

UNIVERSIDADE ABERTA

Mestrado em Bioestatística e Biometria

**Modelação Estatística: Estudo do perfil cardiovascular
de adultos assintomáticos, da cidade de Benguela**

Ana Isabel Ramos Gonçalves Cunha

Dissertação apresentada na Universidade Aberta para obtenção do grau de mestre em
Bioestatística e Biometria.

Orientadora:

Professora Doutora Teresa Paula da Costa Azinheira Oliveira

Lisboa, 2014

Ao Zé... sempre ao meu lado, seja qual for a aventura!

Aos meus maiores tesouros: o Pedro e o André!

Aos meus pais... os meus maiores protetores!

Agradecimentos

À doutora Teresa Oliveira, pela paciência que sempre teve, mesmo nos meus momentos de desvaneios, apoiando-me sempre e ajudando-me a ultrapassar todas as dificuldades, que iam surgindo. Durante estes dois anos sonhámos juntas e acreditámos que seria possível. E foi!

Ao doutor Pedro Catarino, pelo apoio incondicional, num país em que nem tudo é tão fácil quanto parece. Sem ele, teria sido impossível avançar com um estudo destes em Angola.

A todos os meus colegas do Instituto Superior Politécnico de Benguela que sempre se mostraram disponíveis para auxiliar no que fosse preciso e cuja ajuda foi tão necessária na altura da recolha de dados.

Aos mais de 700 Benguelenses que gentilmente se dirigiram ao largo de África e aceitaram participar nesta dissertação, dando-me o privilégio de explorar uma base de dados inédita.

A todos os meus familiares e amigos, por me apoiarem incondicionalmente. Há muito tempo perdido a ser recuperado e a promessa de não falar sobre risco cardiovascular durante, pelo menos, um ano!

À minha irmã de coração, Diana, por ter vibrado comigo desde a primeira página e pela forma bondosa com que partilhou comigo todo o material que tinha das suas aulas de cardiologia e pela paciência com que leu e releu, vezes sem conta, este trabalho.

Ao meu Zé... pelo altruísmo com que assumiu este sonho como “nosso” e por, numa altura em que julguei que não iria ser possível, ele ter arregaçado as mangas e me ter contemplado com a mais linda prova de amor de sempre.

A Deus... por tudo e todos, que referi anteriormente.

Resumo

Há alguns anos que as doenças cardiovasculares lideram o *ranking* das doenças que mais pessoas matam no mundo. Este conjunto de doenças sempre teve maiores prevalências nos países desenvolvidos, enquanto os países não desenvolvidos sempre estiveram associados a maiores prevalências de doenças transmissíveis, como a malária, o HIV ou a tuberculose. Apostando em estratégias de prevenção e controlo, os países desenvolvidos têm conseguido baixar a sua prevalência, mas os países desenvolvidos carregam atualmente um duplo fardo, pois têm grandes prevalências dos dois tipos de doenças, criando um fosso ainda maior entre alguns países.

A luta contra estas doenças não é fácil, pois o seu desenvolvimento é silencioso e quando se manifestam são, geralmente, fatais. Quando não são fatais, acarretam a um encargo bastante pesado tanto para as famílias como para o sistema de saúde, que no caso dos países não desenvolvidos são, ainda, bastante frágeis. É, por isso, mais do que evidente que a melhor estratégia é a prevenção, controlando os fatores de risco, conhecendo as suas prevalências e a forma como interagem entre si e criando modelos preditivos de análise de risco. Tendo em conta o caráter multifatorial destas doenças a estatística multivariada tem-se revelado fundamental, devido a técnicas como as da modelação estatística.

Em Angola, pouco se sabe sobre as doenças cardiovasculares, além de que as suas prevalências já estão bastante elevadas. Nesta dissertação foi estudado o perfil cardiovascular de 600 benguelenses, de acordo com variáveis sociais, antropométricas, clínicas e de estilos de vida, recorrendo a técnicas univariadas, bivariadas e multivariadas. Os modelos de regressão logística estereótipos foram bastante úteis, tendo em conta o caráter ordinal da variável resposta (risco cardiovascular). Com este trabalho pretendeu-se chamar a atenção para as doenças cardiovasculares em Angola e para a gravidade da sua prevalência, bem como conseguir identificar os principais grupos de risco, de forma a serem tomadas medidas que reduzam a exposição aos fatores de risco e, conseqüentemente, o desenvolvimento dessas doenças.

Palavras-chave: Análise de risco; modelo estereótipo; doenças cardiovasculares; fatores de risco; Angola

Summary

For several years that cardiovascular diseases lead the ranking of the diseases that kill the most people in the world. This group of diseases has always had the highest prevalence in developed countries, while the non-developed countries had always been associated with higher prevalence of communicable diseases such as malaria, HIV or tuberculosis. Betting on strategies for prevention and control, developed countries have managed to decreased its prevalence, but, currently developed countries carry a double burden because they have large prevalence of both types of this diseases, creating an even bigger gap between some countries.

The fight against these diseases is not easy, as its development is quiet and when they occur, they are usually fatal. When they're not fatal, carry a very heavy weight on both families and the health system, which in the case of undeveloped countries are still quite fragile. It is therefore abundantly clear that the best strategy is prevention by controlling risk factors, knowing their prevalences and how they interact with each other and creating predictive models for risk analysis. Given the multifactorial nature of these diseases multivariate statistic has proved crucial, due to techniques such as statistical modelling.

In Angola, little is known about cardiovascular diseases, and that their prevalence rates are already quite high. In this dissertation the cardiovascular profile of 600 Benguela residents, according to social, anthropometric , clinical and lifestyle variables was studied using univariate, bivariate and multivariate techniques. The stereotype logistic models were very useful, taking into account the ordinal character of the response variables (cardiovascular risk). With this dissertation is intended to draw attention to cardiovascular disease in Angola and the severity of its prevalence, be able to identify the main risk groups, in a way that measures can be taken to reduce exposure to risk factors and consequently, the development of these diseases.

Keywords: Risk analysis; Stereotype model; Cardiovascular diseases; Risk factors; Angola

Índice

Agradecimentos.....	2
Resumo	3
Índice	5
Índice de Tabelas	9
Índice de Figuras.....	11
Índice de Anexos.....	16
Introdução	17
PARTE I.....	20
1. TÉCNICAS DE INFERÊNCIA ESTATÍSTICA NÃO PARAMÉTRICA	21
1.1. Introdução.....	21
1.2. Teste de ajustamento à normalidade de Kolmogorov-Smirnov	24
1.3. Teste da Binomial.....	27
1.4. Teste de independência de Mann-Whitney	29
1.5. Teste de independência de Kolmogorov-Smirnov.....	33
1.6. Odds Ratio (OR).....	36
2. MODELAÇÃO ESTATÍSTICA.....	38
2.1. Introdução.....	38
2.2. Modelos de Regressão Linear Generalizados	41
2.2.1.Introdução.....	41
2.2.2.Modelo de Regressão Logística Binária	45
2.2.3.Modelos de Regressão Logística Ordinal	51
3. DOENÇAS CARDIOVASCULARES.....	60
3.1. Dimensão do problema.....	60
3.2. Doenças Cardiovasculares.....	62

3.3. Fatores de Risco Cardiovascular	63
3.4. Avaliação do Risco Cardiovascular	71
PARTE II.....	75
1. INTRODUÇÃO	76
2. ANÁLISE UNIVARIADA.....	84
2.1. Indicadores sociais	84
2.2. Indicadores antropométricos.....	86
2.3. Indicadores clínicos	92
2.4. Indicadores de estilo de vida	95
2.5. Avaliação do risco cardiovascular	96
3. ANÁLISE BIVARIADA.....	98
3.1. Comportamento dos indicadores de risco cardiovascular, por género	98
3.1.1. Indicadores antropométricos versus Género	98
3.1.2. Indicadores clínicos versus Género.....	100
3.1.3. Indicadores de estilo de vida versus Género.....	101
3.2. Comportamento dos indicadores de risco cardiovascular, entre os mais velhos e os mais novos	103
3.2.1. Indicadores antropométricos versus Faixa etária.....	103
3.2.2. Indicadores clínicos versus Faixa etária	105
3.2.3. Indicadores de estilo de vida versus Faixa etária	106
3.3. Comportamento dos indicadores de risco cardiovascular, entre os indivíduos com sobrepeso e os sem sobrepeso	107
3.3.1. Indicadores antropométricos versus Sobrepeso	107
3.3.2. Indicadores clínicos versus Sobrepeso.....	108
3.3.3. Indicadores de estilo de vida versus Sobrepeso	109

3.4. Comportamento dos indicadores de risco cardiovascular, entre os indivíduos com obesidade abdominal e sem obesidade abdominal.....	110
3.4.1. Indicadores antropométricos versus Obesidade abdominal.....	110
3.4.2. Indicadores clínicos versus Obesidade abdominal	111
3.4.3. Indicadores de estilo de vida versus Obesidade abdominal.....	112
3.5. Comportamento dos indicadores de risco cardiovascular, entre os hipertensos e os normotensos.....	113
3.5.1. Indicadores antropométricos versus Hipertensão	113
3.5.2. Indicadores clínicos versus Hipertensão	115
3.5.3. Indicadores de estilo de vida versus Hipertensão	116
3.6. Comportamento dos indicadores de risco cardiovascular, entre os fumadores e os não fumadores.....	116
3.6.1. Indicadores antropométricos versus Tabagismo	116
3.6.2. Indicadores clínicos versus Tabagismo	118
3.6.3. Indicadores de estilo de vida versus Tabagismo.....	119
3.7. Comportamento dos indicadores de risco cardiovascular, entre os consumidores de bebidas alcoólicas e os não consumidores.....	119
3.7.1. Indicadores antropométricos versus Álcool	119
3.7.2. Indicadores clínicos versus Álcool.....	120
3.7.3. Indicadores de estilo de vida versus Álcool	121
3.8. Comportamento dos indicadores de risco cardiovascular, entre os consumidores de gorduras saturadas e os não consumidores.....	122
3.8.1. Indicadores antropométricos versus Gorduras saturadas.....	122
3.8.2. Indicadores clínicos versus Gorduras saturadas	123
3.9. Comportamento do Risco Cardiovascular, por grupos de risco	124
3.10. Correlações entre Risco Cardiovascular e alguns indicadores.....	135

3.11. Grau de associação entre os Riscos Cardiovasculares	137
4. ANÁLISE MULTIVARIADA	138
4.1. Modelação do Risco A.....	138
4.2. Modelação do Risco B	148
4.3. Modelação do Risco C	152
5. CONCLUSÕES E RECOMENDAÇÕES	153
REFERÊNCIAS BIBLIOGRÁFICAS.....	159

Índice de Tabelas

Tabela 1: Valores de referência para medição do grau de associação de variáveis, segundo Cohen e Holliday (1982).....	33
Tabela 2: Tabela de contingência (2x2) de apoio ao cálculo do Odds Ratio	37
Tabela 3: Algumas distribuições da família exponencial.....	44
Tabela 4: Percentagens de óbitos causados por doenças cardiovasculares, em 2008, nos países angolanos de língua oficial portuguesa (PALOP)	60
Tabela 5: Estratificação do risco cardiovascular tendo em conta o género, o índice de massa corporal e o perímetro abdominal	73
Tabela 6: Estratificação do risco cardiovascular tendo em conta a condição de hipertensão e o número de fatores de risco adicionais.....	73
Tabela 7: Classificação internacional de um adulto quanto ao peso, pelo índice de massa corporal	77
Tabela 8: Classificação de um adulto quanto aos valores das pressões arteriais.....	78
Tabela 9: Estratificação das variáveis, segundo o risco para Doenças Cardiovasculares	80
Tabela 10: Output (SPSS) do teste da binomial para a variável género (n=600)	84
Tabela 11: Output (SPSS) do teste de ajustamento à normalidade, para a variável idade (n=600).....	85
Tabela 12: Frequências absolutas e relativas da estimação do risco de nos próximos dez anos os participantes desenvolverem um evento cardiovascular, fatal ou não fatal (n=237).....	97
Tabela 13: Output, do SPSS, do teste U de Mann-Whitney, para as variáveis antropométricas (peso; estatura; índice de massa corporal; perímetro abdominal), agrupadas por género (n=600).....	99
Tabela 14: Resumo do teste de Qui-quadrado, combinado com o cálculo do Odds Ratio, para cada um dos indicadores de estilo de vida, agrupados por género	102

Tabela 15: Output (SPSS) do teste de ajustamento à normalidade para cada uma das variáveis antropométricas, agrupadas por idade (n=600)	103
Tabela 16: Output, do SPSS, do teste U de Mann-Whitney, para as variáveis antropométricas (peso; estatura; índice de massa corporal; perímetro abdominal), agrupadas por condição de sobrepeso	108
Tabela 17: Resultados do cálculo do coeficiente de Spearman	137
Tabela 18: Modelos que ajustam significativamente os dados, segundo o teste da razão de verossimilhanças	139
Tabela 19: Resumo das estimativas para os coeficientes do modelo	141
Tabela 20: Coeficientes e Odds Ratio para cada uma das covariáveis do modelo estereótipo, para o Risco A.....	143
Tabela 21: Primeiros dez casos do processo de validação do modelo	146
Tabela 22: Tabela de contingência comparativa de risco observado com o risco modelado (n=100)	147
Tabela 23: Modelos que ajustam significativamente os dados, segundo o teste da razão de verossimilhanças	149
Tabela 24: Coeficientes e Odds Ratio para cada uma das covariáveis do modelo estereótipo, para o Risco B.....	150

Índice de Figuras

Figura 1: Output do SPSS de um teste de aderência à normalidade, de uma variável..	26
Figura 2: Output do R num teste da binomial	29
Figura 3: Fluxograma representativo da Modelação Estatística, considerada como um processo iterativo	40
Figura 4: Mortalidade por doenças cardiovasculares, agrupadas por género, de acordo com o Banco Mundial (por 100 000)	64
Figura 5: Distribuição mundial da mortalidade originada por doenças cardiovasculares, em homens (estratificada por idades, por 100 000)	65
Figura 6: Distribuição mundial da mortalidade originada por doenças cardiovasculares, em mulheres (estratificada por idades, por 100 000)	65
Figura 7: Estimação do risco de, em 10 anos, um africano desenvolver um evento cardiovascular, fatal ou não fatal, de acordo com o género, os hábitos tabágicos, a idade e a pressão arterial sistólica	72
Figura 8: Estatísticas relativas à mortalidade por doenças cardiovasculares (standardizadas por idade), em alguns países desenvolvidos, entre 1980 e 2008	76
Figura 9: Distribuição dos participantes (n=600) por género e por faixas etárias	79
Figura 10: Distribuição do género dos participantes, por faixas etárias	79
Figura 11: Distribuição dos participantes por idades, em anos (n=600)	85
Figura 12: Distribuição dos participantes por grupo de risco, de acordo com a idade (n=600)	86
Figura 13: Distribuição dos pesos dos participantes, em kg (n=600)	87
Figura 14: <i>Normal Q-Q Plot</i> e <i>Detrended Normal Q-Q Plot</i> relativos à variável estatura, em metros (n=600)	87
Figura 15: Distribuição dos índices de massa corporal dos participantes, em kg/m ² (n=600)	88

Figura 16: Q-Q <i>Normal Plot</i> com <i>outliers</i> (esquerda) e sem <i>outliers</i> (direita), para o índice de massa corporal, em kg/m ² (n=600).....	89
Figura 17: Distribuição dos participantes, de acordo com o índice de massa de massa corporal (n=600).....	89
Figura 18: Distribuição dos indivíduos por grupo de risco, de acordo com o índice de massa corporal (n=600).....	90
Figura 19: Distribuição do perímetro abdominal dos participantes, em cm (n=600)....	91
Figura 20: Distribuição dos participantes, de acordo com o perímetro abdominal (n=600).....	91
Figura 21: Distribuição dos participantes por grupo de risco, de acordo com o perímetro abdominal (n=600).....	92
Figura 22: Distribuição da pressão arterial sistólica dos participantes, em mmHg (n=600).....	93
Figura 23: Distribuição da pressão arterial diastólica dos participantes, em mmHg (n=600).....	93
Figura 24: Distribuição dos participantes por classes, de acordo com a pressão arterial (n=600).....	94
Figura 25: Distribuição dos participantes, de acordo com a condição de hipertensão (n=600).....	95
Figura 26: Distribuição dos participantes por grupo de risco, de acordo com os hábitos tabágicos, o consumo de bebidas alcoólicas e o consumo de gorduras saturadas, respetivamente (n=600).....	95
Figura 27: Distribuição dos participantes pelas categorias do risco A (n=600)	96
Figura 28: Distribuição dos participantes pelas categorias do risco B (n=600).....	96
Figura 29: Estimação do risco de nos próximos dez anos os participantes desenvolverem um evento cardiovascular, fatal ou não fatal, considerando os casos válidos (n=237) e os inválidos (n=363)	97

Figura 30: Distribuição dos indicadores antropométricos (peso; estatura; índice de massa corporal; perímetro abdominal), por género (n=600)	98
Figura 31: Distribuição dos indicadores clínicos (pressão arterial sistólica e pressão arterial diastólica), por géneros (n=600)	101
Figura 32: Distribuição dos participantes de acordo com os indicadores de estilo de vida, por género.....	102
Figura 33: Distribuição dos indicadores antropométricos, no grupo dos indivíduos mais velhos e no dos indivíduos mais jovens (n=600)	104
Figura 34: Distribuição dos indicadores clínicos, no grupo dos indivíduos mais velhos e no dos indivíduos mais jovens.....	106
Figura 35: Distribuição da pressão arterial sistólica e da diastólica nos indivíduos com sobrepeso e sem sobrepeso (n=600)	109
Figura 36: Distribuição da pressão arterial sistólica e da pressão arterial diastólica, de acordo com a condição de obesidade abdominal.....	112
Figura 37: Distribuição dos indicadores antropométricos, no grupo dos hipertensos e dos normotensos (n=600)	114
Figura 38: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos hipertensos e dos normotensos (n=600)	115
Figura 39: Distribuição das medidas antropométricas, nos fumadores e não fumadores (n=600).....	117
Figura 40: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos fumadores e dos não fumadores (n=600)	118
Figura 41: Distribuição dos indicadores antropométricos nos consumidores de álcool e nos não consumidores (n=600)	120
Figura 42: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos consumidores de álcool e dos não consumidores (n=600).....	121
Figura 43: Distribuição dos indicadores antropométricos nos consumidores de gorduras saturadas e nos não consumidores (n=600)	122

Figura 44: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos consumidores de gorduras saturadas e dos não consumidores (n=600)	124
Figura 45: Gráfico de barras relativo à distribuição do risco A, de acordo com o género	125
Figura 46: Gráfico de barras relativo à distribuição do risco A, de acordo com a hipertensão.....	127
Figura 47: Gráfico de barras relativo à distribuição do risco B, de acordo com o género	128
Figura 48: Gráfico de barras relativo à distribuição do risco B, de acordo com a faixa etária.....	129
Figura 49: Gráfico de barras relativo à distribuição do risco B, de acordo com o sobrepeso	130
Figura 50: Gráfico de barras relativo à distribuição do risco B, de acordo com a hipertensão.....	131
Figura 51: Gráfico de barras relativo à distribuição do risco B, de acordo com os hábitos tabágicos.....	132
Figura 52: Gráfico de barras relativo à distribuição do risco B, de acordo com os hábitos alcoólicos	133
Figura 53: Output do <i>software</i> R dos gráficos das correlações entre o Risco A e o perímetro abdominal, o índice de massa corporal, o peso e a pressão arterial sistólica	136
Figura 54: Gráfico de resíduos, por contrastes	140
Figura 55: Output do <i>software</i> R da função <i>Coef()</i>	142
Figura 56: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função do perímetro abdominal (cm)	144
Figura 57: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função do índice de massa corporal (kg/m^2).....	145

Figura 58: Gráfico de barras com a percentagem de risco modelado em cada uma das categorias do risco observado.....	146
Figura 59: Distribuição dos desvios entre o risco observado e o risco modelado	148
Figura 60: Output do software R da função <i>Coef()</i>	149
Figura 61: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função da idade	151
Figura 62: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função do peso (kg)	151
Figura 63: Distribuição dos desvios entre o risco observado e o risco modelado	152

Índice de Anexos

Anexo 1: Tipos de doenças cardiovasculares e fatores de riscos correspondentes	164
Anexo 2: Output do SPSS da Análise Univariada	166
Anexo 3: Output do SPSS da análise descritiva, por género	170
Anexo 4: Output do SPSS da análise descritiva, por idade.....	176
Anexo 5: Output do SPSS da análise descritiva, por condição de sobrepeso	182
Anexo 6: Output do SPSS da análise descritiva, por condição de obesidade	188
Anexo 7: Output do SPSS da análise descritiva, por condição de hipertensão.....	194
Anexo 8: Output do SPSS da análise descritiva, de acordo com os hábitos tabágicos	200
Anexo 9: Output do SPSS da análise descritiva, de acordo com os hábitos de ingestão de álcool.....	206
Anexo 10: Output do SPSS da análise descritiva, de acordo com os hábitos de consumo de gorduras saturadas	212
Anexo 11: Tabela de Contingência relativa ao Risco A, por grupos de risco	218
Anexo 12: Tabela de Contingência relativa ao Risco B, por grupos de risco	223
Anexo 13: Tabela de Contingência relativa ao Risco C, por grupos de risco.....	228
Anexo 14: Comandos do <i>software</i> R usados no ajuste dos modelos de Regressão Ordinal	233
Anexo 15: Apresentação em Powerpoint, usada na defesa pública da dissertação (14 de Fevereiro de 2014).....	233

Introdução

As doenças cardiovasculares são a principal causa de morte, em todo o mundo, desenvolvendo-se silenciosamente, no meio do ruído resultante do desenvolvimento e do progresso, e que quando se manifestam são extremamente impiedosas, arruinando a estabilidade de famílias e de países, sobretudo nos casos em que o desfecho não é a morte.

Estas doenças têm sido encaradas como o preço a pagar pelo desenvolvimento mundial. O reverso de moeda, associado a um estilo de vida demasiado sedentário e a uma predominância de açúcares, gorduras saturadas, sal e álcool na dieta, são os problemas de saúde provocados por estes hábitos tais como obesidade, colesterol, hipertensão, diabetes, entre outros.

Com todas as atenções voltadas para esta epidemia, vários países desenvolvidos, recorrendo ao planeamento de experiências e à investigação conseguiram um real conhecimento da sua dinâmica, permitindo-lhes definir estratégias que culminaram no controlo e diminuição do seu impacto. Devido ao carácter dissimulado desta doença, que quando se manifesta é para matar, a melhor estratégia é a prevenção.

Num mundo em que se pensa que o mal bate sempre na porta do vizinho, a análise e a modelação do risco cardiovascular tiveram um lugar de destaque, na redução das taxas de prevalência de fatores de risco e consequente diminuição das taxas de mortalidade, funcionando como uma bola de cristal em que cada indivíduo era associado a uma probabilidade de ocorrência de um evento cardiovascular. Esta individualização do risco e responsabilização de cada um pela sua própria saúde e pela saúde dos que lhe são mais chegados, levou à mudança de estilos de vida e consequente redução do risco cardiovascular, de uma forma geral.

Estas doenças, sempre associadas aos países mais desenvolvidos, têm mudado a sua dinâmica. Em África, na maior parte dos países, as taxas de morte por doenças cardiovasculares já igualam as taxas de morte por HIV e prevê-se que até 2025 se tornem na principal causa de morte dos países da região subsariana. As doenças cardiovasculares chacinam agora os países subdesenvolvidos, que se vêm obrigados a

continuar a luta contra as doenças transmissíveis, como o HIV, a malária e a tuberculose, e a começar a luta contra as doenças cardiovasculares, para a qual a população e o sistema de saúde ainda não estão preparados.

Angola é um desses países que carrega um duplo fardo. O fim da guerra civil, há pouco mais de dez anos, deu início ao desenvolvimento do país. A melhoria gradual das condições socioeconómicas, a influência das multiculturas que se juntaram para erguer o país, a emancipação gradual das mulheres, a facilidade de acesso a *fa(s)t food* e muitos outros fatores formaram um sistema extremamente propício ao desenvolvimento de doenças cardiovasculares e de outras que lhes estão associadas. Por outro lado, a guerra ainda está bastante marcada, com altas prevalências de doenças transmissíveis.

É urgente que as doenças cardiovasculares ocupem um lugar de destaque no sistema de saúde e que haja um compromisso dos seus profissionais para com este desafio. É igualmente urgente apostar na investigação nesta área, estudando os fatores de risco e construindo modelos preditivos específicos para o contexto angolano. A criação da Sociedade Angolana de Cardiologia, no final de 2012, teve um papel bastante importante nesta luta, pois tem chamado a atenção para a crescente necessidade de pesquisa, formação e informação na área das doenças cardiovasculares, bem como para a importância de se cumprirem com os desafios do milénio.

Nesta dissertação, será feito um estudo bastante abrangente sobre este tema, recorrendo aos dados de 600 benguelenses, com o objetivo de se aferir a sua exposição a alguns fatores de risco e tentar encontrar um modelo simples que permita fazer uma predição de risco cardiovascular, para que se comecem a delinear algumas estratégias de prevenção. Foram, ainda recolhidos dados de mais 100 benguelenses, necessários para a validação do modelo.

A dissertação foi dividida em duas partes. A primeira parte é um enfoque teórico sobre os conceitos-chave desta temática, tendo-se começado pelos relacionados com a modelação estatística e posteriormente exploraram-se os relacionados com as doenças cardiovasculares. A segunda parte do trabalho é prática e na qual será

estudado o perfil cardiovascular dos benguelenses, de acordo com alguns grupos considerados de risco e onde será feita uma análise de risco.

Tendo em conta o caráter multifatorial das doenças cardiovasculares, a análise bivariada, apesar de ser bastante importante, é insuficiente para o real conhecimento do perfil cardiovascular destes indivíduos, uma vez que se sabe que a sinergia dos fatores de risco vai muito além da simples adição dos efeitos de cada um. A evolução da análise multivariada, em que podem ser estudados os efeitos conjuntos de múltiplos fatores de risco, mostra-se, assim, como a aliada perfeita para a análise do risco destes indivíduos.

O caminho escolhido para a análise multivariada foi a modelação estatística e, tendo em conta o caráter ordinal da variável dependente (risco cardiovascular), o modelo logístico ordinal foi considerado o ideal. Uma vez que as variáveis dependentes são originalmente ordinais discretas, o modelo estereótipo revelou-se o mais adequado.

Os modelos ordinais, por si só, ainda são pouco usados e, quando o são, a escolha mais encontrada na literatura é o modelo de *odds* proporcionais, mesmo porque é a única alternativa que o SPSS, por exemplo, possui para este grupo de modelos. As alternativas, em termos informáticos, para os modelos estereótipos são escassas e o R, graças a pacotes como o *VGAM* e a comandos como o *rrvglm*, tem-se mostrado uma mais-valia para estudos deste género, em que muitas vezes se cede ao uso, inadequado, do modelo de *odds* proporcionais ou ao modelo multinomial. Neste sentido, na segunda parte, todo o procedimento que esteve associado à construção do modelo estereótipo encontra-se bastante detalhado.

Com esta dissertação, além de se pretender ajudar a abrir caminho no estudo das doenças cardiovasculares em Angola, para que se comecem a delinear estratégias de controlo e prevenção, pretende-se, também, contribuir para a divulgação da relevância do modelo estereótipo na área da epidemiologia e chamar a atenção para algumas lacunas que ainda lhe estão associadas que, juntamente com a sua complexidade matemática, fazem com que ele seja, muitas vezes, menosprezada.

PARTE I

"All models are wrong. Some are useful."

George E. P. Box

1. TÉCNICAS DE INFERÊNCIA ESTATÍSTICA NÃO PARAMÉTRICA

1.1. Introdução

A Estatística é uma ciência que estuda a variabilidade apresentada por dados e que, a partir deles, consegue extrair algumas conclusões com um determinado grau de confiança. Um dos ramos da Estatística é a Inferência Estatística, constituída por um conjunto de técnicas que permitem tirar conclusões sobre uma população, apenas com base numa amostra extraída dessa mesma população e recorrendo, para isso, a técnicas de estimação pontual, a intervalos de confiança e a testes estatísticos paramétricos ou não paramétricos (Oliveira, 2004).

A escolha do tratamento estatístico mais adequado varia consoante o tipo de variáveis em estudo. Há quatro tipos de escalas a ter em conta: nominal, ordinal, intervalo e rácio. Numa escala nominal os elementos são qualidades e os números utilizados servem apenas para codificar esses elementos. Na segunda parte deste trabalho, uma variável nominal estudada foi o género, que assumiu as categorias masculino, que foi associada ao código 0, e feminino, que foi associada ao código 1. Esta variável é um caso particular das escalas nominais que surge quando a característica em estudo tem apenas duas categorias, sendo por isso designadas por variáveis binárias ou dicotómicas.

Numa escala ordinal existe uma relação de ordem entre as diferentes categorias do atributo e os códigos são associados de maneira a que essa ordenação se mantenha (Pestana & Gageiro, 2008). Neste trabalho, uma variável ordinal estudada foi o risco cardiovascular, em que o código 0 foi associado ao grau menos grave de risco e a numeração foi aumentando até atingir a categoria mais grave. Um outro modo de obtenção de uma escala ordinal é a divisão de uma escala contínua em múltiplos intervalos. Neste trabalho foi efetuado esse procedimento com o índice de massa corporal, que por um lado era uma variável contínua, mas por outro também assumia a estrutura de uma variável ordinal, recorrendo a três pontos de corte, cujo procedimento será descrito mais adiante.

Numa escala de intervalo define-se arbitrariamente um ponto zero, obtendo-se depois o outro extremo por operações sucessivas utilizando-se sempre a mesma unidade de medida – os intervalos. As escalas de intervalo incluem o número 0, mas não possuem um zero absoluto.

Uma escala de rácio tem as mesmas propriedades de uma escala por intervalos e adicionalmente apresenta a característica de possuir um zero absoluto como valor mínimo. Nesta escala, os algarismos que a compõem mantêm a mesma distância entre si, conservando a igualdade de razões. Com dados deste tipo, alterações nas unidades de medida não afetam os rácios entre dois valores (Pestana & Gageiro, 2008). Uma variável em escala de rácio utilizada neste trabalho foi, por exemplo, o peso.

As variáveis de escala nominal ou ordinal são classificadas como variáveis qualitativas e as variáveis de escala de intervalo ou rácio são classificadas como quantitativas. As variáveis quantitativas ainda podem ser classificadas em discretas ou contínuas, consoante assumam um número, respetivamente, finito ou infinito de pontos, num intervalo em estudo (Pestana & Gageiro, 2008).

A definição do tipo de variáveis em estudo é essencial para a escolha das técnicas inferenciais e não inferenciais mais adequadas. Na estimação pontual, a partir da amostra, procura-se estimar um único valor de um parâmetro da população, como a média, por exemplo. A escolha do melhor estimador é feita recorrendo a procedimentos como o método dos mínimos quadrados, o método dos momentos ou o método da máxima verosimilhança, que variam consoante o tipo de variáveis. Neste trabalho, nos modelos de regressão logística, recorreu-se ao método da máxima verosimilhança para estimar os parâmetros (Oliveira, 2004).

Uma das grandes limitações dos métodos de estimação pontual deve-se à falta de informação sobre o rigor ou nível de confiança das estimativas que fornecem. A estimação por intervalos de confiança surge, então, como mais vantajosa, uma vez que, em vez de procurar um valor concreto para um determinado parâmetro, determina um intervalo de confiança para esse parâmetro desconhecido, com uma probabilidade $1-\alpha$ (grau de confiança) de que esse intervalo contenha o verdadeiro valor do parâmetro. Note-se que

os intervalos de confiança só podem ser construídos se a distribuição do estimador for conhecida. Nos casos em que a distribuição da variável em estudo é desconhecida podem-se recorrer, por exemplo, a intervalos de confiança empíricos. Os valores mais usuais para o grau de confiança são 99%, 95% e 90%, sendo α a probabilidade de ocorrer o erro tipo I, também designada por *p-value* (Oliveira, 2004). Na segunda parte deste trabalho todos os intervalos de confiança foram construídos com um nível de confiança de 95%, ou seja, $\alpha=.05$.

A definição do nível de significância é importante, pois ele é usado como critério nos testes de hipóteses. Os testes de hipóteses não são mais do que o estabelecimento de uma regra de decisão, determinando as regiões crítica e não crítica em função do nível de confiança escolhido, que irão ajudar na decisão de rejeitar ou não uma determinada hipótese estatística, com base nas observações da amostra. Na notação usual, designa-se por H_0 a hipótese nula, que é a hipótese estatística a ser testada, e por H_1 a hipótese alternativa. A hipótese nula é rejeitada sempre que o valor da estatística do teste pertencer à região crítica (Oliveira, 2004). Sempre que se afirma algo com 95% de confiança, há 5% de probabilidades de se estar a afirmar algo errado. Os riscos inerentes a uma tomada de decisão estão sempre associados a erros. Quando se toma a decisão de rejeitar a hipótese nula, quando ela é verdadeira, comete-se um erro do tipo I e quando se opta por não rejeitar a hipótese nula, quando ela é falsa, comete-se um erro do tipo II (Oliveira, 2004).

Os testes de hipóteses podem ser paramétricos ou não paramétricos e, regra geral, para cada teste paramétrico há sempre um homólogo não paramétrico. Os testes de hipóteses paramétricos apenas podem ser aplicados nos casos em que os valores populacionais seguem uma distribuição normal. Na prática nem sempre é possível ou razoável atingir este pressuposto, pois a amostra pode ter uma dimensão reduzida ou a população em estudo pode revelar uma franca assimetria ou os dados serem de natureza qualitativa (Oliveira, 2004). Nestes casos, os testes não paramétricos revelam-se como a melhor escolha, pois apesar de não incorporam as suposições restritivas dos testes paramétricos, como a normalidade da distribuição da população, muitas vezes permitem chegar a resultados tão exatos como os testes paramétricos, com a vantagem de que são mais facilmente implementados e mais perceptíveis (Oliveira, 2004). Na segunda parte deste

trabalho, apesar da amostra já ter um tamanho considerável ($n=600$), a distribuição da maior parte das variáveis contínuas não era a normal e, por outro lado, havia muitas variáveis qualitativas. Neste caso, optou-se por recorrer a testes não paramétricos e não forçar a aplicação de testes paramétricos, recorrendo ao teorema do limite central, que iria permitir afirmar que se se aumentasse a dimensão da amostra, a distribuição amostral da sua média aproximar-se-ia cada vez mais de uma distribuição normal. A título de curiosidade, foi feito um ensaio, comparando os resultados obtidos pelo teste U de Mann-Whitney com o *Independent Sample t Test*, no estudo de diferenças entre as variáveis contínuas, nos vários grupos de risco em estudo. A *performance* destes dois métodos foi comparada em 56 testes, sendo que em 54 (96.4%) as conclusões foram as mesmas.

1.2. Teste de ajustamento à normalidade de Kolmogorov-Smirnov

Um teste de ajustamento é usado como critério para verificar se um determinado Universo se comporta de acordo com uma determinada distribuição teórica, que se desconfia ajustar bem esse Universo (Oliveira, 2004). A distribuição normal é uma das mais importantes, visto ser pressuposto para a utilização de muitos testes estatísticos e permitir a aplicação de um grande número de estatísticas descritivas.

O teste de Kolmogorov-Smirnov, de ajustamento à normalidade, serve para analisar a aderência à normalidade de uma variável do tipo quantitativa ou qualitativa ordinal, recorrendo à comparação das frequências relativas acumuladas observadas com as esperadas, ou seja, considerando a diferença absoluta máxima entre a função de distribuição acumulada normal e a empírica dos dados e comparando essa diferença com um valor crítico, para que se obtenha um determinado nível de significância.

Considere-se a amostra aleatória simples x_1, x_2, \dots, x_n , de uma população com função de distribuição acumulada desconhecida, designada por F_x . A estatística de teste é

$$D_n = \sup_x |F(x) - F_n(x)| \quad (1)$$

Esta função corresponde à distância vertical máxima entre os gráficos de $F(x)$ e $F_n(x)$ sobre a amplitude dos possíveis valores de x . Em $D_n = \sup_x |F(x) - F_n(x)|$ (1), a função de distribuição acumulada normal está representada por $F(x)$ e a função de distribuição acumulada empírica dos dados está representada por $F_n(x)$. Para testar a hipótese nula ($H_0: F_x = F$), contra a hipótese alternativa ($H_1: F_x \neq F$), consideram-se as observações aleatórias da população, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$, ordenadas de forma crescente; a função de distribuição acumulada normal definida por $F(x_{(i)}) = P(X \leq x_i)$; a função de distribuição acumulada empírica definida por uma função em escada que é dada por

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{]-\infty, x]}(x_{(i)}). \quad (2)$$

Onde I_A é a função indicatriz, definida por

$$I_A = \begin{cases} 1 & ,se \quad x \in A \\ 0 & ,se \quad x \notin A \end{cases} \quad (3)$$

A função (2) corresponde à proporção de valores iguais ou inferiores a x e pode, por isso mesmo, ser descrita por

$$F_n(x) = \begin{cases} 0 & ,se \quad x < x_{(1)} \\ \frac{k}{n} & ,se \quad x_{(k)} \leq x < x_{(k+1)} \\ 1 & ,se \quad x > x_{(n)} \end{cases} \quad (4)$$

Sob H_0 , a distribuição assintótica da estatística de Kolmogorov-Smirnov obtém-se por

$$\lim_{n \rightarrow \infty} P[\sqrt{n}D_n \leq x] = 1 - 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2j^2 x^2) \quad (5)$$

Esta distribuição é válida nos casos em que se tem total conhecimento da distribuição de probabilidades de H_0 , caso contrário pode ser determinada por simulação.

Como a função de distribuição $F_n(x)$ é descontínua e a função de distribuição hipotética é contínua, podem-se considerar duas outras estatísticas para se calcular a estatística de Kolmogorov-Smirnov:

$$D^+ = \sup_{x_{(i)}} \left| F(x_{(i)}) - F_n(x_{(i)}) \right| \quad (6)$$

$$D^- = \sup_{x_{(i)}} \left| F(x_{(i)}) - F_n(x_{(i-1)}) \right| \quad (7)$$

Estas duas estatísticas medem as distâncias verticais entre os gráficos das duas funções nos pontos $x_{(i-1)}$ e $x_{(i)}$, permitindo o cálculo da seguinte estatística de teste

$$D_n = \max(D^+, D^-) \quad (8)$$

Se o valor obtido por (8) for maior do que o valor crítico, rejeita-se a hipótese de normalidade dos dados, com $(1-\alpha)100\%$ de confiança. Atualmente, com a evolução tecnológica, e a criação de programas específicos para tratamento estatístico de dados, estes testes são feitos em segundos. Os comandos para se obter o teste de aderência à normalidade de Kolmogorov-Smirnov, no SPSS, são » *Analyze» Descriptive Statistics» Explore» Plots» Normality plots with tests*. Estes comandos permitem obter uma tabela, com a correcção de Lilliefors, com a estatística de Kolmogorov-Smirnov, os graus de liberdade e a significância e dois gráficos Q-Q (normal e *detrended* normal) onde podem ser analisados os desvios em relação à normalidade, representada por uma reta (Figura 1).

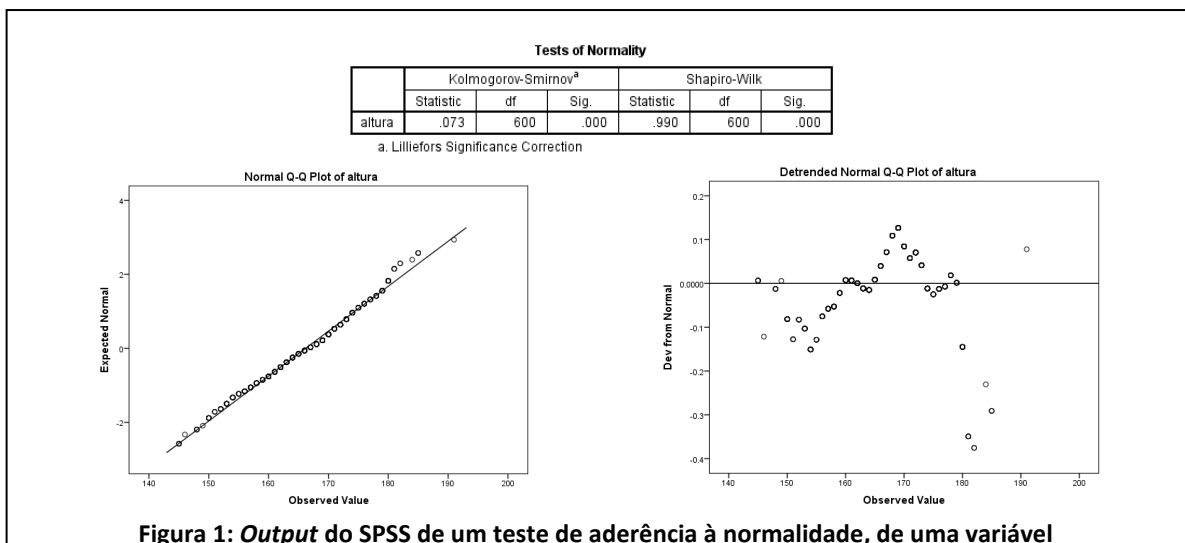


Figura 1: Output do SPSS de um teste de aderência à normalidade, de uma variável

Desta forma, mais do que saber se a distribuição é normal, ou não, é possível saber quais são os dados que se desviam mais da normalidade, que na prática correspondem aos pontos mais afastados de cada uma das retas.

A tabela *Tests of Normality* apresenta o teste de Kolmogorov-Smirnov com a correcção de Lilliefors e ainda outro teste de ajustamento desenvolvido especialmente para a lei normal, o teste de Shapiro-Wilk. Para amostras de dimensão superior ou igual a 30 aconselha-se o uso do teste de Kolmogorov-Smirnov com a correcção de Lilliefors, enquanto que para amostras de dimensão mais reduzida é mais indicado o teste de Shapiro-Wilk. No exemplo da Figura 1, a distribuição das estaturas não segue uma distribuição normal ($\alpha < .05$) e a observação dos gráficos mostra que os maiores desvios da normalidade ocorrem nos valores das estaturas mais elevadas.

Nesta dissertação, na parte II, o teste de Kolmogorov-Smirnov será usado tanto em variáveis quantitativas como qualitativas ordinais. No caso das variáveis ordinais, este teste será usado para verificar se os dados se distribuem igualmente por todas as categorias.

1.3. Teste da Binomial

Nesta dissertação há muitas variáveis. O teste de Kolmogorov-Smirnov apenas pode ser usado para variáveis de natureza no mínimo ordinal, deixando de fora as nominais. Para o caso das variáveis de natureza binomial é usual recorrer-se ao teste de ajustamento do Qui-quadrado para verificar se a distribuição dos dados pelas categorias foi semelhante ao esperado, no entanto, no caso particular em que a natureza da variável é dicotómica há um teste mais específico – o teste da binomial (Pestana & Gageiro, 2008). Como as variáveis nominais estudadas na segunda parte da dissertação são todas dicotómicas, optou-se pelo teste da binomial.

O teste da binomial compara as frequências observadas com aquelas que se espera obter numa distribuição binomial.

Seja $p = \frac{X}{n}$ a proporção de indivíduos com uma certa característica de interesse numa amostra aleatória de dimensão n e seja p_0 a proporção de indivíduos com essa característica na população. As hipóteses que se pretendem testar são: $H_0: p \geq p_0$ versus $H_1: p < p_0$.

Se $n \leq 30$, a estatística de teste será exata e dada por:

$$T = X \underset{\text{sob } H_0}{\sim} \text{Binomial}(n, p_0) \quad (9)$$

Se $n > 30$, a estatística de teste será aproximadamente dada por:

$$T = \frac{p - p_0}{\sqrt{p_0 \frac{1 - p_0}{n}}} = \frac{\frac{X}{n} - p_0}{\sqrt{p_0 \frac{1 - p_0}{n}}} \underset{\text{sob } H_0}{\sim} N(0,1) \quad (10)$$

A hipótese nula deve ser rejeitada se $p\text{-value} = p(Z > T) < (1 - \alpha)100\%$.

Este teste é executado de forma quase intuitiva em qualquer programa computacional. Na segunda parte desta dissertação será feito uma análise em que participaram 600 indivíduos, sendo que 355 eram homens e 245 eram mulheres. Será que a proporção de homens e de mulheres é significativamente a mesma?

As hipóteses a considerar são: $H_0: p = p(\text{homem}) = 0.5$ versus $H_1: p = p(\text{homem}) \neq 0.5$

Para testar estas hipóteses no SPSS usam-se os comandos: »Analyze» *Nonparametric Tests» Legacy Dialogs» Binomial*.

Para testar as mesmas hipóteses no R usam-se os comandos: *binom.test(355,600,0.5,alternative="greater")*.

```
> binom.test(355,600,0.5,alternative="greater")

Exact binomial test

data: 355 and 600
number of successes = 355, number of trials = 600, p-value = 4.066e-06
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
 0.5575437 1.0000000
sample estimates:
probability of success
 0.5916667
```

Figura 2: Output do R num teste da binomial

Neste exemplo (Figura 2) os resultados mostram, para um erro tipo I do analista de .05, que o nível de significância exato do teste foi .000, rejeitando-se H_0 , e concluindo-se que a proporção de homens (.59) é significativamente superior à de mulheres.

1.4. Teste de independência de Mann-Whitney

Numa análise bivariada estuda-se a relação entre duas variáveis. O leque de testes disponíveis é vasto e a escolha do mais adequado deve ser orientada pela natureza e a pela distribuição das variáveis.

Quando se pretende comparar medidas de tendência central, de uma variável quantitativa, em dois grupos diferentes de sujeitos as duas alternativas são o teste paramétrico t, para duas amostras independentes, ou o seu homólogo não paramétrico: o teste de independência de Mann-Whitney. Enquanto o teste t compara as médias das duas populações, com base em duas amostras, o teste de Mann-Whitney compara o centro de localização, como forma de se detetarem diferenças significativas entre as duas populações.

Como teste paramétrico, o teste t apenas pode ser aplicado após verificação da normalidade da distribuição da variável, pelo teste de Kolmogorov-Smirnov, dispensável no caso de a dimensão da amostra ser superior a 30. Há por vezes casos, que serão explorados na segunda parte da dissertação, em que a dimensão da amostra é grande, mas o teste de Kolmogorov-Smirnov revela que não segue a distribuição normal. O que fazer nestes casos? Escolher um teste não paramétrico ou forçar a normalidade, alegando o já referido

teorema do limite central? Pestana & Gageiro (2008) defendem que o teste de Mann-Whitney é de utilização preferível ao teste t, quando há violação da normalidade, pois já foi demonstrado que a perda de eficiência por aplicação do teste de Mann-Whitney é pequena. Este teste, em vez de se basear em parâmetros da distribuição normal, como a média e a variância, baseia-se nas ordenações da variável.

Considerem-se duas amostras independentes com dimensões n_1 e n_2 . O que se pretende é averiguar se os dois grupos independentes de observações foram, ou não, extraídos da mesma população. As hipóteses do teste de Mann-Whitney são

H0: As duas populações são iguais em tendência central.

H1: As duas populações não são iguais em tendência central.

Sejam X_1, X_2, \dots, X_{n_1} e Y_1, Y_2, \dots, Y_{n_2} as duas amostras aleatórias independentes, selecionadas das populações X e Y , respetivamente, admitindo-se que $n_1 \leq n_2$ e que $n = n_1 + n_2$. Considere-se, ainda, que a variável é de natureza contínua e que as observações se traduzem numa escala ou ordem.

Seja $R_x(i)$ a ordem atribuída à i -ésima observação de X e $R_y(j)$ a ordem da j -ésima observação de Y , para $i = 1, \dots, n_1$ e $j = 1, \dots, n_2$. A estatística de teste a considerar será a soma das ordens das observações provenientes da população X , tal que

$$T = \sum_{i=1}^{n_1} R_x(i) \quad (11)$$

A distribuição de probabilidade desta estatística obtém-se pressupondo que as duas populações têm a mesma distribuição e que se conhecem os valores de n_1 e de n_2 .

Para $n_1 \leq 10$ e $n_2 \leq 10$, calcula-se

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - T_1 \quad (12)$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - T_2 \quad (13)$$

Sendo a estatística de teste dada por

$$U = \min(U_1, U_2) \quad (14)$$

O valor de U deve ser comparado com o valor tabelado e se $U < U_c$, então pertence à região crítica, pelo que se deve rejeitar a hipótese nula, a um nível de significância α .

Para $n_1 > 10$ e $n_2 > 10$, é possível efetuar uma aproximação à distribuição normal. Caso a hipótese nula seja válida, pode provar-se que, para grandes amostras, a variável T tem o seguinte comportamento

$$T \approx N(\mu_T, \sigma_T), \text{ com } \mu_T = E[T] = \frac{n_1 n_2}{2} \text{ e } \sigma_T^2 = \frac{n_1 n_2 (n+1)}{12}.$$

Ou seja,

$$Z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n+1)}{12}}} \approx N(0,1) \quad (15)$$

Os valores críticos serão os da tabela da distribuição normal reduzida. Quando a probabilidade associada ao valor z observado for inferior a α , rejeita-se a hipótese nula.

No caso de ocorrerem empates entre duas observações da mesma amostra, o valor de T não é afetado, mas se os empates acontecerem entre duas ou mais observações, envolvendo ambas as amostras, então o valor de T poderá ser afetado e deverá ser feita uma correção no desvio-padrão, tal que

$$Z = \frac{T - \mu_T}{\sigma_T} = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2}{n(n-1)} \left(\frac{n^3 - n}{12} - \sum_{k=1}^q \frac{e_k^3 - e_k}{12} \right)}} \approx N(0,1) \quad (16)$$

Em que e_k representa o número de observações empatadas para um dado posto k e q representa o número de empates.

Os comandos necessários para se obter o teste de Mann-Whitney no SPSS são « *Analyze* » *Nonparametric Tests*» *Legacy Dialogs*» *2 Independent Samples*» *Mann-Whitney*. No R, usa-se a função $>wilcox.test(x\sim y)$, sendo x uma variável contínua e y uma variável dicotómica.

Coefficiente de Correlação Eta

O teste de Mann-Whitney indica apenas se há relação ou não entre as variáveis, mas não mede o grau da associação entre elas. Para medir o grau de associação entre variáveis recorre-se ao cálculo de coeficientes de correlação ou associação. Para medir a correlação entre uma variável contínua e uma dicotómica, recorre-se ao cálculo do coeficiente Eta. Este coeficiente assume valores entre 0 e 1, sendo que valores próximos de 0 indicam uma fraca associação e próximos de 1 indicam uma forte associação entre as variáveis.

O estimador do coeficiente de correlação Eta é a raiz quadrada de

$$\eta_{y,x}^2 = \frac{\text{soma de quadrados entre grupos}}{\text{soma total de quadrados}} \quad (17)$$

O erro padrão de (17) é dado por

$$\sigma_{\eta_{y,x}^2} = \frac{1 - \eta_{y,x}^2}{n - k} \quad (18)$$

Onde: $\sigma_{\eta_{y,x}^2}$ é o erro padrão do quadrado do coeficiente Eta; $\eta_{y,x}^2$ é o quadrado do coeficiente Eta; n é o número de observações da amostra; k é o número de categorias da variável nominal.

O quadrado do Eta é interpretado como a proporção da variação na variável dependente que é explicada pela variável independente, permitindo, na segunda parte deste trabalho, afirmações como “31.8% da variação da estatura dos benguelenses é explicada pelo género”, revelando uma associação moderada entre a estatura e o género.

Existem muitos critérios para classificar o grau de correlação, no entanto, neste trabalho, serão considerados os critérios de Cohen & Holliday (1982), apresentados na Tabela 1.

Tabela 1: Valores de referência para medição do grau de associação de variáveis, segundo Cohen e Holliday (1982)

Valores de Eta	Grau de correlação
$\text{Eta} \leq 0.19$	Correlação muito fraca
$0.20 \leq \text{Eta} \leq 0.39$	Correlação fraca
$0.40 \leq \text{Eta} \leq 0.69$	Correlação moderada
$0.70 \leq \text{Eta} \leq 0.89$	Correlação forte
$0.90 \leq \text{Eta} \leq 1.00$	Correlação muito elevada

No SPSS, o valor do Eta obtém-se através dos comandos » *Analyze*» *Descriptive Statistics*» *Crosstabs*» *Statistics*» *Eta*.

1.5. Teste de independência de Kolmogorov-Smirnov

Para estudar a independência entre duas variáveis qualitativas, o teste de independência do Qui-quadrado é o mais generalista, pois é aplicável tanto em variáveis ordinais como nominais. Este teste pressupõe que nenhuma célula da tabela tenha frequência esperada inferior a 1 e que não mais do que 20% das células da tabela tenham frequência esperada inferior a 5 unidades. Se estes pressupostos não forem garantidos, o nível de significância observado tanto pode ser correto como enganador e, em alternativa, deve ser usado o teste exato de Fisher (Pestana & Gageiro, 2008).

