



ASSOCIAÇÃO ENTRE PRIMING COM GELATINA E HEMORRAGIA PÓS-OPERATÓRIA EM CIRURGIA CARDÍACA

UMA ABORDAGEM COM MODELOS DE REGRESSÃO BINÁRIA EM
DADOS NÃO BALANCEADOS

RAFAEL MANIÉS PEREIRA
Licenciado em Cardiopneumologia

MESTRADO EM ESTATÍSTICA PARA A SAÚDE
Universidade NOVA de Lisboa
Maio, 2024



ASSOCIAÇÃO ENTRE PRIMING COM GELATINA E HEMORRÁGIA PÓS-OPERATÓRIA EM CIRURGIA CARDÍACA

UMA ABORDAGEM COM MODELOS DE REGRESSÃO BINÁRIA EM DADOS
NÃO BALANCEADOS

RAFAEL MANIÉS PEREIRA

Licenciado em Cardiopneumologia

Orientadora: Professora Doutora Isabel Cristina Maciel Natário,
Professora Associada, FCT-NOVA

Júri:

Presidente: Professor Doutor Pedro José dos Santos
Palhinhas Mota,
Professor Associado, FCT-NOVA

Arguentes: Professora Doutora Dora Susana Raposo Prata
Gomes,
Professora Associada, FCT-NOVA

Orientador: Professora Doutora Isabel Cristina Maciel
Natário,
Professora Associada, FCT-NOVA

MESTRADO EM ESTATÍSTICA PARA A SAÚDE

Universidade NOVA de Lisboa

Abril, 2024

Associação entre priming com gelatina e hemorragia pós-operatória em cirurgia cardíaca: uma abordagem com modelos de regressão binária em dados não balanceados

Copyright © Rafael Maniés Pereira, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Este trabalho é dedicado à minha mulher Vera Clemente e aos meus filhos,
Beatriz e Tiago Pereira.

AGRADECIMENTOS

Gostaria de agradecer, em primeiro lugar, à minha orientadora, Professora Doutora Isabel Natário, pela sua orientação sábia, paciência e apoio incansável ao longo deste trabalho. A sua capacidade de fornecer *insights* perspicazes e conselhos práticos foi verdadeiramente inestimável. Por várias ocasiões a sua orientação foi hábil ajudando a superar obstáculos aparentemente insuperáveis.

Gostaria também de expressar o meu agradecimento aos membros do Júri, por dedicarem parte do seu tempo e *expertise* na avaliação deste trabalho, fornecendo assim contribuições valiosas.

Para além disso, sou imensamente grato ao meu amigo Professor Doutor Tiago Velho, pelos seus *insights* e, sobretudo, pelo seu alento ao longo desta jornada.

Não poderia deixar de reconhecer o apoio basilar da minha família durante todo o processo de elaboração desta tese. O seu amor e compreensão foram verdadeiramente marcantes. Agradeço a sua flexibilidade e sacrifícios incomensuráveis nos meus momentos de isolamento e de concentração.

Por fim, agradeço a todos os meus colegas e a toda a equipa do Serviço de Cirurgia Cardiorácica do Hospital de Santa Maria que, através do seu trabalho diário forneceram recursos e geraram dados, permitindo que este trabalho se tornasse uma realidade.

*"Without data,
you're just another person with an opinion."*

(W. Edwards Deming)

RESUMO

O objetivo deste trabalho foi investigar se o *priming* com gelatina está associado a um aumento da hemorragia pós-operatória, após cirurgia cardíaca com Circulação Extracorporeal (CEC), em adultos. Para tal, realizou-se um estudo do tipo observacional, retrospectivo, num único centro cirúrgico.

A regressão binária pode ser estimada como um modelo linear generalizado, aplicando-se a uma variável de resposta binária. A função de regressão *log-log* complementar é uma alternativa de função de ligação utilizada, por exemplo, quando a probabilidade de um evento é muito pequena. Não são expectáveis dados balanceados na maioria das variáveis de interesse aplicadas às Ciências da Saúde. Como tal, existem técnicas de balanceamento, sendo que algumas utilizam a repetição, por aleatorização, de casos da classe minoritária. É o caso da técnica de *oversampling*.

Foram ajustados três modelos de regressão binária, aplicados a cada uma das bases de dados: (i) original; (ii) parcialmente balanceada; e (iii) completamente balanceada. O modelo que obteve o menor valor de AIC, corrigido à dimensão da amostra, foi o segundo. Os resultados sugerem que, os modelos estimados a partir de bases de dados balanceados proporcionaram melhorias significativas na capacidade de precisão na classificação. Contudo, o modelo com dados completamente balanceados não mostrou ser adequado em termos de *deviance*, pelo que poderá ter sido adicionada demasiada complexidade ao modelo. De um modo geral, os resultados sugerem que não existe evidência de que a presença de gelatina no *priming* esteja associada a um aumento da hemorragia, quer do ponto de vista quantitativo, quer do ponto de vista do aumento da probabilidade predita para a sua ocorrência.

Por fim, este trabalho realça o desafio no tratamento das bases de dados reais e aplicadas às áreas da saúde, nomeadamente em dados não balanceados, que tendem a condicionar sobre ajustamento nos modelos estimados.

Palavras chave: Modelos Lineares Generalizados, Modelos de Regressão Binária, Regressão log-log complementar, Balanceamento de Dados, Gelatina, Priming, Circulação Extracorporeal, Cirurgia Cardíaca

ABSTRACT

This study aimed to investigate whether priming with gelatin is associated with an increased risk of postoperative bleeding after cardiac surgery with Extracorporeal Circulation (ECC) in adults. For that, an observational and retrospective study was conducted at a single surgical center.

A binary regression can be estimated through generalized linear models, applied to a binary response variable. The complementary log-log regression function is an alternative link function used, for example, when the probability of an event is very small. Balanced data are not expected in most variables of interest in Health Sciences. As such, there are balancing techniques, some of which involve repeating cases from the minority class, by randomization. The oversampling technique exemplifies this.

Three binary regression models were adjusted for each database: (i) original; (ii) partially balanced; and (iii) completely balanced. The model with the lowest AIC value, adjusted for sample size, was the partially balanced model. The results suggest that models estimated from balanced databases significantly improved classification accuracy. However, the model with completely balanced data was not suitable in terms of deviance, suggesting that too much complexity may have been added to the model. Overall, the results suggest that there is no evidence that gelatin in priming is associated with an increase in bleeding, both quantitatively and in terms of increased predicted probability for its occurrence.

Finally, this work highlights the challenge of managing real databases within the healthcare domain, especially when dealing with unbalanced data. Such Imbalances often lead to overfitting in the models estimated.

Keywords: Generalized Linear Models, Binary Regression Models, Complementary Log-log Regression, Data Balancing, Gelatin, Priming, Extracorporeal Circulation, Cardiac Surgery.

ÍNDICE

1	INTRODUÇÃO	1
2	ESTADO DA ARTE	5
3	MÉTODOS	7
3.1	Testes não paramétricos	7
3.1.1	Teste de Mann-Whitney U	7
3.1.2	Teste de qui-quadrado para duas amostras independentes.....	9
3.1.3	Teste de Fisher.....	10
3.2	Regressão binária multivariada	11
3.2.1	Técnica de <i>Oversampling</i>	16
3.2.2	Validação Cruzada	16
4	DADOS.....	19
4.1	População Estudada	19
4.2	Variáveis analisadas	20
4.2.1	Variáveis dependentes.....	21
4.2.2	Variáveis individuais	23
4.2.3	Variáveis cirúrgicas	25
4.2.4	Variáveis Laboratoriais	26
4.2.5	Variáveis da CEC.....	26
4.3	Procedimento Perioperatório	27
4.4	Dados omissos.....	29
4.5	Análise Estatística	29

5	RESULTADOS	31
5.1	Análise Preliminar dos Dados.....	31
5.1.1	Variáveis dependentes	31
5.1.2	Variáveis individuais	34
5.1.3	Variáveis cirúrgicas	41
5.1.4	Variáveis laboratoriais pré-operatórias	42
5.1.5	Variáveis da CEC.....	44
5.2	Hemorragia pós-operatória e utilização de gelatina no <i>priming</i>	47
5.2.1	Comparação da hemorragia com as restantes variáveis explicativas	54
5.3	Modelo de Regressão Binária Multivariada para Hemorragia Significativa.....	65
5.3.1	Seleção das variáveis explicativas	65
5.3.2	Modelo utilizando a <i>base de dados original</i>	68
5.3.3	Modelo utilizando a <i>base de dados balanceados</i>	72
5.3.4	Interpretação dos modelos	78
5.3.4.1	Modelo utilizando a <i>base de dados original</i>	78
5.3.4.2	Modelo utilizando a <i>base de dados parcialmente balanceados</i>	84
5.3.5	Avaliação dos modelos	88
6	DISCUSSÃO	95
7	CONCLUSÕES.....	99

ÍNDICE DE FIGURAS

Figura 1 - Fluxograma de aplicação de critérios de inclusão e exclusão	20
Figura 2 - Organização das variáveis da amostra.....	21
Figura 3 - Variável dependente - Hemorragia no pós-operatório (12 e 24 horas).....	32
Figura 4 - Variável dependente - Categorização da hemorragia às 12 horas de pós-operatório	32
Figura 5 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório	33
Figura 6 - Variável dependente - Necessidade de Transfusões com UCE e Quantidade de UCE transfundidas no pós-operatório.....	34
Figura 7 - Variável individual - Sexo	35
Figura 8 - Variável individual - Idade	35
Figura 9 - Variável individual - Peso	36
Figura 10 - Variável individual - Altura.....	36
Figura 11 - Variável individual - Dislipidemia e Hipertensão Arterial.....	37
Figura 12 - Variável individual - Diabetes	37
Figura 13 - Variável individual - Tabagismo.....	38
Figura 14 - Variável individual - IMC e IMC por categoria	39
Figura 15 - Variável individual - Insuficiência Renal Crônica	39
Figura 16 - Variável individual - Enfarte e DVP	40
Figura 17 - Variável individual - Função Ventricular Esquerda.....	40
Figura 18 - Variável cirúrgica - Grau de Urgência.....	41
Figura 19 - Variável cirúrgica - Peso efetivo do tipo de cirurgia	42
Figura 20 - Variável laboratorial pré-operatória - Hemoglobina e Hematócrito.....	42
Figura 21 - Variável laboratorial pré-operatória - Plaquetas e Fibrinogénio	43
Figura 22 - Variável laboratorial pré-operatória - Tempo de Protrombina, INR e Tempo de Tromboplastina Parcial Ativado	44
Figura 23 - Variável da CEC - ACT Basal, ACT após Heparina e ACT Final	44

Figura 24 - Variável da CEC - Diferença Categórica entre o ACT Final e o ACT Basal	45
Figura 25 - Variável da CEC - Tempo de CEC e de Clampagem da Aorta	45
Figura 26 - Variável da CEC - Tipos de <i>Priming</i>	46
Figura 27 - Comparação das Categorias de Hemorragia Pós-operatória por tipo de <i>Priming</i> através do teste de Fisher	49
Figura 28 - Comparação da hemorragia às 12 horas de pós-operatório por tipo de <i>priming</i> através do teste de <i>Mann-Whitney U</i>	50
Figura 29 - Comparação da hemorragia às 24 horas de pós-operatório por tipo de <i>priming</i> através do teste de <i>Mann-Whitney U</i>	51
Figura 30 - Comparação das Categorias de Hemorragia Pós-operatória por tipo de <i>Priming</i> através do teste de qui-quadrado.....	52
Figura 31 - Comparação das Categorias da necessidade de transfusão sanguínea por tipo de <i>Priming</i> através do teste de qui-quadrado	53
Figura 32 - Hemorragia significativa no pós-operatória vs. Sexo através do teste de qui-quadrado.....	54
Figura 33 - Hemorragia significativa no pós-operatória vs. Necessidade de Heparina durante a CEC através do teste de qui-quadrado.....	55
Figura 34 - Hemorragia significativa no pós-operatória vs. ACT categórico após heparina através do teste de qui-quadrado.....	56
Figura 35 - Hemorragia significativa no pós-operatória vs. peso efetivo da cirurgia através do teste de <i>Fisher</i>	57
Figura 36 - Hemorragia significativa no pós-operatória vs. IMC categórico através do teste de qui-quadrado	57
Figura 37 - Hemorragia significativa no pós-operatória vs. Função ventricular esquerda através do teste de Fisher	58
Figura 38 - Hemorragia significativa no pós-operatória vs. feve2 através do teste de qui-quadrado.....	59
Figura 39 - Hemorragia significativa no pós-operatória vs. Insuficiência Renal Crônica através do teste de Fisher	60
Figura 40 - Hemorragia significativa no pós-operatória vs. Tabagismo através do teste de qui-quadrado.....	60
Figura 41 - Hemorragia significativa no pós-operatória vs. Doença vascular periférica através do teste de Fisher	61
Figura 42 - Hemorragia significativa no pós-operatória vs. grau de urgência do procedimento cirúrgico através do teste de qui-quadrado.....	62

Figura 43 - Hemorragia significativa no pós-operatória vs. Diferença de ACT categórico através do teste de qui-quadrado.....	62
Figura 44 - Hemorragia significativa no pós-operatório vs. Tempo de CEC através do teste de <i>Mann-Whitney-U</i>	63
Figura 45 - Hemorragia significativa no pós-operatório vs. Idade através do teste de <i>Mann-Whitney-U</i>	64
Figura 46 - Etapas da cirurgia com CEC	66
Figura 47 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório da base de dados de treino	69
Figura 48 - Curvas ROC dos dois modelos estimados com dados de treino	72
Figura 49 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório da base de dados parcialmente balanceada	74
Figura 50 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório da base de dados completamente balanceada	74
Figura 51 - Curvas ROC dos quatro modelos estimados.....	77
Figura 52 - Probabilidade predita de hemorragia significativa por tipo de <i>priming</i> e por tempo de CEC	80
Figura 53 - Probabilidade predita de hemorragia significativa por enfarte e por tempo de CEC	81
Figura 54 - Probabilidade predita de hemorragia significativa por sexo e por tempo de CEC	82
Figura 55 - Probabilidades preditas de acordo com o modelo para todos os cenários de combinações de variáveis categóricas	83
Figura 56 - Probabilidade predita de hemorragia significativa por tipo de <i>priming</i> e por tempo de CEC	86
Figura 57 - Probabilidade predita de hemorragia significativa por enfarte e por tempo de CEC	87
Figura 58 - Probabilidade predita de hemorragia significativa por sexo e cirurgia da aorta por tempo de CEC.....	88
Figura 59 - Gráfico de Resíduos do Modelo com dados originais.....	89
Figura 60 - Gráfico de resíduos por grupo.....	90
Figura 61 - Gráfico de resíduos padronizados do modelo com a base de dados original	91
Figura 62 - Avaliação de observações efetivamente influentes do modelo com dados originais	91
Figura 63 - Gráfico de Resíduos do Modelo com dados parcialmente balanceados	1
Figura 64 - Gráfico de Resíduos do Modelo com dados parcialmente balanceados	2

Figura 65 - Gráfico de resíduos padronizados do modelo com a base de dados parcialmente balanceada	2
Figura 66 - Avaliação de observações efetivamente influentes do modelo com dados parcialmente balanceados	3
Figura 67 - Gráfico de Resíduos do Modelo com dados completamente balanceados	4
Figura 68 - Gráfico de Resíduos do Modelo com dados completamente balanceados	4
Figura 69 - Gráfico de resíduos padronizados do modelo com a base de dados completamente balanceada	5
Figura 70 - Avaliação de observações efetivamente influentes do modelo com base de dados completamente balanceados	5

ÍNDICE DE TABELAS

Tabela 1 - Variáveis dependentes.....	21
Tabela 2 - Classificação categórica da hemorragia às 12 horas de pós-operatório utilizada ...	23
Tabela 3 - Variáveis individuais da amostra	23
Tabela 4 - Classificação categórica do índice de massa corporal.....	25
Tabela 5 - Variáveis cirúrgicas da amostra	25
Tabela 6 - Variáveis laboratoriais pré-operatórias.....	26
Tabela 7 - Variáveis da CEC.....	26
Tabela 8 - Comparação dos dados qualitativos por grupo de tipo de <i>priming</i>	47
Tabela 9 - Comparação dos dados quantitativos por grupo de tipo de <i>priming</i>	48
Tabela 10 - Comparação da hemorragia às 12 e 24 horas de pós-operatório, indexadas à ASC e Peso, por tipo de <i>priming</i>	51
Tabela 11 - Hemorragia significativa vs. restantes variáveis quantitativas	64
Tabela 12 - Fatores de inflação da variância das variáveis para o modelo de regressão binária	66
Tabela 13 - Variáveis incluídas no modelo	69
Tabela 14 - Resultados obtidos com os dois modelos estimados com a base de dados de treino	70
Tabela 15 - Resultados obtidos com os dois modelos estimados após stepwise.....	75
Tabela 16 - Probabilidades preditas por sexo e cirurgia da aorta por tempo de CEC.....	88
Tabela 17 - Validação cruzada com dados de teste de cada um dos modelos.....	93

GLOSSÁRIO

- Circulação Extracorporal** [1] Desvio do fluxo sanguíneo da aurícula direita para a artéria aorta, através de componentes extracorporais, permitindo que o sangue seja desviado do coração e dos pulmões durante uma cirurgia cardíaca.
- Gelatina** [2] Solução coloidal a 3.5% que contém peptídeos unidos por ureia. Tem um peso molecular de aproximadamente 35.000 Dalton e é preparada a partir de gelatina e eletrólitos. Esta solução é usada como expansor plasmático. É um dos componentes que se pode utilizar no priming.
- Heparina** [3] Um mucopolissacarídeo obtido do fígado, pulmão e mastócitos de vertebrados. A sua função, enquanto formulação farmacológica é a prevenção de coágulos sanguíneos *in vivo* e *in vitro*. É um dos componentes que se pode utilizar no *priming*.
- Polieletrólítico** [4] Soluções salinas isotónicas utilizadas na fluidoterapia por forma a promover uma reidratação e restauro do volume sanguíneo. É um dos componentes que se pode utilizar no *priming*.
- Priming** [5] Técnica de expurga do circuito de circulação extracorporal durante a preparação para a sua utilização. Este componente líquido do circuito, aquando da sua conexão ao doente, causa algum grau de hemodiluição por se misturar com o sangue do mesmo.
- Volémia** [6] Volume total de sangue. Corresponde à soma do volume de plasma com o volume dos componentes figurados do plasma.

SIGLAS

ACT	Tempo De Coagulação Ativado (Do Inglês, <i>Activated Clotting Time</i>)
AIC	Critério De Informação De Akaike (Do Inglês, <i>Akaike Information Criterion</i>)
aPTT	Tempo De Tromboplastina Parcial Ativada
AUC	Área Debaixo Da Curva (Do Inglês, <i>Area Under The Curve</i>)
CEC	Circulação Extracorporal
DVP	Doença Vascular Periférica
EAM	Enfarte Agudo Do Miocárdio
FIV	Fatores De Inflação Da Variância
GLM	Modelo Linear Generalizado (Do Inglês, <i>Generalized Linear Model</i>)
Hb	Hemoglobina
Htc	Hematócrito
IC	Intervalo De Confiança
IRC	Insuficiência Renal Crônica
INR	Relação Internacional Normalizada do TP
IT	Insulino-Tratado (Aplicado À <i>Diabetes</i>)
NIT	Não Insulino-Tratado (Aplicado À <i>Diabetes</i>)
TP	Tempo De Protrombina
UCE	Unidade De Concentrado Eritrocitário

SÍMBOLOS

$\tilde{\mu}$	Mediana
ns	Não Significativo
χ^2	Qui-Quadrado

INTRODUÇÃO

De acordo com o *Global Burden of Diseases, Injuries, and Risk Factors Study*, que estuda a evolução das doenças cardiovasculares desde 1990, tem-se verificado um aumento significativo destas, despoletando uma preocupação global com o crescente número de doentes que necessitam de cirurgia cardíaca [7]. A utilização da circulação extracorporeal (CEC) é crucial para a maioria das cirurgias cardíacas, permitindo a manutenção da perfusão dos órgãos e tecidos durante os procedimentos cardíacos [8]. Esta necessidade prende-se com o facto de, na maioria dos procedimentos, ser necessária a interrupção das funções cardíacas, assegurando previamente a adequada perfusão dos restantes órgãos e tecidos - técnica de circulação extracorporeal [8], [9], [10]. No entanto, a CEC não é isenta de complicações sendo que, o aperfeiçoamento contínuo é essencial de modo a minimizar os riscos e a otimizar os resultados [9], [11].

O circuito de CEC deve ser expurgado com uma qualquer solução, por forma a possibilitar que durante o início da técnica se anule o risco de embolia gasosa. A esta solução líquida, independentemente da sua composição, dá-se o nome de *priming*. A composição do *priming* é um assunto que, apesar dos anos de evolução da CEC, continua em permanente discussão [10], [12]. Até à data de publicação deste trabalho, não são conhecidas publicações que orientem os profissionais (Perfusionistas) nesse sentido. Nas mais recentes *guidelines* europeias de Perfusão Cardiovascular é referido que existem diversas combinações de composição de *priming* reportadas, variando de acordo com a tradição das diferentes equipas e departamentos cirúrgicos. Os *experts* referem ainda, que não existe consenso em relação ao tipo de *priming* mais indicado [13]. De um modo geral, o *priming* pode conter: (i) soluções cristaloides, como os polieletrólíticos, que são soluções que contêm água e sais minerais em concentrações isotónicas; (ii) soluções coloides, como por exemplo, as gelatinas; ou (iii) sangue, que pode ser autólogo ou homólogo. Para além disso, estão descritas diversas combinações,

nomeadamente a utilização combinada de soluções eletrolíticas e de soluções coloides [12], a que vulgarmente se dá o nome de *priming* misto.

Independentemente das soluções utilizadas no *priming*, existem algumas preocupações que devem ser tidas em conta no processo de escolha das mesmas [12]. Assim, a escolha do *priming* preocupa-se da melhor estratégia no que concerne à gestão de fluídos, e da sua osmolaridade, por forma a equilibrar a movimentação de fluídos entre os compartimentos intravascular e intersticial, procurando evitar o edema [12], [14]. Também deve ser tida em conta, a melhor estratégia que garanta a manutenção da adequada entrega de oxigénio, evitando a abrupta diminuição do transporte de oxigénio, por diminuição do hematócrito abaixo do seu limiar inferior (na prática clínica comumente designado por nadir) por diluição [8], [14] - afixado nos 24% [15]. Por fim, uma última preocupação acérrima com a escolha das soluções que compõem o *priming*, tem-se a preservação da hemostase, isto é, a capacidade fisiológica de prevenir, ou interromper, a hemorragia (neste caso, pós-operatória) [12], [13], [16].

No presente trabalho, o objetivo foi investigar se o *priming* com gelatina está associado à hemorragia, após cirurgia cardíaca com CEC, em adultos. Foi elaborado um estudo observacional, retrospectivo, num único centro cirúrgico. O objetivo primário foi verificar se o *priming* com gelatina está associado a um aumento da hemorragia às 12 e às 24 horas de pós-operatório. O objetivo secundário do presente estudo, foi estimar se a inclusão de gelatina no *priming* aumenta a probabilidade predita de hemorragia significativa às 12 horas de pós-operatório.

Como a variável de interesse, a hemorragia, não é balanceada na base de dados e na prática clínica, foram utilizados diversos modelos de acordo com três bases de dados distintas: uma original e duas replicadas artificialmente com técnica de *oversampling*, para balanceamento dos dados. Foram, *a posteriori*, analisados os ajustes destes modelos.

Para além deste capítulo introdutório, a presente tese está estruturada em outros 6 capítulos. O segundo capítulo pretende abordar o problema em estudo do ponto de vista de uma revisão da literatura sobre hemorragia e *priming* com gelatina. O terceiro capítulo corresponde aos Métodos, onde são abordadas as principais metodologias aplicadas à análise estatística deste trabalho. Segue-se a apresentação e explicação dos dados utilizados no Capítulo 4. Aqui são segmentadas as diferentes variáveis por tipo de variável e é explicado o protocolo de anestesia, operatório e da CEC, assim como, evidenciado o tipo de teste estatístico

aplicado a cada um dos objetivos desta tese. O quinto capítulo diz respeito aos resultados que estão compartimentados em três subcapítulos: (i) uma análise preliminar e estatisticamente descritiva dos dados por tipo de variável; (ii) uma análise que responde ao objetivo principal do trabalho; e (iii) uma análise dos resultados no objetivo secundário do mesmo. No sexto capítulo discutem-se os resultados e no sétimo, e último, conclui-se apontando possibilidade de trabalhos futuros.

ESTADO DA ARTE

A hemorragia em cirurgia cardíaca está associada a um aumento da mortalidade e morbidade [17], [18], [19], [20]. Karkouti et al. examinou o impacto direto de hemorragias significativas, em doentes submetidos a cirurgia cardíaca, e a sua relação com a taxa de mortalidade. Os resultados destacaram uma associação entre a hemorragia significativa e o aumento da mortalidade após cirurgia cardíaca [20]. Para além da mortalidade, a hemorragia está associada a diversos problemas dos quais se destacam o aumento do risco de acidente vascular cerebral, de insuficiência renal crónica e a um prolongamento da duração do internamento hospitalar [21], [22], que culminam com o aumento dos custos relacionados com a cirurgia cardíaca [23], [24].

Como referido anteriormente a composição do *priming* varia de acordo com a tradição das diferentes equipas e departamentos cirúrgicos, não existindo consenso em relação ao tipo de *priming* mais indicado [13]. Em 2017, e em resultado a um inquérito respondido por 106 centros cirúrgicos, de 18 países europeus, verificou-se que as soluções cristaloides são utilizadas no *priming* em 51.5% dos centros e que em 36% utiliza-se uma combinação de solução cristalóide e colóide [25].

A gelatina tem sido reportada como elemento constituinte do *priming* do circuito de CEC devido às suas melhores capacidades de manutenção do volume no espaço intravascular, quando comparada com um *priming* somente composto por soluções cristaloides [26], [27], [28]. Devido a esta capacidade, este tipo de solução encontra-se dentro da categoria de expansores do plasma sanguíneo [26], [27]. Aqui, o principal objetivo é reduzir o extravasamento de líquidos e a formação de edema, sendo que este efeito foi comprovado em estudos recentes [29]. Contudo, existem outras variáveis, relacionadas com o doente e com o circuito de CEC, que devem ser tidas em conta, nomeadamente, toda a gestão de fluidos perioperatória [13].

Apesar de fornecer vantagens na manutenção da volémia dos doentes, as gelatinas têm sido associadas a alterações da coagulação e da hemostase após cirurgia cardíaca [17], [30]. Schramko et al., num ensaio randomizado, estudaram os efeitos na coagulação, através de tromboelastografia modificada, após a administração de uma dose (7 ml/kg) de soluções coloides. Os autores verificaram que mesmo uma pequena dose de gelatina, prejudicou a força do coágulo após a cirurgia cardíaca de forma dose-dependente. Apesar disso, não se verificou o aumento da hemorragia em nenhum dos grupos [30]. Existem relatos que indicam que o distúrbio causado pela gelatina, na coagulação, só poderá ser revertido pela administração de fibrinogénio [17]. A relação entre o uso de gelatina e a hemorragia pós-operatória parece ser dose-dependente, tal como demonstrado também por Kuitunen et al.. Nos seus trabalhos, doses superiores a 20 mL/Kg parecem comprometer a hemostasia após a cirurgia cardíaca. Os autores referem ainda, de forma subjetiva, que este efeito parece estar relacionado com a formação de um coágulo pouco estável [31]. Já Shanliang e Yu, verificaram que o grupo de doentes a quem foi administrada gelatina, prolongou significativamente o tempo de protrombina (TP) e o tempo de tromboplastina parcialmente ativado (aPTT) e que diminuíram o fator VIII e o fator *von Willebrand* imediatamente após a sua administração [32]. Outro autor, numa revisão dos efeitos de diversos tipos de fluidos na coagulação, refere que os efeitos colaterais na coagulação, provocados pelos coloides, são unicamente dose-dependentes. Assim, refere também que se os coloides forem administrados de forma controlada, os efeitos na força do coágulo são mínimos e não exigem qualquer tipo de reversão [33].

Como descrito anteriormente, existe evidência de que, de uma forma experimental e associada a parâmetros analíticos, a utilização de gelatina se relaciona com alterações em alguns parâmetros da coagulação. De um modo geral, parece que estes distúrbios não estão relacionados com um aumento quantitativo da hemorragia pós-operatória, traduzido por um eventual aumento do volume de sangue contabilizado. No trabalho de pesquisa realizado, não foi possível encontrar nenhum estudo que associe a utilização de gelatina no *priming* com a probabilidade predita de hemorragia significativa após cirurgia cardíaca com circulação extracorporeal.

MÉTODOS

Neste capítulo, será apresentada uma revisão detalhada dos métodos estatísticos aplicados neste trabalho, incluindo a sua formulação matemática, suposições e pressupostos subjacentes, bem como a interpretação dos resultados. Além disso, serão discutidas algumas considerações importantes para a aplicação deste método em estudos nas áreas da saúde.

3.1 Testes não paramétricos

Nos testes não paramétricos, a estatística de teste não depende da distribuição da população em estudo. Para além disso, alguns testes, como o caso do qui-quadrado baseiam-se em frequências de classes e estimativas de pertença a essas classes. Por outro lado, outros testes, como o caso do teste de *Mann-Whitney U* comparam os valores centrais baseando-se em ordenações e não em valores absolutos da variável de interesse. Assim, a distribuição da variável não afeta a execução do teste, não sendo por isso um pressuposto do mesmo [34].

3.1.1 Teste de Mann-Whitney U

O teste de *Mann-Whitney U* (teste não paramétrico) é um teste que, para amostras de grandes dimensões, tem uma potência de teste muito semelhante à de um *teste-t* (teste paramétrico). Este teste não paramétrico considera a mediana ($\tilde{\mu}$) como parâmetro de localização [35]. Existem alguns pressupostos deste teste aplicado a amostras independentes, tais como: (i) cada amostra ser aleatória; (ii) as duas amostras serem mutuamente independentes; (iii) a escala de medição ser, pelo menos, ordinal; e (iv) a distribuição da população subjacente ser simétrica. Este último pressuposto é muito menos rígido do que o

pressuposto da distribuição Normal para os testes paramétricos [36]. Se houver uma diferença nas duas funções de distribuição $F(x)$ e $G(y)$, então essa é uma diferença na localização das distribuições. Por outras palavras, se $F(x)$ não é idêntico a $G(y)$, então $F(x)$ é idêntico a $G(y + c)$, sendo c uma constante [37].

Seja, X_1, X_2, \dots, X_n uma amostra aleatória de tamanho n da população 1 e Y_1, Y_2, \dots, Y_m uma amostra aleatória de tamanho m da população 2. Seja também $n + m = N$. Atribui-se classificações de 1 a N a todas as observações, ordenadas da menor para a maior, independentemente da população de origem, 1 ou 2. Represente $R(X_i)$ e $R(Y_j)$ as classificações atribuídas a X_i e Y_j para todos os valores i e j . Se se obtiverem empates deverá ser atribuído a cada um a média das classificações que teria sido atribuída, caso não existisse empate [36].

Assim, para as duas amostras independentes, o teste pode ser formulado em termos da diferença das medianas ($\tilde{\mu}$) populacionais. Seja H_0 a hipótese nula e H_1 a hipótese alternativa, as seguintes hipóteses devem ser formuladas:

$$\begin{aligned} H_0: \tilde{\mu}_X &= \tilde{\mu}_Y \\ H_1: \tilde{\mu}_X &\neq \tilde{\mu}_Y \\ &\text{(teste bilateral) [36], [37]} \end{aligned}$$

Assim, a estatística de teste é dada pela fórmula 3.1,

$$W' = \frac{W' - \text{média}}{\text{desvio padrão}} = \frac{W^* - n \frac{N+1}{2}}{\sqrt{\frac{nm}{N(N-1)} \sum_{i=1}^N R_i^2 - \frac{nm(N+1)^2}{4(N-1)}}} \quad (3.1)$$

e que se aplica a amostras de grandes dimensões e com muitos empates e onde $\sum R_i^2$ corresponde à soma dos quadrados de todos os N ranks usados nas duas amostras [36], [37].

Sob H_0 , a estatística de teste segue uma distribuição Normal estandardizada. A regra de teste para este caso será, para um nível de significância de α igual a 5%, rejeitar-se a hipótese nula (H_0) se o valor-p for menor que 5% [35], [36], [37].

3.1.2 Teste de qui-quadrado para duas amostras independentes

Quando se pretende analisar a relação existente entre duas variáveis que sejam do tipo qualitativas, dever-se-á aplicar um teste de associação do qui-quadrado. Como se trata de um teste não-paramétrico, não existe necessidade de conhecer a distribuição das variáveis. Este teste baseia-se numa tabela de contingência, que traduz as frequências observadas em cada uma das categorias, por cada célula. Assim, este teste faz a comparação entre as frequências observadas e as frequências esperadas, no caso de as duas variáveis serem independentes [38].

O teste de qui-quadrado também apresenta alguns pressupostos e que devem ser cumpridos, sendo eles: (i) a amostra deve ser aleatória; (ii) os resultados das várias amostras são mutuamente independentes; (iii) cada observação pode ser categorizada exatamente em uma das $r \times c$ categorias (ou seja, as categorias são mutuamente exclusivas e coletivamente exaustivas); (iv) foi determinado o número de observações em cada amostra colhida; (v) a escala de medição deve ser, pelo menos, ordinal; e (vi) cada célula da tabela de contingência deve conter, pelo menos, 5 observações [34], [37].

Para avaliar se duas variáveis qualitativas X e Y são associadas, considere-se a tabela de contingência dos valores observados O_{ij} na classe i da variável X , $i = 1, \dots, r$, e na classe j da variável Y , $j = 1, \dots, c$, organizados numa tabela de contingência $r \times c$. Tem-se que $n_i = O_{i1} + O_{i2} + \dots + O_{ij} + \dots + O_{ic}$ para todo o i [18], [20].

Seja H_0 a hipótese nula e H_1 a hipótese alternativa, as seguintes hipóteses devem ser formuladas:

H_0 : as variáveis são independentes (não estão associadas)

H_1 : as variáveis não são independentes (estão associadas)

A estatística de teste baseia-se nas probabilidades esperadas de cada célula, com base no que deveria ser, se as distribuições de X para cada nível de Y fossem de facto iguais (que corresponde à hipótese nula). Se isso acontecer, as probabilidades estimadas de cada célula aproximam-se do produto das correspondentes probabilidades marginais estimadas, fazendo com que o número observado de cada célula se aproxime do número esperado de observações, contribuindo com uma parcela pequena para a estatística de teste que avalia estas diferenças. Caso não sejam iguais, essa diferença fará com que o valor da estatística do teste aumente. É sabido que um valor de estatística de teste grande tende a rejeitar a hipótese nula (o que

significa que se rejeita a hipótese de as distribuições serem iguais e as variáveis independentes) e a não rejeitar a hipótese alternativa (assumindo que, pelo menos duas das distribuições, sejam diferentes) [35]. Os valores esperados de cada célula E_{ij} , são calculados de acordo com o apresentado em 3.2.

$$E_{ij} = \frac{n_i n_j}{N} \quad (3.2)$$

onde $n_i = \sum_j O_{ij}$ é o total na linha i , e $n_j = \sum_i O_{ij}$ o total da coluna j e N o total da amostra.

A estatística de teste é então dada por:

$$X^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{i=1}^r \sum_{j=1}^c \frac{O_{ij}^2}{E_{ij}} - N \quad (3.3)$$

Sob H_0 , a estatística de teste segue uma distribuição de qui-quadrado (X^2) com $(r - 1)(c - 1)$ graus de liberdade, representado abaixo e, onde r representa o número total de linhas e c o número total de colunas [34], [37], [38].

$$X^2 \sim X^2_{(r-1)(c-1)} \quad (3.4)$$

A regra de teste para este caso será, e para um nível de significância de α igual a 5%, rejeitar-se a hipótese nula (H_0), ou seja, rejeitar a independência das variáveis, se $X^2 > X^2_{(95\%)(gl)}$ ou, se o valor-p for menor que 5%, onde gl corresponde aos graus de liberdade calculados [36].

3.1.3 Teste de Fisher

O teste de *Fisher* é um teste não paramétrico utilizado quando se pretende comparar duas populações a partir de amostras independentes de pequena dimensão e relativas a uma variável apenas dicotômica. Assim, o teste de Fisher apresenta-se como uma alternativa ao teste do qui-quadrado, apenas para tabelas 2x2, sempre que os pressupostos do primeiro não são cumpridos, nomeadamente (i) a dimensão da amostra ($N \geq 20$); (ii) todos os E_{ij} serem superiores a um; e (iii) pelo menos 80% dos E_{ij} serem superiores a 5 [35].

Seja H_0 a hipótese nula e H_1 a hipótese alternativa, as seguintes hipóteses devem ser formuladas:

H_0 : as variáveis são independentes (não estão associadas)

H_1 : as variáveis não são independentes (estão associadas)

O teste exato de *Fisher* não utiliza a estatística qui-quadrado nem a distribuição amostral. Ao invés disso, calcula as distribuições marginais, ou seja, considera todas as tabelas de contingência possíveis, desde os totais de linhas, aos totais de colunas da tabela que registra as observações. Assim, é calculado o valor-p encontrando a proporção de tabelas possíveis que são mais extremas do que a tabela observada [36].

A regra de teste para este caso será também, e para um nível de significância de α igual a 5%, rejeitar-se a hipótese nula (H_0), ou seja, rejeitar a independência das variáveis, se o valor-p for menor que 5% [36].

3.2 Regressão binária multivariada

A regressão logística binária multivariada é uma técnica estatística utilizada para modelar a relação entre uma variável binária, isto é, uma variável que assuma duas categorias de resposta, e um conjunto de variáveis explicativas [39]. Seja Y a variável de interesse e X_1, X_2, \dots, X_p as variáveis explicativas de interesse, o modelo de regressão logística é dado por:

$$\log\left(\frac{P(Y = 1)}{1 - P(Y = 1)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (3.5)$$

onde, $P(Y = 1)$ é a probabilidade do evento de interesse, β_0 é o intercepto, $\beta_1, \beta_2, \dots, \beta_p$ são os coeficientes de regressão que representam o efeito das variáveis explicativas na variável resposta e para p variáveis independentes. Por fim, a função logística $\log\left(\frac{P(Y=1)}{1-P(Y=1)}\right)$, representando a log-chance, tem como função a transformação da escala linear, numa escala de probabilidades, isto é, em valores contínuos dispostos entre 0 e 1 [40]. Em diversos ensaios clínicos ou estudos das áreas da saúde é experimentada a conotação binária como, por exemplo, para conotar *sim* vs. *não* ou *sucesso* vs. *insucesso*. Pretende-se formular uma equação de regressão logística para poder estimar que variáveis influenciam o resultado e

assim, avaliar a probabilidade de ocorrência do evento estudado [41]. Assim, e por conveniência matemática, atribui-se o evento de interesse como $Y = 1$, e o seu contrário $Y = 0$ [42]. Este método estatístico não cumpre os pressupostos para a utilização do método dos mínimos quadrados, pelo que a estimação dos coeficientes é realizada através do método da máxima verosimilhança [35].

