

Este estudo apresenta algumas limitações que pode ter afetado os resultados e, por isso, merecem uma discussão. Primeiramente, os modelos de predição não foram validados externamente. Assim, o modelo carece de generalização que limita sua portabilidade a outros contextos. Ao mesmo tempo que um modelo universal pode apresentar barreiras durante o seu desenvolvimento, visto que além dos fatores de riscos cardiovasculares, o TP depende de políticas institucionais internas que influenciam muito nas práticas de alta.

Em segundo lugar, o estudo incluiu uma amostra de tamanho pequeno quando comparada a outros trabalhos publicados na literatura.

Terceiro, é reconhecido a existência de fatores de risco não clínicos que podem ter influenciado o desfecho de permanência do paciente no hospital e que não foram incluídos neste estudo, devido à indisponibilidade de dados. Para tais fatores incluem os julgamentos dos médicos, a política hospitalar de autorização de transferência ou alta e as considerações de demanda e capacidade.

Uma quarta limitação refere-se aos aspectos restritivos dos modelos de previsão selecionados. Em ambos os casos, pode haver a ocorrência de *overfitting* e a aplicação das técnicas apenas é recomendada quando o pesquisador já tenha identificado todas as variáveis independentes relevantes. Em particular RNAs, é caracterizado pela dificuldade de parametrização da rede quando não aplicado um algoritmo de otimização, devido a ampla dimensão de combinações; além da presença de mínimos locais que impede a modelagem de atingir o erro mínimo global.

Finalmente, a seleção do percentil 75 para separar os dados entre treinamento e validação pode ser vista como um corte arbitrário na ausência de um valor clinicamente

aceitável e validado na literatura.

7.2 SUGESTÕES PARA TRABALHOS FUTUROS

Diante da diversidade de modelos de previsão aplicados ao tempo de permanência pós-operatório e a ampla dimensão dos fatores de risco que podem impactar no período de hospitalização, é então sugerido um estudo amplo na identificação das variáveis de real impacto ao TP e na aplicação dos demais métodos de previsão. Sendo dado destaque aos métodos de regressão e de agrupamento, devido a facilidade de realizar comparações entre as metodologias propostas.

Além disso, para o posicionamento de validação externa e a possibilidade de generalização do modelo, é proposto o desenvolvimento de estudos multicêntricos em pesquisas futuras.

É ainda indicado avaliar a associação dos fatores ‘não-médicos’ ao TP pós-operatório e como é possível minimizar os seus impactos, tal como observado por Azarfarin et al. (2014), visto que estes fatores influenciam na real data de alta. Em alguns casos, a determinação da alta independe de o paciente estar apto ou não de recebe-la, mas sim da disponibilidade de recursos humanos e materiais para a devida autorização.

Em outra linha de pesquisa, recomenda-se avaliar o impacto dos escores de risco de morte ao tempo de permanência pós-operatório, visto que alguns autores como Najafi; Sheikhvatan e Sheikhfathollahi (2014) abordam essa relação.

Propõe-se também o desenvolvimento de escores de risco aditivos, assim como modelado nos escores de risco de mortalidade, para a estimação do tempo de permanência pós-operatória. De modo a garantir resultados a partir de um método de fácil mensuração e sem a necessidade de investimento em recursos materiais. E, por fim, é sugerido a prática deste estudo tempo de permanência no âmbito de outras especialidades cirúrgicas.

REFERÊNCIAS

ADJEI, I. A.; KARIM, R. An application of Bootstrapping in Logistic Regression Model.

Library Journal, v. 3, n. 9, p. 1-9, Set. 2016. Disponível em:

<http://file.scirp.org/pdf/OALibJ_2016092817220271.pdf>. Acesso em: 16 jul. 2016.

ALBORNOS, V. et al. Uso de cadenas de Markov para la predicción de la dinámica del comportamiento de pacientes en una unidad de cuidado intensivo cardiológica.

Revista Chilena de Ingeniería, v.14, n.2, p.153-158, Jun. 2006. Disponível em:

<<http://www.scielo.cl/pdf/ingeniare/v14n2/art09.pdf>>. Acesso em: 16 jul. 2016.

ALMASHRAFI, A. et al. Systematic review of factors influencing length of stay in ICU after adult cardiac surgery. **BioMed Central and Health Services Research**, v. 16, Abr. 2016a, p. 318-326. Disponível em:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4966741/pdf/12913_2016_Article_1591.pdf>. Acesso em: 16 jul. 2016.

ALMASHRAFI, A. et al. Factors associated with prolonged length of stay following cardiac surgery in a major referral hospital in Oman: a retrospective observational study. **British Medical Journal**, v. 6, p. 764-774, Jul. 2016b. Disponível em:

<<http://bmjopen.bmj.com/content/bmjopen/6/6/e010764.full.pdf>>. Acesso em: 16 jul. 2016.

ALTMAN, D. G. et al. Prognosis and prognostic research: validating a prognostic model.

British Medical Journal, v. 338, p 1-11, Maio 2009. Disponível em:

<<http://www.bmj.com/content/338/bmj.b605>>. Acesso em: 16 jul. 2016.

AMERICAN DIABETES ASSOCIATION. Economic costs of diabetes in the USA in 2012.

Diabetes Care, v. 36, p.1033 – 1046, 2013. Disponível em:

<<http://care.diabetesjournals.org/content/36/4/1033.long>>. Acesso em: 16 jul. 2016.

AMERICAN HEART ASSOCIATION. **Ejection fraction heart failure measurement**, 2015. Disponível em:

<http://www.heart.org/HEARTORG/Conditions/HeartFailure/SymptomsDiagnosisofHeartFailure/Ejection-Fraction-Heart-Failure-measurement_UCM_306339_Article.jsp#.V_LWMfkrLZ4>. Acesso em: 18 jul. 2016.