Consoante a variável seja ordinal, nominal ou até no caso particular de ser dicotómica, existem outras alternativas para ultrapassar estes constrangimentos. Neste trabalho, interessa referir duas delas: o teste de independência de Kolmogorov-Smirnov, quando uma das variáveis é ordinal e a outra dicotómica; o *Odds Ratio*, quando as duas variáveis são dicotómicas e uma antecede a outra.

O teste de independência de Kolmogorov-Smirnov analisa a distribuição de uma variável ordinal nas duas categorias de uma variável dicotómica. Na parte II desta dissertação, este

teste foi usado para comparar a distribuição do risco cardiovascular entre fumadores e não fumadores, por exemplo.

As hipóteses do teste de Kolmogorov-Smirnov são:

H0: As variáveis são independentes, isto é, a variável dicotómica não é distinta em relação à variável ordinal.

H1: As variáveis não são independentes, isto é, a variável dicotómica é distinta em relação à variável ordinal.

Quando se verifica a hipótese nula é possível afirmar que as amostras aleatórias independentes foram extraídas de populações idênticas, esperando-se, portanto, que ambas as distribuições de frequências acumuladas sejam também semelhantes.

Procedendo de forma análoga à do teste de Kolmogorov-Smirnov já descrito anteriormente, considerem-se duas amostras aleatórias independentes, X_1, X_2, \dots, X_{n_1} e Y_1, Y_2, \dots, Y_{n_2} , selecionadas das populações X e Y . Para testar a hipótese nula ($H_0: F_x = F_y$), contra a hipótese alternativa ($H_1: F_x \neq F_y$), a estatística de teste será:

$$D_{n_1, n_2} = \max |F_{n_1}(x) - F_{n_2}(x)| \quad (19)$$

Esta função corresponde à distância vertical máxima entre os gráficos de $F_{n_1}(x)$ e $F_{n_2}(x)$ sobre a amplitude dos possíveis valores de x . Em (17), a função de distribuição acumulada empírica dos dados da amostra extraída da população X é representada por $F_{n_1}(x)$ e a função de distribuição acumulada empírica dos dados da amostra extraída da população Y é representada por $F_{n_2}(x)$.

A decisão sobre a rejeição ou não da hipótese nula é tomada com base no valor crítico da estatística Z , com distribuição normal reduzida. Quando o valor obtido em (19) é significativamente grande, rejeita-se a hipótese nula, com $(1 - \alpha)100\%$ de confiança.

Em termos de aplicação deste teste no SPSS, os comandos a seguir são bastante semelhantes aos do teste de Mann-Whitney » *Analyze* » *Nonparametric Tests*» *Legacy Dialogs*» *2 Independent Samples*» *Kolmogorov-Smirnov Z*.

Coeficiente de contingência V de Cramer

Para medir o grau de associação entre uma variável ordinal e uma nominal pode-se recorrer ao cálculo do coeficiente de contingência V de Cramer, que, tal como o coeficiente Eta, assume valores entre 0 e 1. Uma das razões para a escolha deste coeficiente foi o facto de ser válido para tabelas com qualquer número de linhas e de colunas. O seu valor é baseado na estatística do Qui-quadrado (χ^2), tal que

$$V = \sqrt{\frac{\chi^2}{n \times (q-1)}} \quad (20)$$

Onde n representa a dimensão da amostra, q representa o menor número de linhas ou de colunas da tabela de contingência e $\chi^2 = \sum \frac{(o-e)^2}{e}$, sendo o a frequência observada e e a frequência esperada.

No SPSS, o valor do coeficiente de contingência V de Cramer obtém-se através dos comandos » *Analyze*» *Descriptive Statistics*» *Crosstabs*» *Statistics*» *Phi and Cramer's V*.

Coeficiente de correlação ρ de Spearman

O coeficiente ρ de Spearman é usado para medir a intensidade da relação entre variáveis no mínimo ordinais e, em vez do valor observado, usa apenas a ordem das observações. Deste modo, este coeficiente não é sensível a assimetrias na distribuição, nem à presença de *outliers*, não exigindo portanto que os dados provenham de duas populações normais. O coeficiente é calculado recorrendo a

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}, \quad (21)$$

Em que n é o número de pares (x_i, y_i) e $d_i = (\text{postos de } x_i \text{ dentre os valores de } x) - (\text{postos de } y_i \text{ dentre os valores de } y)$. Se os postos de x forem exatamente iguais aos pontos de y , então todos os d_i serão zero e ρ será 1, revelando uma correlação perfeita.

O coeficiente ρ de Spearman varia entre -1 e 1. Quanto mais próximo estiver destes extremos, maior será a associação entre as variáveis. O sinal negativo da correlação

significa que as variáveis variam em sentido contrário, isto é, as categorias mais elevadas de uma variável estão associadas a categorias mais baixas da outra variável.

Os valores de referência considerados para a medição do grau de associação foram os usados para o coeficiente Eta, da Tabela 1, mas considerado o valor absoluto do coeficiente.

No SPSS, o valor do coeficiente de correlação de Spearman obtém-se através dos comandos » *Analyze*» *Correlate*» *Bivariate*» *Spearman*.

1.6. Odds Ratio (OR)

Como foi referido no ponto anterior, o *Odds Ratio*, que será definido mais adiante, é usado para estudar a relação de independência entre duas variáveis nominais, sendo, atualmente, um dos preferidos dos investigadores, para tabelas 2x2, por ser mais facilmente interpretável do que o teste do Qui-quadrado (Pestana & Gageiro, 2008).

O *Odds Ratio* mede a associação entre duas variáveis nominais, sendo uma delas designada por fator e a outra por acontecimento e considerando que o fator é de ocorrência anterior ao acontecimento. Na segunda parte do trabalho usou-se o cálculo do *ratio* de produtos cruzados para medir a associação entre o género e a condição de hipertensão (hipertenso/normotenso), por exemplo.

As hipóteses para aplicação de *Odds Ratio* são:

H0: As variáveis são independentes, ou seja, os rácios são iguais a 1.

H1: Existe relação de dependência entre as variáveis, ou seja, os rácios são diferentes de 1.

Para melhor compreensão desta medida de associação será explorado um exemplo prático, utilizando dados da segunda parte desta dissertação. Pretendeu-se analisar a relação entre o género (fator) e a hipertensão (acontecimento), numa amostra de 600 indivíduos. As categorias do fator foram: 0=homem; 1=mulher. As categorias da hipertensão foram: 0=tem hipertensão; 1=não tem hipertensão.

Os resultados foram organizados na Tabela 2.

Tabela 2: Tabela de contingência (2x2) de apoio ao cálculo do Odds Ratio

	Homem	Mulher	Total marginal
Tem hipertensão	a=130	b=44	a+b=174
Não tem hipertensão	c=225	d=201	c+d=201
Total marginal	a+c=355	b+d=245	n=600

O fator de interesse deve ser sempre indicado na primeira coluna e neste caso foi “ser homem”. A tabela permite tirar muitas conclusões. Algumas são:

O *odds* de ser homem, para quem tem hipertensão, é igual a $\frac{a}{b} = \frac{130}{44} = 2.955$, ou seja, há

2955 hipertensos por cada 1000 normotensos. O *odds* de ser homem, para quem não tem hipertensão, é igual a $\frac{c}{d} = \frac{225}{201} = 1.119$, ou seja, há 1119 hipertensos por cada 1000 normotensos.

A probabilidade de um indivíduo hipertenso ser mulher é $p = \frac{b}{a+b} = \frac{44}{174} = 0.253 = 25.3\%$, ou seja, 25.3% das pessoas hipertensas são mulheres e

7.7% são homens. De forma análoga, conclui-se que 47.2% dos normotensos são mulheres e 52.8% são homens.

O *Odds Ratio* permite analisar os indivíduos com e sem hipertensão, em função de serem homens ou mulheres, sendo dado de uma forma geral por

$$OR = \frac{\frac{a}{a+b} / \frac{b}{a+b}}{\frac{c}{c+d} / \frac{d}{c+d}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \times d}{b \times c} \quad (22)$$

E, no exemplo referido anteriormente, por $OR = \frac{130 \times 201}{44 \times 225} = 2.64$, permitindo concluir que

um homem tem 3 vezes mais probabilidade de ser hipertenso do que uma mulher.

2. MODELAÇÃO ESTATÍSTICA

2.1. Introdução

Genericamente, um modelo é uma tentativa de representação simplificada de algo e a modelação matemática não é mais do que a construção de um conjunto de ferramentas matemáticas que auxiliem essa representação. O início do recurso a modelos matemáticos é usualmente creditado a Galileu Galilei, que se impôs contra a filosofia de Aristóteles, que apenas fazia especulações baseadas em raciocínios, e chamou a atenção para a necessidade de se experimentarem, observarem, medirem, pesarem, calcularem e expressarem essas relações através de fórmulas matemáticas. Desde então, um problema transversal às mais variadas áreas científicas é tentar expressar uma ou mais variáveis (variáveis dependente – y) como função de um conjunto de outras variáveis (variáveis explicativas – x), através de um modelo matemático do tipo $y = g(x)$. Este tipo de modelação chama-se regressão e permite perceber como é que determinadas variáveis influenciam outra variável.

As principais finalidades dos modelos de regressão são: predição; seleção de variáveis; estimação de parâmetros; inferência.

Quando o objetivo é fazer uma predição espera-se que grande parte da variação da variável resposta seja explicada pelas variáveis explicativas, podendo-se, inclusive, utilizar o modelo para obter valores da variável resposta, correspondentes a valores das variáveis explicativas que não fazem parte da base de dados. Normalmente, usam-se valores de variáveis explicativas contidas no intervalo de variação estudado, pois a utilização de valores fora desse intervalo é conhecida por extrapolação e deve ser usada com muito cuidado, pois, o modelo pode não ter um bom ajuste fora do intervalo estudado. Uma aplicação extremamente importante dos modelos de regressão é a predição de um evento futuro que esteja associado, por exemplo, a um determinado grau de risco, através do desenvolvimento de modelos preditivos bastante robustos, rigorosos e abrangentes que

resultam da análise exaustiva de um conjunto de dados atuais e históricos, graças à utilização de Bancos de Dados e *Data Mining*.

Geralmente não se sabe quais são as variáveis que afetam significativamente a variação da variável resposta, então, uma alternativa é estudar um grande número de variáveis. A análise de regressão pode, também, auxiliar no processo de seleção de variáveis, eliminando aquelas cuja contribuição não é significativa.

Dado um modelo e um conjunto de dados referente às variáveis resposta e explicativas, é possível, ainda, estimar parâmetros ou ajustar um modelo aos dados, para obter valores ou estimativas para esses parâmetros, tendo por base o modelo e os dados observados. Mais do que estimar os parâmetros, o ajuste de um modelo de regressão permite realizar inferências sobre eles, tais como testes de hipóteses e intervalos de confiança.

Por vezes, um modelo de regressão é afetado por características perturbadoras, muitas delas não mensuráveis, que dão origem a erro de pesquisa (ε) que, por ser imprevisível, associa aleatoriedade à resposta, e obriga a um tratamento probabilístico, que faz com que estes modelos matemáticos sejam designados, de forma mais específica, por modelos estatísticos. Desta forma, a relação entre as variáveis passa a ter a forma $y = g(x) + \varepsilon$, pressupondo-se um efeito conjunto aditivo, que faz com que seja impossível estudar o efeito das variáveis explicativas na variável resposta, de forma isolada.

A escolha do modelo mais adequado depende de muitos fatores, particularmente do modo como as variáveis dependente e explicativas se relacionam, da natureza das variáveis explicativas e de um conjunto de pressupostos sobre o erro. Esta fase é bastante importante e decisiva e, por isso, convém ter-se os objetivos da pesquisa bem definidos e o máximo de informações sobre o problema em si, pois quanto melhor for a adequação do modelo escolhido, melhor será a qualidade da análise do estudo. Encarar a Modelação Estatística como um processo iterativo (Figura 3) pode revelar-se uma boa alternativa.

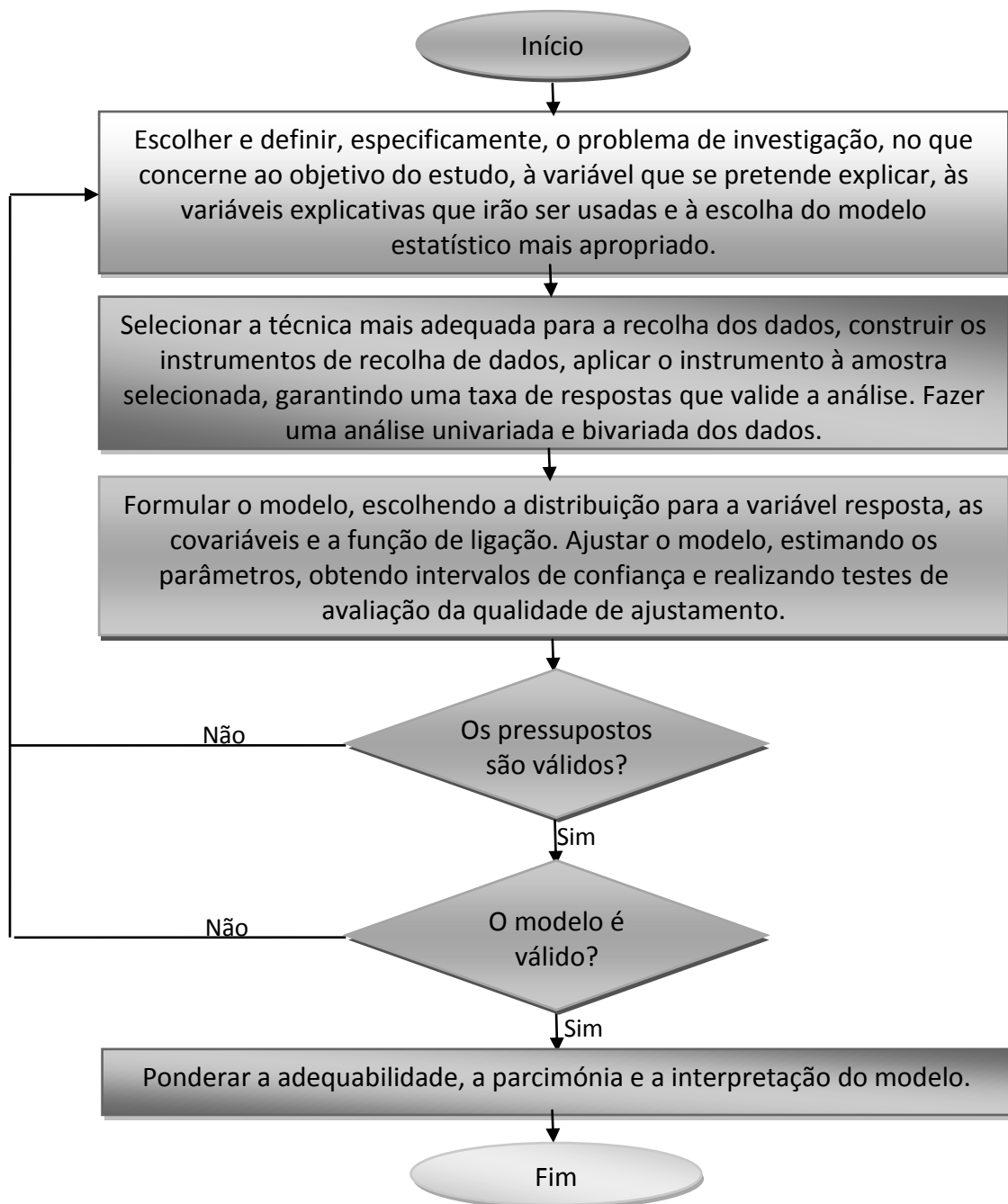


Figura 3: Fluxograma representativo da Modelação Estatística, considerada como um processo iterativo

Convém sempre ter em atenção que não há um modelo único para representar um determinado sistema e, como tal, deve ter-se sempre em conta o princípio da parcimónia aquando da modelação estatística, escolhendo-se sempre o modelo mais simples, quer em termos de número de variáveis explicativas, quer de relação funcional entre variáveis resposta e explicativas e quer em termos de distribuição de probabilidades associadas às

variáveis do modelo, para se evitarem previsões de relações espúrias e desajustadas da realidade.

2.2. Modelos de Regressão Linear Generalizados

2.2.1. Introdução

Em estatística os modelos mais simples e bem estudados são aqueles em que $g(x)$ assume a forma de uma função linear, denominado por modelo de regressão linear, que engloba a regressão linear simples (uma variável explicativa) e a regressão linear múltipla (mais do que uma variável explicativa). Este tipo de modelos, criado no início do século XIX por Legendre e Gauss, dominou a modelação estatística até meados do século XX, embora vários modelos não lineares ou não normais tenham sido entretanto desenvolvidos para fazer face a situações que não eram adequadamente explicadas pelo modelo linear, com distribuição normal.

Numa análise de regressão, a variável resposta pode ser influenciada por variáveis explicativas tanto de natureza quantitativa como qualitativa. As variáveis quantitativas são facilmente mensuradas numa escala, mas o mesmo não acontece com as variáveis qualitativas, uma vez que muitas vezes apenas indicam a presença ou a ausência de uma qualidade ou atributo. Desta forma, um método para interpretar esses atributos é construir variáveis artificiais que assumem os valores 0 ou 1, conhecidas por variáveis *dummy* e que representam a ausência e a presença de um determinado atributo. A introdução destas variáveis tornou o modelo de regressão linear numa ferramenta extremamente flexível, capaz de lidar com muitos problemas encontrados até então, principalmente, em estudos empíricos. A literatura especializada referente à abordagem da análise de regressão sobre variáveis *dummy* desenvolveu-se, principalmente, a partir das décadas de 70 e 80 do século passado, embora já tenha sido objeto de estudos há muitos anos.

Em 1972, Nelder e Wedderburn, introduziram uma generalização desses modelos lineares, designados por modelos lineares generalizados, que passou a admitir a hipótese de a

variável resposta ser da família exponencial, abrindo o leque a uma vasta classe de distribuições e introduzindo uma maior flexibilidade entre a variável explicativa e a média da variável resposta (Silva, 1995).

Um modelo linear generalizado é definido por três componentes: uma componente aleatória, na qual é identificada a distribuição de probabilidade da variável dependente; uma componente sistemática, que especifica uma função linear entre as variáveis explicativas; uma função de ligação que descreve a relação matemática entre a componente sistemática e o valor esperado da componente aleatória.

A componente aleatória diz respeito às observações da variável aleatória, Y , que esteve na origem do interesse pela procura de um modelo e sobre a qual foram recolhidas n observações independentes, ou seja, é um vetor do tipo $Y = (y_1, \dots, y_n)^T$, com uma distribuição da família exponencial e com médias $\mu_1, \mu_2, \dots, \mu_n$. A escolha do tipo de distribuição depende, habitualmente, da natureza dos dados, da assimetria e do seu intervalo de variação.

A componente sistemática é definida através de um vetor $\eta = (\eta_1, \dots, \eta_n)$ que está associado ao conjunto das variáveis explicativas através de um modelo linear do tipo

$$\eta = BX, \quad (23)$$

em que $X = (x_1, \dots, x_r)^T$ é uma matriz composta pelas variáveis explicativas das n observações, que se supõem explicar parte da variabilidade de Y , e $B = (\beta_1, \dots, \beta_r)$ é um vetor formado pelos parâmetros do modelo. Ou seja, havendo r variáveis explicativas e n observações a expressão (22) passa a ser

$$\eta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_r x_{ir}, \quad \forall i \in \{1, 2, \dots, n\} \quad (24)$$

O objetivo da modelação é determinar estimadores para os parâmetros β que produzam um bom ajustamento “à generalidade” das n observações existentes, para efetuar o ajustamento do modelo com robustez suficiente para não ser afetado significativamente por algumas observações atípicas.

A última componente é a função de ligação entre a componente aleatória e a componente sistemática. Seja $\mu_i = E(Y_i | x_i)$, com $i \in \{1, \dots, n\}$, então η_i é definida por $\eta_i = g(\mu_i)$, onde g é uma função monótona e diferenciável. Assim, a função de ligação relaciona os valores esperados das observações com as variáveis explicativas, para $i \in \{1, \dots, n\}$, tal que

$$g(\mu_i) = \sum_{j=1}^r \beta_j x_{ij} \quad (25)$$

A formulação do modelo implica, então, que se defina adequadamente a família da distribuição para a variável resposta em função dos dados disponíveis, se selecionem as covariáveis, se descreva adequadamente a matriz de especificação tendo em conta a natureza das covariáveis e se selecione uma função de ligação compatível com a distribuição do erro proposto para os dados, que satisfaça as propriedades estatísticas pretendidas para o modelo.

Para além do modelo de regressão linear clássico, que admite a distribuição normal, destacam-se, entre outras, duas classes importantes nos modelos lineares generalizados. Uma delas é formada pelos modelos *logit*, nos quais a variável dependente é uma variável aleatória de Bernoulli e pelos modelos *loglinear*, nos quais a variável dependente é associada a uma variável aleatória de Poisson. Na Tabela 3 podem ser consultadas algumas distribuições da família exponencial.

O presente estudo explora os modelos logísticos, tendo em conta o caráter categórico da variável dependente, para as quais este tipo de modelos se têm revelado a melhor alternativa, pois além de serem muito poderosos e flexíveis são de fácil interpretação, quando comparados com outros como os modelos *probit* (Agresti, 2007). Muitas vezes, mesmo quando a variável dependente não é do tipo binário, alguns pesquisadores têm dicotomizado a variável resposta de modo a que a probabilidade de sucesso possa ser ajustada através da regressão logística, devido à facilidade de interpretação dos seus parâmetros e à possibilidade do uso deste tipo de metodologia em análise discriminante.

Tabela 3: Algumas distribuições da família exponencial

Distribuição	Normal	Binomial	Poisson	Gama
Notação	$N(\mu, \sigma^2)$	$B(m, \pi)/m$	$P(\lambda)$	$Ga\left(v, \frac{v}{\mu}\right)$
Suporte	$(-\infty, +\infty)$	$\left\{0, \frac{1}{m}, \dots, 1\right\}$	$\{0, 1, \dots\}$	$(0, +\infty)$
θ	μ	$\ln\left(\frac{\pi}{1-\pi}\right)$	$\ln \lambda$	$-\frac{1}{\mu}$
$a(\phi)$	σ^2	$\frac{1}{m}$	1	$\frac{1}{v}$
ϕ	σ^2	1	1	$\frac{1}{v}$
ω	1	m	1	1
$c(y, \phi)$	$-\frac{1}{2}\left(\frac{y^2}{\phi} + \ln(2\pi\phi)\right)$	$\ln\left(\frac{m}{my}\right)$	$-\ln y!$	$v \ln v - \ln \Gamma(v) + (v-1) \ln y$
$b(\theta)$	$\frac{\theta^2}{2}$	$\ln(1 + \exp(\theta))$	$\exp(\theta)$	$-\ln(-\theta)$
$b'(\theta) = E(Y)$	θ	$\pi = \frac{\exp(\theta)}{1 + \exp(\theta)}$	$\lambda = \exp(\theta)$	$\mu = -\frac{1}{\theta}$
$b''(\theta) = V(\mu)$	1	$\pi(1-\pi)$	λ	μ^2
$\text{var}(Y)$	σ^2	$\frac{\pi}{1-\pi}$	λ	$\frac{\mu^2}{v}$

Fonte: Adaptada de Turkman & Silva(2000)

Nos casos em que a variável dependente possui apenas duas categorias, ou seja, é de natureza binária ou dicotómica recorre-se à regressão binária e nos casos em que a variável dependente possui mais do que duas categorias recorre-se à regressão multinomial, que pode ser nominal ou ordinal.

No presente trabalho serão explorados modelos logísticos ordinais, que são um caso particular da regressão ordinal, em que a função de ligação é a distribuição logística, no entanto, um bom entendimento dos modelos ordinais obriga a que se faça uma revisão dos modelos binários.

2.2.2. Modelo de Regressão Logística Binária

O modelo de regressão logística binária é um caso particular dos modelos *logit*, que é utilizado nos casos em que se procura ajustar um modelo para uma variável dependente binária, recorrendo a mais do que uma variável explicativa (covariável).

Formulação do modelo

Suponha-se que se tem n variáveis resposta independentes binomialmente distribuídas,

$Y_i = B(1, \pi_i)$, tais que

$$f(y_i | \pi_i) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad , y = 0, 1 \quad (26)$$

E que cada indivíduo i está associado a um vetor de especificação $X = (1, x_{i1}, \dots, x_{ir})^T$, que resultou do vetor de covariáveis x_i , $i = 1, \dots, n$, e onde π_i representa a probabilidade de sucesso.

Como $E(Y_i) = \pi_i$, de acordo com a Tabela 3,

$$g(X) = \theta_i = \ln\left(\frac{\pi_i}{1 - \pi_i}\right), \quad (27)$$

A função de ligação é linear, ou seja,

$$\theta_i = \eta_i = \beta X = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_r x_{ir} \quad (28)$$

Por (26) e por (27) conclui-se que a função de ligação será

$$g(X) = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_r x_{ir} \quad (29)$$

Constata-se que a função de ligação canónica é a função *logit*. Neste sentido, a probabilidade de sucesso ($P(Y_i = 1) = \pi_i$) fica relacionado com o vetor X por:

$$\pi_i = \frac{\exp(\beta X)}{1 + \exp(\beta X)} \quad (30)$$

Onde $\beta = (\beta_0, \beta_1, \dots, \beta_r)$ e $i = 1, \dots, n$.

Desta forma, a função $F : \mathbb{R} \rightarrow [0,1]$ fica definida por

$$F(x) = \frac{\exp(x)}{1 + \exp(x)} \quad (31)$$

constituindo a função de distribuição logística, que esteve na origem do nome atribuído a este modelo.

Estimativas dos parâmetros

Os parâmetros β são geralmente estimados pelo método da máxima verossimilhança e obtidos através da resolução, por processos iterativos, das equações

$$\begin{aligned} \sum_{i=1}^n y_i (1 + \exp(g(X_i))) - \sum_{i=1}^n m_i (1 + \exp(g(X_i))) &= 0 \\ \sum_{i=1}^n y_i x_i (1 + \exp(g(X_i))) - \sum_{i=1}^n m_i x_i (1 + \exp(g(X_i))) &= 0 \end{aligned} \quad (32)$$

Conseguidas as estimativas dos parâmetros do modelo, podem-se calcular os *Odds Ratio*

$$OR(\beta_j) = \exp(\beta_j) \quad (33)$$

E as probabilidades ajustadas, ou seja,

$$\pi_i = \frac{\exp(g(X_i))}{1 + \exp(g(X_i))} \quad (34)$$

Em que

$$g(X_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_r x_{ir} \quad (35)$$

Inferência no modelo

Com os parâmetros já estimados, é necessário testar as suas significâncias no modelo, recorrendo ao teste da razão de verossimilhança, ao teste de Wald ou ao teste *score*, considerando as seguintes hipóteses:

H0: $\beta_1 = \beta_2 = \dots = \beta_r = 0$, ou seja, o modelo não é estatisticamente significativo.

H1: $\exists \beta_j \neq 0, \quad j = 1, \dots, r$, ou seja, o modelo é estatisticamente significativo.

A estatística para o teste da razão de verossimilhança

$$G^2 = 2 \sum n_j \log(n_j / n\pi_{j0}) \quad (36)$$

Quanto maior for o valor desta estatística, mais evidências há contra H_0 . Para amostras com grandes dimensões, a estatística G^2 segue a distribuição de Qui-Quadrado, com $c-1$ graus de liberdade, sendo c o número de categorias.

No teste de Wald, a estatística de teste usada para avaliar o valor estimado de β_j é

$$W_j = \frac{\beta_j}{DP(\beta_j)} \quad (37)$$

Em que $DP(\beta_j)$ é o desvio-padrão do valor estimado do coeficiente β_j e equivale à raiz quadrada positiva da variância do mesmo coeficiente. O cálculo da variância de cada coeficiente é feito com base na teoria da máxima verossimilhança, recorrendo à matriz de covariâncias. A variância de β_j é o j -ésimo elemento da diagonal dessa matriz.

Intervalos de confiança

Um outro passo importante no ajustamento dos modelos é a obtenção de intervalos de confiança. O intervalo de confiança de $(1-\alpha)100\%$ para um parâmetro β_j é baseado no respetivo teste de Wald e é dado por

$$IC(\beta_j, 1-\alpha) = \left[\beta_j - z_{1-\frac{\alpha}{2}} DP(\beta_j) , \beta_j + z_{1-\frac{\alpha}{2}} DP(\beta_j) \right] \quad (38)$$

Seleção de Variáveis

Na altura de formular um modelo não se deve cair no erro de considerar todas as covariáveis possíveis. A inserção das covariáveis num modelo deve resultar de uma escolha ponderada, após análise cuidadosa dos dados, pois quantas mais covariáveis tiver um modelo, maior se torna a estimativa do erro e mais dependente dos dados observados fica o modelo. Por outro lado, variáveis consideradas importantes, no contexto do estudo, não devem ser deixadas de lado devido a não terem sido consideradas estatisticamente

significativas. Assim, a seleção de um modelo logístico deve ser um processo conjugado de domínio estatístico e muito bom senso.

Para se selecionarem as covariáveis mais relevantes para o modelo pode-se recorrer a métodos iterativos que se mostram como auxiliares preciosos, sobretudo quando o número de variáveis explicativas candidatas à inclusão no modelo é grande.

Destacam-se o método *standard*, em que todas as variáveis entram ao mesmo tempo no processo iterativo; o método *stepwise*, em que é adicionada uma variável de cada vez, sendo incluída ou não no modelo, de acordo com um critério pré-estabelecido; o método hierárquico, em que, contrariamente ao método *stepwise*, o total controlo da entrada das variáveis é do investigador.

Qualquer um destes três procedimentos é baseado num algoritmo que testa a importância das variáveis, incluindo ou excluindo-as do modelo baseando-se numa determinada regra de decisão. A importância da variável é definida em termos de uma medida de significância estatística, do coeficiente associado à variável para o modelo, e que depende das suposições de cada modelo.

Na altura de selecionar as variáveis também se podem ter em conta todos os modelos possíveis e compará-los entre si, de acordo com outros critérios. Na regressão logística é usual recorrer-se ao critério de informação de Akaike

$$AIC = -2\ln(L(\beta)) + 2r \quad (39)$$

Ou ao critério de informação Bayesiano

$$BIC = -2\ln(L(\beta)) + r\ln(n) \quad (40)$$

nos quais $\ln(L(\beta))$ é o logaritmo da verossimilhança do modelo e r o número de parâmetros do modelo.

Estes critérios não se baseiam em testes de hipótese, mas apenas na comparação das mesmas medidas, recorrendo aos logaritmos das máximas verossimilhanças. Neste sentido, o “melhor” modelo será o que tiver o valor mais baixo em cada um destes critérios. É claro que esta opção não deve ser encarada de uma forma muito rígida, mesmo

porque, segundo Long & Freese (2005) não há evidências, suficientemente convincentes, de que os modelos que maximizam os valores de uma dada medida de ajustamento são melhores do que outros que têm um valor mais elevado para essa medida.

Qualidade do ajuste do modelo

Como já foi referido anteriormente, qualquer modelo é uma tentativa de representar uma determinada realidade e procurar o modelo perfeito é uma utopia. A escolha do modelo final exige que o investigador seja uma espécie de equilibrista, capaz de conseguir encontrar a melhor combinação de variáveis explicativas, baseando-se em conhecimentos estatísticos, conhecimentos informáticos e, no caso da área de epidemiologia, de medicina.

É claro que associado ao acto de escolher, está sempre uma componente de erro e de incerteza, mas o modelo final deverá ser capaz de representar os dados, da melhor forma possível. Nesta fase deve ser avaliada a qualidade do ajustamento do(s) modelo(s), estudando os desvios (resíduos) entre os dados observados e os estimados pelo modelo de regressão.

Se estes desvios forem grandes, então o modelo não está bem ajustado, não representa bem os dados. Existem imensos métodos para se fazer esta avaliação. A qualidade do ajuste do modelo pode ser avaliada construindo um gráfico do modelo ajustado ou recorrendo ao cálculo de estatísticas, como a *Deviance*, o Qui-quadrado de Pearson ou a de Hosmer e Lemeshow.

Considere-se: J que representa o número distinto de valores que x assume na amostra; m_j que representa o número de indivíduos na amostra com valores iguais $x = x_j$; π_j que representa a probabilidade ajustada dos indivíduos em j ; y_j que representa o valor observado em j ; \hat{y}_j que representa o valor ajustado em j . A estatística *Deviance* é dada por

$$D = \sum_{j=1}^J d(y_j, \pi_j)^2 \quad (41)$$

Em que,

$$d(y_j, \pi_j) = \pm \sqrt{2 \left[y_j \ln \left(\frac{y_j}{m_j \pi_j} \right) + (m_j - y_j) \ln \left(\frac{m_j - y_j}{m_j \pi_j (1 - \pi_j)} \right) \right]} \quad (42)$$

O sinal positivo ou negativo a considerar deverá ser o mesmo sinal de $(y_j - m_j \pi_j)$. A estatística D, sob a suposição que o modelo ajustado é correto, tem distribuição assintótica χ^2 , com $J - (r + 1)$ graus de liberdade, sendo J????????????????

A estatística do Qui-quadrado de Pearson é dada por:

$$\chi^2 = \sum_{j=1}^J \left(\frac{(y_j - m_j \pi_j)}{\sqrt{m_j \pi_j (1 - \pi_j)}} \right)^2 \quad (43)$$

A estatística χ^2 , sob a suposição que o modelo ajustado é correto, tem distribuição assintótica χ^2 , com $J - (r + 1)$ graus de liberdade.

O recurso às duas estatísticas anteriores, nos casos em que $J \approx n$, sob n-assintótico, devolve p-valores incorretos. Hosmer & Lemeshow (1980) propuseram a ordenação dos casos de acordo com as probabilidades estimadas e a sua posterior divisão em g grupos de tamanho aproximadamente igual. A recomendação é de dez grupos. Para cada um dos grupos calcula-se a estatística de Hosmer e Lemeshow:

$$C = \sum_{k=1}^g \frac{(O_k - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)}, \quad (44)$$

Em que: n_k é o número de indivíduos no k-ésimo grupo; $\bar{\pi}_k = \sum_{j=1}^{C_k} \frac{m_j \pi_j}{n_k}$; C_k é o número total de combinações de níveis dentro do k-ésimo decil; $O_k = \sum_{j=1}^{C_k} y_j$ é o número total de respostas dentro do grupo k . A estatística do teste de Hosmer e Lemeshow tem distribuição Qui-quadrado, com $g - 2$ graus de liberdade.

Avaliação da capacidade discriminativa do modelo

Quando a variável resposta é binária é necessário escolher uma regra de predição ($Y = 0$ ou $Y = 1$). Tendo em conta que π varia entre 0 e 1, é intuitivo pensar que se o valor de π_i for grande, então $Y_i = 1$ e se o valor de π_i for pequeno, então $Y_i = 0$. Para determinar o ponto de corte que separa um evento do outro recorre-se muitas vezes à curva ROC (*Receiver Operating Curve*), na qual é representada a sensibilidade ($P(Y = 1 | Y = 1)$) e a especificidade ($1 - P(Y = 0 | Y = 0)$), para todos os possíveis pontos de corte entre 0 e 1.

A escolha do ponto de corte deve ser baseada numa combinação ótima tanto da sensibilidade quanto da especificidade, pois classificar o indivíduo como possuidor de um dos eventos, sem que ele o seja, ou não o classificar como possuidor, quando ele realmente o é, poderá trazer malefícios. Recorrendo, então, à análise da curva ROC, deve-se escolher o ponto de corte referente à melhor combinação entre sensibilidade e especificidade o que, na prática, equivale a escolher a combinação que mais se aproxima do canto superior esquerdo do gráfico.

Validação do modelo

Para validar o modelo recorre-se a uma nova amostra de dados para verificar se o modelo obtido pode ser satisfatoriamente ajustado a estes novos dados. Ou seja, é preciso determinar se os coeficientes e os erros-padrão obtidos a partir dos dados utilizados para elaboração do modelo são similares aos obtidos para os dados de validação.

2.2.3. Modelos de Regressão Logística Ordinal

A variável dependente ordenada, na forma de *score*, tem sido cada vez mais utilizada, fazendo com que os modelos de regressão ordinal estejam a ser aplicados cada vez mais, com sucesso, em variados estudos epidemiológicos, de qualidade de vida, indicadores de saúde e mesmo de gravidade e risco de doenças (Abreu et al, 2009). Estes modelos, dependendo do delineamento do estudo, permitem também calcular a estatística *odds ratio* ou prever a probabilidade de ocorrência de um evento (Ananth & Kleinbaum, 1997).

Formulação dos modelos

Na literatura são referidos quatro modelos ordinais: o modelo de *odds* proporcionais; o modelo de *odds* proporcionais parciais; o modelo de razão contínua; modelo estereótipo.

Entre os quatro modelos referidos, o modelo de *odds* proporcionais é o mais utilizado e, talvez por isso, também é o que está mais desenvolvido em termos de técnicas e até em termos de *software* de apoio. Estes modelos, quando comparados com os modelos de regressão linear para variáveis dependentes contínuas e até com os modelos logísticos binários, estão muito pouco explorados ainda. Apesar de, tal como os binários, fazerem parte dos modelos logísticos, a maior parte das técnicas utilizadas não podem ser as mesmas, uma vez que nos modelos ordinais se tem em conta a ordenação das categorias da variável dependente. O facto de haver carência no desenvolvimento de técnicas de análise para alguns destes modelos, leva a que muitas vezes os pressupostos sejam ignorados e se force a aplicação de modelos que geram estimativas de coeficientes desajustadas.

Na construção de qualquer um dos modelos ordinais, Hosmer & Lemeshow (2000) recomendam que se comece por fazer uma análise univariada, para seleção dos efeitos principais, para que seguidamente se incluam no modelo apenas as covariáveis significantes, de acordo com um nível de significância pré estabelecido. Posteriormente, o modelo deve ser ajustado, a sua adequação deve ser avaliada, recorrendo a testes adequados e gráficos de resíduos, e, finalmente, deve ser interpretado, através do cálculo de estimativas de *Odds Ratio*.

No modelo estudado anteriormente, a variável dependente era binária. A generalização dessa situação permite modelar variáveis com mais de duas categorias originando os modelos de regressão logística multinomial. Quando as categorias estão naturalmente ordenadas, como o grau de risco cardiovascular, estes modelos designam-se por modelos de regressão logística ordinal e baseiam-se no uso da probabilidade acumulada.

Seja Y a variável resposta, com k categorias codificadas em $1, 2, \dots, k$, e $X = (x_1, x_2, \dots, x_r)$ o vetor de variáveis explicativas. As k categorias de Y ocorrem com probabilidades

p_1, p_2, \dots, p_k , isto é, $p_j = P(Y = j)$, $j = 1, 2, \dots, k$. Considere-se, ainda, que α se refere ao intercepto do modelo e que β corresponde aos efeitos das covariáveis na variável dependente.

Modelo de Odds Proporcionais

O modelo de *odds* proporcionais é usado nos casos em que a suposição de *odds* proporcionais é válida e a variável dependente resultou de uma variável que originalmente era contínua.

Neste modelo são considerados $(k-1)$ pontos de corte das categorias, sendo que o j -ésimo ($j = 1, 2, \dots, k - 1$) ponto de corte é baseado na comparação de probabilidades acumuladas, ou seja,

$$\begin{aligned} \lambda_j(X) &= \ln \left(\frac{P(Y=1|X) + P(Y=2|X) + \dots + P(Y=j|X)}{P(Y=j+1|X) + P(Y=j+2|X) + \dots + P(Y=k|X)} \right) = \\ &= \ln \left(\frac{\sum_{j=1}^j P(Y=j|X)}{\sum_{j+1}^k P(Y=j|X)} \right) = \alpha_j + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_r x_r, \quad j = 1, 2, \dots, k - 1 \end{aligned} \quad (45)$$

Na qual o termo α_j varia para cada uma das k categorias e cada β não depende de índice i , levando a que a relação entre X e Y seja independente da categoria.

Como já foi referido, este modelo possui a suposição de *odds* proporcionais relativamente aos $(k-1)$ pontos, também designada por suposição de regressão paralela, que deve ser assumida por cada uma das variáveis explicativas incluídas no modelo. Para tal, o recurso a este modelo obriga a que esta suposição seja testada, separadamente, em cada covariável e no modelo final.

O modelo de *odds* proporcionais tanto pode ser estudado no SPSS como no R. No caso do SPSS o modelo das *odds* proporcionais obtido pelos comandos » *Analyze» Regression» Ordinal*. No *software* R, o modelo de *odds* proporcionais pode ser ajustado por meio da função *polr*, integrante do pacote *MASS*. Tanto num *software* como noutro, existe uma boa variedade de funções, que agilizam muito este tipo de modelação.

Modelo de Odds Proporcionais Parciais

O modelo de *odds* proporcionais parciais deve ser usado nos casos em que a suposição de *odds* proporcionais não é verificada, impossibilitando a aplicação do modelo referido anteriormente.

Este modelo permite que algumas covariáveis sejam modeladas com a suposição de *odds* proporcionais, mas para aquelas em que essa suposição não se verifica incrementa-se um coeficiente γ que representa o efeito associado a cada j -ésimo logit cumulativo, ajustado às restantes covariáveis.

A forma geral do modelo é quase a mesma do modelo de *odds* proporcionais, mas com os referidos coeficientes associados a cada categoria da variável resposta, ou seja,

$$\lambda_j(X) = \ln \left(\frac{P(Y=1|X)+P(Y=2|X)+\dots+P(Y=j|X)}{P(Y=j+1|X)+P(Y=j+2|X)+\dots+P(Y=k|X)} \right) = \ln \left(\frac{\sum_{j=1}^j P(Y=j|X)}{\sum_{j=1}^k P(Y=j|X)} \right) = \quad (46)$$

$$= \alpha_j + (\beta_1 + \gamma_{j1})x_1 + (\beta_2 + \gamma_{j2})x_2 + \dots + \beta_q x_q + \dots + \beta_r x_r, \quad j = 1, 2, \dots, k-1$$

Normalmente, espera-se que haja algum tipo de tendência linear entre cada *odds* ratio dos pontos específicos de corte e a variável dependente e para isso há restrições que podem ser incluídas no modelo para esclarecer essa linearidade, resultando em

$$\lambda_j(X) = \ln \left(\frac{P(Y=1|X)+P(Y=2|X)+\dots+P(Y=j|X)}{P(Y=j+1|X)+P(Y=j+2|X)+\dots+P(Y=k|X)} \right) = \ln \left(\frac{\sum_{j=1}^j P(Y=j|X)}{\sum_{j=1}^k P(Y=j|X)} \right) = \quad (47)$$

$$= \alpha_j + \tau_j \left((\beta_1 + \gamma_{j1})x_1 + \dots + \beta_q x_q \right) + \beta_{q+1} x_{q+1} + \dots + \beta_r x_r, \quad j = 1, 2, \dots, k-1$$

O modelo acima é conhecido por modelo de *odds* proporcionais parciais restrito e os parâmetros τ_j são escalares fixos que representam restrições alocadas aos parâmetros independentes dos pontos de corte. Até o momento, este modelo não se encontra disponível no *software* R nem no SPSS, fazendo com que esteja, ainda, pouco explorado em termos de literatura.

Modelo de Razão Contínua

O modelo de razão contínua deve ser usado nos casos em que há um interesse intrínseco numa categoria específica da variável dependente, pois a forma

$$\lambda_j(X) = \ln \left(\frac{P(Y = j | X)}{P(Y = j+1 | X) + \dots + P(Y = k | X)} \right) = \ln \left(\frac{P(Y = j | X)}{\sum_{j+1}^k P(Y = j | X)} \right) = \alpha_j + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jr}x_r, \quad j = 1, \dots, k \quad (48)$$

Permite comparar a probabilidade de uma resposta igual à categoria de um determinado *score* com a probabilidade de uma resposta maior. No caso do *score* y_j , permite comparar a probabilidade $Y = j$ com a probabilidade $Y > y_j$. Este modelo possui diferentes interceptos e coeficientes para cada comparação e pode ser ajustado por k modelos de regressão logística binária.

O modelo de razão contínua pode ser implementado no *software* R por meio de uma reestruturação dos dados que é feita com o comando *cr.setup*, integrante do pacote *Design*. Este modelo deve ser usado com cautela, pois esta bipartição leva a perda de informação e deve ser apenas usada nos casos em que o objetivo do estudo incide sobre uma determinada categoria da variável dependente.

Modelo Estereótipo

O modelo estereótipo, originalmente proposto por Anderson (1984) é considerado um meio-termo entre o modelo multinomial e o modelo de *odds* proporcionais. Segundo Hosmer & Lemeshow (2000), o modelo estereótipo é uma extensão do modelo de regressão multinomial, na qual cada categoria da variável dependente é comparada com uma categoria de referência, que normalmente é a primeira ou a última. Este modelo deve ser utilizado quando a variável dependente é ordinal, com $k+1$ categorias discretas, e nunca quando a variável resultou da categorização de uma variável contínua (Anderson, 1984).

Devido ao caráter ordinal dos dados, deve ser imposta uma estrutura linear aos β_{jl} ($j = 0, \dots, k$ e $l = 1, \dots, r$) e serem atribuídos pesos, ω_j , aos coeficientes. Estes pesos estão

diretamente relacionados com os efeitos das covariáveis e por isso o *odds ratio* formado terá uma tendência de crescimento, pois os pesos, normalmente, são construídos com ordenação, ou seja, $0 = \omega_0 \leq \omega_1 \leq \dots \leq \omega_k = 1$.

A forma funcional do modelo estereótipo é

$$P(y_i = s | x_{i1}, \dots, x_{ir}) = \frac{\exp\left(\beta_{0s} + \omega_s \sum_{j=1}^r x_{ij} \beta_{sj}\right)}{\sum_{t=1}^k \exp\left(\beta_{0t} + \omega_t \sum_{j=1}^r x_{ij} \beta_{tj}\right)} \quad (49)$$

A interpretação dos coeficientes, β , está intimamente relacionada com ω_j , pois os logaritmos dos *Odds Ratio* da forma funcional do modelo (49) são dados por $\omega_j \beta$.

Na segunda parte desta dissertação, tendo em conta a natureza original da variável dependente e o objetivo do estudo, o modelo estereótipo foi o eleito. A vantagem do recurso a este modelo é a possibilidade de estudar a variável dependente em toda a sua essência, tendo em conta a ordenação das categorias e tendo obrigação de selecionar uma delas como a mais importante. As grandes desvantagens deste modelo, e que fazem com que o recurso a este modelo seja raro, são a sua complexidade matemática, atribuída ao produto dos parâmetros no preditor linear, e a falta de *software* de apoio.

Ajuste do modelo

O ajuste de um modelo ordinal é feito iterativamente recorrendo ao método da máxima verossimilhança. Uma vez que o modelo pressupõe que as curvas de probabilidade das $k - 1$ categorias da variável dependente sejam iguais para todas as classes e sejam calculadas de forma cumulativa, a função de verossimilhança é (Agresti, 2002)

$$L = \prod_{i=1}^n \left[\prod_{k=1}^K \left(P(Y \leq k | x_i) - P(Y \leq k-1 | x_i) \right)^{Y_{ki}} \right] =$$

$$= \prod_{i=1}^n \left[\prod_{k=1}^K \left(\frac{\exp(\alpha_k - \beta x_i)}{1 + \exp(\alpha_k - \beta x_i)} - \left(\frac{\exp(\alpha_{k-1} - \beta x_i)}{1 + \exp(\alpha_{k-1} - \beta x_i)} \right)^{Y_{ki}} \right) \right] \quad (50)$$

Em que Y_{ki} representa a classe dicotómica da variável Y para o indivíduo i .

No caso dos modelos estereótipos, com devido ao produto $\omega_j\beta$, o uso da função de verossimilhança está longe de ser algo trivial. Anderson (1984) propôs que, no caso dos modelos estereótipos, a função fosse diretamente estimada por

$$L(x_i, \alpha_j, \omega_j, \beta_i) = \prod_{i=1}^N \prod_{j=1}^J P_j(x_i)^{I(Y=j)} \quad (51)$$

Como em qualquer tipo de análise de regressão é importante avaliar a qualidade do ajuste dos modelos, pois a falta de ajuste pode, por exemplo, levar ao enviesamento da estimação dos efeitos. Assim, avaliação do ajuste é bastante útil no sentido em que permite detetar: covariáveis importantes; interações omitidas; funções de ligação desapropriadas; formas funcionais incorretas, usadas na modelação de covariáveis; violação na suposição de *odds ratio* proporcional (Brant, 1990).

Embora tenham sido desenvolvidos muitos métodos para avaliar a qualidade do ajuste da regressão logística binária, poucos desses métodos podem ser estendidos para a maior parte dos modelos ordinais (Hosmer & Lemeshow, 2000). Normalmente, a qualidade do ajuste dos modelos ordinais é verificada pelos testes de *Pearson* ou *Deviance* ou recorrendo à criação de uma tabela de contingência na qual as linhas consistem de todas as possíveis configurações das covariáveis do modelo e as colunas são as categorias da variável dependente ordinal (Pukstenis & Robinson, 2004). As contagens esperadas dessa tabela são expressas por

$$E_{lj} = \sum_{l=1}^{N_l} p_{lj}, \quad (52)$$

Onde N_l é o número total de indivíduos classificados na linha l e p_{lj} representa a probabilidade de um indivíduo na linha l ter a resposta j calculada a partir do modelo adotado (Pukstenis & Robinson, 2004).

O teste de *Pearson* para avaliar a adequação do ajuste compara essas contagens esperadas com as observadas, recorrendo à fórmula

$$\chi^2 = \sum_{l=1}^L \sum_{j=1}^k \frac{(O_{lj} - E_{lj})^2}{E_{lj}} \quad (53)$$

A estatística *Deviance* também compara contagens observadas e esperadas, mas recorrendo à fórmula

$$D = 2 \sum_{l=1}^L \sum_{j=1}^k O_{lj} \log \frac{O_{lj}}{E_{lj}} \quad (54)$$

Os testes usados para avaliar a qualidade do ajuste do modelo são baseados nas aproximações das estatísticas (53) e (54) para a distribuição Qui-quadrado, com $(L-1)(k-1)p$ graus de liberdade, em que L é o número de linhas, k é o número de colunas da tabela de contingência e p é o número de covariáveis do modelo.

O recurso ao teste de Pearson e ao teste *Deviance* não é consensual. Pukstenis & Robinson (2004) defendem que estas duas estatísticas não fornecem boas aproximações da distribuição Qui-quadrado nos casos em que as variáveis explicativas ajustadas são contínuas. Hosmer & Lemeshow (2000) consideram que, uma vez que ainda há uma escassez muito grande de ferramentas para avaliação de ajuste dos modelos ordinais, se deve recorrer a regressões binárias separadas para cada ponto de corte para que sejam criadas as estatísticas de diagnóstico para os modelos ordinais.

Geralmente, para estudar o ajuste de modelos de *odds* proporcionais, constroem-se gráficos de resíduos usando o ajuste desses modelos para prever uma série de eventos binários do tipo $Y > j$, $j = 1, 2, \dots, k$. Desta forma, para a variável indicadora, o resíduo *score*, para o caso i e covariável r é dado por:

$$U_{ir} = X_{ir} \left((Y_i \geq j) - P_{ij} \right), \quad (55)$$

em que

$$P_{ij} = \frac{1}{1 + \exp\left(-(\alpha_j + \beta X)\right)} \quad (56)$$

Nos gráficos dos resíduos *score* colocam-se a média, $\bar{U}_{.r}$, e os respetivos intervalos de confiança no eixo vertical e as categorias da variável dependente são colocadas no eixo horizontal. Se a suposição de *odds* proporcionais for válida, para cada covariável, os intervalos de confiança para cada categoria da resposta devem ser parecidos. Podem-se

também estudar os resíduos parciais para verificar se todas as covariáveis do modelo se comportam de forma linear, calculando modelos de regressão binária para todos os pontos de corte da variável dependente. Os gráficos de resíduos parciais geram estimativas de como cada covariável se relaciona com cada categoria da variável resposta, permitindo verificar a necessidade de transformações das covariáveis (linearidade) e a validade da suposição de *odds* proporcionais (teste das linhas paralelas).

Odds Ratio para dados ordinais

A estimação do *Odds Ratio* é feita considerando uma categoria da variável dependente como sendo a referência e compara-se essa com as restantes ou agrupando as categorias maiores e comparando-as com as menores.

Suponha-se que a variável dependente tem k categorias ordenadas ($Y_j, j = 1, 2, \dots, k$) e que se pretende comparar dois grupos A e B. Neste caso, para a categoria j , o *Odds Ratio* é dado por

$$OR_j = \frac{\frac{P(Y \leq Y_j | x^{(A)})}{1 - P(Y \leq Y_j | x^{(A)})}}{\frac{P(Y \leq Y_j | x^{(B)})}{1 - P(Y \leq Y_j | x^{(B)})}} = \frac{\frac{P(Y \leq Y_j | x^{(A)})}{P(Y > Y_j | x^{(A)})}}{\frac{P(Y \leq Y_j | x^{(B)})}{P(Y > Y_j | x^{(B)})}} = \frac{odds^{(A)}}{odds^{(B)}} \quad (57)$$

Por definição, *Odds Ratio* é a razão entre duas *chances* (*odds*), mas neste caso a chance é definida em termos de probabilidades cumulativas. Para a sua interpretação, basta considerar que a resposta foi dicotomizada com ponto de corte j . Se A e B representam respetivamente a exposição ou não a um fator de risco, a razão de odds quantifica a *chance* de um indivíduo do grupo exposto ser classificado até uma determinada categoria comparada com a chance do grupo não exposto.