Para que os resultados da regressão logística binária sejam válidos, é necessário o cumprimento de diversas suposições e pressupostos, tais como, (i) a independência, ou seja, a ocorrência de um evento para um indivíduo não deve influenciar a ocorrência do evento para outro indivíduo; (ii) a linearidade na log-chance, ou seja, a relação entre as variáveis independentes e a log-chance da variável dependente deve ser linear; e (iii) ausência de multicolinearidade: as variáveis independentes não devem estar fortemente correlacionadas entre si. A presença de multicolinearidade pode levar a estimativas imprecisas dos coeficientes de regressão [39].

No caso de uma regressão logística simples a razão de chances (*Odds Ratio*) de uma variável mede o grau de associação, dessa mesma variável, com a única variável explicativa. Contudo, numa regressão logística múltipla, em que existe mais do que uma variável explicativa, a razão de chances de uma variável mede o grau de associação entre a variável dependente e uma dada variável explicativa, mas apenas depois de controladas ou fixos os efeitos de todas as outras variáveis independentes e que constam no modelo estimado [43], [44].

Por forma a avaliar o modelo, dever-se-á proceder ao teste da razão de verosimilhança em que se averigua a hipótese de pelo menos um dos parâmetros estimados ser diferente de zero, à exceção do intercepto. A estatística de teste corresponde à *Deviance* e tem distribuição de qui-quadrado com $k-1$ graus de liberdade. Para a avaliação da significância de cada um dos parâmetros estimados é comum utilizar-se a estatística de *Wald*. Assim, é testada a hipótese nula de que o coeficiente seja igual a zero contra a hipótese alternativa de que o parâmetro é diferente de zero. A estatística de teste de Wald é definida como o quadrado da razão entre o estimador de máxima verosimilhança tanto para o próprio coeficiente, como para o seu desvio-padrão (DP). Importa ainda realçar que a estatística de teste segue uma distribuição de qui-quadrado com um grau de liberdade. A equação 3.6 representa a estatística de teste anteriormente mencionada [39].

$$W = \left(\frac{\hat{\beta}_j}{DP(\hat{\beta}_j)} \right)^2 \sim \chi_{(1)}^2 \quad (3.6)$$

Quando o modelo estimado tem uma qualquer variável independente que seja categórica e com k níveis de resposta, o k -ésimo nível será assumido como base de comparação, ou outro de acordo com a preferência e, portanto, podem estabelecer-se $k - 1$ relações, contrastando cada nível de resposta com o nível de referência [43]. Assim, tem-se:

$$g_j(X) = \ln \left[\frac{P(y|X_j)}{P(y|X_k)} \right] = \beta_0 + \beta_j X_j, \text{ com } 1 \leq j \leq k - 1 \quad (3.7)$$

Exponenciando os dois lados da igualdade, tem-se:

$$e^{g_j(X)} = \frac{P(y|X_j)}{P(y|X_k)} = e^{\beta_0 + \beta_j X_j}, \text{ com } 1 \leq j \leq k - 1 \quad (3.8)$$

Assim, compara-se a probabilidade de sucesso para a categoria j desta covariável com o seu nível de referência, k neste caso, através da quantidade e^{β_j} que corresponde à razão entre as respetivas probabilidades. Portanto, para uma observação que pertença à categoria j desta variável explicativa categórica, estima-se que a sua probabilidade de sucesso seja e^{β_j} a probabilidade de sucesso que se obteria se pertencesse à classe de referência [39], [44].

A curva ROC (*Receiver Operating Characteristics*) representa a relação entre a sensibilidade e a especificidade de um teste de diagnóstico quantitativo. Quanto maior é o ponto de corte, maior é a especificidade do teste, mas menor é a sua sensibilidade. A área sob a curva ROC representa o poder que o teste tem de discriminar os positivos dos negativos. Quanto mais perto de 1, maior é a sua capacidade discriminante, sendo desejável que não seja inferior a 70% [35]. Duas outras medidas são importantes na avaliação do modelo de regressão logística binária: a sensibilidade e a especificidade. A sensibilidade é a capacidade de um teste identificar os verdadeiros positivos nos indivíduos que têm efetivamente o evento. Calcula-se a sensibilidade, portanto, pelo quociente entre o número de casos positivos e com eventos e o número total de casos positivos do modelo. Por outro lado, a especificidade é a capacidade de um teste identificar os verdadeiros negativos nos indivíduos nos quais não se verificou o evento. Calcula-se a especificidade, portanto, pelo quociente entre o número de casos negativos e sem o evento e o número total de casos negativos estimados pelo modelo [39], [43], [44], [45].

Este modelo pertence ao grupo dos modelos lineares generalizados (GLM). Neste tipo de modelos, devem ser especificados três componentes: (i) um componente aleatório; (ii) um componente sistemático; e (iii) uma função de ligação. Um modelo de regressão logística nesse caso, utiliza uma função de probabilidade binomial, ou de Bernoulli, como componente aleatório. Esta componente corresponde a uma distribuição de probabilidade discreta de variáveis aleatórias que assumem o valor de 1 com probabilidade p e o valor de 0 com a probabilidade igual a $1 - p$. Para o segundo componente, são utilizadas funções preditoras lineares $X\beta$ (onde X é uma matriz das covariáveis e β é um vetor com os respectivos coeficientes). Trata-se de uma ponderação da regressão pelos seus coeficientes. Aqui, esta função deve ser transformada utilizando, por exemplo, a função logística de modo que se obtenham probabilidades. A esta função de ligação corresponde uma função logística com um formato sigmoide e transforma a combinação linear do componente sistemático numa probabilidade no intervalo de valores 0 a 1. Assim, para este tipo de modelo, a função de ligação, logística, é definida como o descrito na equação 3.9, onde z corresponde à combinação linear de preditores [39], [43], [44].

$$g(z) = \frac{1}{1 + e^{-z}} \quad (3.9)$$

Esses três componentes em conjunto formam o modelo de regressão logística binária, que é comumente usado para modelar a probabilidade de uma variável resposta binária em relação a um conjunto de variáveis explicativas [45]. Uma das principais vantagens de se usar modelos de regressão logística neste tipo de trabalhos é que a probabilidade da variável resposta é então modelada diretamente, tendo em conta a natureza aleatória do fenómeno de interesse [39], [44].

Existem diversas funções de ligação, que devem ser utilizadas de forma criteriosa. Apresenta-se a regressão *log-log complementar* ("*cloglog*"). A regressão *cloglog* é também uma técnica de modelação estatística utilizada para analisar variáveis de resposta binária. Trata-se de uma alternativa à regressão logística em situações características, como por exemplo, quando a probabilidade de um evento é muito pequena ou muito grande. Na maioria das vezes, a regressão de *cloglog* é aplicada para lidar com eventos raros ou em situações em que a base de dados é não balanceada [46]. Uma base diz-se não balanceada quando as classes ou categorias de resposta não se encontram igualmente representadas. A regressão *cloglog* usa a função log-log complementar que gera uma curva em forma de S , mas assimétrica. A

regressão cloglog pode ser representada conforme consta na equação 3.10, em que o lado esquerdo da equação corresponde a essa mesma transformação [47].

$$\ln(-\ln(1 - P(Y = 1))) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (3.10)$$

Assim, o modelo também pode ser descrito como consta na equação 3.11 [47].

$$P(Y = 1) = 1 - e^{-e^{\sum_{i=1}^k \beta_i X_i}} \quad (3.11)$$

O Critério de Informação de Akaike (AIC) é uma medida comumente empregada na seleção de modelos estatísticos, especialmente quando se comparam modelos com diferentes números de parâmetros. Este critério combina a adequação do modelo aos dados, medida pela log-verossimilhança (ℓ), com uma penalização associada à complexidade do modelo, representada pelo número de coeficientes de regressão para covariáveis não constantes (p). A fórmula do AIC é dada pelo expresso na equação 3.12 [39], [43], [48].

$$AIC = -2\ell + 2(p + 1) \quad (3.12)$$

O fator de correção $2(p + 1)$ é adicionado à log-verossimilhança multiplicada por -2. Um valor menor de AIC, por comparação, indica um melhor ajuste do modelo aos dados. Portanto, na seleção de modelos, o objetivo é minimizar o valor do AIC, pois isso sugere que o modelo é mais adequado aos dados observados. Em resumo, o AIC oferece uma forma objetiva de comparar modelos estatísticos e selecionar o mais apropriado para um determinado fim [48].

Hurvich e Tsai (1989) derivaram um AIC corrigido (AICc) que inclui uma correção para o tamanho da amostra como o expresso na equação 3.13, onde AIC é o AIC obtido, k é o número de parâmetros no modelo e n é o número de observações, ou seja, a dimensão da amostra [49], [50].

$$AICc = AIC + \frac{2k(k + 1)}{nk - 1} \quad (3.13)$$

No contexto de modelos de regressão multivariada, o método de stepwise [45] é um procedimento utilizado para selecionar as variáveis mais importantes para o modelo. Isto acontece, removendo ou adicionando variáveis independentes de acordo com um critério de informação, frequentemente o AIC [39]. Existem duas abordagens principais dentro do

método de *stepwise: forward* e *backward*. Na modalidade de *forward*, as variáveis são adicionadas ao modelo uma de cada vez, começando com a variável mais significativa. As variáveis que não são estatisticamente significativas são eliminadas e, entre as sobrantes, são escolhidas aquelas que estão mais fortemente relacionadas com a variável dependente. Esse processo continua até que a adição de mais uma variável não melhore significativamente o modelo. Na abordagem *backward*, todas as variáveis são inicialmente incluídas no modelo. Em seguida, a variável menos significativa é removida e o processo é repetido até que todas as variáveis sobrantes estejam significativamente relacionadas com a variável dependente. [45].

3.2.1 Técnica de *Oversampling*

A técnica de *oversampling*, em geral, é uma abordagem utilizada para lidar com conjuntos de dados não balanceados, ou seja, quando uma das classes da variável dependente é muito mais provável do que outra. Algumas métricas de avaliação de desempenho dos modelos podem ser suscetíveis de erros, provocados por uma maior probabilidade de classificação para a classe majoritária, em detrimento da classe minoritária [51], [52]. Existem diversos métodos de *oversampling*, contudo nem todos se aplicam a uma base de dados com variáveis categóricas. Aplica-se *oversampling* com o objetivo de aumentar a quantidade de dados na classe minoritária, por forma a tentar equilibrar a distribuição das classes. Assim, a técnica mais simples, para aumentar o tamanho da classe minoritária, é a sobreamostragem aleatória, ou seja, a replicação aleatória de exemplos dessa mesma classe [53].

3.2.2 Validação Cruzada

Para avaliar o desempenho dos modelos estimados, utilizou-se a técnica de *Holdout*. Esta técnica propõe a divisão aleatória do conjunto de dados num conjunto de treino do modelo, que representará 80% da base de dados, e um conjunto de teste, que compreenderá os 20% restantes. Assim, a estimação dos modelos será realizada com recurso aos dados de treino. Em seguida, avalia-se o desempenho de cada modelo utilizando o conjunto de dados de teste. Por forma a que as amostras das bases de treino e de teste contenham uma adequada representatividade de casos da variável de interesse, procede-se a uma aleatorização estratificada. Em vez de selecionarem observações aleatórias para compor a base de dados de

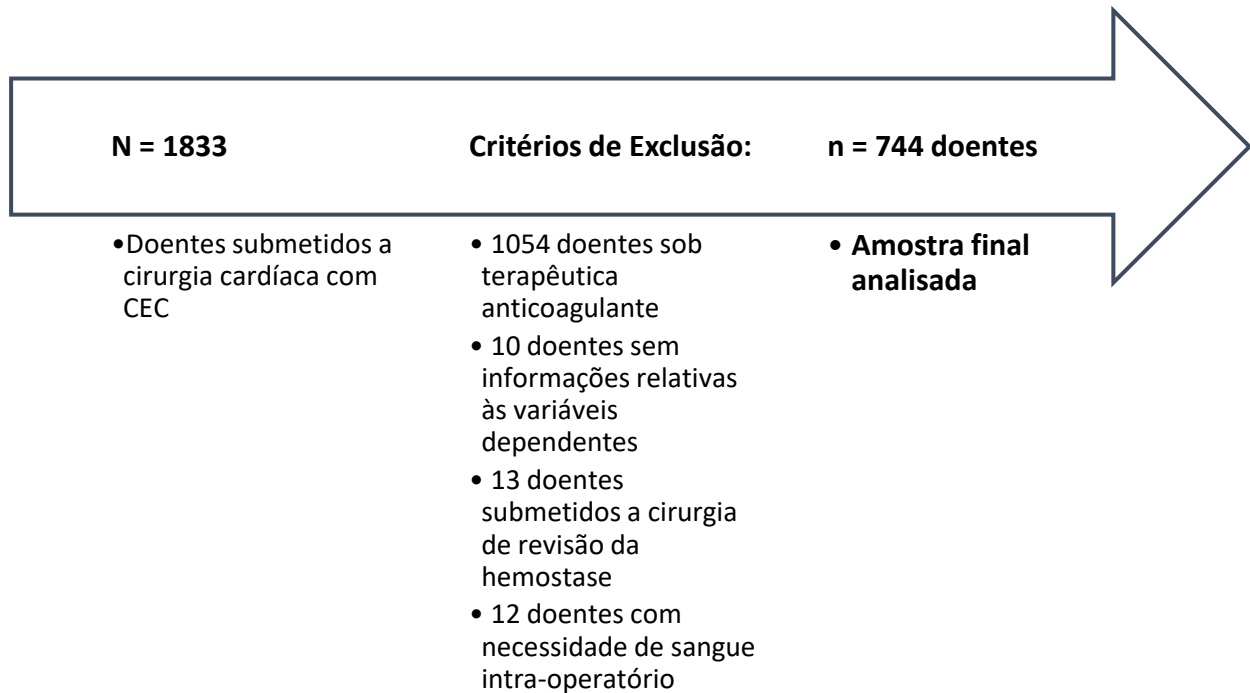
treino e as restantes para a base de dados de teste, com a aleatorização estratificada, opta-se por seleccionar uma proporção específica de observações de cada classe (da variável de interesse) para a base de dados de treino e de teste [54], [55].

4.1 População Estudada

Trata-se de um estudo observacional, retrospectivo, realizado num único centro de cirurgia cardíaca e de um hospital terciário e universitário da região de Grande Lisboa. Foi recolhida a informação de 1833 doentes submetidos a cirurgia cardíaca entre janeiro de 2018 e outubro de 2021. Os critérios de inclusão utilizados na recolha dos dados foram: (i) doentes maiores de idade (idade superior ou igual a 18 anos); e (ii) cirurgia cardíaca com necessidade de técnica de circulação extracorporeal (CEC). Foram excluídos os doentes: (i) com terapêutica anticoagulante em ambulatório; (ii) com complicação hemorrágica pós-operatória de causa cirúrgica - confirmada por revisão cirúrgica da hemostase; (iii) doentes sem registos clínicos referentes à hemorragia, (iv) com necessidade de aplicação intraoperatória de técnica de reaproveitamento sanguíneo - *Cell-Saver*, e (v) doentes com necessidade de transfusão sanguínea intraoperatória. Assim, e após a aplicação dos critérios acima referidos, procedeu-se à inclusão de 744 doentes, como mostra a Figura 1.

Obteve-se aprovação do Comité de Ética da Instituição onde se desenrolou o presente estudo (Referência N° 23/18) e o consentimento informado foi assinado por todos os doentes.

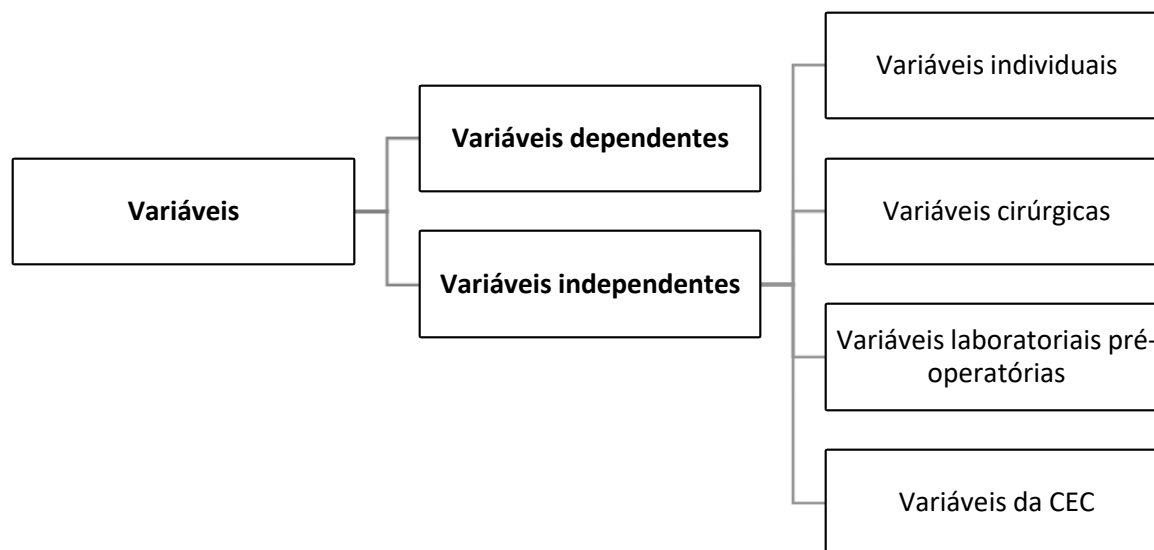
Figura 1 - Fluxograma de aplicação de critérios de inclusão e exclusão



4.2 Variáveis analisadas

Foi recolhida informação clínica pré-operatória dos doentes, tais como, dados demográficos, fatores de risco cardiovasculares, medicação e comorbilidades. A base de dados inicial é composta por 48 variáveis, as quais foram organizadas em dois grupos: variáveis dependentes e variáveis independentes. As variáveis independentes foram, por sua vez, reunidas em quatro grupos de acordo com a sua origem ou afetação: variáveis individuais, variáveis cirúrgicas, variáveis laboratoriais pré-operatórias e variáveis da CEC (Figura 2).

Figura 2 - Organização das variáveis da amostra



4.2.1 Variáveis dependentes

Nas seguintes tabelas estão listadas as variáveis que compõem cada um dos grupos e subgrupos de dados identificados. Pretende-se também identificar os tipos das variáveis e os respetivos valores ou unidades de medida, no caso de corresponder a uma variável contínua. A Tabela 1 contém todas as variáveis dependentes de interesse para o trabalho. Os dados relativos à hemorragia, em mililitros de sangue contabilizado nos drenos cirúrgicos, foram colhidos às 12 e 24 horas de pós-operatório, tendo sido exportadas através do *software Anaesthesia Manager® (Picis Clinical Solutions, Estados Unidos da América)*. Também foi daqui recolhida informação acerca da quantidade de uso de sangue e hemoderivados.

Tabela 1 - Variáveis dependentes

Nome da Variável	Descrição	Tipo	Valores
<i>drenagem_12h</i>	Contabilização das drenagens torácicas às 12 horas de pós-operatório	Contínua	ml
<i>drenagem_12h_ASC</i>	Contabilização das drenagens torácicas às 12 horas de pós-operatório indexada à área de superfície corporal	Contínua*	ml/m ²
<i>drenagem_12h_catg5</i>	Categorização da hemorragia às 12 horas de pós-operatório	Categórica*	1 = "Insignificante" 2 = "Ligeira" 3 = "Moderada" 4 = "Grave" 5 = "Extrema"

Tabela 1 - Variáveis dependentes

Nome da Variável	Descrição	Tipo	Valores
<i>drenagem_12h_Peso</i>	Contabilização das drenagens torácicas às 12 horas de pós-operatório indexada ao peso	Contínua*	ml/Kg
<i>drenagem_24h</i>	Contabilização das drenagens torácicas às 24 horas de pós-operatório	Contínua	ml
<i>drenagem_24h_ASC</i>	Contabilização das drenagens torácicas às 24 horas de pós-operatório indexada à área de superfície corporal	Contínua*	ml/m2
<i>drenagem_24h_Peso</i>	Contabilização das drenagens torácicas às 24 horas de pós-operatório indexada ao peso	Contínua*	ml/Kg
<i>drenagem_sig</i>	Categorização binária da hemorragia às 12 horas de pós-operatório	Categórica*	0 = "Não significativa" 1 = "Significativa"
<i>UCE</i>	Quantidade de Unidades de Concentrado Eritrocitários transfundidos no pós-operatório	Discreta	Unidades
<i>UCE_uso</i>	Necessidade de transfusão pós-operatória com Concentrado Eritrocitário	Categórica*	0 = "Não" 1 = "Sim"

*variáveis criadas a partir de outras variáveis existentes

A medição da hemorragia, dada pelo nível de sangue nos reservatórios de drenagens torácicas dos doentes, foi categorizado segundo os critérios sugeridos por Dyke *et al.* na definição universal de hemorragia perioperatória em cirurgia cardíaca em adultos [56]. Estes autores categorizam a hemorragia em cinco graus de gravidade de acordo com a classificação representada na Tabela 2. Assim, a variável *drenagem_12h* foi transformada em *drenagem_12h_catg5* respeitando os critérios anteriormente definidos. Para além disso, foi criada uma segunda variável, a partir desta, e que corresponde à transformação de *drenagem_12h_catg5* em apenas duas categorias: (i) drenagem não significativa - todas aquelas que apresentam uma hemorragia insignificante; e (ii) drenagem significativa - aglutinação de todos os restantes graus de hemorragia (de ligeira a extrema). A esta nova variável deu-se o nome de *drenagem_sig*. Ainda a partir das variáveis *drenagem_12h* e *drenagem_24h* foram criadas outras duas novas variáveis, para cada uma, através da respetiva indexação à área de superfície corporal ou apenas ao peso dos doentes. Trata-se das variáveis, *drenagem_12h_ASC*, *drenagem_12h_Peso*, *drenagem_24h_ASC* e *drenagem_24h_Peso*, respetivamente.

Ainda dentro das variáveis dependentes, mas relacionada com a necessidade de transfusão, foi criada uma variável, *UCE_uso*, a partir da categorização binária do número de UCE transfundidas. Assim, obtiveram-se as seguintes categorias: (i) Sim - caso o doente tivesse

sido transfundido com pelo menos uma UCE; e (ii) Não - caso o doente não tivesse sido transfundido de todo.

Tabela 2 - Classificação categórica da hemorragia às 12 horas de pós-operatório utilizada

Grau de hemorragia	Drenagem pós-operatória às 12 horas (mL)
<i>insignificante</i>	< 600
<i>ligeira</i>	601 - 800
<i>moderada</i>	801 - 1000
<i>grave</i>	1001 - 2000
<i>extrema</i>	> 2000

adaptado de: Dyke et al. [56]

4.2.2 Variáveis individuais

Quanto às variáveis independentes, ou explicativas, tem-se as variáveis individuais, representadas na Tabela 3 que são variáveis relacionadas com características individuais e demográficas de cada indivíduo incluído na amostra. No presente trabalho foram recolhidas 18 variáveis individuais das quais 12 são do tipo categóricas e 6 contínuas.

Tabela 3 - Variáveis individuais da amostra

Nome da Variável	Descrição	Tipo	Valores
<i>ID</i>	Identificação do indivíduo	Categórica	
<i>sexo</i>	Sexo do indivíduo	Categórica	1 = "Masculino" 2 = "Feminino"
<i>idade</i>	Idade do indivíduo	Contínua	anos
<i>Peso</i>	Peso do indivíduo	Contínua	Kg
<i>Altura</i>	Altura do indivíduo	Contínua	cm
<i>IMC</i>	Índice de Massa Corporal do indivíduo	Contínua	Kg/m ²
<i>IMC_catg</i>	Índice de Massa Corporal do indivíduo	Categórica*	1 = "Baixo peso" 2 = "Peso normal" 3 = "Excesso de peso" 4 = "Obesidade"
<i>IMC_catg2</i>	Índice de Massa Corporal modificado	Categórica*	1 = "Baixo peso" ou "Peso normal" 2 = "Excesso de peso" 3 = "Obesidade"
<i>ASC</i>	Área de Superfície Corporal do indivíduo	Contínua*	m ²

Tabela 3 - Variáveis individuais da amostra

Nome da Variável	Descrição	Tipo	Valores
<i>diabetes</i>	Historial clínico de diabetes do indivíduo	Categórica	0 = "Sem Diabetes" 1 = "Diabetes NIT" 2 = "Diabetes IT"
<i>dislipidemia</i>	Historial clínico de dislipidemia do indivíduo	Categórica	0 = "Não" 1 = "Sim"
<i>hta</i>	Historial clínico de hipertensão arterial do indivíduo	Categórica	0 = "Não" 1 = "Sim"
<i>fumador</i>	Hábitos tabágicos	Categórica	0 = "Não Fumador" 1 = "Ex-fumador" 2 = "Fumador"
<i>enfarte</i>	Antecedentes clínicos de enfarte agudo do miocárdio	Categórica	0 = "Não" 1 = "Sim"
<i>dvp</i>	Antecedentes clínicos de doença venosa periférica	Categórica	0 = "Não" 1 = "Sim"
<i>insuf_renal</i>	Antecedentes clínicos de insuficiência renal crônica	Categórica	0 = "Sem IRC" 1 = "IRC Ligeira" 2 = "IRC Moderada" 3 = "IRC Grave" 4 = "IRC em Hemodiálise"
<i>feve</i>	Função ventricular esquerda	Categórica	1 = "Normal" 2 = "Ligeiramente diminuída" 3 = "Moderadamente diminuída" 4 = "Diminuída" 5 = "Bastante diminuída"
<i>volemia</i>	Volémia estimada do indivíduo	Contínua*	ml

*variáveis criadas a partir de outras variáveis existentes

Através da variável relativa ao índice de massa corporal do indivíduo (*IMC*), foi criada uma variável categórica relativa à definição de excesso de peso e de obesidade de acordo com a classificação do Centro para o Controle e Prevenção de Doenças (CDC, do inglês, *Centers for Disease Control and Prevention*) [57]. As categorias e respectivos intervalos de valores estão listado na Tabela 4 abaixo. A esta variável deu-se o nome de *IMC_catg*. Foram posteriormente aglutinadas as categorias de baixo peso e de peso normal, passando a existir apenas três categorias de resposta (baixo peso ou normal; excesso de peso; obesidade) - variável *IMC_catg2*.

Tabela 4 - Classificação categórica do índice de massa corporal

Grau	IMC (Kg/m²)
<i>Baixo peso</i>	< 18.5
<i>Normal</i>	18.5 - < 25
<i>Excesso de peso</i>	25 - <30
<i>Obesidade</i>	> 30

adaptado de: CDC [57]

A área de superfície corporal dos indivíduos foi calculada através da fórmula de cálculo simplificado de *Mosteller* [58]. Esta variável foi criada essencialmente com o intuito de indexar o volume de hemorragia, conforme explicado no ponto anterior. Quanto ao cálculo da volémia, este foi estimado para cada indivíduo pela proporção de 70 mL/Kg [59].

4.2.3 Variáveis cirúrgicas

As variáveis cirúrgicas, representadas na Tabela 5 são variáveis que estão relacionadas com o procedimento cirúrgico. No presente trabalho foram recolhidas 3 variáveis do tipo cirúrgicas, sendo que todas são categóricas. As informações relacionadas com o procedimento cirúrgico (Tabela 5), bem como, as relacionadas com a CEC (Tabela 7) foram recolhidas através da base de dados concebida pelo sistema informático de registos existente no serviço (*CardioBase*®).

Tabela 5 - Variáveis cirúrgicas da amostra

Nome da Variável	Descrição	Tipo	Valores
<i>cirurgia_aorta</i>	Procedimento cirúrgico que envolve a substituição da aorta ascendente	Categórica	0 = "Não" 1 = "Sim"
<i>peso_cirurgia</i>	Peso efetivo da cirurgia	Categórica*	1 = "Cirurgia Coronária" 2 = "1 procedimento não-Coronário" 3 = "2 procedimentos" 4 = "3 procedimentos"
<i>urgencia</i>	Grau de urgência do procedimento	Categórica	1 = "Programada" 2 = "Urgente"

* variáveis criadas a partir de outras variáveis existentes

A variável *peso_cirurgia* resulta da conversão dos procedimentos efetuados nos registos cirúrgicos de acordo com o *score* de risco pré-operatório *EuroSCORE II* [60].

4.2.4 Variáveis Laboratoriais

Foram recolhidos dados laboratoriais pré-operatórios (Tabela 6) de parâmetros relacionados com o hemograma e coagulação. Os dados analíticos foram todos processados no mesmo laboratório, o do Serviço de Patologia Clínica da instituição.

Tabela 6 - Variáveis laboratoriais pré-operatórias

Nome da Variável	Descrição	Tipo	Valores
<i>Hb_Pre</i>	Valor de hemoglobina no pré-operatório	Contínua	g/dL
<i>Htc_Pre</i>	Valor de hematócrito no pré-operatório	Contínua	%
<i>Fib_Pre</i>	Valor de fibrinogénio no pré-operatório	Contínua	mg/dL
<i>Plaquetas_Pre</i>	Valor de plaquetas no pré-operatório	Contínua	$\times 10^9/L$
<i>TP_Pre</i>	Tempo de Protrombina no pré-operatório	Contínua	seg.
<i>INR_Pre</i>	Valor de INR no pré-operatório	Contínua	n.a.
<i>aPTT_Pre</i>	Tempo de tromboplastina ativado no pré-operatório	Contínua	seg.

n.a. - Não aplicável

INR - Relação Internacional Normalizada do TP (do inglês, *International Normalized Ratio*)

4.2.5 Variáveis da CEC

Por fim, foram estudadas as variáveis da CEC, que se encontram resumidas na Tabela 7. As mesmas estão relacionadas com a preparação ou realização da técnica supramencionada. Foram incluídas 10 variáveis deste tipo, sendo que 3 são categóricas e 7 contínuas.

Tabela 7 - Variáveis da CEC

Nome da Variável	Descrição	Tipo	Valores
<i>ACT_Basal</i>	Tempo de coagulação ativado antes da administração de heparina	Contínua	segundos
<i>ACT_Heparina</i>	Tempo de coagulação ativado após a administração de heparina	Contínua	segundos
<i>ACT_Final</i>	Tempo de coagulação ativado após a administração de protamina	Contínua	segundos

Tabela 7 - Variáveis da CEC

Nome da Variável	Descrição	Tipo	Valores
<i>Tempo_CEC</i>	Tempo de circulação extracorporeal	Contínua	minutos
<i>Tempo_Clamp</i>	Tempo de clampagem da aorta	Contínua	minutos
<i>Volume_Priming</i>	Volume de <i>priming</i> do circuito à entrada de CEC	Contínua	mililitros
<i>Gelatina</i>	Utilização de Gelatina durante a CEC	Catagórica	0 = "Sem Gelatina" 1 = "Com Gelatina"
<i>Heparina_CEC</i>	Necessidade de reforço de Heparina durante a CEC	Catagórica	0 = "Não" 1 = "Sim"
<i>TaxaDiluicao</i>	Taxa de diluição provocada pela CEC	Contínua*	percentagem
<i>ACT_480</i>	Avaliação subjetiva da sensibilidade à heparina	Catagórica*	0 = "ACT após heparina < 480 segundos" 1 = "ACT após heparina ≥ 480 segundos"

Foi calculada a taxa de diluição provocada pelo *priming* do circuito de CEC, correspondendo à variável *TaxaDiluicao* da Tabela 7. Esta variável foi calculada com recurso à fórmula transcrita na equação 4.1 e que foi utilizada nos trabalhos de Muraki *et al.* [59].

$$TaxaDiluicao = \frac{Volume_Priming}{Volume_Priming + Volemia} \quad (4.1)$$

Por fim, foi convertida a variável *ACT_Heparina* que traduz o valor de ACT obtido após a administração da Heparina, numa variável catagórica que pretende avaliar, de forma subjetiva, a eventual sensibilidade ou resistência, do doente à heparina. A esta variável deu-se o nome de *ACT_480*, pois categorizou-se *ACT_Heparina* da seguinte forma: (i) ACT após heparina inferior a 480 segundos; e (ii) ACT após heparina superior ou igual a 480 segundos. Utilizou-se o *cutoff* de ACT após heparina igual aos 480 segundos de acordo com as recomendações do *European Board of Cardiovascular Perfusion* [13].

4.3 Procedimento Perioperatório

Todas as cirurgias foram realizadas seguindo o mesmo protocolo anestésico e cirúrgico, ajustado em caso de necessidade pontual e dependendo das exigências particulares de cada

doente, de acordo com a prática clínica do Serviço. O protocolo de anestesia incluiu fentanil e propofol, sendo a intubação orotraqueal facilitada com brometo de rocurônio. A manutenção da anestesia foi realizada totalmente por via intravenosa, mantendo a perfusão de propofol através do cateter venoso central colocado percutaneamente via veia jugular interna direita. A monitorização da pressão arterial invasiva foi realizada através de cateter na artéria radial esquerda. Os circuitos de CEC utilizados incluíam os dois oxigenadores presentes no serviço [Inspire 6® (Sorin Group, Itália) ou Quadrox-i® (Getinge, Suécia)] com reservatório único e rígido com filtro de cardiectomia, filtro arterial integrado e respetivos tubos do circuito em policloreto de vinil. O *priming* do circuito foi realizado com aproximadamente 700 mililitros de solução eletrolítica balanceada (Polieletrólítico) com ou sem a adição de 500 mililitros de Gelatina (GelaFundina a 4%), no grupo de estudo e com apenas 1200 mL de Polieletrólítico no grupo controlo. Para iniciação da CEC os doentes foram anticoagulados com heparina sódica à dose de 300 UI/Kg com o objetivo de obter um ACT-alvo superior a 480 segundos. Foi utilizada CEC com fluxo não-pulsátil em todos os doentes com um fluxo indexado à superfície corporal dos doentes de 2.4 L/min/m². O controlo da temperatura foi realizado com o objetivo de manter o doente entre uma hipotermia ligeira a normotermia (32 a 36 °C, medidos na orofaringe). Para proteção miocárdica foi utilizada cardioplegia hipotérmica do tipo sanguínea por via anterógrada para indução ou manutenção da assistolia e normotérmica para a reperfusão, diluída com uma solução (*Buckberg*) hipercaliémica à proporção de 4:1. A glicémia dos doentes no perioperatório foi mantida abaixo de 250 mg/dL. O valor mínimo permitido de hematócrito, durante a CEC, foi de 24%. Em caso de hipotensão durante a CEC, o tratamento de primeira-linha foi o incremento do fluxo indexado à superfície corporal para 2.6 L/min/m². Em caso de persistência da hipotensão arterial, foi administrado bólus único de noradrenalina, seguida de iniciação de uma perfusão intravenosa.

Após término da CEC, a heparina foi revertida com sulfato de protamina à relação de 1:1. A drenagem do sangue mediastínico no pós-operatório foi realizada por intermédio de pelo menos duas drenagens, uma no mediastino anterior e outra no pericárdio, e com tubos de drenagem de silicone com calibre 28 French. Os drenos foram ambos conectados a um balão de drenagem sob sifonagem, com soro, e conectados sob forma contínua ao sistema de vácuo. Todos os doentes, foram transferidos para a unidade de cuidados intensivos a seguir à cirurgia e permaneceram com as drenagens mediastínicas durante, pelo menos, 24 horas.

4.4 Dados omissos

Verificou-se que não existem dados omissos para todas as categorias de variáveis recolhidas. Trata-se de uma constatação deveras positiva.

Orchard e Woodbury afirmariam na década de 70 que "*Obviously the best way to treat missing data is not to have them.*"[61].

4.5 Análise Estatística

Não foi executado qualquer tipo de cálculo de dimensão da amostra, limitando-se este trabalho a apresentar apenas os dados disponíveis. As variáveis contínuas são expressas sob a forma de mediana e intervalo interquartil (IQR, do inglês *Interquartile Range*). As variáveis categóricas são apresentadas sob a forma de frequência absoluta e respetivas percentagens.

Para comparar as medianas das variáveis quantitativas entre os dois grupos, fossem eles associados ao tipo de *priming* (com vs. sem gelatina) ou relacionados com o *outcome* hemorrágico (hemorragia significativa vs. não significativa) foi utilizado o teste de *Mann-Whitney U*. Por outro lado, para estudar a associação entre as variáveis categóricas e as variáveis dependentes utilizou-se o teste de qui-quadrado com correção de continuidade de *Yates* ou o teste exato de *Fisher*, de acordo com a homogeneidade de distribuição por classes. Todos os testes realizados foram bilaterais.

Por forma a analisar a associação entre diversas variáveis potencialmente explicativas com a variável dependente (em termos de probabilidades preditas), foi executado um Modelo de Regressão Binária Multivariada. Por se tratar de uma variável dependente com uma probabilidade de sucesso baixa, optou-se por ajustar um modelo de regressão complementar log-log. Foram estimados três modelos principais utilizando: (i) base de dados original; (ii) uma base de dados parcialmente balanceada; e (iii) uma base de dados completamente balanceada. Para o balanceamento das bases de dados foi utilizada a técnica de *oversampling* a diferentes percentagens de balanceamento, 20 e 50%, respetivamente. Para cada um destes modelos foi aplicada a técnica de *stepwise* em ambos os sentidos com o objetivo de apenas incluir as variáveis que reduzissem o AIC ao seu menor valor. Os modelos elaborados foram avaliados do ponto de vista da sua especificidade e sensibilidade com recurso a curvas de ROC

através do cálculo da AUC. Para além disso, utilizou-se a validação cruzada, tendo sido, cada base de dados anteriormente mencionada, dividida em duas partes: (i) dados de treino - correspondentes a 80% da respetiva amostra, e que serviram para treinar os modelos de regressão; e (ii) dados de teste, correspondente aos restantes 20% da respetiva amostra e que serviram para a obtenção das métricas de desempenho dos modelos. A técnica de divisão da amostra utilizada foi a técnica de *Holdout*.

Para todas as análises que advêm destes dados, o intervalo de confiança selecionado foi de 95% e o nível de significância utilizado foi de 5%. Toda análise estatística deste trabalho foi feita com recurso ao software *R*, em ambiente *R-Studio* [62]. Os *packages* do *R* utilizados para as demais análises e construção de grafismos foram: *ggplot2* [63], *ggpubr* [64], *pROC* [65], *dplyr* [66], *tidyr* [67], *naniar* [68], *paletteer* [69], *ROSE* [53], *car* [70] e *caret* [71].