ANDRADE, J. P. et al. Sociedade Brasileira de Cardiologia: carta do Rio de Janeiro. **Arquivo Brasileiro Cardiologia**, v. 100, n. 1, p. 3-5, Jan. 2013. Disponível em:

<<http://www.scielo.br/pdf/abc/v100n1/v100n1a02.pdf>>. Acesso em: 18 jul. 2016.

AO, H. et al. The impact of body mass index on short- and long-term outcomes in patients undergoing Coronary Artery Graft Bypass. **PLOS ONE**, v. 9, n. 4, p. 1-6, Abr. 2014.

Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0095223>>.

Acesso em: 18 jul. 2016.

AWAD, A.; BADER-EL-DEN, M.; MCNICHOLAS, J. Modeling and predicting patient length of stay: a survey. **International Journal of Advanced Scientific Research and Management**, v. 1, n. 8, p. 90 – 102, Ago. 2016. Disponível em: <http://ijasrm.com/wp-content/uploads/2016/08/IJASRM_V1S8_100_90_101.pdf>. Acesso em: 03 ago. 2016.

AZARFARIN, R. et al. Factors influencing prolonged ICU stay after open heart surgery. **Research in Cardiovascular Medicine**, v. 3, n. 4, p.1-5, Out. 2014. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4347792/>>. Acesso em: 03 ago. 2016.

BANU, M. A. N.; GOMATHY, B. Disease forecasting system using data mining methods. **International Conference on Intelligent Computing**, v. 5, n. 1, p. 41-45, Nov. 2014. Disponível em: <<http://ieeexplore.ieee.org/document/6965026/>>. Acesso em: 03 ago. 2016.

BARBINI, P. et al. A straightforward approach to designing a scoring system for predicting length-of-stay of cardiac surgery patients. **BioMed Central Medical Informatics and Decision Making**, v. 14, n. 89, p. 1-8, Out. 2014. Disponível em: <<https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-14-89>>. Acesso em: 03 ago. 2016.

BARBOSA, R. R.; et al. Impact of renal failure on in-hospital outcomes after Coronary Artery Bypass Surgery. **Arquivo Brasileiro de Cardiologia**, v. 97, n. 3, p. 249-253, Jun. 2011. Disponível em: <http://www.scielo.br/pdf/abc/v97n3/en_aop06711.pdf>. Acesso em: 03 ago. 2016.

BARRETT-CONNOR, E. Sex differences in coronary heart disease: why are women so superior?. **Circulation**, v. 95, n. 1, p. 1-11, Jan. 1997. Disponível em: <<http://circ.ahajournals.org/content/95/1/252.long>>. Acesso em: 03 ago. 2016.

BABYAK, M. A. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. **Statistical Corner**, v. 66, p. 411–421, Jun. 2004. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/15184705>>. Acesso em: 03 ago. 2016.

BERNSTEIN, A. D.; PARSONNET, V. Bedside estimation of risk as an aid for decision-making in cardiac surgery. **Annals of Thoracic Surgery**, v. 69, n. 3, p. 823-828, Mar. 2000. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(99\)01424-1/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(99)01424-1/pdf)>. Acesso em: 03 ago. 2016.

BILLAH, B. et al. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. **European Journal Cardiothoracic Surgery**, v. 37, p. 1086-1092, Maio 2010. Disponível em: <<https://academic.oup.com/ejcts/article-lookup/doi/10.1016/j.ejcts.2009.11.021>>. Acesso em: 03 ago. 2016.

BOOTSMA, I. T. et al. Right ventricular ejection fraction in postoperative cardiac surgery patients is independently associated with ICU morbidity and mortality. **Intensive Care Medicine Experimental**, v. 3, p. 112-128, Dez. 2015, (Suplemento 1). Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4798364/pdf/40635_2015_Article_258.pdf>. Acesso em: 03 ago. 2016.

BORGHANS, I. et al. Fifty ways to reduce length of stay: an inventory of how hospital staff would reduce the length of stay in their hospital. **Health Policy**, v. 104, p.222–233, Mar. 2012. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168851011002892>>. Acesso em: 03 ago. 2016.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: teoria e prática**. 2. Ed. Rio de Janeiro: LTC, 2011.

BROCK, G. N. How to handle mortality when investigating length of hospital stay and time to clinical stability. **British Medical Journal and Medical Research Methodology**, v. 11, n.144, p. 1-14, Out. 2011. Disponível em:
<<https://bmcmredresmethodol.biomedcentral.com/articles/10.1186/1471-2288-11-144>>.
Acesso em: 03 ago. 2016.

BUTTERWORTH, J. et al. Female gender associates with increased duration of intubation and length of stay after coronary artery surgery. **Anesthesiology**, v. 92, n.2, p.414-424, Fev. 2000. Disponível em:< <http://anesthesiology.pubs.asahq.org/article.aspx?articleid=1946009>>.
Acesso em: 14 ago. 2016.

CABRAL, C. I. S. **Aplicação do modelo de regressão logística num estudo de mercado**. 2013. 59 f. Dissertação (Mestrado em Matemática Aplicada à Economia e Gestão) – Universidade de Lisboa, Dez. 2013. Disponível em:
<http://repositorio.ul.pt/bitstream/10451/10671/1/ulfc106455_tm_Cleidy_Cabral.pdf>.
Acesso em: 16 ago. 2016.

CARMONA, M. J. C.; et al. Avaliação pré-operatória do paciente cardiopata. In: CAVALCANTI, I. L.; CANTINHO, F. A. F.; ASSAD, A. (Org.). **Medicina perioperatória**. 1. ed. Rio de Janeiro: SAERJ (Sociedade de Anestesiologia do Estado do Rio de Janeiro), 2005, p. 193-237.