3. DOENÇAS CARDIOVASCULARES

3.1. Dimensão do problema

As doenças cardiovasculares são, atualmente, responsáveis pela maior parte dos óbitos em todo o mundo. Em 2008, mais de 17 milhões de pessoas morreram devido a estas doenças, entre os quais, 18% tinham menos de 60 anos e prevalências de fatores de risco modificáveis (WHO, 2011b). Nesse mesmo ano, 27.2% dos óbitos registados em Angola foram causados por doenças cardiovasculares em adultos com idades entre os 30 e 70 anos (Tabela 4).

Tabela 4: Percentagens de óbitos causados por doenças cardiovasculares, em 2008, nos países angolanos de língua oficial portuguesa (PALOP)

País	Angola	Cabo Verde	Guiné-Bissau	Moçambique	São Tomé e Príncipe
Percentagem	27.2	30.4	26.0	23.0	28.0

Fonte: WHO, 2013a

Comparativamente, em Portugal, em 2008, foram registados 22.8% de mortes devido a doenças cardiovasculares, ou seja, uma percentagem inferior a todos os países angolanos de língua oficial portuguesa.

Até há bem pouco tempo, as doenças cardiovasculares eram uma epidemia associada a países desenvolvidos e considerada o fardo a carregar pelo aumento da qualidade de vida, originado pelas Revoluções Industrial e Tecnológica e pelas transformações económicas e sociais associadas. No entanto a dinâmica tem mudado e, recentemente, a percentagem de mortes prematuras nos países desenvolvidos, após a implementação de algumas estratégias de prevenção reduziu para 4%, enquanto que em alguns países sub-desenvolvidos já atingiu os 42%, devido ao aumento da ingestão de gorduras e redução da actividade física e ao aumento da esperança média de vida. Esta alteração aumentou, ainda mais, a disparidade entre vários países, uma vez que os menos desenvolvidos, que ainda estão a tentar baixar a prevalência das doenças transmissíveis, se vêem numa situação de duplo fardo socioeconómico (WHO, 2011b). Em Angola, entre 1990 e 2011 a

esperança média de vida subiu de 41 para 51 anos, sendo de 50 anos para os homens e 53 anos para as mulheres. Em 2008, 46.2% dos óbitos foram provocados por doenças transmissíveis, como a HIV/AIDS, a malária e a tuberculose e 47.5% por doenças não transmissíveis, como as doenças cardiovasculares entre outras. A mortalidade causada por este novo fardo é já bastante alta, apesar de ainda não ter havido redução da mortalidade por doenças herdadas na fase de pestilência e fome (WHO, 2013a).

Além da mortalidade, as doenças cardiovasculares são também uma das principais causas de morbilidade responsável por hospitalização, incapacidade física e, conseqüentemente, desarticulação socioeconómica de inúmeras famílias de pessoas que sobreviveram a eventos cardiovasculares. Em Angola, as carências em saúde e os problemas que o Serviço Nacional de Saúde enfrenta, atualmente, relacionados sobretudo com limitações de recursos humanos (quantitativas e qualitativas), de recursos materiais (reduzido acesso a água potável, saneamento e energia) e de recursos financeiros (insuficientes, devido a um modelo de financiamento inadequado) agravam ainda mais este problema e validam a urgência em se apostar na mobilização geral dos angolanos na luta pela promoção e preservação da saúde, atacando estas doenças muito antes de elas se manifestarem (Ministério da Saúde de Angola, 2012). Segundo a WHO (2011b), o combate deve ser multifacetado: vigilância; prevenção; controlo dos casos identificados.

A pesquisa relacionada com os vários aspetos das doenças cardiovasculares, em Angola, é escassa e é urgente vigiar e produzir conhecimento sobre fatores de risco, prevalências, custos económicos e sociais e o impacto da globalização e urbanização sobre os comportamentos. Após o conhecimento da dinâmica desta enfermidade no contexto angolano, é preciso apostar em estratégias de prevenção primária que levem à redução da exposição aos fatores de risco. Há uma clara evidência de que a combinação destas duas abordagens, aplicadas tanto individual como globalmente, contribuem para a redução do peso deste fardo e, conseqüente, declínio da mortalidade e morbilidade que lhe estão inerentes, sendo, por isso, uma solução bastante sustentável.

3.2. Doenças Cardiovasculares

As doenças cardiovasculares representam um amplo grupo de patologias que afetam o aparelho cardiovascular, constituído pelo coração e pelos vasos sanguíneos (sistema arterial e venoso).

As doenças cardiovasculares devem-se essencialmente a um acúmulo de depósitos de gordura nas paredes internas dos vasos sanguíneos, designado por aterosclerose, um fenómeno que tem início numa fase precoce da vida, progredindo silenciosamente durante anos, e que habitualmente já está bastante avançado no momento em que aparecem as primeiras manifestações clínicas. Contudo, regra geral, não causa qualquer sintomatologia até aos 50/70 anos, embora possa atingir adultos jovens (30/40 anos), principalmente se forem fumadores intensivos.

O coração pode ser afetado por uma grande diversidade de doenças como por exemplo: cardiopatia congénita (malformações presentes no momento do nascimento); doenças valvulares (como estenoses ou prolapsos nas válvulas cardíacas); doença reumática (secundária à febre reumática) e arritmias.

O sistema arterial pode ser afetado por uma patologia dilatante como é o caso dos aneurismas (dilatação localizada e permanente de uma artéria mais frequentemente localizados na artéria aorta abdominal) ou de patologia obliterante. A patologia obliterante pode afetar os territórios vasculares que irrigam o cérebro (levando à ocorrência de Acidente Vascular Cerebral ou Acidente Isquémico Transitório), as coronárias (responsáveis pela irrigação do músculo cardíaco, podendo levar à ocorrência de Enfarte Agudo do Miocárdio ou Angina de Peito); pode ainda afetar a vascularização do rim, do intestino ou dos membros inferiores.

No que respeita ao sistema venoso a complicação mais temível é a trombose venosa profunda, consequência da oclusão das veias profundas dos membros e que secundariamente pode dar origem a uma tromboembolia pulmonar.

As doenças cardiovasculares podem desenvolver-se mais precocemente, por vezes de forma bastante dramática, e muitos doentes poderão sofrer um Acidente Vascular

Cerebral, um enfarte agudo de miocárdio, morte súbita ou desenvolver insuficiência cardíaca, que é o resultado final de muitas doenças do coração que têm na sua origem, em grande parte, a negligência no controlo dos fatores de risco.

Há alguns sintomas que podem constituir sinais de alerta, principalmente em pessoas mais idosas, e que devem ser do conhecimento geral da população, como: dificuldade em respirar, sobretudo se acontecer em repouso; sensação de peso, aperto ou opressão por detrás do esterno, que por vezes se estende até ao pescoço, ao braço esquerdo ou ao dorso (angina de peito); alterações do ritmo cardíaco; enfarte do miocárdio, associado a dor prolongada no peito, surgindo muitas vezes em repouso e, por vezes, acompanhada de ansiedade, sudação, falta de força e vômitos; insuficiência cardíaca, que costuma estar associada a fadiga, grande debilidade, falta de ar em repouso, distensão do abdómen e pernas inchadas.

3.3. Fatores de Risco Cardiovascular

3.3.1. Introdução

Nos últimos 50 anos, os estudos relacionados com fatores de risco cardiovascular proliferaram em todo o mundo e contribuíram bastante para o conhecimento e compreensão da dinâmica entre os fatores de risco e as doenças cardiovasculares (Libby, 2005). A interpretação das relações entre a multiplicidade dos fatores de risco cardiovascular não é fácil, pois embora qualitativamente seja similar em todas as regiões do globo e em todas as populações, há variações significativas e lógicas nas prevalências dos diversos elementos (Yusuf, et al., 2004). O número de fatores de risco cardiovascular tem aumentado ao longo dos últimos anos, baseado na consistência demonstrada em vários estudos distintos, porém nem todos mostraram ter o mesmo peso no panorama das doenças cardiovasculares.

Apesar de alguns fatores de risco serem inalteráveis, como a idade e o género, a maior parte das doenças cardiovasculares estão também relacionadas com fatores de risco modificáveis, como o tabagismo, a alimentação inadequada, o sedentarismo e o uso nocivo

de álcool. Estes comportamentos pouco saudáveis provocam outras complicações como hipertensão arterial, excesso de peso/obesidade, diabetes e dislipidémia. A adoção destes comportamentos na infância desencadeia um processo silencioso que, passados alguns anos, se manifesta em eventos cardiovasculares, muitas vezes fatais, sendo, portanto imprescindível a sua prevenção, desde tenra idade. Em termos de mortalidade mundial, os principais fatores de risco são a hipertensão, o tabagismo, a diabetes, o sedentarismo e o excesso de peso/obesidade, responsáveis, respetivamente, por 13%, 9%, 6%, 6% e 5% dos óbitos (WHO, 2011b). Para se consultar os fatores específicos para cada tipo de doença cardiovascular pode-se consultar o Anexo 1.

3.3.2. Género

De uma forma global, os fatores de risco cardiovascular afetam de forma idêntica os homens e as mulheres. Na literatura detetam-se alguns equívocos em relação à dinâmica das doenças cardiovasculares entre géneros, pois o género masculino é globalmente associado a uma maior prevalência, mas a realidade é que o facto de a esperança média de vida ser maior nas mulheres do que nos homens, leva a que as doenças cardiovasculares se desenvolvam 7 a 10 anos mais tarde nas mulheres, fazendo com que ganhem anos de vida sem que elas se manifestem, quando comparadas com homens da mesma idade.

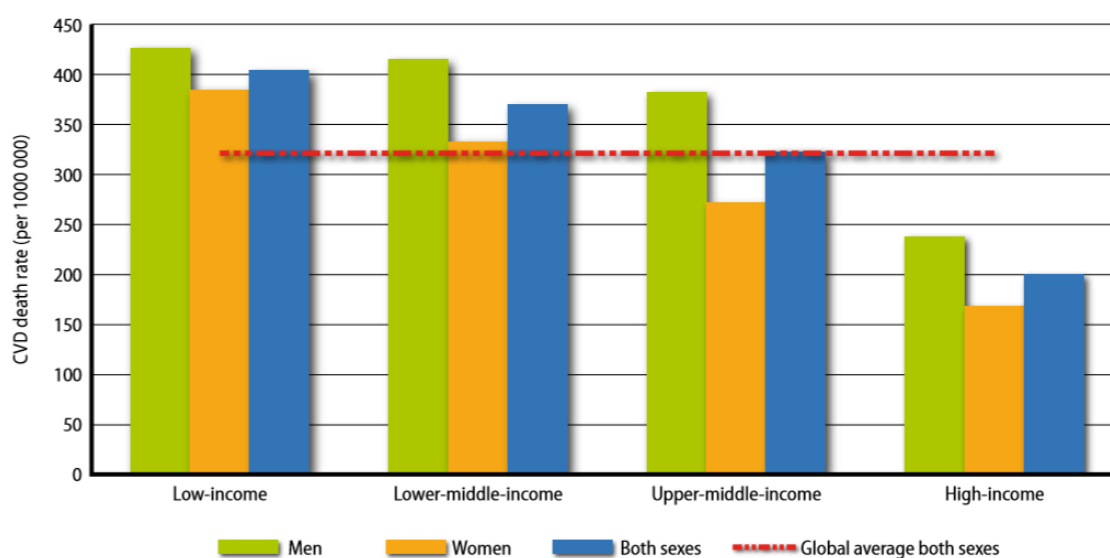


Figura 4: Mortalidade por doenças cardiovasculares, agrupadas por género, de acordo com o Banco Mundial (por 1000 000)

Fonte: WHO, 2011b

A mortalidade por doenças cardiovasculares não é significativamente maior nos homens do que nas mulheres (Figura 4). Esta concepção faz com que o risco de doenças cardiovasculares em mulheres seja frequentemente sobestimado (WHO, 2011b).

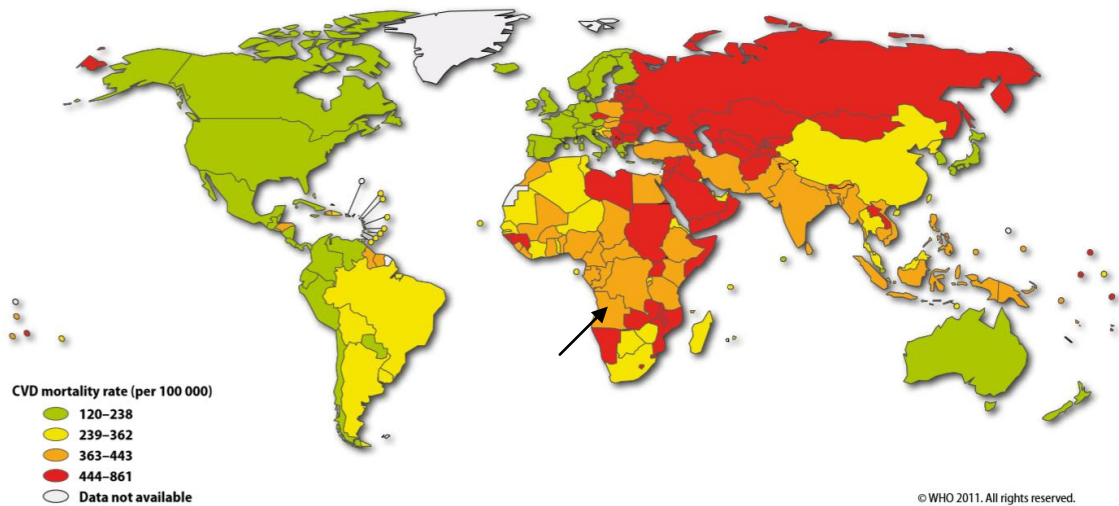


Figura 5: Distribuição mundial da mortalidade originada por doenças cardiovasculares, em homens (estratificada por idades, por 100 000)

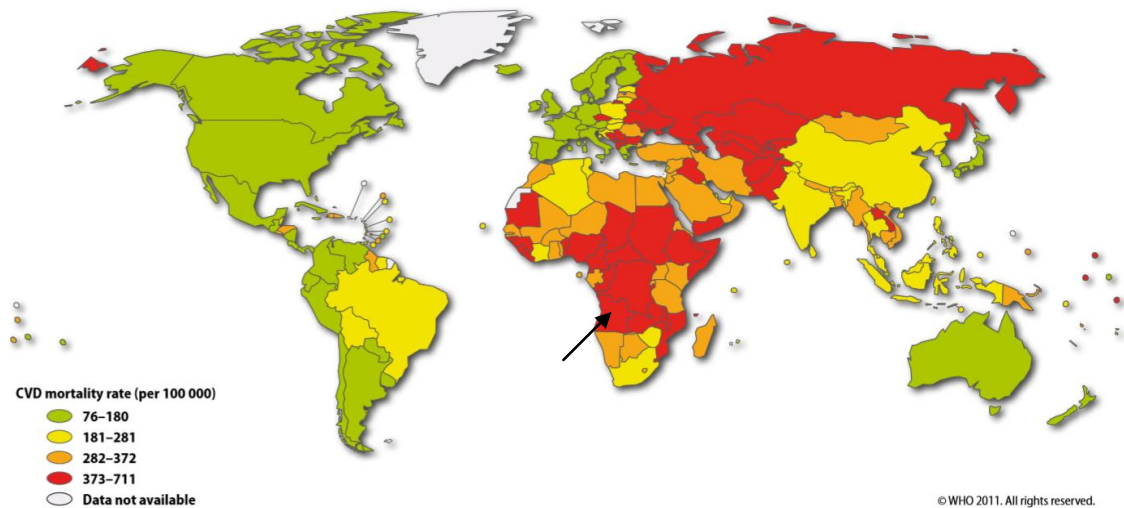


Figura 6: Distribuição mundial da mortalidade originada por doenças cardiovasculares, em mulheres (estratificada por idades, por 100 000)

Em Angola, e de uma forma geral na região africana, comparando a Figura 5 e a Figura 6 constata-se precisamente o oposto: a mortalidade originada por doenças cardiovasculares

afeta mais as mulheres do que os homens, apesar da esperança média de vida das mulheres angolanas (53 anos) ser três anos superior à dos homens (WHO, 2011b).

Uma fundamentação para que o facto da taxa de mortalidade nas mulheres, por doenças cardiovasculares, ser maior do que nos homens pode ser a desigualdade entre as normas e papéis das mulheres, específicos de cada sociedade, que influenciam a sua exposição a fatores de risco. Para um conhecimento mais aprofundado sobre as condições de saúde das mulheres africanas, pode-se consultar WHO (2012). Numa sociedade em desenvolvimento, como Angola, o problema agrava-se nas mulheres com risco de doenças cardiovasculares, pois além das diferenças referidas, ainda experimentam desafios específicos no acesso à prevenção, deteção e tratamento, devido à enorme desigualdade de género, responsabilidades familiares e os fracos recursos financeiros. Estes fatores são agravados pelos sistemas de saúde que não conseguem responder às necessidades específicas destas mulheres, quando as deveria considerar uma peça fundamental na luta contra as doenças cardiovasculares, através da promoção de comportamentos saudáveis, uma vez que podem ser importantes agentes de mudança, enquanto responsáveis pela educação e nutrição das suas famílias (WHO, 2011b).

3.3.3. Idade

Universalmente, o aumento da esperança média de vida tem provocado um aumento explosivo da prevalência de fatores de risco, nas populações, e conseqüente da ocorrência de eventos cardiovasculares. Vários estudos mostram que a prevalência de doenças cardiovasculares, tanto em homens como em mulheres, aumenta com o avançar da idade, tornando-o num fator de risco cardiovascular inquestionavelmente poderoso, pois não pode ser alterado e, na presença de outros fatores de riscos, parece ter um efeito intensificador (WHO, 2011b).

As doenças cardiovasculares, muitas vezes pensadas ser um problema "masculino", são a principal causa de morte de pessoas idosas de ambos os sexos em quase todo o mundo. Uma das grandes discussões dos epidemiologistas, desde a década de 70, é se a menopausa pode ser considerada um fator de risco a mais, já que a idade acima de 55 anos nas mulheres (idade na qual as mulheres em média já apresentaram a menopausa) e acima

de 45 anos nos homens são fatores de risco independentes (Grundy, Pasternak, Greenland, Smith Jr, & Fuster, 1999).

Na verdade, a cada ano a doença cardiovascular provoca um maior número de mortes em mulheres mais velhas do que em homens mais velhos (7.5 milhões de mulheres com mais de 60 anos de idade, em comparação com 6.6 milhões de homens em 2008), sobretudo nos países sub-desenvolvidos, pois nos países desenvolvidos tendo diminuído significativamente nos últimos anos, devido à redução de comportamentos de risco, a uma melhor gestão da medicação direcionada para alguns fatores de risco e à melhoria do tratamento de doenças cardiovasculares, de uma forma geral. Em contrapartida, nos países sub-desenvolvidos, como em Angola, na maior parte das vezes estas doenças não são reconhecidas nem tratadas, sobretudo nas mulheres. E este panorama agrava-se pelo facto delas serem menos propensas a procurar ajuda médica, não obtendo os cuidados adequados atempadamente (Global Health Observatory, 2013).

3.3.4. Hábitos alimentares inadequados

Contrariamente ao género e à idade, que são fatores de risco que não podem ser modificados, há um conjunto de outros fatores que podem ser facilmente modificados e está provado que a redução da exposição a esses fatores de risco reduz o risco de doenças cardiovasculares. Algumas dessas evidências podem ser consultadas no relatório *Prevention of Cardiovascular Disease- Guidelines for assessment and management of cardiovascular risk*, da WHO (2007).

O papel que bons hábitos alimentares têm para o desenvolvimento de doenças cardiovasculares ou de outras doenças que lhes estão associadas (como a obesidade, a hipertensão ou a diabetes) é inquestionável (WHO, 2011a). É sabido que a transição económica que alguns países em desenvolvimento, como Angola, estão a atravessar está a provocar um aumento do consumo (tanto em quantidade como em fraca qualidade nutricional) de bebidas alcoólicas e de alimentos com gorduras saturadas e com excesso de sal e de açúcar e uma diminuição do consumo de frutas, legumes e fibras. O risco de desenvolver problemas cardiovasculares aumenta ainda mais quando, aliados a estes hábitos, se associam hábitos de sedentarismo, que também têm tendência a aumentar

com o desenvolvimento económico. A combinação destes dois maus hábitos, geralmente, culmina em excesso de peso e até em obesidade.

3.3.5. Excesso de peso/obesidade

Segundo a Organização Mundial da Saúde, a obesidade é uma doença em que o excesso de gordura corporal acumulada pode atingir graus capazes de afectar a saúde. A prevalência da obesidade, a nível mundial, é tão elevada que a considerou como a epidemia global do século XXI. Neste século, a obesidade tem uma prevalência igual ou superior à da desnutrição e das doenças infecciosas. Por tal facto, se não se tomarem medidas drásticas para prevenir e tratar a obesidade, mais de 50% da população mundial será obesa em 2025.

Em todo o mundo, mais de 2.8 milhões de pessoas morrem anualmente devido a excesso de peso/obesidade e, em 2008, 34% dos adultos, com mais de 20 anos de idade, estavam acima do peso, com um índice de massa corporal igual ou superior a 25 kg/m^2 (33.6% nos homens e 35% das mulheres), entre os quais 9.8 % dos homens e 13.8 % das mulheres eram obesos, ou seja, tinham índice de massa corporal igual ou superior a 30 kg/m^2 (WHO, 2011b). Em Angola, nesse mesmo ano, a prevalência de obesidade era de 3.8% nos homens e 10.2% nas mulheres, mostrando uma diferença de 6.4 pontos percentuais, entre os dois géneros (WHO, 2013a).

A associação entre o índice de massa corporal e a hipertensão foi também enfatizada nos estudos de Framingham, onde foi observado que excesso de 20% do peso ideal aumenta em oito vezes a incidência de hipertensão arterial (Framingham Heart Study, 2013). Num estudo realizado por Carneiro et al (1992) a prevalência de hipertensão aumentou de 23% no grupo com sobrepeso para 67.1% no grupo com obesidade grau 3, com índice de massa corporal igual ou superior a 40 kg/m^2 .

O índice de massa corporal não distingue entre diferentes tipos de adiposidade, alguns dos quais podem estar mais associados a doença cardiovascular.

Outro aspecto importante na avaliação da obesidade é a distribuição da gordura corporal. Quando o tecido adiposo se acumula na metade superior do corpo, sobretudo no abdómen, diz-se que a obesidade é abdominal, sendo típica do homem obeso; quando a

gordura se distribui, sobretudo, na metade inferior do corpo, particularmente na região glútea e coxas, diz-se que é do tipo ginóide, sendo típica da mulher obesa.

Estudos mais recentes dos diferentes tipos de tecido adiposo têm demonstrado, por exemplo, que a obesidade abdominal tem uma correlação com a doença cardiovascular maior do que o índice de massa corporal, por si só. A Organização Mundial de Saúde propõe que valores de perímetro abdominal inferiores a 80cm nas mulheres e a 94cm nos homens sejam considerados normais e os 88cm nas mulheres e 102cm nos homens sejam considerados pontos de corte entre risco aumentado e risco elevado.

Vários estudos epidemiológicos prospetivos têm demonstrado uma forte relação entre o excesso de peso/obesidade e o risco de doenças cardiovasculares (WHO, 2011b). Além dos seus efeitos individuais, a obesidade favorece ainda a ocorrência de outros fatores de risco cardiovascular, sendo que a distribuição central da gordura corporal se destaca especialmente como fator importante no desenvolvimento da hipertensão arterial (Carneiro et al., 1992).

Em Angola, poucos estudos há sobre a prevalência do excesso de peso/obesidade. Segundo Simão *et al* (2008), num estudo realizado a 667 estudantes universitários angolanos, 69.1% dos indivíduos apresentaram índice de massa corporal normal, 10.6% estavam com baixo peso, 17.1% com pré-obesidade e apenas 3.2% eram obesos.

3.3.6. Hipertensão

A hipertensão é uma das causas evitáveis mais importantes de morte e é definida como uma pressão arterial sistólica acima de 140 mmHg ou pressão arterial diastólica acima de 90 mmHg. Na maioria dos países, até 30% dos adultos sofrem com a hipertensão e mais de metade deles estariam com melhores condições de saúde, se reduzissem a sua pressão arterial, aumentado a atividade física, mantendo um peso saudável e comendo mais frutas e legumes. Para pessoas com idade até 50 anos, tanto a pressão arterial sistólica como a diastólica estão associadas a risco cardiovascular, mas acima desta idade a pressão arterial sistólica tem tido lugar de destaque como preditora. A pressão arterial geralmente aumenta com a idade, a menos que a ingestão de sal seja baixa, a atividade física seja alta e não haja quadro de obesidade (WHO, 2002).

Estima-se que a pressão arterial elevada seja responsável por cerca de 7.5 milhões de mortes no mundo, cerca de 12,8 % do total de todas as mortes anuais além de ser um fator de risco preponderante para o desenvolvimento de doenças cardiovasculares, uma vez que, em determinados grupos etários, o risco de doenças cardiovasculares duplica a cada aumento incremental de 20/10mmHg de pressão arterial, a partir de pressões baixas, como 115/75 mmHg (WHO, 2011b).

Globalmente, a prevalência global de hipertensão arterial, em adultos com 25 ou mais anos, rondava os 40% em 2008. Em Angola, nesse mesmo ano, a prevalência de hipertensão era de 39.6% nos homens e 33.8% nas mulheres (WHO, 2013a). Apesar do reconhecimento da sua gravidade, muitos países africanos, como Angola, dedicam pouca atenção à sua prevenção e controlo e praticamente não há estudos publicados sobre as suas prevalências, se bem que localmente, pelo menos em Benguela, se nota que começam a surgir algumas iniciativas direccionadas para esta doença, sobretudo em termos de rastreios nos quais são aferidas as pressões arteriais e as pessoas são alertadas para a importância de se prevenirem contra este tipo de doenças.

3.3.7. Hábitos tabágicos

Existem atualmente cerca de um bilião de fumadores no mundo, sendo a prevalência mais baixa (10%) na região Africana (WHO, 2011b). Atualmente ainda não existem muitas referências sobre a prevalência de hábitos tabágicos em Angola, nem sobre os pontos de corte a partir dos quais o risco cardiovascular aumenta (WHO, 2013b) (WHO, 2013c).

Angola aderiu à Convenção-Quadro para o controlo do tabaco, através da resolução número 48/05, de 3 de Outubro de 2005, da Assembleia Nacional, comprometendo-se a realizar iniciativas para controlar o tabagismo: proibição da propaganda; educação e conscientização da população; proibição de fumo em ambientes fechados; controlo do mercado ilegal de cigarros; tratamento da dependência da nicotina; inserção de mensagens de advertências sanitárias fortes e contundentes nas embalagens dos produtos de tabaco; regulação dos produtos de tabaco quanto aos seus conteúdos e emissões, entre outras. Em 2009, o Governo aprovou um Decreto que proíbe o fumo nos locais públicos. Recentemente, a WHO (2013c) publicou os primeiros dados sobre o tabagismo em Angola,

mas apenas foi estudada a prevalência de hábitos tabágicos em jovens entre os 13 e os 15 anos de idade, tendo-se constatado uma prevalência de 20% e sem diferenças significativas entre os dois géneros. Arrais, Lusitano & Neto (2004) estudaram 173 adultos angolanos e verificaram uma prevalência de 20% de fumadores. Uma prevalência muito diferente dos 4% encontrados num outro estudo, a estudantes universitários do Lubango (Simão, Hayashida, Santos, Cesarino, & Nogueira, 2008).

3.4. Avaliação do Risco Cardiovascular

Muitas vezes, coexistem na mesma pessoa mais do que um fator de risco e os efeitos da conglomeração (*clusters*) vão além de uma simples soma do contributo de cada um, pois a sinergia entre eles induz, por potenciação mútua, aumentos exponenciais no risco individual de doenças cardiovasculares, tornando ainda mais difícil a quantificação do contributo de cada um (Vaz, Santos, & Carneiro, 2005). Os estudos de grandes coortes populacionais, como o de Framingham (Framingham Heart Study, 2013) ou o SCORE Project (Conroy, et al., 2002), são uma mais valia e têm alertado para a importância da identificação individual de fatores de risco para posterior modelação e predição do risco de doenças cardiovasculares, em pessoas assintomáticas. Se as pessoas em risco de desenvolver doenças cardiovasculares forem identificadas antecipadamente, podem ser tomadas medidas para redução do risco e a maior parte dos eventos cardiovasculares poderão ser prevenidos (WHO, 2011a,b).

Em Angola não há estudos sobre a prevalência da pressão arterial alta, nem do consumo de tabaco, nem da glicose aumentada no sangue, nem do sedentarismo, nem do excesso de peso/obesidade e, evidentemente, também não foi feito qualquer estudo populacional que tenha permitido a construção de um modelo preditivo do risco cardiovascular, específico para o contexto angolano (WHO, 2011a,b). No sentido de colmatar esta lacuna, um grupo de especialistas, baseando-se em inúmeros estudos de coorte disponíveis na literatura, criou os modelos de predição de risco cardiovascular da OMS, recorrendo à coleta padronizada, à análise da prevalência de fatores de risco em cada região do mundo e à criação de coortes hipotéticos para cada uma das 14 sub-regiões epidemiológicas. Estas

escalas colorimétricas (Figura 7) possibilitam prever o risco de, em 10 anos, um angolano assintomático desenvolver um evento cardiovascular, fatal ou não fatal, de acordo com idade, género, pressão arterial, tabagismo, colesterol total do sangue (facultativo) e presença ou ausência de diabetes (WHO, 2007). Se o risco for inferior a 10%, significa que é baixo; se for entre 10 e 20% significa que é moderado; se for entre 20 e 30% significa que é alto; se for superior a 30% significa que é muito alto (WHO, 2007).

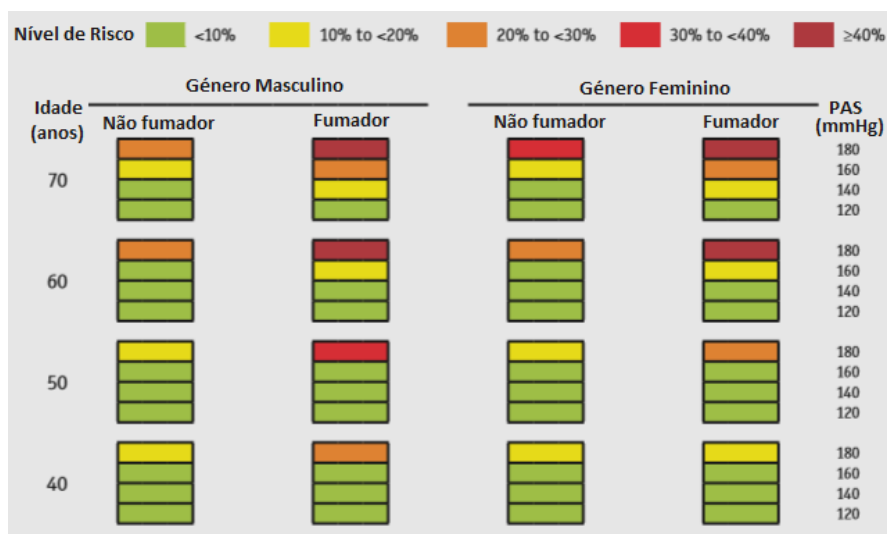


Figura 7: Estimação do risco de, em 10 anos, um africano desenvolver um evento cardiovascular, fatal ou não fatal, de acordo com o género, os hábitos tabágicos, a idade e a pressão arterial sistólica

Fonte: WHO, 2011a

O modelo descrito acima apenas é aplicável a indivíduos com idades superiores a 40 anos. Tendo em conta que este estudo foi feito apenas onze anos depois do fim da guerra civil angolana e que essa guerra durou 27 anos (antecedida pela guerra pela independência, que durou treze anos), este modelo estuda o risco cardiovascular de pessoas que viveram quase toda a vida num clima de pestilência e de fome e, como tal, considerou-se que o estudo do perfil cardiovascular deveria abranger todas as faixas etárias dos benguelenses adultos.

Para tal, foram considerados mais dois modelos para avaliação do risco cardiovascular, tendo medidas antropométricas e, no caso de medidas clínicas, apenas as pressões arteriais, pois a aferição de outras medidas clínicas, como o índice de glicémia, seria mais dispensioso.

Um dos modelos selecionados é proposto pela Organização Mundial de Saúde para estratificação do risco cardiovascular, de acordo com a avaliação do índice de massa corporal, do perímetro abdominal e do género, de acordo com a Tabela 5.

Tabela 5: Estratificação do risco cardiovascular tendo em conta o género, o índice de massa corporal e o perímetro abdominal

Índice de massa corporal (kg/m ²)		Perímetro abdominal (cm)	
		Homem: 94-102 Mulher: 80-88	Homem: >102 Mulher: >88
Baixo peso	<18.5	Normal	Normal
Peso saudável	18.5-24.9	Normal	Aumentado
Sobrepeso	25.0-29.9	Aumentado	Alto
Obesidade	>30	Alto	Muito alto

Fonte: OMS

Uma vez que no modelo anterior não é tida em conta a condição de hipertensão, optou-se por estratificar, também, o risco cardiovascular de acordo com a condição de hipertensão e o número de outro fatores de risco, de acordo com a Tabela 6.

Tabela 6: Estratificação do risco cardiovascular tendo em conta a condição de hipertensão e o número de fatores de risco adicionais

Hipertensão	Número de fatores de risco adicionais		
	0	1-2	>3
Normal	Normal	Baixo acrescido	Moderado acrescido
Normal alta	Normal	Baixo acrescido	Alto acrescido
Hipertensão de grau I	Baixo acrescido	Moderado acrescido	Alto acrescido
Hipertensão de grau II	Moderado acrescido	Moderado acrescido	Alto acrescido
Hipertensão de grau III	Alto acrescido	Muito alto acrescido	Muito alto acrescido

Estes três modelos são bastante frágeis e nenhum deles parece ser suficientemente abrangente: o primeiro é comum para muitos países africanos, não tem em conta a obesidade e apenas é aplicável a pessoas com mais de 40 anos; o segundo não tem em conta a condição de hipertensão nem a idade; o terceiro é simples demais, considerando que todos os fatores, além da hipertensão, têm o mesmo peso no modelo.

Porém, tendo em conta as altas prevalências de doenças cardiovasculares, a falta de condições do sistema de saúde e a falta de informações sobre os fatores de risco cardiovasculares considera-se que o estudo do risco cardiovascular numa perspectiva multimodelada será uma mais valia, para a atual realidade angolana.

Desta forma, neste trabalho, o risco cardiovascular será classificado de três formas distintas, recorrendo aos três modelos descritos anteriormente.

PARTE II

1. INTRODUÇÃO

O perfil epidemiológico das doenças crónicas não transmissíveis é insuficientemente conhecido em Angola, por não fazer parte do sistema de notificação de doenças obrigatórias e não existirem estudos sobre as mesmas. Todavia, atualmente, é já reconhecido, pelos profissionais de saúde, o seu aumento progressivo e acelerado, lado a lado com as ainda altas prevalências das principais doenças transmissíveis, constituindo assim um duplo fardo para o país que urge ser minimizado (Ministério da Saúde de Angola, 2012).

Perante este flagelo, de dimensão mundial, e o número reduzido de dados epidemiológicos sobre doenças cardiovasculares e fatores de risco em Angola é urgente que alguns estudos se comecem a direccionar para esta temática, no sentido de se definirem estratégias, para a prevenção e o controlo, específicas para o contexto angolano, tomando como exemplo tudo aquilo que já foi feito em países desenvolvidos e que resultaram em reduções significativas das taxas de mortalidade associadas a essas doenças (Figura 8).

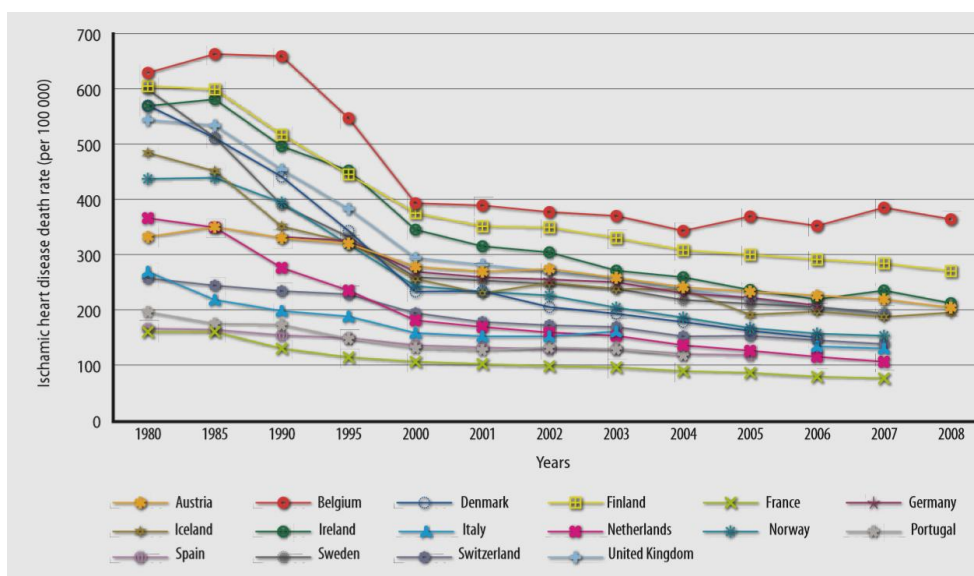


Figura 8: Estatísticas relativas à mortalidade por doenças cardiovasculares (standardizadas por idade), em alguns países desenvolvidos, entre 1980 e 2008

Fonte: WHO, 2011b

Os dados utilizados neste estudo foram recolhidos durante uma feira de saúde realizada em Novembro de 2013, pelo Instituto Superior Politécnico de Benguela, no largo de África, em Benguela (ANGOP, 2013).

Apesar de haver muito ainda por ser desvendado sobre os determinantes das doenças cardiovasculares em Angola, parece haver alguns fatores de risco cardiovascular universais, bem delineados na literatura. Com base nesse conhecimento, além do género e da idade, todos os indivíduos foram avaliados quanto aos seguintes parâmetros:

- Indicadores antropométricos: peso; estatura, perímetro abdominal.
- Indicadores clínicos: pressão arterial sistólica; pressão arterial diastólica.
- Indicadores de estilo de vida: hábitos tabágicos; ingestão de álcool; consumo de gorduras saturadas.

Para a coleta dos dados referentes ao estilo de vida, utilizou-se um inquérito por entrevista. Após a entrevista, os indivíduos foram submetidos a uma avaliação antropométrica, seguida da aferição de pressão arterial. Nenhum dos indivíduos afirmou ter qualquer doença cardiovascular.

O peso foi aferido numa balança digital com capacidade máxima de 150kg. A estatura foi aferida com fita métrica colada a uma parede sem rodapé, com extensão de dois metros, dividida em centímetros. Com as medidas de peso e estatura, calculou-se o índice de massa corporal, de acordo com a fórmula [IMC = peso (kg) / estatura² (m²)]. Para uma melhor avaliação do sobrepeso do indivíduo, o índice de massa corporal foi estratificado de acordo com a Tabela 7.

Tabela 7: Classificação internacional de um adulto quanto ao peso, pelo índice de massa corporal

Classificação	Valor do índice de massa corporal (kg/m²)
Baixo peso	Inferior a 18.50
Peso saudável	Entre 18.5 e 24.99
Excesso de peso	Entre 25.00 e 29.99
Obesidade	Maior ou igual a 30.00

Fonte: WHO, 2006

O perímetro abdominal foi medido na menor curvatura localizada entre as costelas e a crista ilíaca com fita métrica flexível e inelástica. Para uma melhor avaliação do risco associado foi feita uma estratificação do perímetro abdominal. Segundo a Organização

Mundial de Saúde, nos indivíduos do sexo feminino, perímetros abdominais superiores a 80 cm são considerados como um fator de risco aumentado de complicações metabólicas e acima dos 88 cm como um fator de risco elevado; nos indivíduos do sexo masculino esses valores são 94 e 102 cm, respetivamente.

Para medição das pressões arteriais, foram utilizados esfigmomanómetros aneroides e estetoscópios duplos, além de jogos de manguitos com larguras variáveis. Durante a aferição da pressão arterial, o indivíduo permaneceu em posição sentada, relaxado, com as costas apoiadas, pés pousados no chão, pernas descruzadas e com o braço apoiado sobre uma mesa e à altura do precórdio.

Tabela 8: Classificação de um adulto quanto aos valores das pressões arteriais

Classificação	Valor da pressão arterial sistólica (mmHg)	Valor da pressão arterial diastólica (mmHg)
Ótima	<120	<80
Normal	120-129	80-84
Normal alta	130-139	85-89
Hipertensão de grau I	140-159	90-99
Hipertensão de grau II	160-179	100-109
Hipertensão de grau III	>180	>110

Fonte: Mancia, 2007

A amostra utilizada neste estudo é composta por 600 benguelenses assintomáticos, adultos, que participaram na feira e que autorizaram a avaliação da sua condição cardiovascular, entre os quais 355 (59.2%) eram do género masculino e 245 (40.8%) eram do género feminino (Figura 9).

Em termos de idade, 199 (33.2%) tinham menos de 30 anos; 164 (27.3%) tinham [30,40[anos; 164 (27.3%) tinham [40,50[anos; 56 (9.3%) tinham [50,60[anos; 17 (2.8%) tinham 60 anos ou mais (Figura 9). Em quase todas as faixas etárias, o número de participantes do género masculino era superior ao número de participantes do género feminino.

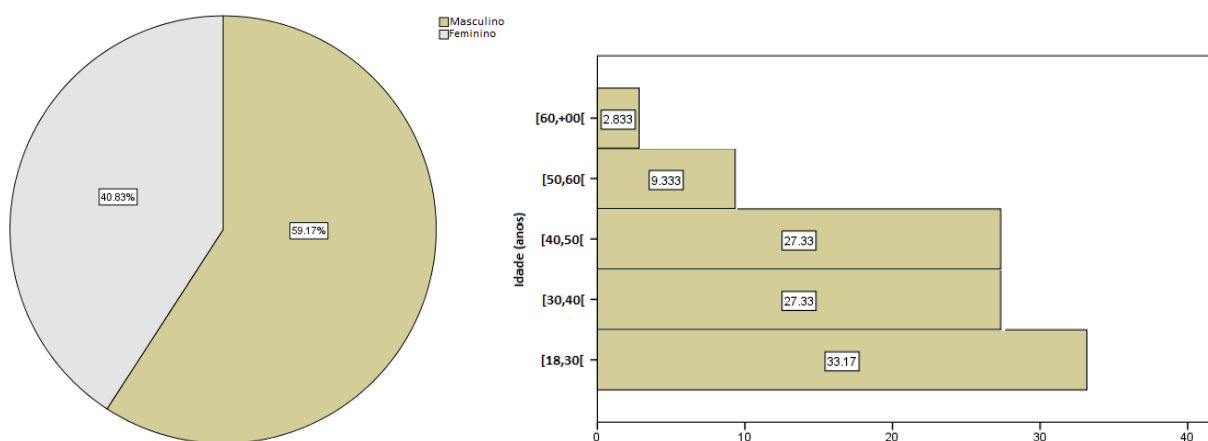


Figura 9: Distribuição dos participantes (n=600) por género e por faixas etárias

Na Figura 10 é possível observar a proporção de participantes em cada faixa etária, agrupados por género.

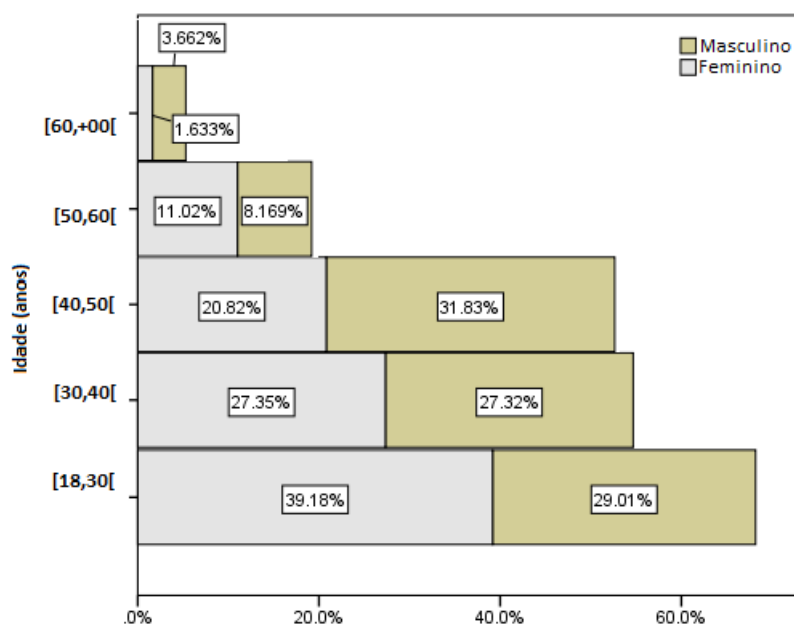


Figura 10: Distribuição do género dos participantes, por faixas etárias (n₁=355; n₂=245)

Posteriormente, as variáveis foram estratificadas pela forma como estão associadas ao risco de doenças cardiovasculares, considerando os pontos de corte já referidos na primeira parte da dissertação (Tabela 9).

Tabela 9: Estratificação das variáveis, segundo o risco para Doenças Cardiovasculares

Variável	Estratificação	Risco aumentado
Género	Feminino	
	Masculino	+
Idade (anos)	Inferior a 45 anos nos homens e a 55 anos nas mulheres	
	45 anos ou mais nos homens e 55 anos ou mais nas mulheres	+
IMC (kg/m ²)	Inferior a 25 kg/m ²	
	25kg/m ² ou mais	+
Perímetro abdominal (cm)	Inferior a 94cm nos homens e a 80cm nas mulheres	
	94cm ou mais nos homens e 80cm ou mais nas mulheres	+
Hipertensão arterial (mmHg)	Normotenso- Pressão arterial sistólica inferior a 140 ou pressão arterial diastólica inferior a 90mmHg	
	Hipertenso- Pressão arterial sistólica igual ou superior a 140 ou pressão arterial diastólica igual ou superior a 90mmHg	+
Hábitos tabágicos	Não fumador	
	Fumador	+
Ingestão de bebidas alcoólicas	Não consumidor de bebidas alcoólicas	
	Consumidor de bebidas alcoólicas	+
Consumo de gorduras saturadas	Não consumidor de gorduras saturadas	
	Consumidor de gorduras saturadas	+

Tedo em conta o que foi referido na primeira parte da dissertação, sobre o facto de ainda não haver um modelo preditivo específico para o contexto angolano que englobe a maior parte dos fatores de risco referidos na Tabela 9, e as limitações dos poucos modelos que são válidos para esta população, optou-se por, nesta dissertação se avaliar o risco cardiovascular de acordo com as Tabela 5 e a Tabela 6. Também foi feita a estimação do risco de, em 10 anos, um africano, com 40 anos ou mais, desenvolver um evento cardiovascular, fatal ou não fatal, de acordo com a Figura 7.

Neste trabalho, recorrendo a metodologias estatísticas e ferramentas computacionais avançadas, pretende-se avaliar o grau de associação entre os diferentes fatores de risco estabelecidos, com o grau de risco de doenças cardiovasculares, de acordo com os critérios descritos na PARTE I, colmatando nos seguintes objetivos:

- Descrever a amostra de benguelenses, quanto ao género, à idade, ao peso, à estatura, ao índice de massa corporal, ao perímetro abdominal, à pressão arterial, aos hábitos tabágicos, ao consumo de bebidas alcoólicas, ao consumo de gorduras saturadas e ao risco cardiovascular;
- Comparar o grau de risco cardiovascular entre homens e mulheres;
- Comparar o grau de risco cardiovascular entre indivíduos mais velhos e mais novos;
- Comparar o grau de risco cardiovascular entre indivíduos com sobrepeso e sem sobrepeso;
- Comparar o grau de risco cardiovascular entre indivíduos com obesidade abdominal e sem obesidade abdominal;
- Comparar o grau de risco cardiovascular entre hipertensos e não hipertensos;
- Comparar o grau de risco cardiovascular entre fumadores e não fumadores;
- Comparar o grau de risco cardiovascular entre indivíduos que ingerem álcool e que não ingerem álcool;
- Comparar o grau de risco cardiovascular entre consumidores de gorduras saturadas e não consumidores;
- Construir um modelo estatístico que permita prever o grau de risco cardiovascular, de acordo com alguns indicadores de risco, que se tenham mostrado mais significativos.

Em qualquer estudo a análise dos dados pode (e deve) ser feita numa perspetiva univariada, bivariada ou multivariada, consoante o número de variáveis tratadas em simultâneo. Numa análise univariada, cada variável é tratada isoladamente e deve sempre ser o primeiro passo da exploração dos dados (Pestana & Gageiro, 2008). Para as variáveis quantitativas foram estudadas medidas descritivas (mínimo; máximo; amplitude; média; desvio-padrão; mediana; valores aberrantes ou *outliers*; média aparada a 5%;

enviesamento; curtose), foram construídos gráficos e tabelas e estudado o ajustamento à distribuição normal, recorrendo ao teste de Kolmogorov-Smirnov, com correção de Lilliefors, e à análise dos gráficos *Q-Q* e *Detrended normal QQ plots*. Para as variáveis ordinais foram estudadas as frequências absolutas e relativas, foram construídos gráficos e foram comparadas as proporções de indivíduos nas diferentes categorias das variáveis, aplicando o teste de Kolmogorov-Smirnov. Para as variáveis nominais, que neste estudo são todas dicotómicas (Tabela 9), foram estudadas as frequências absolutas e relativas, foram construídos gráficos e foram comparadas as proporções de indivíduos nas diferentes categorias das variáveis, aplicando o teste da binomial, que confronta as frequências obtidas com as frequências esperadas.

Algumas das variáveis foram dicotomizadas para que os indivíduos fossem associados a grupos de risco, o que será útil para as análises bivariada e multivariada.

O risco cardiovascular foi avaliado por três modelos diferentes e, para facilitar a análise, foram designados da seguinte forma:

Risco A – Tabela 5: Estratificação do risco cardiovascular tendo em conta o género, o índice de massa corporal e o perímetro abdominal (Tabela 5).

Risco B – Tabela 6: Estratificação do risco cardiovascular tendo em conta a condição de hipertensão e o número de fatores de risco adicionais (Tabela 6).

Risco C – Figura 7: Estimação do risco de, em 10 anos, um africano desenvolver um evento cardiovascular, fatal ou não fatal, de acordo com o género, os hábitos tabágicos, a idade e a pressão arterial sistólica (Figura 7).

Posteriormente foi feita a análise bivariada, na qual foram estudadas as relações entre os vários indicadores (sociais; antropométricos; clínicos; de estilo de vida) e cada um dos grupos associados a risco, de acordo com a Tabela 9 e também as relações entre os riscos e cada um dos grupos associados a risco. O teste de Mann-Whitney foi usado no caso em que as variáveis, divididas por grupos de risco, eram quantitativas. Nos casos em que foram detetadas relações entre as variáveis foi calculado o Eta e o quadrado do Eta, para determinar a percentagem de uma variável que era explicada pela outra. Para os casos em que as variáveis agrupadas eram ordinais optou-se pelo teste de independência de Kolmogorov-Smirnov, pois tem sido apontado como mais eficiente do que o teste de

independência do Qui-quadrado, por ser menos generalista (Pestana & Gageiro, 2008). Nos casos em que foram detetadas relações foi calculado o coeficiente V de Cramer, para medir o grau de associação. No que diz respeito às variáveis nominais, uma vez que as deste estudo são dicotómicas, optou-se pelo cálculo do *Odds Ratio* (OR), tendo em conta que atualmente muitos investigadores defendem que, para tabelas 2x2, a análise através de produtos cruzados é mais facilmente interpretável do que o teste de independência do Qui-quadrado (Pestana & Gageiro, 2008)).

Em termos de análise multivariada foram usados os modelos estereótipos, tendo em conta a natureza original da variável dependente (ordinal discreta) e o objetivo do estudo. As ferramentas disponíveis para a aplicação deste modelo são bastante reduzidas e o processo de modelação foi bastante moroso, tanto em termos de experimentação de técnicas como em termos de pesquisa bibliográfica. O *software* R revelou-se o melhor aliado, graças ao pacote VGAM e à função *rrvglm*. A função referida anteriormente, porém, mostrou-se incompatível com algumas funções que pareciam ser transversais a quase todos os tipos de modelação. Por exemplo, se conseguiu encontrar referências sobre o recurso ao método *stepwise* para auxiliar na seleção das variáveis para o modelo, nem se conseguiu usar a função *confint()* para determinarem intervalos de confiança, devido à distribuição considerada para tal ser a normal. Uma vez que a vertente mais prática da aplicação deste modelo está pouco documentada na literatura, sobretudo na nacional, optou-se por se tentar dar um contributo nesse sentido e todos os passos do processo de modelação foram explicados exhaustivamente.