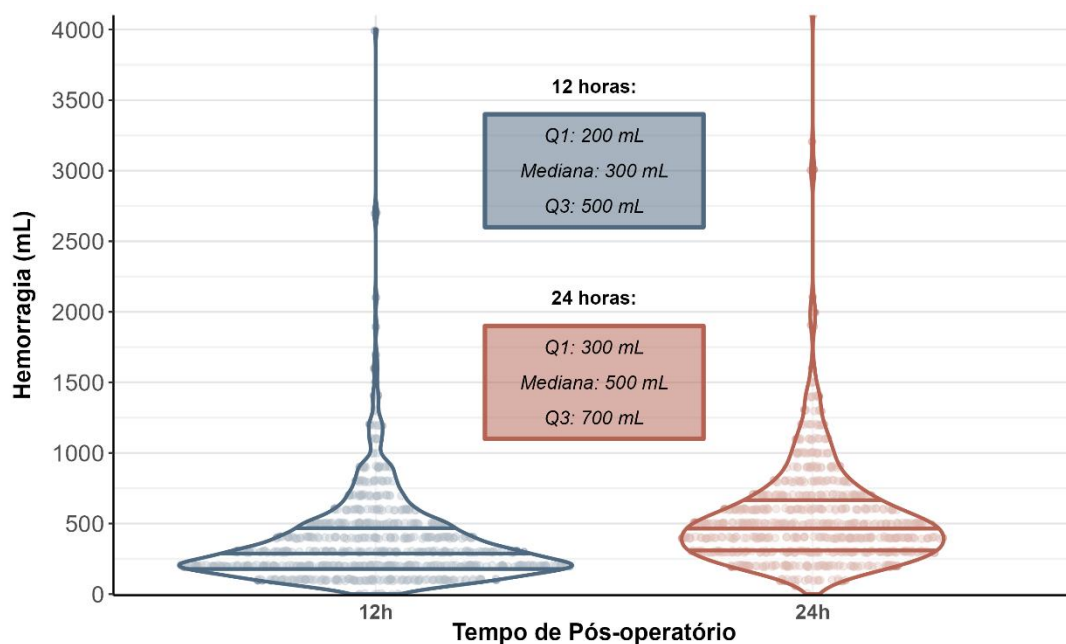
RESULTADOS

5.1 Análise Preliminar dos Dados

5.1.1 Variáveis dependentes

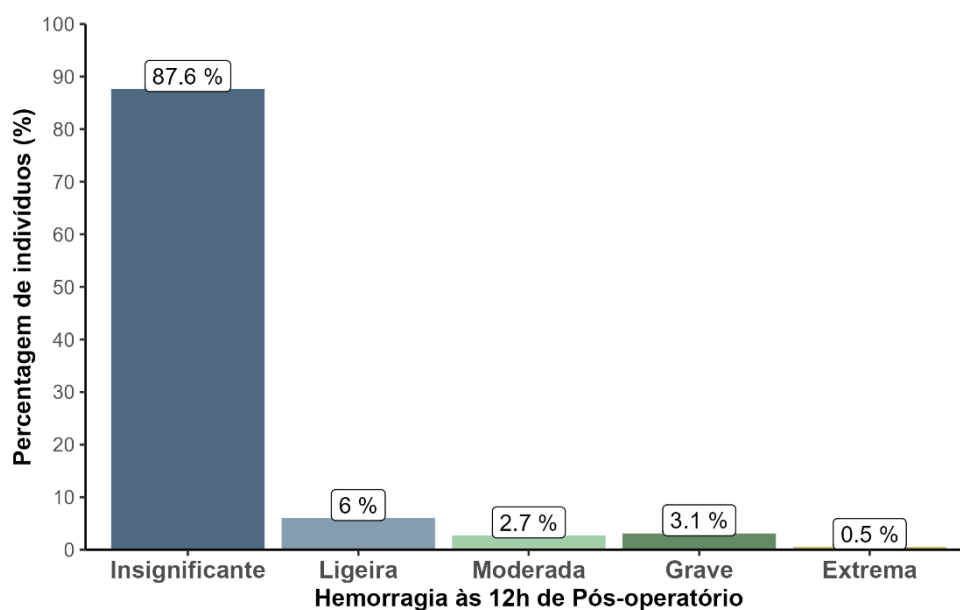
A Figura 3 mostra os diagramas de extremos e quartis no que diz respeito a uma das variáveis dependentes deste trabalho, ou seja, a hemorragia quantitativa às 12 e às 24 horas de pós-operatório, medidas nos reservatórios de drenagem torácicos. Como se pode verificar, às 12 horas de pós-operatório metade da amostra tem uma hemorragia até 300 ml enquanto às 24 horas metade da amostra tem uma hemorragia que totaliza 500 ml. Na figura estão descritas as seguintes estatísticas descritivas (mediana e primeiro e terceiro quartil). Por fim, verifica-se que apesar da maior concentração de pontos estarem em níveis reduzidos de hemorragia, existem ainda alguns pontos dispersos para valores superiores de hemorragia para ambas as unidades temporais.

Figura 3 - Variável dependente - Hemorragia no pós-operatório (12 e 24 horas)



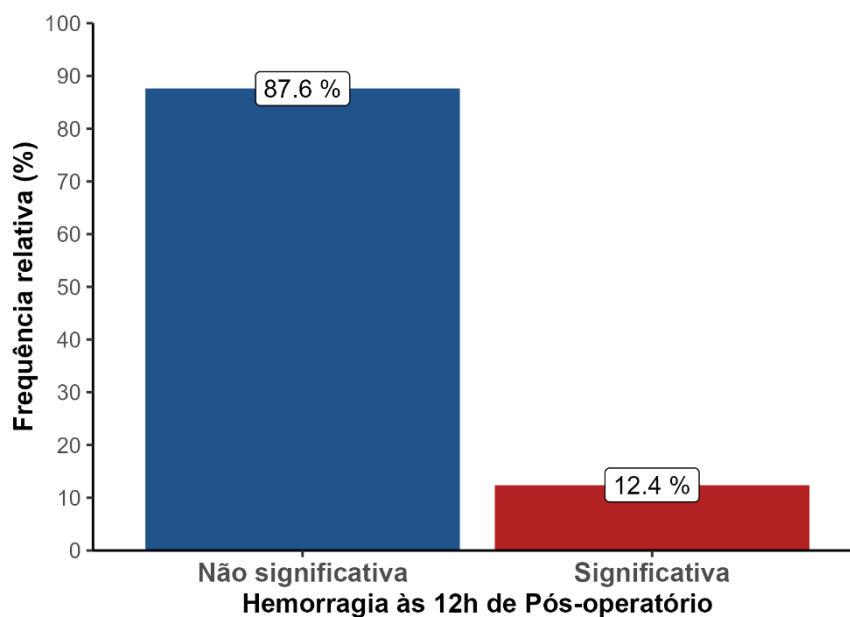
A Figura 4 representa as frequências relativas à categorização da hemorragia pós-operatória às 12 horas, de acordo com a classificação utilizada por Dyke *et al* [56]. Como se pode verificar, 87.6 % dos doentes tem uma hemorragia considerada insignificante.

Figura 4 - Variável dependente - Categorização da hemorragia às 12 horas de pós-operatório



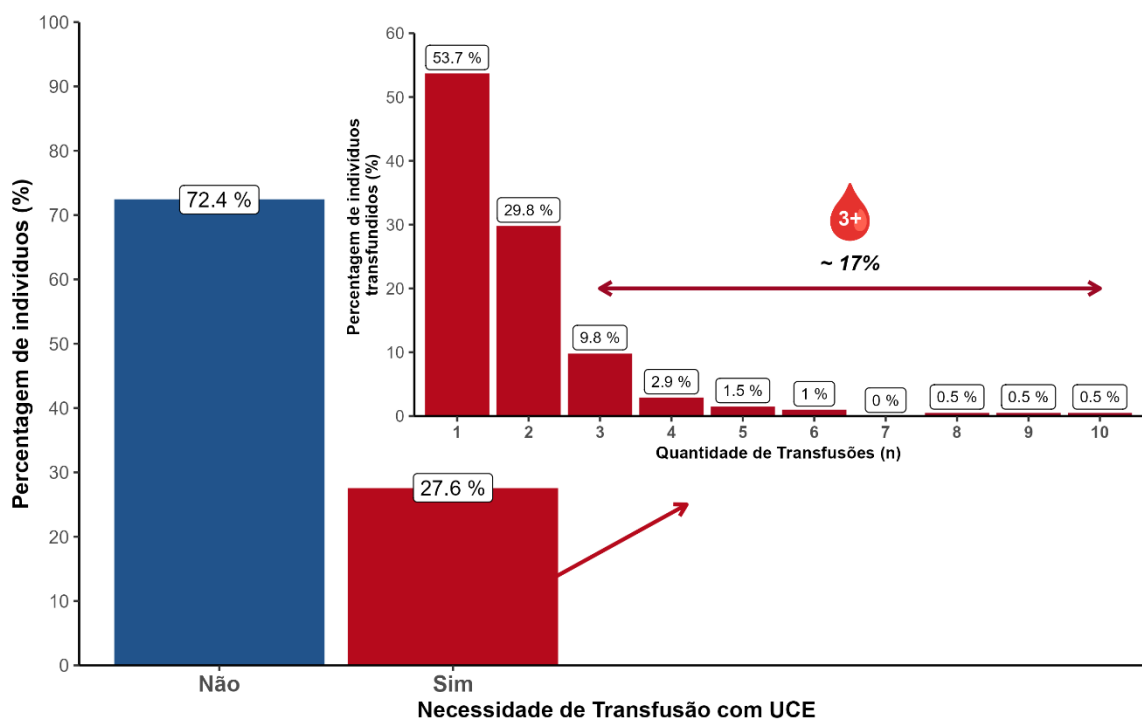
A Figura 5 representa a categorização binária da hemorragia às 12 horas de pós-operatório. A totalidade das quatro categorias de hemorragias significativas (desde ligeira a extrema) totalizam aproximadamente 12 %.

Figura 5 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório



Relativamente à necessidade de transfusão com Unidades de Concentrado Eritrocitário (UCE), a Figura 6 mostra que aproximadamente 28%, dos doentes submetidos a cirurgia, necessitaram de pelo menos uma transfusão. Quando analisados estes doentes, os transfundidos, verificou-se que metade necessitam de apenas uma UCE. Apesar disso, é verdadeiramente importante salientar que cerca de 17% dos doentes com necessidade de transfusão utilizaram três ou mais UCE.

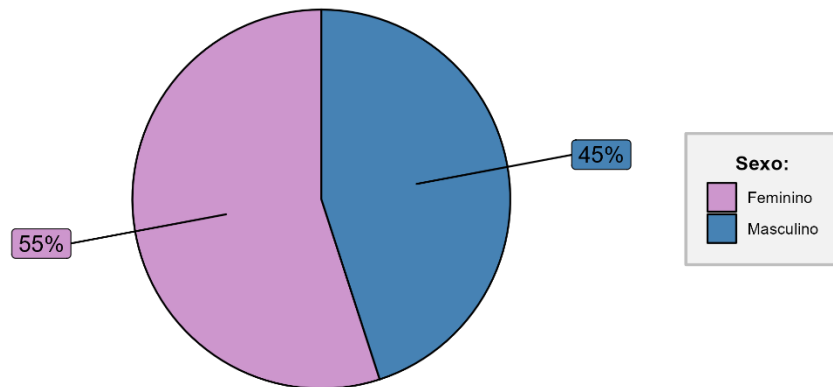
Figura 6 - Variável dependente - Necessidade de Transfusões com UCE e Quantidade de UCE transfundidas no pós-operatório



5.1.2 Variáveis individuais

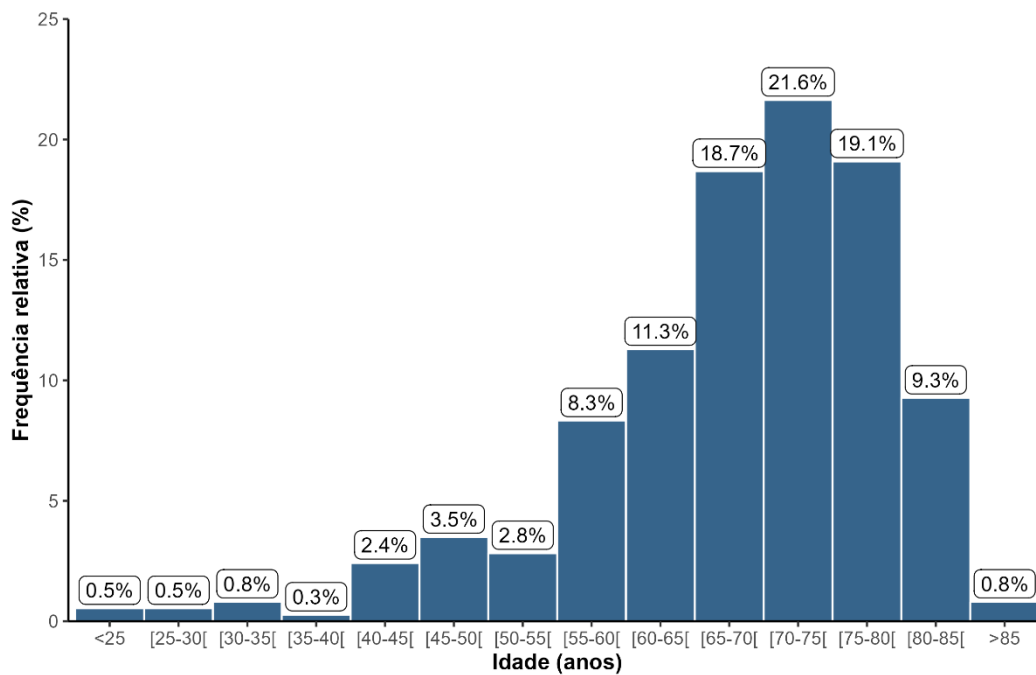
Na Figura 7, representa-se a distribuição por sexo dos indivíduos incluídos na amostra, onde se verifica que existe uma representação do sexo feminino ligeiramente maior em relação ao sexo masculino (55% e 45%, respetivamente).

Figura 7 - Variável individual - Sexo



A Figura 8 representa o histograma relativo à variável idade do indivíduo. Como se pode verificar a categoria maioritária é a de indivíduos com idades compreendidas entre 70 e 75 anos, representando 21.6% da amostra. Também se verifica que existe um desvio negativo, ou seja, existe uma longa cauda no início do gráfico, isto é, das categorias relativas às idades mais jovens, e frequências mais elevadas para categorias relativas a idades mais avançadas.

Figura 8 - Variável individual - Idade



A Figura 9 representa o histograma relativo à variável peso do indivíduo. Como se pode verificar a categoria maioritária é a de indivíduos com pesos compreendidas entre 70 e

80 quilogramas, representando 30.6% da amostra. Por outro lado, a Figura 10 representa a distribuição das alturas dos indivíduos da amostra, onde se verifica uma distribuição homogênea pela totalidade das categorias.

Figura 9 - Variável individual - Peso

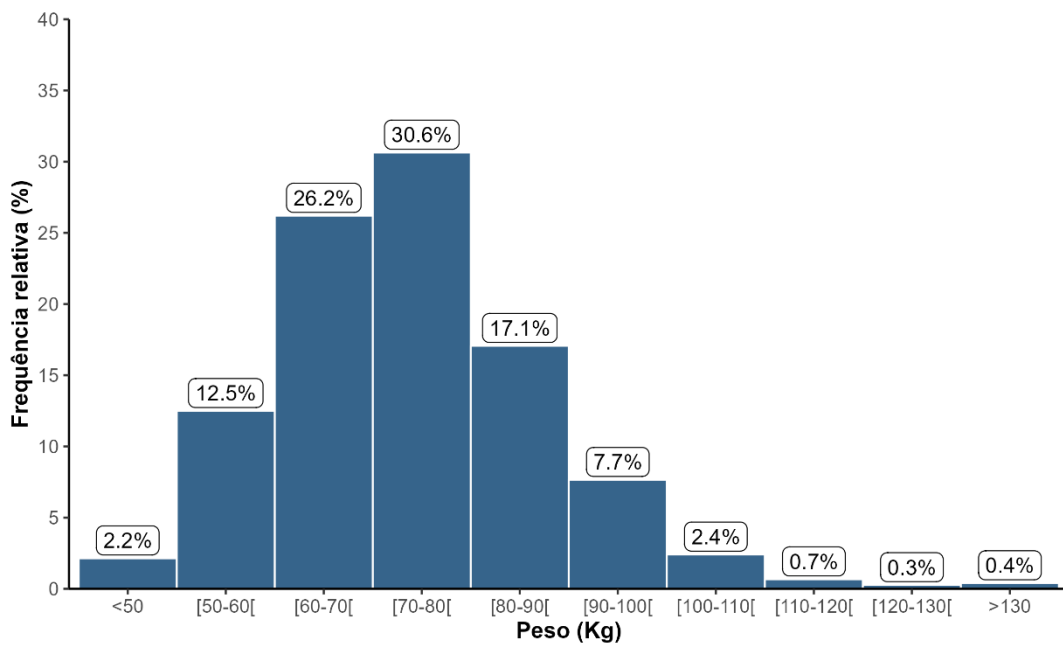
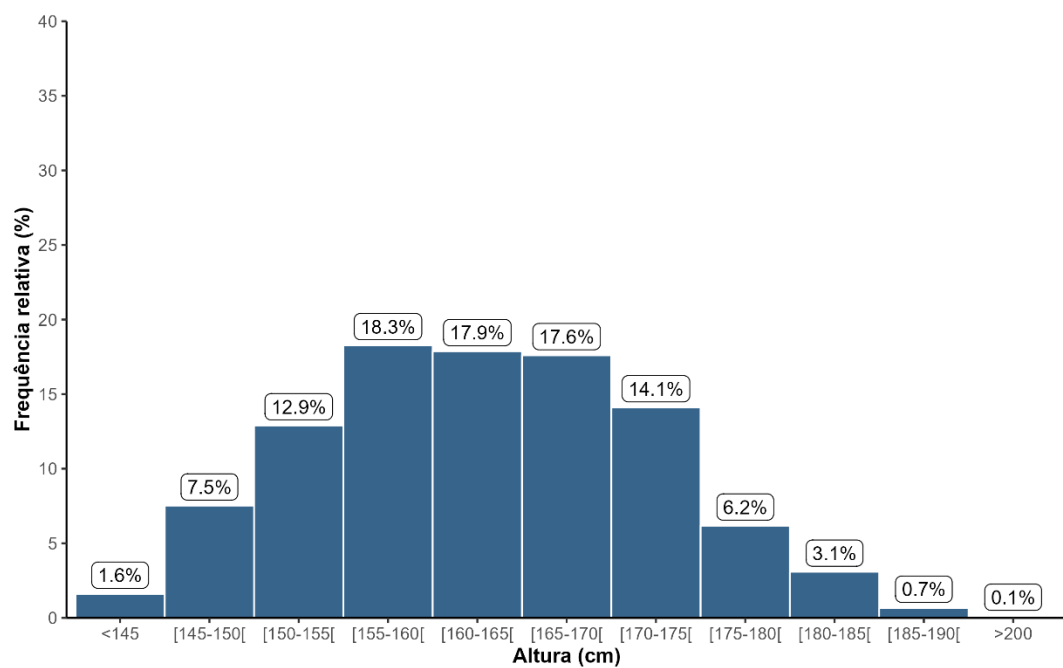
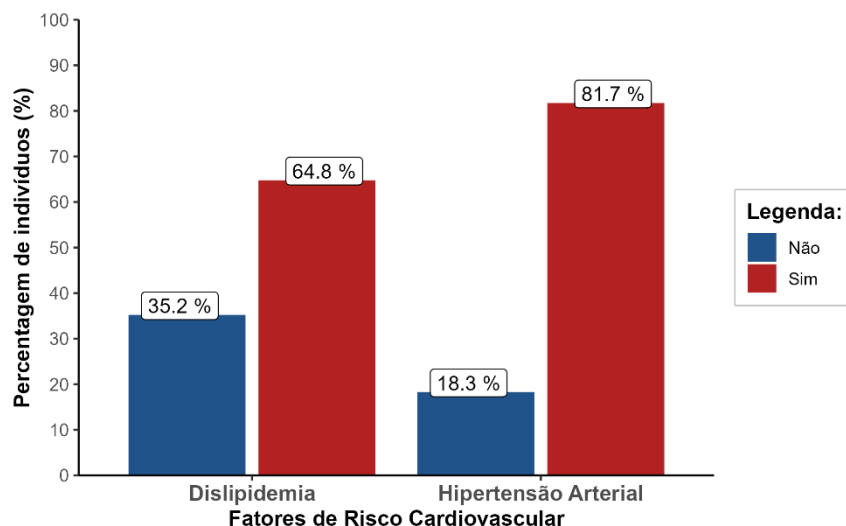


Figura 10 - Variável individual - Altura



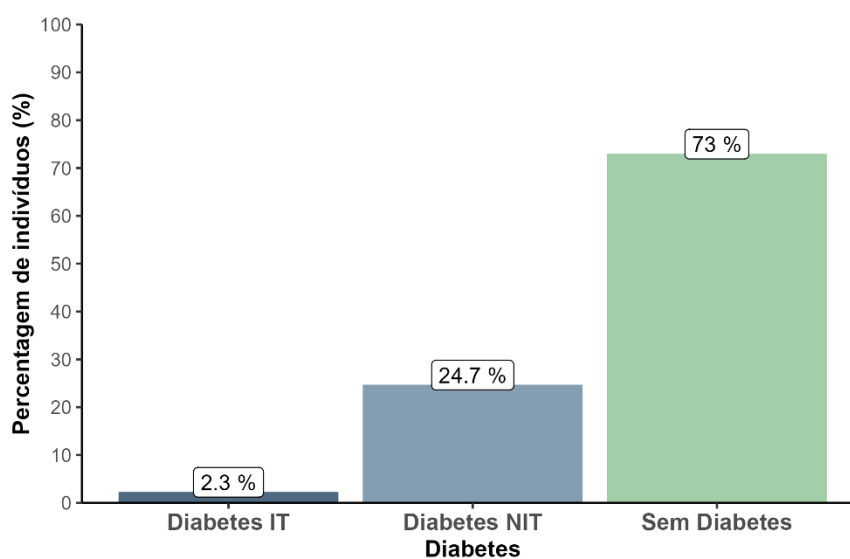
Como se verifica na Figura 11, 64.8% dos indivíduos apresenta como fator de risco dislipidemia. Para além disso, observa-se que 81.7% dos doentes têm Hipertensão Arterial.

Figura 11 - Variável individual - Dislipidemia e Hipertensão Arterial



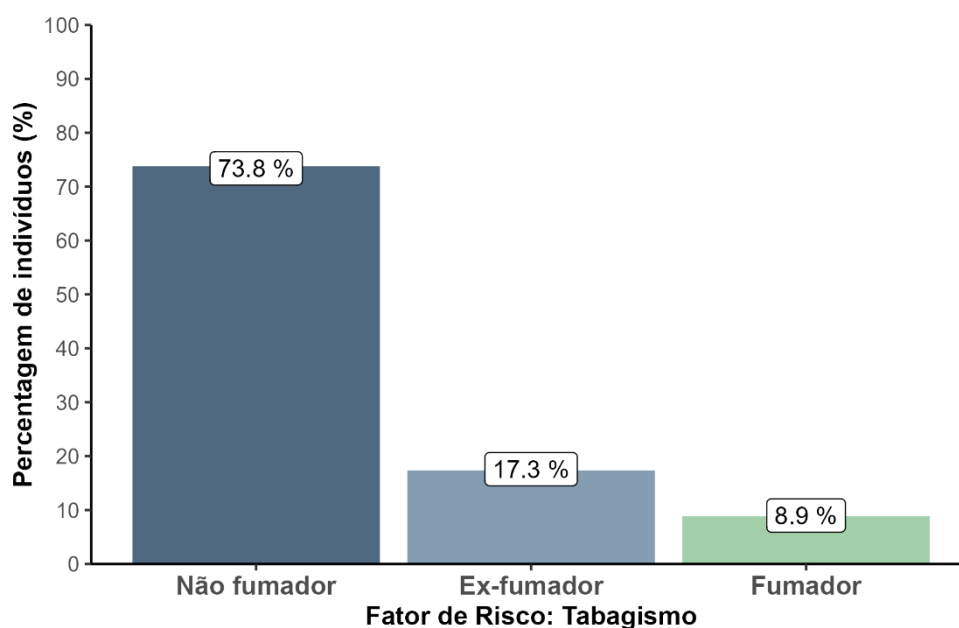
Na Figura 12 encontra-se a informação relativa à variável individual *diabetes*. A maioria dos indivíduos da amostra, 73%, não são diabéticos, enquanto os restantes 27% dos doentes apresentam diabetes não insulino-tratados ou insulino-tratados, 24,7% e 2,3%, respetivamente.

Figura 12 - Variável individual - Diabetes



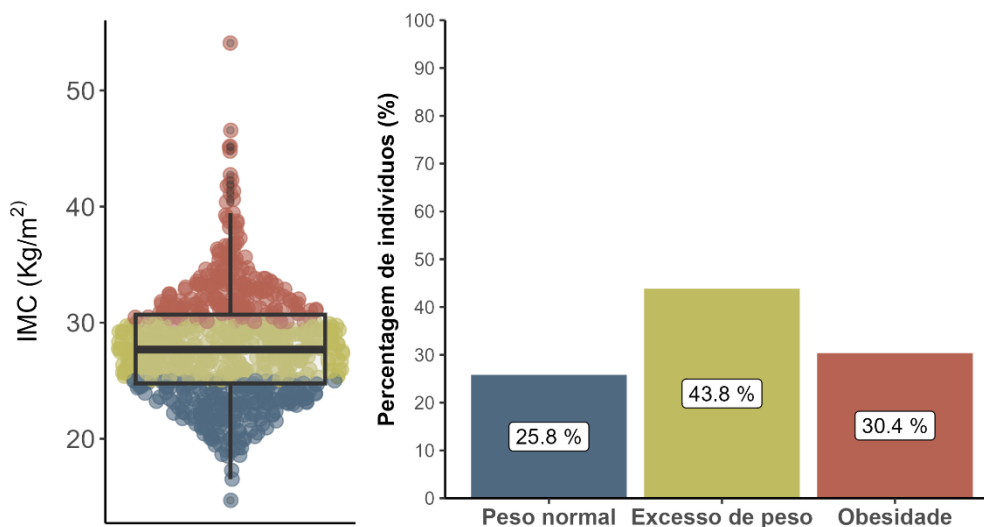
Na Figura 13 encontra-se a informação relativa à variável individual associada ao fator de risco Tabagismo. A maioria dos indivíduos da amostra, 73.8 %, não são, nem nunca foram fumadores. Para além disso, 17.3 % dos doentes são ex-fumadores e 8.9 % são fumadores.

Figura 13 - Variável individual - Tabagismo



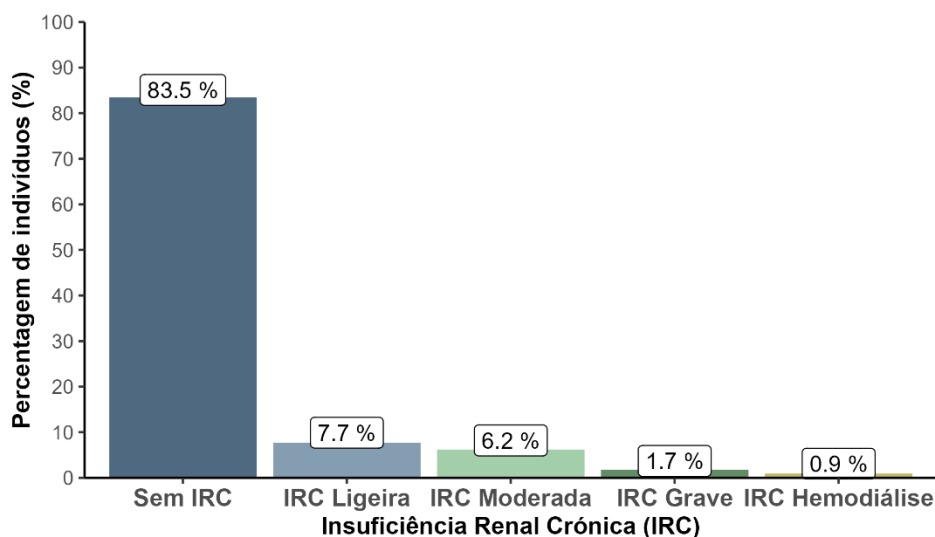
A Figura 14 contém a distribuição dos Índices de Massa Corporal (IMC) dos indivíduos sob a forma de diagrama de extremos e quartis, à esquerda, e a sua conversão para variável categórica e respetivas frequência, à direita. Verifica-se que a maioria da amostra apresenta excesso de peso como fator de risco cardiovascular (43.8 %). Esta observação também poderia ser corroborada pela análise visual da *boxplot* uma vez que a maioria dos pontos que estão contidos na caixa representativa dos 50% da amostra (do primeiro ao terceiro quartil) são pontos representativos desta categoria - Excesso de peso. Para além disso, apenas 25.8 % dos indivíduos da amostra tinham peso normal à data da cirurgia e 30.4 % apresentavam algum grau de obesidade.

Figura 14 - Variável individual - IMC e IMC por categoria



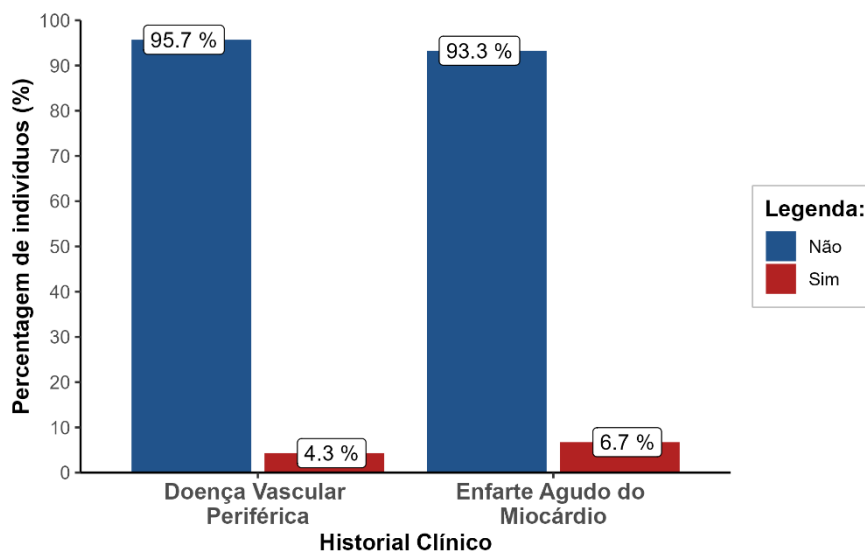
Ainda dentro das variáveis individuais, mas relativamente ao historial e antecedentes clínicos dos doentes, tem-se na Figura 15 a distribuição dos indivíduos relativamente aos graus de Insuficiência Renal Crónica (IRC). A maioria dos doentes, 83.5 % não tem qualquer grau de IRC, o que significa que os restantes 16.5 % dos doentes apresenta algum grau de IRC.

Figura 15 - Variável individual - Insuficiência Renal Crónica



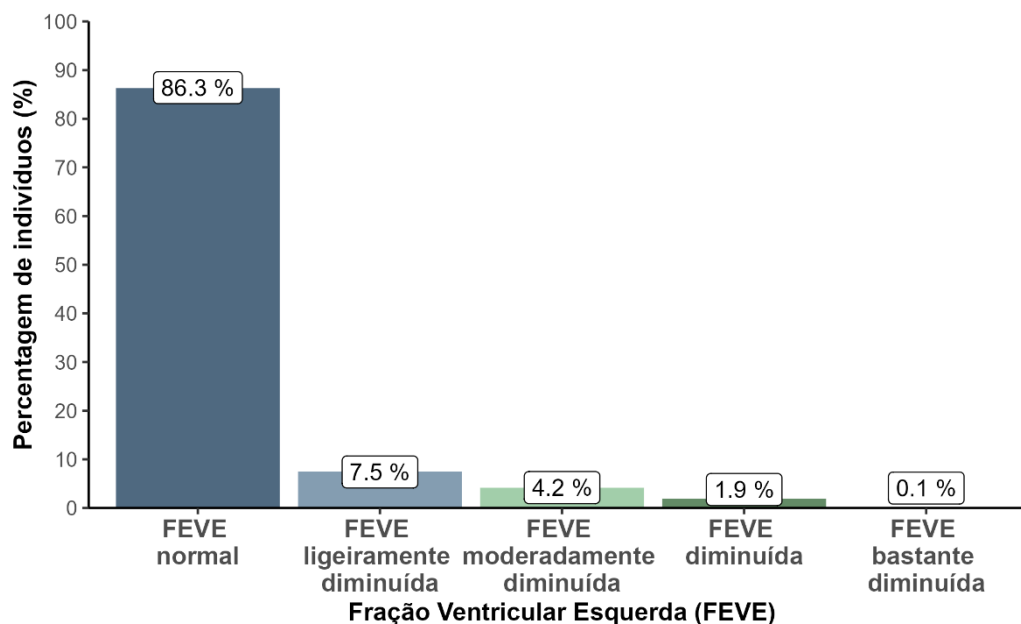
Relativamente a outros dois antecedentes clínicos, encontram-se esquematizadas na Figura 16 as frequências relativas de Enfarte Agudo do Miocárdio (EAM) e de Doença Vascular Periférica (DVP). Deste modo, verifica-se que 4.3% dos doentes apresenta DVP e 6.7% dos doentes já teve pelo menos um EAM.

Figura 16 - Variável individual - Enfarte e DVP



Por fim, na Figura 17 encontram-se as frequências da avaliação pré-operatória da função ventricular esquerda através da fração de ejeção do ventrículo esquerdo avaliada por ecocardiograma. A maioria dos doentes (86.3 %) apresenta-se à cirurgia com função ventricular esquerda normal.

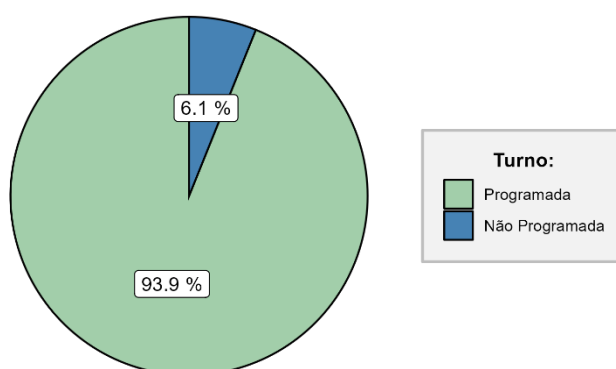
Figura 17 - Variável individual - Função Ventricular Esquerda



5.1.3 Variáveis cirúrgicas

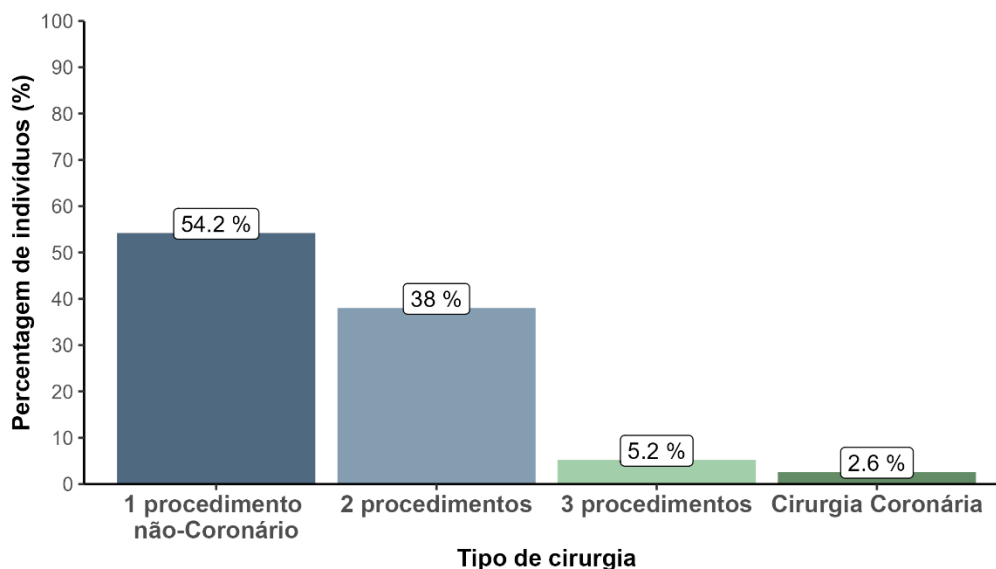
Relativamente à variável cirúrgica do grau de urgência do procedimento, verifica-se na Figura 18 que aproximadamente 94% das cirurgias foram realizadas eletivamente (programada), contendo a amostra apenas 6% de cirurgias urgentes e/ou emergentes (não programada).

Figura 18 - Variável cirúrgica - Grau de Urgência



Como se verifica na Figura 19, aproximadamente metade da amostra corresponde a uma cirurgia com apenas um procedimento que não a cirurgia de revascularização do miocárdio (54,2 %). O segundo grupo com maior representatividade na amostra corresponde a cirurgias que englobem 2 procedimentos distintos (38%).

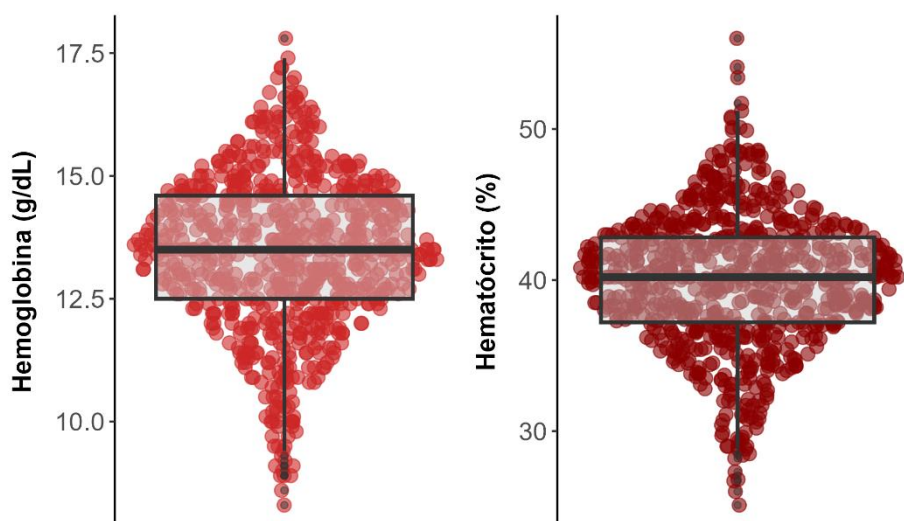
Figura 19 - Variável cirúrgica - Peso efetivo do tipo de cirurgia



5.1.4 Variáveis laboratoriais pré-operatórias

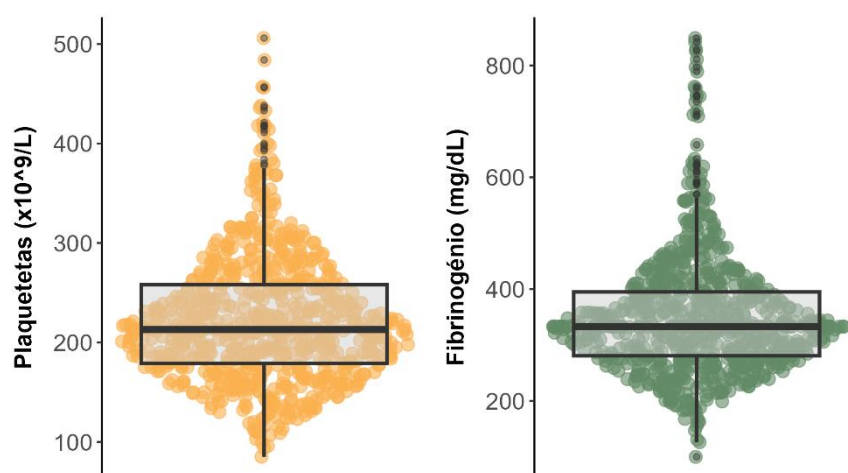
A Figura 20 representa os diagramas de extremos e quartis para os parâmetros do hemograma referentes à hemoglobina e ao hematócrito medidos no pré-operatório. Verifica-se que a mediana e intervalo interquartil para a hemoglobina (gráfico à esquerda) são de 13.5 (12.5 - 14.60) g/dl. Para além disso, verifica-se que estes valores de hemoglobina correspondem a valores de hematócritos (gráfico à direita) de 40.20 (37.20 - 42.90) %.