CARVALHO JUNIOR, L. H. et al. Reducing the length of hospital stay after total knee arthroplasty: influence of femoral and sciatic nerve block. **Revista da Associação de Medicina Brasileira**, v. 61, n. 1, p. 40-43, Fev. 2015. Disponível em:
<<http://www.scielo.br/pdf/ramb/v61n1/0104-4230-ramb-61-01-0040.pdf>>. Acesso em: 16 ago. 2016.

CARVALHO, M. R. M; et al. Complicações e tempo de internação na revascularização miocárdica em hospitais públicos no Rio de Janeiro. **Revista Brasileira de Terapia Intensiva**, v. 23, n. 3, p.312-320, Set. 2011. Disponível em:
<<http://www.scielo.br/pdf/rbti/v23n3/v23n3a09.pdf>>. Acesso em: 14 ago. 2016.

CACCIATORE, F. et al. Determinants of prolonged intensive care unit stay after cardiac surgery in the elderly. **Aging clinical and experimental research**, v. 24, n. 6, p. 627-634, Dez. 2012. Disponível em:< <https://www.ncbi.nlm.nih.gov/pubmed/22828570>>. Acesso em: 16 ago. 2016.

CALLOU, M. R. A.; RAMOS, P. R. M. Hipertensão arterial pulmonar. **Arquivos Brasileiros de Cardiologia**, v. 93, n. 6, Dez. 2009. Disponível em:
<<http://www.scielo.br/pdf/abc/v93n6s1/v93n6s1a16.pdf>>. Acesso em: 14 ago. 2016.

CASELLA, G.; BERGER, R. L. Modelos de regressão. In: **Inferência Estatística**. Tradução Visconde, S. A. 2. ed. São Paulo: Cenage Learning, 2014, p. 527 – 536.

CASTILLO, M. G. **Modelling patient length of stay in public hospitals in Mexico**. Maio

2012. 319 f. Tese (Doutor em Filosofia) – Universidade de Southampton, Reino Unido, 2012. Disponível em: <https://eprints.soton.ac.uk/345554/1/Final_PhD_thesis_-_Maria_Castillo.pdf>. Acesso em: 16 ago. 2016.

COCKER, J. et al. Preoperative prediction of intensive care unit stay following cardiac surgery. **European Journal of Cardiothoracic Surgery**, v. 39, p. 60-67, Jan. 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/20627608>>. Acesso em: 15 ago. 2016.

COONEY, M. T.; DUDINA, A. L.; GRAHAM, I. M. Value and limitations of existing scores for the assessment of cardiovascular risk: a review for clinicians. **Journal of the American College of Cardiology**, v. 59, p. 1209–1227, Set. 2009. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0735109709025029>>. Acesso em: 14 ago. 2016.

COSTA, V. E. A. et al. Impact of body mass index on outcome in patients undergoing coronary artery by-pass grafting and/or valve replacement surgery. **Brazilian Journal of Cardiovascular Surgery**, v. 30, n. 3, p. 335-342, Set. 2015. Disponível em: <<http://www.scielo.br/pdf/rbccv/v30n3/0102-7638-rbccv-30-03-0335.pdf>>. Acesso em: 14 ago. 2016.

COUTINHO, K. M. V. et al. Modelo Fuzzy estimando tempo de internação por doenças cardiovasculares. **Ciência & Saúde Coletiva**, v. 20, n. 8, p. 2585-2590, Jan. 2014. Disponível em: <<http://www.scielo.br/pdf/csc/v20n8/1413-8123-csc-20-08-2585.pdf>>. Acesso em: 15 ago. 2016.

D'AGOSTINO, R. B. et al. (CHD Risk Prediction Group). Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. **Journal of the American Medical Association (JAMA)**, v. 286, p. 180–187, Jul. 2001. Disponível em: <<http://jamanetwork.com/journals/jama/fullarticle/193997>>. Acesso em: 15 ago. 2016.

DeRIENZO, C. M. et al. A discrete event simulation tool to support and predict hospital and clinic staffing. **Health Informatics Journal**, v.1 n. 1, p. 1-10, Fev. 2016. Disponível em: <<http://journals.sagepub.com/doi/abs/10.1177/1460458216628314?journalCode=jhib>>. Acesso em: 15 ago. 2016.

DIAS FILHO, J. M.; CORRAR, L. J. Regressão Logística. In: CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M (coord.). **Análise multivariada: para cursos de administração, ciências contábeis, atuarias e financeiras**. 1. ed. São Paulo: Atlas, 2007, p. 280 – 317.

ENGEL, A. M.; MCDONOUGH, S.; SMITH, J. M. Does an obese body mass index affect hospital outcome after coronary artery bypass graft surgery?. **Annals of Thoracic Surgery**, v. 88, n. 6, p.1793-1800, Dez. 2009. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(09\)01581-1/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(09)01581-1/pdf)>. Acesso em: 15 ago. 2016.

EFRON, B.; TIBSHIRANI, R. J. **An introduction to the bootstrap**. New York: Chapman and Hall, 1. ed., 1993, p. 13-33.

EFRON, B. Bootstrap Methods: another look at the jackknife, **Annals of Statistics**, v. 7, p. 1-

26, Maio 1979. Disponível em: <<http://www.stat.cmu.edu/~fienberg/Statistics36-756/Efron1979.pdf>>. Acesso em: 15 ago. 2016.

FEDOROV, V; MANNINO, F; ZHANG, R. Consequences of dichotomization. **Pharm Stat Journal**, v. 8, p. 50 – 61, Mar. 2009. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/18389492>>. Acesso em: 15 ago. 2016.