A fundamentação teórica das metodologias estatísticas utilizadas foi apresentada na primeira parte desta dissertação, para que nesta segunda parte se dê especial atenção à parte mais prática.

As ferramentas computacionais usadas neste trabalho foram o SPSS (*Statistical Package for the Social Sciences*), versão 20.0, o *software* R e o *software* Geogebra.

2. ANÁLISE UNIVARIADA

2.1. Indicadores sociais

Género

Entre os 600 benguelenses que participaram no estudo, 355 (59.2%) eram do género masculino e 245 (40.8%) eram do género feminino (Figura 9). A proporção de homens e de mulheres parecia ser a mesma, porque as frequências observadas estavam relativamente próximas de 50% (frequências esperadas), mas essa suposição tinha que ser verificada estatisticamente. Para isso, recorreu-se ao teste binomial e às seguintes hipóteses:

H0: A proporção entre homens e mulheres é igual.

H1: A proporção entre homens e mulheres é diferente.

Tabela 10: Output (SPSS) do teste da binomial para a variável género (n=600)

	Categoria	N	Proporções observadas	Proporções esperadas	Significância exata
Género	Masculino	355	.59	.50	.000
	Feminino	245	.41		
	Total	600	1.00		

Os resultados do teste binomial (Tabela 10) mostraram que o número observado de mulheres (245) era significativamente inferior ao número observado de homens, para um erro tipo I do analista de 0.05, na medida em que o nível de significância exato do teste foi 0.000, levando à rejeição da hipótese nula.

Idade

Os parâmetros relativos à idade dos participantes no estudo podem ser consultados no Anexo 2.

A distribuição da idade dos participantes era assimétrica positiva, pois o quociente $0.422/0.100=4.22$ tinha um valor absoluto superior a 1.96. Esta assimetria positiva parecia estar evidenciada no histograma da Figura 11, mostrando haver uma predominância das idades mais novas.

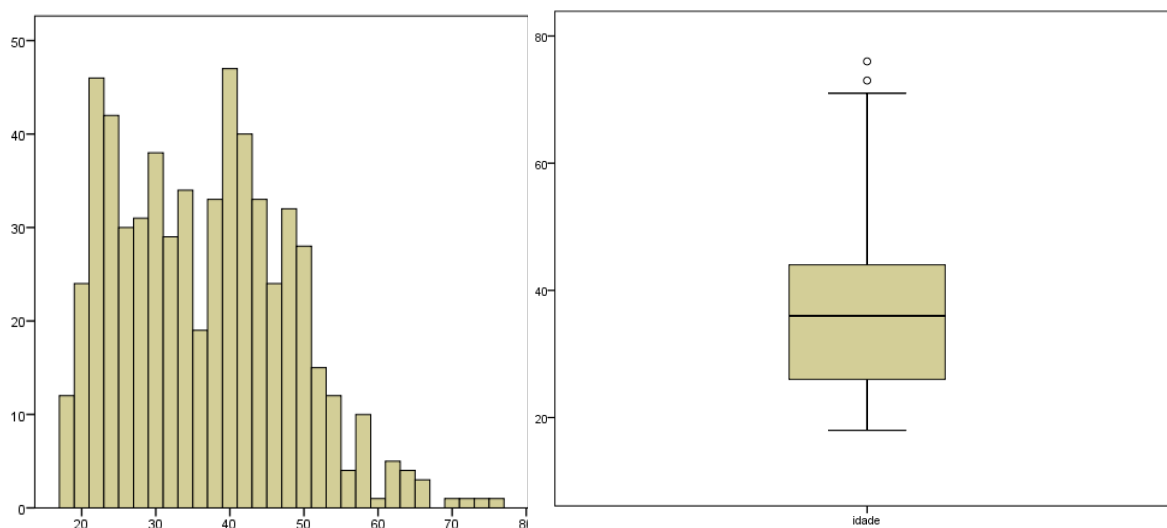


Figura 11: Distribuição dos participantes por idades, em anos (n=600)

O ajustamento à distribuição normal foi verificada pelo teste de Kolmogorov-Smirnov, com as seguintes hipóteses:

H0: A variável idade tem distribuição normal.

H1: A variável idade não tem distribuição normal.

Tabela 11: Output (SPSS) do teste de ajustamento à normalidade, para a variável idade (n=600)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	gl	Significância	Estatística	gl	Significância
Idade	.099	600	.000	.958	600	.000

^aLilliefors Significance Correction
gl – graus de liberdade

O teste de Kolmogorov-Smirnov, com $p=.000$, levou à rejeição da hipótese nula, pelo que se concluiu que a idade dos participantes não estava normalmente distribuída (Tabela 11). A análise dos gráficos *Normal Q-Q* e *Detrended Normal QQ* mostrou que o maior afastamento da distribuição normal ocorreu nas observações correspondentes a idades inferiores a 40 anos e nas idades mais elevadas.

A idade mediana deste conjunto de benguelenses era de 36 anos.

A idade avançada é considerada um indicador de risco cardiovascular e por isso foi dicotomizada de acordo com a Tabela 9. Entre os 600 benguelenses que participaram no estudo, 105 (17.5%) faziam parte do grupo de risco, em relação à idade, e 495 (82.5%) não (Figura 12).

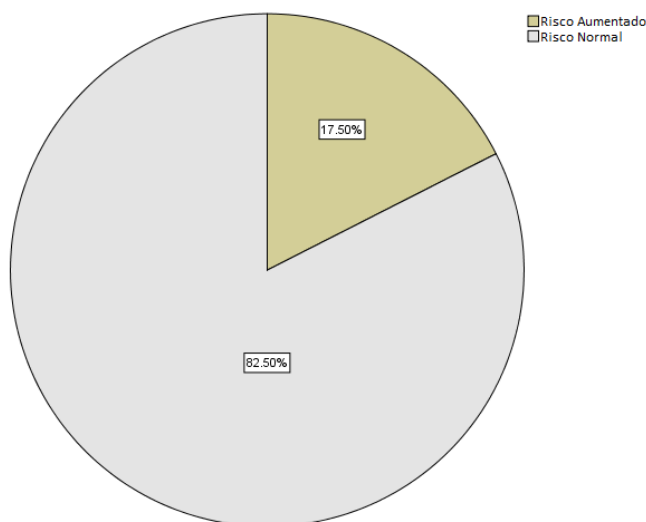


Figura 12: Distribuição dos participantes por grupo de risco, de acordo com a idade (n=600)

Os resultados do teste da binomial mostraram que o número observado de participantes em idade de risco aumentado era significativamente inferior ao número observado de participantes em idade de risco normal ($p=.000$).

2.2. Indicadores antropométricos

Peso

Os parâmetros relativos ao peso da amostra foram sintetizados no Anexo 2.

A assimetria positiva, evidenciada no histograma da Figura 13, foi confirmada pelo cálculo do quociente $0.367/0.100=3.67$ que devolveu um valor absoluto superior a 1.96, mostrando uma predominância de pesos mais baixos.

O teste de Kolmogorov-Smirnov, com $p=.000$, levou à rejeição da hipótese nula, pelo que se concluiu que o peso dos participantes não estava normalmente distribuído.

O peso mediano deste grupo de benguelenses era de 64.00kg.

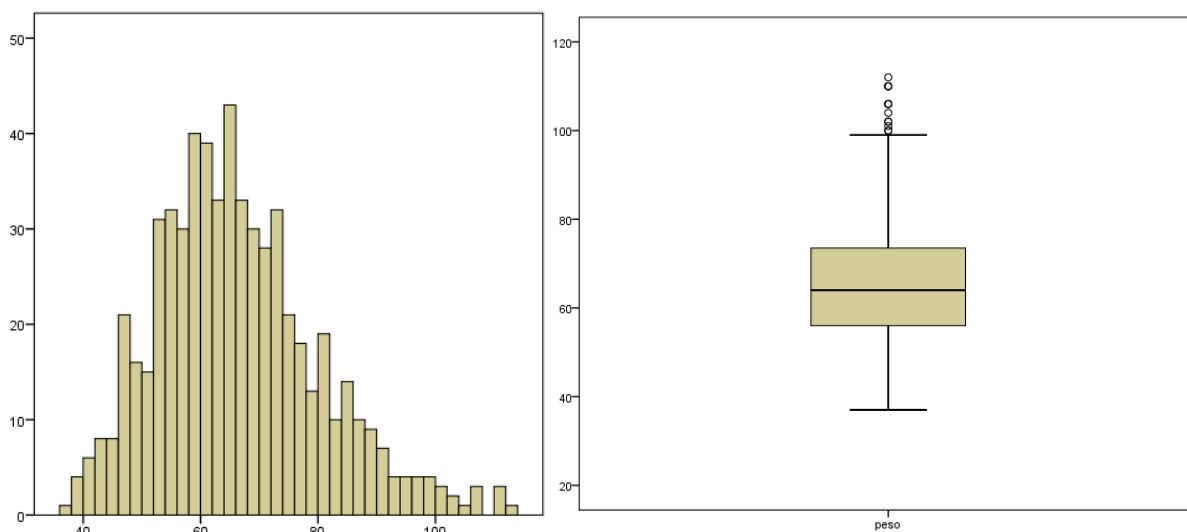


Figura 13: Distribuição dos pesos dos participantes, em kg (n=600)

Estatura

No Anexo 2 foram sintetizados os parâmetros relativos à estatura dos participantes no estudo.

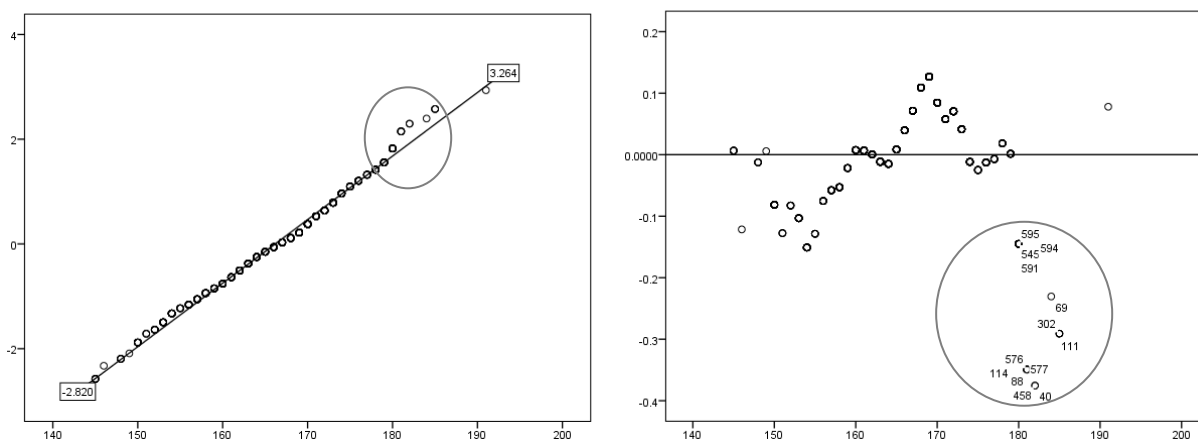


Figura 14: Normal Q-Q Plot e Detrended Normal Q-Q Plot relativos à variável estatura, em metros (n=600)

A distribuição da estatura dos participantes era simétrica, para $p=0.05$, pois o rácio $-.196/0.100=-1.96$ tinha um valor absoluto inferior ou igual a 1.96. Como o rácio $-0.396/0.199=-1.99$ tem valor absoluto superior a 1.96 a distribuição das estaturas não é mesocúrtica, como numa normal, mas sim platicúrtica. O teste de Kolmogorov-Smirnov, com $p=0.000$, levou à rejeição da hipótese nula, pelo que a estatura dos participantes não estava normalmente distribuída. A análise dos gráficos *Normal Q-Q* e *Detrended Normal*

QQ mostrou um grande afastamento da distribuição normal em observações de estaturas acima de 1.80, como se pode observar no gráfico da Figura 14.

A estatura mediana deste grupo de benguelenses era de 1.67 metros.

Índice de massa corporal

No Anexo 2 foram sintetizados os parâmetros relativos ao índice de massa corporal dos participantes no estudo.

O índice de massa corporal dos participantes variou entre 14.0 e 46.0 kg/m², equivalendo a uma amplitude de 32kg/m². O índice de massa corporal mediano foi de 23.0kg/m² e o médio foi de 23.8kg/m², com um desvio-padrão de 4.6kg/m². No gráfico *boxplot*, da Figura 15, foram identificados 11 *outliers*, sendo 3 deles considerados severos, pelo que se foi calcular a média aparada a 5%, obtendo-se um índice de massa corporal médio de 23.5kg/m², mais próximo do valor mediano.

Verificou-se, com 95% de confiança, que o verdadeiro valor médio de índice de massa corporal dos benguelenses se situava entre 23.4 e 24.1kg/m². O valor mediano estava mais próximo do extremo inferior do intervalo de confiança, indiciando uma assimetria positiva que foi confirmada pelo cálculo do rácio $1.012/0.100=10.12$. Esta assimetria positiva também foi evidenciada no histograma da Figura 15, onde se observava uma predominância de índices de massa corporal mais baixos.

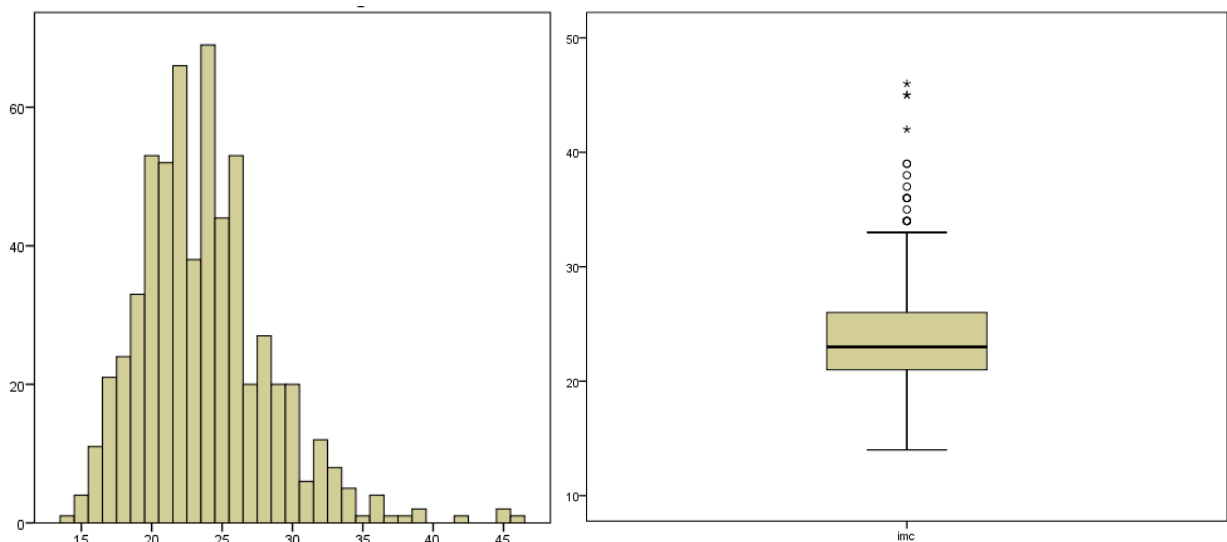


Figura 15: Distribuição dos índices de massa corporal dos participantes, em kg/m² (n=600)

O teste de Kolmogorov-Smirnov, com $p=.000$, levou à rejeição da hipótese nula, pelo que o índice de massa corporal dos participantes não estava normalmente distribuído. A análise dos gráficos *Q-Q* e *Detrended normal QQ plots* mostrou um grande afastamento da distribuição normal nas observações mais extremas, pelo que foi analisado o índice de massa corporal sem os 11 *outliers*. A distribuição manteve-se não normal ($p=.000$) e assimétrica positiva. Na Figura 16 é possível observar-se os gráficos *Q-Q* Normal obtidos com e sem os 11 *outliers*. Estas 11 observações não foram excluídas da análise, pois índices de massa corporal tão elevados eram relevantes para o objetivo deste estudo.

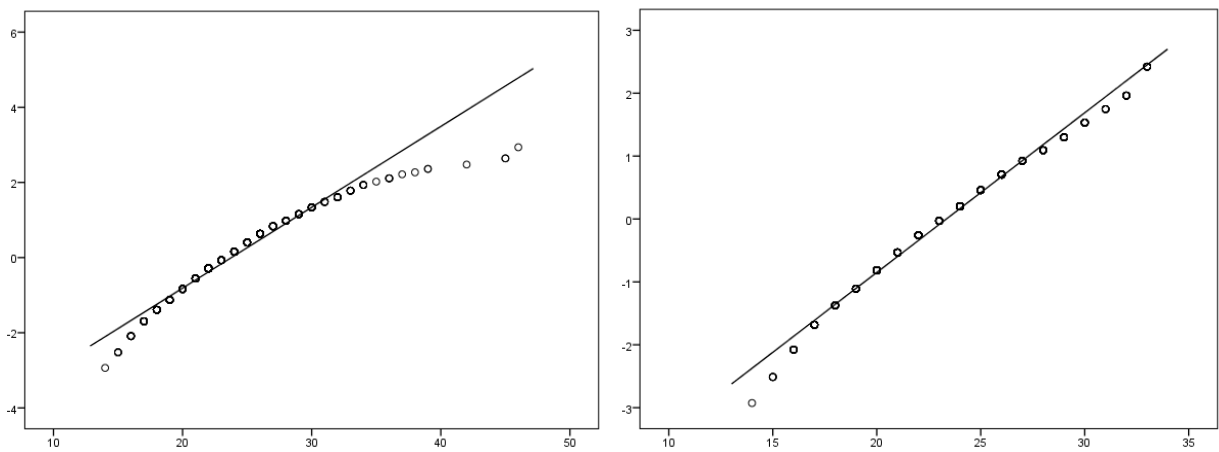


Figura 16: *Q-Q Normal Plot* com *outliers* (esquerda) e sem *outliers* (direita), para o índice de massa corporal, em kg/m^2 ($n=600$)

Para uma melhor avaliação do estado de sobrepeso (excesso de peso ou obesidade), o índice de massa corporal foi categorizado em quatro classes.

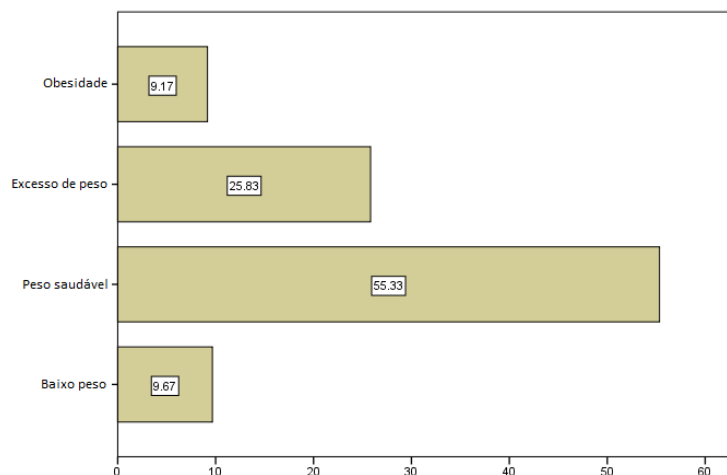


Figura 17: Distribuição dos participantes, de acordo com o índice de massa de massa corporal ($n=600$)

Observando a Figura 17, constatou-se que mais de metade dos participantes (55.3%) apresentava peso saudável e que 25.83% (155) apresentavam excesso de peso. Salientou-se o facto de a proporção de participantes com baixo peso (9.7%) ser praticamente igual à de participantes com obesidade (9.2%).

O sobrepeso é um indicador de risco cardiovascular e, por isso, houve necessidade de dicotomizar o índice de massa corporal. Entre os 600 benguelenses que participaram no estudo, 210 (35.0%) estavam com excesso de peso ou obesidade (Figura 18).

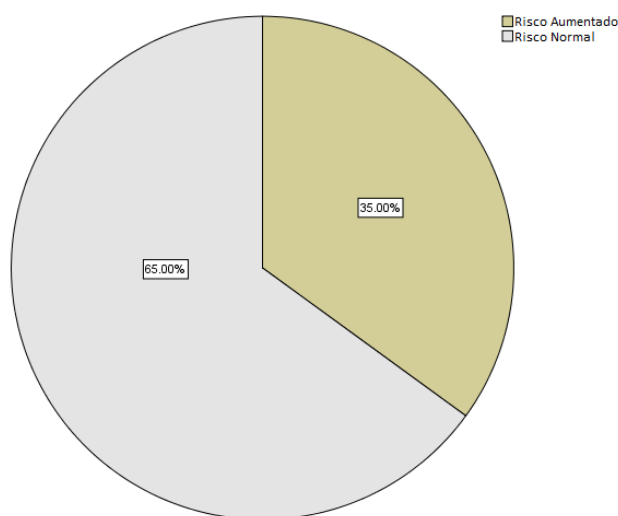


Figura 18: Distribuição dos indivíduos por grupo de risco, de acordo com o índice de massa corporal (n=600)

Os resultados do teste da binomial mostraram que o número observado de participantes com sobrepeso era significativamente inferior ao número observado de participantes sem sobrepeso, para um erro tipo I do analista de 0.05, na medida em que o nível de significância exato do teste foi .000, levando à rejeição da hipótese nula.

Perímetro abdominal

Os parâmetros relativos ao perímetro abdominal da amostra foram resumidos no Anexo 2.

Parecia haver uma assimetria positiva na distribuição dos pesos, constatada pelo cálculo do quociente $0.674/0.100=6.74$ e pelo histograma da Figura 19, mostrando uma predominância de perímetros abdominais mais baixos.

O teste de Kolmogorov-Smirnov, com $p=.000$, levou à rejeição da hipótese nula, pelo que o perímetro abdominal dos participantes não estava normalmente distribuído.

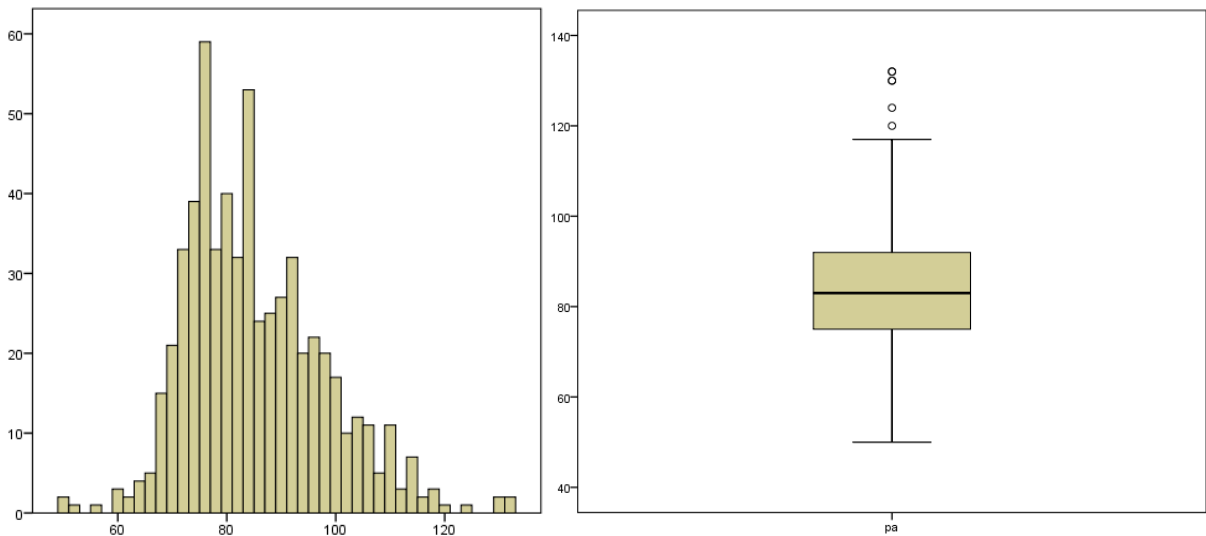


Figura 19: Distribuição do perímetro abdominal dos participantes, em cm (n=600)

O perímetro abdominal mediano deste grupo de benguelenses era de 83.00 centímetros.

Para uma melhor análise da obesidade abdominal e posterior avaliação do risco cardiovascular, o perímetro abdominal foi categorizado em três classes. A aplicação do teste de Kolmogorov-Smirnov revelou que a distribuição dos valores de perímetro abdominal não era igual pelas três classes ($p=.000$).

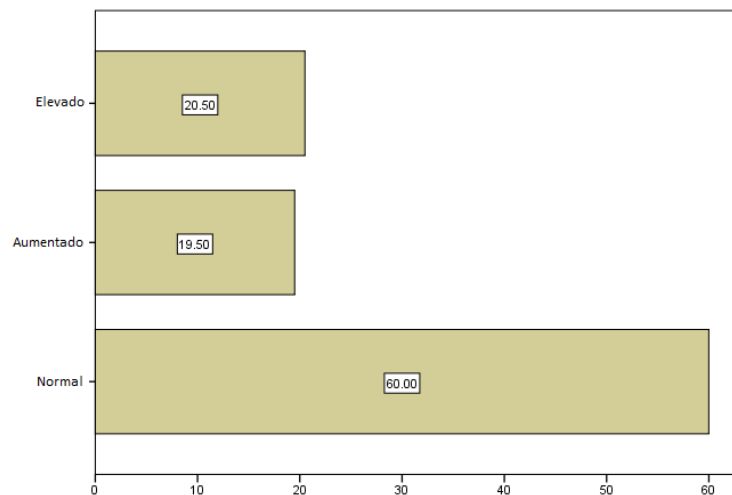


Figura 20: Distribuição dos participantes, de acordo com o perímetro abdominal (n=600)

Observando a Figura 20, constatou-se que mais de metade dos participantes (60.0%) apresentava perímetro abdominal dentro dos valores normais. Os restantes participantes

apresentavam perímetros abdominais aumentados ou elevados, correspondendo a, respetivamente, 19.5% (117) e a 20.5% (123).

O perímetro abdominal tem sido apontado como um indicador do risco cardiovascular e, por isso, foi dicotomizado. Entre os 600 benguelenses estudados, 240 (40.0%) estavam com excesso de gordura abdominal (Figura 21).

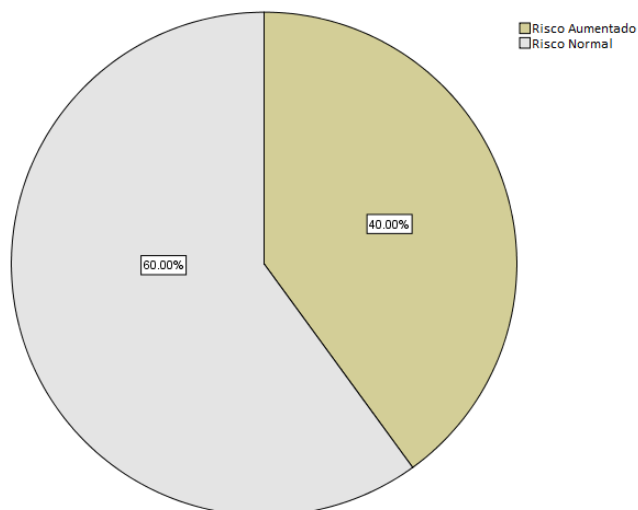


Figura 21: Distribuição dos participantes por grupo de risco, de acordo com o perímetro abdominal (n=600)

Os resultados do teste da binomial mostraram que o número observado de participantes com excesso de gordura abdominal era significativamente inferior ao de participantes sem excesso de gordura abdominal ($p=.000$).

2.3. Indicadores clínicos

Pressão arterial sistólica

Os parâmetros relativos à pressão arterial sistólica dos participantes foram sintetizados no Anexo 2.

A distribuição da pressão arterial sistólica dos participantes era assimétrica positiva ($0.933/0.100=9.33$), revelando uma predominância dos valores mais baixos (Figura 22). Foram salientados 8 indivíduos com valores elevados aberrantes (*outliers*).

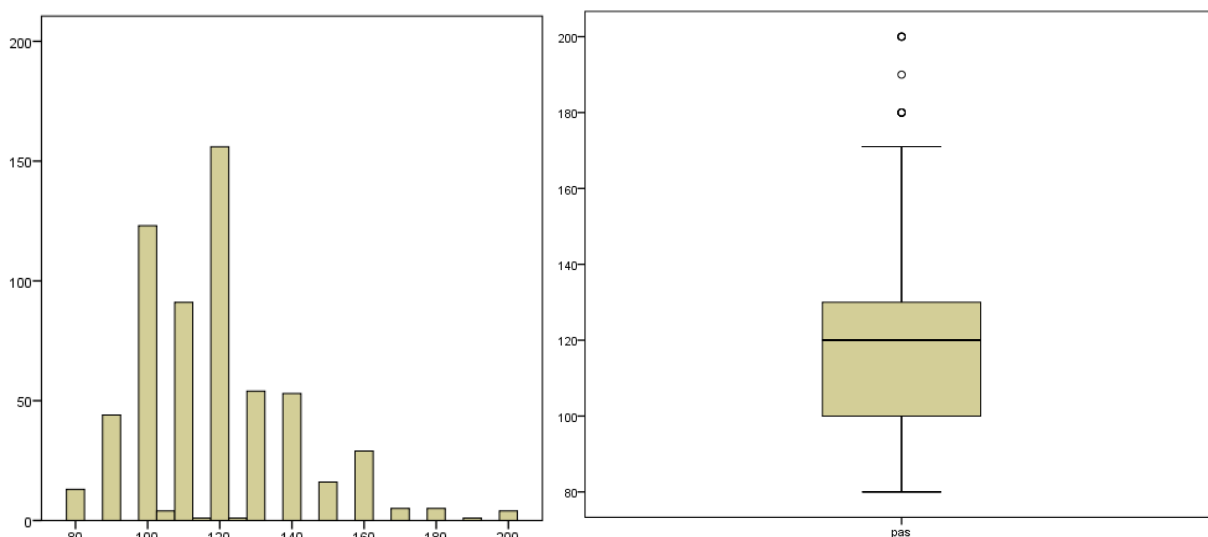


Figura 22: Distribuição da pressão arterial sistólica dos participantes, em mmHg (n=600)

O teste de Kolmogorov-Smirnov, com $p=.000$, mostrou que a pressão arterial sistólica média dos participantes não estava normalmente distribuída.

A pressão arterial sistólica mediana deste grupo de benguelenses era de 120mmHg.

Pressão arterial diastólica

Os parâmetros relativos à pressão arterial diastólica da amostra foram sintetizados no Anexo 2. No gráfico *boxplot*, da Figura 23, foram identificados 6 *outliers* superiores, entre os quais 2 eram severos.

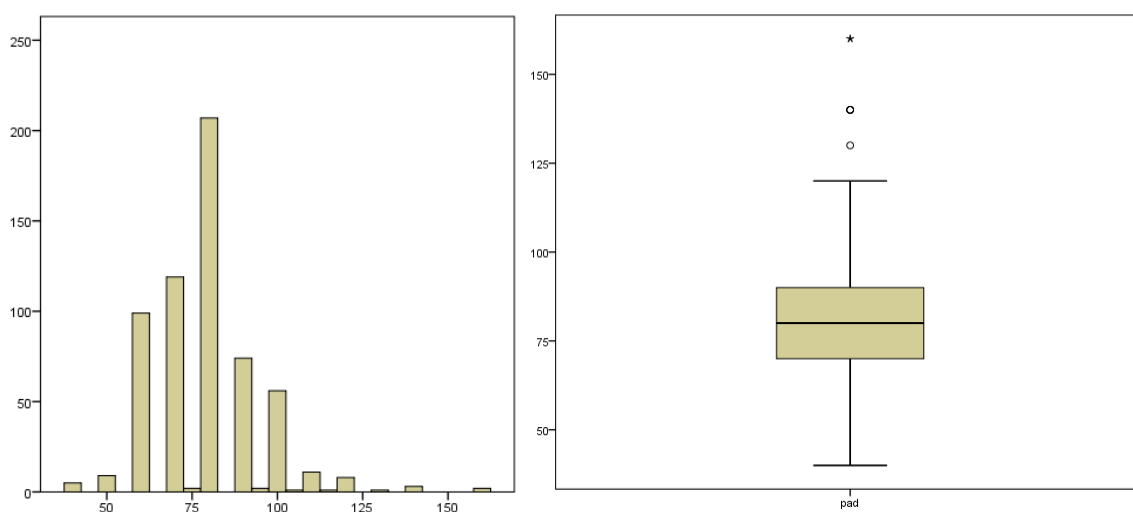


Figura 23: Distribuição da pressão arterial diastólica dos participantes, em mmHg (n=600)

A distribuição da pressão arterial diastólica dos participantes era assimétrica positiva ($0.956/0.100=9.56$), mostrando uma supremacia de valores mais baixos (Figura 23).

O teste de Kolmogorov-Smirnov, com $p=.000$, levou à rejeição da hipótese nula, pelo que a pressão arterial diastólica dos participantes não estava normalmente distribuída.

A pressão arterial diastólica mediana deste grupo de benguelenses era de 80mmHg.

Hipertensão

Observando a Figura 24 verificou-se que 13.8% (83) hipertensão arterial grau I, 10.5% (63) apresentavam hipertensão grau II, 5.0% (30) apresentavam pressão arterial normal alta e 4.7% (28) apresentavam hipertensão grau III.

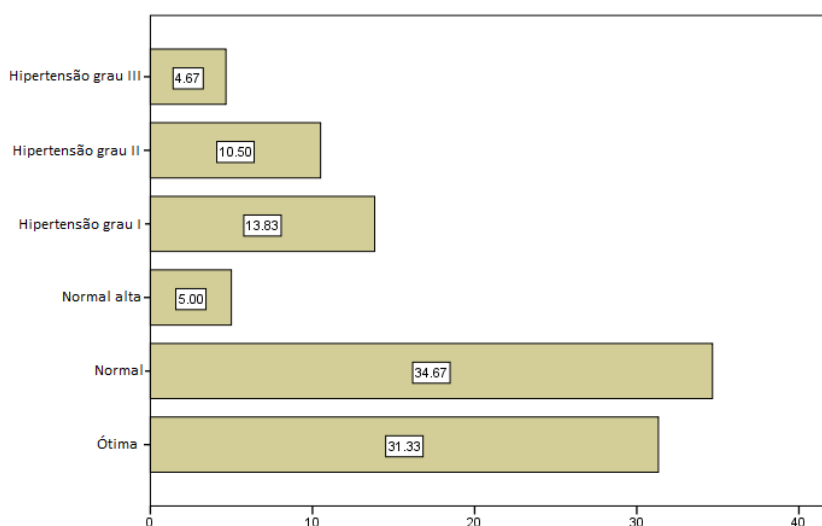


Figura 24: Distribuição dos participantes por classes, de acordo com a pressão arterial (n=600)

A condição de hipertensão é apontada como um indicador de risco cardiovascular. Entre os 600 participantes, 174 (29.0%) estavam hipertensos (Figura 25).

O teste da binomial mostrou que o número observado de hipertensos era significativamente inferior ao de normotensos ($p=.000$).

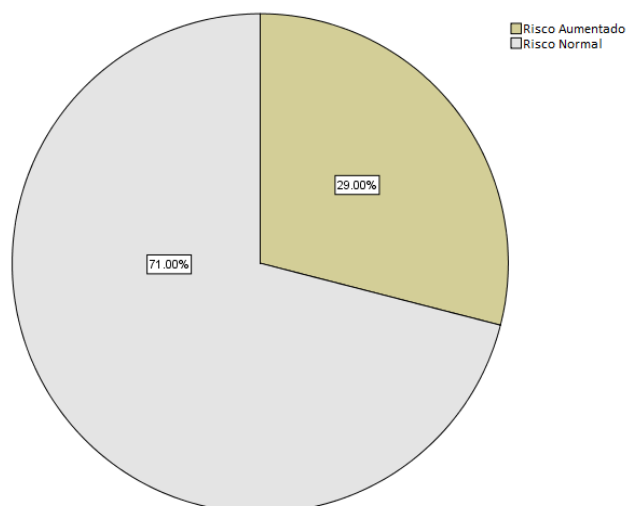


Figura 25: Distribuição dos participantes, de acordo com a condição de hipertensão (n=600)

2.4. Indicadores de estilo de vida

Entre os 600 participantes, 47 (7.8%) eram fumadores, 197 (32.8%) ingeriam bebidas alcoólicas e 500 (83.3%) consumiam gorduras saturadas (Figura 26).

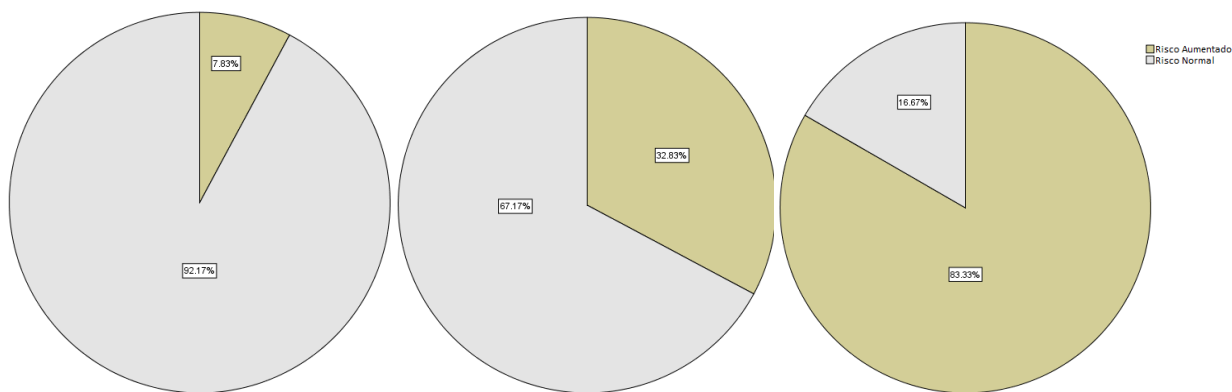


Figura 26: Distribuição dos participantes por grupo de risco, de acordo com os hábitos tabágicos, o consumo de bebidas alcoólicas e o consumo de gorduras saturadas, respetivamente (n=600)

Os resultados do teste binomial mostraram, com $p=.000$, que o número observado de fumadores era significativamente inferior ao de não fumadores; o número de participantes que ingeriam bebidas alcoólicas era significativamente inferior ao dos que não o faziam; o de participantes que consumiam gorduras saturadas era significativamente superior ao número de participantes que não o faziam.

2.5. Avaliação do risco cardiovascular

Relativamente ao risco A concluiu-se que 425 (70.8%) tinham um risco normal; 11.2% (67) tinham risco cardiovascular aumentado; 9.5% (57) tinham risco cardiovascular alto; 8.5% (51) tinham risco cardiovascular muito alto (Figura 27).

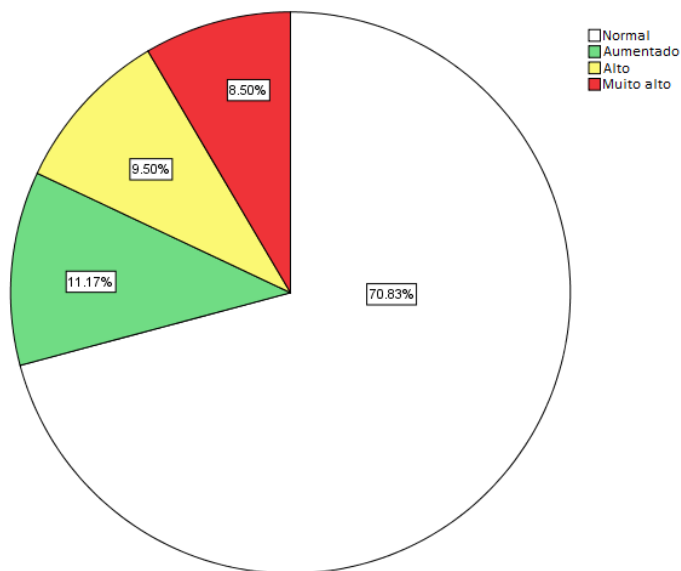


Figura 27: Distribuição dos participantes pelas categorias do risco A (n=600)

No que diz respeito ao risco B, constatou-se que 49.2% (295) tinham risco cardiovascular normal; 14.8% (89) tinham risco cardiovascular acrescido, mas baixo; 20.0% (120) tinham risco cardiovascular moderado; 14.0% (84) tinham risco cardiovascular alto; 2.0% (12) tinham risco cardiovascular muito alto.

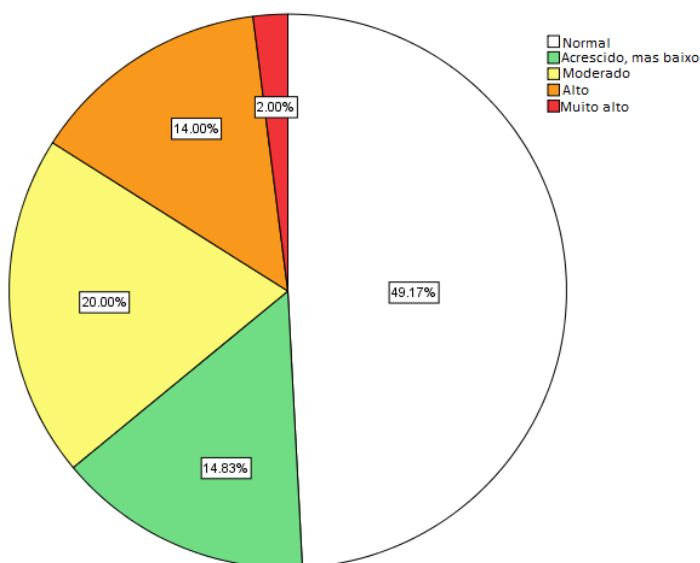


Figura 28: Distribuição dos participantes pelas categorias do risco B (n=600)

A estratificação dos participantes, de acordo com o risco C apenas foi possível nos participantes com 40 anos ou mais, ou seja, em 39.5% (237) dos indivíduos. De acordo com a Tabela 12, entre os 237 indivíduos que foram abrangidos por este modelo preditivo, 95.4% (226) tinham uma probabilidade inferior a 10% de, nos próximos dez anos, desenvolverem um evento cardiovascular, fatal ou não fatal; 3.8% (9) tinham uma probabilidade entre 10 e 20%; 0.8% (2) tinham uma probabilidade entre 20 e 30%. Nenhum dos participantes foi associado a uma probabilidade superior a 30%.

Tabela 12: Frequências absolutas e relativas da estimação do risco de nos próximos dez anos os participantes desenvolverem um evento cardiovascular, fatal ou não fatal (n=237)

Observações	Frequências	Percentagens	Percentagens válidas	Percentagens acumuladas
Válidas	<10.0%	226	37.7	95.4
	10.0-19.9%	9	1.5	99.2
	20.0-29.9%	2	.3	100.0
	Total	237	39.5	100.0
Inválidas	363	60.5		
Total	600	100.0		

A título de curiosidade, considerando todos os participantes no estudo (Figura 29), verificou-se que 37.7% (226) tinham uma probabilidade inferior a 10% de, nos próximos dez anos, desenvolverem um evento cardiovascular; 1.5% (9) tinham uma probabilidade entre 10 e 20%; 0.3% (2) tinham uma probabilidade entre 20 e 30%.

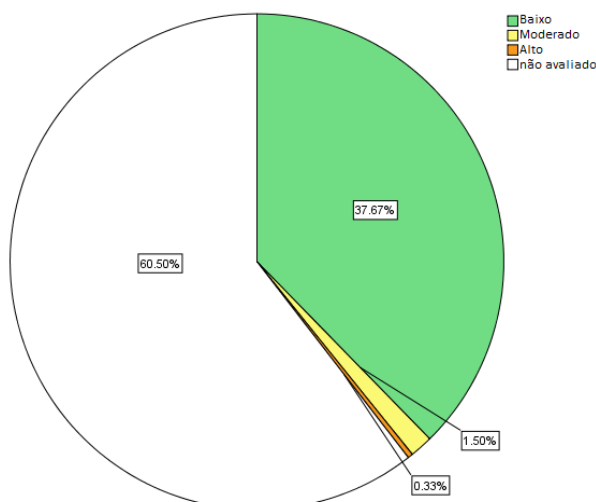


Figura 29: Estimação do risco de nos próximos dez anos os participantes desenvolverem um evento cardiovascular, fatal ou não fatal, considerando os casos avaliados (n=237) e os não avaliados (n=363)

3. ANÁLISE BIVARIADA

3.1. Comportamento dos indicadores de risco cardiovascular, por género

3.1.1. Indicadores antropométricos versus Género

Para analisar o comportamento dos indicadores antropométricos (peso; estatura; índice de massa corporal; perímetro abdominal) no grupo dos homens em comparação com o do grupo das mulheres, calcularam-se algumas medidas descritivas (Anexo 3) e construíram-se os gráficos *boxplot* da Figura 30.

O teste de ajustamento à normalidade de Kolmogorov-Smirnov mostrou que as distribuições dos pesos, das estaturas, dos índices de massa corporal e dos perímetros abdominais, não seguiam uma distribuição normal ($p=.000$, nos homens; $p=.000$, nas mulheres).

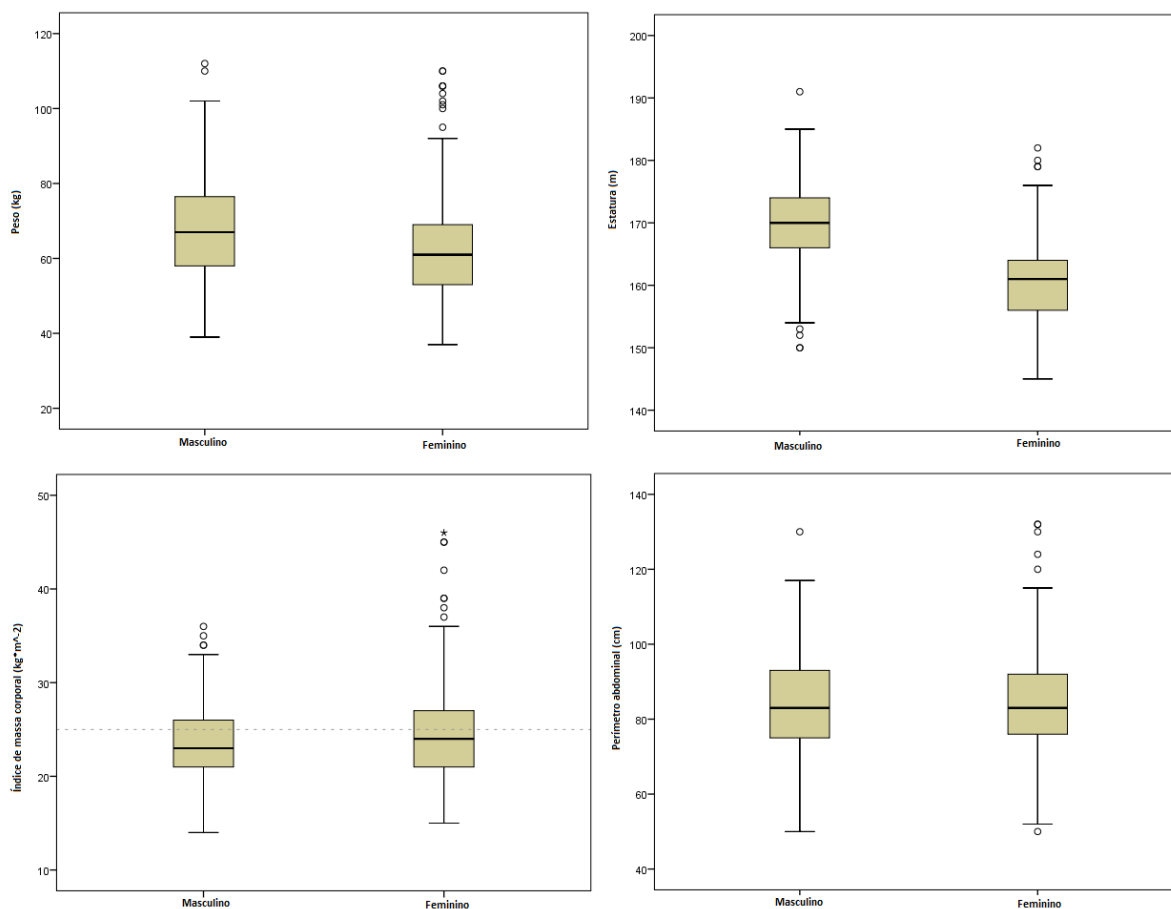


Figura 30: Distribuição dos indicadores antropométricos (peso; estatura; índice de massa corporal; perímetro abdominal), por género (n=355; n=245)

Uma vez que as dispersões pareciam ter igual forma, aplicou-se o teste não paramétrico de Mann-Whitney, para verificar se existiam diferenças significativas entre os homens e as mulheres, em relação às variáveis antropométricas. As hipóteses testadas foram:

H0: As duas populações são iguais em tendência central, relativamente às variáveis antropométricas.

H1: As duas populações não são iguais em tendência central, relativamente às variáveis antropométricas.

Tabela 13: Output, do SPSS, do teste U de Mann-Whitney, para as variáveis antropométricas (peso; estatura; índice de massa corporal; perímetro abdominal), agrupadas por género (n=600)

	Peso	Estatura	Índice de massa corporal	Perímetro abdominal
Mann-Whitney U	32712.500	14285.500	40893.500	42891.500
<i>Asymp. Sig. (2-tailed)</i>	.000	.000	.213	.775

Relativamente ao peso, os resultados apresentados na Tabela 13 (U=32712.500; p=.000) levaram à rejeição da hipótese nula. As duas distribuições eram diferentes em tendência central, pelo que se concluiu que a diferença de pesos entre os homens e as mulheres era significativa, ou seja, os homens eram significativamente mais pesados do que as mulheres (Anexo 3). Uma vez detetada uma relação entre as variáveis peso e género procedeu-se à quantificação da associação, recorrendo ao cálculo do Eta que devolveu o valor .189. Desta forma, constatou-se que o género explicava 3.6% da variação do peso, revelando uma associação muito fraca.

No que diz respeito à estatura, a Tabela 13 (U=14285.500; p=.000) mostrou que também a diferença de estatura entre os homens e as mulheres era significativa, ou seja, os homens eram significativamente mais altos do que as mulheres (Anexo 3). Recorrendo ao cálculo do Eta, considerando a estatura como variável dependente, obteve-se o valor .564. Desta forma, pode-se afirmar que 31.8% da variação da estatura era explicada pelo género, o que denota uma associação moderada.

Os resultados apresentados na Tabela 13 (U=40893.500; p=.213) mostraram que não existiam diferenças significativas entre homens e mulheres em relação ao índice de massa corporal.

A categorização do índice de massa corporal permitiu a divisão dos participantes em dois grupos: um associado a sobrepeso e outro associado a não sobrepeso. A prevalência de sobrepeso observada foi de 34.4% no grupo dos homens e de 35.9% no grupo das mulheres. O grupo com sobrepeso tinha 58.1% de mulheres e 41.9% de homens. O cálculo do *Odds Ratio* devolveu o valor .9, que confirmou não existirem diferenças significativas entre homens e mulheres, em relação à condição de sobrepeso ($p=.728$).

Os resultados apresentados na Tabela 13 ($U=42891.500$; $p=.775$) mostraram que também não existiam diferenças significativas entre homens e mulheres em relação ao perímetro abdominal.

O perímetro abdominal foi dicotomizado, dividindo os participantes em dois grupos: grupo com obesidade abdominal; grupo sem obesidade abdominal. A prevalência de obesidade abdominal foi de 24.2% no grupo dos homens e de 62.9% no grupo das mulheres. No conjunto de indivíduos com obesidade abdominal, 35.8% eram homens e 64.2% eram mulheres. Calculou-se o *Odds Ratio* ($OR=.189$) e constatou-se que o género masculino funcionava como um fator de proteção em relação ao desenvolvimento de obesidade abdominal. O *Odds Ratio*, considerando o grupo das mulheres como fator de referência, foi $OR=(154 \times 269)/(91 \times 86)=5.29$, que revelou que a probabilidade de uma mulher desenvolver obesidade abdominal era cinco vezes superior à dos homens.

3.1.2. Indicadores clínicos versus Género

Para analisar o comportamento das pressões arteriais no grupo dos homens em comparação com o do grupo das mulheres, calcularam-se algumas medidas descritivas (Anexo 3) e construíram-se os gráficos de linha, da Figura 31.

O teste de ajustamento à normalidade de Kolmogorov-Smirnov mostrou que nenhuma das pressões arteriais aferidas seguia uma distribuição normal ($p=.000$), nem nos homens nem nas mulheres.

Uma vez que, graficamente, as dispersões de cada uma das pressões aferidas pareciam ter algumas semelhanças (Figura 31), aplicou-se o teste de Mann-Whitney, que revelou que a pressão arterial sistólica ($U=31441.000$; $p=.000$) e a pressão arterial diastólica ($U=31723.500$; $p=.000$) diferiam entre homens e mulheres, no entanto, o cálculo do η^2

mostrou que as variações da pressão arterial sistólica e da pressão arterial diastólica explicadas pelo género eram, respetivamente, 4.9% e 6.0%, denotando uma associação fraca.