Figura 20 - Variável laboratorial pré-operatória - Hemoglobina e Hematócrito



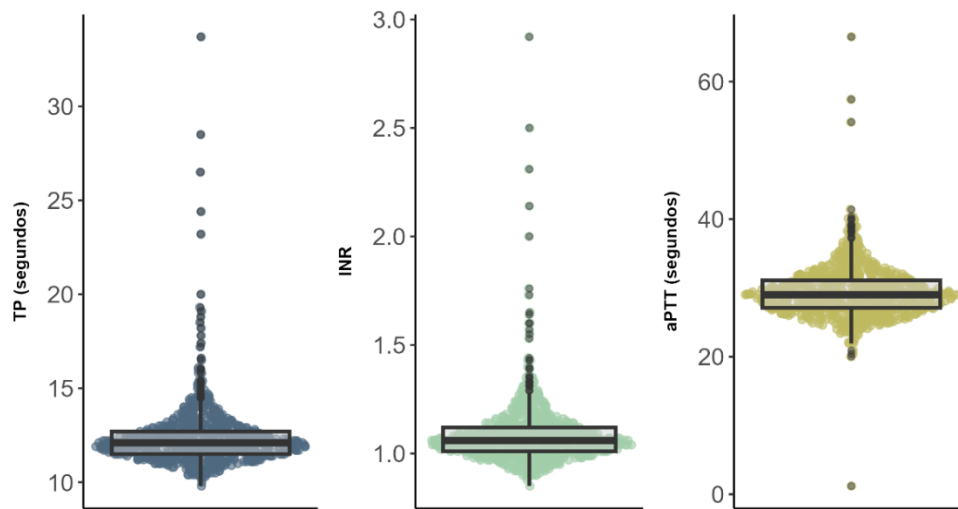
Para representar os valores analíticos de plaquetas e de fibrinogénio, ambos no pré-operatório, foi construída a Figura 21. A amostra apresenta contagens de plaquetas (gráfico à esquerda) com mediana e intervalo interquartil de 213 (178 - 258) $\times 10^9/L$. Quanto ao fibrinogénio (gráfico à direita), os valores obtidos foram de 333 (281 - 397) mg/dl. Verifica-se que para ambas as variáveis existem alguns doentes que apresentam valores que se comportam como *outliers* (pontos cinzentos) e que os mesmos apenas existem para valores superiores de cada uma das respetivas variáveis. Por outras palavras, não foram registados *outliers* para valores reduzidos de plaquetas ou de fibrinogénio.

Figura 21 - Variável laboratorial pré-operatória - Plaquetas e Fibrinogénio



Por fim, e por forma a explorar os valores laboratoriais pré-operatórios dos parâmetros analíticos da coagulação, tais como, o tempo de protrombina, INR e tempo de tromboplastina parcial ativada, foi elaborada a Figura 22. Os valores de mediana e intervalo interquartil para o tempo de protrombina foram 12.1 (11.5 - 12.8) segundos. Os valores de INR obtidos são iguais a 1.06 (1.01 - 1.13). Por fim, os valores de aPTT obtidos foram de 29 (27.1 - 31.1) segundos. Para as variáveis de TP e INR verifica-se a presença de valores *outliers* embora apenas para os valores superiores da variável. Quanto ao aPTT, existem alguns valores *outliers* tanto para valores superiores como inferiores da variável.

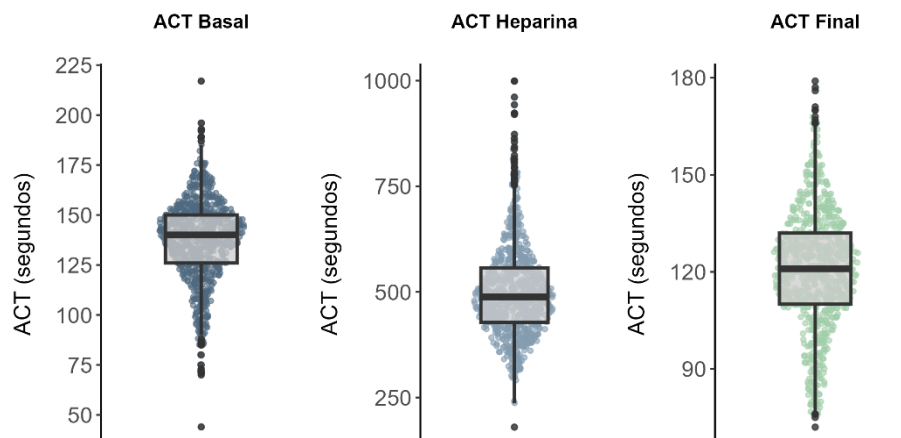
Figura 22 - Variável laboratorial pré-operatória - Tempo de Protrombina, INR e Tempo de Tromboplastina Parcial Ativado



5.1.5 Variáveis da CEC

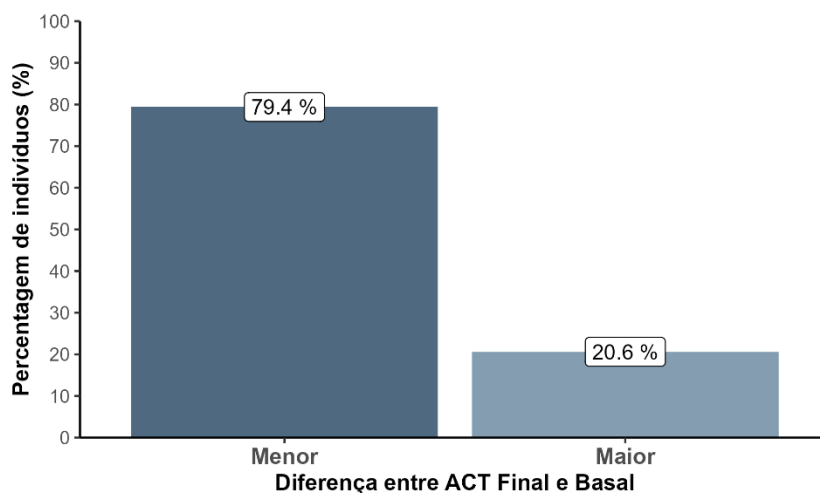
A Figura 23 apresenta os gráficos correspondentes à distribuição dos valores de ACT obtidos em três momentos distintos: (i) antes da administração da Heparina - ACT Basal; (ii) após a administração de Heparina - ACT Heparina; e (iii) Após a administração de protamina - ACT Final. Para estes três parâmetros de coagulação os valores obtidos para mediana e intervalo interquartil foram 140 (126 -150), 488 (428 - 558) e 121 (110 - 132) segundos, respectivamente. Para as três variáveis, verifica-se a presença de alguns valores que se comportam como *outliers*.

Figura 23 - Variável da CEC - ACT Basal, ACT após Heparina e ACT Final



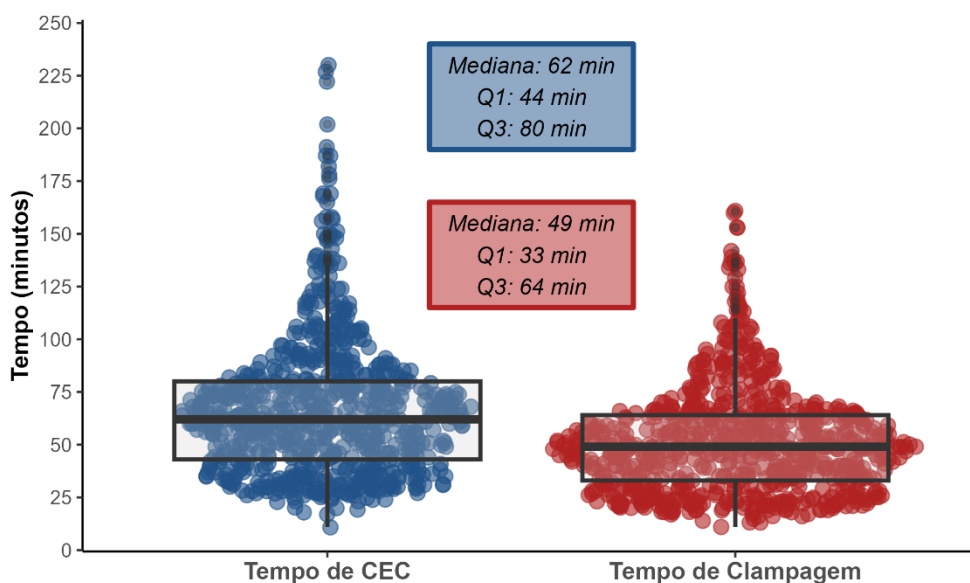
Por forma a apresentar as frequências relativas à variável da diferença categórica entre o ACT Final e o ACT Basal, elaborou-se a Figura 24. A maioria dos doentes (cerca de 80%) apresenta um ACT Final menor do que o ACT Basal, isto é, após a reversão com protamina o ACT é menor do que o Basal. Esta variável havia sido abordada num trabalho previamente publicado na revista *Interdisciplinary CardioVascular and Thoracic Surgery (ICVTS)* [72].

Figura 24 - Variável da CEC - Diferença Categórica entre o ACT Final e o ACT Basal



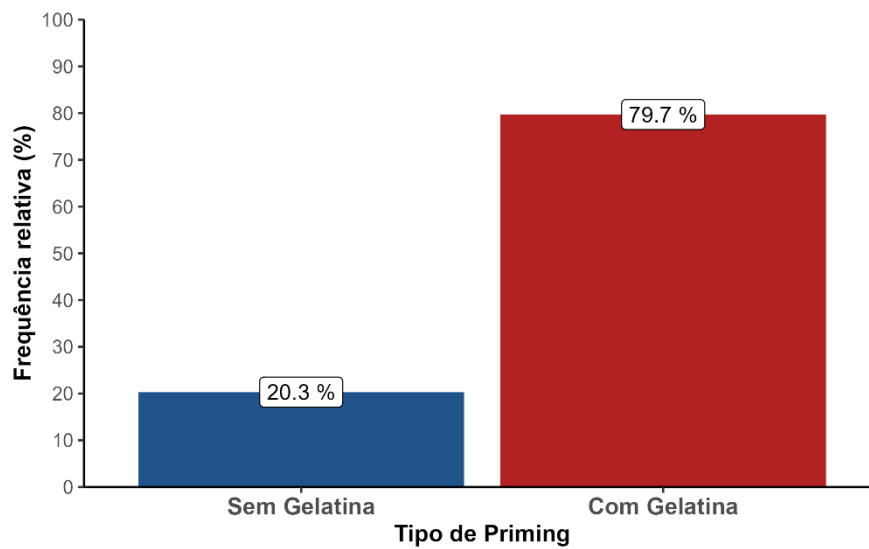
Quanto aos tempos de CEC e de clampagem da aorta, estes encontram-se sumarizados na Figura 25. O tempo de CEC obtido foi de 62 (44 - 80) minutos e o de clampagem da aorta foi de 49 (33 - 64) minutos.

Figura 25 - Variável da CEC - Tempo de CEC e de Clampagem da Aorta



Quanto aos tipos de *priming*, verifica-se que 79.7 % dos doentes realizou Gelatina no *priming* do circuito de CEC (Figura 26).

Figura 26 - Variável da CEC - Tipos de *Priming*



5.2 Hemorragia pós-operatória e utilização de gelatina no *priming*

A Tabela 8 resume os dados demográficos e clínicos dos doentes envolvidos neste estudo, estratificados de acordo com o tipo de *priming* utilizado no circuito de CEC: *priming* sem gelatina ou *priming* com gelatina. O objetivo principal desta análise é avaliar se existe uma associação entre o tipo de *priming* e a ocorrência de hemorragia no pós-operatório. Deste modo, pretende-se consubstanciar se os dois grupos de estudo são homogêneos no que concerne às variáveis explicativas categóricas presentes na tabela. Os resultados são expressos como o número de contagens (frequência absoluta). Os valores de p apresentados dizem respeito aos resultados do teste de qui-quadrado, à exceção da variável Insuficiência Renal Crónica em que foi utilizado um teste de *Fisher*.

Tabela 8 - Comparação dos dados qualitativos por grupo de tipo de *priming*

	Todos os doentes	<i>Priming</i> sem gelatina	<i>Priming</i> com gelatina	p-value*
Sexo feminino, n (%)	406 (54.6)	87 (57.6)	319 (53.8)	0.453
IMC categórico, n (%)				
Peso normal	192 (25.8)	38 (25.2)	154 (26)	0.870
Excesso de peso	326 (43.8)	69 (45.7)	257 (43.3)	
Obesidade	226 (30.4)	44 (29.1)	182 (30.7)	
Diabetes, n (%)				
Não Diabéticos	543 (73)	110 (72.8)	433 (73)	0.280
Diabetes Tipo II NIT	184 (24.7)	35 (23.2)	149 (25.1)	
Diabetes Tipo II IT	17 (2.3)	6 (4)	11 (1.9)	
Dislipidemia, n (%)	482 (64.8)	101 (66.9)	381 (64.2)	0.610
Hipertensão arterial, n (%)	608 (81.7)	119 (78.8)	489 (82.5)	0.358
Tabagismo				
Não fumador	549 (73.8)	102 (67.5)	447 (75.4)	0.065
Ex-fumador	129 (17.3)	29 (19.2)	100 (16.9)	
Fumador	66 (8.9)	20 (13.2)	46 (7.8)	
Enfarte agudo do miocárdio, n (%)	50 (6.7)	14 (9.3)	36 (6.1)	0.222
Doença vascular periférica, n (%)	32 (4.3)	12 (7.9)	20 (3.4)	0.025
Insuficiência Renal Crónica, n (%)				
Sem IRC	621 (83.5)	131 (86.8)	490 (82.6)	0.059
IRC Ligeira	57 (7.7)	4 (2.6)	53 (8.9)	
IRC Moderada	46 (6.2)	12 (7.9)	34 (5.7)	
IRC Grave	13 (1.7)	3 (2)	10 (1.7)	
IRC Hemodialise	7 (0.9)	1 (0.7)	6 (1)	
Função ventricular Esquerda, n (%)				
Normal	642 (86.3)	126 (83.4)	516 (87)	0.237
Ligeiramente diminuída	56 (7.5)	14 (9.3)	42 (7.1)	
Moderadamente diminuída	31 (4.2)	10 (6.6)	21 (3.5)	
Diminuída	14 (1.9)	1 (0.7)	13 (2.2)	
Muito diminuída	1 (0.1)	0 (0)	1 (0.2)	
Peso efetivo da cirurgia, n(%)				
Cirurgia Coronária	19 (2.6)	9 (6)	10 (1.7)	0.026
1 procedimento	403 (54.2)	82 (54.3)	321 (54.1)	

Tabela 8 - Comparação dos dados qualitativos por grupo de tipo de *priming*

	2 procedimentos	3 procedimentos			
	283 (38)	39 (5.2)	52 (34.4)	231 (39)	
			8 (5.3)	31 (5.2)	
Cirurgia da aorta, n(%)	128 (17.2)		28 (18.5)	100 (16.9)	0.713
Cirurgia urgente, n(%)	44 (5.9)		10 (6.6)	34 (5.7)	0.826

*Foi utilizado o teste de qui-quadrado para a comparação das amostras, à exceção da variável Insuficiência Renal Crónica em que foi utilizado um teste de Fisher

Em suma, os resultados indicam que, em geral, os grupos de *priming* com e sem gelatina são homogêneos em relação às variáveis explicativas categóricas incluídas, com exceção da presença de doença vascular periférica e ao peso efetivo da cirurgia, onde se observou uma diferença significativa nas respetivas contagens.

A Tabela 9 apresenta as variáveis quantitativas dos doentes submetidos a cirurgia cardíaca com e sem o uso de gelatina no *priming* do circuito CEC. Os resultados são expressos como mediana (intervalo interquartil) para todas as variáveis. O teste de *Mann-Whitney U* foi utilizado para comparar as medianas entre os grupos. Para praticamente todas as variáveis não se verificaram diferenças significativas. No entanto, foi observada uma diferença estatisticamente significativa no tempo de protrombina (TP) entre os grupos (p-value = 0,022), com o grupo de *priming* sem gelatina apresentando um TP ligeiramente maior em comparação com o grupo de *priming* com gelatina. Da mesma forma, o índice internacional normalizado (INR) também foi significativamente diferente entre os grupos (p-value = 0,005). Em resumo, os resultados sugerem que, as características clínicas e laboratoriais dos doentes são idênticas em ambos os grupos. Para além disso, verifica-se que apesar de na Tabela 8 se ter verificado diferenças entre os pesos efetivos da cirurgia, doença vascular periférica, em termos de tempo de CEC e de Clampagem da Aorta os dois grupos são homogêneos.

Tabela 9 - Comparação dos dados quantitativos por grupo de tipo de *priming*

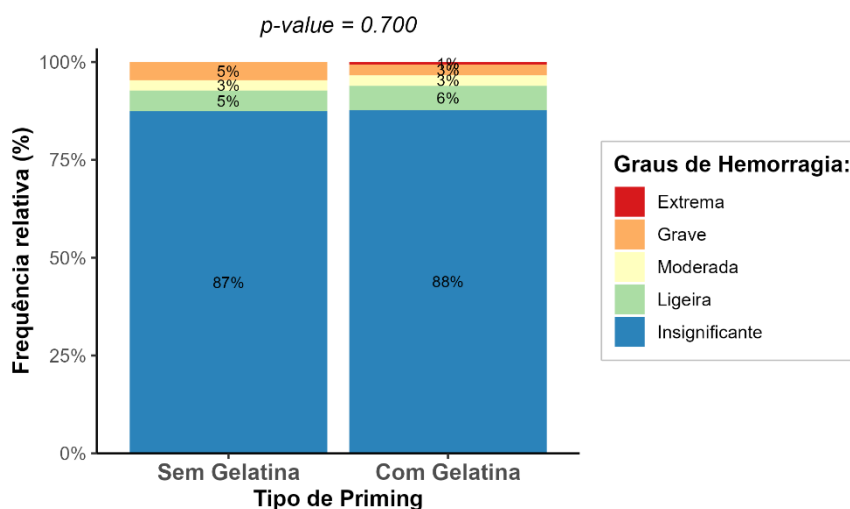
	Todos	<i>Priming</i> sem gelatina	<i>Priming</i> com gelatina	p-value*
Idade, anos	71 (63 – 76)	70 (61 – 76)	71 (64 – 76)	0.193
ASC, m²	1.8 (1.7 – 2)	1.8 (1.7 – 2)	1.8 (1.7 – 2)	0.351
Volémia, mL	5180 (4550 – 5810)	5250 (4550 – 5950)	5180 (4550 – 5740)	0.389
Hemoglobina, g/dL	13.5 (12.5 – 14.6)	13.5 (12.4 – 14.6)	13.5 (12.5 – 14.6)	0.715
Hematócrito, %	40.2 (37.2 – 42.8)	40.5 (37.3 – 43)	40.1 (37.1 – 42.8)	0.512
Fibrinogénio, mg/dL	333 (280.8 – 395)	333 (290 – 411)	333 (280 – 393)	0.178
Plaquetas, x10⁹/L	213 (179 – 258)	213 (176.5 – 243)	214 (179 – 261)	0.210
TP, segundos	12.1 (11.5 – 12.7)	12.2 (11.7 – 12.9)	12.1 (11.5 – 12.7)	0.022
INR	1.1 (1 – 1.1)	1.1 (1 – 1.1)	1.1 (1 – 1.1)	0.005
aPTT, segundos	29 (27.1 – 31.1)	29 (27.6 – 31.3)	29 (27 – 31)	0.122

ACT Basal , segundos	140 (126 – 150)	137 (122.5 – 150)	140 (126 – 150)	0.398
ACT Heparina , segundos	488 (427.8 – 557.2)	493 (437 – 583)	488 (424 – 551)	0.195
ACT Final , segundos	121 (110 – 132)	120 (109.5 – 129.5)	122 (110 – 134)	0.150
Tempo de CEC , minutos	62 (43 – 80)	62 (44 – 78.5)	62 (43 – 80)	0.843
Tempo de Clamp , minutos	49 (33 – 64)	49 (34 – 63)	49 (33 – 64)	0.632

*Foi utilizado o teste de *Mann-Whitney U* para a comparação das amostras independentes
Os dados laboratoriais dizem respeito às medições pré-operatórias

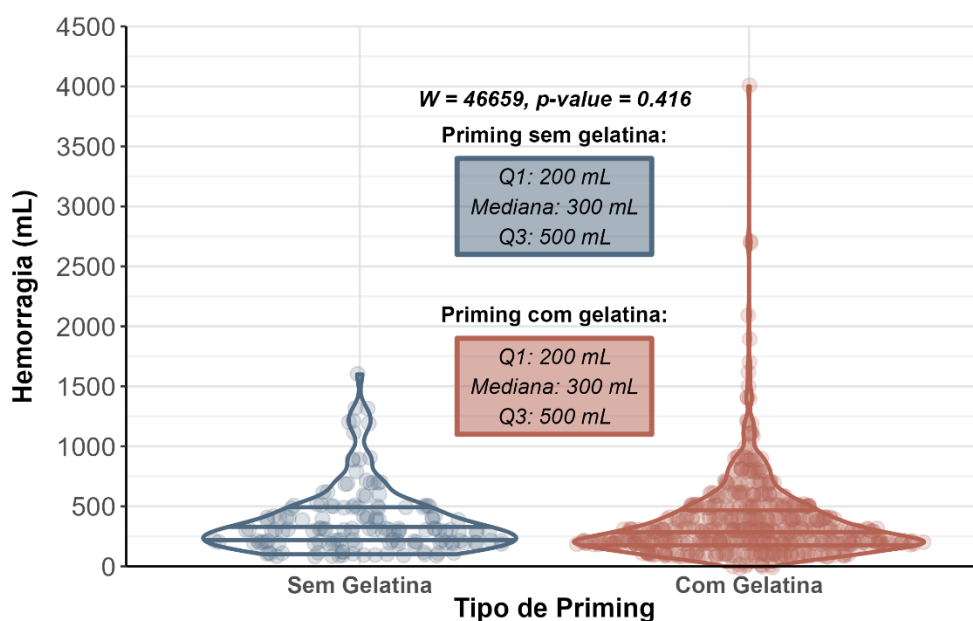
O gráfico da Figura 27 representa as frequências relativas das diferentes categorias de hemorragia para os dois grupos de *priming*. Por forma a avaliar a associação entre o tipo de *priming* (com vs. sem gelatina) e a gravidade da hemorragia às 12 horas de pós-operatório (insignificante, ligeira, moderada, grave ou extrema), foi realizado um teste de *Fisher*. Os resultados indicam que não existe evidência estatística de associação significativa entre o tipo de *priming*, quando à adição de gelatina, e a gravidade da hemorragia (p -value = 0.700).

Figura 27 - Comparação das Categorias de Hemorragia Pós-operatória por tipo de *Priming* através do teste de Fisher



Para comparar a variável quantitativa relacionada com a hemorragia às 12 horas de pós-operatório entre os dois grupos de doentes, conforme definido pela utilização de gelatina no *priming*, foi elaborado o gráfico presente na Figura 28. Foi aplicado um teste de *Mann-Whitney U* para investigar possíveis diferenças na hemorragia entre os grupos. Os resultados indicam que não existe evidência estatisticamente significativa que sugira que a quantidade de hemorragia às 12 horas difira entre os grupos (p -value = 0.416).

Figura 28 - Comparação da hemorragia às 12 horas de pós-operatório por tipo de *priming* através do teste de *Mann-Whitney U*



Para comparar a variável quantitativa relacionada com a hemorragia às 24 horas de pós-operatório entre os dois grupos de doentes foi elaborado o gráfico presente na Figura 29. Foi também aplicado um teste de *Mann-Whitney U*. Os resultados indicam que existe evidência estatisticamente significativa para sugerir que a hemorragia difere entre os grupos (p-value = 0.024). Através da análise das estatísticas descritivas presentes nos dois gráficos verifica-se que em termos de mediana e de intervalos interquartílicos, não existem diferenças entre os grupos. Apenas se verificam diferenças numa maior cauda de valores superiores existentes no grupo de doentes que efetuou gelatina no *priming*:

Como foi descrito na secção correspondente às Variáveis dependentes (Tabela 1 ,página 21) foram indexados os valores da hemorragia às 12 e às 24 horas de pós-operatório ao peso e à área de superfície corporal. Assim, foi construída a Tabela 10, onde foram realizados vários testes de *Mann-Whitney U* para investigar possíveis diferenças nos valores indexados de hemorragia entre os grupos (com vs. sem gelatina). Às 12 horas de pós-operatório, não foram observadas diferenças estatisticamente significativas em relação à drenagem indexada à ASC (p-value = 0.569) e à drenagem indexada ao peso corporal (p-value = 0.639). Já para as 24 horas de pós-operatório, embora tenha sido observada uma tendência de redução na hemorragia indexada à ASC no grupo de *priming* com gelatina em comparação com o grupo *priming* sem gelatina (268.7 mL/m² vs. 247.2 mL/m², respetivamente), contudo essa diferença não alcançou a significância estatística (p-value = 0.075). Da mesma forma, a drenagem indexada ao peso corporal do indivíduo também apresentou uma tendência de redução no grupo com

priming com gelatina em comparação com o grupo sem gelatina no *priming* (6.7 mL/Kg vs. 6.1 mL/Kg, respectivamente), mas novamente essa diferença não foi estatisticamente significativa (p-value = 0.147).

Figura 29 - Comparação da hemorragia às 24 horas de pós-operatório por tipo de *priming* através do teste de *Mann-Whitney U*

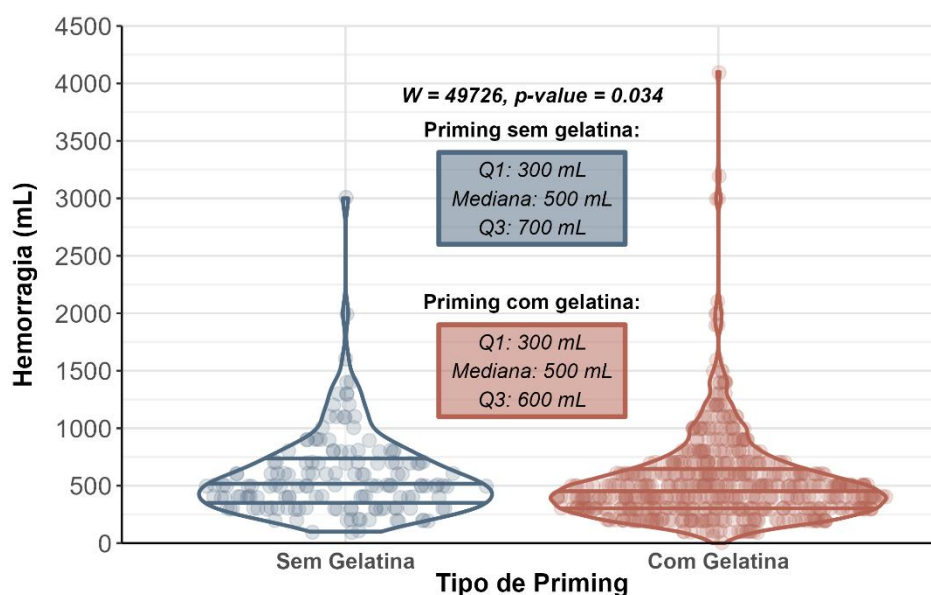


Tabela 10 - Comparação da hemorragia às 12 e 24 horas de pós-operatório, indexadas à ASC e Peso, por tipo de *priming*

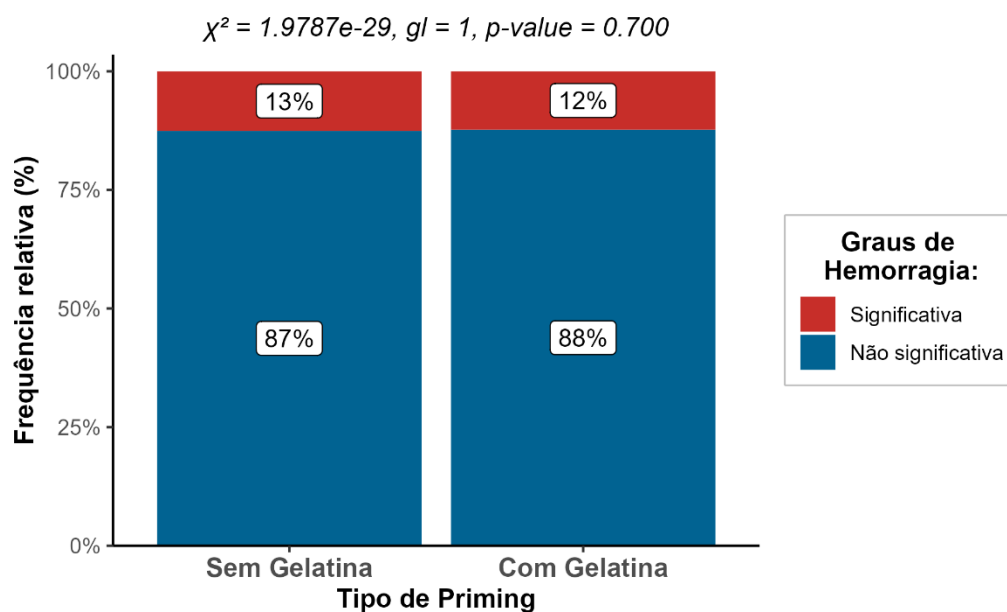
	<i>Priming sem Gelatina</i>	<i>Priming com Gelatina</i>	p-value*
12 horas pós-operatório			
<i>drenagem/ASC (mL/m²)</i>	163.3 (104.5 - 252.1)	155.9 (102.9 - 248.9)	0.569
<i>drenagem/Peso (mL/Kg)</i>	4.1 (2.5 - 6.3)	3.8 (2.5 - 6.3)	0.639
24 horas pós-operatório			
<i>drenagem/ASC (mL/m²)</i>	268.7 (187.6 - 386.9)	247.2 (167.9 - 348.4)	0.075
<i>drenagem/Peso (mL/Kg)</i>	6.7 (4.3 - 10)	6.1 (4.1 - 9)	0.147

*Foi utilizado o teste de Mann-Whitney U para a comparação das amostras

O gráfico da Figura 30 representa as frequências das duas categorias de hemorragia (correspondente à variável *drenagem_sig*) para os dois grupos de *priming* relativos à utilização de gelatina no mesmo. Por forma a avaliar a associação entre o tipo de *priming* (com gelatina vs. sem gelatina) e a presença de uma hemorragia significativa às 12 horas de pós-

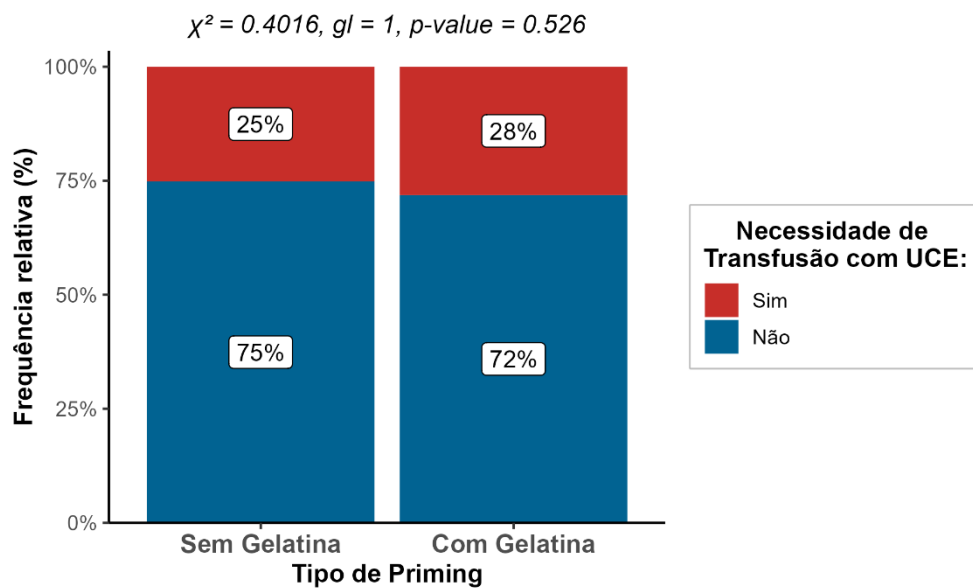
operatório (ligeira, moderada, grave ou extrema), foi realizado um teste de qui-quadrado. Os resultados indicam que não existe evidência estatisticamente significativa de possível associação entre o tipo de *priming*, quanto à adição de gelatina e a existência de hemorragia significativa (p-value = 0.700).

Figura 30 - Comparação das Categorias de Hemorragia Pós-operatória por tipo de *Priming* através do teste de qui-quadrado



Foi investigada a relação entre a necessidade de transfusão sanguínea com o uso de UCE e o tipo de *priming* utilizado. As distribuições das frequências por grupo encontram-se representadas na Figura 31. Por forma a avaliar esta associação foi realizado um teste de qui-quadrado. Os resultados indicam que não existe evidência estatisticamente significativa de possível associação entre o tipo de *priming*, quanto à adição de gelatina, e a necessidade de transfusão sanguínea (p-value = 0.526). Isto sugere que o tipo de *priming* não parece ter influenciado de forma significativa a necessidade de transfusão sanguínea no pós-operatório.

Figura 31 - Comparação das Categorias da necessidade de transfusão sanguínea por tipo de *Priming* através do teste de qui-quadrado

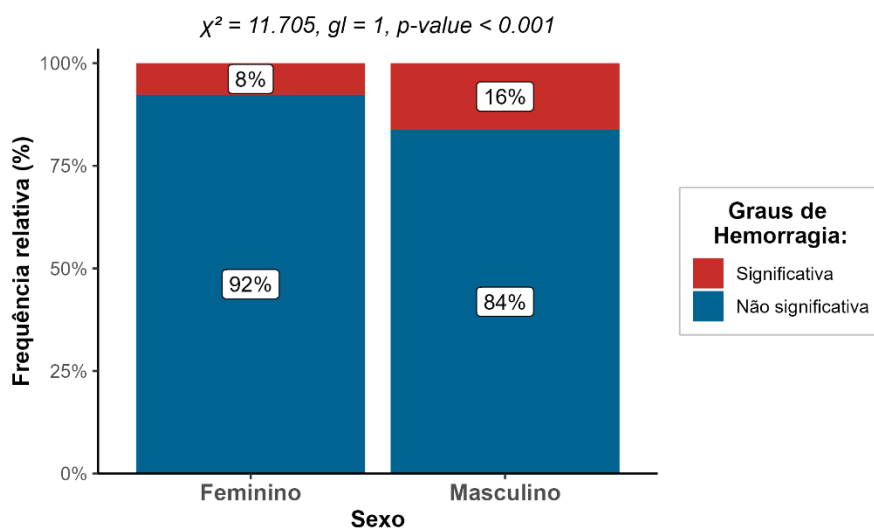


5.2.1 Comparação da hemorragia com as restantes variáveis explicativas

Foram realizadas diversas tabelas de contingência para investigar a associação entre as demais variáveis potencialmente explicativas e a hemorragia significativa. Os dados foram agrupados de acordo com cada uma das variáveis e a categorização da hemorragia.

Em primeiro lugar, estudou-se a associação com a variável referente ao sexo do indivíduo (Figura 32). Os resultados revelaram que, entre os doentes do sexo feminino, 340 (83.7%) apresentaram hemorragia não significativa, enquanto 66 (16.3%) tiveram hemorragia significativa. Por outro lado, entre os doentes do sexo masculino, 312 (92.3%) tiveram hemorragia não significativa e 26 (7.7%) tiveram hemorragia significativa. Um teste de qui-quadrado foi realizada para avaliar a associação entre as variáveis. O teste revelou uma associação estatisticamente significativa entre essas variáveis ($\chi = 11.705$, $gl = 1$, $p < 0.001$), indicando que o sexo dos doentes está associado de forma significativa à hemorragia pós-operatória.

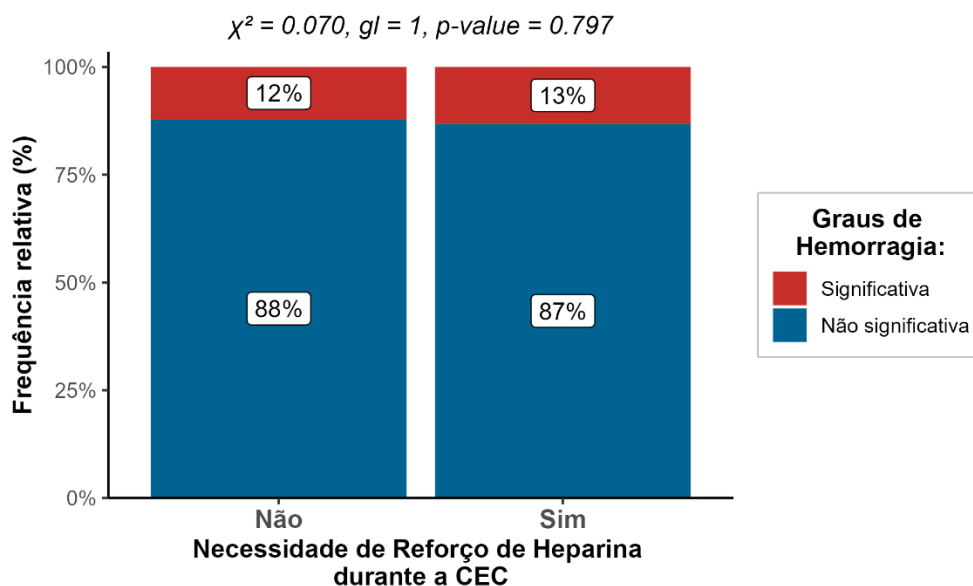
Figura 32 - Hemorragia significativa no pós-operatória vs. Sexo através do teste de qui-quadrado



Foi realizada uma tabela de contingência para examinar a associação entre a necessidade de reforço de Heparina durante a CEC e a hemorragia significativa (Figura 33). Os resultados mostraram que, entre os doentes que não necessitaram de Heparina durante a CEC, 494 (87.9%) apresentaram hemorragia não significativa, enquanto 68 (12.1%) tiveram hemorragia significativa. Por outro lado, entre os doentes que receberam Heparina durante a CEC, 158 (86.8%) tiveram uma hemorragia não significativa e 24 (13.2%) significativa. O teste de qui-

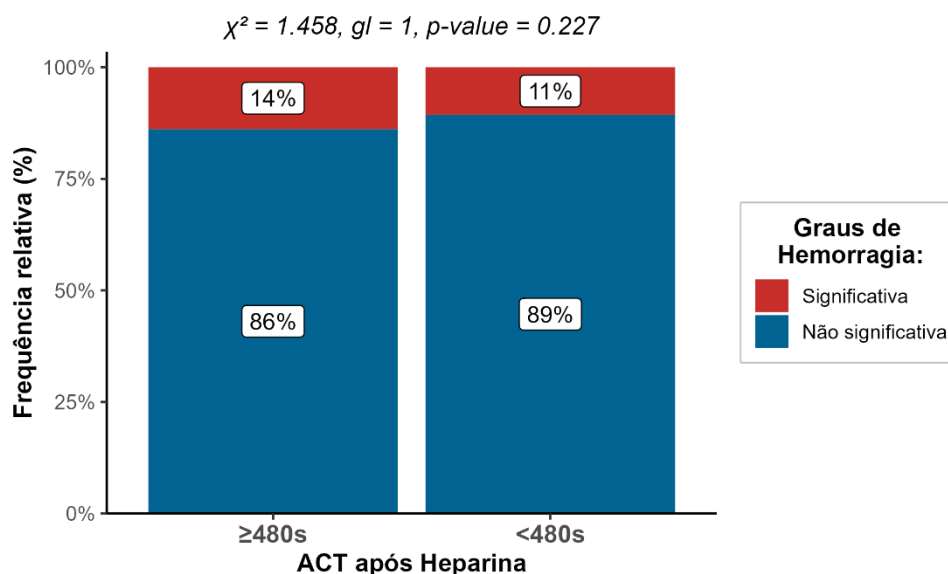
quadrado realizado indicou que não existe uma associação estatisticamente significativa entre as variáveis ($\chi = 0.070$, $gl = 1$, $p\text{-value} = 0.797$).