FERRERO, C. A.; MALETZKE, A. G. ZALEWSKI, W. Modelos de regressão para a previsão de séries temporais por meio do algoritmo KNN-TSP. In: Encontro Nacional de Inteligência Artificial e Computacional (ENIAC), 12., **Anais...** São Carlos: BRACIS, Out. 2014. p. 1-8. Disponível em: <https://www.researchgate.net/publication/272086195_Modelos_de_Regressao_para_a_Previ_sao_de_Series_Temporais_por_meio_do_Algoritmo_kNN-TSP>. Acesso em: 09 set. 2016.

FLEGAL, K. M.; CARROLL, M. D.; OGDEN, C. L. Prevalence of obesity and trends in the distribution of body mass index among US adults in 1999-2010. **Journal of the American Medical Association (JAMA)**, v. 307, n. 5, p.491-509, Jan. 2012. Disponível em: <<http://jamanetwork.com/journals/jama/fullarticle/1104933>>. Acesso em: 09 set. 2016.

FIELD, A. **Discovering Statistics Using IBM SPSS Statistics**. 4. ed. Condado de Waukesha, EUA: SAGE, 2013. p. 720-743. Disponível em: <<https://us.sagepub.com/en-us/nam/discovering-statistics-using-ibm-spss-statistics/book238032>>. Acesso em: 09 set. 2016.

FOX, C. S. et al. The significant effect of diabetes duration on coronary heart disease mortality: the Framingham Heart Study. **Diabetes Care**, v. 27, n. 3, p.704-708, Mar. 2004. Disponível em: <<http://care.diabetesjournals.org/content/27/3/704.long>>. Acesso em: 09 set. 2016.

FREITAS, A. et al. Factors influencing hospital high length of stay outliers. **BioMed Central and Health Services Research**, v. 12, p. 265 – 275, Ago. 2012. Disponível em: <<https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-265>>. Acesso em: 09 set. 2016.

HACHESU, P. R. et al. Use of data mining techniques to determine and predict length of stay of cardiac patients. **Healthcare Informatics Research**, v.19, n. 2, p.121-129, Jun. 2013. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3717435/pdf/hir-19-121.pdf>>. Acesso em: 09 set. 2016.

HAIR JÚNIOR, J. F. et al. **Análise multivariada de dados**. Tradução Adonai Schlup Sant’Anna. 9. ed. Porto Alegre: Bookman, 2009. 120 f.

HARRELL, F. **Problems Caused by Categorizing Continuous Variables**. Universidade de Vanderbilt. Departamento de Bioestatística, Jan.2016. Disponível em: <<http://biostat.mc.vanderbilt.edu/wiki/Main/CatContinuous>>. Acesso em: 03 jan. 2017.

HAYKIN, S. **Neural Networks: a comprehensive foundation**. 2 Ed. Porto Alegre: Bookman, 2003. 205 f.

HOUTHOOFT, R. et al. Predictive modelling of survival and length of stay in critically ill

patients using sequential organ failure scores. **Artificial Intelligence in Medicine**, v. 63, p. 191–207, Mar. 2015. Disponível em: <[http://www.aiimjournal.com/article/S0933-3657\(14\)00147-X/](http://www.aiimjournal.com/article/S0933-3657(14)00147-X/)>. Acesso em: 23 ago. 2016.

HUANG, J. Q.; HOOPER, P. M.; MARRIE, T. J. Factors associated with length of stay in hospital for suspected community-acquired pneumonia. **Canadian Respiratory Journal**, v.13, n. 6, p. 317-324, Set. 2006. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2683319/>>. Acesso em: 23 ago. 2016.

JEREMIC, A.; TAN, K. Predicting the length of stay for neonates using heart-rate Markov models. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 30., Mar. 2008, Vancouver. **Anais...** Vancouver: EMBC, 2008. p.2912-2915. Disponível em: <ieeexplore.ieee.org/iel5/4636107/4649055/04649812.pdf>. Acesso em: 04/08/2016.

KALMAN, P. G.; et al. Preoperative factors that predict hospital length of stay after distal arterial bypass. **Journal of Vascular Surgery**, v. 20, n. 1, p. 70–75, Abr. 1994. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0741521494901775>>. Acesso em: 23 ago. 2016.

KAPLAN, R. C. et al. Body mass index, sex, and cardiovascular disease risk factors among hispanic/latino adults: hispanic community health study/study of latinos. **Journal of the American Heart Association**, v. 3, p. 1-12, Jul. 2014. Disponível em: <<http://jaha.ahajournals.org/content/3/4/e000923>>. Acesso em: 23 ago. 2016.

KARTHIK, S. et al. Effect of avoiding cardiopulmonary bypass in non-elective coronary artery bypass surgery: a propensity score analysis. **European Journal Cardiothoracic Surgery**, v. 24, n. 1, p. 66 – 71, Jul. 2003. Disponível em: <<https://academic.oup.com/ejcts/article/24/1/66/377633/Effect-of-avoiding-cardiopulmonary-bypass-in-non>>. Acesso em: 23 ago. 2016.

KHAIRUDIN, Z. Determinants of prolonged stay after coronary artery bypass graft surgery. **Procedia Social and Behavioral Sciences**, v. 36, n. 1, p. 87-95, Jun. 2011. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1877042812004776>>. Acesso em: 23 ago. 2016.

KHAIRUDIN, Z.; MOHD, N.; HAMID, H. Predictive models of prolonged stay after coronary artery bypass surgery. **Procedia - Social and Behavioral Sciences**, v. 36, Abr. 2012. p. 87-95. Disponível em: <<http://ieeexplore.ieee.org/document/6396536/>>. Acesso em: 04 jul. 2016.