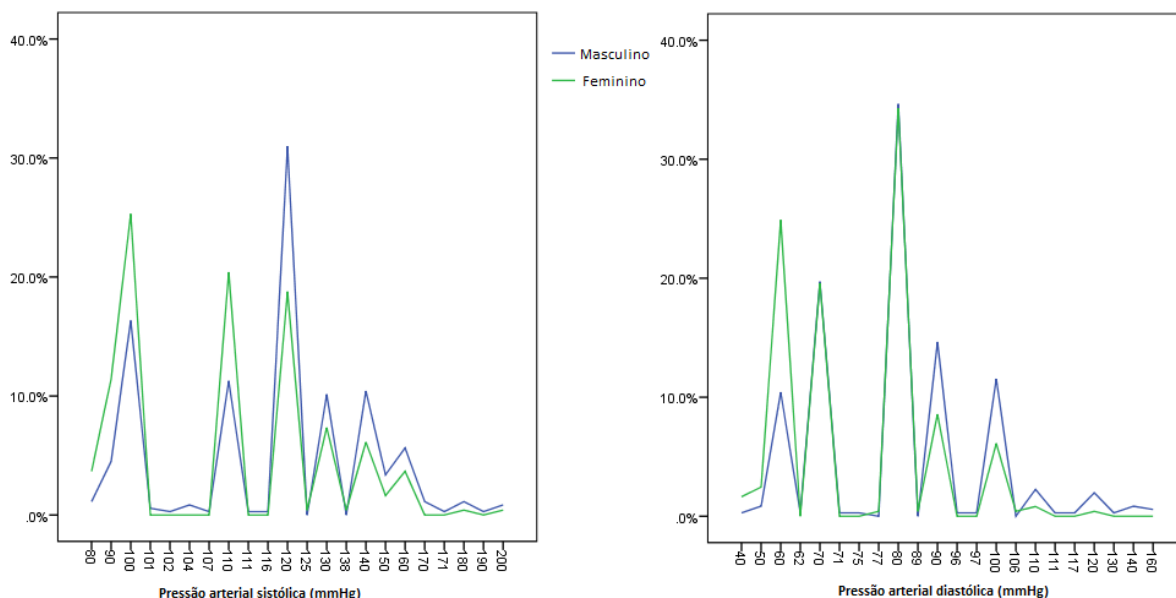


Figura 31: Distribuição dos indicadores clínicos (pressão arterial sistólica e pressão arterial diastólica), por géneros (n=355; n=245)

A conjugação destes dois indicadores origina a divisão dos participantes em dois grupos (hipertensos e não hipertensos). 36.6% dos homens estavam hipertensos e 18.0% das mulheres estavam hipertensas. O grupo de hipertensos era composto em 74.7% por homens e 25.3% por mulheres. Foi calculado o *Odds Ratio*, tendo-se obtido $OR=2.639$, pelo que se concluiu que o género masculino funcionava como um fator de risco para o desenvolvimento de hipertensão e que a probabilidade de um homem desenvolver hipertensão era quase o triplo da das mulheres.

3.1.3. Indicadores de estilo de vida *versus* Género

11.5% dos homens eram fumadores e 2.4% das mulheres eram fumadoras; 43.9% dos homens consumiam álcool e 16.7% das mulheres consumiam álcool; 84.5% dos homens consumiam gorduras saturadas e 81.6% das mulheres consumiam gorduras saturadas. As proporções atribuídas a cada um destes indicadores, por género, em relação ao total dos participantes foi sintetizada na Figura 32, na qual sobressaiu a alta prevalência de consumo de gorduras saturadas, tanto nos homens como nas mulheres.

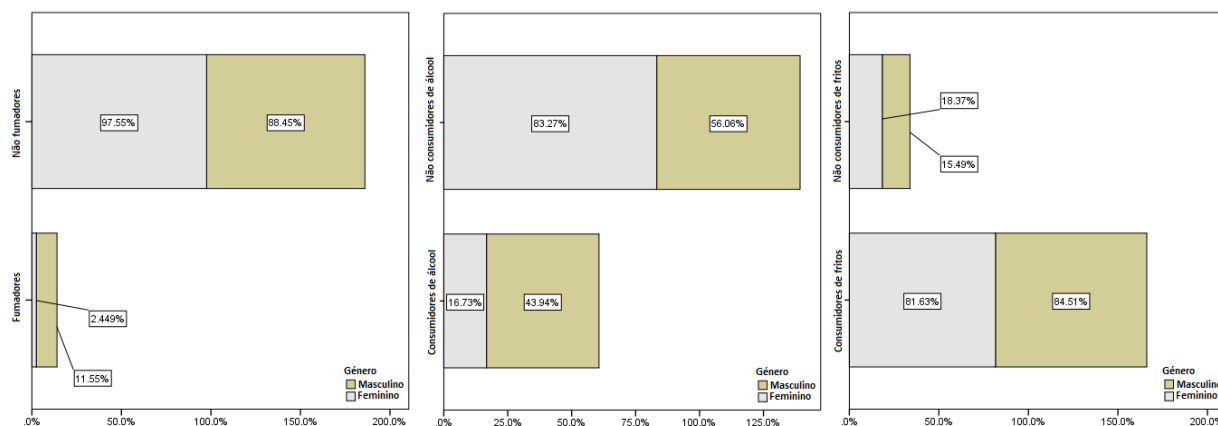


Figura 32: Distribuição dos participantes de acordo com os indicadores de estilo de vida, por género

Com o objetivo de estudar a associação entre cada um destes três indicadores de estilo de vida e o género recorreu-se ao cálculo do *Odds Ratio*, para os casos em que o teste do Qui-Quadrado apresentou um nível de significância inferior a .05 (caso em que as variáveis estavam relacionadas). Os resultados foram resumidos na Tabela 14.

Tabela 14: Resumo do teste de Qui-quadrado, combinado com o cálculo do Odds Ratio, para cada um dos indicadores de estilo de vida, agrupados por género

Indicador	Qui-quadrado de Pearson	Significância	Odds Ratio (IC 95%)
Fumadores vs não fumadores	16.628	.000	5.201 (2.172-12.453)
Consumidores de álcool vs não consumidores de álcool	48.663	.000	3.900 (2.626-5.793)
Consumidores de gorduras saturadas vs não consumidores	.862	.374	-

IC – intervalo de confiança

O género mostrou estar significativamente relacionado com os hábitos tabágicos e com o consumo de gorduras saturadas e, tanto num caso como noutro, o género masculino funcionava como um fator de risco para o desenvolvimento desses hábitos: a probabilidade de um homem fumar era cinco vezes maior do que a probabilidade de uma mulher fumar; a probabilidade de um homem consumir álcool era quatro vezes superior à de uma mulher consumir álcool. No que diz respeito ao consumo de gorduras saturadas, não foram detetadas diferenças entre géneros ($p=.374$).

3.2. Comportamento dos indicadores de risco cardiovascular, entre os mais velhos e os mais novos

3.2.1. Indicadores antropométricos *versus* Faixa etária

Os parâmetros de cada um dos indicadores, no grupo dos mais velhos (homens com 45 anos ou mais e mulheres com 55 anos ou mais) e dos mais novos, foram sintetizadas no Anexo 4.

A aplicação do teste de Kolmogorov-Smirnov (Tabela 15) mostrou que o peso, o índice de massa corporal e o perímetro abdominal dos mais velhos estão normalmente distribuídos, mas as restantes distribuições não.

Tabela 15: Output (SPSS) do teste de ajustamento à normalidade para cada uma das variáveis antropométricas, agrupadas por idade (n=600)

Variável		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estatística	gl	Significância	Estatística	gl	Significância
Peso	Mais velhos	.073	105	.200	.975	105	.048
	Mais novos	.075	495	.000	.971	495	.000
Estatura	Mais velhos	.107	105	.005	.979	105	.099
	Mais novos	.068	495	.000	.990	495	.002
IMC	Mais velhos	.073	105	.198	.989	105	.570
	Mais novos	.116	495	.000	.934	495	.000
Perímetro abdominal	Mais velhos	.058	105	.200	.988	105	.495
	Mais novos	.101	495	.000	.960	495	.000

^aLilliefors Significance Correction

gl – graus de liberdade; IMC- índice de massa corporal

Na Figura 33, foram assinalados muitos valores aberrantes no grupo dos mais novos, em relação à distribuição do peso (8 *outliers* superiores), do índice de massa corporal (11 *outliers* superiores, entre os quais 3 eram severos) e do perímetro abdominal (8 *outliers* superiores).

O teste de Mann-Whitney (U=21023.000; p=.002) mostrou existir uma relação significativa de dependência entre o peso e a idade, mas o cálculo do coeficiente Eta (.131) mostrou que era muito fraca, pois a pertença ao grupo dos mais velhos ou dos mais novos apenas explicava 1.7% da variação do peso. E também mostrou existir uma relação significativa de dependência entre a estatura e a pertença ao grupo dos mais velhos ou dos mais novos

($U=20129.000$; $p=.000$), mas também era muito fraca, pois a pertença ao grupo dos mais velhos ou dos mais novos apenas explicava 2.3% da variação da estatura ($Eta=.151$).

A dependência entre o índice de massa corporal e a faixa etária também foi validada ($U=22787.000$; $p=.047$), classificada como muito fraca, pois a pertença ao grupo dos mais velhos ou dos mais novos apenas explicava .3% da variação do índice de massa corporal.

A avaliação do índice de massa corporal permitiu a categorização dos indivíduos em dois grupos, de acordo com a condição de sobrepeso ou não. Desta forma, constatou-se que no grupo com sobrepeso mais de três quartos dos indivíduos eram do grupo mais jovem (77.1%) e que apenas 22.9% eram indivíduos mais velhos. A tabela de contingência mostrou, ainda, que 45.7% dos mais velhos e 32.7% dos mais novos estavam com sobrepeso. Com $p=.013$, foi possível afirmar que uma pessoa mais velha tinha duas vezes mais probabilidades de ter sobrepeso, quando comparada com uma pessoa mais nova, pois o *Odds Ratio* foi 1.7, com 95% de confiança na pertença ao intervalo 1.1-2.7.

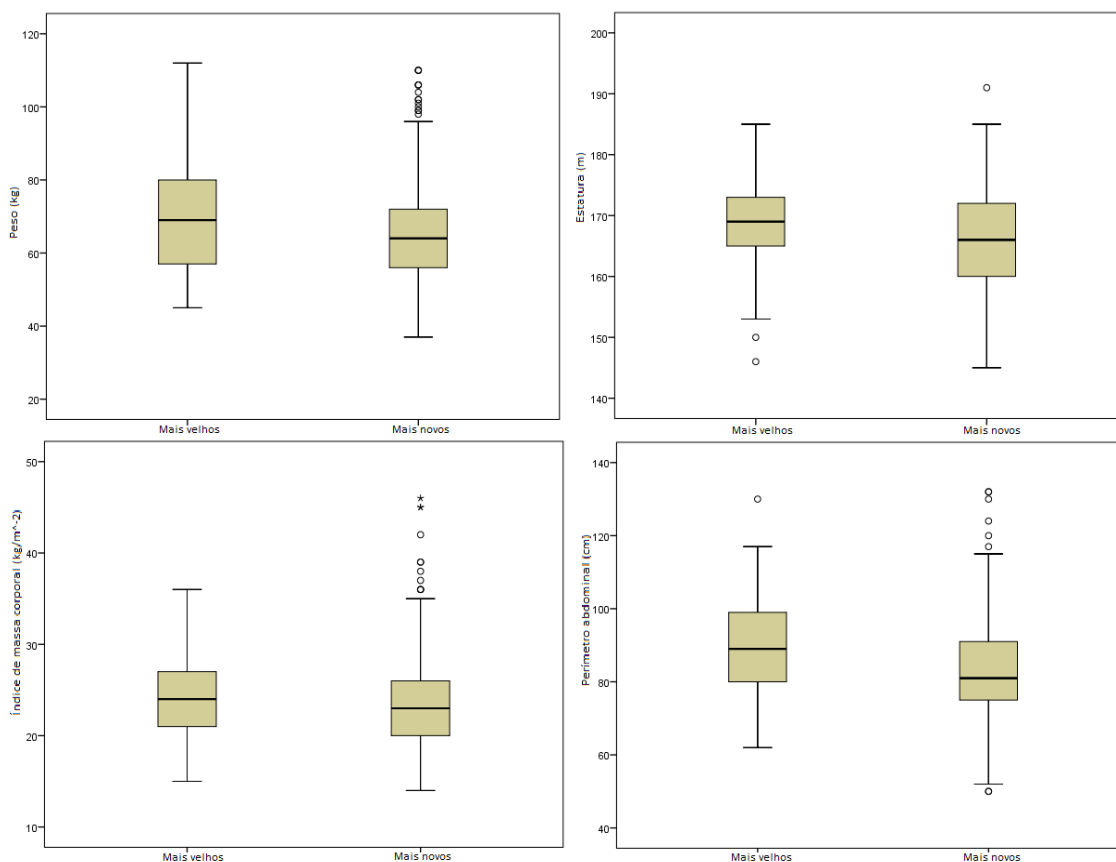


Figura 33: Distribuição dos indicadores antropométricos, no grupo dos indivíduos mais velhos e no dos indivíduos mais jovens (n=105; n=495)

O teste de Mann-Whitney ($U=18820.000$; $p=.000$) mostrou existir uma relação significativa de dependência entre os valores do perímetro abdominal e a pertença ao grupo dos mais velhos ou dos mais novos e o cálculo do coeficiente Eta permitiu a medição do grau de associação ($\text{Eta}=.180$) e mostrou que era muito fraca, pois a pertença ao grupo dos mais velhos ou dos mais novos apenas explicava 3.2% da variação dos perímetros abdominais.

Os indivíduos foram divididos em dois grupos (com obesidade abdominal *versus* sem obesidade abdominal) e constatou-se que a prevalência de obesidade abdominal era de 44.8% nos mais velhos e de 39.0% nos mais novos. A tabela de contingência revelou ainda que, no grupo com obesidade abdominal, 80.4% eram indivíduos mais novos. O *Odds Ratio* foi 1.3, com 95% de confiança na pertença ao intervalo .8-1.9, não havendo evidências estatísticas que garantissem a existência de diferenças em relação à obesidade abdominal, entre os mais velhos e os mais novos.

3.2.2. Indicadores clínicos *versus* Faixa etária

Para comparar o comportamento das pressões arteriais nos mais velhos e nos mais novos, foram calculadas algumas medidas estatísticas, por grupo (Anexo 4), e construídos os gráficos *boxplot* da Figura 34.

O *boxplot* da pressão arterial sistólica, no grupo dos mais novos, apresentou 11 *outliers* superiores, entre os quais 3 eram severos, e o da pressão arterial diastólica apresentou 14 *outliers* superiores, entre os quais 7 eram severos (Figura 34).

Os comportamentos das pressões arteriais em cada um dos grupos não pareciam ser normais, nem iguais. A aplicação dos testes de Kolmogorov-Smirnov ($p=.000$, para os quatro testes realizados) confirmaram a suspeita da não normalidade da distribuição das pressões arteriais pelos dois grupos e o teste de Mann-Whitney revelou existir relação significativa entre a pressão arterial sistólica ($U=17534.000$; $p=.000$) ou a pressão arterial diastólica ($U= 16921.000$; $p=.000$) e a pertença ao grupo dos mais novos ou dos mais velhos. Essas relações foram medidas através do cálculo do Eta que mostrou que as associações eram fracas, pois a pertença ao grupo dos mais velhos ou dos mais novos apenas explicava 5.7% da variação da pressão arterial sistólica e 6.9% da pressão arterial diastólica.

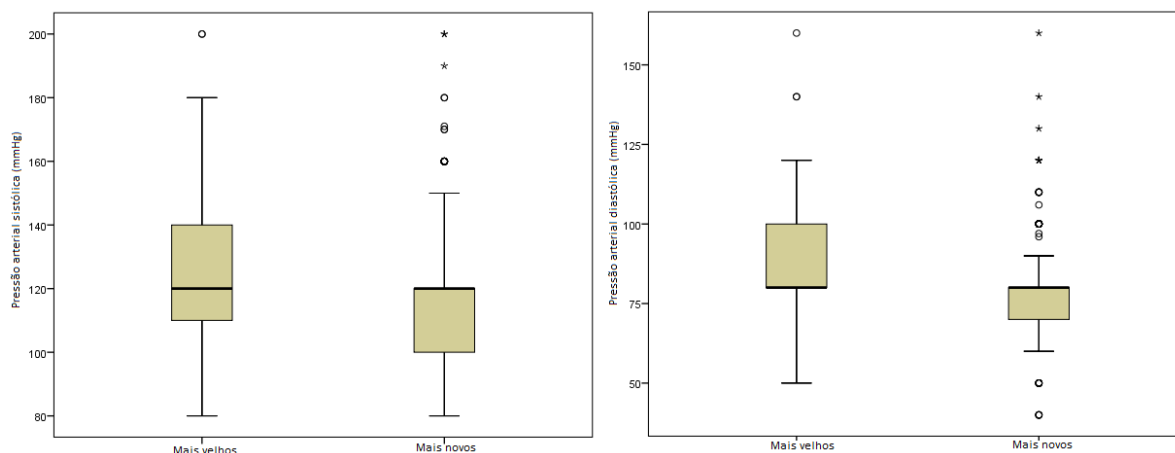


Figura 34: Distribuição dos indicadores clínicos, no grupo dos indivíduos mais velhos e no dos indivíduos mais jovens (n=105; n=495)

Em termos de pressão arterial, foi ainda estudada a relação entre os grupos hipertensos/normotensos e mais velhos/mais novos. Constatou-se que 45.7% dos mais velhos estavam hipertensos e que 32.7% dos mais novos estavam hipertensos. O grupo dos hipertensos era constituído em 31.6% por pessoas mais velhas e 68.4% por pessoas mais novas. O *Odds Ratio* foi 3.5 (2.3-5.4), mostrando que uma pessoa mais velha tinha quatro vezes mais probabilidades de ser hipertenso, do que uma pessoa mais nova.

3.2.3. Indicadores de estilo de vida versus Faixa etária

O grupo velhos/jovens foi ainda comparado com os grupos fumador/não fumador, consumidor de álcool/não consumidor e consumidor de frutos/não consumidor. Constatou-se que a prevalência de fumadores no grupo dos mais velhos era de 12.4% e no grupo dos mais jovens era de 6.9%. No grupo dos fumadores, 27.7% eram indivíduos mais velhos e 72.3% eram mais novos. O *Odds Ratio* foi 1.9 (1.0-3.8), no entanto não havia evidências estatísticas ($p=.07$) que permitissem afirmar que os mais velhos e os mais novos tinham comportamentos significativamente diferentes em relação a hábitos tabágicos.

Verificou-se, ainda, que 51.4% dos mais velhos consumiam bebidas alcoólicas, enquanto no grupo dos mais jovens a prevalência foi de 28.9%. No grupo de consumidores de bebidas alcoólicas, a prevalência de indivíduos mais jovens foi de 72.6%. O cálculo do *Odds Ratio* mostrou que a probabilidade de uma pessoa mais velha consumir álcool é o triplo da de uma pessoa mais jovem, pois, com 95% de confiança, o valor pertence ao intervalo 1.7-4.0.

No que diz respeito ao consumo de gorduras saturadas, as prevalências foram bastante altas tanto num grupo como noutro: 83.8% dos mais velhos consumiam gorduras saturadas regularmente, tal como 83.2% dos mais jovens. Analisando apenas o grupo de risco (os consumidores de gorduras saturadas) constatou-se que 82.4% do grupo era formado por indivíduos mais jovens. O *Odds Ratio* foi aproximadamente 1.0, mostrando que a probabilidade de consumir gorduras saturadas, com regularidade, não era significativamente diferente nos mais velhos e nos mais novos, o que vem de encontro ao já afirmado anteriormente.

3.3. Comportamento dos indicadores de risco cardiovascular, entre os indivíduos com sobrepeso e os sem sobrepeso

3.3.1. Indicadores antropométricos versus Sobrepeso

O teste de Kolmogorov-Smirnov, com $p=.072$, mostrou que a distribuição do perímetro abdominal no grupo com sobrepeso seguia a normal, mas no grupo sem sobrepeso não.

Para testar a relação de dependência entre cada uma das variáveis antropométricas e a pertença ao grupo com sobrepeso ou sem sobrepeso, aplicou-se o teste de Mann-Whitney (Tabela 16), que provou existirem diferenças significativas entre os pesos, os índices de massa corporal e os perímetros abdominais dos indivíduos com sobrepeso e os dos indivíduos sem sobrepeso, mas não existirem diferenças significativas entre as estaturas destes dois grupos.

Uma vez que foram detetadas três variáveis antropométricas relacionadas com a pertença ao grupo com sobrepeso ou sem sobrepeso, procedeu-se à medição do grau dessa associação através do cálculo da medida Eta, tendo-se constatado que existia uma associação forte entre o sobrepeso e o perímetro abdominal (Eta=.657), além da associação forte já esperada com o peso (.703) e com o índice de massa corporal (Eta=.771). Verificou-se, então, que 43.2% da variação do perímetro abdominal era explicada pela condição de sobrepeso.

Tabela 16: Output, do SPSS, do teste U de Mann-Whitney, para as variáveis antropométricas (peso; estatura; índice de massa corporal; perímetro abdominal), agrupadas por condição de sobrepeso

	Peso	Estatura	Índice de massa corporal	Perímetro abdominal
Mann-Whitney U	5755.000	40674.000	234.000	8585.000
Asymp. Sig. (2-tailed)	.000	.892	.000	.000

O perímetro abdominal, além de ter sido considerado uma variável contínua, também foi categorizado, dando origem a dois grupos: o grupo com obesidade abdominal e sem obesidade abdominal. A construção da tabela de contingência mostrou que 74.8% dos indivíduos com sobrepeso tinham obesidade abdominal e que 21.3% dos indivíduos com peso normal tinham obesidade abdominal. No grupo com obesidade abdominal, 65.4% dos indivíduos apresentavam sobrepeso. O *Odds Ratio* obtido foi 11.0 (7.4-16.3, com 95% de confiança), mostrando que o facto de um indivíduo ter sobrepeso fazia com que tivesse 11 vezes mais probabilidades de também ter obesidade abdominal e vice-versa.

3.3.2. Indicadores clínicos versus Sobrepeso

O teste de Kolmogorov-Smirnov, com $p=.000$, provou que nenhuma das pressões arteriais estava normalmente distribuída, nem no grupo com sobrepeso nem no sem sobrepeso. Salientaram-se os gráficos *boxplot* da Figura 35. Relativamente à pressão arterial sistólica, foram assinalados 4 *outliers* (correspondentes a indivíduos com 200mmHg) no grupo com sobrepeso e 8 *outliers* (sendo um deles severo, com 190mmHg) no grupo sem sobrepeso. No que concerne à pressão arterial diastólica foram detetados 5 *outliers* (entre os quais dois eram severos, com 160mmHg) no grupo com sobrepeso e 18 *outliers* no grupo sem sobrepeso. Entre estes 18 *outliers*, 7 eram inferiores (40mmHg) e 11 eram superiores, entre os quais 4 eram severos, com valores acima de 120mmHg.

Pela observação dos gráficos *boxplot*, pareciam existir diferenças nas pressões arteriais entre os grupos, no entanto, apenas um teste como o de Mann-Whitney poderia provar que existiam diferenças significativas. O teste de Mann-Whitney mostrou que existiam, realmente, diferenças significativas nos valores da pressão arterial sistólica ($U=29881.000$; $p=.000$) e nos valores da pressão arterial diastólica ($U=30964.000$; $p=.000$) entre os indivíduos que tinham sobrepeso e os que não tinham. Pelo cálculo do coeficiente *Eta* concluiu-se que apenas 6.0% ($.244^2$) da variação da pressão arterial sistólica era explicada

pelo sobrepeso e apenas 4.8% (.219²) da variação da pressão arterial diastólica era explicada pelo sobrepeso.

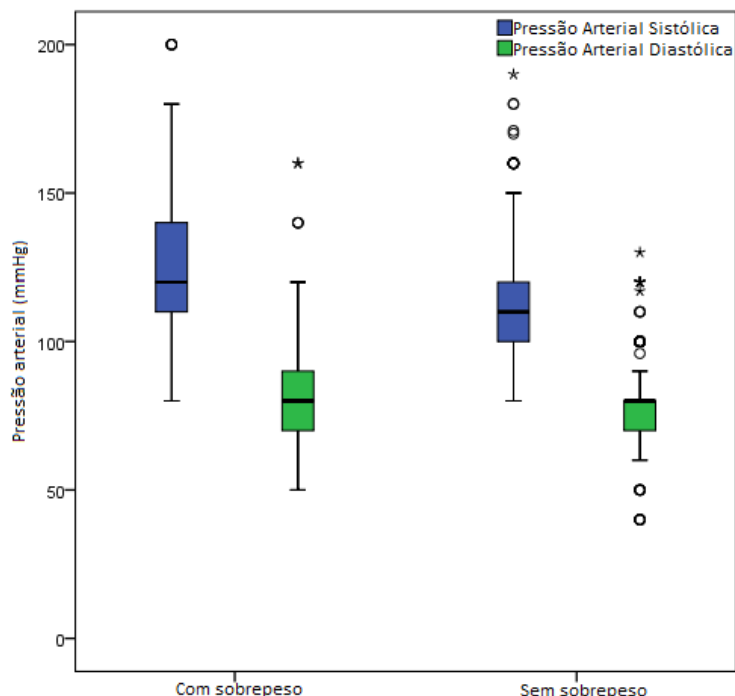


Figura 35: Distribuição da pressão arterial sistólica e da diastólica nos indivíduos com sobrepeso e sem sobrepeso (n=600)

A análise conjunta da pressão arterial sistólica e da diastólica originou a categorização dos indivíduos de acordo com a sua condição de hipertensão e constatou-se que 38.6% dos indivíduos com sobrepeso estavam hipertensos e 23.8% dos indivíduos sem sobrepeso também estavam hipertensos. No grupo com obesidade abdominal, 65.4% dos indivíduos estavam com sobrepeso. O cálculo do *Odds Ratio* devolveu o valor 2.0 (1.4-2.9), revelando que os indivíduos com sobrepeso tinham o dobro de risco de ser hipertensos, quando comparados com os que não tinham sobrepeso.

3.3.3. Indicadores de estilo de vida versus Sobrepeso

A prevalência de fumadores, entre os 210 indivíduos com sobrepeso, foi de 3.3% e no grupo sem sobrepeso foi de 10.3%. A tabela de contingência mostrou, ainda, que 85.1% dos fumadores não tinham sobrepeso. O *Odds Ratio* foi .3 (.1-.7), revelando que o consumo de tabaco parecia ser um fator de proteção em relação ao desenvolvimento de sobrepeso (p=.002). Considerando os fumadores como grupo de referência, o *Odds Ratio* foi

$OR=(40 \times 203)/(7 \times 350)=3.3$. Conclui-se, então, que os indivíduos fumadores têm três vezes mais probabilidades de não terem sobrepeso.

A prevalência de consumidores de álcool entre os indivíduos com sobrepeso foi de 36.7% e no grupo sem sobrepeso foi de 30.8%. No conjunto de indivíduos consumidores de álcool, 39.1% tinham sobrepeso e 60.9% não tinham sobrepeso. O *Odds Ratio* foi 1.3 (.9-1.9, com 95% de confiança), mostrando que não existia uma relação significativa entre as duas variáveis, ou seja, o sobrepeso era tão provável nos consumidores de álcool como nos não consumidores.

No que diz respeito ao consumo de gorduras saturadas, a sua prevalência foi de 84.3% no grupo com sobrepeso e de 82.8% no grupo sem sobrepeso. No conjunto de pessoas que consumiam regularmente fritos a maior prevalência foi dos indivíduos sem sobrepeso (64.6%). O *Odds Ratio* foi 1.1 (.7-1.8, com 95% de confiança), mostrando que não existia uma relação significativa no consumo de gorduras entre os indivíduos com sobrepeso e sem sobrepeso, ou seja, tanto os consumidores de gorduras saturadas como os não consumidores tinham igual probabilidade de terem sobrepeso.

3.4. Comportamento dos indicadores de risco cardiovascular, entre os indivíduos com obesidade abdominal e sem obesidade abdominal

3.4.1. Indicadores antropométricos versus Obesidade abdominal

Segundo o teste de Kolmogorov-Smirnov, apenas a distribuição dos pesos nos indivíduos com obesidade abdominal seguia a distribuição normal ($p=.080$).

Para averiguar se existia uma relação de dependência entre o peso e a pertença ao grupo dos indivíduos com ou sem obesidade abdominal foi aplicado o teste de Mann-Whitney que mostrou existir uma relação de dependência ($U=21428.000$; $p=.000$). O coeficiente Eta foi .457, mostrando uma relação forte entre o peso e a obesidade abdominal, pois 20.9% da variação do peso era explicada pelo facto de o indivíduo ter obesidade abdominal ou não. Também foi assinalada uma relação de dependência entre a estatura e a obesidade

abdominal ($U=35947.000$; $p=.000$), mas o coeficiente Eta mostrou que era uma relação fraca (Eta=.139).

O teste de Mann-Whitney mostrou existir uma relação de dependência entre o índice de massa corporal e a obesidade abdominal ($U=13991.000$; $P=.000$) e o coeficiente Eta mostrou que era uma relação moderada, uma vez que 32.4% ($.569^2$) da variação do índice de massa corporal era explicada pela obesidade abdominal.

3.4.2. Indicadores clínicos versus Obesidade abdominal

Os parâmetros destes indicadores, nos grupos com e sem obesidade abdominal, foram condensados no Anexo 6.

Os gráficos *boxplot*, da Figura 36, estavam repletos de *outliers* e quase todos eram elevados. Em termos de pressão arterial sistólica foram assinalados 5 *outliers* superiores, no grupo com obesidade abdominal, e 9 *outliers* (entre os quais 2 são severos e superiores a 180mmHg). Na distribuição da pressão arterial diastólica foram assinalados 3 *outliers* (sendo 1 deles severo e equivalente a 160mmHg), no grupo com obesidade abdominal, e 18 *outliers* no grupo sem obesidade abdominal, sendo 7 deles inferiores e os restantes 11 superiores (entre os quais, 5 são severos com valores superiores a 119mmHg).

As distribuições das pressões arteriais pareciam ser diferentes entre os indivíduos com obesidade abdominal e os sem obesidade abdominal, mas apenas um teste, como o de Mann-Whitney, permitiu afirmar que a pressão arterial sistólica era dependente da condição de obesidade abdominal ($U=35253.000$; $p=.000$), tal como a pressão arterial diastólica ($U=35922.000$; $p=.000$). A força destas duas associações foi medida pelo coeficiente Eta, tendo-se constatado que eram fracas, pois a obesidade abdominal apenas explicava 3.0% ($.172^2$) da variação da pressão arterial sistólica e 2.0% ($.141^2$) da variação da pressão arterial diastólica.

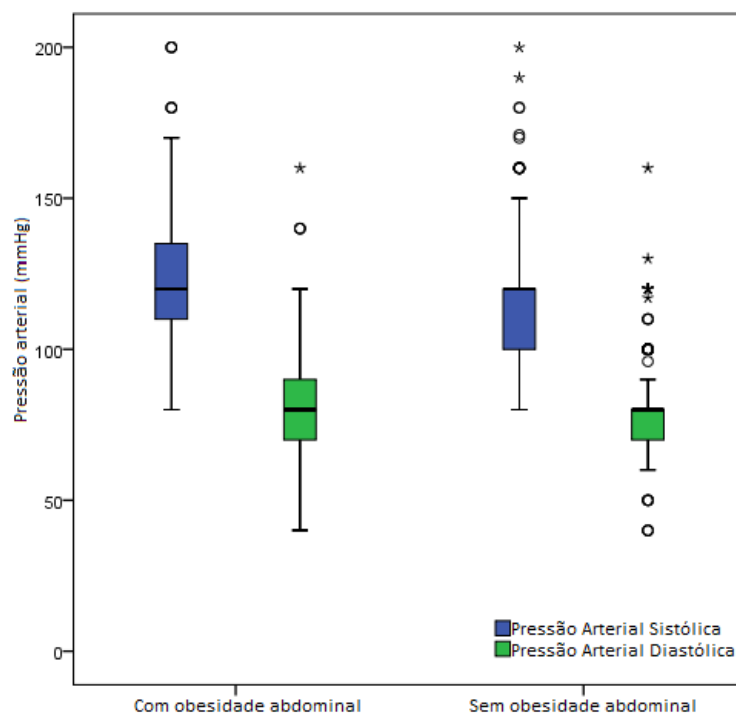


Figura 36: Distribuição da pressão arterial sistólica e da pressão arterial diastólica, de acordo com a condição de obesidade abdominal

De acordo com os valores combinados da pressão arterial sistólica e diastólica, os indivíduos foram classificados como hipertensos e normotensos. No grupo de indivíduos com obesidade abdominal, 35.0% estavam hipertensos e no grupo sem obesidade abdominal a prevalência de hipertensão era de 25.0%. O grupo de indivíduos hipertensos era composto por 48.3% de obesos e 51.7% de não obesos. Com $p=.010$, havia evidências estatísticas que garantiam a dependência entre estes dois fatores de risco. O *Odds Ratio* foi 1.6 (1.1-2.3, com 95% de confiança), mostrando que a obesidade abdominal funcionava como um fator de risco para o desenvolvimento de hipertensão e fazendo com que os obesos tivessem duas vezes mais probabilidades de desenvolver hipertensão do que os não obesos.

3.4.3. Indicadores de estilo de vida versus Obesidade abdominal

Nos indivíduos com obesidade abdominal, a prevalência de consumo de tabaco era de 4.2% e no grupo sem obesidade abdominal era de 10.3%. O grupo de fumadores era composto por 21.3% de obesos e 78.7% de não obesos. O *Odds Ratio* foi .4 (.2-.8, com 95% de confiança) revelando que a condição de obesidade abdominal predispunha o indivíduo

para o não consumo de tabaco e vice-versa. Considerando os fumadores como grupo de referência, obteve-se $OR=1/.380=2.6$, mostrando que os fumadores tinham três vezes mais probabilidades de não desenvolverem obesidade abdominal.

Em relação ao consumo de álcool, a sua prevalência foi de 31.2% no grupo com obesidade e 33.9% no grupo sem obesidade abdominal. No conjunto de consumidores de álcool encontravam-se 38.1% de indivíduos com obesidade e 61.9% sem obesidade. Com $p=.535$, não existiam evidências estatísticas para afirmar que o consumo de álcool fosse diferente entre os obesos e os não obesos ($OR=.9$).

A prevalência do consumo de gorduras saturadas era a mesma nos indivíduos obesos e não obesos (83.3%). Considerando apenas os consumidores de gorduras saturadas, constatou-se que 40% eram obesos e 60% não eram. O *Odds Ratio* foi de 1.0, confirmando que tanto os obesos como os não obesos tinham a mesma probabilidade de serem consumidores de gorduras saturadas.

3.5. Comportamento dos indicadores de risco cardiovascular, entre os hipertensos e os normotensos

3.5.1. Indicadores antropométricos versus Hipertensão

Os parâmetros de cada uma das variáveis, no grupo dos hipertensos e dos normotensos, foram sintetizados no Anexo 7. O teste de Kolmogorov-Smirnov apenas detetou uma distribuição que seguia a normal: distribuição do perímetro abdominal dos hipertensos ($p=.059$). Todas as restantes distribuições não seguiam a distribuição normal ($p<.05$).

No grupo dos normotensos, apesar de haver uma grande predominância dos pesos mais baixos, foram identificados 7 *outliers* superiores (Figura 37). O teste de Mann-Whitney mostrou existir uma relação de dependência entre o peso e a hipertensão ($U=26376.000$; $p=.000$), no entanto o coeficiente Eta mostrou que era uma associação fraca ($Eta=.248$). O mesmo se passou com a estatura ($Eta=.195$).

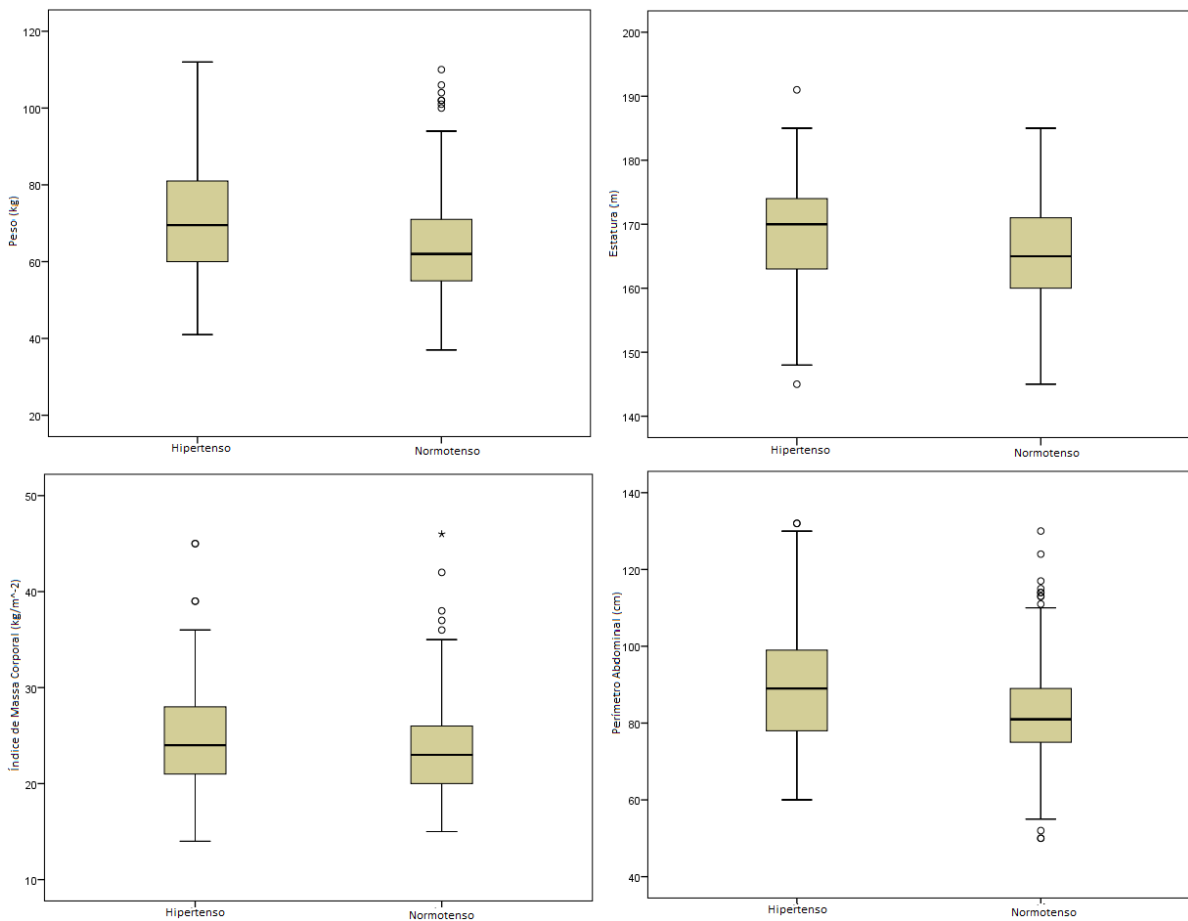


Figura 37: Distribuição dos indicadores antropométricos, no grupo dos hipertensos e dos normotensos (n=174; n=426)

No *boxplot* relativo ao índice de massa corporal dos hipertensos (Figura 38) foram assinalados 4 *outliers* superiores, relativos a indivíduos com índices acima de 38kg/m^2 , por seu lado, no grupo dos normotensos, foram assinalados 5 *outliers*, com valores acima dos 35kg/m^2 , sendo que um deles foi considerado severo. O teste de Mann-Whitney comprovou a existência de uma relação de dependência entre o índice de massa corporal e a hipertensão ($U=30160.000$; $p=.000$), porém o coeficiente Eta mediou essa associação como sendo fraca ($\text{Eta}=.165$).

Relativamente ao perímetro abdominal, o *boxplot* dos normotensos destacou-se com um grande número de *outliers*, sendo 3 inferiores e 7 superiores (Figura 38). O teste de Mann-Whitney mostrou existir uma relação significativa entre o perímetro abdominal e a hipertensão, no entanto o coeficiente Eta revelou que era uma relação fraca ($\text{Eta}=.240$). Os

boxplot mostram, ainda, que os valores medianos dos hipertensos, para cada uma das variáveis, é sempre superior ao valor mediano dos normotensos.

3.5.2. Indicadores clínicos versus Hipertensão

A distribuição da pressão arterial sistólica dos normotensos era simétrica, platicúrtica e no *boxplot* foram assinalados 10 *outliers*, sendo que 1 era inferior e os restantes 10 eram superiores, entre os quais havia 4 considerados severos (Figura 38). A distribuição da pressão arterial diastólica dos normotensos era assimétrica negativa e mesocúrtica (Anexo 7).

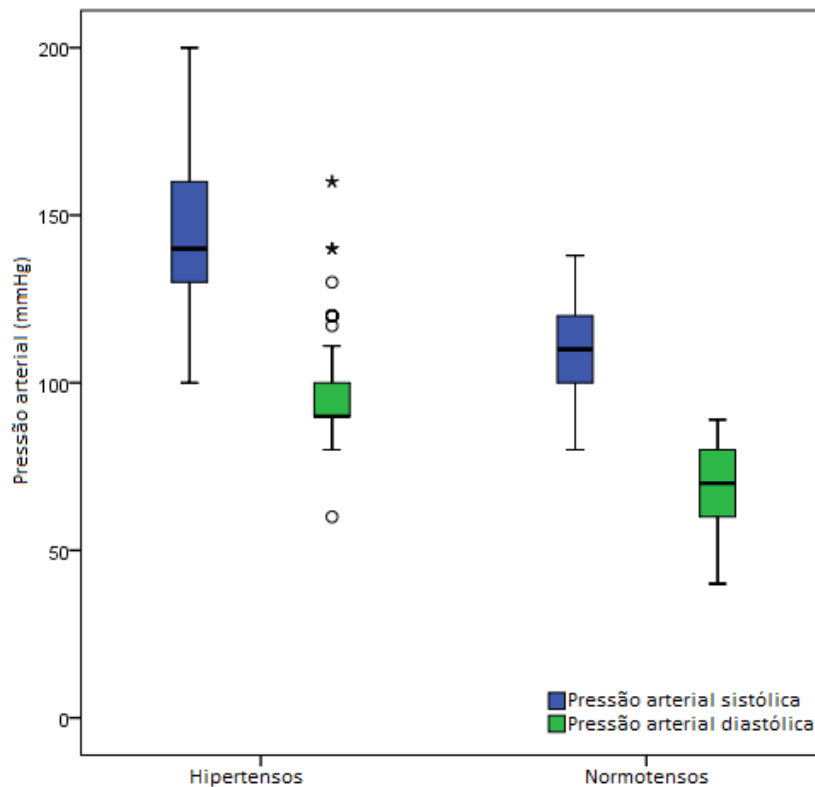


Figura 38: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos hipertensos e dos normotensos (n=600)

As distribuições das pressões arteriais pareciam ser dependentes da condição de hipertensão. A desconfiança anterior foi confirmada pelo teste de Mann-Whitney que mostrou que a pressão arterial sistólica ($U=4710.000$; $p=.000$) e a pressão arterial diastólica ($U=95661.000$; $p=.000$) estavam relacionadas com condição de hipertensão. O coeficiente Eta foi usado para medir o grau de associação e mostrou que 51.3% ($\text{Eta}=.716$) da variação da pressão arterial sistólica era explicada pela condição de hipertensão e 54.5% ($\text{Eta}=.738$)

da variação da pressão arterial sistólica era explicada pela condição de hipertensão, revelando associações fortes.

3.5.3. Indicadores de estilo de vida versus Hipertensão

A prevalência de fumadores entre os hipertensos foi de 10.3% e entre os normotensos foi de 6.8%. Considerando apenas o grupo de fumadores, verificou-se que 38.3% estavam hipertensos e 61.7% estavam normotensos. Com $p=.179$, não havia evidências estatísticas para afirmar que existiam diferenças significativas entre os hipertensos e os normotensos, em relação aos hábitos tabágicos ($OR=1.6$).

Em relação ao consumo de álcool, a prevalência foi maior no grupo dos hipertensos, com 39.7%. No grupo dos normotensos a prevalência de consumo de bebidas alcoólicas foi de 30.0%. Analisando apenas os consumidores de álcool. Verificou-se que 35.0% estavam hipertensos e 65.0% estavam normotensos. Com $p=.027$, existiam evidências estatísticas para afirmar que o consumo de álcool era diferente nestes dois grupos. O *Odds Ratio* foi de 1.5 (1.1-2.2, com 95% de confiança) revelando que os consumidores de álcool tinham duas vezes mais probabilidades de serem hipertensos do que os não consumidores de álcool.

Mais uma vez, o consumo de gorduras saturadas atingiu altas prevalências, sendo de 85.6% no grupo dos hipertensos e de 82.4% no grupo dos normotensos. Estudando apenas os consumidores de gorduras saturadas, verificou-se que 29.8% estavam hipertensos e 70.2% estavam normotensos. Com $p=.398$, não se verificou uma relação significativa entre a condição de hipertensão e o consumo de gorduras saturadas ($OR=1.3$), ou seja, a condição de hipertensão não era explicada pelo fato do indivíduo ingerir gorduras saturadas ou não.

3.6. Comportamento dos indicadores de risco cardiovascular, entre os fumadores e os não fumadores

3.6.1. Indicadores antropométricos versus Tabagismo

Os parâmetros calculados para cada uma das variáveis, para os fumadores e para os não fumadores, foram condensados no Anexo 8.

No gráfico *boxplot* dos não fumadores foram assinalados 5 *outliers*, com pesos superiores a 100kg (Figura 39). O teste de Mann-Whitney mostrou que existia uma relação significativa de dependência entre o peso e os hábitos tabágicos ($U=10715.000$; $p=.046$), no entanto o cálculo do coeficiente Eta identificou esta associação como sendo muito fraca (Eta=.081).

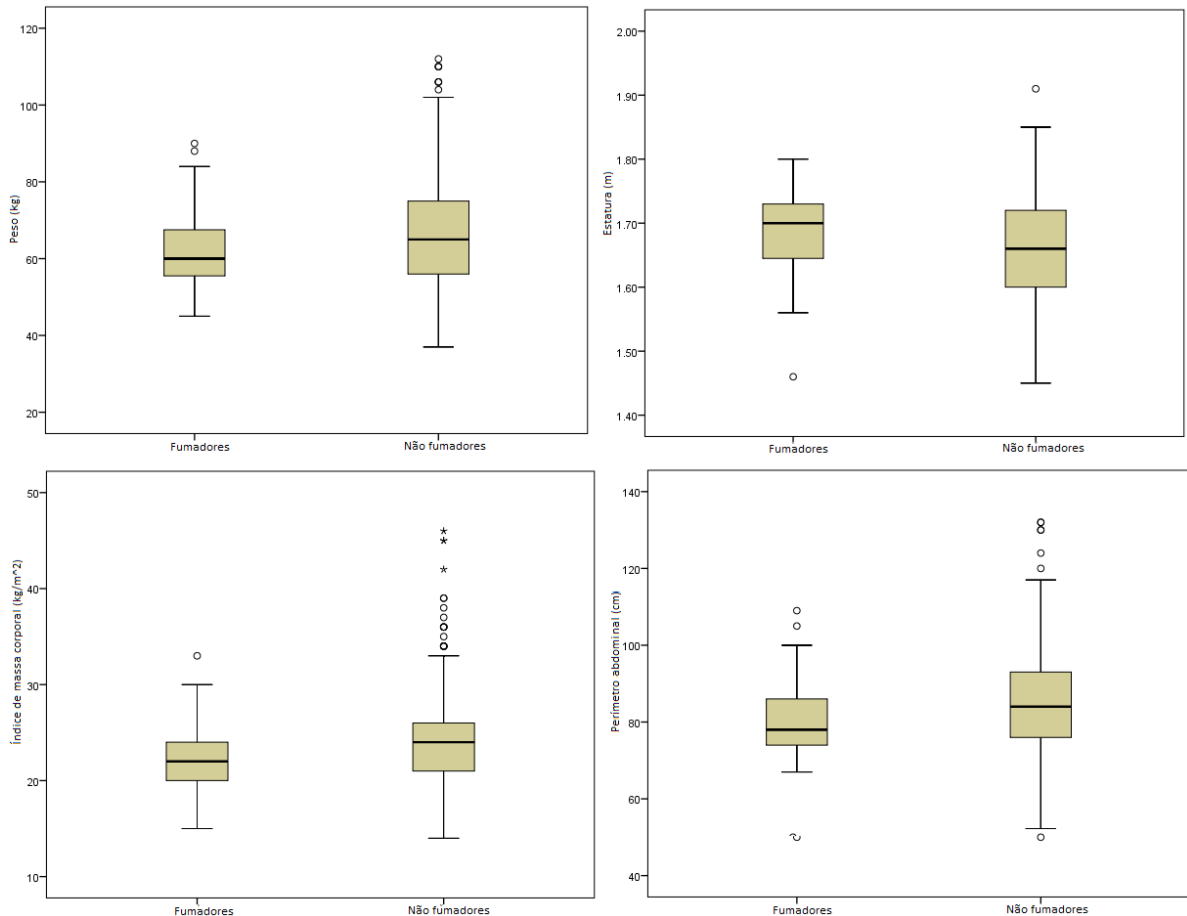


Figura 39: Distribuição das medidas antropométricas, nos fumadores e não fumadores (n=47; n=553)

O teste de Mann-Whitney mostrou evidências estatísticas ($p=.014$) que permitiram constatar a existência de uma relação de dependência entre a estatura e os hábitos tabágicos, no entanto o coeficiente Eta mostrou que essa associação era fraca (Eta=.093).

O *boxplot* da distribuição do índice de massa corporal dos não fumadores apresentou 15 *outliers*, sendo que 4 eram de nível severo, com índices acima de 41kg/m^2 (Figura 39). O teste de Mann-Whitney mostrou existir uma relação de dependência entre o índice de massa corporal e os hábitos tabágicos ($U=9220.500$; $p=.001$), no entanto o coeficiente Eta avaliou essa relação como sendo fraca (Eta=.128).

O *boxplot* da distribuição do perímetro abdominal dos não fumadores destacou-se por ter 6 *outliers* superiores, correspondentes a indivíduos não fumadores com perímetros abdominais iguais ou superiores a 120cm. O teste de Mann-Whitney mostrou existir uma relação de dependência entre os perímetros abdominais e os hábitos tabágicos ($U=9807.500$; $p=.005$), no entanto o coeficiente Eta revelou que essa associação era fraca ($\text{Eta}=.109$). O teste de Kolmogorov-Smirnov mostrou que nenhuma das duas distribuições anteriores seguia a distribuição normal ($p<.005$).

3.6.2. Indicadores clínicos versus Tabagismo

As distribuições das pressões arteriais, tanto dos fumadores como dos não fumadores, mostraram uma prevalência dos valores mais baixos (Anexo 8) e o teste de Kolmogorov-Smirnov mostrou que as distribuições das pressões arteriais dos dois grupos não seguiam a distribuição normal, com $p=.000$. Os gráficos *boxplot* da distribuição das pressões arteriais, tanto dos fumadores como dos não fumadores, evidenciaram a existência de muitos *outliers* superiores (Figura 40).

O teste de Mann-Whitney mostrou não haver evidências estatísticas para afirmar que os comportamentos das pressões arteriais sistólica ($U=11955.500$; $p=.354$) e diastólica ($U=11108.500$; $p=.089$) estavam relacionados dos hábitos tabágicos.

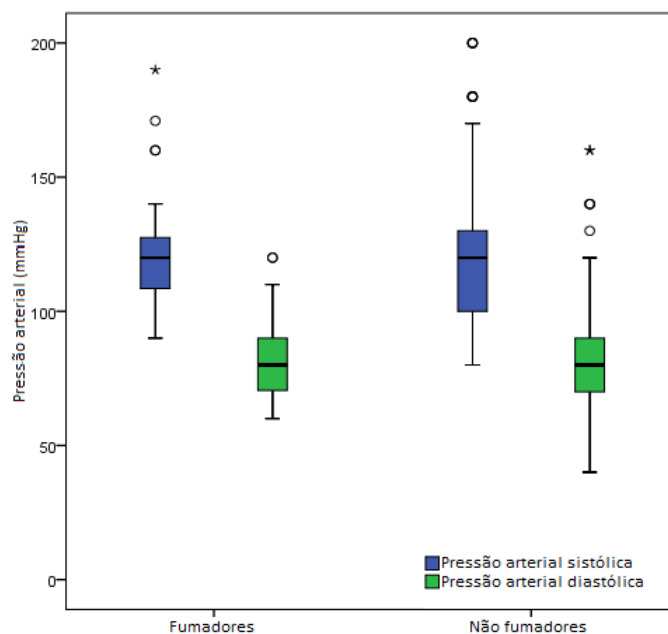


Figura 40: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos fumadores e dos não fumadores (n=47; n=553)

3.6.3. Indicadores de estilo de vida versus Tabagismo

A prevalência de consumo de álcool no grupo dos fumadores foi de 66.0% e menos de metade (30.0%) no grupo dos não fumadores. Analisando apenas os consumidores de álcool verificou-se que 15.7% eram também fumadores e os restantes 84.3% não eram fumadores. Com $p=.000$, existiam evidências estatísticas para considerar uma relação entre estes dois fatores. O *Odds Ratio* foi 4.5 (2.4-8.5, com 95% de confiança) e mostrou que o consumo de bebidas alcoólicas era cinco vezes mais provável nos fumadores, do que nos não fumadores.