Figura 33 - Hemorragia significativa no pós-operatória vs. Necessidade de Heparina durante a CEC através do teste de qui-quadrado



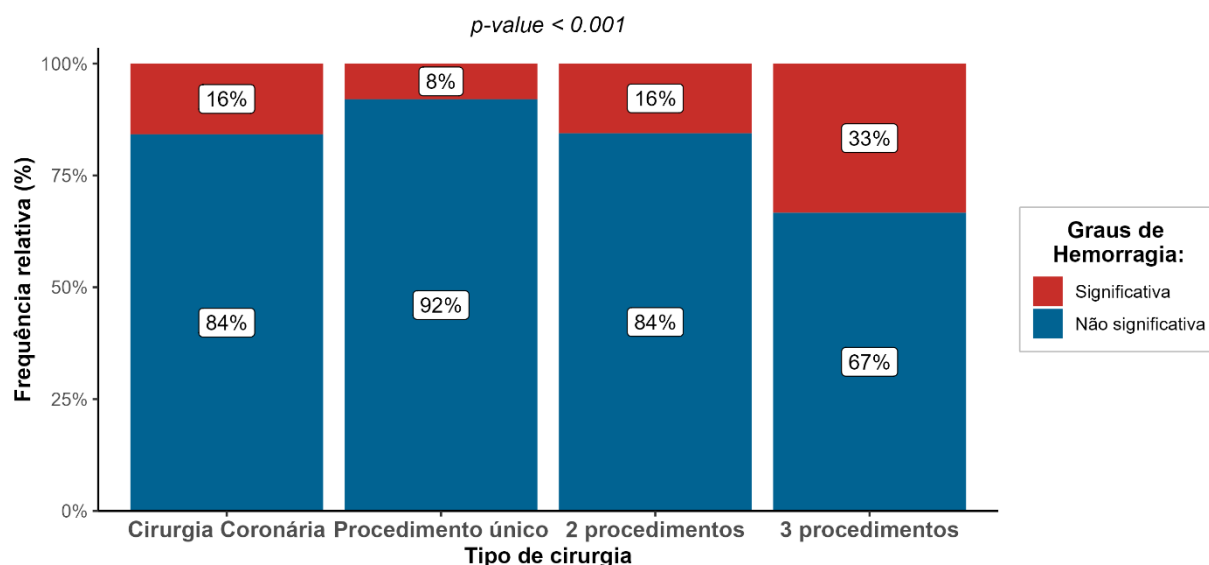
Foi realizada uma tabela de contingência sobre a associação entre o ACT categórico após a heparina ($ACT \geq 480$ ou < 480 segundos) e a significância da hemorragia pós-operatória (Figura 34). O teste de qui-quadrado foi realizado para avaliar a associação entre ambas as variáveis. Os resultados indicam que não parece existir uma associação estatisticamente significativa entre as variáveis ($\chi = 1.458$, $gl = 1$, $p\text{-value} = 0.227$), sugerindo que esta divisão categórica não está significativamente associada à hemorragia pós-operatória.

Figura 34 - Hemorragia significativa no pós-operatória vs. ACT categórico após heparina através do teste de qui-quadrado



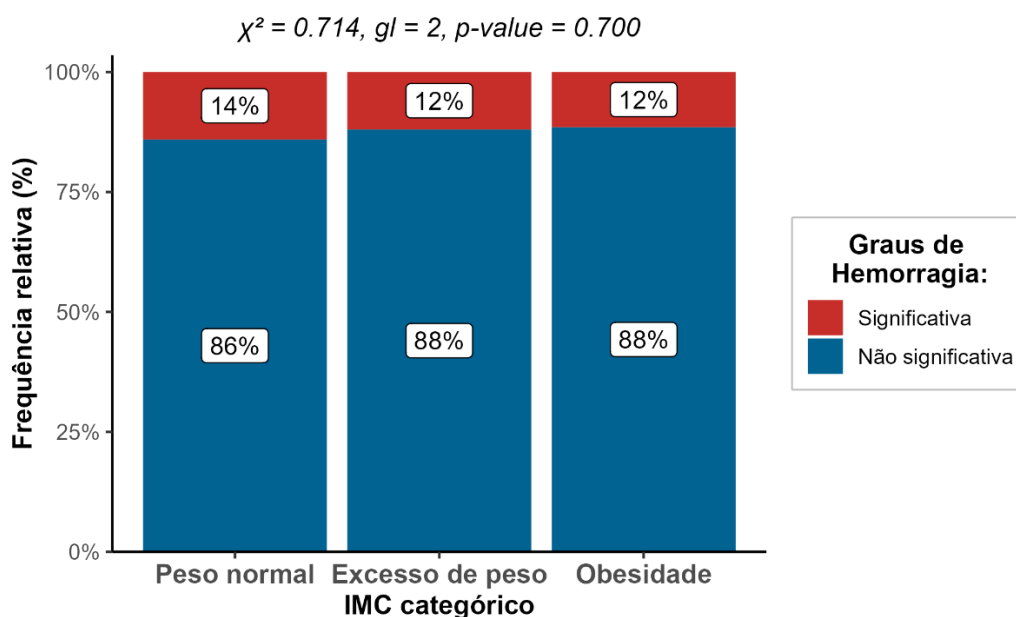
Foi realizada uma tabela de contingência para examinar a associação entre o tipo de procedimento cirúrgico (cirurgia coronária, procedimento único não coronário, 2 procedimentos, 3 procedimentos) e a presença de hemorragia pós-operatória significativa (Figura 35). Os resultados revelaram que entre os doentes submetidos apenas a cirurgia coronária, 16 (84.2%) tiveram hemorragia não significativa, enquanto 3 (15.8%) tiveram hemorragia significativa. Para doentes submetidos a um procedimento único, não coronário, 371 (92.1%) tiveram hemorragia não significativa e 32 (7.9%) tiveram hemorragia significativa. Para doentes submetidos a 2 procedimentos, 239 (84.5%) tiveram hemorragia não significativa e 44 (15.5%) tiveram hemorragia significativa. Finalmente, entre doentes submetidos a 3 procedimentos cirúrgicos, 26 (66.7%) tiveram hemorragia não significativa e 13 (33.3%) tiveram hemorragia significativa. Foi utilizado o teste de *Fisher*. Os resultados indicaram uma associação estatisticamente significativa entre estas variáveis ($p\text{-value} < 0.001$), sugerindo que o tipo de procedimento cirúrgico está significativamente associado à hemorragia pós-operatória.

Figura 35 - Hemorragia significativa no pós-operatória vs. peso efetivo da cirurgia através do teste de Fisher



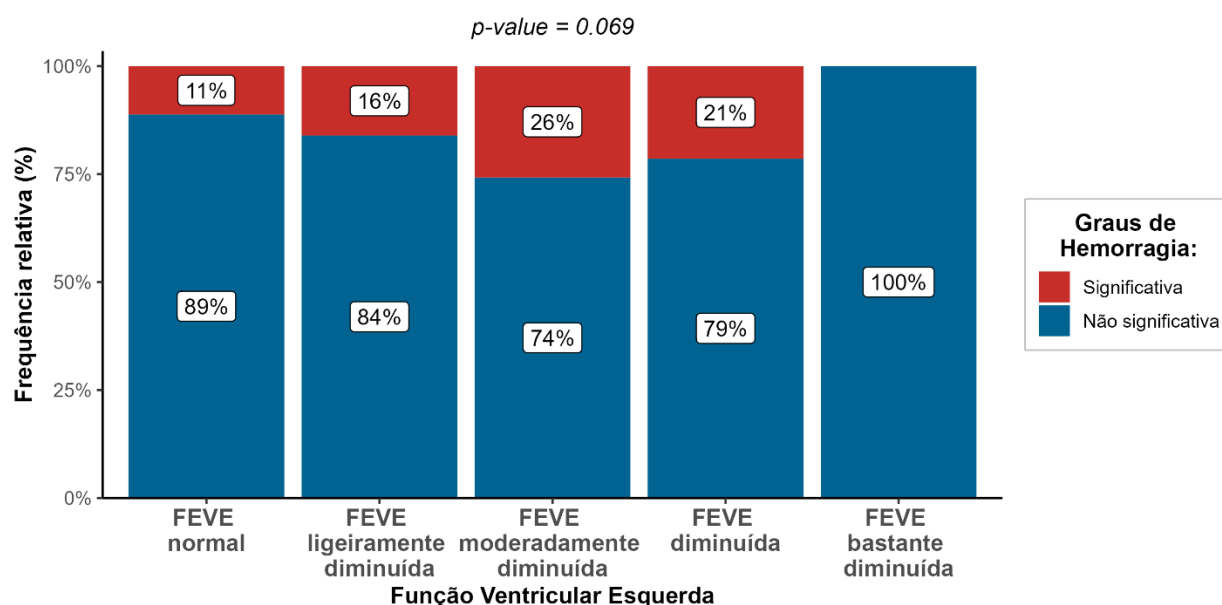
Foi também realizada uma tabela de contingência sobre a associação entre o IMC categórico e a ocorrência de uma hemorragia significativa no pós-operatório (Figura 36). O teste de qui-quadrado foi realizado para avaliar a associação entre ambas as variáveis. Os resultados indicam que parece não existir uma associação estatisticamente significativa entre as variáveis ($\chi^2 = 0.714$, $gl = 2$, $p\text{-value} = 0.700$).

Figura 36 - Hemorragia significativa no pós-operatória vs. IMC categórico através do teste de qui-quadrado



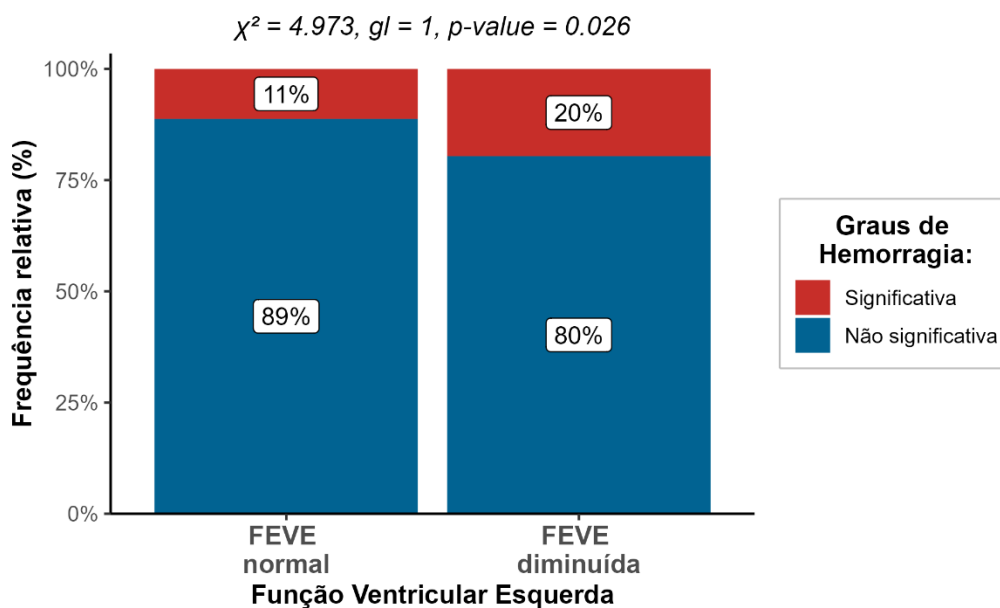
Uma tabela de contingência para examinar a associação entre a fração de ejeção ventricular esquerda (FEVE) e a significância da hemorragia pós-operatória foi efetuada (Figura 37). Os resultados revelaram que entre os doentes com FEVE normal (>50%), 570 (88.8%) tiveram hemorragia não significativa, enquanto 72 (11.2%) tiveram hemorragia significativa. Para doentes com FEVE ligeiramente diminuída (41-50%), 47 (83.9%) tiveram hemorragia não significativa e 9 (16.1%) tiveram hemorragia significativa. Para doentes com FEVE moderadamente diminuída (31-40%), 23 (74.2%) tiveram hemorragia não significativa e 8 (25.8%) tiveram hemorragia significativa. Para doentes com FEVE diminuída (21-30%), 11 (78.6%) tiveram hemorragia não significativa e 3 (21.4%) tiveram hemorragia significativa. Por fim, entre os doentes com FEVE bastante diminuída (<21%), 1 (100%) teve hemorragia não significativa e nenhum paciente apresentou hemorragia significativa. Para testar a associação entre a FEVE e a significância da hemorragia, foi utilizado o *Teste Exato de Fisher*. Os resultados indicaram uma associação com uma tendência para a significância estatística ($p\text{-value} = 0.069$) mas que ao nível de significância escolhido não se poderá assumir que, a FEVE pareça estar associada à significância da hemorragia pós-operatória.

Figura 37 - Hemorragia significativa no pós-operatória vs. Função ventricular esquerda através do teste de Fisher



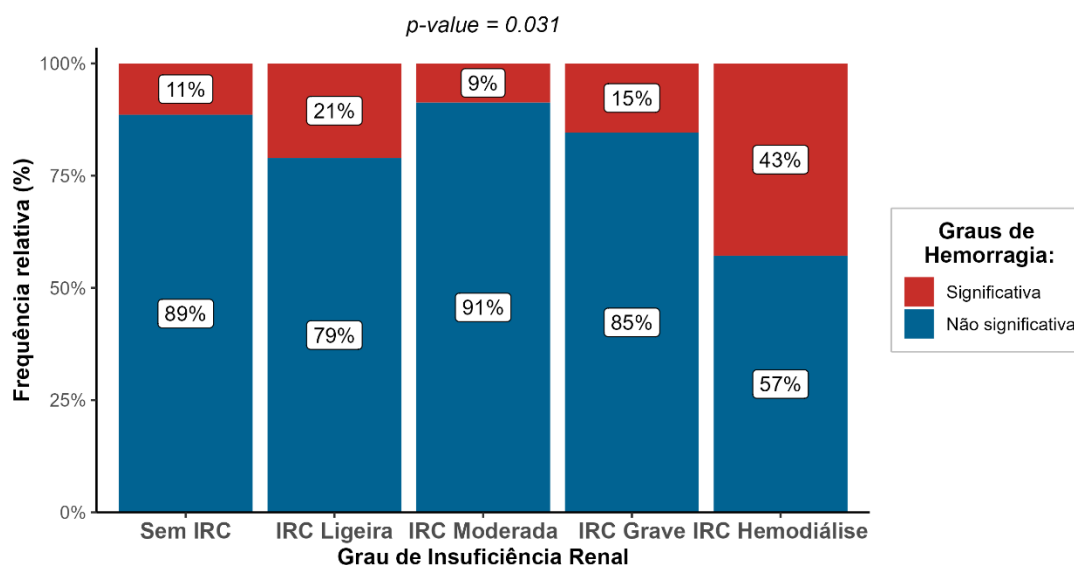
Por forma a simplificar o modelo futuramente estimado procedeu-se à categorização da variável *feve* numa nova variável categórica com a aglutinação de todas as categorias que representam FEVE não normal. Os resultados (Figura 38) indicam que parece existir uma associação estatisticamente significativa entre as variáveis ($\chi = 4.973$, $gl = 1$, $p = 0.026$).

Figura 38 - Hemorragia significativa no pós-operatória vs. feve2 através do teste de qui-quadrado



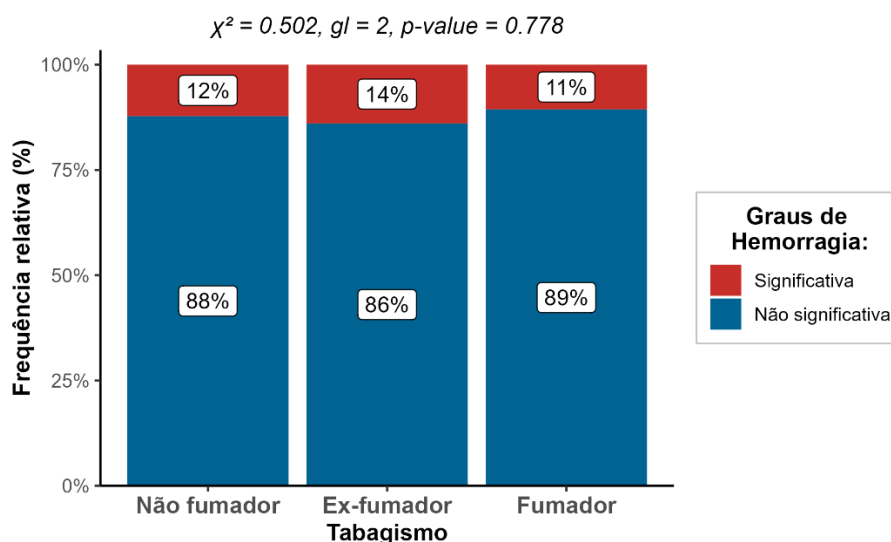
Procedeu-se agora à mesma análise, mas para a variável relacionada com os antecedentes de insuficiência renal crónica (Figura 39). Os resultados revelaram que, entre os doentes sem IRC, 550 (88.6%) não apresentaram hemorragia significativa, enquanto 71 (11.4%) tiveram hemorragia significativa. No gráfico da Figura 39 é possível visualizar as restantes frequências para cada uma das categorias, que representam os diferentes graus de IRC. Para testar a associação entre a IRC e a presença de hemorragia, foi utilizado o *Teste Exato de Fisher*. Os resultados indicaram uma associação estatisticamente significativa ($p\text{-value} = 0.031$), sugerindo que a IRC parece estar associada com a ocorrência de hemorragia pós-operatória.

Figura 39 - Hemorragia significativa no pós-operatória vs. Insuficiência Renal Crónica através do teste de Fisher



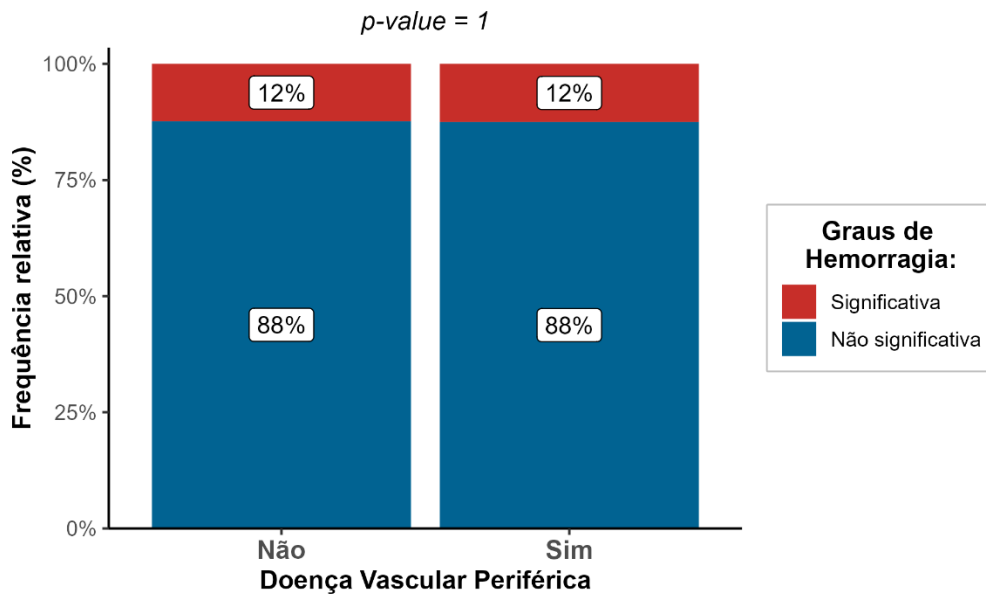
Foi investigada também a associação entre a variável *fumador* e a ocorrência de hemorragia (Figura 40). A análise foi realizada por intermédio de um teste de qui-quadrado. Os resultados indicam que parece não existir uma associação estatisticamente significativa entre as variáveis ($\chi^2 = 0.502$, $gl = 2$, $p\text{-value} = 0.778$).

Figura 40 - Hemorragia significativa no pós-operatória vs. Tabagismo através do teste de qui-quadrado



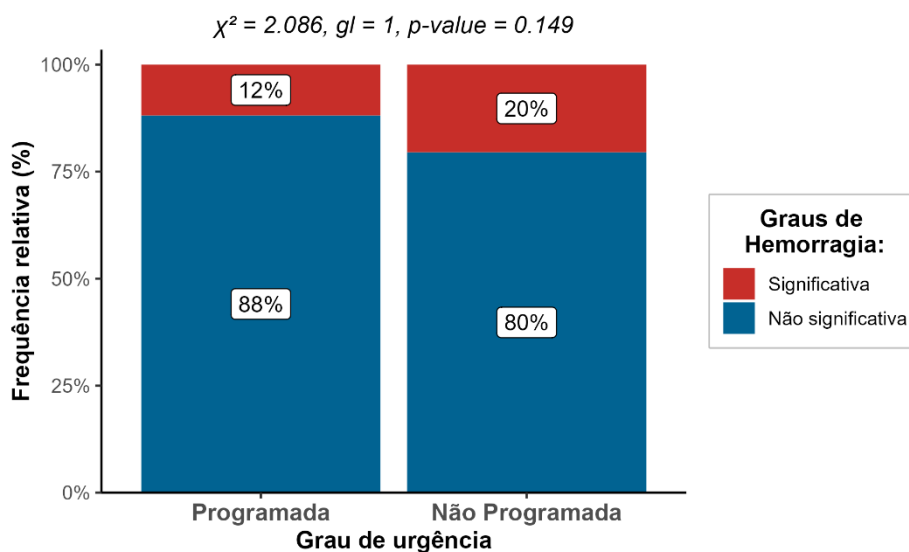
Quanto à doença vascular periférica (Figura 41), o teste Exato de *Fisher* indica que parece não existir uma associação estatisticamente significativa entre as variáveis ($p\text{-value} = 1$), sugerindo que a DVP não está significativamente associada à hemorragia pós-operatória significativa.

Figura 41 - Hemorragia significativa no pós-operatória vs. Doença vascular periférica através do teste de Fisher



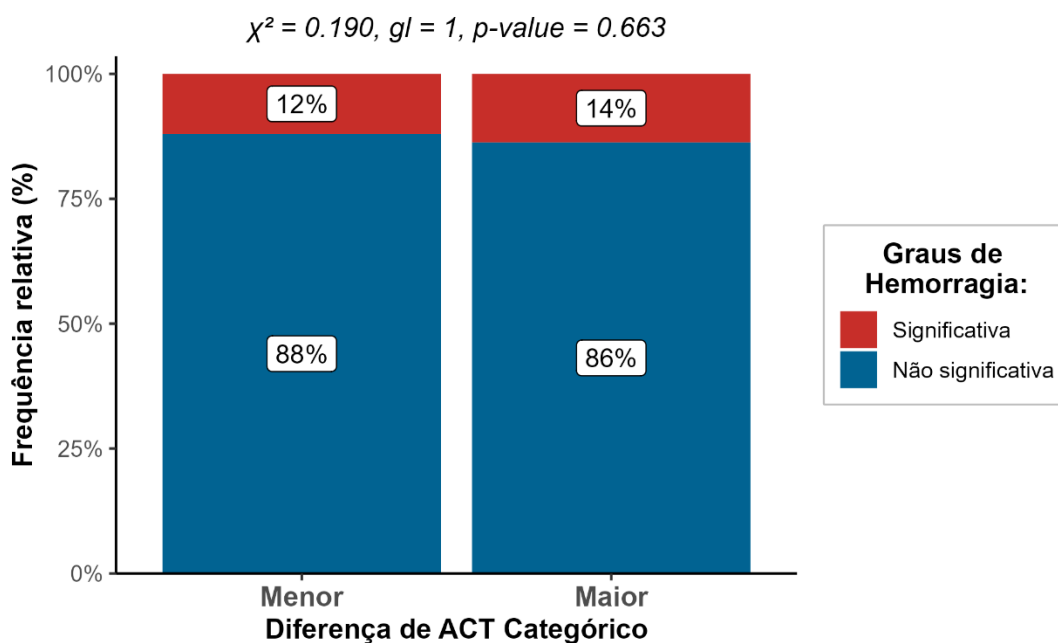
Segue-se uma tabela de contingência para o grau de urgência da cirurgia (programada vs. não programada) e a significância da hemorragia pós-operatória (Figura 42). Os resultados revelaram que entre as cirurgias programadas, 617 (88.1%) não resultaram em hemorragia, enquanto 83 (11.9%) resultaram em hemorragia. Para as cirurgias não programadas, ou urgentes, 35 (79.5%) não resultaram em hemorragia e 9 (20.5%) resultaram em hemorragia. Uma análise de qui-quadrado revelou que parece não existir evidência de uma associação estatisticamente significativa entre estas variáveis ($\chi = 2.086$, $gl = 1$, $p\text{-value} = 0.149$).

Figura 42 - Hemorragia significativa no pós-operatória vs. grau de urgência do procedimento cirúrgico através do teste de qui-quadrado



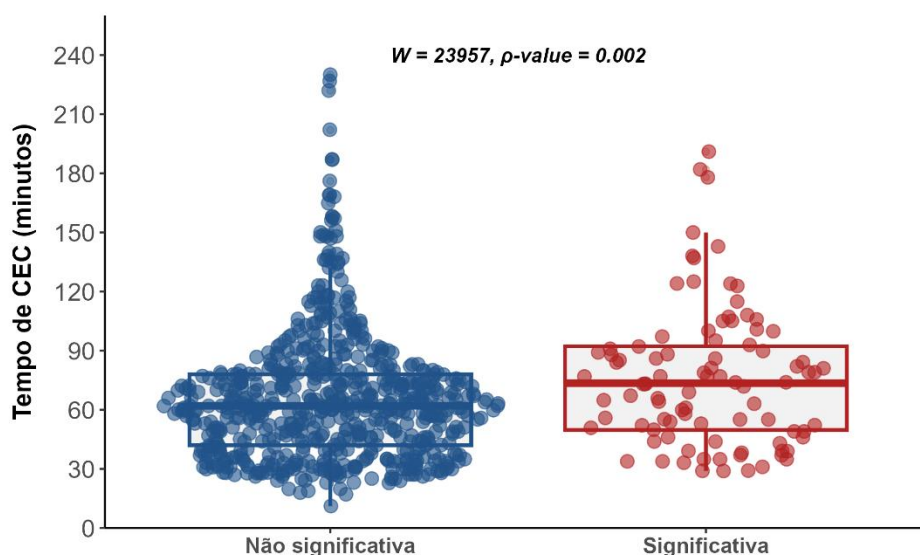
Foi analisada agora a diferença categórica entre o ACT Final e o ACT Basal (Figura 43). A análise foi realizada através de um teste de qui-quadrado por forma a avaliar a associação entre as variáveis. Os resultados indicam que parece não existir uma associação estatisticamente significativa entre as variáveis e a ocorrência de hemorragia ($p\text{-value} = 0.663$).

Figura 43 - Hemorragia significativa no pós-operatória vs. Diferença de ACT categórico através do teste de qui-quadrado



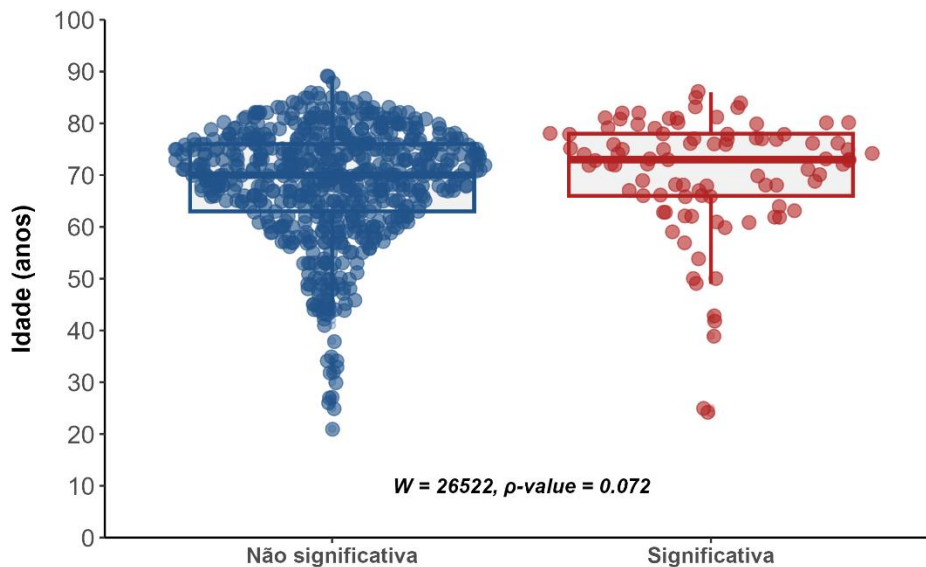
Por forma a comparar a variável relativa ao tempo de CEC entre os dois grupos de doentes, foi construído o gráfico presente na Figura 44. Foi aplicado um teste de *Mann-Whitney-U* para investigar possíveis diferenças no tempo de CEC entre os grupos. Os resultados indicam que existe evidência estatisticamente significativa que sugere que os tempos de CEC diferem entre os dois grupos ($p\text{-value} = 0.002$). Através da análise das estatísticas descritivas presentes nos dois gráficos verifica-se que o grupo de doentes com hemorragia significativa tem um valor mediano superior (73.5, IQR: 49.75 - 92.25 minutos) aos sem hemorragia significativa (62, IQR: 42 - 78 minutos).

Figura 44 - Hemorragia significativa no pós-operatório vs. Tempo de CEC através do teste de *Mann-Whitney-U*



A mesma comparação foi realizada para a variável idade (Figura 45). Os resultados indicam que não existe evidência estatisticamente significativa para sugerir que os as idades diferem entre os dois grupos ($p\text{-value} = 0.072$), pelo menos para o nível de significância escolhido. Através da análise do gráfico verifica-se que o grupo de doentes com hemorragia significativa têm uma ligeira tendência para um valor mediano superior (73, IQR: 66 - 78 anos), quando comparada aos doentes sem hemorragia significativa (70, IQR: 63 - 76 anos).

Figura 45 - Hemorragia significativa no pós-operatório vs. Idade através do teste de *Mann-Whitney-U*



Por fim, e por forma a comparar as restantes variáveis quantitativas e que dizem respeito a variáveis laboratoriais pré-operatórias ou às medições de ACT intraoperatórias, foi construída a Tabela 11. Aplicou-se a mesma metodologia que a anteriormente descrita, mas a este novo conjunto de variáveis. Para nenhuma das variáveis apresentadas verificaram-se diferenças estatisticamente significativas entre os grupos.

Tabela 11 - Hemorragia significativa vs. restantes variáveis quantitativas

	Hemorragia não significativa	Hemorragia significativa	<i>p</i> -value*
Hemoglobina , g/dL	13.5 (12.5-14.6)	13.4 (12.4 – 14.3)	0.158
Hematócrito , %	40.2 (37.2 – 42.9)	39.3 (36.9 – 42.3)	0.138
Fibrinogénio , mg/dL	333 (284 – 397)	329.5 (260 – 387.8)	0.216
Plaquetas , $\times 10^9$ /L	213 (179 – 261)	210 (173.8 – 235.2)	0.065
TP , segundos	12.1 (11.5 – 12.7)	12.1 (11.6 – 12.9)	0.288
INR	1.1 (1.0 – 1.1)	1.1 (1.0 – 1.1)	0.272
aPTT , segundos	29 (27.1 – 31.1)	29 (27.3 – 30.9)	0.965
ACT Basal , segundos	140 (126 – 150)	139.5 (130 – 151.2)	0.593
ACT Heparina , segundos	488 (424 – 553)	498 (453.8 -574.2)	0.134
ACT Final , segundos	121 (110 – 132)	123 (112 – 135)	0.376

*Foi utilizado o teste de Mann-Whitney U para a comparação das amostras

5.3 Modelo de Regressão Binária Multivariada para Hemorragia Significativa

5.3.1 Seleção das variáveis explicativas

Numa primeira fase, para selecionar as variáveis explicativas para o modelo de regressão binária multivariada foram calculados os fatores de inflação da variância (FIV) através de um modelo mais complexo e que incluiu todas as variáveis explicativas presentes na base de dados. A base de dados utilizada para esta análise foi a base de dados original. Os valores obtidos encontram-se resumidos na Tabela 12. Esta decisão prende-se com o facto de existir uma preocupação com a possibilidade de multicolinearidade com algumas variáveis da amostra. Algumas variáveis, tais como, a área de superfície corporal e o índice de massa corporal estão intrinsecamente relacionadas e derivam da variável relativa ao peso do indivíduo, pelo que se antevem problemas de multicolinearidade. Deste modo, optou-se por esta abordagem por forma a decidir quais as variáveis a serem removidas. Como seria de esperar, foram encontrados valores bastante elevados de FIV para algumas variáveis. Foi dada especial atenção às variáveis com valores de FIV superiores a 5. As variáveis *Peso*, *Altura*, *IMC*, *IMC_catg2* e *ASC* obtiveram valores elevados de FIV, motivo pelo qual se optou por excluir todas as variáveis à exceção da variável categórica *IMC_catg2*, pois resume os graus relativos à definição de excesso de peso e de obesidade universalmente aceites [57], [58]. Para além disso, no que concerne às variáveis laboratoriais pré-operatórias verificou-se a existência de valores FIV aumentados para dois pares de variáveis. O primeiro diz respeito às variáveis *Hb_Pre* e *Htc_Pre*. De facto, o valor de hematócrito é diretamente proporcional aos valores analíticos de hemoglobina. Deste modo, optou-se pela manutenção da variável hematócrito, em detrimento da variável hemoglobina. Em segundo lugar, tem-se o par de variáveis *TP_Pre* e *INR_Pre*. É sabido que o INR é uma relação entre o tempo de protrombina do doente e um valor padrão do mesmo [73]. Deste modo, optou-se por manter a variável *INR_Pre*. Existem ainda preocupações entre o tempo de CEC e de clampagem da aorta. Uma vez que num procedimento cirúrgico com CEC todos os doentes têm ambos os tempos e, para além disso, o tempo de clampagem está contido no tempo de CEC, conforme mostra o diagrama da Figura 46, optou-se por eliminar a variável *Tempo_Clamp*. Por fim, existem preocupações acerca dos valores de FIV entre *Volume_Priming* e *TaxaDiluicao*. Tal poderá eventualmente dever-se ao facto de o volume de *priming* ser parte integrante da fórmula da taxa de diluição, como mostra a equação 4.1. Por este motivo, manteve-se a variável *TaxaDiluicao*. Em suma, as variáveis

Peso, Altura, IMC, ASC, Hb_Pre, TP_Pre, Tempo_Clamp e Volume_Priming foram removidas desta análise.

Figura 46 - Etapas da cirurgia com CEC

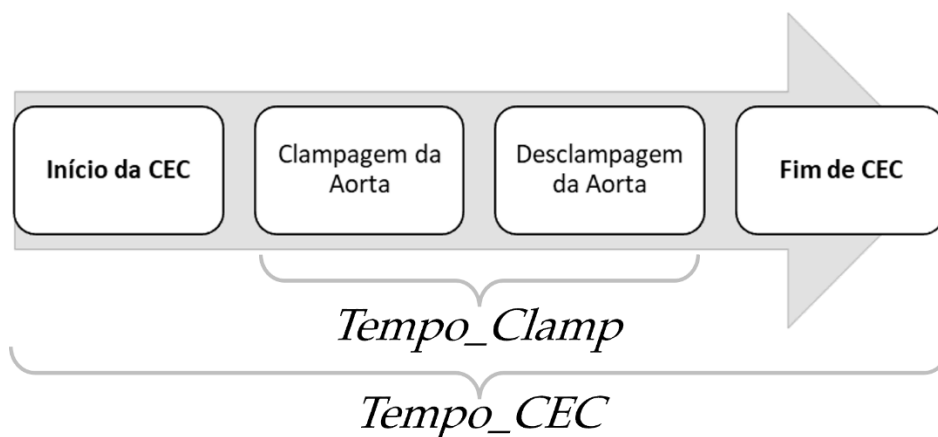


Tabela 12 - Fatores de inflação da variância das variáveis para o modelo de regressão binária

	GVIF	gl	GVIF ^{1/(2*Df)}
Variáveis individuais			
sexo	2,405	1	1,551
idade	1,574	1	1,254
Peso	401,900	1	20,047
Altura	102,009	1	10,100
IMC	86,237	1	9,286
IMC_catg2	6,745	2	1,612
ASC	590,164	1	24,293
diabetes	1,518	2	1,110
dislipidemia	1,335	1	1,155
hta	1,382	1	1,176
fumador	1,535	2	1,113
enfarte	1,290	1	1,136
dvp	1,128	1	1,062
irc2	1,811	1	1,077
feve2	1,635	1	1,063
Variáveis cirúrgicas			
cirurgia_aorta	1,499	1	1,224
peso_cirurgia	1,950	3	1,118
urgencia	1,359	1	1,166
Variáveis laboratoriais pré-operatórias			
Hb_Pre	15,563	1	3,945
Htc_Pre	14,651	1	3,828
Fib_Pre	1,883	1	1,372
Plaquetas_Pre	1,646	1	1,283
TP_Pre	234,676	1	15,319
INR_Pre	231,071	1	15,201

	aPTT_Pre		
	1,320	1	1,149
Variáveis da CEC			
ACT_Basal	2,564	1	1,601
ACT_Heparina	2,405	1	1,551
ACT_Final	2,289	1	1,513
Tempo_CEC	7,528	1	2,744
Tempo_Clamp	7,559	1	2,749
Dif_ACT_catg	2,632	1	1,622
Volume_Priming	46,579	1	6,825
Gelatina	1,233	1	1,110
Heparina_CEC	1,180	1	1,086
TaxaDiluicao	45,106	1	6,716
ACT_480	2,331	1	1,527

GVIF - Fator de Inflação de Variância Generalizado

gl - Graus de liberdade

GVIF^{1/(2*Df)} - Transformação do GVIF para os graus de liberdade

Para avaliar a eficácia do modelo, que visa prever a probabilidade de hemorragia significativa (*drenagem_sig*), inicialmente, foi desenvolvido um modelo nulo. O modelo nulo é uma ferramenta fundamental na análise estatística, pois fornece uma base de comparação para avaliar se o modelo de regressão posteriormente desenvolvido é estatisticamente significativo e diferente deste. O modelo nulo foi construído considerando apenas o intercepto, sem a inclusão de quaisquer variáveis preditoras. A estatística de teste z para o intercepto é altamente significativa ($z = -13.97, p < 0.001$), indicando que existem evidências de uma relação entre a variável resposta *drenagem_sig* e pelo menos uma das variáveis explicativas (ou seja, não incluídas neste modelo nulo).