KIM, H. D. et al. Resistant hypertension caused by stenosis of the aorta in elderly women: three case reports. **Clinical Hypertension**, v. 20, n. 5, p. 1-5, Nov. 2014. Disponível em: <<https://clinicalhypertension.biomedcentral.com/articles/10.1186/s40885-014-0005-2>>. Acesso em: 01 set. 2016.

KRELL. R. W.; GIROTTI, M. E.; DIMICK, J. B. Extended length of stay after surgery: complications, inefficient practice, or sick patients?. **Journal of the American Medical Association (JAMA)**, v. 149, n. 8, p. 815–820, Aug. 2014. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4245077/pdf/nihms643773.pdf>>. Acesso

em: 01 set. 2016.

KURKI, T. S. et al. Evaluation of the relationship between preoperative risk scores, postoperative and total length of stays and hospital costs in coronary bypass surgery. **European Journal of Cardiothoracic Surgery**, v. 20, p. 1183–1187, Dez. 2001. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/11717025>>. Acesso em: 01 set. 2016.

LaFARO, R. J. et al. Neural Network prediction of ICU length of stay following cardiac surgery based on pre-incision variables. **Plos One**, v. 10, n. 12, p. 1-19, Dez. 2015. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0145395>>. Acesso em: 01 set. 2016.

LaPAR, D. J. et al. A contemporary cost analysis of postoperative morbidity after coronary artery bypass grafting with and without concomitant aortic valve replacement to improve patient quality and cost-effective care. **Annals of Thoracic Surgery**, v. 96, p. 1620–1627, Ago. 2013. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(13\)01137-5/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(13)01137-5/pdf)>. Acesso em: 01 set. 2016.

LAWRENCE, D. R. et al. Parsonnet score is a good predictor of the duration of intensive care unit stay following cardiac surgery. **Heart**, v. 83, n. 4, p. 429 – 432, Abr. 2000. Disponível em: <<http://heart.bmj.com/content/heartjnl/83/4/429.full.pdf>>. Acesso em: 01 set. 2016.

LAWTON, J. S. et al. Magnetic resonance imaging detects significant sex differences in human myocardial strain. **BioMedical Engineering Online**, v. 10, p.76-87, Ago. 2011. Disponível em: <<https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-10-76>>. Acesso em: 01 set. 2016.

LeMOND, L.; ALLEN, L. A palliative care and hospice in advanced heart failure. **Progress in Cardiovascular Diseases**, v. 54, n. 2, p. 168 – 178, Out. 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3786355/pdf/nihms398149.pdf>>. Acesso em: 01 set. 2016.

MACHADO, L. M. Inference for non-Markov multi-state models: an overview. **Statistical Journal**, v. 9, n.1, p. 83–98, Mar. 2011. Disponível em: <<https://www.ine.pt/revstat/pdf/rs110105.pdf>>. Acesso em: 01 set. 2016.

MCCULLOCH W.; PITTS W. a logical calculus of the ideas immanent in nervous activity, **Bulletin of Mathematical Biophysics**, 5, p. 115-133, Mar. 1943. Disponível em: <<https://pdfs.semanticscholar.org/5272/8a99829792c3272043842455f3a110e841b1.pdf>>. Acesso em: 01 set. 2016.

MEDHI, M. et al. HbA1c Predicts length of stay in patients admitted for coronary artery bypass surgery. **Heart Disease**, v. 3, n. 2, p. 77-79, Abr. 2001. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/11975774>>. Acesso em: 01 set. 2016.

MEJÍA, O. A. V. et al. InsCor: avaliação de risco em cirurgia cardíaca. **Arquivo Brasileiro de Cardiologia**, v. 100, n. 3, p. 246-254, Mar. 2013. Disponível em:<<http://www.scielo.br/pdf/abc/v100n3/v100n3a06.pdf>>. Acesso em: 01 set. 2016.

MEJÍA, O. A. V. et al. The 2000 Bernstein-Parsonnet score and EuroSCORE were similar in

predicting mortality at the Heart Institute, USP. **Revista Brasileira de Cirurgia Cardiovascular**, v. 26, n.1, p. 1-6, Dez. 2011. Disponível em: <<http://www.scielo.br/pdf/rbccv/v26n1/v26n1a04.pdf>>. Acesso em: 01 set. 2016.

MELERO-LACASIA, A. Fragility in severe aortic stenosis patient: a multidimensional assessment from nursing. **Enfermeria Clinica**, v. 26, n. 2, p. 142-145, Dez. 2016. Disponível em: <<http://www.elsevier.es/es-revista-enfermeria-clinica-35-articulo-fragilidad-el-paciente-con-estenosis-S113086211500162X>>. Acesso em: 01 set. 2016.

MIHAJLOVIĆ, B. et al. Correlation between EuroSCORE and intensive care unit length of stay after coronary surgery. **Medicinski pregled Journal**, v. 44, n. 1-2, p. 46-50, Fev. 2011. Disponível em: <<http://www.doiserbia.nb.rs/img/doi/0025-8105/2011/0025-81051102046M.pdf>>. Acesso em: 01 set. 2016.

MINISTÉRIO DA SAÚDE. **Procedimentos hospitalares do SUS por local de internação no Brasil: DATASUS**, 2016. Disponível em: <<http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sih/cnv/qiuf.def>>. Acesso em 20 out. 2016.

MINISTÉRIO DA SAÚDE. **Prevenção e cuidado a doenças crônicas não-transmissíveis, 2008**. Disponível em: <http://dab.saude.gov.br/docs/geral/doenca_renal1.pdf>. Acesso em 20 out. 2016.