Quanto ao consumo de gorduras saturadas, as prevalências foram de 89.4% no grupo dos fumadores e 82.8% no grupo dos não fumadores. Analisando apenas os consumidores de gorduras saturadas, verificou-se que 8.4% eram fumadores e 91.6% não era fumadores. Com $p=.311$, não se verificou uma relação significativa entre os hábitos tabágicos e o consumo de gorduras saturadas.

3.7. Comportamento dos indicadores de risco cardiovascular, entre os consumidores de bebidas alcoólicas e os não consumidores.

3.7.1. Indicadores antropométricos versus Álcool

Tanto para os consumidores de álcool como para os não consumidores, foram determinados parâmetros estatísticos, sintetizados no Anexo 9. A normalidade da distribuição do peso, da estatura, do índice de massa corporal e do perímetro abdominal foi testada pelo teste de Kolmogorov-Smirnov que mostrou que nenhuma das distribuições seguia a normal ($p<.05$), em nenhum dos grupos.

Entre os *boxplot*, da Figura 41, destacaram-se o da distribuição dos pesos dos não consumidores de álcool, com 8 *outliers* relativos a indivíduos com mais de 100kg, e os das distribuições dos índices de massa corporal. No grupo dos consumidores de álcool foram assinalados 4 *outliers*, sendo 1 severo, e no grupo dos não consumidores assinalados 10 *outliers*, sendo 2 severos.

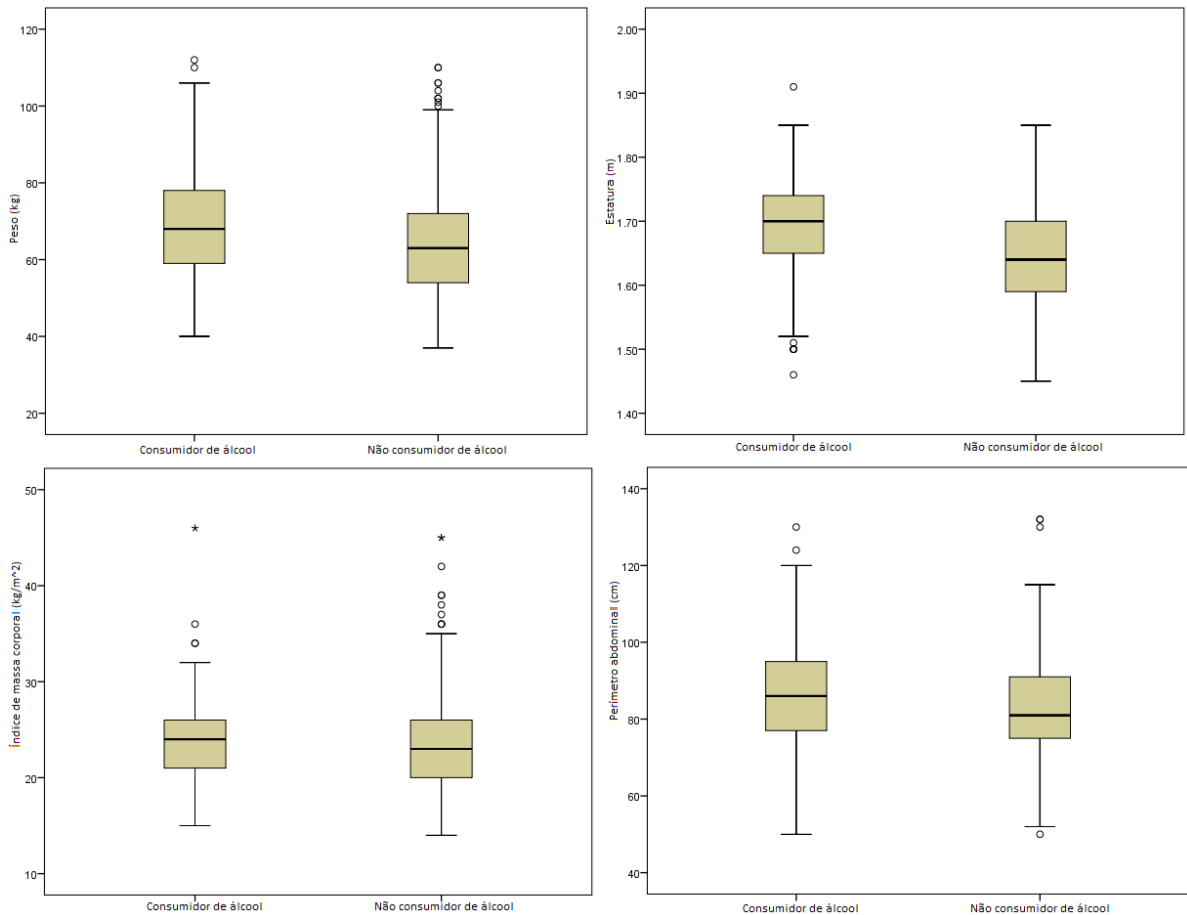


Figura 41: Distribuição dos indicadores antropométricos nos consumidores de álcool e nos não consumidores (n=197; n=403)

O teste de Mann-Whitney apontou uma relação entre o peso ($U=31629.500$; $p=.000$), a estatura ($U=26395.500$; $p=.000$), o perímetro abdominal ($U=33692.500$; $p=.003$) e o facto de consumir ou não álcool, no entanto as três relações foram classificadas como fracas ($Eta=.158$; $Eta=.260$; $Eta=.124$). Não foi assinalada qualquer relação de dependência entre o índice de massa corporal e o facto de o indivíduo consumir, ou não, álcool ($U=36743.500$; $p=.138$).

3.7.2. Indicadores clínicos versus Álcool

As pressões arteriais foram descritas para o grupo dos consumidores de álcool e para o grupo dos não consumidores (Anexo 9). A normalidade das quatro distribuições foi testada pelo teste de Kolmogorov-Smirnov que provou que nenhuma delas seguia a normal ($p<.05$). Foram, ainda, construídos gráficos *boxplot* para cada uma das distribuições (Figura

42), nos quais se destacaram os inúmeros *outliers* relativos a pressões arteriais muito elevadas, dentro de cada grupo.

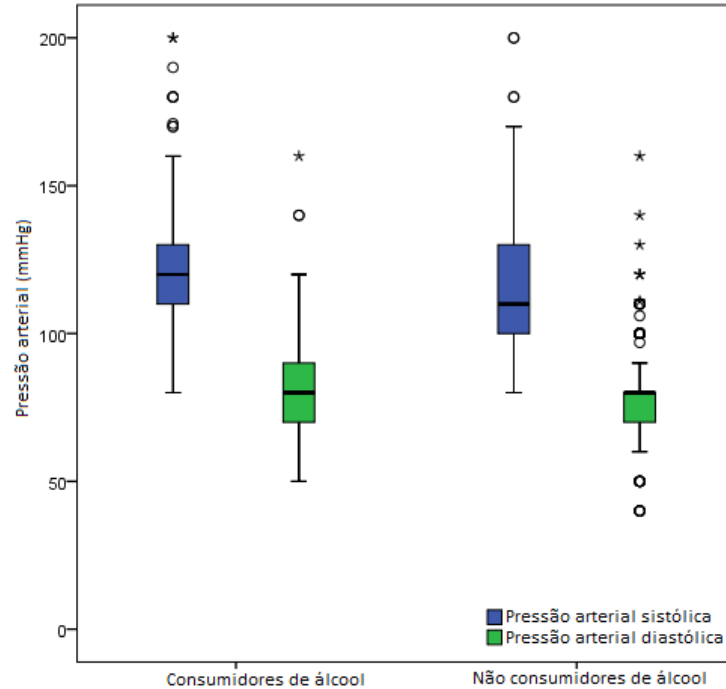


Figura 42: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos consumidores de álcool e dos não consumidores (n=197; n=403)

O teste de Mann-Whitney mostrou haver evidências estatísticas para afirmar que os comportamentos das pressões arteriais sistólica ($U=33465.000$; $p=.002$) e diastólica ($U=33326.500$; $p=.001$) eram dependentes do consumo de álcool, no entanto o coeficiente Eta mostrou que essas duas associações eram fracas (Eta=.123 e Eta=.138, respetivamente).

3.7.3. Indicadores de estilo de vida *versus* Álcool

Em relação ao consumo de gorduras saturadas, tanto os consumidores de álcool como os não consumidores apresentaram grandes prevalências (89.3% e 40.4%, respetivamente). Considerando apenas os consumidores de gorduras saturadas, verificou-se 35.2% ingeriam bebidas alcoólicas e 64.8% não. Com $p=.005$, constatou-se que existia uma relação significativa entre estes dois fatores. Como o *Odds Ratio* foi de 2.0 (1.2-3.4), os consumidores de gorduras saturadas tinham duas vezes mais probabilidades de também serem consumidores de bebidas alcoólicas.

3.8. Comportamento dos indicadores de risco cardiovascular, entre os consumidores de gorduras saturadas e os não consumidores

3.8.1. Indicadores antropométricos *versus* Gorduras saturadas

As estatísticas descritivas de cada variável, por grupo de consumidores ou não, foram sintetizadas no Anexo 10.

O teste de Kolmogorov-Smirnov mostrou que apenas a distribuição dos pesos dos não consumidores de gorduras segue a distribuição normal ($p=.160$).

O gráfico *boxplot* dos índices de massa corporal, de cada um dos grupos, apresentaram muitos *outliers*, sendo 12 no dos consumidores de gorduras saturadas e 5 no dos não consumidores (Figura 43).

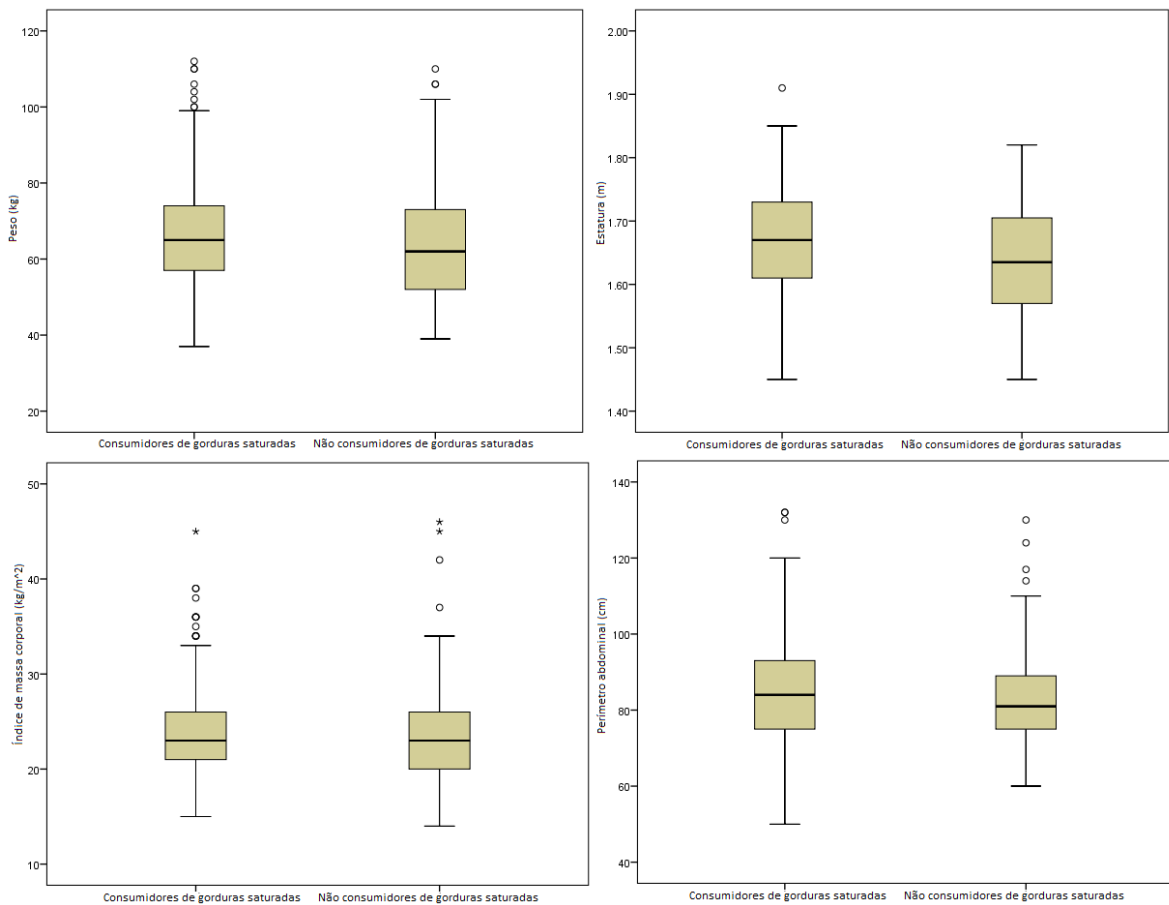


Figura 43: Distribuição dos indicadores antropométricos nos consumidores de gorduras saturadas e nos não consumidores (n=500; n=100)

O teste de Mann-Whitney mostrou que a distribuição dos pesos ($U=24927.500$; $p=.963$), dos índices de massa corporal ($U=24245.000$; $p=.632$) e dos perímetros abdominais ($U=22927.000$; $p=.190$) não estão relacionados com o facto de um indivíduo ser ou não consumidor de gorduras saturadas.

Por outro lado, mostrou que a estatura dos indivíduos dependia do facto de consumir ou não gorduras saturadas ($U=20860.500$; $p=.009$), no entanto o coeficiente Eta mostrou que essa associação é bastante fraca ($\text{Eta}=.124$).

3.8.2. Indicadores clínicos versus Gorduras saturadas

As pressões arteriais foram estudadas para o grupo de consumidores de gorduras saturadas e para o grupo dos não consumidores e os resultados foram sintetizados no Anexo 10.

Foram construídos gráficos *boxplot* para analisar as distribuições de cada uma das variáveis, em cada grupo (Figura 44), tendo-se salientado os inúmeros *outliers*, correspondentes a indivíduos com valores de pressão bastante mais elevados do que o grupo a que pertencem.

O teste de Kolmogorov-Smirnov comprovou que nenhuma das distribuições de pressões segue a distribuição normal ($p<.05$).

O teste de Mann-Whitney mostrou haver evidências estatísticas para afirmar que os comportamentos das pressões arteriais sistólica ($U=23572.500$; $p=.359$) e diastólica ($U=22696.500$; $p=.134$) eram independentes do consumo de gorduras saturadas.

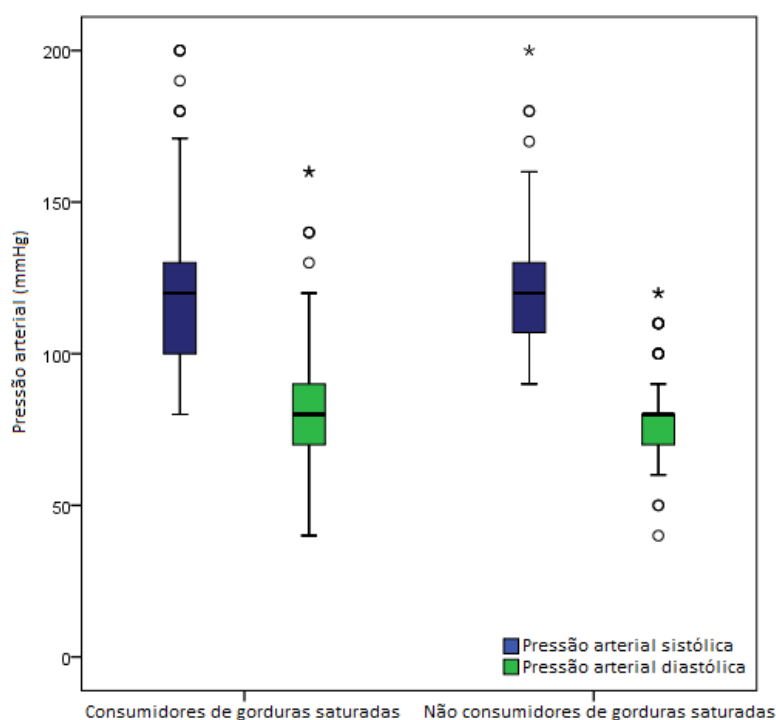


Figura 44: Distribuição da pressão arterial sistólica e da pressão arterial diastólica dos consumidores de gorduras saturadas e dos não consumidores (n=500; n=100)

3.9. Comportamento do Risco Cardiovascular, por grupos de risco

As prevalências dos riscos cardiovasculares em cada grupo de risco podem ser consultadas nas tabelas de contingência do Anexo 11, do Anexo 12 e do Anexo 13, sendo que as principais para este estudo serão descritas de seguida, por tipo de risco.

Risco A

No que diz respeito ao género, 20.8% dos homens e 41.2% das mulheres estavam com risco A aumentado, alto ou muito alto. Entre os 67 indivíduos que estavam com risco aumentado, 50.7% eram homens e 49.3% eram mulheres.

Entre os 57 indivíduos que estavam com risco alto, 35.1% eram homens e 64.9% eram mulheres. Entre os 51 indivíduos que estavam com risco muito alto, 39.2% eram homens e 60.8% eram mulheres.

O género foi significativamente relacionado com o risco A ($Z=2.454$; $p=.000$), pelo teste de independência de Kolmogorov-Smirnov, no entanto o coeficiente V de Cramer mostrou

que era uma relação fraca ($V=0.233$; $p=0.000$). A análise do gráfico de barras (Figura 45) revelou que, no que diz respeito ao género, o feminino era um fator de risco cardiovascular, uma vez que, contrariamente ao que se passava com os homens, havia uma maior percentagem de mulheres nos níveis mais graves e menos percentagem nos níveis menos graves de risco.

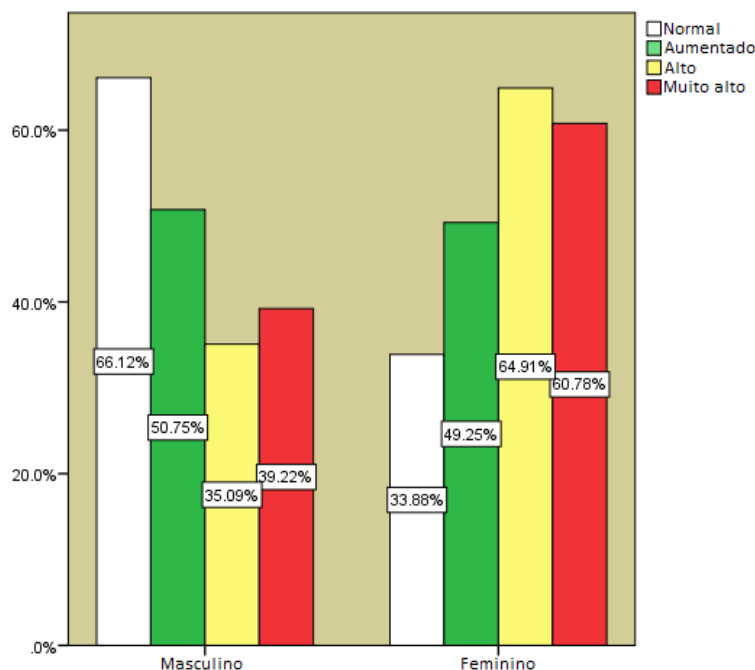


Figura 45: Gráfico de barras relativo à distribuição do risco A, de acordo com o género (n=355; n=245)

No que diz respeito à faixa etária, 38.1% dos mais velhos e 27.3% dos mais novos estavam com risco A anormal. Entre os 67 indivíduos que estavam com risco aumentado, 20.9% eram mais velhos e 79.1% eram mais novos. Entre os 57 indivíduos que estavam com risco alto, 28.1% eram mais velhos e 71.9% eram mais novos. Entre os 51 indivíduos que estavam com risco muito alto, 19.6% eram mais velhos e 80.4% eram mais novos. A faixa etária não foi significativamente relacionada com o risco A ($Z=1.007$; $p=.262$), pelo teste de independência de Kolmogorov-Smirnov.

74.8% dos indivíduos com sobrepeso e 4.6% dos indivíduos sem sobrepeso estavam com risco A aumentado, alto ou muito alto. Entre os 67 indivíduos que estavam com risco aumentado, 73.1% estavam com sobrepeso e 26.9% não estavam com sobrepeso. Todos os indivíduos que estavam com risco alto ou muito alto estavam com sobrepeso. O teste de independência Kolmogorov-Smirnov mostrou existir uma relação significativa entre a

condição de sobrepeso e o risco A ($Z=8.195$; $p=.000$) e o coeficiente V de Cramer avaliou essa relação como sendo forte ($V=.751$; $p=.000$), como seria de esperar, tendo em conta o modelo do risco A.

72.9% dos indivíduos com obesidade abdominal e 0.0% dos indivíduos sem obesidade abdominal estavam com risco A aumentado, alto ou muito alto. Todos os indivíduos associados ao risco aumentado, alto ou muito alto eram obesos, a nível abdominal. O teste de independência de Kolmogorov-Smirnov mostrou que existia uma relação de dependência entre a obesidade abdominal e o risco A ($Z=8.750$; $p=.000$) e o coeficiente V de Cramer apontou que era uma relação forte ($V=.789$; $p=.000$), como seria de esperar, tendo em conta o modelo do risco A.

40.8% dos indivíduos hipertensos e 24.4% dos indivíduos normotensos estavam com risco A aumentado, alto ou muito alto. Entre os 67 indivíduos com risco aumentado, 34.3% eram hipertensos e 65.7% eram normotensos. Entre os 57 indivíduos com risco alto, 42.1% era hipertensos e 57.9% eram normotensos. Entre os 51 indivíduos com risco muito alto, 47.1% eram hipertensos e 52.9% eram normotensos.

O teste de independência de Kolmogorov-Smirnov mostrou existir uma relação de dependência entre a hipertensão e o risco A ($Z=1.822$; $p=.003$), no entanto, segundo o coeficiente V de Cramer, era uma relação muito fraca ($V=.175$; $p=.000$). A análise do gráfico de barras (Figura 46) revelou que no grupo dos hipertensos havia uma maior percentagem de indivíduos nos níveis mais graves de risco A do que nos níveis menos graves, contrariamente ao que se passava no grupo dos normotensos.

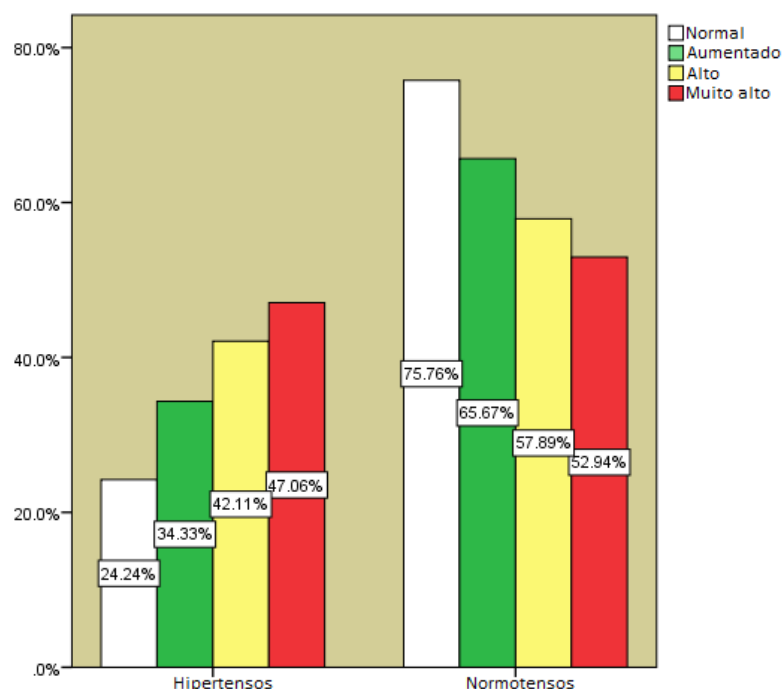


Figura 46: Gráfico de barras relativo à distribuição do risco A, de acordo com a hipertensão (n=174; n=426)

10.6% dos fumadores e 30.7% dos não fumadores estavam com risco A aumentado, alto ou muito alto. Entre os 67 indivíduos com risco aumentado, 4.5% eram fumadores e 95.5% eram não fumadores. Entre os 57 indivíduos com risco alto, 1.8% eram fumadores e 98.2% eram não fumadores. Entre os 51 indivíduos com risco muito alto, 2.0% eram fumadores e 98.0% eram não fumadores. O teste de independência de Kolmogorov-Smirnov mostrou não existir uma relação significativa de dependência entre os hábitos tabágicos e o risco A ($Z=1.323$; $p=.060$).

30.5% dos consumidores regulares de álcool e 28.5% dos não consumidores estavam com risco A aumentado, alto ou muito alto. Entre os 67 indivíduos com risco aumentado, 35.8% eram consumidores de álcool e 64.2% eram não consumidores. Entre os 57 indivíduos com risco alto, 35.1% eram consumidores de álcool e 64.9% eram não consumidores. Entre os 51 indivíduos com risco muito alto, 31.4% eram consumidores de álcool e 68.6% não eram. O teste de independência de Kolmogorov-Smirnov mostrou não existir uma relação significativa de dependência entre os hábitos alcoólicos e o risco A ($Z=.221$; $p=1.000$).

30.0% dos consumidores de gorduras saturadas e 25.0% dos não consumidores estavam com risco A aumentado, alto ou muito alto. Entre os 67 indivíduos com risco aumentado, 88.1% eram consumidores de gorduras e 11.9% não eram. Entre os 57 indivíduos com risco

alto, 86.0% eram consumidores de gorduras e 14.0% não eram. Entre os 51 indivíduos com risco muito alto, 82.4% eram consumidores de gorduras e 17.6% não eram. O teste de independência de Kolmogorov-Smirnov mostrou não existir uma relação significativa de dependência entre o consumo de gorduras saturadas e o risco A ($Z=.456$; $p=.985$).

Risco B

No que diz respeito ao género, 56.9% dos homens e 42.0% das mulheres estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos que estavam com risco acrescido, mas baixo, 41.6% eram homens e 58.4% eram mulheres. Entre os 120 indivíduos que estavam com risco moderado, 70.8% eram homens e 29.2% eram mulheres. Entre os 84 indivíduos que estavam com risco alto, 81.0% eram homens e 19.0% eram mulheres. Os 12 indivíduos que estavam com risco muito alto eram homens.

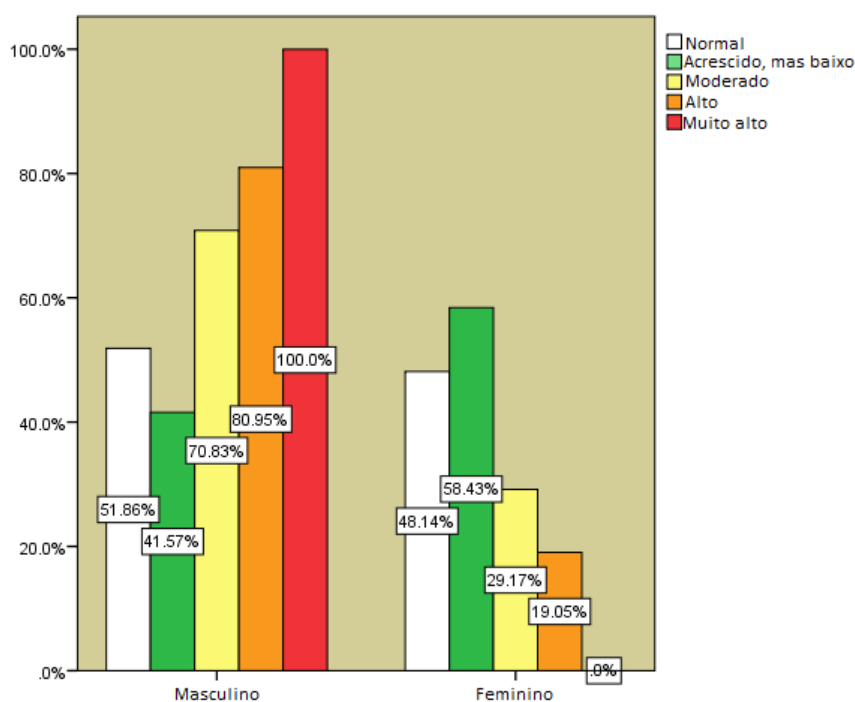


Figura 47: Gráfico de barras relativo à distribuição do risco B, de acordo com o género

De acordo com o teste de independência de Kolmogorov-Smirnov, existia uma relação significativa de dependência entre o género e o risco B ($Z=3.090$; $p=.000$), mas o cálculo do coeficiente V de Cramer mostrou que essa relação, apesar de significativa, era fraca ($V=.287$; $p=.000$). A análise do gráfico de barras (Figura 47) mostrou que, em termos de género, os homens definiam o grupo de risco, uma vez que havia uma maior percentagem

de homens nos níveis de risco mais graves e uma menor percentagem nos níveis menos graves, contrariamente ao que se passava no grupo das mulheres.

No que diz respeito à faixa etária, 74.3% dos mais velhos e 45.9% dos mais novos estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos que estavam com risco acrescido, mas baixo, 20.2% eram mais velhos e 79.8% eram mais novos. Entre os 120 indivíduos que estavam com risco moderado, 25.0% eram mais velhos e 75.0% eram mais novos. Entre os 84 indivíduos que estavam com risco alto, 28.6% eram mais velhos e 71.4% eram mais novos. Entre os 12 indivíduos com risco muito alto, metade eram mais velhos e a outra metade eram mais novos.

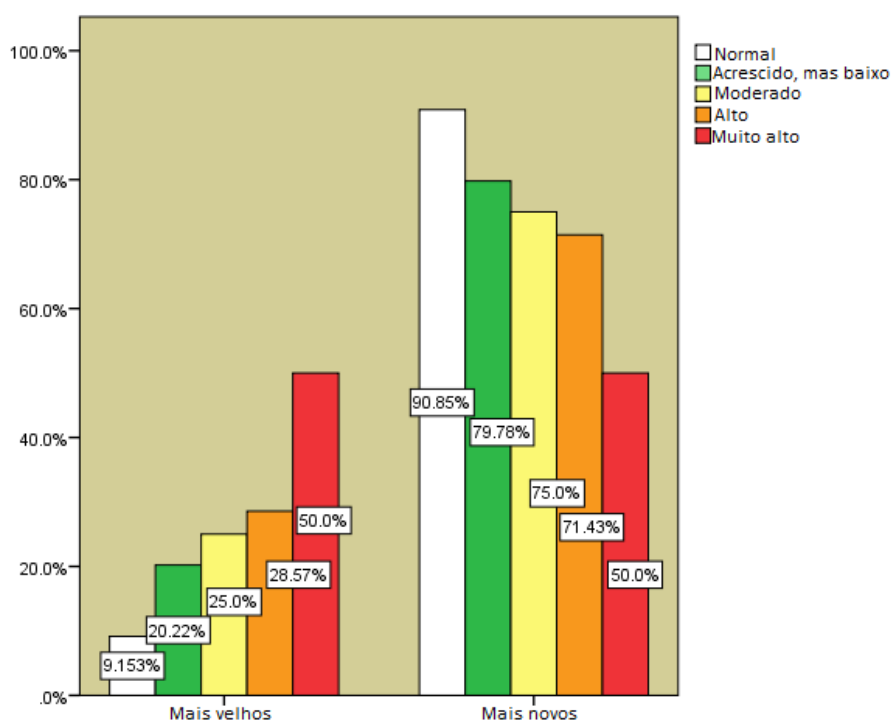


Figura 48: Gráfico de barras relativo à distribuição do risco B, de acordo com a faixa etária (n=105; n=495)

Pelo teste de independência de Kolmogorov-Smirnov, existia uma relação significativa entre a faixa etária e o risco B ($Z=2.646$; $p=.000$), mas o cálculo do coeficiente V de Cramer mostrou que essa relação, apesar de significativa, era fraca ($V=.242$; $p=.000$). A observação do gráfico de barras (Figura 48) mostrou que a percentagem de indivíduos mais velhos aumentava nos níveis mais graves de risco, contrariamente ao que se passava no grupo dos mais novos.

60.0% dos indivíduos com sobrepeso e 45.9% dos indivíduos sem sobrepeso estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos com risco acrescido, mas baixo, 27.0% tinham sobrepeso e 73.0% não tinham sobrepeso. Entre os 120 indivíduos com risco moderado, 41.7% tinham sobrepeso e 58.3% não tinham sobrepeso. Entre os 84 indivíduos com risco alto, 53.6% tinham sobrepeso e 46.4% não tinham sobrepeso. Entre os 12 indivíduos com risco muito alto, 58.3% tinham sobrepeso e 41.7% não tinham sobrepeso.

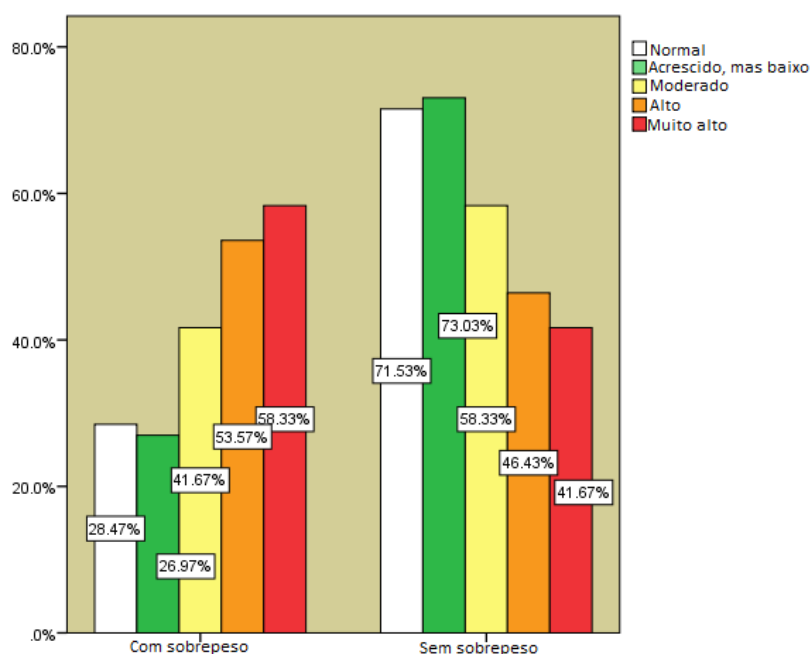


Figura 49: Gráfico de barras relativo à distribuição do risco B, de acordo com o sobrepeso (n=210; n=390)

O teste de independência de Kolmogorov-Smirnov mostrou existir uma relação significativa de dependência entre o sobrepeso e o risco B ($Z=2.260$; $p=.000$), mas o coeficiente V de Cramer mostrou ser uma relação fraca ($V=.208$; $p=.000$). A análise do gráfico de barras (Figura 49) mostrou haver uma maior percentagem de indivíduos com sobrepeso nos níveis mais graves de risco, contrariamente ao que se passou nos indivíduos sem sobrepeso.

57.1% dos indivíduos com obesidade abdominal e 53.3% dos indivíduos sem obesidade abdominal estavam com Risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos com risco acrescido, mas baixo, 43.8% tinham obesidade abdominal e 56.2% não tinham. Entre os 120 indivíduos com risco moderado, 40.8% tinham obesidade abdominal e 59.2% não tinham. Entre os 84 indivíduos com risco alto, 52.4% tinham obesidade

abdominal e 47.6% não tinham. Entre os 12 indivíduos com risco muito alto, 41.7% tinham obesidade abdominal e 58.3% não tinham. O teste de Kolmogorov-Smirnov para a independência mostrou que não existiam evidências estatísticas ($Z=1.250$; $p=.088$) que permitissem afirmar a existência de dependência entre a obesidade abdominal e o risco B.

97.1% dos indivíduos hipertensos e 31.9% dos indivíduos normotensos estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos com risco acrescido, mas baixo, 42.7% estavam hipertensos e 57.3% estavam normotensos. Entre os 120 indivíduos que estavam com risco moderado, 38.3% estavam hipertensos e 61.7% estavam normotensos. Entre os 84 indivíduos que estavam com risco alto, 86.9% estavam hipertensos e 13.1% estavam normotensos. Todos os 12 indivíduos que estavam com risco muito alto estavam hipertensos.

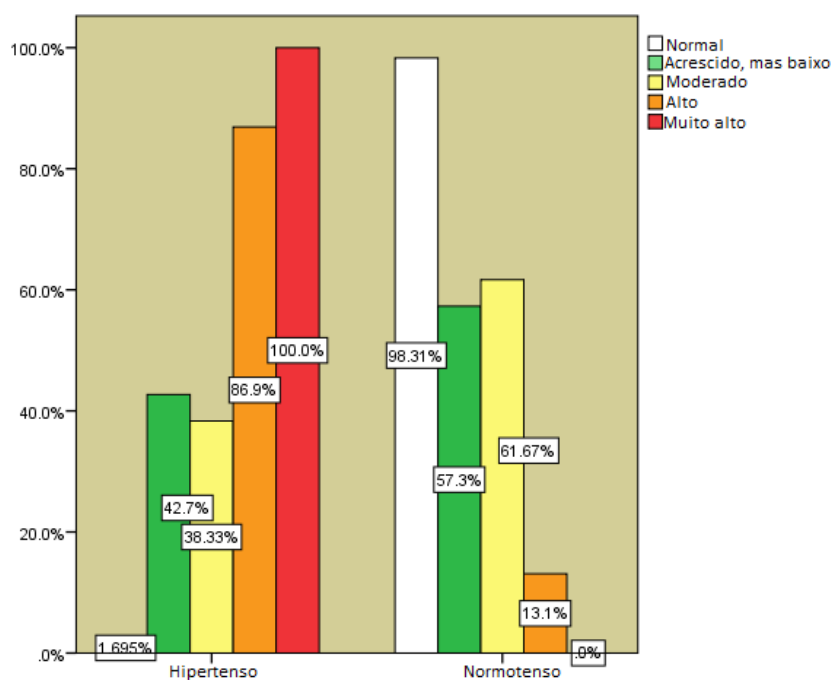


Figura 50: Gráfico de barras relativo à distribuição do risco B, de acordo com a hipertensão

O teste de independência de Kolmogorov-Smirnov mostrou existir uma relação significativa de dependência entre a hipertensão e o risco B ($Z=7.247$; $p=.000$) e o coeficiente V de Cramer mostrou que essa relação era forte ($V=.691$; $p=.000$). A análise do gráfico de barras (Figura 50) mostrou haver uma maior percentagem de hipertensos nos níveis mais graves de risco, contrariamente ao que se passava no grupo dos normotensos.

70.2% dos fumadores e 49.2% dos não fumadores estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos com risco acrescido, mas baixo, 4.5% eram fumadores e 95.5% eram não fumadores. Entre os 120 indivíduos com risco moderado, 12.5% eram fumadores e 87.5% eram não fumadores. Entre os 84 indivíduos com risco alto, 14.3% eram fumadores e 85.7% eram não fumadores. Entre os 12 indivíduos com risco muito alto, 16.7% eram fumadores e 83.3% eram não fumadores.

O teste de independência de Kolmogorov-Smirnov revelou uma relação de dependência entre os hábitos tabágicos e o risco B ($Z=1.835$; $p=.002$), no entanto o coeficiente V de Cramer classificou essa relação como sendo muito fraca ($V=.158$; $p=.005$). A análise do gráfico de barras (Figura 51) mostrou que, contrariamente ao que se passava no grupo dos não fumadores, no grupo dos fumadores a percentagem de indivíduos nos níveis mais graves de risco era maior do que a dos níveis menos graves.

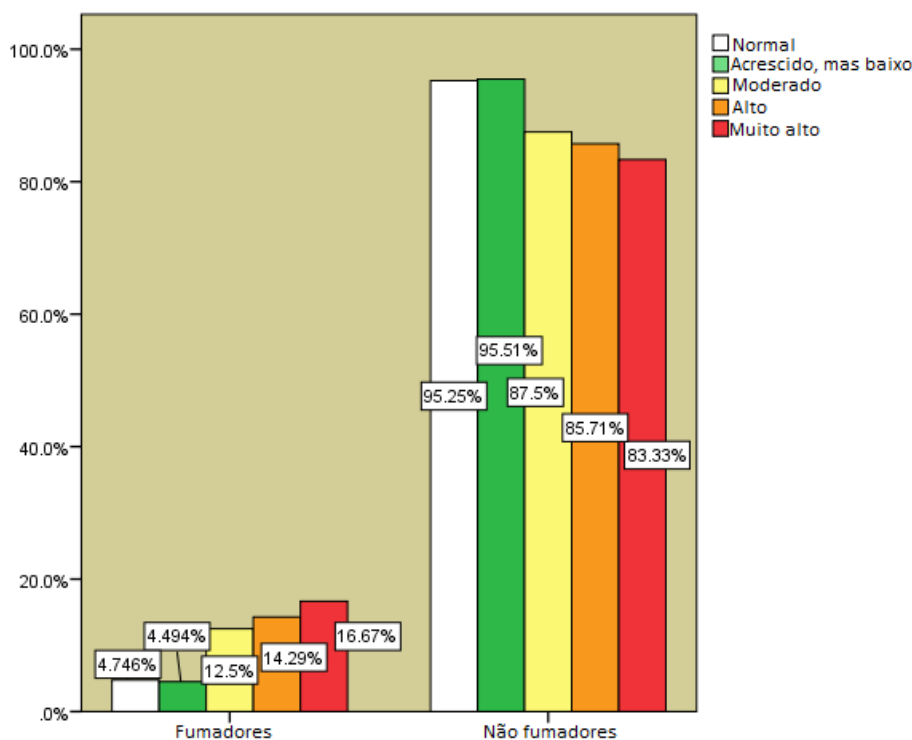


Figura 51: Gráfico de barras relativo à distribuição do risco B, de acordo com os hábitos tabágicos (n=47; n=553)

58.9% dos consumidores de álcool e 46.9% dos não consumidores de álcool estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos com risco acrescido, mas baixo, 16.9% consumiam álcool e 83.1% não consumiam álcool. Entre os 120

indivíduos com risco moderado, 37.5% consumiam álcool e 62.5% não consumiam álcool. Entre os 84 indivíduos com risco alto, 57.1% consumiam álcool e 42.9% não consumiam. Entre os 12 indivíduos com risco muito alto, 66.7% consumiam álcool e 33.3% não consumiam.

O teste de independência de Kolmogorov-Smirnov mostrou existir uma relação significativa de dependência ($Z=2.615$; $p=.000$) entre os hábitos alcoólicos e o risco B, no entanto o coeficiente V de Cramer mostrou que a relação era fraca ($V=.271$; $p=.000$). A análise do gráfico de barras (Figura 52) mostrou que os consumidores de gorduras saturadas tinham maiores percentagens associadas aos níveis mais graves de risco, contrariamente ao que se passava no grupo dos não consumidores de gorduras saturadas.

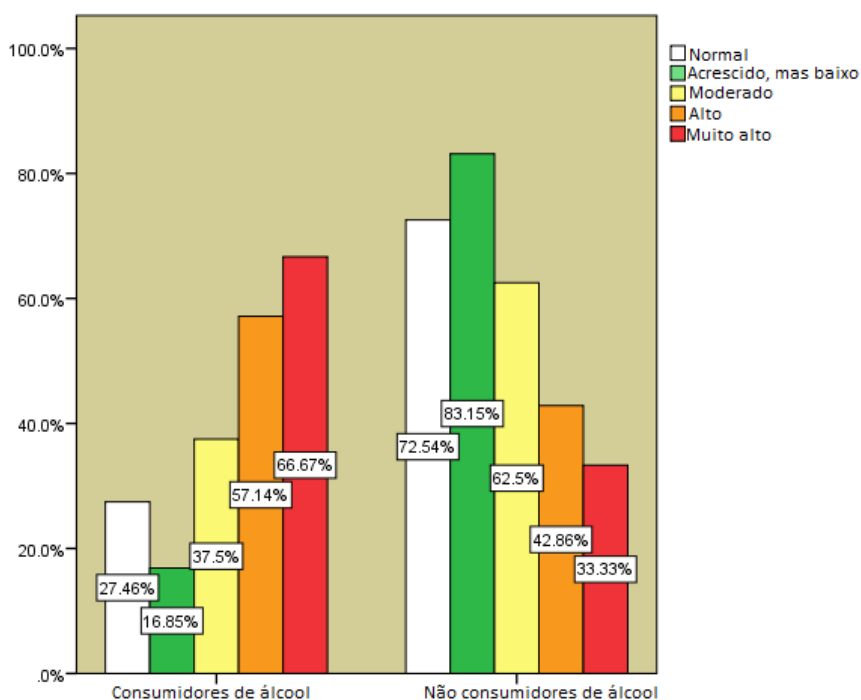


Figura 52: Gráfico de barras relativo à distribuição do risco B, de acordo com os hábitos alcoólicos (n=197; n=403)

50.6% dos consumidores de gorduras saturadas e 52.0% dos não consumidores de gorduras saturadas estavam com risco B acrescido, moderado, alto ou muito alto. Entre os 89 indivíduos com risco acrescido, mas baixo, 71.9% consumiam gorduras saturadas e 28.1% não consumiam. Entre os 120 indivíduos com risco moderado, 82.5% consumiam gorduras saturadas e 17.5% não consumiam. Entre os 84 indivíduos com risco alto, 92.9% consumiam gorduras saturadas e 7.1% não consumiam. Os 12 indivíduos que estavam com

risco muito alto eram consumidores de gorduras saturadas. O teste de independência de Kolmogorov-Smirnov não mostrou evidências estatísticas que permitam afirmar existir uma relação de dependência entre o consumo de gorduras saturadas e o risco B ($Z=1.095$; $p=.181$).

Risco C

Entre os 226 indivíduos com risco baixo, 64.2% eram homens e 35.8% eram mulheres. Entre os 9 indivíduos com risco moderado, 88.9% eram homens e 11.1% eram mulheres. Os dois indivíduos com risco alto eram homens. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre o género e o risco C ($Z=.383$; $p=.999$).

Entre os 226 indivíduos com risco baixo, 42.9% eram mais velhos e 57.1% eram mais novos. Entre os 9 indivíduos com risco moderado, 66.7% eram mais velhos e 33.3% eram mais novos. Os dois indivíduos com risco alto eram mais velhos. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre a faixa etária e o risco C ($Z=.409$; $p=.996$).

Entre os 226 indivíduos com risco baixo, 47.8% tinham sobrepeso e 52.2% não tinham sobrepeso. Entre os 9 indivíduos com risco moderado, 77.8% tinham sobrepeso e 22.2% não tinham sobrepeso. Entre os 2 indivíduos com risco alto, 1 deles tinha sobrepeso e o outro não. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre o sobrepeso e o risco C ($Z=.340$; $p=1.000$).

Entre os 226 indivíduos com risco baixo, 54.4% tinham obesidade abdominal e 45.6% não tinham obesidade abdominal. Entre os 9 indivíduos com risco moderado, 66.7% tinham obesidade abdominal e 33.3% não tinham obesidade abdominal. Entre os 2 indivíduos com risco alto, 1 deles tinha obesidade abdominal e o outro não. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre a obesidade abdominal e o risco C ($Z=.126$; $p=1.000$).

Entre os 226 indivíduos com risco baixo, 40.7% estavam hipertensos e 59.3% estavam normotensos. Os 9 indivíduos com risco moderado e os 2 indivíduos com risco muito alto

estavam todos hipertensos. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre a hipertensão e o risco C ($Z=.815$; $p=.520$).

Entre os 226 indivíduos com risco baixo, 11.9% eram fumadores e 88.1% não. Os 9 indivíduos com risco moderado e os 2 indivíduos com risco alto eram todos não fumadores. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre os hábitos tabágicos e o risco C ($Z=.256$; $p=1.000$).

Entre os 226 indivíduos com risco baixo, 39.8% consumiam bebidas alcoólicas e 60.2% não. Entre os 9 indivíduos com risco moderado, 55.6% consumiam bebidas alcoólicas e 44.4% não. Entre os 2 indivíduos com risco alto, um deles consumia gorduras saturadas e outro não. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre o consumo de álcool e o risco C ($Z=.204$; $p=1.000$).

Entre os 226 indivíduos com risco baixo, 85.0% consumiam gorduras saturadas e 15.0% não. Entre os 9 indivíduos com risco moderado, 77.8% consumiam gorduras saturadas e 22.2% não. Os dois indivíduos com risco alto eram consumidores de gorduras saturadas. O teste de independência de Kolmogorov-Smirnov não detetou evidências estatísticas que garantissem uma relação de dependência entre o consumo de gorduras saturadas e o risco C ($Z=.060$; $p=1.000$).

3.10. Correlações entre Risco Cardiovascular e alguns indicadores

Para cada um dos riscos em análise foi medido o grau de associação com alguns indicadores de natureza quantitativa, como a idade, o peso, a estatura, o índice de massa corporal e as pressões arteriais. Para tal, calcularam-se os coeficientes de correlação ρ de Spearman para cada uma destas combinações (Tabela 17).

O risco A foi significativamente correlacionado com todos os indicadores, mas com graus de associação diferentes. Os indicadores com correlações fortes com este risco foram o perímetro abdominal ($\rho_S=.728$) e o índice de massa corporal ($\rho_S=.727$). O risco A foi

moderadamente correlacionado com o peso ($\rho_S = 597$). As correlações com os restantes indicadores foram todas fracas ou muito fracas, pois foram associadas a coeficientes com valores a quatro décimas (Tabela 17).

Observando os gráficos da Figura 53, obtidos pelo comando *plot* do *software* R, é possível verificar as três correlações positivas, fortes e moderada, referidas no parágrafo anterior e também verificar como é o aspeto gráfico da correlação fraca entre o risco A e a pressão arterial sistólica ($\rho_S = .238$). Na verdade, à medida que se vai “subindo” na gravidade do risco A, os valores do perímetro abdominal, do índice de massa corporal e do peso vão aumentando. Em contrapartida, no gráfico da pressão arterial sistólica não é notória qualquer tendência.

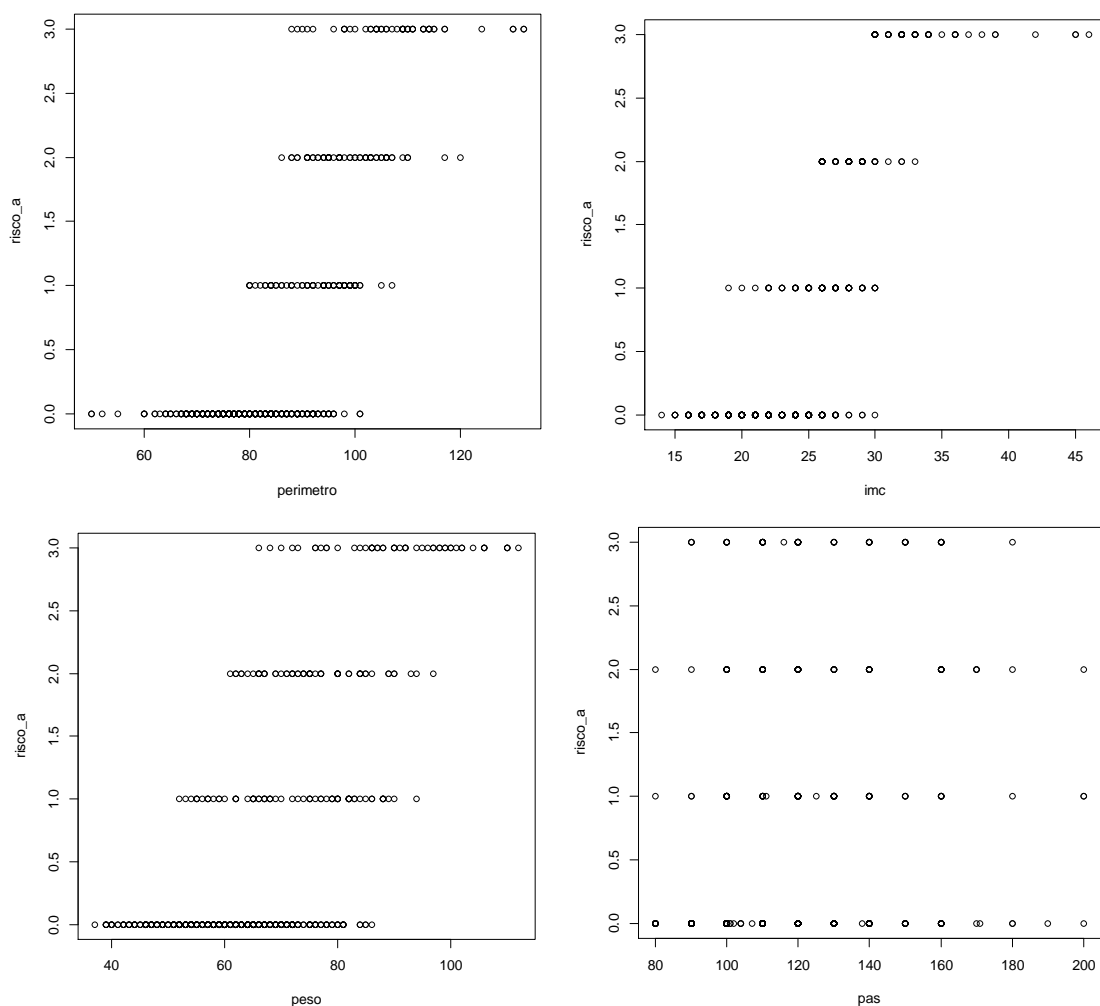


Figura 53: Output do *software* R dos gráficos das correlações entre o Risco A e o perímetro abdominal, o índice de massa corporal, o peso e a pressão arterial sistólica

O risco B também estava positiva e significativamente correlacionado com todos os indicadores, no entanto apenas estava fortemente correlacionado com a pressão arterial diastólica ($\rho_S=.725$) e moderadamente correlacionado com a pressão arterial sistólica ($\rho_S=.637$). Todas as restantes correlações eram fracas ou muito fracas. O risco C apenas foi significativamente correlacionado com a pressão arterial sistólica e com a pressão arterial diastólica, com coeficientes inferiores a .39 que revelaram que as duas correlações eram fracas.