Além disso, a análise da *deviance* revelou que tanto a *deviance* nula quanto a *deviance* residual são iguais a 670.55, com 657 graus de liberdade e com um AIC igual a 672.55. Isso indica que o modelo nulo não melhora a explicação dos dados em comparação com um modelo que simplesmente prevê a média da variável resposta.

Portanto, o próximo passo será desenvolver um modelo de regressão mais elaborado, incorporando variáveis explicativas potencialmente relevantes para prever a probabilidade de hemorragia significativa, e avaliar a sua adequação em comparação com este modelo nulo através de uma análise de *deviance*.

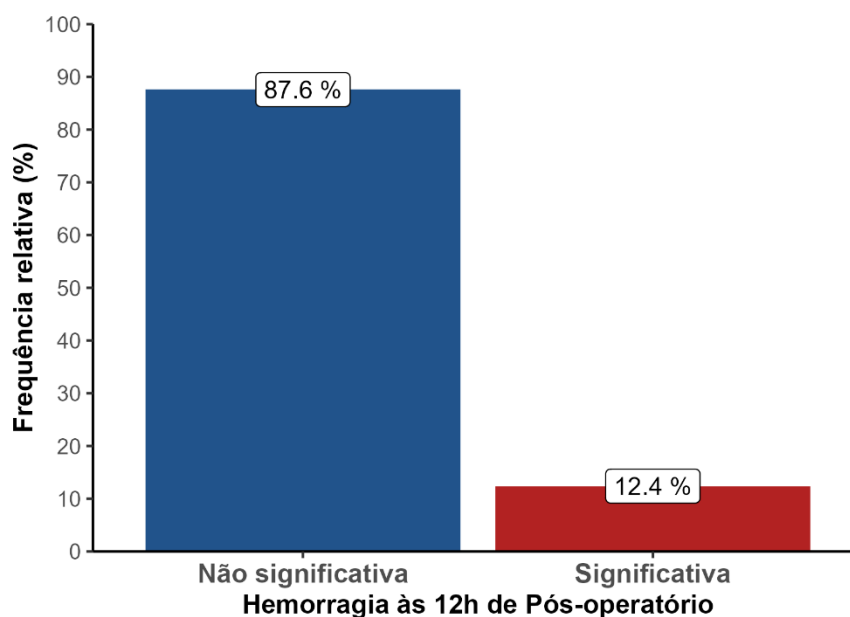
5.3.2 Modelo utilizando a *base de dados original*

Para a realização da modelação estatística procedeu-se à divisão dos dados em dois subconjuntos: dados de treino do modelo e dados de teste do modelo. Esta divisão teve como objetivo fundamental avaliar a capacidade de generalização do modelo para novos dados não utilizados no modelo de treino, evitando o *overfitting*, onde o modelo se ajusta excessivamente aos dados de treino, mas não consegue generalizar para novas observações.

Foi então aplicado o método de *Holdout* para a divisão dos dados nos dois subconjuntos. Foi definida a proporção dos dados que serão atribuídos ao conjunto de treino. Neste caso, 80% dos dados serão usados para o treino do modelo e passam a configurar a base de dados de treino. Foi definida uma semente para garantir a reprodutibilidade dos resultados, que neste caso foi utilizado o valor 2024. Foi utilizada a função *createDataPartition* do pacote *caret* para criar os índices que dividem os dados em conjuntos de treino e de teste. A variável *drenagem_sig* foi utilizada como a variável de resposta para garantir que a divisão dos dados seja estratificada em relação a esta variável.

Assim, foi criada a base de dados de treino, seleccionando as linhas dos dados originais correspondentes aos índices gerados para o treino do modelo (índices *holdout* de treino). Para além disso, foi criado o conjunto de teste, seleccionando apenas as linhas dos dados originais que não foram incluídas no conjunto de treino. Em suma, o subconjunto de treino foi a porção dos dados utilizado para treinar o modelo preditivo. Isto significa que o modelo foi ajustado aos padrões presentes neste subconjunto, capturando as relações entre as variáveis independentes (preditoras) e a variável dependente (alvo). Tal verificação comprova-se com a informação presente na Figura 47, onde se pode verificar que a frequência absoluta da categoria de casos (hemorragia significativa) mantém-se inalterada quando comparada à da Figura 5, antes da divisão da base de dados.

Figura 47 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório da base de dados de treino



Após a análise e interpretação da Tabela 12 relativa aos fatores de inflação da variância, procedeu-se à construção de um modelo de regressão complementar log-log binária com as variáveis explicativas que constam na Tabela 13. Os dados utilizados constam na base de dados de treino obtida através dos dados originais.

Tabela 13 - Variáveis incluídas no modelo

	Variáveis contínuas	Variáveis categóricas
Variáveis individuais	idade	sexo
		IMC_catg2
		diabetes
		dislipidemia
		hta
		fumador
		enfarte
		dvp
		irc2
feve2		
Variáveis cirúrgicas		cirurgia_aorta
		urgencia
Variáveis laboratoriais pré-operatórias	Htc_Pre	
	Fib_Pre	
	Plaquetas_Pre	

Variáveis da CEC	INR_Pre	
	aPTT_Pre	
	Tempo_CEC	Gelatina
	TaxaDiluicao	Heparina_CEC

Realizou-se então um modelo de regressão binária para investigar a relação entre a hemorragia pós-operatória significativa, isto é, uma drenagem superior ou igual a 600 mL às 12 horas de pós-operatório (variável dependente) e um conjunto alargado de potenciais variáveis explicativas. Os resultados encontram-se na Tabela 14. Utilizou-se a função *glm()* (do inglês, *Generalized Linear Models*) disponível na linguagem de programação R.

A primeira abordagem passou pela construção de um *modelo completo* em que foram incluídas todas as variáveis explicativas anteriormente enumeradas. Este primeiro modelo foi ajustado utilizando a distribuição binomial com a função de ligação *loglog*, adequada para variáveis de resposta binárias desbalanceadas. A segunda abordagem passou pela seleção automática de variáveis para identificar um modelo mais parcimonioso e com melhor ajuste aos dados, o qual se apelidou de *modelo stepwise*. Foi então utilizado o método de seleção *stepwise*, onde as variáveis foram incluídas e excluídas do modelo com base no critério de informação de *Akaike* (AIC). O procedimento *stepwise* reduziu o modelo original, removendo variáveis que não contribuíam significativamente para explicar a variabilidade na variável resposta. Os resultados de ambos os modelos encontram-se na Tabela 14.

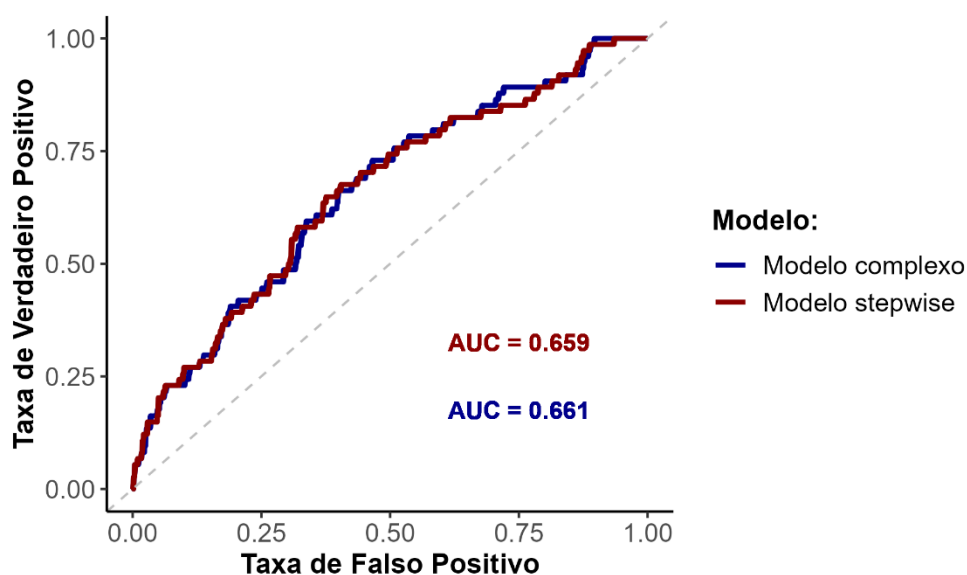
Tabela 14 - Resultados obtidos com os dois modelos estimados com a base de dados de treino

	Modelo complexo				Modelo stepwise		
	Estimativa parâmetro	Erro padrão	Valor-p		Estimativa parâmetro	Erro padrão	Valor-p
(Intercepto)	-1,5766	2,0623	0,445		-1,3546	1,0358	0,191
Variáveis Individuais							
sexo (Masculino)	0,7799	0,2931	0,008	**	0,8678	0,2740	0,002
idade	0,0027	0,0129	0,836				
IMC_catg2 (Excesso de peso)	-0,2207	0,3063	0,471				
IMC_catg2 (Obesidade)	-0,1559	0,3419	0,648				
diabetes (Diabetes NIT)	-0,1295	0,3035	0,700				
diabetes (Diabetes IT)	0,4911	0,6446	0,446				
dislipidemia (Sim)	0,3639	0,3000	0,225				
hta (Sim)	0,3600	0,3765	0,339				
fumador (Ex-fumador)	-0,0371	0,3181	0,907				
fumador (Fumador)	-0,2307	0,4670	0,621				
enfarte (Sim)	0,7706	0,3792	0,042	*	0,8095	0,3411	0,018

dvpc (Sim)	-0,1771	0,5379	0,742				
irc2 (Com IRC)	0,1848	0,3238	0,568				
feve2 (FEVE diminuída)	0,4508	0,3135	0,150				
Variáveis cirúrgicas							
cirurgia_aorta (Sim)	0,3320	0,3192	0,298				
urgencia (Não programada)	0,6111	0,4216	0,147				
Variáveis laboratoriais pré-operatórias							
Htc_Pre	-0,0456	0,0286	0,111		-0,0489	0,0256	0,056
Fib_Pre	-0,0003	0,0013	0,817				
Plaquetas_Pre	-0,0018	0,0021	0,383				
INR_Pre	0,8987	0,6692	0,179				
aPTT_Pre	-0,0361	0,0296	0,222				
Variáveis da CEC							
Tempo_CEC	0,0069	0,0035	0,048	*	0,0076	0,0031	0,015
Heparina_CEC (Com Heparina)	-0,1895	0,2874	0,510				
Gelatina (Com Gelatina)	0,1905	0,3106	0,540		0,1390	0,2998	0,643
TaxaDiluicao	0,0007	0,0151	0,964				
Null deviance (gl)		447.16 (595)				447.16 (595)	
Residual deviance (gl)		409.05 (570)				423.50 (590)	
AIC		461.05				435.5	
1-pchisq(deviance, df.residual)		0.999				1	

Ambos os modelos não apresentam significância estatística para o intercepto, conforme evidenciado pelos valores-p dos testes de significância elevados (acima de 0.05). Nos dois modelos a variável *sexo* e *enfarte* estão significativamente associadas a uma maior probabilidade de hemorragia pós-operatória significativa (estimativas de parâmetro positivas e valores-p < 0.05). Assim, os modelos obtidos sugerem que o sexo masculino e o antecedente clínico de enfarte agudo do miocárdio estão associados a uma maior probabilidade de hemorragia significativa. O tempo de CEC foi também significativo em ambos os modelos, apresentando uma estimativa de parâmetro positiva. Todas as restantes variáveis utilizadas na modelação aparentam não ser significativas para explicar a variável dependente, verificando-se que foram removidas pela seleção *stepwise*. A variável referente à utilização de gelatina do *priming*, no qual se centra o problema do presente trabalho, foi mantida propositadamente no modelo *stepwise* por forma a dar resposta ao mesmo na secção referente à sua interpretação e estimação. Por fim, os dois modelos apresentam valores de *deviance* e AIC semelhantes embora se verifique uma redução com o modelo *stepwise*, como seria de esperar, indicando que o modelo se ajusta bem aos dados de treino. Para além disso, e conforme se verifica na Figura 48, os dois modelos originam AUCs muito aproximados (66.1 % para o modelo complexo e 65.9 % para o modelo *stepwise*). Dado que o modelo *stepwise* tem melhor desempenho em termos de estatísticas de ajuste, tais como o AIC e de resíduos, optou-se por prosseguir a análise com este modelo.

Figura 48 - Curvas ROC dos dois modelos estimados com dados de treino



Foi efetuada uma análise de *deviance* para comparar o modelo *stepwise* com o modelo nulo, anteriormente mencionado. A análise revelou uma significativa redução na *deviance residual* do modelo *stepwise* em comparação com o modelo nulo (Δ Deviance = 23.664, $\chi = 23.664$, $df = 5$, $p < 0.001$), indicando que as variáveis explicativas adicionadas contribuíram de forma significativa para a explicação da variabilidade da variável dependente, isto é, na explicação da hemorragia pós-operatória significativa às 12 horas de pós-procedimento. Os resultados sugerem assim, que o modelo *stepwise* é estatisticamente superior a um modelo nulo na predição da hemorragia.

5.3.3 Modelo utilizando a *base de dados balanceados*

Como se verifica na Figura 5, existe um desequilíbrio entre as duas categorias de hemorragia (significativa vs. não significativa). Por forma a lidar com o desequilíbrio destas classes, foi utilizado o método de *oversampling*. O objetivo foi aumentar o número de observações na classe minoritária, isto é, da classe correspondente à hemorragia significativa. Assim, propõem-se ajustar um melhor equilíbrio entre as duas classes sem que sejam perdidas informações importantes da classe maioritária, motivo pelo qual não se optou por recorrer a técnicas de *undersampling*. Esta outra técnica iria remover informação à categoria de hemorragia não significativa.

O procedimento de *oversampling* foi realizado utilizando a função *ovun.sample()* da biblioteca *ROSE* do *software*. Esta função aplica tanto técnicas de *undersampling* como de *oversampling* para equilibrar as classes de acordo com alguns parâmetros especificados. Neste caso concreto, foi utilizado o método "*over*" para aumentar o número de observações na classe minoritária. Foram criadas duas bases de dados balanceadas. Uma primeira, onde foi especificado um parâmetro "p" igual a 0.2, o que significa que a classe minoritária passará a ter, aproximadamente, 20% do tamanho da amostra após o *oversampling*. Este valor representa uma decisão de não exagerar na duplicação das observações da classe minoritária que na base de dados original é de aproximadamente 12%. Uma segunda base de dados foi criada com a especificação do parâmetro "p" igual a 0.5, o que significa que a classe minoritária passará a ter, aproximadamente, 50% do tamanho da amostra após o *oversampling*. Esta base de dados pretende exemplificar como o modelo de regressão binária se comportaria com a replicação de dados para uma base de dados completamente balanceada. O processo de *oversampling* foi realizado utilizando uma semente (*seed*) específica e igual a 2024 por forma a garantir a reprodutibilidade dos resultados. Após o procedimento, o conjunto de dados foi sobre amostrado, resultando em dois conjuntos de dados mais equilibrados em termos da distribuição de classes. À primeira deu-se o nome de *base de dados parcialmente balanceados* e à segunda, deu-se o nome de *base de dados completamente balanceados*.

Também aqui para a realização da modelação estatística com dados balanceados procedeu-se à divisão dos dados em dois subconjuntos: dados de treino do modelo e dados de teste do modelo. Foi também aplicado o método de *Holdout* para a divisão dos dados nos dois subconjuntos, mantendo-se a proporção 80/20 e foi definida a mesma semente, ou seja, o valor 2024. De seguida procedeu-se da mesma forma que a mencionada na página 68.

Tal verificação comprova-se com a informação presente na Figura 49 e na Figura 50, onde se pode verificar que a frequência relativa da categoria de casos (hemorragia significativa) aumentou para 20% na *base de dados parcialmente balanceados* e para 50% na *base de dados completamente balanceados*, respetivamente. Constata-se assim, que estes valores são superiores aos encontrados na base de dados original e listados na Figura 5.

Figura 49 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório da base de dados parcialmente balanceada

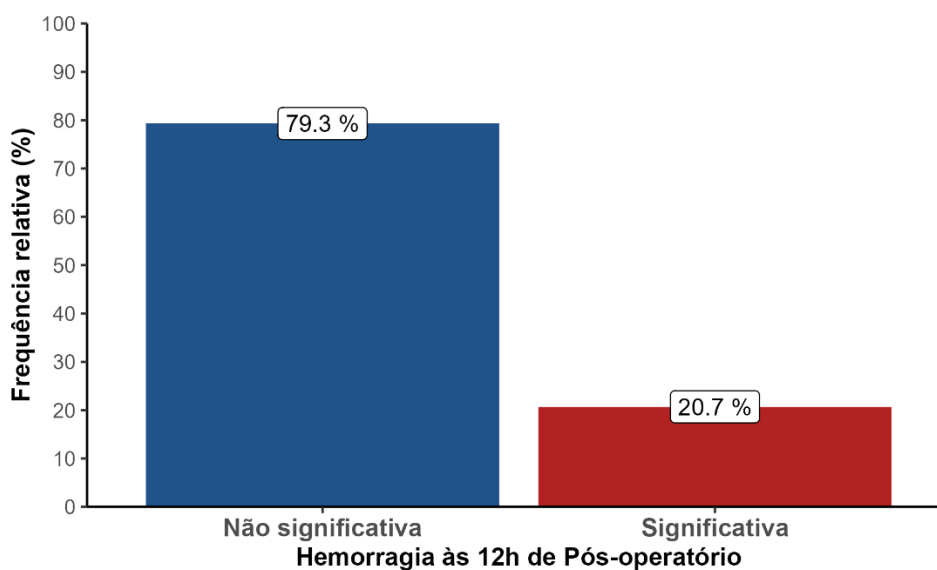
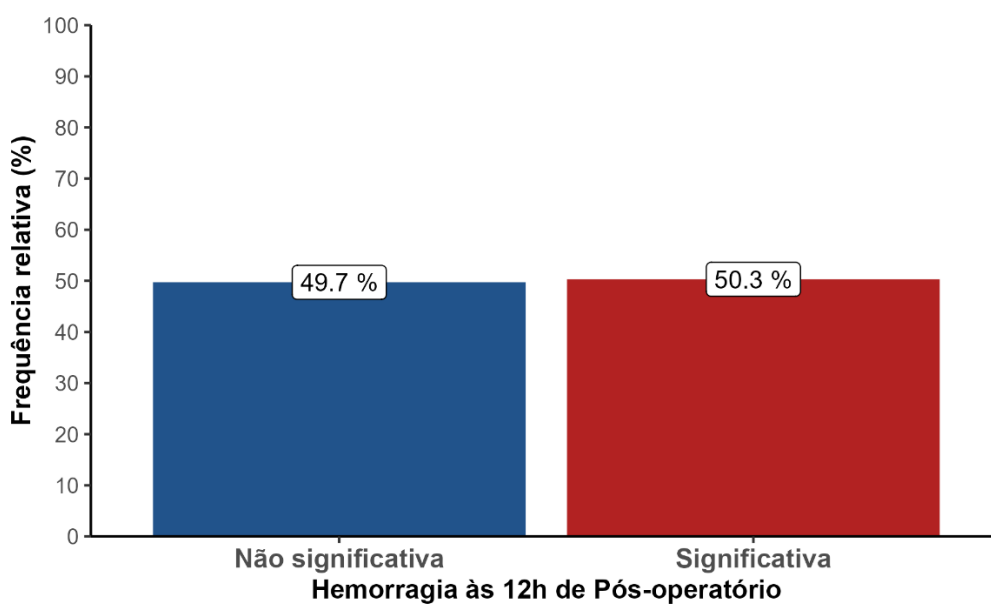


Figura 50 - Variável dependente - Categorização binária da hemorragia às 12 horas de pós-operatório da base de dados completamente balanceada



Realizaram-se então dois modelos de regressão binária para investigar a relação entre a hemorragia pós-operatória significativa às 12 horas de pós-operatório (variável dependente) e o mesmo conjunto alargado de potenciais variáveis explicativas e que constam na Tabela 13. Em ambos os modelos foi utilizada a função *glm*.

A primeira abordagem passou pela construção de um *modelo completo* em que foram incluídas todas as variáveis explicativas que constam na Tabela 13. Este primeiro modelo foi ajustado utilizando a distribuição binomial com a função de ligação *cloglog*, por forma a permitir a posterior comparação com o modelo em que se usou a *base de dados original*. A segunda abordagem passou pela seleção automática de variáveis para identificar um modelo mais parcimonioso e com melhor ajuste aos dados ao qual se apelidou também de *modelo stepwise*. O procedimento *stepwise* reduziu o modelo original, removendo variáveis que não contribuíam significativamente para explicar a variabilidade na variável resposta. Esta metodologia aplicou-se de igual forma para as duas bases de dados balanceadas. Contudo, por forma a resumir melhor a informação deste trabalho, apenas foram resumidos, na Tabela 15, os dados provenientes dos modelos *stepwise* de cada uma das bases de dados balanceadas.

Tabela 15 - Resultados obtidos com os dois modelos estimados após stepwise

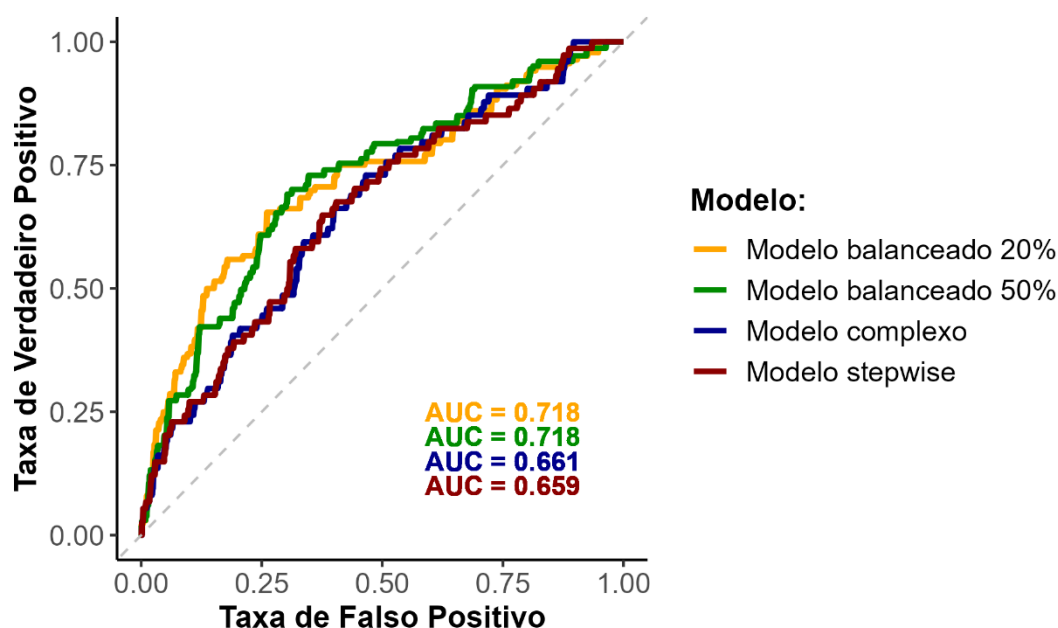
	Base de dados parcialmente balanceados (Modelo stepwise)			Base de dados completamente balanceados (Modelo stepwise)				
	Estimativa parâmetro	Erro padrão	Valor-p	Estimativa parâmetro	Erro padrão	Valor-p		
(Intercepto)	-2,3830	1,3243	0,072	.	1,5174	0,7360	0,039	*
Variáveis Individuais								
sexo (Masculino)	0,9397	0,2120	<0.001	***	0,5373	0,1070	<0.001	***
idade	0,0175	0,0095	0,065	.	0,0091	0,0049	0,063	.
hta (Sim)	0,4031	0,2790	0,148		0,3249	0,1443	0,024	*
enfarte (Sim)	0,9704	0,2448	<0.001	***	0,6262	0,1456	<0.001	***
irc2 (Com IRC)	0,4322	0,2085	0,038	*	0,1745	0,1179	0,139	
feve2 (Diminuída)					0,1745	0,1179	0,139	
Variáveis cirúrgicas								
cirurgia_aorta (Sim)	0,4470	0,2183	0,041	*	0,3607	0,1213	0,003	**
urgencia (Não programada)					0,3689	0,1679	0,028	*
Variáveis laboratoriais pré-operatórias								
Htc_Pre	-0,0342	0,0204	0,094	.	-0,0570	0,0112	<0.001	***
Fib_Pre					-0,0009	0,0005	0,056	.
Plaquetas_Pre					-0,0014	0,0008	0,081	.
aPTT_Pre	-0,0255	0,0227	0,261		-0,0257	0,0138	0,062	.
Variáveis da CEC								
Tempo_CEC	0,0062	0,0027	0,021	*	0,0053	0,0015	<0.001	***
Gelatina (Com Gelatina)	0,3292	0,2514	0,190		-0,1306	0,1147	0,255	
TaxaDiluicao	-0,0160	0,0106	0,134					
Null deviance (gl)		670.55 (657)			1455.6 (1049)			
Residual deviance (gl)		598.69 (646)			1312.4 (1035)			
AIC		622.69			1342.1			

$1 - \text{pchisq}(\text{deviance}, \text{df.residual})$	0,908	<0,001
--	-------	--------

Os resultados obtidos revelam diferenças significativas entre os dois modelos. No modelo que deriva da *base de dados parcialmente balanceados*, observa-se que a variável *sexo*, *enfarte*, *irc2*, *cirurgia_aorta* e *Tempo_CEC* foram significativas para prever a variável de interesse, enquanto a *idade* e o *Htc_Pre* mostraram apenas uma tendência de associação à variável dependente. Por outro lado, no modelo construído a partir da base de dados completamente balanceada, além das variáveis já mencionadas como significativas, as variáveis *idade*, *hta*, *Htc_Pre*, *Fib_Pre*, *aPTT_Pre*, *urgencia* e *feve2* também surgem como significativas na explicação da hemorragia pós-operatória. Para além disso, verificou-se que, para o modelo que tem origem na *base de dados completamente balanceados* o intercepto é estatisticamente significativo. Os dois modelos diferem bastante no que concerne aos valores de *deviance* e *AIC*, com o modelo da *base de dados completamente balanceada* a apresentar os maiores valores. Por fim, o teste de adequabilidade não é igual para os dois modelos. Mais uma vez, o modelo da *base de dados completamente balanceados* indica que o modelo não se ajusta bem aos dados observados, ao contrário do que se observou no modelo da *base de dados parcialmente balanceados*.

Conforme se verifica na Figura 51, os dois modelos estimados a partir das duas bases de dados balanceadas originam AUCs iguais (71.8% para ambos os modelos). Para além disso, verifica-se que estes dois modelos têm valores de AUC superiores aos do modelo com a base de dados original quer para o modelo complexo, quer para o *stepwise*.

Figura 51 - Curvas ROC dos quatro modelos estimados



A análise de *deviance* foi realizada para comparar dois modelos de regressão binária no contexto da previsão da variável de resposta *drenagem_sig*. O primeiro modelo utilizado para comparar foi o modelo nulo, enquanto o segundo modelo utilizado nesta análise foi o modelo da *base de dados parcialmente balanceados*. A análise revelou uma significativa redução na *deviance residual* deste modelo em comparação com o modelo nulo (Δ Deviance = 71.853, $\chi = 71.853$, $df = 11$, $p < 0.001$), indicando que as variáveis explicativas adicionadas contribuíram de forma significativa para a explicação da variabilidade da variável dependente. Os resultados sugerem assim, que o modelo é estatisticamente superior a um modelo nulo na predição da hemorragia. Posteriormente foi realizada outra análise de *deviance* mas entre o respetivo modelo nulo e o modelo que utilizou a *base de dados completamente balanceados*. A análise revelou também uma significativa redução na *deviance residual* deste modelo em comparação com o modelo nulo (Δ Deviance = 152.79, $\chi = 152.79$, $df = 25$, $p < 0.001$), indicando que as variáveis explicativas adicionadas contribuíram de forma significativa para a explicação da variabilidade da variável dependente.

5.3.4 Interpretação dos modelos

5.3.4.1 Modelo utilizando a *base de dados original*

A interpretação dos resultados obtidos a partir do modelo incluirá a análise dos coeficientes estimados, os seus intervalos de confiança e significância estatística. Isso permitirá identificar quais as variáveis que são estatisticamente significativas na predição da ocorrência de hemorragia significativa pós-operatória, bem como a sua contribuição relativa para o modelo.

O modelo final (Modelo *stepwise*, presente na Tabela 14) incluiu apenas as seguintes variáveis explicativas, ordenadas pela sua categoria: (i) 2 variáveis individuais: Sexo (**, $p = 0.015$) e enfarte agudo do miocárdio (*, $p = 0.018$); (ii) 1 variável laboratorial pré-operatória: Hematócrito (., $p = 0.056$); e (iii) 2 variáveis da CEC: Tempo de CEC (*, $p = 0.015$) e utilização de gelatina no *priming* (ns, $p = 0.643$). Destas variáveis apenas *sexo*, *enfarte* e tempo de CEC são estatisticamente significativas na explicação da hemorragia significativa. A variável referente ao hematócrito pré-operatório encontra-se no limite da significância ao nível de alfa igual a 0.05.

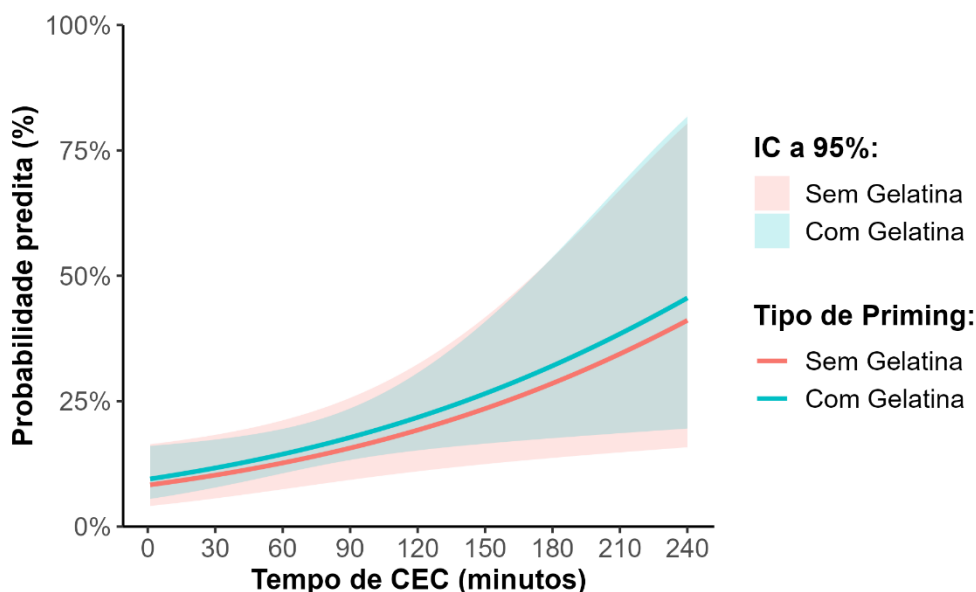
Para as variáveis qualitativas *sexo* e *enfarte* obtiveram-se coeficientes positivos para as respetivas categorias de referência (*sexo* masculino e presença de *enfarte*) o que evidencia que estas variáveis têm, de forma independente, um risco hemorrágico superior, quando comparado com a outra categoria. A variável explicativa relativa ao tempo de CEC também produziu um coeficiente positivo pelo que se pode concluir, desde já, que o aumento do tempo de CEC está associado a um aumento do risco de hemorragia significativa. Todas as outras variáveis não foram significativas no modelo. A variável de interesse para dar resposta ao problema levantado por este trabalho (tipo de *priming*) não se revelou explicativa da hemorragia pós-operatória (ns, $p = 0.643$). De um modo geral, os resultados sugerem que o *sexo* masculino (em comparação ao feminino), a presença de antecedente de *enfarte* agudo do miocárdio e o tempo de CEC podem aumentar o risco de hemorragia pós-operatória. A utilização de Gelatina no *priming* do circuito de CEC não aparenta ser um preditor significativo de hemorragia com este modelo.

Por forma a dar resposta ao problema levantado por este trabalho, procedeu-se à estimação das probabilidades preditas para o desfecho relacionado com a variável dependente deste modelo. Esta comparação foi realizada por grupo de *priming*, isto é, procedeu-se à

comparação das probabilidades preditas de hemorragia para o grupo de doentes com e sem gelatina no *priming*. Para tal, foi criado um subconjunto de dados que consiste em combinações de valores fixos e outros variáveis das variáveis explicativas para fazer a previsão para diferentes tempos de CEC. As variáveis com valores fixos incluíram o sexo masculino e a ausência de enfarte. Pretendeu-se com esta seleção reduzir ao máximo o enviesado do aumento da probabilidade provocado por estas variáveis que mostraram serem significativas e positivas na sua contribuição para o desfecho de hemorragia significativa. Para além disso, fixou-se o hematócrito no seu valor mediano na base de dados de treino (isto é, 40.2%). Fizeram-se então variar apenas as variáveis do tipo de *priming* e do tempo de CEC. A primeira, o tipo de *priming*, variou entre as duas categorias de resposta da variável (com vs. sem gelatina) e para a segunda, o tempo de CEC, foi criada uma sequência de valores entre 1 e 240 minutos com variação de uma unidade. Após a criação deste subconjunto de dados foi aplicado o modelo previamente ajustado (modelo *stepwise* da *base de dados original*) para prever a probabilidade do evento "hemorragia significativa" ocorrer. O tipo de previsão utilizada foi de "*link*", o que significa que foram obtidas as probabilidades na escala *cloglog*. Além de prever as probabilidades, também foram calculados os intervalos de confiança para as mesmas. Os intervalos de confiança foram estimados para assinalar a incerteza associada às estimativas das probabilidades. Uma vez que as probabilidades foram obtidas numa escala *cloglog*, procedeu-se à sua transformação de volta para a escala de probabilidades programando o *software* R para a fórmula apresentada nos métodos na equação 3.11. Esta conversão foi executada por forma a interpretar as probabilidades de uma perspectiva mais intuitiva, pois as probabilidades variam de 0 a 1, convertidos, de 0 a 100%.

Com o descrito anteriormente, obteve-se o gráfico de probabilidades preditas presente na Figura 52. Como se pode verificar o gráfico é composto por 2 linhas que representam as probabilidades preditas para cada um dos grupos. Ao longo de todos os valores de tempo de CEC verifica-se que as 2 linhas se acompanham mutuamente assim como os respetivos intervalos de confiança a 95%. Verificando, por exemplo, aos 60 minutos de CEC, um doente com as restantes características descritas no parágrafo anterior e que não tenha recebido gelatina no *priming* tem uma probabilidade de hemorragia significativa às 12 horas de pós-procedimento de 12.7% (IC 95%: 7.4 - 21.1%), enquanto um doente com gelatina tem 14.4% (IC 95%: 10.6 - 19.5%). Como se pode verificar, os intervalos de confiança encontram-se bastante sobrepostos em ambos os grupos dando a percepção de que não existe diferenças na probabilidade de hemorragia significativa entre os grupos.

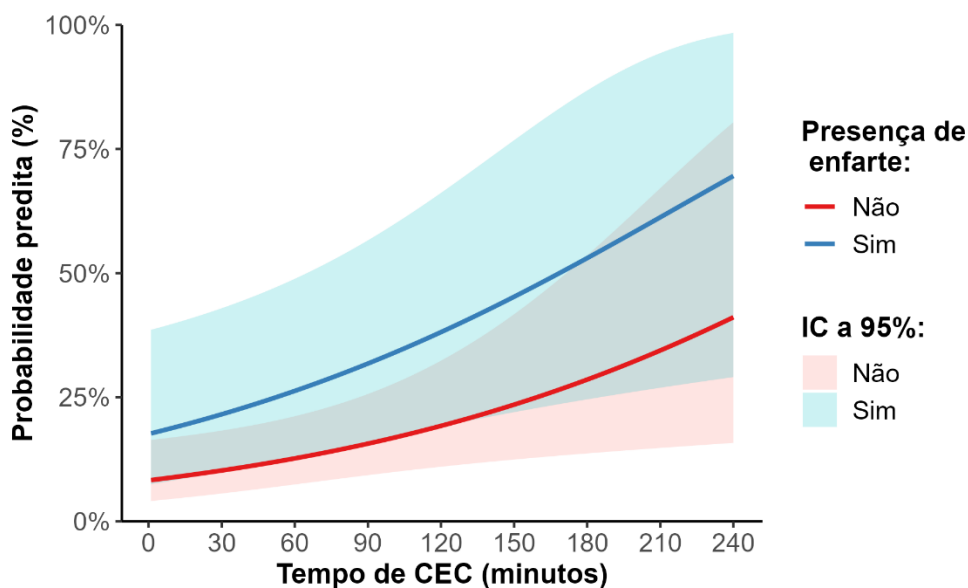
Figura 52 - Probabilidade predita de hemorragia significativa por tipo de *priming* e por tempo de CEC



Em segundo lugar, procedeu-se à estimação das probabilidades preditas para uma das variáveis com maior impacto em termos de *Odds Ratio*, o *enfarte*. Como tal, foi criado um segundo subconjunto de dados. As variáveis fixas incluíram o sexo masculino e a ausência de gelatina no *priming*. Para além disso, fixou-se o hematócrito nos seus valores medianos dentro da amostra de dados de treino (isto é, 40.2%). Assim, fizeram-se variar apenas as variáveis *enfarte* e tempo de CEC. O *enfarte* variou entre as duas categorias de resposta da variável (com vs. sem *enfarte* nos antecedentes clínicos) e para o tempo de CEC, foi criada uma sequência de valores entre 1 e 240 minutos com variação de uma unidade. Após a criação deste subconjunto de dados foi aplicada a mesma metodologia anteriormente descrita.