NAJAFI, M.; GOODARZYNEJAD, H. Determinants of length of stay in surgical ward after coronary bypass surgery: glycosylated hemoglobin as a predictor in all patients, diabetic or non-diabetic. **Journal of Tehran University Heart Center**, v. 7, n. 4, p. 170-176, Nov. 2012. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3537203/pdf/jthc-7-170.pdf>>. Acesso em: 01 set. 2016.

NAJAFI, M.; SHEIKHVATAN, M.; SHEIKHFATHOLLAHI, M. Discriminative power of EuroSCORE in predicting morbidity and prolonged hospital stay in an iranian sample population. **The Journal of Tehran University Heart Center**, v. 9, n.1, p. 15-20, Dez. 2014. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4277786/pdf/JTHC-9-15.pdf>>. Acesso em: 01 set. 2016.

NASHEF S. A. M. et al. European system for cardiac operative risk evaluation (EuroSCORE). **European Journal of Cardio-Thoracic Surgery**, v. 16, p. 9-13, Mai. 1999. Disponível em: <[http://euroscore.org/9\(1\).pdf](http://euroscore.org/9(1).pdf)>. Acesso em: 01 set. 2016.

NASHEF, S. A. M. et al. EuroSCORE II. **European Journal of Cardio-Thoracic Surgery**, v. 41, p. 734–745, Jun. 2012. Disponível em: <<http://www.endocardites-aquitaine.fr/fileadmin/documents/documents/2012nashef.pdf>>. Acesso em: 01 set. 2016.

NIKOLIC, A. EuroSCORE as a predictor of extended intensive care unit stay after cardiac surgery. **International Journal of Medicine**, v. 3, n. 1, p. 1-7, Fev. 2015. Disponível em: <<http://razavijournal.com/41120.pdf>>. Acesso em: 04 set. 2016.

NILSSON, J. et al. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. **Annals of Thoracic Surgery**, v. 78, p.1534–1535, Nov. 2004. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(04\)00998-1/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(04)00998-1/pdf)>. Acesso em: 04 set. 2016.

OLIVEIRA, E. K. et al. Risk factors for prolonged hospital stay after isolated coronary artery bypass grafting. **Brazilian Journal of Cardiovascular Surgery**, v. 28, n. 3, p. 353-363, Set. 2013. Disponível em: <<http://www.scielo.br/pdf/rbccv/v28n3/v28n3a10.pdf>>. Acesso em: 04 set. 2016.

ORTIGA, B.; et al. Standardizing admission and discharge processes to improve patient flow: A cross sectional study. **BioMed Central and Health Services Research**, v. 12, n. 180, p. 1-6, Jun. 2012. Disponível em: <<https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-12-180>>. Acesso em: 04 set. 2016.

OSNABRUGGE, R. L. et al. Prediction of costs and length of stay in coronary artery bypass grafting. **Annals Thoracic Surgery**, v. 98, p. 1286 – 1293, Ago. 2014. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(14\)01271-5/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(14)01271-5/pdf)>. Acesso em: 04 set. 2016.

PAGOWSKA-KLIMEK, I. et al. Predictors of long intensive care unit stay following cardiac surgery in children. **European Journal of Cardio-thoracic Surgery**, v. 40, p. 179-184, Jan. 2011. Disponível em: <<https://academic.oup.com/ejcts/article-lookup/doi/10.1016/j.ejcts.2010.11.038>>. Acesso em: 04 set. 2016.

PAPAKONSTANTINO, A. N. et al. Sex differentiation with regard to coronary artery disease. **Journal of Cardiology**, v. 62, p. 4–11, Jul. 2013. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0914508713000919>>. Acesso em: 04 set. 2016.

PENA, F. M. et al. Análise de um modelo de risco pre-operatório específico para cirurgia valvar e a relação com o tempo de internação em unidade de terapia intensiva. **Revista Brasileira de Terapia Intensiva**, v. 22, n.4, p. 339-345, Nov. 2010. Disponível em: <<http://www.scielo.br/pdf/rbti/v22n4/05.pdf>>. Acesso em: 04 set. 2016.

PENG, S. Y.; PENG, S. K. Predicting adverse outcomes of cardiac surgery with the application of artificial neural networks. **Anaesthesia**, v. 63, n. 7, p. 705-713, Jul. 2008. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2044.2008.05478.x/abstract;jsessionid=5A430C6163A176C262F0226ED5098661.f04t03>>. Acesso em: 04 set. 2016.

PETERSON, E. D. et al. Hospital variability in length of stay after coronary artery bypass surgery: results from the Society of Thoracic Surgeon's National Cardiac database. **Annals Thoracic Surgery**, v. 74, p. 464 –473, Ago. 2002. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(02\)03694-9/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(02)03694-9/pdf)>. Acesso em: 04 set. 2016.

PIEGAS, L. S.; BITTAR, O. J. N. V.; HADDAD, N. Cirurgia de revascularização miocárdica resultados do Sistema Único de Saúde. **Arquivo Brasileiro de Cardiologia**, v. 93, n. 5, p. 555-560, Nov. 2010. Disponível em: <http://www.scielo.br/pdf/abc/v93n5/en_a18v93n5.pdf>. Acesso em: 04 set. 2016.

POURHOSEINGHOLI, M. A. et al. Comparing linear regression and quantile regression to analyze the associated factors of length of hospitalization in patients with gastrointestinal tract

cancers. **Italian Journal Public Health**, v. 6, n. 2, p. 136-140, Ago. 2010. Disponível em: <<http://ijphjournal.it/article/view/5787/5531>>. Acesso em: 04 set. 2016.