Tabela 17: Resultados do cálculo do coeficiente de Spearman

Indicador		Risco A	Risco B	Risco C
Idade	Coeficiente de Spearman	.296	.325	.062
	<i>p-value</i>	.000	.000	.346
Peso	Coeficiente de Spearman	.597	.238	.122
	<i>p-value</i>	.000	.000	.061
Estatura	Coeficiente de Spearman	-.092	.227	.051
	<i>p-value</i>	.024	.000	.435
Índice de massa corporal	Coeficiente de Spearman	.727	.150	.102
	<i>p-value</i>	.000	.000	.116
Perímetro abdominal	Coeficiente de Spearman	.728	.195	.107
	<i>p-value</i>	.000	.000	.101
Pressão arterial sistólica	Coeficiente de Spearman	.238	.637	.362
	<i>p-value</i>	.000	.000	.000
Pressão arterial diastólica	Coeficiente de Spearman	.217	.725	.342
	<i>p-value</i>	.000	.000	.000

3.11. Grau de associação entre os Riscos Cardiovasculares

O coeficiente de Spearman foi ainda usado para averiguar se existia alguma correlação entre os três tipos de riscos.

Os riscos A e B foram correlacionados positivamente, mas a correlação era muito fraca ($\rho=.179$). Os riscos B e C também estavam positivamente correlacionados, mas a correlação era fraca ($\rho=.283$).

4. ANÁLISE MULTIVARIADA

4.1. Modelação do Risco A

Entre os 600 indivíduos a quem foram avaliados os riscos A (considerando o género, o índice de massa corporal e o perímetro abdominal), 425 estavam com risco normal, 67 com risco aumentado, 57 com risco alto e 51 com risco muito alto.

Da análise cruzada do risco A com os indicadores detetaram-se as seguintes relações:

- Correlações fortes: perímetro abdominal; índice de massa corporal.
- Correlações moderadas: peso.
- Correlações fracas: idade; pressão arterial sistólica; pressão arterial diastólica; género.
- Correlações muito fracas: estatura.

Considerando estas 8 covariáveis, que foram significativamente correlacionas com o risco A, ter-se-ia 255 modelos possíveis a ajustar e comparar recorrendo ao critério de informação de Akaike e/ou ao critério de informação Bayesiano.

Não foi encontrada nenhuma função no R associada a um algoritmo do género de *stepwise* compatível com a função *rrvglm*. Logo numa primeira análise ficou decidido que a estatura não iria ser considerada no modelo, pois a sua relação estatística com o risco foi muito fraca e, na pesquisa bibliográfica não foi encontrada qualquer referência à estatura como indicador de risco cardiovascular.

As três variáveis que estavam mais correlacionadas com o risco A foram as que serviram de indicadores para a avaliação do risco e, como tal, considerou-se que deveriam entrar no modelo inicial algumas das variáveis além destas, mesmo estando fracamente correlacionadas com o risco A. A alternativa encontrada para minimizar o número de modelos possíveis foi retirar as duas pressões arteriais, enquanto variáveis contínuas, e inserir a variável hipertensão dicotomizada, conseguindo reduzir o número de modelos possíveis para 63.

Um procedimento, recomendado pelo autor, para o recurso à função *rrvglm* é a uniformização e negação das variáveis binárias, para que sejam tratadas como variáveis numéricas. Para isso, recorreu-se à função *transform()* e *scale()* (Yee T. , 2010).

Cada um dos 63 modelos foi ajustado pela função *rrvglm* do *package VGAM* do *software R* e a cada um deles foi aplicado o teste da razão de verosimilhanças e calculada a sua significância. Como complemento, optou-se por calcular a estatística para aplicação do critério de informação Bayesiano.

Nesses 63 modelos, 7 foram associados a um bom ajustamento dos dados, com $p > .05$. (Tabela 18).

Tabela 18: Modelos que ajustam significativamente os dados, segundo o teste da razão de verosimilhanças

Modelo	Covariáveis	Estatística	<i>p-value</i>	BIC
M1	X1+X2+X4+X5+X6	.763	.382	299.604
M2	X1+X2+X3+X5+X6	2.205	.138	301.046
M3	X1+X2+X3+X4+X6	1.128	.288	299.970
M4	X1+X2+X5+X6	3.088	.214	295.532
M5	X1+X2+X4+X6	1.836	.399	294.280
M6	X1+X2+X3+X6	3.890	.143	296.335
M7	X1+X2+X3	4.691	.196	290.738

x1- perímetro abdominal; x2-índice de massa corporal; x3- peso; x4- idade; x5- hipertensão arterial; x6- género

Um dos erros que muitas vezes se comete é considerar que quantas mais covariáveis tiver o modelo, melhor ajustamento ele terá. Na verdade, quanto mais variáveis tiver um modelo, maior será a estimativa do erro e mais dependente ele ficará dos dados observados. Desta forma, já que qualquer um destes sete mostrou estar significativamente ajustado aos dados, optou-se por não se recorrer aos modelos M1, nem M2, nem M3, por estarem a apenas uma variável de diferença do modelo completo.

Por outro lado, a variável peso (X3) acaba por estar representada na variável índice de massa corporal (X2) e, por isso, foram desconsiderados os modelos em que aparecia X3. Desta forma, restaram os modelos M4 e M5.

A diferença, em termos de variáveis explicativas, entre os modelos M4 e M5 é que um inclui a variável hipertensão (dicotómica) e o outro a variável idade (contínua). A hipertensão é um fator de risco modificável, importantíssimo e com grandes prevalências, por outro lado a idade é um fator de risco não modificável. Segundo George Box, “*All models are wrong. Some are useful*”. Neste alinhamento, optou-se pelo modelo M4, com quatro variáveis explicativas: perímetro abdominal; índice de massa corporal; hipertensão; género.

A análise da qualidade do ajuste foi ainda feita recorrendo a um gráfico com os resíduos para cada um dos três contrastes, recorrendo à função *residuals()* e à função *scatterplot3d()*, sendo obtido o gráfico da Figura 54.

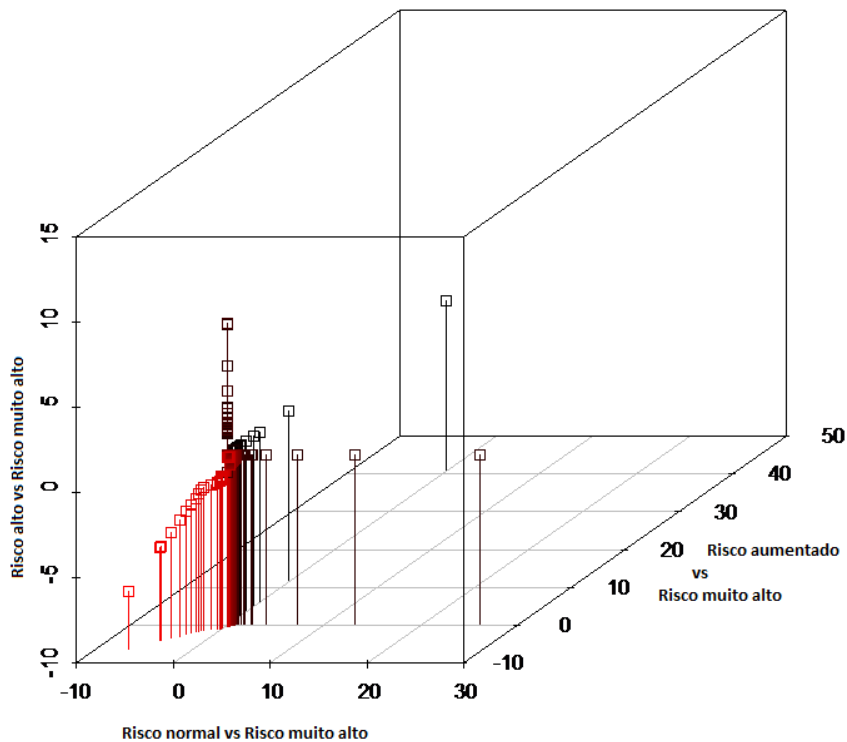


Figura 54: Gráfico de resíduos, por contrastes

A análise do gráfico com os resíduos confirmou o bom ajustamento do modelo selecionado, uma vez que os resíduos, dentro de cada contraste, estavam todos muito próximos de zero. O recurso à função *summary()* para os resíduos mostrou que todos os valores estão bastante próximos de zero.

A mesma função `summary()` para o modelo selecionado devolveu um resumo com os dados dos coeficientes (Tabela 19). Os valores de z são os quocientes entre a estimativa e o correspondente erro padrão. Geralmente, esta é a estatística usada para testar a significância de cada uma das variáveis, recorrendo à distribuição normal. Em termos de literatura, há autores que forçam a aplicação da distribuição normal, tendo em conta a grande dimensão da amostra. Neste caso, e tendo em conta o método usado para seleção do modelo, talvez fosse desnecessário o cálculo do nível de significância. De qualquer das formas, a autora optou por o fazer e partilhar. Constatou-se que todas as variáveis eram importantes para este modelo, uma vez que $p > .05$.

Tabela 19: Resumo das estimativas para os coeficientes do modelo

Variável	Estimativa β (IC)	Erro padrão	z	p -value
Peso (Aumentado vs Muito alto)	.498 -	.024	21.105	1.000
Peso (Alto vs Muito alto)	.165 -	.009	17.867	1.000
Intercepto (Normal)	201.252 -	-	-	-
Intercepto (Aumentado)	105.316 (85.860,124.771)	9.926	10.610	1.000
Intercepto (Alto)	36.910 (26.828,46.991)	5.144	7.176	1.000
Perímetro abdominal	-1.34888 -	-	-	-
Índice de massa corporal	-2.711 (-3.021,-2.401)	.159	-17.058	1.000
Hipertensão (hipertenso)	-.591 (-1.494,.311)	.461	-1.283	0.900
Género (masculino)	9.284 -	-	-	-

IC- intervalo de confiança.

O pacote VGAM também se mostrou incompatível com a função `confint()`, que geralmente é transversal a qualquer modelo de regressão e que permite obter intervalos de confiança. Considerando a aproximação à distribuição, construíram-se os intervalos de confiança, de acordo com a fórmula para esse efeito, nos casos em que o erro era conhecido.

Recorrendo à função `Coef()`, obteve-se o *output* mostrado na Figura 55.

```

> Coef(ajuste48)
A matrix:
                latvar
log(mu[,1]/mu[,4]) 1.0000000
log(mu[,2]/mu[,4]) 0.4978150
log(mu[,3]/mu[,4]) 0.1651452

C matrix:
                latvar
x1  -1.3488832
x2  -2.7109651
sx5 -0.5911457
sx6  9.2836380

B1 matrix:
                log(mu[,1]/mu[,4]) log(mu[,2]/mu[,4]) log(mu[,3]/mu[,4])
(Intercept)          201.2518          105.3158          36.90955

```

Figura 55: Output do software R da função Coef()

As matrizes A e C, da Figura 55, estavam interligadas pela variável *latvar*, que corresponde ao parâmetro ω_j da forma funcional de um modelo estereótipo (49), representando o peso dos efeitos das 4 covariáveis do modelo. Confirmou-se que também os ω_j seguiam uma ordenação: $1.000 > .498 > .165 > .000$.

Uma análise da matriz A, em termos de monotonia revelou que à medida que o nível de gravidade de risco ia aumentando, o peso dos efeitos das covariáveis ia diminuindo. Por sua vez, a matriz mostrou que, para as variáveis perímetro abdominal (X1), índice de massa corporal (X2) e hipertensão (X5), por terem coeficientes negativos, à medida que os pesos de ω_j iam diminuindo (e o grau de risco aumentando) os valores da covariáveis iam aumentando. Ou seja, um aumento do perímetro abdominal e do índice de massa corporal estava relacionado com um aumento do grau de risco cardiovascular.

Na parte inferior da Figura 55 podem ser consultados os três interceptos da forma funcional do modelo estereótipo.

Os valores de $\omega_j\beta$ podem ser calculados pela função *coef(...,matrix=T)* do software R.

Estes valores são logaritmos de Odds e, para facilitar a interpretação, foram calculadas exponenciais dos valores de $\omega_j\beta$. Estes resultados foram sintetizados na Tabela 20.

Tabela 20: Coeficientes e Odds Ratio para cada uma das covariáveis do modelo estereótipo, para o Risco A

	Condição de Risco A					
	Normal vs Muito alto		Aumentado vs Muito alto		Alto vs Muito alto	
	$\omega_j\beta$	OR	$\omega_j\beta$	OR	$\omega_j\beta$	OR
Perímetro abdominal	-1.349	.259	-.671	.511	-.223	.800
Índice de massa corporal	-2.711	.066	-1.350	.259	-.448	.639
Hipertensão (hipertenso)	-.591	.553	-.294	.745	-.098	.907
Género (masculino)	9.284	10764	4.622	101.697	1.533	4.632

OR-Odds Ratio; vs- versus

Pela Tabela 20, verificou-se que

$$\frac{P(\text{Risco normal})}{P(\text{Risco muito alto})} = .553 \Rightarrow P(\text{Risco muito alto}) = 1.808P(\text{Risco normal})$$

Ou seja, um indivíduo hipertenso, tinha duas vezes mais probabilidades de estar com risco muito alto do que com risco normal. Fazendo raciocínios análogos, verificou-se que a probabilidade de um indivíduo hipertenso ter risco aumentado ou muito alto, ou ter risco alto ou muito alto é praticamente a mesma.

No que diz respeito ao género, verificou-se que o género masculino funciona como um fator de proteção, contrariamente ao que se considera para os países desenvolvidos.

No que diz respeito às variáveis contínuas foram construídos gráficos com as probabilidades, recorrendo à forma funcional de um modelo estereótipo (49) e à função *plot()* do software R.

$$P(y_i = 1 | x_{i1}, \dots, x_{ir}) = \frac{\exp(\nu_1)}{\exp(\nu_1) + \exp(\nu_2) + \exp(\nu_3) + 1}$$

$$P(y_i = 2 | x_{i1}, \dots, x_{ir}) = \frac{\exp(\nu_2)}{\exp(\nu_1) + \exp(\nu_2) + \exp(\nu_3) + 1}$$

$$P(y_i = 3 | x_{i1}, \dots, x_{ir}) = \frac{\exp(\nu_3)}{\exp(\nu_1) + \exp(\nu_2) + \exp(\nu_3) + 1}$$

$$P(y_i = 4 | x_{i1}, \dots, x_{ir}) = \frac{1}{\exp(\nu_1) + \exp(\nu_2) + \exp(\nu_3) + 1}$$

Em que, de acordo com a Tabela 20:

$$\nu_1 = 201.252 - 1.349X_1 - 2.711X_2 - .591X_5 + 9.284X_6$$

$$\nu_2 = 105.316 - .671X_1 - 1.350X_2 - .294X_5 + 4.622X_6$$

$$\nu_3 = 36.910 - .223X_1 - .448X_2 - .098X_5 + 1.533X_6$$

Cada uma das probabilidades referidas atrás era função de quatro variáveis. Para construir gráficos de probabilidades para cada uma das variáveis substituíram-se as restantes pelas suas médias. As variáveis binárias foram substituídas, portanto, por zero. Os gráficos foram realizados com o auxílio do *software* Geogebra.

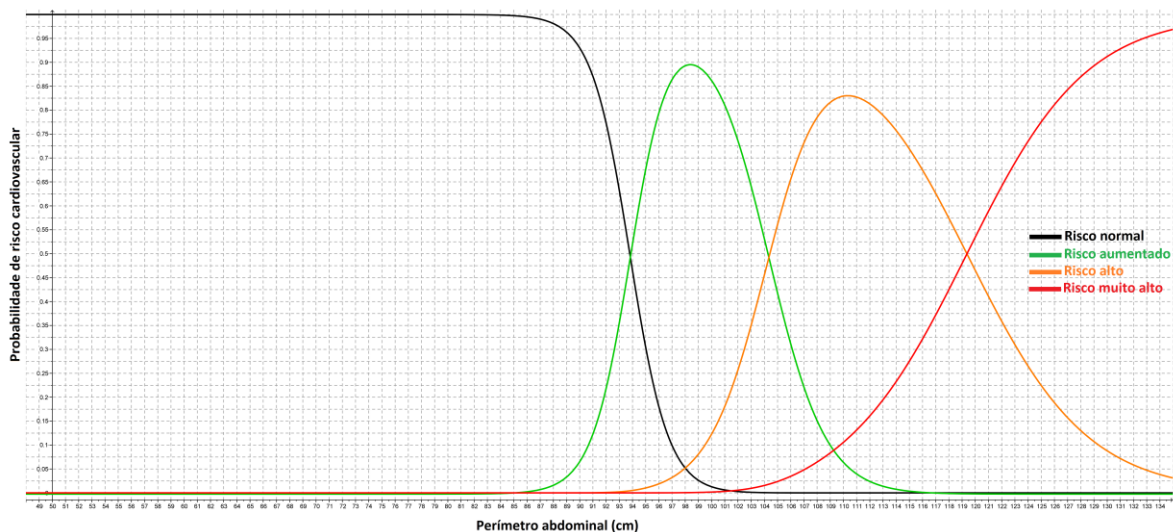


Figura 56: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função do perímetro abdominal (cm)

Estes gráficos são bastante úteis para fazer inferência sobre o risco cardiovascular, pois o seu manuseio é bastante simples. Por exemplo, um indivíduo que tenha 107cm de perímetro abdominal tem mais probabilidades de estar com risco alto, do que com qualquer outro nível de risco. A probabilidade de estar com risco normal é de 0.0%; a probabilidade de estar com risco aumentado é de 22.5%; a probabilidade de estar com risco alto é de 72.5%; a probabilidade de estar com risco muito alto é de 5.0% (Figura 56).

Na Figura 57, pode-se observar um gráfico semelhante que foi construído para a outra variável preditiva, índice de massa corporal, e cuja interpretação é análoga à do gráfico anterior.

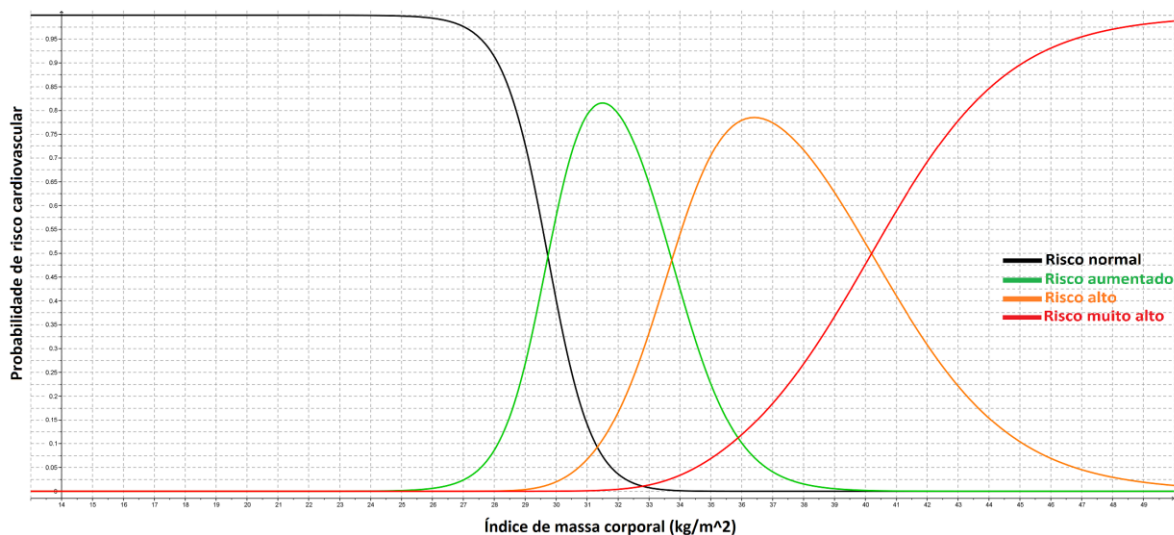


Figura 57: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função do índice de massa corporal (kg/m^2)

Este modelo preditivo, porém, só poderá ser utilizado para fazer previsões de risco se for comprovada a sua validade. Para isso, utilizou-se a amostra de 100 indivíduos que havia sido reservada para este efeito.

Recorrendo ao R e aos valores do perímetro, do índice de massa corporal, da hipertensão (uniformizada e negada) e do género (uniformizado e negado), estimaram-se as probabilidades de ocorrência de cada uma das categorias do risco. Entre as quatro categorias, considerou-se a que tinha a probabilidade mais elevada para ser a categoria de risco desse indivíduo.

Na Tabela 21 foram apresentados os valores das probabilidades de cada categoria, o risco previsto pelo modelo e o risco observado, dos primeiros 10 indivíduos.

Tabela 21: Primeiros dez casos do processo de validação do modelo

ID	P(normal)	P(aumentado)	P(alto)	P(muito alto)	Risco Modelado	Risco Observado
1	0.960	0.040	0.000	0.000	Normal	Normal
2	0.080	0.889	0.030	0.000	Aumentado	Aumentado
3	0.999	0.001	0.000	0.000	Normal	Normal
4	1.000	0.000	0.000	0.000	Normal	Normal
5	1.000	0.000	0.000	0.000	Normal	Normal
6	1.000	0.000	0.000	0.000	Normal	Normal
7	1.000	0.000	0.000	0.000	Normal	Normal
8	0.000	0.000	0.022	0.978	Muito alto	Muito alto
9	0.000	0.012	0.697	0.292	Alto	Muito alto
10	1.000	0.000	0.000	0.000	Normal	Normal

Para se saber se o modelo podia ser usado para fazer previsões, foi feita uma comparação dos dois riscos referidos. Para tal, construiu-se um gráfico de barras (Figura 58).

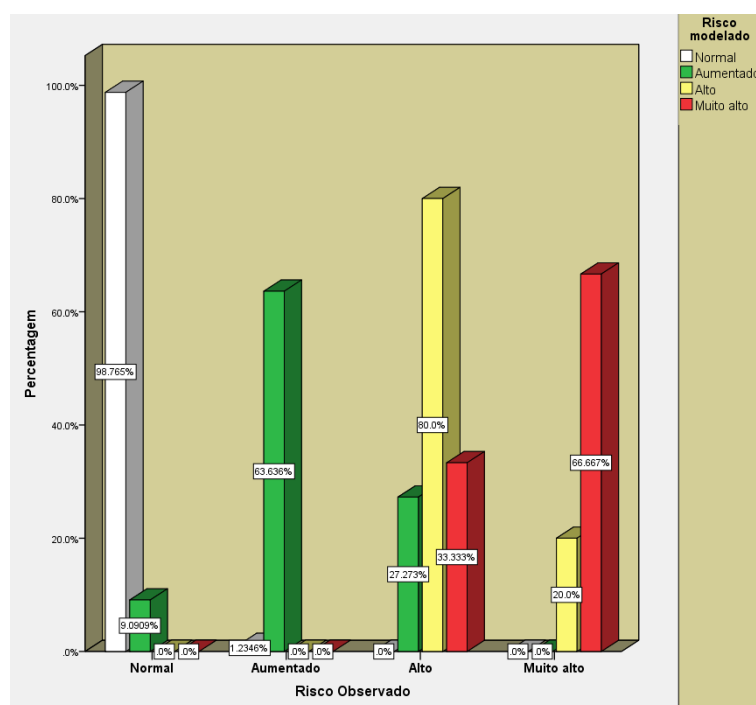


Figura 58: Gráfico de barras com a percentagem de risco modelado em cada uma das categorias do risco observado

O gráfico da Figura 58 mostrou que o modelo parecia ser válido, sendo que apenas parecia haver discrepâncias entre as categorias de risco alto e muito alto.

A tabela de contingência (Tabela 22) ideal seria uma em que as percentagens da diagonal fossem 100% e as restantes fossem 0%.

Tabela 22: Tabela de contingência comparativa de risco observado com o risco modelado (n=100)

		Risco modelado			
		Normal	Aumentado	Alto	Muito alto
Normal	Contagem	80	1	0	0
	% de risco observado	98.8%	1.2%	0.0%	0.0%
Aumentado	Contagem	1	7	0	0
	% de risco observado	12.5%	87.5%	0.0%	0.0%
Alto	Contagem	0	3	4	1
	% de risco observado	0.0%	37.5%	50.0%	12.5%
Muito alto	Contagem	0	0	1	2
	% de risco observado	0.0%	0.0%	33.3%	66.7%
Total	Contagem	81	11	5	3
	% de risco observado	81.0%	11.0%	5.0%	3.0%

A tabela mostra que dos 81 indivíduos com risco normal, apenas em um deles falhou a previsão, pois o indivíduo foi associado a risco aumentado. No entanto, nos restantes 80 a previsão foi acertada. A assertividade do modelo, para o risco normal, foi de 98.8%.

Nos 8 indivíduos com risco aumentado, 7 foram associados a risco aumentado, mas para o outro o risco foi subvalorizado. A assertividade do modelo, para o risco aumentado, foi de 87.5%.

Entre os 8 indivíduos com risco alto, o modelo apenas modelou metade com esse tipo de risco. Para 3 indivíduos o risco foi subvalorizado e para 1 deles foi sobrevalorizado. A assertividade do modelo, para o risco alto, foi de 50%.

No que diz respeito aos 3 indivíduos que estavam com risco muito alto, o modelo subvalorizou o risco de um deles, tendo uma assertividade de 66.7%.

A simples observação da tabela não permite concluir se o modelo é válido ou não, pois a sua capacidade preditiva parece ser melhor para os níveis menos graves de risco. No

entanto, esta validação deveria ser repetida para uma amostra maior, pois o número de indivíduos nas últimas categorias era bastante reduzido.

Foram ainda analisados os desvios entre o risco observado e o risco modelado e constatou-se, mais uma vez, que na maior parte dos indivíduos o risco foi bem modelado (Figura 59). Nos casos em que isso não aconteceu, em 71.4% dos indivíduos o risco foi subvalorizado e em 28.6% dos indivíduos o risco foi sobrevalorizado.

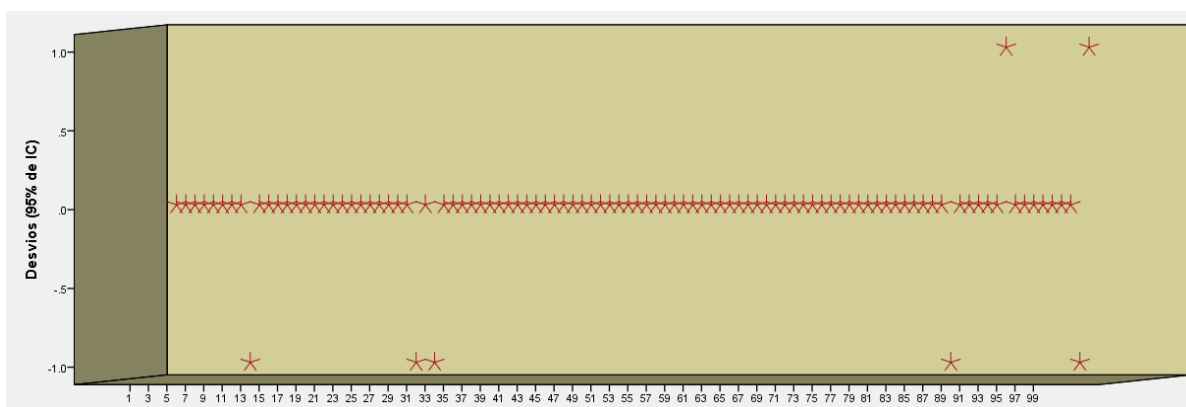


Figura 59: Distribuição dos desvios entre o risco observado e o risco modelado

Tentar encontrar um modelo perfeito é uma utopia, pelo que o objetivo de qualquer modelação deverá ser encontrar uma boa relação entre o risco observado e o risco modelado. De acordo com o coeficiente de Spearman, existe uma correlação significativa e extremamente forte ($\rho=.945$; $p=.000$) entre estes dois riscos, pelo que se concluiu que este modelo é válido e o origina boas identificações de fatores de risco, sobretudo nos casos de risco normal ou risco aumentado.

4.2. Modelação do Risco B

Entre os 600 participantes no estudo, relativamente ao risco B (considerando a hipertensão e o número de fatores de risco adicionais), 295 estavam com risco normal, 89 com risco acrescido (mas baixo), 120 com risco moderado, 84 com risco alto e 12 com risco muito alto. Para a sua modelação procedeu-se de forma análoga à do risco A (Anexo 14).

Da análise bivariada constatou-se que este risco foi fortemente correlacionado com a pressão arterial diastólica e com a hipertensão, moderadamente correlacionado com a pressão arterial sistólica, fracamente correlacionado com a idade, o peso, a estatura e o

género e muito fracamente correlacionado com o índice de massa corporal e com o perímetro abdominal.

Optou-se por não inserir no modelo variáveis muito fracamente correlacionadas com o risco B, mas considerou-se que seria necessário testar a inserção no modelo de algumas variáveis fracamente correlacionadas com esse risco. Foram consideradas para possíveis variáveis explicativas duas variáveis contínuas (idade e peso) e duas variáveis binárias (hipertensão e género). Entre os 15 modelos possíveis, houve 3 que representavam os dados de forma significativa (Tabela 23).

Tabela 23: Modelos que ajustam significativamente os dados, segundo o teste da razão de verosimilhanças

Modelo	Covariáveis	Estatística	p-value	BIC
M1	X1+X2+X3	-610.372	.653	1284.713
M2	X1+X2+X4	-610.853	.491	1285.675
M3	X1+X2	-612.480	.246	1282.533

x1- hipertensão; x2-idade; x3- peso; x4- género

Considerou-se que o modelo mais adequado seria o M1, por ser o que tem mais variáveis explicativas modificáveis pelos indivíduos. Para análise do ajustamento do modelo foram calculados os resíduos para cada um dos quatro contrastes e, recorrendo à função *summary()* verificou-se que a maior parte dos valores estão muito próximos de zero, indiciando um bom ajustamento global do modelo aos dados observados.

```
> Coef(ajustell)
A matrix:
          latvar
log(mu[,1]/mu[,5]) 1.0000000
log(mu[,2]/mu[,5]) 0.4924753
log(mu[,3]/mu[,5]) 0.5032344
log(mu[,4]/mu[,5]) 0.1902324

C matrix:
          latvar
sx1 -3.25728625
x2  -0.06401248
x3  -0.02762889

B1 matrix:
          log(mu[,1]/mu[,5]) log(mu[,2]/mu[,5]) log(mu[,3]/mu[,5]) log(mu[,4]/mu[,5])
(Intercept) 8.654899          5.965606          6.319477          3.723318
```

Figura 60: Output do software R da função *Coef()*

Tal como para o risco A, à medida que o nível de gravidade do risco B ia aumentando, o peso dos efeitos das covariáveis ia diminuindo. As três covariáveis foram associadas a

coeficientes negativos, permitindo concluir que o avançar da idade ou o aumento de peso estavam relacionados com um aumento do grau de risco cardiovascular (Figura 60).

Tabela 24: Coeficientes e Odds Ratio para cada uma das covariáveis do modelo estereótipo, para o Risco B

	Condição de Risco A							
	Normal vs Muito alto		Acrescido vs Muito alto		Aumentado vs Muito alto		Alto vs Muito alto	
	$\omega_j\beta$	OR	$\omega_j\beta$	OR	$\omega_j\beta$	OR	$\omega_j\beta$	OR
Hipertensão (Hipertenso)	8.655	.038	-1.604	.201	-1.639	.194	-.620	.538
Idade	-.064	.938	-.032	.969	-.032	.968	-.012	.988
Peso	-.028	.973	-.014	.986	-.014	.986	-.005	.995

OR-Odds Ratio; vs- versus

Pela Tabela 24, procedendo de forma análoga à do risco A, concluiu-se que um indivíduo hipertenso tinha 26 vezes mais probabilidade de estar com risco muito alto do que com risco normal; 5 vezes mais probabilidades de estar com risco muito alto do que com risco acrescido (mas baixo); 5 vezes mais probabilidades de estar com risco muito alto do que com risco moderado; duas vezes mais probabilidades de estar com risco muito alto do que com risco alto. A condição de hipertensão revelou-se como um fator de risco, como seria de esperar, tendo em conta o modelo preditivo do risco B.

No que diz respeito às variáveis contínuas foram construídos gráficos com as probabilidades, recorrendo à forma funcional de um modelo estereótipo (49), na qual:

Com o auxílio do *software* Geogebra, obtiveram-se os gráficos das Figura 61 e da Figura 62.

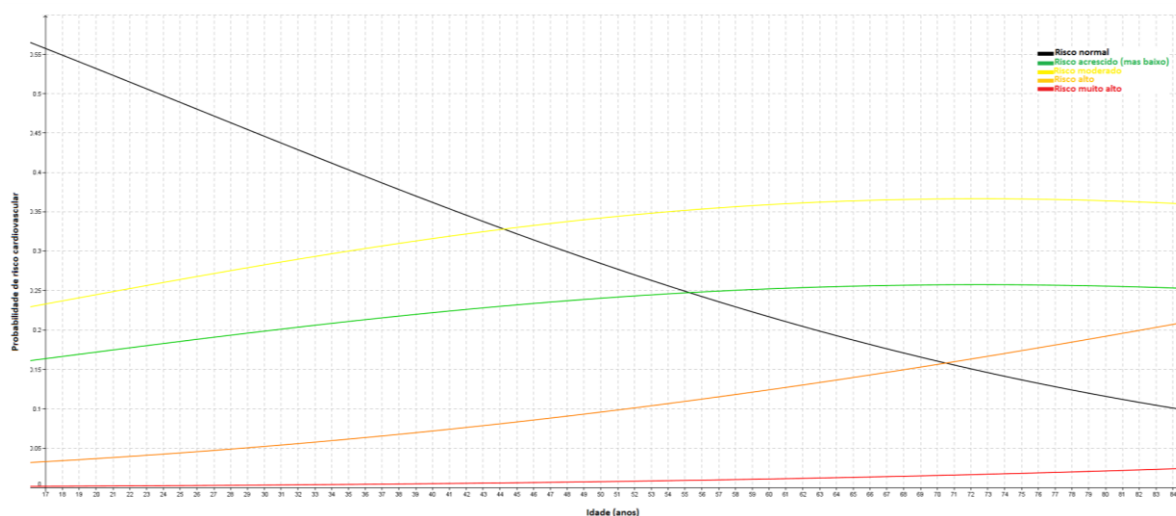


Figura 61: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função da idade

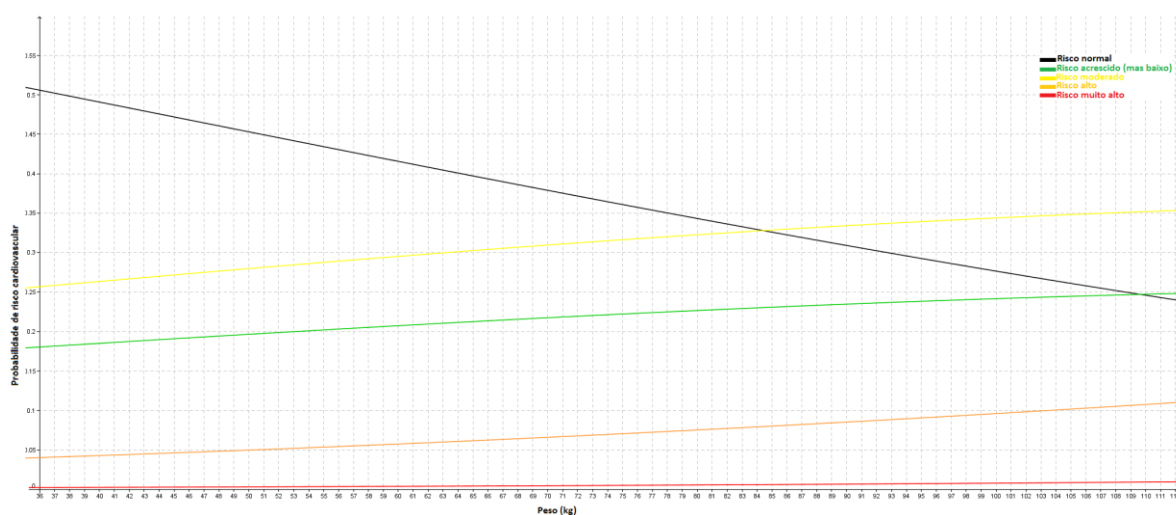


Figura 62: Gráficos com as probabilidades dos vários graus de risco cardiovascular, em função do peso (kg)

Este modelo apenas pode ser utilizado depois de ser validado. Recorrendo-se à amostra de validação e à forma funcional do modelo estereótipo, foram calculados os riscos B para cada um dos indivíduos da amostra de validação e os níveis foram comparados com os níveis obtidos pelo outro modelo preditivo.

O modelo preditivo utilizado para no risco B apresentou alguns desvios na previsão do risco para os 100 indivíduos da amostra de validação. O cálculo do coeficiente de Spearman mostrou que existe uma correlação significativa ($p=.000$) e moderada (.671) entre os níveis de risco B obtidos por estes dois modelos.

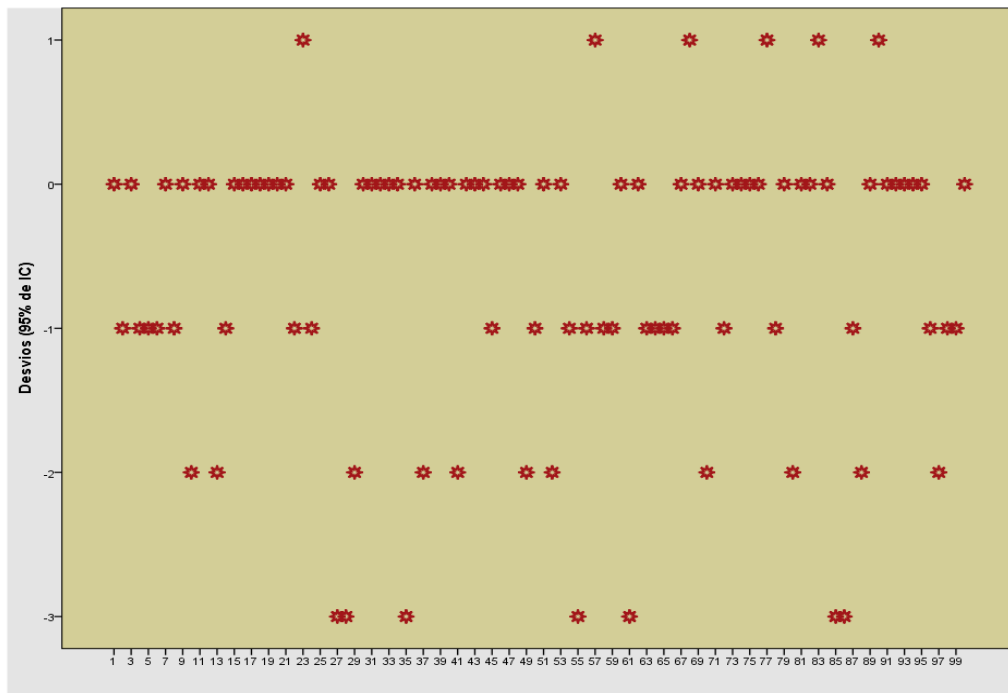


Figura 63: Distribuição dos desvios entre o risco observado e o risco modelado

Observando a Figura 63 constata-se que a maior parte das diferenças entre os dois modelos preditivos estão associadas a subvalorizações do risco, pelo que não é recomendável o seu uso.

4.3. Modelação do Risco C

O risco C calculado a partir do modelo preditivo que a OMS propôs para alguns países africanos não pode ser usado para modelação, uma vez que na análise bivariada este risco não foi significativamente correlacionado com nenhum indicador de risco, além das pressões arteriais, mostrando que este modelo preditivo parece não estar bem ajustado à realidade benguelense.

5. CONCLUSÕES E RECOMENDAÇÕES

Comparação do risco cardiovascular entre homens e mulheres

A primeira questão colocada no início da dissertação prendia-se com a comparação do perfil cardiovascular de homens e de mulheres. É importante salientar que o número de homens e de mulheres que participaram no estudo não foi igual, pois o número de homens que mostraram interesse em conhecer o seu perfil cardiovascular foi significativamente superior ao de mulheres. Na maior parte dos indicadores não foram encontrados motivos para que se fizessem afirmações sobre diferenças entre géneros, mas houve alguns que se destacaram.

Apesar de não terem sido apontadas diferenças significativas entre o perímetro abdominal de homens e mulheres, a verdade é que, considerando que os valores de referência não são os mesmos entre géneros, as mulheres tinham cinco vezes mais obesidade abdominal do que os homens.

Por seu lado, o género masculino foi apontado como grupo de risco para o desenvolvimento de hipertensão, de hábitos tabágicos e de hábitos alcoólicos. O consumo de tabaco era cinco vezes superior nos homens e o consumo de álcool era quatro vezes superior nos homens. Estes dois hábitos estão relacionados com o desenvolvimento de hipertensão e, como seria de esperar, a prevalência de hipertensão nos homens era o triplo da das mulheres.

Não foram encontradas evidências suficientes que justifiquem a afirmação de que um género tem maior risco cardiovascular do que o outro. Aparentemente, a prevalência de risco é a mesma nos dois grupos e apenas o perfil cardiovascular é distinto. É importante que as mulheres benguelenses sejam aconselhadas a reduzir o perímetro abdominal e que os homens benguelenses alterem alguns hábitos de vida.

Comparação do risco cardiovascular entre indivíduos mais velhos e mais novos

Os 600 indivíduos, no que concerne à idade, foram divididos em dois grupos. O ponto de corte nos homens foi nos 45 anos e nas mulheres foi nos 55 anos. Por uma questão de

facilitismo da comunicação, um dos grupos foi sendo referido como o dos mais velhos e o outro o dos mais novos. O número de indivíduos mais novos era significativamente superior ao de mais velhos.

As maiores diferenças encontradas entre estes dois grupos apontaram sempre o grupo dos mais velhos como grupo de risco. A prevalência de sobrepeso nos mais velhos era o dobro da dos mais novos. Detetou-se também uma prevalência de hipertensão nos mais velhos quatro vezes superior à dos mais novos. Os hábitos alcoólicos também se destacaram, sendo três vezes superior nos mais velhos, porém é importante salientar que no grupo de indivíduos que afirmaram ingerir álcool com regularidade, 72.6% são jovens. Da mesma forma, entre os indivíduos que consomem gorduras saturadas com regularidade, 82.4% são mais novos.

Apesar das informações do parágrafo anterior, não há fortes evidências que mostrem que os benguelenses mais velhos estejam com maior risco cardiovascular do que os mais novos. É importante referir a existência de muitos indivíduos mais novos com valores de índice de massa corporal, de perímetro abdominal e de pressões arteriais aberrantes, quando comparados com indivíduos do mesmo grupo. Estes *outliers* podem ser indicadores de uma tendência para o aumento destes fatores de risco entre os mais novos.

A sensibilização dos benguelenses para a redução de peso e dos valores da pressão arterial deve ser generalizada, no que diz respeito às idades. Os mais novos precisam de ter noção que a idade não está a funcionar como um fator de proteção.

Comparação do risco cardiovascular entre indivíduos com sobrepeso e sem sobrepeso

Os indivíduos foram também separados de acordo com o peso, os indivíduos com excesso de peso ou com obesidade formaram o grupo com sobrepeso e os restantes o grupo sem sobrepeso. O número observado de indivíduos com sobrepeso era significativamente inferior.

A relação entre sobrepeso e perímetro abdominal/obesidade abdominal foi muito forte, de tal forma que 74.8% dos indivíduos com sobrepeso também tinham obesidade abdominal. A prevalência de obesidade abdominal no grupo com sobrepeso era onze vezes superior à observada no grupo sem sobrepeso. As pessoas com sobrepeso também parecem estar

mais propensas a desenvolver hipertensão, pois a prevalência de hipertensão foi duas vezes superior no grupo com sobrepeso.

Os indivíduos com sobrepeso apresentaram maior risco cardiovascular (riscos A e B).

Em relação a este fator de risco deixam-se três fatos, para reflexão:

- A prevalência de sobrepeso registada neste conjunto de benguelenses foi de 35.0%, entre os quais 9.2% estavam obesos, quase tantos quantos os com baixo peso (9.7%);
- 83.3% dos benguelenses afirmaram consumir regularmente gorduras saturadas;
- Segundo previsões da OMS, em 2025, metade da população mundial estará obesa.

Comparação do risco cardiovascular entre indivíduos com obesidade abdominal e sem obesidade abdominal

O grupo em estudo foi também dividido de acordo com a obesidade abdominal, tendo-se obtido o grupo com obesidade abdominal e o grupo sem essa condição. O número de indivíduos com excesso de gordura abdominal era significativamente inferior ao que não tinham obesidade abdominal.

Além da relação, já referida anteriormente, entre o sobrepeso e a obesidade, apenas outro fator de risco apresentou diferenças significativas entre estes dois grupos: a hipertensão. Tal como havia acontecido no grupo com sobrepeso, o grupo com obesidade abdominal apresentou uma prevalência de hipertensão duas vezes superior à do grupo com indivíduos não obesos.

Não foram encontradas evidências que permitam afirmar que o grupo com obesidade abdominal apresente maior risco cardiovascular do que o grupo sem obesidade abdominal.

Comparação do risco cardiovascular entre hipertensos e não hipertensos

Os participantes foram divididos em dois grupos, de acordo com a pressão arterial. 29% dos indivíduos estavam hipertensos e os restantes estavam normotensos, tendo-se verificado que o número de hipertensos era significativamente inferior ao número de normotensos.

A hipertensão mostrou ser uma condição propensa ao agravamento do risco cardiovascular. Como já foi referido anteriormente, os homens, os mais velhos, os indivíduos com sobrepeso, os indivíduos com obesidade abdominal e os consumidores de álcool são grupos com maior risco de desenvolvimento de hipertensão.

Excluindo a idade e o género, que são condições inalteráveis, a redução do risco de hipertensão (e conseqüente redução do risco cardiovascular) pode ser bem-sucedida devido à alteração de alguns hábitos pouco saudáveis. É importante que toda a população tenha conhecimento dos fatores de risco cardiovasculares e do efeito bola de neve entre alguns deles.

Comparação do risco cardiovascular entre fumadores e não fumadores

Entre os 600 participantes no estudo encontravam-se 47 (7.8%) fumadores. Este valor foi inferior aos 10% registados em 2011 na região africana, pela OMS.

Além da relação já referida anteriormente com o género masculino verificou-se que os fumadores ingerem cinco vezes mais álcool do que os não fumadores e que os fumadores têm um terço das probabilidades de terem sobrepeso ou obesidade abdominal, comparativamente com os não fumadores.

Não foram encontradas evidências suficientes que permitissem afirmar que os fumadores tinham maior risco cardiovascular do que os não fumadores.

Comparação do risco cardiovascular entre indivíduos que ingerem álcool e que não ingerem álcool

32.8% dos indivíduos afirmaram ingerir álcool regularmente.

Além da relação já referida anteriormente entre o consumo de álcool e o género masculino, a idade mais avançada, a hipertensão e o tabagismo, verificou-se que os indivíduos que ingerem álcool consomem o dobro de gorduras saturadas, quando comparados com os que não ingerem.

Não foram encontradas evidências suficientes que permitissem afirmar que os consumidores de álcool tinham maior risco cardiovascular do que os não consumidores, no

entanto é importantíssimo que sejam definidas estratégias que culminem na adoção de estilos de vida mais saudáveis.

Modelação do risco cardiovascular

Tal como o conhecimento dos fatores de risco cardiovascular dos angolanos, também a análise de risco cardiovascular é bastante frágil. Até ao momento não se conseguiu encontrar na bibliografia consultada qualquer tentativa de análise de risco específica para o contexto angolano. A única análise de risco foi proposta pela OMS que, neste trabalho, correspondeu ao risco C. Constatou-se que este modelo preditivo parece não se adequar à realidade benguelense. Neste trabalho foi construído um modelo preditivo bastante simples através do qual, conhecendo o género, a condição de hipertensão, o perímetro abdominal e o índice de massa corporal se consegue associar um indivíduo a uma probabilidade de risco cardiovascular. Pretende-se que este modelo preditivo seja uma pequena amostra daquilo que é possível fazer, em grade escala, e que já foi feito, com sucesso, em países desenvolvidos, como o Framingham Heart Study, nas Américas, ou o SCORE Project, na Europa.

Assim, surge a necessidade de se planearem estratégias mais adequadas para diminuir o risco cardiovascular na população em geral, estudando mecanismos que apontem melhores formas de detetar atempadamente a doença e descobrindo fatores de predição que sejam fulcrais para a redução do desenvolvimento e da ocorrência de doenças cardiovasculares. É urgente que a investigação epidemiológica, em Angola, se volte agora para a aquisição de conhecimentos importantes no combate às doenças cardiovasculares. Há que conhecer o real impacto que esta enfermidade está a ter em Angola, conhecer o perfil cardiovascular dos angolanos, identificar os fatores de risco *major* desta população, promover estratégias para a redução da exposição aos fatores de risco modificáveis e desenvolver modelos preditivos, específicos, que permitam uma boa gestão do risco cardiovascular, em Angola.

A melhor estratégia é portanto mista, incluindo a populacional, de maior impacto, na qual a população é permanentemente sensibilizada para a importância da adoção de estilos de vida mais saudáveis, mas que deve ser associada a uma identificação do risco individual,

recorrendo a modelos preditivos de análise de risco, especificamente construídos para o contexto angolano. Só deste modo, será possível aumentar a qualidade e os anos de vida.

O modelo preditivo deste trabalho é bastante simplista, mas adaptado à realidade benguelense e aos recursos financeiros disponíveis para o mesmo. É importante que se construam técnicas de análise de risco simples e com baixo custo, compatíveis com a realidade do sistema de saúde, mas também que se incentive e financie a construção de modelos preditivos mais abrangentes (e, inevitavelmente, mais caros), tanto em termos de população como de fatores de risco, considerando outros indicadores importantes como os níveis de glicémia, de colesterol e de triglicéridos.

Sugere-se o início de um estudo de coorte, na província de Benguela, selecionando um conjunto de indivíduos assintomáticos, expostos a alguns fatores de risco, e acompanhando-os para comparar a ocorrência de eventos cardiovasculares.

A realização deste trabalho foi uma experiência extremamente gratificante, apesar de, em alguns momentos, a escassez de técnicas e de *software* específicos para o modelo estereótipo, ter sido desesperante. É urgente que mais atenções se voltem para a construção de técnicas e de *software* de apoio ao modelo estereótipo, pois é extremamente útil em análise de risco.

Espera-se que este estudo, bastante modesto, possa de alguma forma contribuir para o conhecimento da dinâmica das doenças cardiovasculares e chamar a atenção para a importância do desenvolvimento de metodologias para a análise do risco cardiovascular, em Angola.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu, M., Siqueira, A., & Caiaffa, W. (2009). Regressão logística ordinal em estudos epidemiológicos. *Rev. Saúde Pública*, 43, pp. 183-194.
- Adimari, G., & Ventura, L. (2001). Robust inference for generalized linear model with application to logistic regression. *Statistics & Probability Letters*, 55, pp. 413-419.
- Agresti, A. (2002). *Categorical Data Analysis* (2ª ed.). Gainesville: University of Florida.
- Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Nova Jersey: John Wiley and Sons.
- Ananth, C., & Kleinbaum, D. (1997). Regression models for ordinal responses: a review of methods and applications. *International Journal of Epidemiology* 26(6), 1323-33.
- Anderson, J. (1984). *Regression and ordered categorical variable*. . J R Stat Soc B.
- ANGOP. (21 de Novembro de 2013). *Feira de higiene bocal decorre em Benguela*. Obtido de Agência Angola Press:
http://www.portalangop.co.ao/angola/pt_pt/noticias/saude/2013/10/47/Feira-higiene-bocal-decorre-Benguela,90d346d1-f095-4e08-8642-0d8870738e68.html?fb_action_ids=679663338734921&fb_action_types=og.recommends&fb_source=aggregation&fb_aggregation_id=28838148
- Arrais, M., Lusitano, A., & Neto, F. (2004). *Tabagismo entre Profissionais de Saúde do HMP/IS*. (F. d. UAN, Ed.) Obtido de Revista Angolana de Educação Médica:
<http://www.fmuan.ao/resumodcinco.html>
- Bedogni, G., Kahn, H., Bellentani, S., & Tiribelli, C. (2010). A simple index of lipid overaccumulation is a good marker of liver steatosis. *BMC Gastroenterology*, 10:98. Obtido de <http://www.biomedcentral.com/1471-230X/10/98>
- Brant, R. (1990). *Assessing proportionality in the proportional odds model for ordinal logistic regression*. *Biometrics*.