Obteve-se, assim, o gráfico de probabilidades preditas presente na Figura 53. Verificando, uma vez mais, que com 60 minutos de CEC, um doente com as restantes características descritas no parágrafo anterior e que não tenha antecedentes de *enfarte* agudo do miocárdio tem uma probabilidade de hemorragia significativa às 12 horas de pós-procedimento de 12.7% (IC 95%: 7.4 - 21.2%), enquanto um doente com esse antecedente tem 26.3% (IC 95%: 12.9 - 48.8%).

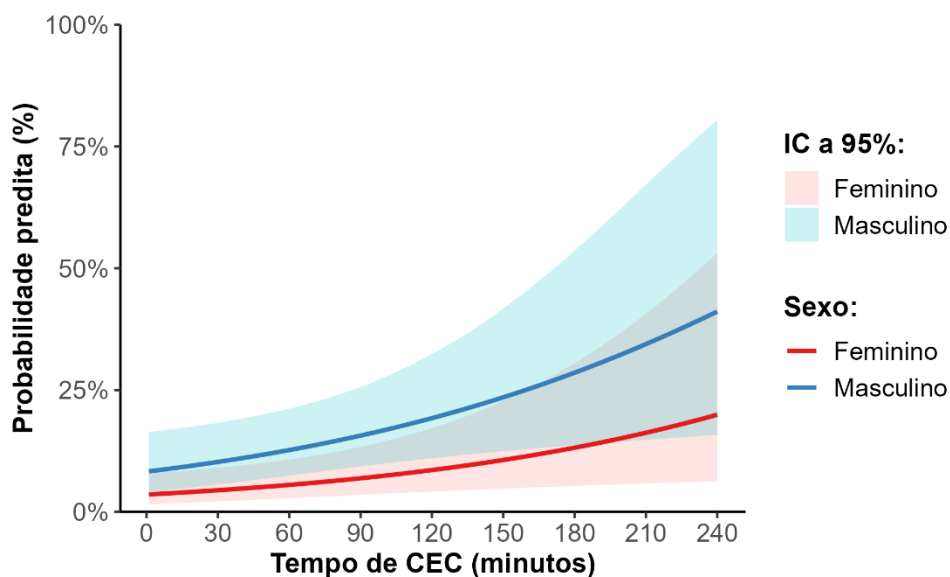
Figura 53 - Probabilidade predita de hemorragia significativa por enfarte e por tempo de CEC



Procedeu-se também à estimação das probabilidades preditas, considerando o sexo do doente. Como tal, foi criado um terceiro subconjunto de dados. As variáveis com valores fixados incluíram a ausência de enfarte e de gelatina no *priming*. Para além disso, fixou-se o hematócrito no seu valor mediano. Assim, fizeram-se variar apenas as variáveis relativas ao sexo do doente e ao tempo de CEC. A primeira, varia entre as duas categorias de resposta de cada uma (feminino vs. masculino) e para a última, o tempo de CEC, foi também utilizada a sequência de valores entre 1 e 240 minutos.

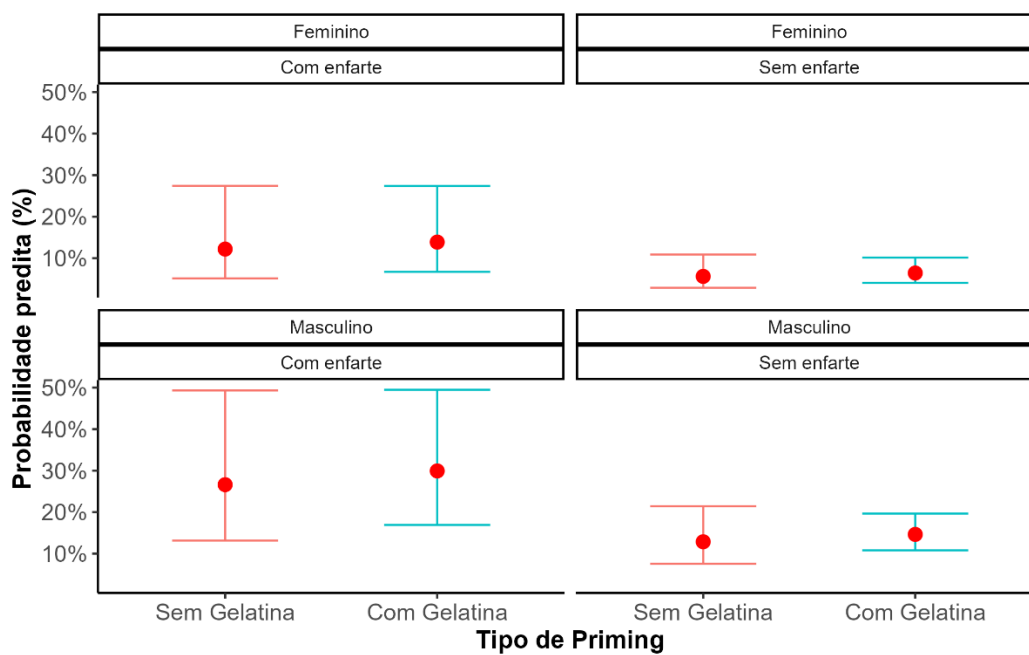
Obteve-se, assim, o gráfico de probabilidades preditas presente na Figura 54. Assim sendo, aos 60 minutos de CEC, um doente com as restantes características descritas no parágrafo anterior, mas do sexo feminino tem uma probabilidade de hemorragia significativa de 5.5% (IC 95%: 2.8 - 10.7%), enquanto um doente do sexo masculino tem 12.7% (IC 95%: 7.4 - 21.1%).

Figura 54 - Probabilidade predita de hemorragia significativa por sexo e por tempo de CEC



O gráfico da Figura 55 representa as probabilidades preditas para hemorragia significativa e os respectivos intervalos de confiança para quatro cenários distintos de acordo com as combinações possíveis entre as variáveis sexo e enfarte agudo do miocárdio. Para além disso, fixaram-se os restantes parâmetros (hematócrito e tempo de CEC) no seu valor mediano. Assim, apresenta-se para cada uma das combinações as probabilidades por tipo de *priming*. Uma vez mais é possível confirmar que o *priming* com gelatina não está associado a um aumento da probabilidade predita para hemorragia significativa às 12 horas de pós-operatório para nenhuma das combinações apresentadas.

Figura 55 - Probabilidades previstas de acordo com o modelo para todos os cenários de combinações de variáveis categóricas



5.3.4.2 Modelo utilizando a *base de dados parcialmente balanceados*

O modelo estimado com a *base de dados completamente balanceados* não mostrou ser adequado apesar de ter originado uma AUC satisfatória e ter tido um bom comportamento no que diz respeito à análise de *deviance*. Como tal apenas será utilizado para interpretação e estimação o modelo da *base de dados parcialmente balanceados*. Contudo, foram calculados os resultados relativos à validação cruzada e de outras métricas, pelo que serão apresentados mais à frente na secção relativa à avaliação dos modelos.

O modelo final incluiu apenas as seguintes variáveis explicativas ordenadas pela sua categoria de variáveis: (i) 5 variáveis individuais: sexo (***, $p < 0.01$), idade (., $p = 0.06475$), hipertensão arterial (ns, $p = 0.015$), enfarte agudo do miocárdio (***, $p < 0.01$) e insuficiência renal crónica (*, $p = 0.038$); (ii) 1 variável cirúrgica: cirurgia da aorta (*, $p = 0.040$); (iii) 2 variáveis laboratoriais pré-operatórias: Hematócrito (., $p = 0.093$) e aPTT (ns, $p = 0.261$); e (iv) 3 variáveis da CEC: tempo de CEC (*, $p = 0.021$), Taxa de diluição (ns, $p = 0.134$) e utilização de gelatina no *priming* (ns, $p = 0.190$). Destas variáveis apenas sexo, enfarte, insuficiência renal crónica, cirurgia da aorta e tempo de CEC são estatisticamente significativas na explicação da hemorragia.

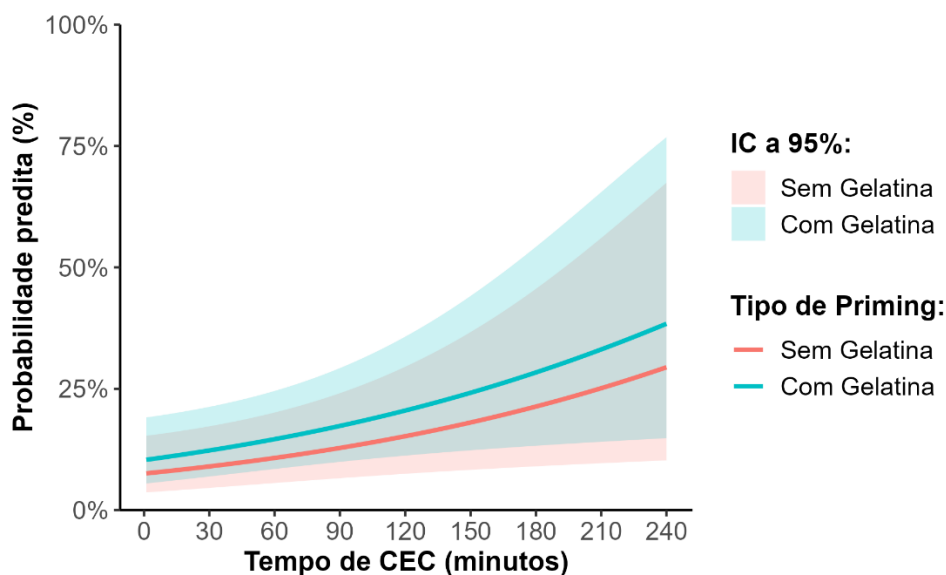
Para as variáveis qualitativas sexo, enfarte, IRC e cirurgia da aorta obtiveram-se coeficientes positivos para as respetivas categorias de referência (sexo masculino, presença de enfarte, presença de IRC e cirurgia que envolva a artéria aorta) o que evidencia que estas variáveis têm, de forma independente, um risco hemorrágico superior, quando comparado com a outra categoria. A variável explicativa relativa ao tempo de CEC gerou um coeficiente positivo pelo que se pode concluir, desde já, que o aumento do tempo de CEC está associado a um aumento do risco de hemorragia significativa. Todas as outras variáveis não foram significativas no modelo final. A variável de interesse para dar resposta ao problema levantado por este trabalho (tipo de *priming*) não se revelou explicativa da hemorragia pós-operatória (ns, $p = 0.190$). A utilização de gelatina no *priming* do circuito de CEC não aparenta ser preditora de hemorragia significativa no pós-operatório.

Uma vez mais, procedeu-se à estimação das probabilidades preditas para o desfecho relacionado com a variável dependente deste modelo (hemorragia significativa). Esta comparação foi realizada por grupo de *priming*, isto é, procedeu-se à comparação das probabilidades preditas de hemorragia para o grupo de doentes com e sem gelatina no *priming*. Para tal, foi criado um subconjunto de dados que consiste em combinação de valores fixos e outros variáveis das variáveis explicativas para fazer a previsão, para diferentes tempos

de CEC. As variáveis com valores fixos incluíram o sexo masculino, a ausência de hipertensão arterial, enfarte, IRC e de cirurgia da aorta. Para além disso, fixou-se a idade, o hematócrito, o aPTT, a taxa de diluição nos seus valores medianos dentro da amostra de dados de treino utilizada neste modelo (isto é, 71 anos, 40.2%, 29 segundos e 14.6%, respetivamente). Fizeram-se então variar apenas as variáveis tipo de *priming* e tempo de CEC. A primeira, o tipo de *priming* variou entre as duas categorias de resposta da variável (com vs. sem gelatina) e para a segunda, o tempo de CEC, foi criada uma sequência de valores entre 1 e 240 minutos com variação de uma unidade. Após a criação deste subconjunto de dados foi aplicado o modelo previamente ajustado (modelo *stepwise* da *base de dados parcialmente balanceados*) para prever as probabilidades do evento "hemorragia significativa" ocorrer. O tipo de previsão utilizada foi de "link", o que significa que foram obtidas as probabilidades na escala *cloglog*. Além de prever as probabilidades, também foram calculados os intervalos de confiança para as mesmas. Uma vez que as probabilidades foram obtidas numa escala *cloglog*, procedeu-se à sua transformação de volta para a escala de probabilidades utilizando a mesma metodologia enunciada para o modelo anterior.

Com o descrito anteriormente, obteve-se o gráfico de probabilidades preditas presente na Figura 56. Como se pode verificar o gráfico é composto por 2 linhas que representam as probabilidades preditas para cada um dos grupos. Ao longo de todos os valores de tempo de CEC verifica-se que as 2 linhas se acompanham mutuamente assim como os respetivos intervalos de confiança a 95%. Verificando, por exemplo, que aos 60 minutos de CEC, um doente com as restantes características descritas no parágrafo anterior e que não tenha recebido gelatina no *priming* tem uma probabilidade de hemorragia às 12 horas de pós-procedimento de 10.8% (IC 95%: 5.6 - 20.1%), enquanto um doente com gelatina tem 14.6% (IC 95%: 8.5 - 24.5%). Como se pode verificar, os intervalos de confiança encontram-se bastante sobrepostos em ambos os grupos dando a perceção de que não existem diferenças na probabilidade de hemorragia significativa entre os 2 grupos.

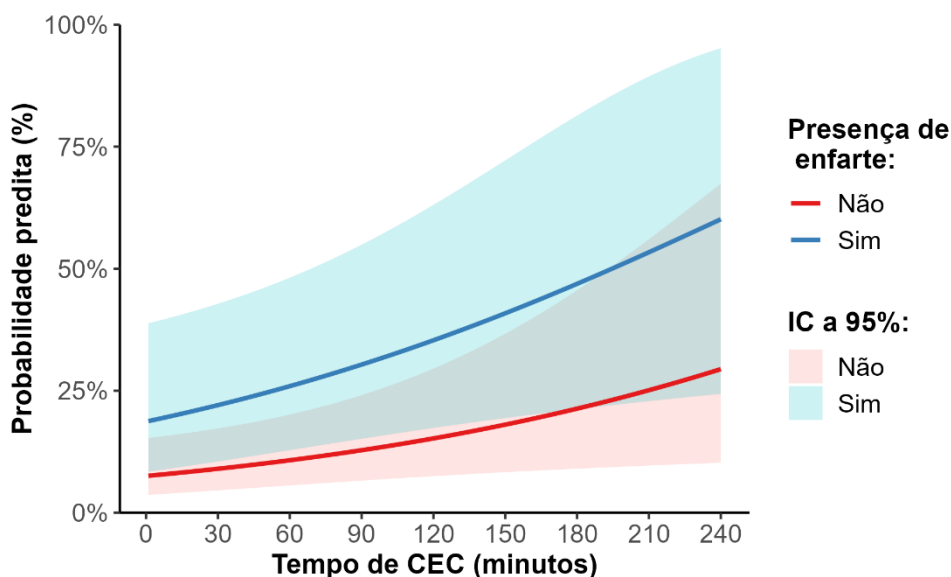
Figura 56 - Probabilidade predita de hemorragia significativa por tipo de *priming* e por tempo de CEC



Procedeu-se também à estimação das probabilidades preditas para o desfecho relacionado com a variável dependente deste modelo (hemorragia significativa) mas, para o enfarte. As variáveis com valores fixados incluíram o sexo masculino, a ausência de hipertensão arterial, IRC e de cirurgia da aorta e sem gelatina no *priming*. Para além disso, fixou-se a idade, o hematócrito, o aPTT e a taxa de diluição nos seus valores medianos dentro da amostra de dados de treino da base de *dados parcialmente balanceados* (isto é, 71 anos, 40.2%, 29 segundos e 14.6%, respetivamente). Assim, fizeram-se variar apenas as variáveis enfarte e tempo de CEC. O enfarte variou entre as duas categorias de resposta da variável (com vs. sem enfarte nos antecedentes clínicos) e para o tempo de CEC, foi criada uma sequência de valores entre 1 e 240 minutos com variação de uma unidade. Após a criação deste subconjunto de dados foi aplicada a mesma metodologia anteriormente descrita.

Obteve-se, assim, o gráfico de probabilidades preditas presente na Figura 57. Verificando, uma vez mais, que aos 60 minutos de CEC, um doente com as restantes características descritas no parágrafo anterior e que não tenha antecedentes de enfarte agudo do miocárdio tem uma probabilidade de hemorragia significativa às 12 horas de pós-procedimento de 10.8% (IC 95%: 5.6 - 20.1%), enquanto um doente com esse antecedente tem 25.9% (IC 95%: 12.8 - 48.1%).

Figura 57 - Probabilidade predita de hemorragia significativa por enfarte e por tempo de CEC



Procedeu-se posteriormente à estimação das probabilidades preditas com as variáveis explicativas sexo e cirurgia da aorta. As variáveis com valores fixados incluíram a ausência de enfarte, de hipertensão arterial, de insuficiência renal crónica e de gelatina no *priming*. Para além disso, fixou-se a idade, o hematócrito, o aPTT e a taxa de diluição nos seus valores medianos dentro da amostra de dados de treino. Assim, fizeram-se variar apenas as variáveis relativas à cirurgia da aorta, sexo e tempo de CEC. A duas primeiras, variam entre as duas categorias de resposta de cada uma (com vs. sem cirurgia da aorta e feminino vs. masculino, respetivamente) e para a última, o tempo de CEC, foi também utilizada a sequência de valores entre 1 e 240 minutos.

Obteve-se, assim, o gráfico de probabilidades preditas presente na Figura 58. Podem ser apuradas quatro combinações distintas: (i) doente do sexo masculino com cirurgia da aorta; (ii) doente do sexo masculino sem cirurgia da aorta; (iii) doente do sexo feminino com cirurgia da aorta; e (iv) doente do sexo feminino sem cirurgia da aorta. As probabilidades preditas e os respetivos intervalos de confiança para cada uma destas possíveis combinações notadas na Figura 58, encontram-se esquematizadas na Tabela 16. Esta verificação e cálculo foram realizados, verificando as probabilidades para os 60 minutos de CEC.

Figura 58 - Probabilidade predita de hemorragia significativa por sexo e cirurgia da aorta por tempo de CEC

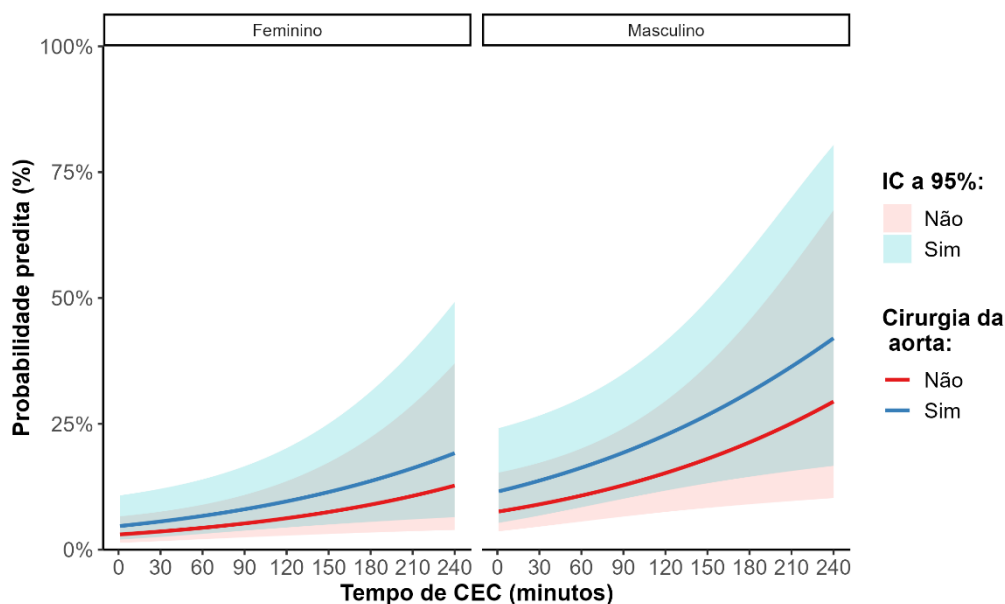


Tabela 16 - Probabilidades preditas por sexo e cirurgia da aorta por tempo de CEC

		Cirurgia da Aorta	
		Não	Sim
Sexo	Feminino	4.3 (2.1 - 8.9%)	6.7 (3.2 - 14%)
	Masculino	10.8 (5.6 - 20.1%)	16.3 (8.4 - 30.1%)

5.3.5 Avaliação dos modelos

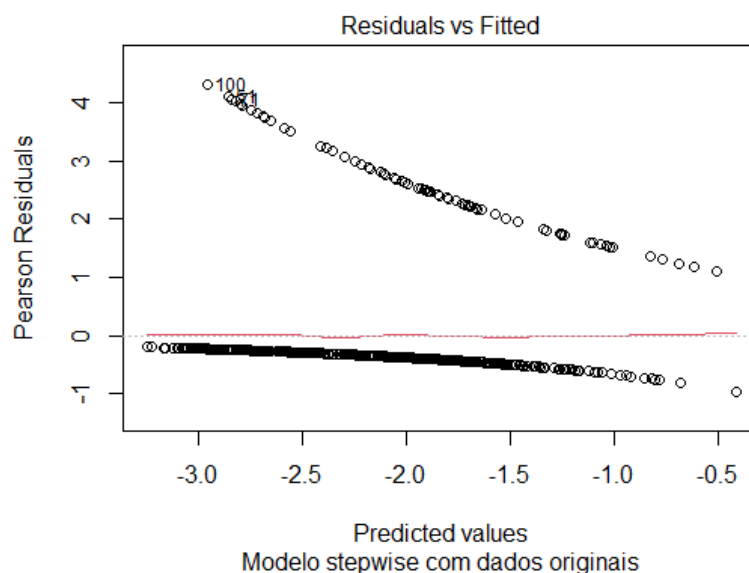
Foram verificados os pressupostos para a realização da Regressão Binária: (i) a variável dependente utilizada é binária o que significa que tem duas categorias mutuamente exclusivas (hemorragia significativa vs. hemorragia não significativa); (ii) existe independência das observações, na medida em que não foram incluídas medidas repetidas (por exemplo, os parâmetros analíticos referem-se apenas ao momento pré-operatório) e (iii) verificou-se a ausência de multicolinearidade das variáveis independentes utilizadas no modelo, com recurso ao valor de FIV que se encontra acima, na Tabela 12.

Iniciando a avaliação dos modelos com a avaliação dos resíduos optou-se por apresentar aqui apenas os gráficos referentes a esta avaliação, mas aplicada ao modelo obtido com a *base de dados originais* e após a seleção automática das variáveis, ou seja, o modelo representado na Tabela 12 (*modelo stepwise*). Apesar disso, nesta secção será feita uma

comparação com os outros dois modelos (modelos com a base de dados parcialmente e completamente balanceada) cujos gráficos se encontram na secção Apêndices.

Em primeiro lugar, foi elaborado o gráfico representado na Figura 59 que esquematiza os resíduos em função dos valores preditos pelo modelo, ou seja, o preditor linear. O modelo aqui usado foi o modelo de regressão binária e criado após a seleção automática das variáveis. A linha a vermelho indica que não parece existir diferenças nas magnitudes provocadas pelos resíduos em comparação ao zero. Não existem, portanto, grandes desvios dos pontos do gráfico face ao eixo horizontal do zero. Apesar disso, verifica-se que existem duas observações classificadas como *outliers*, as observações 100 e 71, cujos resíduos são elevados e ligeiramente superiores a 4. Comparando esta análise com os mesmos gráficos, mas para os outros dois modelos estimados, verificou-se que para o modelo obtido com a *base de dados parcialmente balanceados* (Figura 63) as constatações são as mesmas. Contudo, para o modelo obtido com a *base de dados completamente balanceados* (Figura 67) verificou-se a situação inversa. As observações que foram classificadas como *outliers* e com maiores valores de resíduos, pertencem agora ao grupo de observações associadas aos pontos negativos do gráfico que representam as observações 502, 623 e 40.

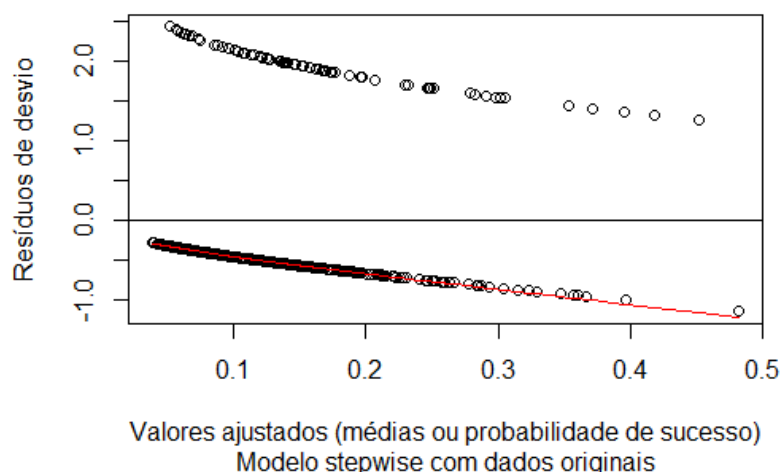
Figura 59 - Gráfico de Resíduos do Modelo com dados originais



O gráfico da Figura 60 apresenta os valores de resíduos em função dos valores ajustados. Observamos dois grupos de pontos, um no eixo superior (pontos positivos), associados à presença de hemorragia significativa, e um no eixo inferior (pontos negativos), estes associados à presença de hemorragia não-significativa. Observa-se também que a maior

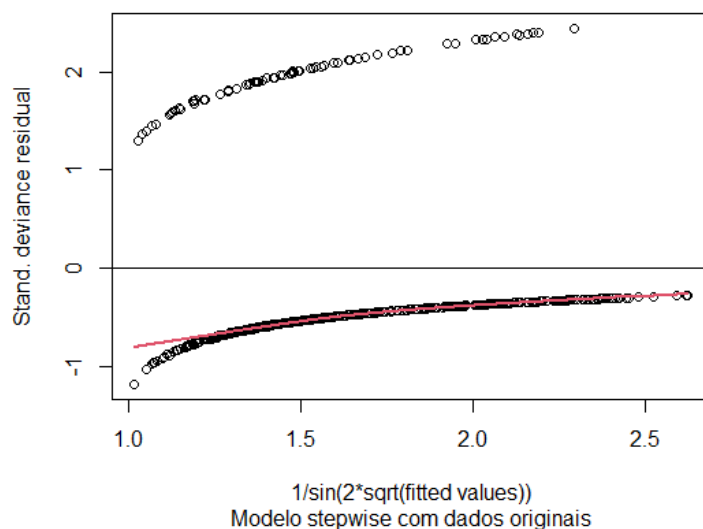
densidade de pontos positivos se concentra mais à esquerda do gráfico, ou seja, para valores ajustados menores, o que também se observa nos pontos inferiores. Esta observação não seria desejável, mas poder-se-á assumir como expectável uma vez que existe uma elevada discrepância entre o número de doentes nos dois grupos. A linha a vermelho, mostra isso mesmo, que existe uma tendência para o modelo estimado atribuir a previsão de o doente não ter uma hemorragia significativa. Para além disso, também se verifica que os resíduos positivos do modelo são superiores aos negativos. Uma vez mais, as descrições anteriormente realizadas aplicam-se igualmente ao gráfico da Figura 67, que representa a mesma situação, mas para a *base de dados parcialmente balanceados*. Contudo, na Figura 68, que representa a mesma situação, mas para o modelo com a *base de dados completamente balanceados*, verifica-se que desaparece a tendência para o modelo estimado atribuir a previsão de o doente não ter uma hemorragia significativa. Para além disso, os pontos apresentam-se melhor distribuídos no eixo horizontal.

Figura 60 - Gráfico de resíduos por grupo



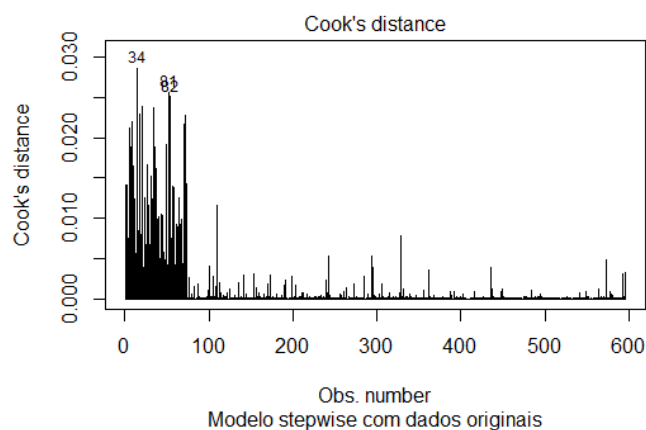
O gráfico seguinte (Figura 61) representa os desvios residuais contra uma transformação adequada para o modelo binomial, neste caso, $2\sin^{-1}\sqrt{\hat{\mu}}$. A linha a vermelho representa a tendência dos pontos desenhados no gráfico dos resíduos de desvios padronizados contra a transformação dos valores citada anteriormente. Como se pode verificar a linha vermelha de tendência está muito próxima de zero, ou seja, do eixo horizontal, à exceção dos valores mais à esquerda do eixo dos xx . A mesma constatação verifica-se para o modelo da *base de dados parcialmente balanceados* que consta na Figura 65. Um padrão diferente é verificado na Figura 69 e que diz respeito ao modelo que utilizou a *base de dados completamente balanceados*.

Figura 61 - Gráfico de resíduos padronizados do modelo com a base de dados original



Por forma a verificar se existem observações efetivamente influentes no modelo ajustado foi construído o gráfico da Figura 62, que representa as distâncias de *Cook*, para cada uma das 744 observações. Esta informação (distância de *Cook*) representa a medida utilizada para avaliar a influência efetiva das observações, sendo que quanto maior for o seu valor, mais influente é essa mesma observação e vice-versa. Verifica-se, nessa mesma figura, que não existem observações com distâncias de *Cook* superiores a 0.5. Por não existirem distâncias de *Cook* aumentadas, pode-se concluir que não existem observações efetivamente influentes, o que significa que a sua inclusão no modelo não terá produzido, à partida, alterações significativas aos estimadores e parâmetros do mesmo. A mesma conclusão verifica-se para os outros dois modelos, apesar de as observações mais influentes não serem as mesmas (Figura 66 e Figura 70).

Figura 62 - Avaliação de observações efetivamente influentes do modelo com dados originais



Foi efetuado o teste de *Hosmer & Lemeshow* para avaliar a adequabilidade dos modelos. Os valores ajustados do modelo de regressão foram divididos em oito grupos usando o método de divisão em percentis. Para cada grupo, foram calculados os valores esperados e observados da variável resposta. Os valores esperados foram obtidos calculando a média dos valores ajustados dentro de cada grupo, enquanto os valores observados foram obtidos calculando a média das respostas observadas (ajustadas para a escala binária 0 e 1) dentro de cada grupo. Foi calculada uma estatística de qui-quadrado com base na comparação entre os valores esperados e observados. Os resultados dos testes de *Hosmer & Lemeshow* foram os seguintes: (i) *modelo com os dados originais* - estatística de teste igual a 0.229, com 6 graus de liberdade pelo que se obteve um valor-p igual a 0.999; (ii) *modelo com os dados parcialmente balanceados* - estatística de teste igual a 0.775, com 6 graus de liberdade pelo que se obteve um valor-p igual a 0.993; e (iii) *modelo com os dados completamente balanceados* - estatística de teste igual a 0.027, com 6 graus de liberdade pelo que se obteve um valor-p igual a 0.999; Com base na significância estatística desse resultado pode-se inferir que existem evidências estatísticas suficientes para rejeitar a hipótese nula de que não há discrepância significativa entre os valores observados e esperados, o que indica uma boa adequação do modelo aos dados.

Utilizaram-se os modelos de regressão ajustados para prever as probabilidades de ocorrência de hemorragia significativa às 12 horas de pós-operatório para cada observação no respetivo conjunto de dados de teste, correspondentes aos 20% da amostra inicial conforme mencionado anteriormente. Com base nas probabilidades previstas, classificaram-se as observações como "*não significativa*" ou "*significativa*" usando um limiar de probabilidade igual a 0,5. Se a probabilidade prevista foi maior que 0,5, a observação foi classificada como "*significativa*", caso contrário, foi classificada como "*não significativa*". Como tal, foi construída a matriz de confusão para avaliar o desempenho de cada um dos modelos. A matriz de confusão (Tabela 17) compara as classificações previstas com as classificações reais e efetivamente comprovadas nos registos clínicos, e que constam no conjunto de teste. A seguir, foi calculada a taxa de acerto como correspondendo à proporção de observações classificadas corretamente em relação ao total de observações no conjunto de teste. Esta análise foi realizada para cada um dos modelos utilizados neste trabalho, sendo que se encontram todos resumidos na Tabela 17.

Tabela 17 - Validação cruzada com dados de teste de cada um dos modelos

		Classificações do modelo com dados originais		
		Não significativa	Significativa	
Classificações dados teste	Não significativa	130	0	
	Significativa	18	0	
			Classificações do modelo com dados parcialmente balanceados	
	Não significativa	128	2	
	Significativa	31	2	
			Classificações do modelo com dados completamente balanceados	
Não significativa	90	40		
Significativa	52	80		

Com base nos resultados apresentados, pode-se concluir que o modelo foi capaz de classificar corretamente (taxa de acerto): (i) *modelo com os dados originais* - 87.84%; (ii) *modelo com os dados parcialmente balanceados* - 79.75%; e (iii) *modelo com os dados completamente balanceados* - 64.85%. A matriz de confusão mostra que o modelo teve sucesso principalmente na classificação de observações com hemorragia não significativa. Enquanto isso, teve um mau desempenho na classificação de observações como hemorragia *significativa* no caso do modelo que usou a base de dados original, uma vez que não classificou corretamente nenhum dos 18 casos de hemorragia significativa presentes nos dados de teste.

DISCUSSÃO

Após a apresentação dos resultados, importa resumir que foram estimados 3 modelos de regressão binária que partiram de 3 bases de dados diferentes. A discussão cabe agora, à comparação dos resultados obtidos com os 3 modelos, isto é, o modelo estimado com a *base de dados original*, o modelo estimado com a *base de dados parcialmente balanceados* e com a *base de dados completamente balanceados*. Os modelos aqui discutidos dizem apenas respeito aos 3 modelos obtidos com a função *stepwise*. Resumidamente: (i) o modelo estimado a partir da *base de dados original* obteve um AIC igual a 435,5, uma AUC igual a 66,1% e uma taxa de acerto igual a 87,84%; (ii) o modelo estimado a partir da *base de dados parcialmente balanceados* obteve um AIC igual a 622,69, uma AUC igual a 71,8% e uma taxa de acerto igual a 79,75%; e, por fim, (iii) o modelo estimado a partir da base de dados completamente balanceada obteve um AIC de 1342,1, uma AUC igual a 71,8% e uma taxa de acerto igual a 64,85%.

Os dados não balanceados representam um grande desafio na análise estatística e no tratamento dos dados. De acordo com Zhang et al., a classe minoritária (ou seja, o evento) costuma ser sempre a classe de interesse e a mais angustiante caso seja classificada erradamente por um qualquer modelo estimado. O desafio é que a maioria dos métodos estatísticos têm viés em relação à classe majoritária e não podem prever com precisão a classe minoritária, incluindo o modelo de regressão, que é favorecido pela sua elevada interpretabilidade [74]. Tal foi sugerido por Yang et al., no seu estudo de impacto e performance de bases de dados balanceados por métodos aleatórios de *oversampling* com dados de saúde. Estes autores verificaram que um modelo de previsão desenvolvido usando dados desequilibrados pode se tornar tendencioso para a classe majoritária e que pode assim, ter maior probabilidade de classificar incorretamente a classe minoritária [75], [76]. A análise de regressão é a abordagem de modelação mais frequentemente usada para analisar variáveis de resposta binárias ao relacioná-la estatisticamente com variáveis preditoras [77]. Tal também

se verificou para a realização deste trabalho. Foram encontradas complexidades na modelação dos dados presentes na amostra estudada. Os dados são reais, e como tal, por conveniência clínica não são expectáveis dados balanceados no que concerne a uma variável de desfecho relacionada com a hemorragia pós-operatória significativa. Aliás, todos os esforços clínicos prestados para com os doentes vão exatamente no sentido contrário - restringir ao máximo a ocorrência deste desfecho. Portanto, o objetivo aqui proposto foi deveras desafiante.

O *critério de Informação de Akaike* é uma medida comumente aplicada na avaliação da qualidade do ajuste do modelo, tendo em consideração a sua complexidade [48], [78], [79]. Note-se que o modelo estimado a partir da *base de dados original* apresentou o menor valor de AIC (435.5), parecendo indicar um melhor ajuste, em relação aos outros modelos. Contudo, tal constatação não deverá ser tida, uma vez que o AIC está intrinsecamente relacionado com a dimensão da amostra [50]. As três bases de dados utilizadas, a original e as duas derivadas desta por *oversampling* a 20 e 50% não têm a mesma dimensão. As dimensões da amostra são de 596, 658 e 1050, respetivamente. Isto referente à dimensão da base de dados de treino de cada um dos modelos [78], [79], [80]. Alguns autores, tais como Hurvich & Tsai, sugerem a uniformização e a aplicação do AIC corrigido (AICc) para a dimensão da amostra por forma a lidar com esta questão [49]. Aplicando este método aos resultados deste trabalho verifica-se que os valores de AIC brutos eram iguais a 435.5, 622.7 e 1342.1. Aplicando a correção à dimensão da amostra obtiveram-se os seguintes valores, pela mesma ordem anteriormente enumerada, 141 079.5, 77 883.8 e 156 562.4. Verifica-se que o valor de AICc mais baixo foi, com uma diferença substancial, encontrado no modelo com a *base de dados parcialmente balanceados*. Também é conhecido que o tamanho da base de dados afeta o desempenho do modelo. Para amostras de maior dimensão são esperados tamanhos de efeitos mais pequenos e intervalos de confiança mais estreitos [79]. Este resultado sugere que, embora o modelo com a *base de dados completamente balanceados* tenha sido ajustado para lidar com o desequilíbrio da *base de dados original*, este pode ter introduzido uma complexidade adicional que não melhorou significativamente o ajuste do modelo.

A taxa de acerto é uma medida de precisão do modelo nomeadamente na correta classificação das observações. Surpreendentemente, o modelo original obteve a taxa de acerto mais alta (87.84%), seguido do modelo com a *base de dados parcialmente balanceados* (79,75%) e, por fim, do modelo com a *base de dados completamente balanceados* (64,85%). Contudo, é necessário ter em atenção que na classificação obtida da validação cruzada para o modelo com os dados originais, o mesmo não foi capaz de classificar como significativas as 18 observações (Tabela 17) que na verdade correspondem a hemorragias de grau significativo.