RANUCCI, M. et al. Postoperative hypoxia and length of intensive care unit stay after cardiac surgery: the underweight paradox? (group for the surgical and clinical outcome research). **PLoS ONE**, v. 9, n. 4, p. 1-12, Abr. 2014. Disponível em: <<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0093992>>. Acesso em: 04 set. 2016.

RAVANGARD, R.; et al. Comparison of the results of Cox Proportional Hazards Model and parametric models in the study of length of stay in a tertiary teaching hospital in Tehran, Iran. **Acta Medica Iranica**, v. 49, n. 10, p. 650-659, Abr. 2011. Disponível em: <<http://acta.tums.ac.ir/index.php/acta/article/view/3812>>. Acesso em: 04 set. 2016.

ROCHA, H. A. L.; CARVALHO, E. R. **Conceitos Básicos em Epidemiologia e Bioestatística**. Faculdade de Medicina – Universidade Federal do Ceará, Mar. 2001 (Apostila). Disponível em: <<http://www.epidemiologia.ufc.br/files/ConceitosBasicosemEpidemiologiaeBioestatistica.pdf>>. Acesso em: 20 ago. 2016.

SCHMIDT, R.; GEISLER, S.; SPRECKELSEN, C. Decision support for hospital bed management using adaptable individual length of stay estimations and shared resources. **BioMed Central Medical Informatics and Decision Making**, v. 13, n.3, p. 1-19, Jan. 2013. Disponível em: <<https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-13-3>>. Acesso em: 04 set. 2016.

ROUZBAHMAN, M. JOVICIC, A. CHIGNELL, M. Can Cluster-Boosted Regression improve prediction: death and length of stay in the ICU?. **IEEE Journal Biomed Health Information**, v. --, n. 99, p. 1-11, Fev. 2016. Disponível em: <<http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel7%2F6221020%2F6363502%2F07397813.pdf&authDecision=-203>>. Acesso em: 04 set. 2016.

ROWAN, M. et al. The use of Artificial Neural Networks to stratify the length of stay of cardiac patients based on preoperative and initial postoperative factors. **Artificial Intelligence in Medicine**, v. 40 p.211-221, Jun. 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/17580112>>. Acesso em: 04 set. 2016.

SAUERBREI, W.; ROYSTON, P. Continuous variables: to categorize or to model? International Conference on Teaching Statistics, 8., **Invited Paper**, Jun. 2010, p. 1-5. Disponível em: <http://iase-web.org/documents/papers/icots8/ICOTS8_6D1_SAUERBREI.pdf>. Acesso em: 04 set. 2016.

SCHMIDT, R.; GEISLER, S. SPRECKELSEN, C. Decision support for hospital bed management using adaptable individual length of stay estimations and shared resources. **BioMed Central Medical Informatics and Decision Making**, v. 13, n. 3, p. 1-19, Jan. 2013. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3621822/pdf/1472-6947-13-3.pdf>>. Acesso em: 04 set. 2016.

SINGH, R.; MUKHOPADHYAY, K. Survival analysis in clinical trials: basics and must

know areas. **Perspectives in Clinical Research**, v. 2, n. 4, p. 145–148, Dez. 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3227332/>>. Acesso em: 04 set. 2016.

SHINJO, D.; FUSHIMI, K. Preoperative factors affecting cost and length of stay for isolated off-pump coronary artery bypass grafting: hierarchical linear model analysis. **British Medical Journal**, v. 5, n. 11, p. 1-10, Mar. 2015. Disponível em: <<http://bmjopen.bmj.com/content/bmjopen/5/11/e008750.full.pdf>>. Acesso em: 04 set. 2016.

SILBER, J. H. et al. Length of stay, conditional length of stay, and prolonged stay in pediatric Asthma. **Health Services Research**, v. 38, n. 3, p. 867-887, Jun. 2003. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360920/pdf/hesr_150.pdf>. Acesso em: 04 set. 2016.

SIMÃO, A. F. et al. Sociedade Brasileira de Cardiologia. I Diretriz Brasileira de Prevenção Cardiovascular. **Arquivos Brasileiros de Cardiologia**, v. 101 (6Supl. 2): 1-63, Dez. 2013. Disponível em: <http://publicacoes.cardiol.br/consenso/2013/Diretriz_Prevencao_Cardiovascular.pdf>. Acesso em: 04 set. 2016.

SILVA, G. S. et al. Avaliação do tempo de permanência hospitalar em cirurgia de revascularização miocárdica segundo a fonte pagadora. **Revista da Associação Médica Brasileira**, v. 59, n. 3, p. 248–253, Jun. 2013. Disponível em: <<http://www.scielo.br/pdf/ramb/v59n3/v59n3a10.pdf>>. Acesso em: 04 set. 2016.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais: para engenharia e ciências aplicadas**. São Paulo: Artliber, 1. ed. rev., p. 19-169, 2016.

SOCIEDADE BRASILEIRA DE DIABETES. **Diretrizes da Sociedade Brasileira de Diabetes (2015-2016)**, Jan 2016. Disponível em: <<http://www.diabetes.org.br/profissionais/images/docs/DIRETRIZES-SBD-2015-2016.pdf>>. Acesso em: 01 set. 2016.

TANIGUCHI, F. P.; SOUZA, A. R.; MARTINS, A. S. Tempo de circulação extracorpórea como fator risco para insuficiência renal aguda. **Revista Brasileira de Cirurgia Cardiovascular**, v. 22, n. 2, Jun. 2007. Disponível em: <<http://www.scielo.br/pdf/rbccv/v22n2/v22n2a08.pdf>>. Acesso em: 04 set. 2016.

TERADA, T. et al. Severe obesity is associated with increased risk of early complications and extended length of stay following coronary artery bypass grafting surgery. **Journal of the American Heart Association**, v. 5, p. 1-12, Jun. 2016. Disponível em: <<http://jaha.ahajournals.org/content/5/6/e003282.long>>. Acesso em: 04 set. 2016.

TOUMPOULIS, I. K. et al. Does EuroSCORE predict length of stay and specific postoperative complications after cardiac surgery?. **European Journal of Cardiothoracic Surgery**, v. 27, n. 1, p. 128-33, Jan. 2005. Disponível em: <<https://academic.oup.com/ejcts/article/27/1/128/342317/Does-EuroSCORE-predict-length-of-stay-and-specific>>. Acesso em: 04 set. 2016.

SCHUMER, E. M. et al. Emergency coronary artery bypass grafting: indications and outcomes from 2003 through 2013. **Texas Heart Institute Journal**, v. 43, n. 3, p. 214-220,

Jun. 2016. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4894699/>>. Acesso em: 04 set. 2016.

SILBERBACH, M. et al. Predicting hospital charge and length of stay for congenital heart disease surgery. **The American Journal of Cardiology**, v. 72, n. 12, p.958-963, Out. 1993. Disponível em: <<http://www.sciencedirect.com/science/article/pii/000291499391114W>>. Acesso em: 04 set. 2016.

SHAHIAN, D. M., et al. The Society of Coronary Artery Bypass Grafting (CABG): risk adjusted prolonged postoperative length of stay. **Annals of Thoracic Surgery**, v. 88, p. 1-20, Abr. 2009. Disponível em: <[http://www.annalsthoracicsurgery.org/article/S0003-4975\(09\)01056-X/pdf](http://www.annalsthoracicsurgery.org/article/S0003-4975(09)01056-X/pdf)>. Acesso em: 04 set. 2016.

SHEELA, K. G.; DEEPA, S. N. Review on methods to fix number of hidden neurons in neural networks. **Mathematical Problems in Engineering**, v. 1, p. 1-11, Mai. 2013. Disponível em: <<https://www.hindawi.com/journals/mpe/2013/425740/>>. Acesso em: 04 set. 2016.

SULLIVAN, P. G.; WALLACH, J. D.; IOANNIDIS, J. P. A. Meta-analysis comparing established risk prediction models (EuroSCORE II, STS Score, and ACEF Score) for perioperative mortality during cardiac surgery. **The American Journal of Cardiology**, v. 118, n. 10, p.1574-1582, Ago. 2016. Disponível em: <[https://linkinghub.elsevier.com/retrieve/pii/S0002-9149\(16\)31384-4](https://linkinghub.elsevier.com/retrieve/pii/S0002-9149(16)31384-4)>. Acesso em: 04 set. 2016.

TAFNER, M. A. **Redes Neurais Artificiais: Aprendizado e Plasticidade**. Universidade Estadual de Campinas, Núcleo de Informática Biomédica, Maio 2008. Disponível em: <<http://www.cerebromente.org.br/n05/tecnologia/rna.html>>. Acesso em: 20 out. 2016.

TSAI, P-F. J. et al. length of hospital stay prediction at the admission stage for cardiology patients using Artificial Neural Network. **Journal of Healthcare Engineering**, v. 16, p. 1 – 13, Jun. 2016. Disponível em:<<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5058566/>>. Acesso em: 20 out. 2016.

VAN DER BOON, R. M. A. et al. Effect of body mass index on short- and long-term outcomes after trans catheter aortic valve implantation. **American Journal of Cardiology**, v. 111, n. 2, p. 231–236, Out. 2013. Disponível em: <[http://www.ajconline.org/article/S0002-9149\(12\)02200-X/pdf](http://www.ajconline.org/article/S0002-9149(12)02200-X/pdf)>. Acesso em: 20 out. 2016.

VEGNI, R. et al. Postoperative cardiac artery bypass complications in elderly patients. **Revista Brasileira Terapia Intensiva**, v. 20, n. 3, p.226 – 234, Set. 2008. Disponível em: <http://www.scielo.br/pdf/rbti/v20n3/en_v20n3a04.pdf>. Acesso em: 20 out. 2016.

VIVACQUA, A. et al. Morbidity of bleeding after cardiac surgery: is it blood transfusion, reoperation for bleeding or both?. **Annals Thoracic Surgery**, v. 91, n. 6, p. 1780–1790, Jun. 2011. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/21619974>>. Acesso em: 04/09/2016.

WAGNER, T. H. et al. On-pump versus off-pump coronary artery bypass surgery: cost-effectiveness analysis alongside a multisite trial. **Annals Thoracic Surgery**, v. 96, p.770 – 777, Set. 2013. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0003->

4975(13)00894-1>. Acesso em: 20 out. 2016.

WANG, J. et al. Reducing Length of Stay in Emergency Department: A Simulation Study at a Community Hospital. **IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans**, 42., Out. 2012. Disponível em:

<<http://ieeexplore.ieee.org/document/6330037/>>. Acesso em 20/08/2016.

WANKE, P.; JULIANELLI, L. **Previsão de vendas: processos organizacionais e métodos quantitativos e qualitativos**. 2. ed. São Paulo: Atlas, 2011.

WILKS, D. S. **Statistical Methods in the Atmospheric Sciences**. 3. ed. Amsterdam: Academic Press, v. 100, 2011, p. 215-298.

WORLD HEALTH ORGANIZATION. **Report 2002: reducing risks, promoting healthy life**. Geneve: WHO, 2002. Disponível em: <<http://www.who.int/whr/2002/en/>>

Data de acesso: 18 ago. 2016.

APÊNDICE A – Validação cruzada das variáveis investigadas

Figura 19 – Prevalência do tempo de permanência pós-operatório por fator de risco



Fonte: Autora