Aqui, poder-se-á deduzir que o facto de a amostra estar completamente desbalanceada gerou um modelo que se poderá assumir como tendo constituído *overfitting* à classe maioritária, isto é, à classe correspondente à hemorragia não significativa. De salientar, que estes dados obtidos pela validação cruzada mostram que este modelo obteve uma sensibilidade de 0% e uma especificidade de 100%. O modelo que usou a *base de dados originais* mostrou ser adequado, apesar de não ter alcançado a capacidade de identificar os verdadeiros positivos, ou seja, os doentes com hemorragia significativa provada. Para além disso, este modelo teve uma capacidade de identificar todos os verdadeiros negativos, ou seja, de identificar corretamente todos os indivíduos que não tiveram hemorragia significativa provada. Esta conclusão não é a mais favorável uma vez que na prática clínica teria mais proveito um modelo que tivesse um grau de sensibilidade do teste aceitável, por forma a ocuparem-se, os profissionais de saúde, a detetar precocemente estas situações. Também o segundo modelo, o modelo gerado a partir da *base de dados parcialmente balanceados*, mostrou uma sensibilidade muito reduzida, mas diferente de zero (6.1%), sendo que continuou a apresentar uma especificidade alta (98.5%). Para o último modelo, obtido através da *base de dados completamente balanceados*, houve um aumento significativo da sensibilidade (60,6%) mas à custa de uma redução da especificidade (69,2%). Contudo, deve ser recapitulado que este modelo não é adequado em termos do teste da sua *deviance*.

Os resultados podem sugerir que, os modelos estimados a partir de bases de dados balanceados artificialmente proporcionaram melhorias significativas na capacidade de discriminação ou na precisão da classificação, em comparação com o modelo obtido com os dados originais. Isso pode indicar que a estratégia de balanceamento utilizada, ao introduzir mais observações correspondentes ao desfecho de hemorragia significativa, melhorou a capacidade de o modelo diagnosticar corretamente essa mesma variável. Através da Tabela 17 é possível constatar que para o primeiro modelo existiam apenas 18 observações relacionadas com hemorragia significativa de entre 148 observações. Para o segundo modelo existia uma proporção de 33 para 163 e, no último modelo, de 132 para 262 observações, o que está alinhado com as frequências verificadas para cada uma destas 3 bases de dados e que constam nas Figura 5, Figura 49 e Figura 50, respetivamente.

Os resultados deste trabalho sugerem ainda que a inclusão de mais dados por intermédio de uma técnica de balanceamento, tal como o *oversampling*, pode influenciar a seleção de variáveis e a precisão do modelo de regressão binária, ressaltando a importância do impacto causado pelo desbalanceamento de dados ao realizar análises estatísticas. Contudo, não são poucas as vezes em que os dados em saúde, em qualquer uma das áreas, são natural e clinicamente desbalanceados [81], [82], [83], [84].

A significância estatística do intercepto num modelo de regressão indica que a média da variável dependente (ou a *log-odds* da variável dependente, no caso da regressão) não é zero quando todas as outras variáveis independentes no modelo são zero [85]. Por outras palavras, o intercepto representa o valor esperado da variável dependente quando todas as variáveis independentes do modelo são zero [39], [43]. Neste trabalho verificou-se que a *base de dados completamente balanceados* gerou um modelo em que o intercepto era significativo e, para o qual, o teste de adequabilidade mostrou que o modelo não era adequado. Isto poderá querer significar que para esse modelo existe uma tendência ou probabilidade significativa para a ocorrência do desfecho hemorragia significativa que foi capturada pelo intercepto e que, não foi bem explicada, ou corretamente capturada, por nenhuma das outras variáveis. Também pode estar em causa o facto de as variáveis incluídas nos modelos possam não ser as melhores, ou possam não ser suficientes, para classificar ou prever corretamente a hemorragia significativa. De um modo geral, poderá faltar uma variável que vá ao encontro da possibilidade de ocorrência de algum grau de iatrogenia associada ao procedimento cirúrgico. Nas diversas *guidelines* de gestão de produtos sanguíneos e hemoderivados e de controlo da hemorragia perioperatória, para além de todas as variáveis laboratoriais e de antecedentes clínicos associadas aos doentes, é várias vezes referida a iatrogenia do procedimento [86], [87].

CONCLUSÕES

O presente trabalho explorou a construção e avaliação de modelos de regressão binária multivariada para prever a ocorrência de hemorragia significativa após cirurgia cardíaca com CEC, em adultos. Este teve por objetivo a investigação da associação do *priming* com gelatina com a hemorragia. Para tal, foram utilizadas três bases de dados distintas: uma original e duas derivadas artificial e aleatoriamente, sendo uma parcialmente balanceada e outra completamente balanceada. A técnica de balanceamento utilizada foi o *oversampling*.

Foram verificados os pressupostos necessários para a realização da regressão binária e por fim, foram aplicadas diversas métricas e métodos de análise da sua adequabilidade. Foram utilizadas métricas, tais como, AIC, AUC e taxa de acerto. Os resultados revelaram que o modelo estimado a partir da base de dados original obteve o menor valor de AIC, sugerindo um ajuste melhor em relação aos outros modelos, embora o modelo com a *base de dados parcialmente balanceados* obteve o menor valor de AIC corrigido para a dimensão da amostra (AICc). No entanto, a análise da taxa de acerto revelou que o modelo com os dados originais, e, portanto, completamente desbalanceados, teve dificuldade em classificar corretamente os casos de hemorragia significativa, apresentando uma sensibilidade de 0%. Já os modelos derivados das bases de dados balanceadas apresentaram sensibilidade maior, mas à custa de uma redução na especificidade.

A discussão dos resultados salientou o impacto dos dados não balanceados na análise estatística, assim como a influência da técnica de balanceamento na seleção de variáveis e na precisão do modelo. Foi destacado o desafio representado pelo desbalanceamento de dados em contextos clínicos reais, onde é comum encontrar frequências desproporcionais entre as duas classes de dados binários.

Finalmente, com base nos resultados e discussões apresentadas, é possível concluir que a inclusão de mais dados por meio do balanceamento pode influenciar positivamente a capacidade de predição do modelo, mas também pode introduzir complexidades adicionais. Além disso, a seleção adequada de variáveis e a compreensão dos resultados são aspectos essenciais na construção de modelos de regressão binária para prever hemorragia significativa após cirurgia.

Em suma, os resultados mostraram que não houve evidência estatística de associação significativa entre o tipo de *priming* (com vs. sem gelatina) e a gravidade da hemorragia às 12 horas de pós-operatório. Embora tenha sido observada uma diferença estatisticamente significativa na hemorragia às 24 horas de pós-operatório entre os grupos, essa diferença não foi considerada clinicamente relevante, pois as estatísticas descritivas indicam que as diferenças são mínimas. Não houve evidência estatisticamente significativa de associação entre o tipo de *priming* (com gelatina vs. sem gelatina) e a necessidade de transfusão sanguínea no pós-operatório. Portanto, com base nos resultados apresentados, parece que a escolha do tipo de *priming* (com ou sem gelatina) não teve um impacto significativo nas variáveis clínicas e desfechos avaliados neste trabalho. Em termos de probabilidades preditas, não houve evidência, em nenhum dos modelos, de que a gelatina no *priming* possa estar associada a um aumento da probabilidade de ocorrência de hemorragia. No entanto, é importante considerar que outras variáveis não incluídas nesta análise também podem influenciar esses resultados, e mais pesquisas podem ser necessárias para uma avaliação mais abrangente.

Os dados reais constituem um desafio na aplicação de modelos matemáticos, apesar destes constituírem uma parte fundamental da análise de dados em busca de avanços na melhoria dos resultados em saúde, nomeadamente na melhoria da morbilidade e da mortalidade. A Perfusão Cardiovascular é uma área em permanente evolução, e o desenvolvimento de novas abordagens com este tipo de modelos aplicados à prática, constitui uma área vital para os desafios da medicina cardiovascular no futuro.

BIBLIOGRAFIA

- [1] «Cardiopulmonary Bypass - MeSH - NCBI». Acedido: 23 de Março de 2024. [Em linha]. Disponível em: <https://www.ncbi.nlm.nih.gov/mesh/?term=bypass%2C+cardiopulmonary>
- [2] «Polygeline - MeSH - NCBI». Acedido: 23 de Março de 2024. [Em linha]. Disponível em: <https://www.ncbi.nlm.nih.gov/mesh/?term=gelofusine>
- [3] «Heparin - MeSH - NCBI». Acedido: 23 de Março de 2024. [Em linha]. Disponível em: <https://www.ncbi.nlm.nih.gov/mesh/68006493>
- [4] «Crystalloid Solutions - MeSH - NCBI». Acedido: 23 de Março de 2024. [Em linha]. Disponível em: <https://www.ncbi.nlm.nih.gov/mesh/2028000>
- [5] K. Yang *et al.*, «Modified cardiopulmonary bypass with low priming volume for blood conservation in cardiac valve replacement surgery», *J Cardiothorac Surg*, vol. 18, n. 1, p. 56, Fev. 2023, doi: 10.1186/s13019-023-02175-8.
- [6] «Blood Volume - MeSH - NCBI». Acedido: 28 de Março de 2024. [Em linha]. Disponível em: <https://www.ncbi.nlm.nih.gov/mesh/68001810>
- [7] G. A. Roth *et al.*, «Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study», *J Am Coll Cardiol*, vol. 76, n. 25, p. 2982, Dez. 2020, doi: 10.1016/J.JACC.2020.11.010.
- [8] M. H. L. Souza e D. O. Elias, *Fundamentos da Circulação Extracorpórea*, Segunda edição. Rio de Janeiro: Centro Editorial Alfa Rio, 2006.
- [9] R. M. Bojar, *Manual of Perioperative Care in Adult Cardiac Surgery*, Sixth Edition. Worcester, Massachusetts, USA: Wiley, 2021.
- [10] G. P. Gravlee, R. F. Davis, J. W. Hammon, e B. D. Kussman, «Cardiopulmonary bypass and mechanical support: Principles and practice: Fourth edition», *Cardiopulmonary Bypass and Mechanical Support: Principles and Practice: Fourth Edition*, pp. 1–912, 2015.
- [11] L. H. Cohn e D. H. Adams, *Cardiac Surgery in the Adult*, Fifth Edition. New York: Mc Graw Hill Education, 2018.

- [12] S. Ghosh, F. Falter, e D. J. Cook, Eds., *Cardiopulmonary Bypass*, First edition. United Kingdom: Cambridge University Press, 2009.
- [13] L. Puis *et al.*, «2019 EACTS/EACTA/EBCP guidelines on cardiopulmonary bypass in adult cardiac surgery», *Interact Cardiovasc Thorac Surg*, vol. 30, n. 2, pp. 161–202, 2019, doi: 10.1093/icvts/ivz251.
- [14] G. J. Myers, Ed., *Predictive Hemodilution - The Principles and Practice of Hemodilution during Cardiac Surgery*, First edition. Italy: LivaNova, 2017.
- [15] J. T. Magruder *et al.*, «Nadir oxygen delivery on bypass and hypotension increase acute kidney injury risk after cardiac operations», *Annals of Thoracic Surgery*, vol. 100, n. 5, pp. 1697–1703, Nov. 2015, doi: 10.1016/j.athoracsur.2015.05.059.
- [16] S. Kozek-Langenecher, *Colloidal Volume Replacement Therapy*, First edition. London: International Medical Publishers, 2011.
- [17] A. A. SCHRAMKO, A. H. KUITUNEN, R. T. SUOJARANTA-YLINEN, e T. T. NIEMI, «Role of fibrinogen-, factor VIII- and XIII-mediated clot propagation in gelatin haemodilution», *Acta Anaesthesiol Scand*, vol. 53, n. 6, pp. 731–735, Jul. 2009, doi: 10.1111/j.1399-6576.2008.01901.x.
- [18] C. D. Okello *et al.*, «Mortality and its associated factors in transfused patients at a tertiary hospital in Uganda», *PLoS One*, vol. 17, n. 9, p. e0275126, Set. 2022, doi: 10.1371/journal.pone.0275126.
- [19] C. G. Koch *et al.*, «Morbidity and mortality risk associated with red blood cell and blood-component transfusion in isolated coronary artery bypass grafting», *Crit Care Med*, vol. 34, n. 6, pp. 1608–1616, 2006, doi: 10.1097/01.CCM.0000217920.48559.D8.
- [20] K. Karkouti *et al.*, «The independent association of massive blood loss with mortality in cardiac surgery», *Transfusion (Paris)*, vol. 44, n. 10, pp. 1453–1462, Out. 2004, doi: 10.1111/j.1537-2995.2004.04144.x.
- [21] S. St-Onge *et al.*, «Active clearance vs conventional management of chest tubes after cardiac surgery: a randomized controlled study», *J Cardiothorac Surg*, vol. 16, n. 1, p. 44, Dez. 2021, doi: 10.1186/s13019-021-01414-0.
- [22] M. C. Christensen, F. Dziewior, A. Kempel, e C. von Heymann, «Increased Chest Tube Drainage Is Independently Associated With Adverse Outcome After Cardiac Surgery», *J Cardiothorac Vasc Anesth*, vol. 26, n. 1, pp. 46–51, Fev. 2012, doi: 10.1053/j.jvca.2011.09.021.
- [23] V. Rancati *et al.*, «Challenges in Patient Blood Management for Cardiac Surgery: A Narrative Review», *J Clin Med*, vol. 10, n. 11, p. 2454, Jun. 2021, doi: 10.3390/jcm10112454.

- [24] M. C. Christensen, S. Krapf, A. Kempel, e C. von Heymann, «Costs of excessive postoperative hemorrhage in cardiac surgery», *J Thorac Cardiovasc Surg*, vol. 138, n. 3, pp. 687–693, Set. 2009, doi: 10.1016/j.jtcvs.2009.02.021.
- [25] V. Protsyk, B. S. Rasmussen, F. Guarracino, J. Erb, E. Turton, e J. Ender, «Fluid Management in Cardiac Surgery: Results of a Survey in European Cardiac Anesthesia Departments», *J Cardiothorac Vasc Anesth*, vol. 31, n. 5, pp. 1624–1629, Out. 2017, doi: 10.1053/J.JVCA.2017.04.017.
- [26] D. N. Lobo *et al.*, «Effect of volume loading with 1 liter intravenous infusions of 0.9% saline, 4% succinylated gelatine (Gelofusine) and 6% hydroxyethyl starch (Voluven) on blood volume and endocrine responses: A randomized, three-way crossover study in healthy volunteers», *Crit Care Med*, vol. 38, n. 2, pp. 464–470, Fev. 2010, doi: 10.1097/CCM.0b013e3181bc80f1.
- [27] A. A. Schramko, R. T. Suojaranta-Ylinen, A. H. Kuitunen, P. M. Raivio, S. I. Kukkonen, e T. T. Niemi, «Comparison of the effect of 6% hydroxyethyl starch and gelatine on cardiac and stroke volume index: a randomized, controlled trial after cardiac surgery», <http://dx.doi.org/10.1177/0267659110377819>, vol. 25, n. 5, pp. 283–291, Jul. 2010, doi: 10.1177/0267659110377819.
- [28] R. G. Hahn, «Understanding volume kinetics», *Acta Anaesthesiol Scand*, vol. 64, n. 5, pp. 570–578, Mai. 2020, doi: 10.1111/aas.13533.
- [29] Ø. S. Svendsen, M. Farstad, A. Mongstad, R. Haaverstad, P. Husby, e V. L. Kvalheim, «Is the use of hydroxyethyl starch as priming solution during cardiac surgery advisable? A randomized, single-center trial», *Perfusion*, vol. 33, n. 6, pp. 483–489, Set. 2018, doi: 10.1177/0267659117746235.
- [30] A. Schramko, R. Suojaranta-Ylinen, A. Kuitunen, P. Raivio, S. Kukkonen, e T. Niemi, «Hydroxyethylstarch and gelatin solutions impair blood coagulation after cardiac surgery: A prospective randomized trial», *Br J Anaesth*, vol. 104, n. 6, pp. 691–697, 2010, doi: 10.1093/BJA/AEQ084.
- [31] A. H. Kuitunen, M. J. Hynynen, E. Vahtera, e M. T. Salmenperä, «Hydroxyethyl starch as a priming solution for cardiopulmonary bypass impairs hemostasis after cardiac surgery», *Anesth Analg*, vol. 98, n. 2, pp. 291–297, 2004, doi: 10.1213/01.ANE.0000096006.60716.F6.
- [32] S. L. Jin e B. W. Yu, «Effects of acute hypervolemic fluid infusion of hydroxyethyl starch and gelatin on hemostasis and possible mechanisms», *Clin Appl Thromb Hemost*, vol. 16, n. 1, pp. 91–98, Fev. 2010, doi: 10.1177/1076029608330474.
- [33] S. A. Kozek-Langenecker, «Fluids and coagulation», *Curr Opin Crit Care*, vol. 21, n. 4, pp. 285–291, Jul. 2015, doi: 10.1097/MCC.0000000000000219.

- [34] S. Wassertheil-Smoller e J. Smoller, *Biostatistics and Epidemiology, A Primer for Health and Biomedical Professionals*, Fourth Edition. New York: Springer, 2015.
- [35] F. M. Mello e R. C. Guimarães, *Métodos Estatísticos para o Ensino e a Investigação nas Ciências da Saúde*, 1ª edição. Lisboa: Edições Sílabo, Lda., 2015.
- [36] W. J. Conover, *Practical Nonparametric Statistics*, Third Edition. New York: John Wiley & Sons, 1999.
- [37] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Second Edition. New York: CRC Press, 1997.
- [38] W. J. Conover, *Practical Nonparametric Statistics*, First edition. New York: Wiley, 1999.
- [39] J. M. Hilbe, *Logistic Regression Models*, First edition. Florida: Chapman & Hall/CRC press, 2009.
- [40] D. W. Hosmer, S. Lemeshow, e R. X. Sturdivant, *Applied Logistic Regression*. Wiley, 2013. doi: 10.1002/9781118548387.
- [41] E. Núñez, E. W. Steyerberg, e J. Núñez, «[Regression modeling strategies]», *Rev Esp Cardiol*, vol. 64, n. 6, pp. 501–507, Jun. 2011, doi: 10.1016/J.RECESP.2011.01.019.
- [42] J. A. N. P. McCullagh, *Generalized linear models*, Second Edition., vol. 16. New York: Chapman & Hall, 1989. doi: 10.1016/0377-2217(84)90282-0.
- [43] A. J. Dobson e A. G. Barnett, *An Introduction to Generalized Linear Models*, Third Edition. Florida: Chapman & Hall/CRC press, 2008.
- [44] D. W. Hosmer, S. Lemeshow, e R. X. Sturdivant, *Applied Logistic Regression*, Third Edition. New Jersey: Wiley, 2013.
- [45] J. L. Peacock e P. J. Peacock, *Oxford Handbook of Medical Statistics*, First edition. New York: Oxford University Press, 2011.
- [46] J. S. B. Alves, J. L. Bazán, e R. B. Arellano-Valle, «Flexible cloglog links for binomial regression models as an alternative for imbalanced medical data», *Biometrical Journal*, vol. 65, n. 3, p. 2100325, Mar. 2023, doi: 10.1002/BIMJ.202100325.
- [47] «A Gentle Introduction to Complementary Log-Log Regression | by Akif Mustafa | Towards Data Science». Acedido: 24 de Março de 2024. [Em linha]. Disponível em: <https://towardsdatascience.com/a-gentle-introduction-to-complementary-log-log-regression-8ac3c5c1cd83>
- [48] A. Chakrabarti e J. K. Ghosh, «AIC, BIC and Recent Advances in Model Selection», *Philosophy of Statistics: Volume 7 in Handbook of the Philosophy of Science*, vol. 7, pp. 583–605, Jan. 2011, doi: 10.1016/B978-0-444-51862-0.50018-6.
- [49] C. M. HURVICH e C.-L. TSAI, «Regression and time series model selection in small samples», *Biometrika*, vol. 76, n. 2, pp. 297–307, 1989, doi: 10.1093/biomet/76.2.297.

- [50] K. P. Burnham e D. R. Anderson, «Multimodel inference: Understanding AIC and BIC in model selection», *Sociol Methods Res*, vol. 33, n. 2, pp. 261–304, Nov. 2004, doi: 10.1177/0049124104268644.
- [51] S. M. A. Elrahman e A. Abraham, «A Review of Class Imbalance Problem», 2014.
- [52] N. U. Niaz, K. M. N. Shahariar, e M. J. A. Patwary, «Class Imbalance Problems in Machine Learning: A Review of Methods And Future Challenges», *International Conferences on Computing Advancements*, pp. 485–490, Mar. 2022, doi: 10.1145/3542954.3543024.
- [53] N. Lunardon, G. Menardi, e N. Torelli, «{ROSE}: a {P}ackage for {B}inary {I}mbalanced {L}earning», *{R} Journal*, vol. 6, n. 1, pp. 82–92, 2014.
- [54] K. Ramasubramanian e A. Singh, *Machine Learning Using R*, First edition. Nova Deli, Índia: Apress, 2017.
- [55] B. Lantz, *Machine Learning with R*, First edition. Birmingham, UK: Packt Publishing Ltd., 2013.
- [56] C. Dyke *et al.*, «Universal definition of perioperative bleeding in adult cardiac surgery», *J Thorac Cardiovasc Surg*, vol. 147, n. 5, pp. 1458-1463.e1, Mai. 2014, doi: 10.1016/j.jtcvs.2013.10.070.
- [57] «Defining Adult Overweight & Obesity | Overweight & Obesity | CDC». Acedido: 2 de Março de 2024. [Em linha]. Disponível em: <https://www.cdc.gov/obesity/basics/adult-defining.html>
- [58] R. D. Mosteller, «Simplified Calculation of Body-Surface Area», *New England Journal of Medicine*, vol. 317, n. 17, pp. 1098–1098, Out. 1987, doi: 10.1056/NEJM198710223171717.
- [59] R. Muraki *et al.*, «Novel method for estimating the total blood volume: the importance of adjustment using the ideal body weight and age for the accurate prediction of haemodilution during cardiopulmonary bypass», *Interact Cardiovasc Thorac Surg*, vol. 27, n. 6, pp. 802–807, Dez. 2018, doi: 10.1093/icvts/ivy173.
- [60] S. A. M. Nashef *et al.*, «EuroSCORE II», *European Journal of Cardio-Thoracic Surgery*, vol. 41, n. 4, pp. 734–745, Abr. 2012, doi: 10.1093/ejcts/ezs043.
- [61] S. van Buuren, *Flexible Imputation of Missing Data*, Second Edition. Boca Raton, FL.: Chapman & Hall/CRC press, 2018.
- [62] R Core Team, «R: A Language and Environment for Statistical Computing». R Foundation for Statistical Computing, Vienna, Austria, 2022. [Em linha]. Disponível em: <https://www.r-project.org/>
- [63] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
- [64] A. Kassambara, *ggpubr: «ggplot2» Based Publication Ready Plots*. 2023.

- [65] X. Robin *et al.*, «pROC: an open-source package for R and S+ to analyze and compare ROC curves», *BMC Bioinformatics*, vol. 12, 2011.
- [66] H. Wickham, R. François, L. Henry, K. Müller, e D. Vaughan, *dplyr: A Grammar of Data Manipulation*. 2023.
- [67] H. Wickham, D. Vaughan, e M. Girlich, *tidyr: Tidy Messy Data*. 2023.
- [68] N. Tierney e D. Cook, «Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations», *J Stat Softw*, vol. 105, n. 7, 2023, doi: 10.18637/jss.v105.i07.
- [69] E. Hvitfeldt, *paletteer: Comprehensive Collection of Color Palettes*. 2021.
- [70] J. Fox e S. Weisberg, *An {R} Companion to Applied Regression*. Sage, 2019.
- [71] M. Kuhn, «Building Predictive Models in R Using the caret Package», *J Stat Softw*, vol. 28, n. 5, 2008, doi: 10.18637/jss.v028.i05.
- [72] R. Manies Pereira *et al.*, «Activated clotting time value as an independent predictor of postoperative bleeding and transfusion», *Interdisciplinary CardioVascular and Thoracic Surgery*, vol. 38, n. 5, Mai. 2024, doi: 10.1093/ICVTS/IVAE092.
- [73] «International Normalized Ratio - MeSH - NCBI». Acedido: 24 de Março de 2024. [Em linha]. Disponível em: <https://www.ncbi.nlm.nih.gov/mesh/68019934>
- [74] L. Zhang, T. Geisler, H. Ray, e Y. Xie, «Improving logistic regression on the imbalanced data by a novel penalized log-likelihood function», *J Appl Stat*, vol. 49, n. 13, p. 3257, 2022, doi: 10.1080/02664763.2021.1939662.
- [75] C. Yang, E. A. Fridgeirsson, J. A. Kors, J. M. Reys, e P. R. Rijnbeek, «Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data», *J Big Data*, vol. 11, n. 1, pp. 1–17, Jan. 2024, doi: 10.1186/S40537-023-00857-7/FIGURES/6.
- [76] C. Salas-Eljatib, A. Fuentes-Ramirez, T. G. Gregoire, A. Altamirano, e V. Yaitul, «A study on the effects of unbalanced data when fitting logistic regression models in ecology», *Ecol Indic*, vol. 85, pp. 502–508, Fev. 2018, doi: 10.1016/j.ecolind.2017.10.030.
- [77] D. I. Warton e F. K. C. Hui, «The arcsine is asinine: the analysis of proportions in ecology», *Ecology*, vol. 92, n. 1, pp. 3–10, Jan. 2011, doi: 10.1890/10-0340.1.
- [78] W. Pan, «Akaike's information criterion in generalized estimating equations», *Biometrics*, vol. 57, n. 1, pp. 120–125, 2001, doi: 10.1111/J.0006-341X.2001.00120.X.
- [79] H. Bozdogan, «Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions», *Psychometrika*, vol. 52, n. 3, pp. 345–370, Set. 1987, doi: 10.1007/BF02294361/METRICS.

- [80] H. Bozdogan, «Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions», *Psychometrika*, vol. 52, n. 3, pp. 345–370, Set. 1987, doi: 10.1007/BF02294361/METRICS.
- [81] M. Aktar, J. Reyes, D. Tampieri, H. Rivaz, Y. Xiao, e M. Kersten-Oertel, «Deep learning for collateral evaluation in ischemic stroke with imbalanced data», *Int J Comput Assist Radiol Surg*, vol. 18, n. 4, pp. 733–740, Abr. 2023, doi: 10.1007/S11548-022-02826-6.
- [82] K. C. Chen, S. W. Kuo, R. H. Shie, e H. Y. Yang, «Advancing accuracy in breath testing for lung cancer: strategies for improving diagnostic precision in imbalanced data», *Respir Res*, vol. 25, n. 1, Dez. 2024, doi: 10.1186/S12931-024-02668-7.
- [83] A. Mitrovska, P. Safari, K. Ritter, B. Shariati, e J. K. Fischer, «Secure federated learning for Alzheimer's disease detection», *Front Aging Neurosci*, vol. 16, 2024, doi: 10.3389/FNAGI.2024.1324032.
- [84] F. Wang, X. Yang, X. Zhang, e F. Hu, «Monitoring the after-effects of ischemic stroke through EEG microstates», *PLoS One*, vol. 19, n. 3, p. e0300806, Mar. 2024, doi: 10.1371/JOURNAL.PONE.0300806.
- [85] «Introduction to Regression Methods for Public Health Using R». Acedido: 24 de Março de 2024. [Em linha]. Disponível em: <https://bookdown.org/rwnahhas/RMPH/>
- [86] S. Kietaihl *et al.*, «Management of severe peri-operative bleeding: Guidelines from the European Society of Anaesthesiology and Intensive Care: Second update 2022», *Eur J Anaesthesiol*, vol. 40, n. 4, pp. 226–304, Abr. 2023, doi: 10.1097/EJA.0000000000001803.
- [87] S. Halvorsen, J. Mehilli, S. Choorapoikayil, e K. Zacharowski, «Extract from the 2022 ESC Guidelines on cardiovascular assessment and management of patients undergoing non-cardiac surgery - Patient Blood Management», *Blood Transfusion*, vol. 22, n. 2, p. 122, Mar. 2024, doi: 10.2450/BLOODTRANSFUS.708.

APÊNDICE

A.1 Gráficos da análise do modelo com base de dados parcialmente balanceados

Figura 63 - Gráfico de Resíduos do Modelo com dados parcialmente balanceados

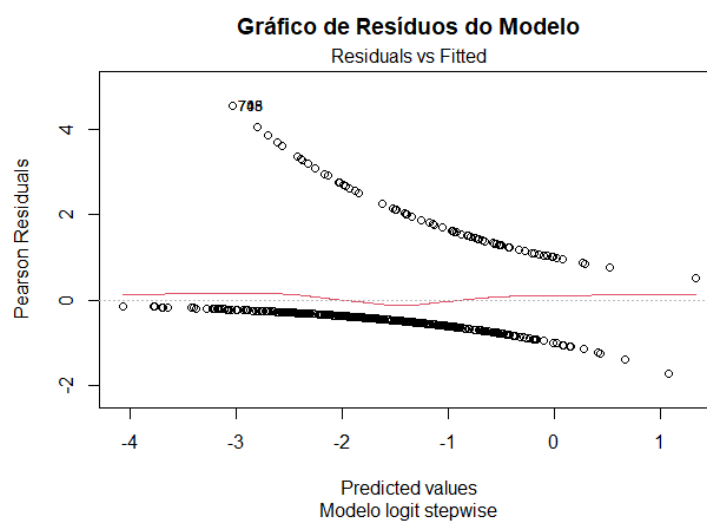


Figura 64 - Gráfico de Resíduos do Modelo com dados parcialmente balanceados

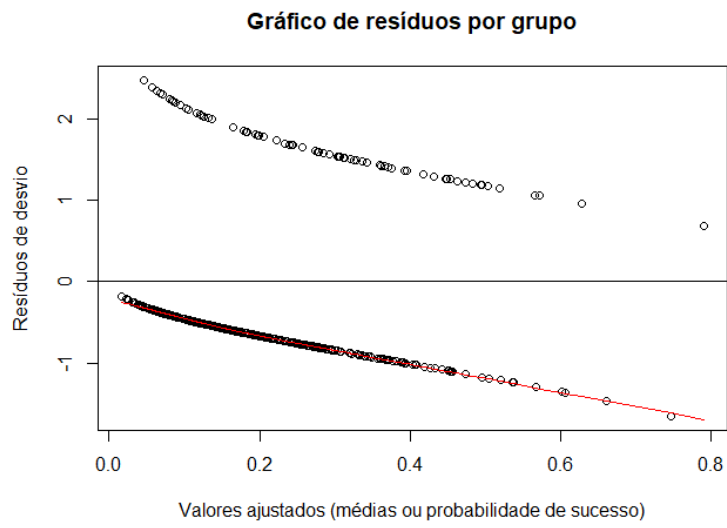


Figura 65 - Gráfico de resíduos padronizados do modelo com a base de dados parcialmente balanceada

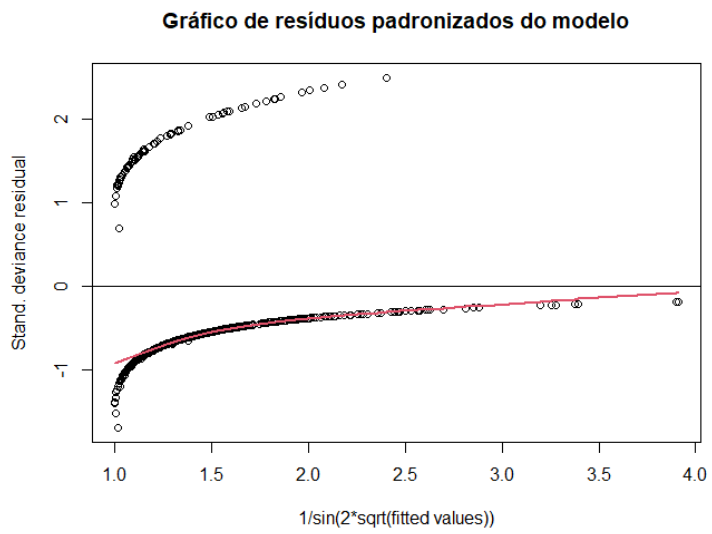
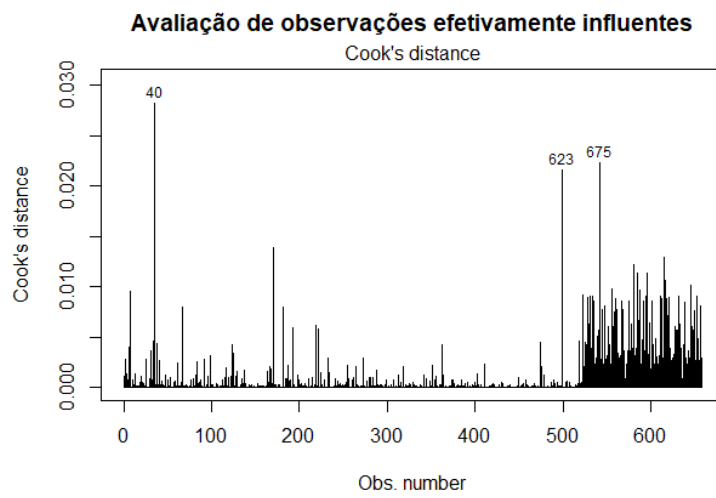


Figura 66 - Avaliação de observações efetivamente influentes do modelo com dados parcialmente balanceados



A.2 Gráficos da análise do modelo com base de dados completamente balanceados

Figura 67 - Gráfico de Resíduos do Modelo com dados completamente balanceados

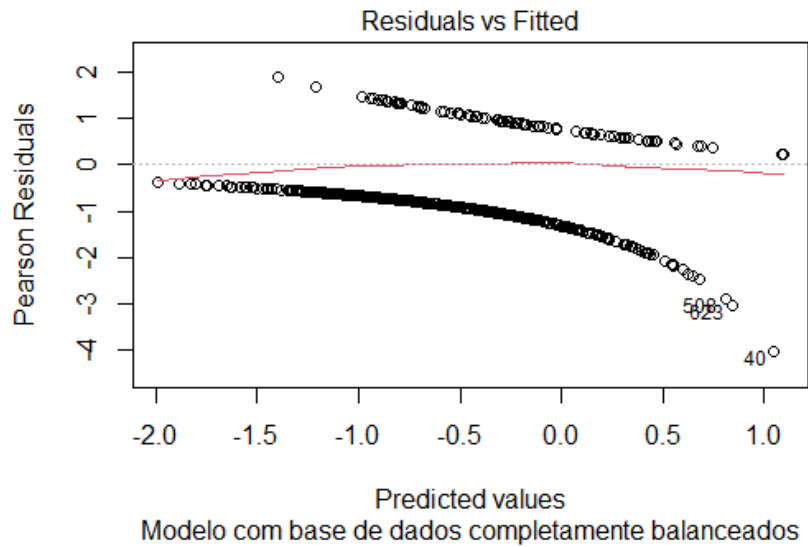


Figura 68 - Gráfico de Resíduos do Modelo com dados completamente balanceados

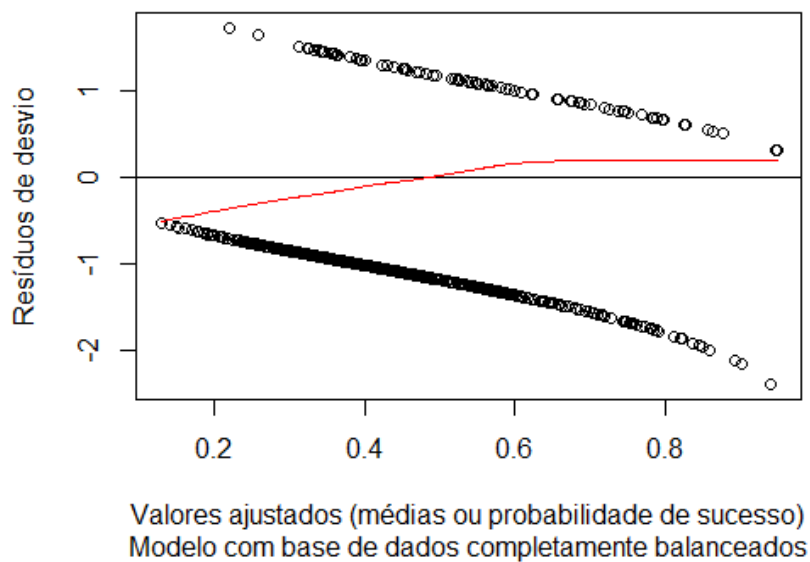


Figura 69 - Gráfico de resíduos padronizados do modelo com a base de dados completamente balanceada

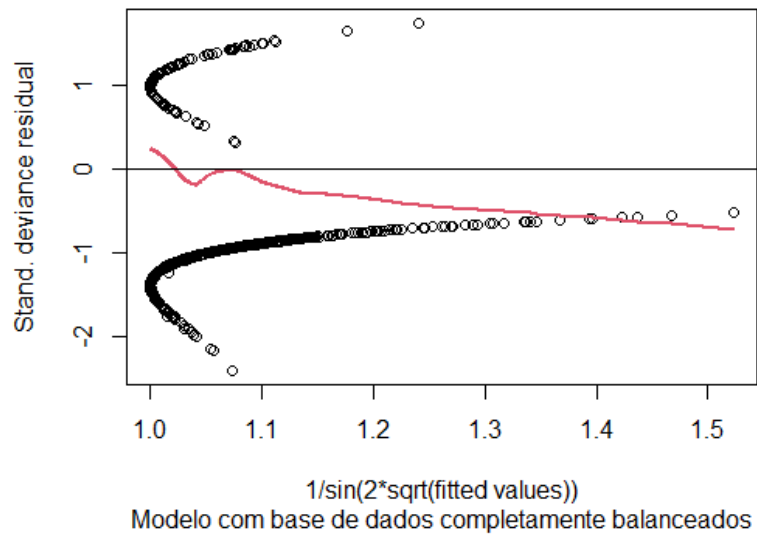
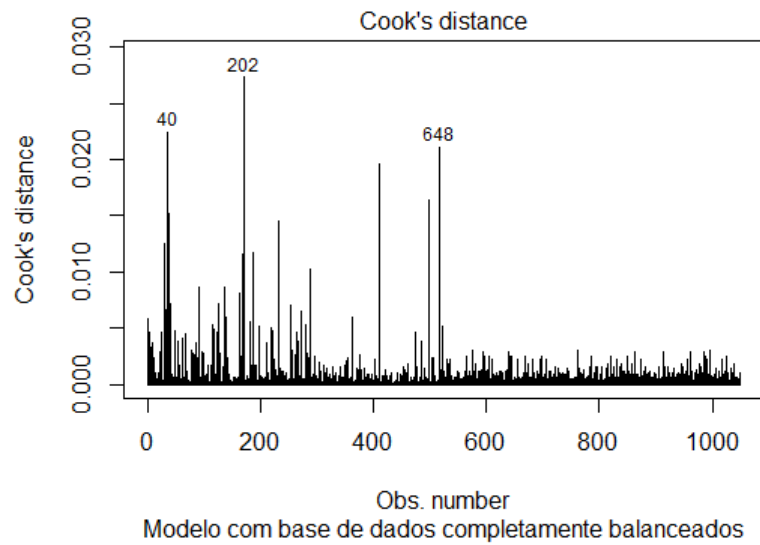


Figura 70 - Avaliação de observações efetivamente influentes do modelo com base de dados completamente balanceados





2024

RAFAEL MANIÉS PEREIRA

ASSOCIAÇÃO ENTRE PRIMING COM GELATINA E
HEMORRAGIA PÓS-OPERATÓRIA EM CIRURGIA CARDÍACA