- Carneiro, G., Faria, A. N., Ribeiro Filho, F. F., Guimarães, A., Lerário, D., Ferreira, S. R., & Zanella, M. T. (1992). Influência da distribuição da gordura corporal sobre a prevalência de hipertensão arterial e outros fatores de risco cardiovascular em indivíduos obesos. *Revista da Associação Médica Brasileira*. Obtido de <http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IscScript=iah/iah.xis&src=google&base=LILACS&lang=p&nextAction=Ink&exprSearch=349567&indexSearch=ID>
- Cohen, L., & Holliday, M. (1982). *Statistics for the social sciences*. London: Harper & Row.
- Conroy, R., Pyorola, K., Fitzgerald, A., Sans, S., Menotti, A., Backer, G., . . . Graham, I. (2002). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *European Heart Journal*. Obtido de <http://eurheartj.oxfordjournals.org/content/24/11/987.abstract>
- Dobson, A. (1990). *An Introduction to Generalized Linear Models*. London: Chapman & Hall.
- Framingham Heart Study*. (03 de Novembro de 2013). Obtido de A Project of the National Heart, Lung and Blood Institute and Boston University: <http://www.framinghamheartstudy.org/>
- Global Health Observatory (GHO). (2013). *10 leading causes of death in females*. Obtido de World Health Organization: http://www.who.int/gho/women_and_health/mortality/causes_death_text/en/
- Grundy, S., Pasternak, R., Greenland, P., Smith Jr, S., & Fuster, V. (1999). *Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: A statement for healthcare professionals from the American Heart Association and the American College of Cardiology*. American Heart Association and the American College of Cardiology.
- Hosmer, D., & Lemeshow, S. (1980). A goodness-of-fit test for the logistic regression model. *Communications in Statistics*, 1043-69.
- Hosmer, D., & Lemeshow, S. (2000). *Applied Logistic Regression*. New York: Wiley.

- Júnior, J., Mendes, J., Barbosa, D., & Lopes, A. (2011). Fatores de risco cardiovascular em adolescentes: prevalência e associação com fatores sociodemográficos. *Revista Brasileira de Epidemiologia*. Obtido de http://www.scielo.br/scielo.php?pid=S1415-790X2011000100005&script=sci_arttext&tIng=es
- Libby, P. (2005). *The vascular biology of atherosclerosis. Braunwald's Heart Diseases: A Textbook os Cardiovascular Medicine* (7ª ed.). Philadelphia: Elsevier Saunders.
- Long, J., & Freese, J. (2005). *Regression Models for Categorical Dependent Variables Using Stata*. STATA Press.
- Mancia, G. e. (2007). Guidelines for the Management of Arterial Hypertension - The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology. *Eur Heart Journal*.
- Marôco, J. (2007). *Análise Estatística com o SPSS Statistics* (3ª ed.). Edições Sílabo.
- Ministério da Saúde de Angola. (2012). *Plano Nacional de Desenvolvimento Sanitário 2012-2025*. Luanda: República de Angola. Obtido em 02 de Novembro de 2013, de <http://www.minsa.gov.ao/VerPublicacao.aspx?id=1083>
- Newton, R., & Rudestam, E. (1999). *Your statistical consultant: Answers to your data analysis questions*. Sage Publications, Inc.
- Oliveira, T. (2004). *Estatística Aplicada*. Lisboa: Universidade Aberta.
- Pestana, M., & Gageiro, J. (2008). *Análise de Dados para Ciências Sociais* (5ª ed.). Lisboa: Edições Sílabo.
- Pukstennis, E., & Robinson, T. (2004). *Goodness-of-ft tests for ordinal response regression models*. *Stat Med*.
- Silva, J. (1995). *Modelos Lineares Generalizados*. Lisboa: Universidade Técnica de Lisboa. Obtido de <http://run.unl.pt/bitstream/10362/7596/1/WP0024.pdf>
- Simão, M., Hayashida, M., Santos, C., Cesarino, E., & Nogueira, M. (julho-agosto de 2008). Hipertensão Arterial entre universitários da cidade do Lubango, Angola. *Rev Latino-am Enfermagem*. Obtido de www.scielo.br/pdf/rlae/v16n4/pt_04.pdf

Turkman, M., & Silva, G. (2000). *Modelos Lineares Generalizados - da teoria à prática*. Lisboa: Edições SPE.

Vaz, D., Santos, L., & Carneiro, A. V. (Janeiro de 2005). Factores de Risco: Conceitos e Implicações Práticas. *Cardiologia Baseada na Evidência*, pp. 121-131. Obtido de http://www.spc.pt/spc/default.aspx?redir=http://www.spc.pt/spc/AreaCientifica/Publicacoes/Revista/rpc_det.aspx?id=210&tipo=0&pesq=factores%20de%20risco

WHO. (2002). *The Atlas of Heart Disease and Stroke*. Geneve: WHO. Obtido de http://www.who.int/cardiovascular_diseases/resources/atlas/en/

WHO. (2006). *BMI Classification*. Obtido de WHO: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html

WHO. (2007). *Prevention of Cardiovascular Disease Pocket- Guidelines for Assessment and Management of Cardiovascular Risk*. Geneva: World Health Organization. Obtido de http://www.who.int/cardiovascular_diseases/publications/Pocket_GL_information/en/index.html

WHO. (2011a). *A Prioritized Research Agenda for Prevention and Control of Noncommunicable Diseases*. Geneva: World Health Organization. Obtido de http://www.who.int/cardiovascular_diseases/publications/ncd_agenda2011/en/index.html

WHO. (2011b). *Global Atlas on cardiovascular disease prevention and control*. Geneva: World Health Organization. Obtido de http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/index.html

WHO. (2012). *Enfrentar o Desafio da Saúde da Mulher em África*. Geneve: WHO. Obtido de <http://apps.who.int/iris/handle/10665/79598>

WHO. (2013a). *World Health Statistics*. Geneva: World Health Organization. Obtido de http://www.who.int/gho/publications/world_health_statistics/en/index.html

WHO. (2013b). *WHO Report on the Global Tobacco Epidemic, 2013 Enforcing bans on tobacco advertising, promotion and sponsorship*. Geneva: World Health

Organization. Obtido de

http://www.who.int/tobacco/global_report/2013/en/index.html

WHO. (2013c). *WHO Report on the Global Tobacco Epidemic, 2013 (Angola)*. Geneva:

World Health Organization. Obtido de

http://www.who.int/tobacco/surveillance/policy/country_profile/en/

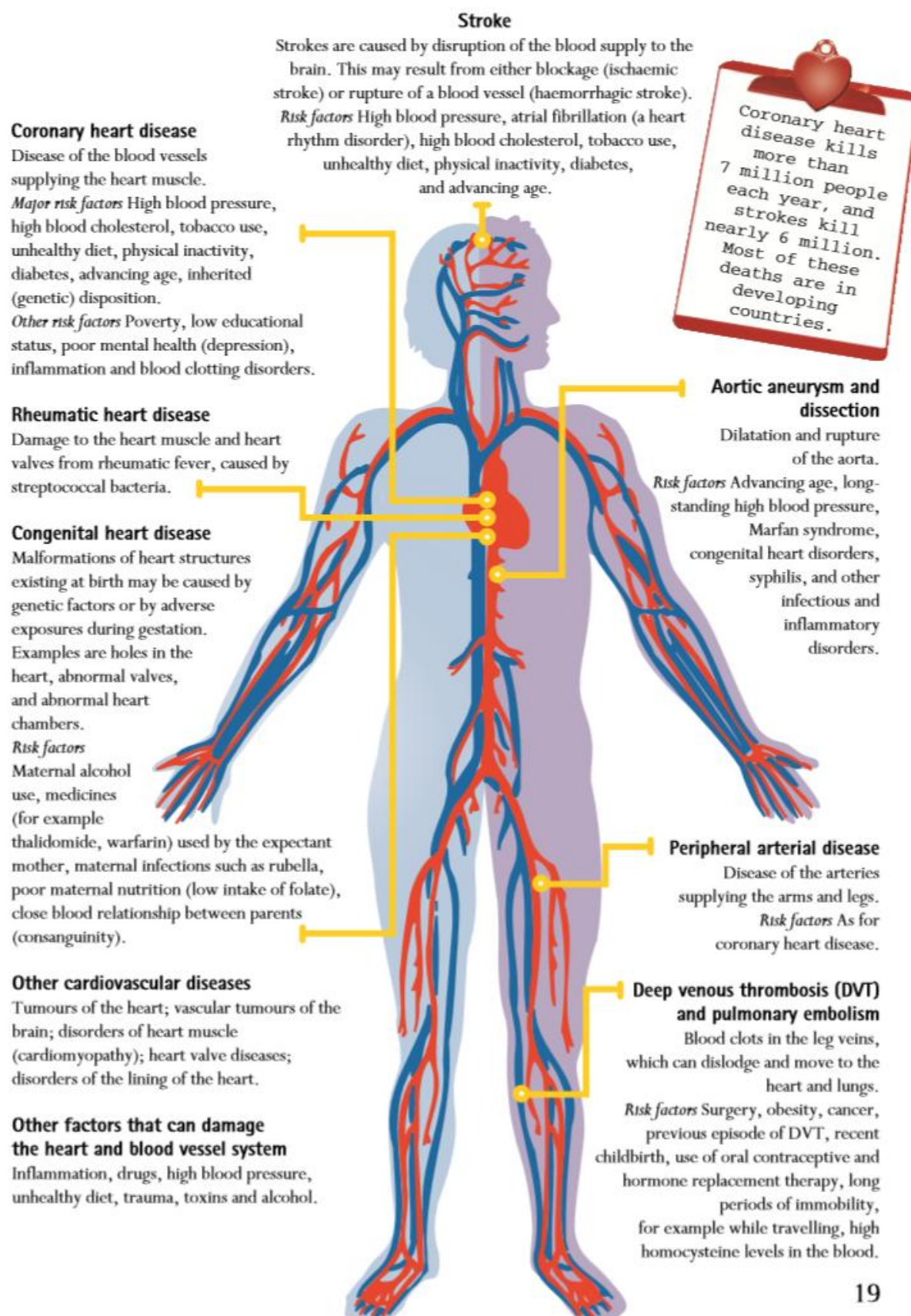
Yee, T. (Janeiro de 2010). The VGAM Package for Categorical Data Analysis. *Journal of*

Statistical Software, 32. Obtido de <http://www.jstatsoft.org/>

Yee, T., & Hastie, T. (2003). *Reduced-rank vector generalized linear models*. *Statistic Model*.

Yusuf, S., Hawken, S., Ôunpuu, S., Dans, T., Avezum, A., Lanas, F., & al, e. (2004). *Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study*. *Lancet*.

Anexo 1: Tipos de doenças cardiovasculares e fatores de riscos correspondentes



Anexo 2: Output do SPSS da Análise Univariada

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

Idade (anos)													
Statistic	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
		Lower Bound	Upper Bound										
Statistic	36.07	35.15	36.98	35.60	36.00	130.39	11.42	18	76	58	18	.422	-.309
Std. Error	.466											.100	.199
Peso (kg)													
Statistic	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
		Lower Bound	Upper Bound										
Statistic	65.75	64.64	66.86	65.18	64.00	191.74	13.85	37	112	75	18	.625	.367
Std. Error	.565											.100	.199
Estatura (m)													
Statistic	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
		Lower Bound	Upper Bound										
Statistic	1.66	1.66	1.67	1.66	1.67	0.01	0.08	1.45	1.91	.46	.11	-.196	-.396
Std. Error	.003											.100	.199

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

Índice de Massa Corporal (kg/m ²)													
Statistic	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
	Lower Bound	Upper Bound											
Statistic	23.78	23.41	24.15	23.53	23.00	21.59	4.65	14	46	32	5	1.012	2.264
Std. Error	.190											.100	.199

Perímetro abdominal (cm)													
Statistic	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
	Lower Bound	Upper Bound											
Statistic	84.76	83.74	85.79	84.24	83.00	162.90	12.76	50	132	82	17	.674	.643
Std. Error	.521											.100	.199

Pressão arterial sistólica (mmHg)													
Statistic	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
	Lower Bound	Upper Bound											
Statistic	118.19	116.50	119.88	117.04	120.00	444.01	21.07	80	200	120	30	.933	1.285
Std. Error	.860											.100	.199

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

Pressão arterial diastólica (mmHg)													
	Mean	95% Confidence Interval for Mean		5% Trimmed Mean	Median	Variance	Std. Deviation	Minimum	Maximum	Range	Interquartile Range	Skewness	Kurtosis
		Lower Bound	Upper Bound										
Statistic	78.92	77.66	80.18	78.11	80.00	247.19	15.72	40	160	120	20	.956	2.900
Std. Error	.642											.100	.199

Anexo 3: Output do SPSS da análise descritiva, por género

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

Género		Statistic	Std. Error			
Idade	Masculino	Mean	37.15	.600		
		95% Confidence Interval for Mean	Lower Bound	35.97		
			Upper Bound	38.33		
		5% Trimmed Mean		36.73		
		Median		38.00		
		Variance		127.977		
		Std. Deviation		11.313		
		Minimum		18		
		Maximum		76		
		Range		58		
		Interquartile Range		17		
		Skewness	.375	.129		
		Kurtosis	-.180	.258		
		Idade	Feminino	Mean	34.49	.729
				95% Confidence Interval for Mean	Lower Bound	33.05
					Upper Bound	35.93
				5% Trimmed Mean		33.98
				Median		34.00
Variance				130.218		
Std. Deviation				11.411		
Minimum				18		
Maximum				73		
Range				55		
Interquartile Range				18		
Skewness	.521			.156		
Kurtosis	-.408			.310		
Peso	Masculino			Mean	67.92	.706
				95% Confidence Interval for Mean	Lower Bound	66.53
					Upper Bound	69.31
				5% Trimmed Mean		67.55
				Median		67.00
		Variance		177.177		
		Std. Deviation		13.311		
		Minimum		39		
		Maximum		112		
		Range		73		
		Interquartile Range		19		
		Skewness	.446	.129		
		Kurtosis	-.029	.258		

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Mean	62.61	.896	
Feminino	95% Confidence Interval for Mean	Lower Bound	60.84		
		Upper Bound	64.37		
	5% Trimmed Mean		61.68		
	Median		61.00		
	Variance		196.895		
	Std. Deviation		14.032		
	Minimum		37		
	Maximum		110		
	Range		73		
	Interquartile Range		17		
	Skewness		1.002	.156	
	Kurtosis		1.464	.310	
	Masculino	Mean		1.7003	.00346
		95% Confidence Interval for Mean	Lower Bound	1.6935	
		Upper Bound	1.7071		
5% Trimmed Mean			1.7020		
Median			1.7000		
Variance			.004		
Std. Deviation			.06519		
Minimum			1.50		
Maximum			1.91		
Range			.41		
Interquartile Range			.08		
Skewness			-.289	.129	
Kurtosis			.351	.258	
Feminino		Mean		1.6058	.00461
	95% Confidence Interval for Mean	Lower Bound	1.5968		
		Upper Bound	1.6149		
	5% Trimmed Mean		1.6047		
	Median		1.6100		
	Variance		.005		
	Std. Deviation		.07209		
	Minimum		1.45		
	Maximum		1.82		
	Range		.37		
	Interquartile Range		.08		
	Skewness		.261	.156	
	Kurtosis		.014	.310	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

Índice de massa corporal	Masculino	Mean	23.42	.215		
		95% Confidence Interval for Mean	Lower Bound	23.00		
			Upper Bound	23.84		
		5% Trimmed Mean		23.31		
		Median		23.00		
		Variance		16.391		
		Std. Deviation		4.049		
		Minimum		14		
		Maximum		36		
		Range		22		
		Interquartile Range		5		
		Skewness		.383	.129	
		Kurtosis		-.097	.258	
		Mean		24.30	.343	
		Índice de massa corporal	Feminino	95% Confidence Interval for Mean	Lower Bound	23.63
	Upper Bound			24.98		
5% Trimmed Mean				23.91		
Median				24.00		
Variance				28.769		
Std. Deviation				5.364		
Minimum				15		
Maximum				46		
Range				31		
Interquartile Range				6		
Skewness				1.258	.156	
Kurtosis				2.468	.310	
Mean				84.62	.664	
Perímetro abdominal	Masculino			95% Confidence Interval for Mean	Lower Bound	83.31
					Upper Bound	85.92
		5% Trimmed Mean		84.22		
		Median		83.00		
		Variance		156.333		
		Std. Deviation		12.503		
		Minimum		50		
		Maximum		130		
		Range		80		
		Interquartile Range		18		
		Skewness		.519	.129	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Kurtosis		-.041	.258
		Mean		84.98	.840
		95% Confidence Interval	Lower Bound	83.32	
		for Mean	Upper Bound	86.63	
		5% Trimmed Mean		84.30	
		Median		83.00	
	Feminino	Variance		173.024	
		Std. Deviation		13.154	
		Minimum		50	
		Maximum		132	
		Range		82	
		Interquartile Range		16	
		Skewness		.865	.156
		Kurtosis		1.448	.310
		Mean		122.06	1.131
		95% Confidence Interval	Lower Bound	119.83	
		for Mean	Upper Bound	124.28	
		5% Trimmed Mean		120.98	
		Median		120.00	
	Masculino	Variance		454.451	
		Std. Deviation		21.318	
		Minimum		80	
		Maximum		200	
		Range		120	
		Interquartile Range		20	
		Skewness		.905	.129
		Kurtosis		1.171	.258
Pressão arterial sistólica		Mean		112.58	1.241
		95% Confidence Interval	Lower Bound	110.14	
		for Mean	Upper Bound	115.03	
		5% Trimmed Mean		111.40	
		Median		110.00	
	Feminino	Variance		377.351	
		Std. Deviation		19.426	
		Minimum		80	
		Maximum		200	
		Range		120	
		Interquartile Range		20	
		Skewness		1.030	.156
		Kurtosis		1.744	.310

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Mean	82.11	.865
Pressão arterial diastólica	Masculino	95% Confidence Interval for Mean	Lower Bound 80.41	Upper Bound 83.81
		5% Trimmed Mean	81.04	
		Median	80.00	
		Variance	265.324	
		Std. Deviation	16.289	
		Minimum	40	
		Maximum	160	
	Range	120		
	Interquartile Range	20		
	Skewness	1.193	.129	
	Kurtosis	3.315	.258	
	Feminino	Mean	74.29	.870
		95% Confidence Interval for Mean	Lower Bound 72.58	Upper Bound 76.01
		5% Trimmed Mean	74.09	
		Median	80.00	
Variance		185.610		
Std. Deviation		13.624		
Minimum		40		
Maximum		120		
Range		80		
Interquartile Range		20		
Skewness		.235	.156	
Kurtosis		.227	.310	

Anexo 4: Output do SPSS da análise descritiva, por idade

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Faixa etária	Statistic	Std. Error	
Idade	Mean	52.33	.674	
	95% Confidence Interval for Mean	Lower Bound 51.00 Upper Bound 53.67		
	5% Trimmed Mean	51.72		
	Median	50.00		
	Variance	47.647		
	Mais velhos	Std. Deviation	6.903	
	Minimum	45		
	Maximum	76		
	Range	31		
	Interquartile Range	10		
	Skewness	1.247	.236	
	Kurtosis	1.140	.467	
	Mean	32.61	.402	
	95% Confidence Interval for Mean	Lower Bound 31.82 Upper Bound 33.40		
	5% Trimmed Mean	32.39		
	Median	32.00		
	Variance	79.889		
	Mais novos	Std. Deviation	8.938	
	Minimum	18		
	Maximum	54		
Range	36			
Interquartile Range	16			
Skewness	.195	.110		
Kurtosis	-.948	.219		
Peso	Mean	69.70	1.434	
	95% Confidence Interval for Mean	Lower Bound 66.85 Upper Bound 72.54		
	5% Trimmed Mean	69.29		
	Median	69.00		
	Variance	215.925		
	Mais velhos	Std. Deviation	14.694	
	Minimum	45		
	Maximum	112		
	Range	67		
	Interquartile Range	23		
	Skewness	.354	.236	
	Kurtosis	-.422	.467	
	Mais novos	Mean	64.91	.608

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	63.72
		for Mean	Upper Bound	66.11
		5% Trimmed Mean		64.34
		Median		64.00
		Variance		183.023
		Std. Deviation		13.529
		Minimum		37
		Maximum		110
		Range		73
		Interquartile Range		16
		Skewness		.682 .110
		Kurtosis		.652 .219
Estatura		Mean		1.6888 .00684
		95% Confidence Interval	Lower Bound	1.6752
		for Mean	Upper Bound	1.7023
		5% Trimmed Mean		1.6905
		Median		1.6900
		Variance		.005
		Std. Deviation		.07013
	Mais velhos	Minimum		1.46
		Maximum		1.85
		Range		.39
		Interquartile Range		.08
		Skewness		-.390 .236
		Kurtosis		.717 .467
		Mean		1.6560 .00376
		95% Confidence Interval	Lower Bound	1.6486
		for Mean	Upper Bound	1.6634
		5% Trimmed Mean		1.6568
		Median		1.6600
		Variance		.007
		Std. Deviation		.08370
Mais novos	Minimum		1.45	
	Maximum		1.91	
	Range		.46	
	Interquartile Range		.12	
	Skewness		-.123 .110	
	Kurtosis		-.505 .219	
Índice de massa corporal		Mean		24.31 .426
	Mais velhos	95% Confidence Interval	Lower Bound	23.47
		for Mean	Upper Bound	25.16

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	5% Trimmed Mean		24.26
	Median		24.00
	Variance		19.045
	Std. Deviation		4.364
	Minimum		15
	Maximum		36
	Range		21
	Interquartile Range		6
	Skewness	.133	.236
	Kurtosis	-.321	.467
	Mean		23.67
	95% Confidence Interval	Lower Bound	23.25
	for Mean	Upper Bound	24.08
	5% Trimmed Mean		23.38
	Median		23.00
	Variance		22.101
Mais novos	Std. Deviation		4.701
	Minimum		14
	Maximum		46
	Range		32
	Interquartile Range		6
	Skewness	1.175	.110
	Kurtosis	2.778	.219
	Mean		89.74
	95% Confidence Interval	Lower Bound	87.18
	for Mean	Upper Bound	92.31
	5% Trimmed Mean		89.51
	Median		89.00
	Variance		175.501
Mais velhos	Std. Deviation		13.248
	Minimum		62
	Maximum		130
	Range		68
	Interquartile Range		20
	Skewness	.264	.236
	Kurtosis	-.236	.467
	Mean		83.71
	95% Confidence Interval	Lower Bound	82.61
	for Mean	Upper Bound	84.80
	5% Trimmed Mean		83.13
	Median		81.00

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum		50
	Maximum		160
	Range		110
	Interquartile Range		20
	Skewness	.993	.236
	Kurtosis	1.993	.467
	Mean	77.02	.649
	95% Confidence Interval for Mean	Lower Bound	75.75
		Upper Bound	78.30
	5% Trimmed Mean		76.48
	Median		80.00
	Variance		208.589
Mais novos	Std. Deviation		14.443
	Minimum		40
	Maximum		160
	Range		120
	Interquartile Range		10
	Skewness	.793	.110
	Kurtosis	2.886	.219

Anexo 5: Output do SPSS da análise descritiva, por condição de sobrepeso

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Sobrepeso	Statistic	Std. Error	
Idade	Com sobrepeso	Mean	39.40	.708	
		95% Confidence Interval for Mean	Lower Bound 38.01 Upper Bound 40.80		
		5% Trimmed Mean	39.38		
		Median	40.00		
		Variance	105.180		
		Std. Deviation	10.256		
		Minimum	18		
		Maximum	63		
		Range	45		
		Interquartile Range	15		
		Skewness	-.090	.168	
		Kurtosis	-.561	.334	
		Sem sobrepeso	Mean	34.27	.588
			95% Confidence Interval for Mean	Lower Bound 33.11 Upper Bound 35.42	
	5% Trimmed Mean		33.56		
	Median		33.00		
	Variance		135.008		
	Std. Deviation		11.619		
	Minimum		18		
	Maximum		76		
	Range		58		
	Interquartile Range		18		
	Skewness	.746	.124		
Kurtosis	.230	.247			
Peso	Com sobrepeso	Mean	79.00	.801	
		95% Confidence Interval for Mean	Lower Bound 77.42 Upper Bound 80.58		
		5% Trimmed Mean	78.59		
		Median	77.50		
		Variance	134.660		
		Std. Deviation	11.604		
		Minimum	53		
		Maximum	112		
		Range	59		
		Interquartile Range	15		
		Skewness	.550	.168	
		Kurtosis	.116	.334	
		o br	Mean	58.62	.445

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	57.74
		for Mean	Upper Bound	59.49
		5% Trimmed Mean		58.64
		Median		59.00
		Variance		77.086
		Std. Deviation		8.780
		Minimum		37
		Maximum		82
		Range		45
		Interquartile Range		12
		Skewness		-.041 .124
		Kurtosis		-.353 .247
		Mean		1.6601 .00602
		95% Confidence Interval	Lower Bound	1.6483
		for Mean	Upper Bound	1.6720
		5% Trimmed Mean		1.6613
	Com sobrepeso	Median		1.6600
		Variance		.008
		Std. Deviation		.08725
		Minimum		1.45
		Maximum		1.91
		Range		.46
		Interquartile Range		.12
		Skewness		-.215 .168
		Kurtosis		-.317 .334
Estatura		Mean		1.6626 .00404
		95% Confidence Interval	Lower Bound	1.6546
		for Mean	Upper Bound	1.6705
		5% Trimmed Mean		1.6636
		Median		1.6700
		Variance		.006
		Std. Deviation		.07975
		Minimum		1.45
		Maximum		1.85
		Range		.40
		Interquartile Range		.11
		Skewness		-.177 .124
		Kurtosis		-.481 .247
		Com sobrepeso	Mean	
Índice de massa corporal		95% Confidence Interval	Lower Bound	28.15
		for Mean	Upper Bound	29.17

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		5% Trimmed Mean	28.22	
		Median	28.00	
		Variance	14.006	
		Std. Deviation	3.742	
		Minimum	25	
		Maximum	46	
		Range	21	
		Interquartile Range	4	
		Skewness	1.972	.168
		Kurtosis	5.228	.334
		Mean	21.15	.124
		95% Confidence Interval for Mean	Lower Bound Upper Bound	20.91 21.40
		5% Trimmed Mean	21.24	
	Sem sobrepeso	Median	21.00	
		Variance	5.971	
		Std. Deviation	2.444	
		Minimum	14	
		Maximum	25	
		Range	11	
		Interquartile Range	3	
		Skewness	-.457	.124
		Kurtosis	-.498	.247
		Mean	96.18	.816
		95% Confidence Interval for Mean	Lower Bound Upper Bound	94.57 97.79
		5% Trimmed Mean	96.08	
	Com sobrepeso	Median	96.00	
		Variance	139.737	
		Std. Deviation	11.821	
		Minimum	63	
		Maximum	132	
		Range	69	
		Interquartile Range	14	
		Skewness	.215	.168
		Kurtosis	.738	.334
		Mean	78.62	.416
		95% Confidence Interval for Mean	Lower Bound Upper Bound	77.80 79.43
		5% Trimmed Mean	78.56	
	Sem sobrepeso	Median	78.00	
Perímetro abdominal				

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Variance	67.497		
		Std. Deviation	8.216		
		Minimum	50		
		Maximum	107		
		Range	57		
		Interquartile Range	10		
		Skewness	.116	.124	
		Kurtosis	.938	.247	
		Mean	125.19	1.603	
		95% Confidence Interval	Lower Bound	122.03	
		for Mean	Upper Bound	128.34	
		5% Trimmed Mean	123.91		
	Com sobrepeso	Median	120.00		
		Variance	539.310		
		Std. Deviation	23.223		
		Minimum	80		
		Maximum	200		
		Range	120		
		Interquartile Range	30		
		Skewness	.907	.168	
		Kurtosis	.832	.334	
Pressão arterial sistólica			Mean	114.42	.952
			95% Confidence Interval	Lower Bound	112.55
			for Mean	Upper Bound	116.29
			5% Trimmed Mean	113.46	
			Median	110.00	
		Variance	353.309		
		Std. Deviation	18.797		
		Minimum	80		
		Maximum	190		
		Range	110		
		Interquartile Range	20		
		Skewness	.776	.124	
		Kurtosis	1.004	.247	
			Mean	83.60	1.194
	Com sobrepeso	95% Confidence Interval	Lower Bound	81.25	
		for Mean	Upper Bound	85.95	
		5% Trimmed Mean	82.49		
		Median	80.00		
		Variance	299.227		
		Std. Deviation	17.298		
Pressão arterial diastólica					

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum	50	
	Maximum	160	
	Range	110	
	Interquartile Range	20	
	Skewness	1.288	.168
	Kurtosis	3.678	.334
	Mean	76.40	.719
	95% Confidence Interval for Mean	Lower Bound Upper Bound	74.98 77.81
	5% Trimmed Mean	75.92	
Sem sobrepeso	Median	80.00	
	Variance	201.669	
	Std. Deviation	14.201	
	Minimum	40	
	Maximum	130	
	Range	90	
	Interquartile Range	10	
	Skewness	.485	.124
	Kurtosis	.851	.247

Anexo 6: Output do SPSS da análise descritiva, por condição de obesidade

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Obesidade abdominal	Statistic	Std. Error	
Idade	Com obesidade	Mean	39.60	.684	
		95% Confidence Interval for Mean	Lower Bound 38.25	Upper Bound 40.94	
		5% Trimmed Mean	39.55		
		Median	40.00		
		Variance	112.284		
		Std. Deviation	10.596		
		Minimum	18		
		Maximum	65		
		Range	47		
		Interquartile Range	16		
		Skewness	-.069	.157	
		Kurtosis	-.532	.313	
		Sem obesidade	Mean	33.71	.598
			95% Confidence Interval for Mean	Lower Bound 32.53	Upper Bound 34.89
			5% Trimmed Mean	32.99	
	Median		32.00		
	Variance		128.919		
	Std. Deviation		11.354		
	Minimum		18		
	Maximum		76		
	Range		58		
	Interquartile Range		18		
	Peso	Com obesidade	Mean	73.49	.974
			95% Confidence Interval for Mean	Lower Bound 71.57	Upper Bound 75.41
			5% Trimmed Mean	73.21	
Median			72.50		
Variance			227.523		
Std. Deviation			15.084		
Minimum			39		
Maximum			112		
Range			73		
Interquartile Range			22		
Skewness			.249	.157	
Kurtosis			-.379	.313	
b			Mean	60.59	.531

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	59.54
		for Mean	Upper Bound	61.63
		5% Trimmed Mean		60.52
		Median		60.00
		Variance		101.669
		Std. Deviation		10.083
		Minimum		37
		Maximum		86
		Range		49
		Interquartile Range		14
		Skewness		.080 .129
		Kurtosis		-.428 .256
		Mean		1.6478 .00552
		95% Confidence Interval	Lower Bound	1.6369
		for Mean	Upper Bound	1.6586
		5% Trimmed Mean		1.6476
	Com obesidade	Median		1.6400
		Variance		.007
		Std. Deviation		.08552
		Minimum		1.45
		Maximum		1.91
		Range		.46
		Interquartile Range		.13
		Skewness		.057 .157
		Kurtosis		-.473 .313
		Mean		1.6710 .00416
		95% Confidence Interval	Lower Bound	1.6628
		for Mean	Upper Bound	1.6792
		5% Trimmed Mean		1.6730
	Sem obesidade	Median		1.6800
		Variance		.006
		Std. Deviation		.07898
		Minimum		1.45
		Maximum		1.85
		Range		.40
		Interquartile Range		.11
		Skewness		-.357 .129
		Kurtosis		-.174 .256
		Mean		27.02 .316
Índice de massa corporal	Com obesidad	95% Confidence Interval	Lower Bound	26.40
		for Mean	Upper Bound	27.64

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		5% Trimmed Mean	26.81	
		Median	26.50	
		Variance	23.891	
		Std. Deviation	4.888	
		Minimum	15	
		Maximum	46	
		Range	31	
		Interquartile Range	6	
		Skewness	.739	.157
		Kurtosis	1.857	.313
		Mean	21.62	.153
		95% Confidence Interval for Mean	Lower Bound Upper Bound	21.32 21.92
		5% Trimmed Mean	21.63	
	Sem obesidade	Median	22.00	
		Variance	8.453	
		Std. Deviation	2.907	
		Minimum	14	
		Maximum	30	
		Range	16	
		Interquartile Range	4	
		Skewness	-.013	.129
		Kurtosis	-.441	.256
		Mean	95.60	.710
		95% Confidence Interval for Mean	Lower Bound Upper Bound	94.20 96.99
		5% Trimmed Mean	94.99	
	Com obesidade	Median	95.00	
		Variance	120.961	
		Std. Deviation	10.998	
		Minimum	80	
		Maximum	132	
		Range	52	
		Interquartile Range	17	
		Skewness	.684	.157
		Kurtosis	.432	.313
Perímetro abdominal		Mean	77.54	.410
		95% Confidence Interval for Mean	Lower Bound Upper Bound	76.74 78.35
		5% Trimmed Mean	77.65	
	Sem obesidade	Median	76.00	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Variance	60.533	
		Std. Deviation	7.780	
		Minimum	50	
		Maximum	94	
		Range	44	
		Interquartile Range	10	
		Skewness	-.104	.129
		Kurtosis	.411	.256
		Mean	122.63	1.448
		95% Confidence Interval	Lower Bound	119.78
		for Mean	Upper Bound	125.49
		5% Trimmed Mean	121.54	
	Com obesidade	Median	120.00	
		Variance	503.170	
		Std. Deviation	22.431	
		Minimum	80	
		Maximum	200	
		Range	120	
		Interquartile Range	28	
		Skewness	.907	.157
		Kurtosis	.886	.313
Pressão arterial sistólica		Mean	115.23	1.033
		95% Confidence Interval	Lower Bound	113.20
		for Mean	Upper Bound	117.26
		5% Trimmed Mean	114.11	
	Sem obesidade	Median	120.00	
		Variance	383.870	
		Std. Deviation	19.593	
		Minimum	80	
		Maximum	200	
		Range	120	
		Interquartile Range	20	
		Skewness	.891	.129
		Kurtosis	1.487	.256
		Mean	81.63	1.080
		95% Confidence Interval	Lower Bound	79.50
		for Mean	Upper Bound	83.76
		5% Trimmed Mean	80.93	
	Com obesidade	Median	80.00	
		Variance	279.908	
		Std. Deviation	16.730	
Pressão arterial diastólica				

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum	40	
	Maximum	160	
	Range	120	
	Interquartile Range	20	
	Skewness	.872	.157
	Kurtosis	2.675	.313
	Mean	77.11	.778
	95% Confidence Interval for Mean	Lower Bound Upper Bound	75.58 78.64
	5% Trimmed Mean	76.31	
Sem obesidade	Median	80.00	
	Variance	217.915	
	Std. Deviation	14.762	
	Minimum	40	
	Maximum	160	
	Range	120	
	Interquartile Range	10	
	Skewness	.981	.129
	Kurtosis	3.128	.256

Anexo 7: Output do SPSS da análise descritiva, por condição de hipertensão

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Hipertensão	Statistic	Std. Error	
Idade	Hipertensos	Mean	40.92	.868	
		95% Confidence Interval for Mean	Lower Bound 39.21	Upper Bound 42.63	
		5% Trimmed Mean	40.80		
		Median	42.00		
		Variance	131.126		
		Std. Deviation	11.451		
		Minimum	18		
		Maximum	71		
		Range	53		
		Interquartile Range	14		
		Skewness	-.065	.184	
		Kurtosis	-.128	.366	
		Normaltensos	Mean	34.08	.524
			95% Confidence Interval for Mean	Lower Bound 33.05	Upper Bound 35.11
	5% Trimmed Mean		33.57		
	Median		32.00		
	Variance		116.810		
	Std. Deviation		10.808		
	Minimum		18		
	Maximum		76		
	Range		58		
	Interquartile Range		16		
	Peso	Hipertensos	Mean	71.11	1.155
95% Confidence Interval for Mean			Lower Bound 68.83	Upper Bound 73.39	
5% Trimmed Mean			70.68		
Median			69.50		
Variance			232.167		
Std. Deviation			15.237		
Minimum			41		
Maximum			112		
Range			71		
Interquartile Range			21		
Skewness			.428	.184	
Kurtosis			-.264	.366	
al te c Mean			63.56	.611	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	62.36	
		for Mean	Upper Bound	64.76	
		5% Trimmed Mean		63.11	
		Median		62.00	
		Variance		159.169	
		Std. Deviation		12.616	
		Minimum		37	
		Maximum		110	
		Range		73	
		Interquartile Range		16	
		Skewness		.598 .118	
		Kurtosis		.604 .236	
Estatura	Hipertensos	Mean		1.6868 .00604	
		95% Confidence Interval	Lower Bound	1.6749	
		for Mean	Upper Bound	1.6987	
		5% Trimmed Mean		1.6890	
		Median		1.7000	
		Variance		.006	
		Std. Deviation		.07972	
		Minimum		1.45	
		Maximum		1.91	
		Range		.46	
		Interquartile Range		.11	
		Skewness		-.346 .184	
		Kurtosis		.122 .366	
		Normaltensos	Mean		1.6515 .00394
			95% Confidence Interval	Lower Bound	1.6437
	for Mean		Upper Bound	1.6592	
	5% Trimmed Mean			1.6522	
	Median			1.6500	
	Variance			.007	
	Std. Deviation			.08134	
	Minimum			1.45	
	Maximum			1.85	
	Range			.40	
	Interquartile Range			.11	
	Skewness			-.149 .118	
Kurtosis			-.527 .236		
Índice de massa corporal	Hiperten s0s		Mean		24.98 .398
			95% Confidence Interval	Lower Bound	24.19
		for Mean	Upper Bound	25.76	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		5% Trimmed Mean	24.71		
		Median	24.00		
		Variance	27.560		
		Std. Deviation	5.250		
		Minimum	14		
		Maximum	45		
		Range	31		
		Interquartile Range	7		
		Skewness	.918	.184	
		Kurtosis	1.568	.366	
		Mean	23.29	.208	
		95% Confidence Interval for Mean	Lower Bound Upper Bound	22.88 23.70	
		5% Trimmed Mean	23.07		
	Normaltensos	Median	23.00		
		Variance	18.390		
		Std. Deviation	4.288		
		Minimum	15		
		Maximum	46		
		Range	31		
		Interquartile Range	6		
		Skewness	.960	.118	
		Kurtosis	2.445	.236	
			Mean	89.55	1.072
			95% Confidence Interval for Mean	Lower Bound Upper Bound	87.43 91.66
			5% Trimmed Mean	89.12	
		Hipertensos	Median	89.00	
	Variance		199.822		
	Std. Deviation		14.136		
	Minimum		60		
	Maximum		132		
	Range		72		
	Interquartile Range		21		
	Skewness		.430	.184	
	Kurtosis		.024	.366	
			Mean	82.81	.563
		95% Confidence Interval for Mean	Lower Bound Upper Bound	81.70 83.92	
		5% Trimmed Mean	82.32		
	Normaltensos	Median	81.00		

Perímetro abdominal

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Variance	135.067		
		Std. Deviation	11.622		
		Minimum	50		
		Maximum	130		
		Range	80		
		Interquartile Range	14		
		Skewness	.683	.118	
		Kurtosis	.962	.236	
		Mean	141.79	1.456	
		95% Confidence Interval for Mean	Lower Bound Upper Bound	138.91 144.66	
		5% Trimmed Mean	140.81		
		Median	140.00		
	Hipertensos	Variance	368.816		
		Std. Deviation	19.205		
		Minimum	100		
		Maximum	200		
		Range	100		
		Interquartile Range	30		
		Skewness	.696	.184	
		Kurtosis	.738	.366	
Pressão arterial sistólica			Mean	108.55	.602
			95% Confidence Interval for Mean	Lower Bound Upper Bound	107.37 109.74
		5% Trimmed Mean	108.71		
		Median	110.00		
		Variance	154.577		
		Std. Deviation	12.433		
		Minimum	80		
		Maximum	138		
		Range	58		
		Interquartile Range	20		
		Skewness	-.190	.118	
		Kurtosis	-.664	.236	
		Mean	97.05	1.014	
		95% Confidence Interval for Mean	Lower Bound Upper Bound	95.05 99.05	
		5% Trimmed Mean	95.89		
		Median	90.00		
		Variance	178.951		
		Std. Deviation	13.377		
Pressão arterial diastólica	Hipertensos				

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum	60	
	Maximum	160	
	Range	100	
	Interquartile Range	10	
	Skewness	1.839	.184
	Kurtosis	6.072	.366
	Mean	71.51	.449
	95% Confidence Interval for Mean	Lower Bound Upper Bound	70.63 72.39
	5% Trimmed Mean	72.15	
Normaltensos	Median	70.00	
	Variance	85.945	
	Std. Deviation	9.271	
	Minimum	40	
	Maximum	89	
	Range	49	
	Interquartile Range	20	
	Skewness	-.808	.118
	Kurtosis	.100	.236

Anexo 8: Output do SPSS da análise descritiva, de acordo com os hábitos tabágicos

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Tabagismo	Statistic	Std. Error	
Idade	Fumadores	Mean	40.28	1.672	
		95% Confidence Interval for Mean	Lower Bound 36.91	Upper Bound 43.64	
		5% Trimmed Mean	40.03		
		Median	41.00		
		Variance	131.335		
		Std. Deviation	11.460		
		Minimum	20		
		Maximum	65		
		Range	45		
		Interquartile Range	18		
		Skewness	.175	.347	
		Kurtosis	-.332	.681	
		Não fumadores	Mean	35.71	.483
			95% Confidence Interval for Mean	Lower Bound 34.76	Upper Bound 36.66
	5% Trimmed Mean		35.25		
	Median		35.00		
	Variance		128.910		
	Std. Deviation		11.354		
	Minimum		18		
	Maximum		76		
	Range		58		
	Interquartile Range		18		
	Skewness		.448	.104	
	Kurtosis		-.273	.207	
	Peso	Fumadores	Mean	61.91	1.498
			95% Confidence Interval for Mean	Lower Bound 58.90	Upper Bound 64.93
5% Trimmed Mean			61.35		
Median			60.00		
Variance			105.471		
Std. Deviation			10.270		
Minimum			45		
Não fumadores		Maximum	90		
		Range	45		
		Interquartile Range	15		
		Skewness	.767	.347	
		Kurtosis	.717	.681	
		fu m ^o Mean	66.08	.598	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	64.90
		for Mean	Upper Bound	67.25
		5% Trimmed Mean		65.52
		Median		65.00
		Variance		197.915
		Std. Deviation		14.068
		Minimum		37
		Maximum		112
		Range		75
		Interquartile Range		19
		Skewness		.592 .104
		Kurtosis		.296 .207
Estatura	Fumadores	Mean		1.6879 .01037
		95% Confidence Interval	Lower Bound	1.6670
		for Mean	Upper Bound	1.7088
		5% Trimmed Mean		1.6911
		Median		1.7000
		Variance		.005
		Std. Deviation		.07111
		Minimum		1.46
		Maximum		1.80
		Range		.34
		Interquartile Range		.09
		Skewness		-.806 .347
	Kurtosis		1.063 .681	
	Não fumadores	Mean		1.6595 .00353
		95% Confidence Interval	Lower Bound	1.6526
		for Mean	Upper Bound	1.6664
		5% Trimmed Mean		1.6604
		Median		1.6600
		Variance		.007
		Std. Deviation		.08295
		Minimum		1.45
		Maximum		1.91
		Range		.46
		Interquartile Range		.12
Skewness			-.148 .104	
Kurtosis		-.425 .207		
Índice de massa corporal	Fumador es	Mean		21.74 .513
		95% Confidence Interval	Lower Bound	20.71
		for Mean	Upper Bound	22.78

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		5% Trimmed Mean	21.57		
		Median	22.00		
		Variance	12.368		
		Std. Deviation	3.517		
		Minimum	15		
		Maximum	33		
		Range	18		
		Interquartile Range	4		
		Skewness	.779	.347	
		Kurtosis	1.516	.681	
		Mean	23.95	.200	
		95% Confidence Interval for Mean	Lower Bound Upper Bound	23.56 24.34	
		5% Trimmed Mean	23.71		
	Não fumadores	Median	24.00		
		Variance	22.020		
		Std. Deviation	4.692		
		Minimum	14		
		Maximum	46		
		Range	32		
		Interquartile Range	5		
		Skewness	.998	.104	
		Kurtosis	2.220	.207	
			Mean	79.98	1.587
			95% Confidence Interval for Mean	Lower Bound Upper Bound	76.78 83.17
			5% Trimmed Mean	79.65	
		Fumadores	Median	78.00	
	Variance		118.413		
	Std. Deviation		10.882		
	Minimum		50		
	Maximum		109		
	Range		59		
	Interquartile Range		13		
	Skewness		.501	.347	
	Kurtosis		1.124	.681	
			Mean	85.17	.546
		95% Confidence Interval for Mean	Lower Bound Upper Bound	84.10 86.24	
		5% Trimmed Mean	84.65		
	Não fumadores	Median	84.00		

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Variance	164.790	
		Std. Deviation	12.837	
		Minimum	50	
		Maximum	132	
		Range	82	
		Interquartile Range	17	
		Skewness	.671	.104
		Kurtosis	.594	.207
		Mean	120.72	3.059
		95% Confidence Interval for Mean	Lower Bound Upper Bound	114.57 126.88
		5% Trimmed Mean	119.28	
		Median	120.00	
	Fumadores	Variance	439.813	
	Fumadores	Std. Deviation	20.972	
	Fumadores	Minimum	90	
	Fumadores	Maximum	190	
	Fumadores	Range	100	
	Fumadores	Interquartile Range	23	
	Fumadores	Skewness	1.144	.347
	Fumadores	Kurtosis	1.837	.681
Pressão arterial sistólica		Mean	117.97	.897
		95% Confidence Interval for Mean	Lower Bound Upper Bound	116.21 119.74
		5% Trimmed Mean	116.85	
		Median	120.00	
	Não fumadores	Variance	444.575	
	Não fumadores	Std. Deviation	21.085	
	Não fumadores	Minimum	80	
	Não fumadores	Maximum	200	
	Não fumadores	Range	120	
	Não fumadores	Interquartile Range	30	
	Não fumadores	Skewness	.921	.104
	Não fumadores	Kurtosis	1.263	.207
		Mean	82.04	2.037
		95% Confidence Interval for Mean	Lower Bound Upper Bound	77.94 86.14
		5% Trimmed Mean	81.24	
		Median	80.00	
Pressão arterial diastólica	Fumadores	Variance	195.085	
	Fumadores	Std. Deviation	13.967	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum	60	
	Maximum	120	
	Range	60	
	Interquartile Range	20	
	Skewness	.759	.347
	Kurtosis	.970	.681
	Mean	78.65	.674
	95% Confidence Interval for Mean	Lower Bound Upper Bound	77.33 79.98
	5% Trimmed Mean	77.85	
Não fumadores	Median	80.00	
	Variance	251.082	
	Std. Deviation	15.846	
	Minimum	40	
	Maximum	160	
	Range	120	
	Interquartile Range	20	
	Skewness	.983	.104
	Kurtosis	3.040	.207

Anexo 9: Output do SPSS da análise descritiva, de acordo com os hábitos de ingestão de álcool

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Alcoolismo	Statistic	Std. Error	
Idade	Consumidor de álcool	Mean	38.34	.763	
		95% Confidence Interval for Mean	Lower Bound 36.84 Upper Bound 39.84		
		5% Trimmed Mean	38.09		
		Median	39.00		
		Variance	114.675		
		Std. Deviation	10.709		
		Minimum	18		
		Maximum	69		
		Range	51		
		Interquartile Range	18		
		Skewness	.192	.173	
		Kurtosis	-.419	.345	
		Não consumidor de álcool	Mean	34.95	.578
			95% Confidence Interval for Mean	Lower Bound 33.82 Upper Bound 36.09	
	5% Trimmed Mean		34.39		
	Median		34.00		
	Variance		134.602		
	Std. Deviation		11.602		
	Minimum		18		
	Maximum		76		
	Range		58		
	Interquartile Range		19		
	Skewness		.568	.122	
	Kurtosis		-.128	.243	
	Peso	Consumidor de álcool	Mean	68.87	.974
			95% Confidence Interval for Mean	Lower Bound 66.95 Upper Bound 70.79	
5% Trimmed Mean			68.47		
Median			68.00		
Variance			187.020		
Std. Deviation			13.676		
Minimum			40		
Maximum			112		
Range			72		
Interquartile Range			19		
Skewness			.523	.173	
Kurtosis			.189	.345	
Mean			64.22	.682	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	62.88
		for Mean	Upper Bound	65.56
		5% Trimmed Mean		63.58
		Median		63.00
		Variance		187.398
		Std. Deviation		13.689
		Minimum		37
		Maximum		110
		Range		73
		Interquartile Range		18
		Skewness		.710 .122
		Kurtosis		.587 .243
Estatura	Consumidor de álcool	Mean		1.6923 .00540
		95% Confidence Interval	Lower Bound	1.6817
		for Mean	Upper Bound	1.7030
		5% Trimmed Mean		1.6953
		Median		1.7000
		Variance		.006
		Std. Deviation		.07586
		Minimum		1.46
		Maximum		1.91
		Range		.45
		Interquartile Range		.09
		Skewness		-.529 .173
	Kurtosis		.362 .345	
	Não consumidor de álcool	Mean		1.6467 .00405
		95% Confidence Interval	Lower Bound	1.6388
		for Mean	Upper Bound	1.6547
		5% Trimmed Mean		1.6470
		Median		1.6400
		Variance		.007
		Std. Deviation		.08138
		Minimum		1.45
		Maximum		1.85
		Range		.40
		Interquartile Range		.11
Skewness			-.041 .122	
Kurtosis		-.436 .243		
Índice de massa corporal	Consumidor de álcool	Mean		24.02 .309
		95% Confidence Interval	Lower Bound	23.40
		for Mean	Upper Bound	24.63

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Variance	153.539	
		Std. Deviation	12.391	
		Minimum	50	
		Maximum	132	
		Range	82	
		Interquartile Range	16	
		Skewness	.776	.122
		Kurtosis	1.042	.243
Pressão arterial sistólica	Consumidor de álcool	Mean	121.91	1.589
		95% Confidence Interval for Mean	Lower Bound	118.78
			Upper Bound	125.04
		5% Trimmed Mean	120.66	
		Median	120.00	
		Variance	497.257	
		Std. Deviation	22.299	
		Minimum	80	
		Maximum	200	
		Range	120	
		Interquartile Range	20	
		Skewness	.919	.173
		Kurtosis	1.301	.345
	Não consumidor de álcool	Mean	116.37	1.008
		95% Confidence Interval for Mean	Lower Bound	114.39
			Upper Bound	118.35
		5% Trimmed Mean	115.28	
		Median	110.00	
		Variance	409.070	
		Std. Deviation	20.225	
Minimum		80		
Maximum		200		
Range		120		
	Interquartile Range	30		
	Skewness	.917	.122	
	Kurtosis	1.196	.243	
Pressão arterial diastólica	Consumidor de álcool	Mean	82.02	1.155
		95% Confidence Interval for Mean	Lower Bound	79.74
			Upper Bound	84.29
		5% Trimmed Mean	80.85	
		Median	80.00	
		Variance	262.668	
	Std. Deviation	16.207		

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum	50	
	Maximum	160	
	Range	110	
	Interquartile Range	20	
	Skewness	1.229	.173
	Kurtosis	3.370	.345
	Mean	77.40	.761
	95% Confidence Interval	Lower Bound	75.91
	for Mean	Upper Bound	78.90
	5% Trimmed Mean	76.82	
	Median	80.00	
	Variance	233.266	
	Std. Deviation	15.273	
Não consumidor de álcool	Minimum	40	
	Maximum	160	
	Range	120	
	Interquartile Range	10	
	Skewness	.804	.122
	Kurtosis	2.559	.243

Anexo 10: Output do SPSS da análise descritiva, de acordo com os hábitos de consumo de gorduras saturadas

		Gorduras Saturadas	Statistic	Std. Error	
Idade	Consumidor de gorduras saturadas	Mean	36.09	.515	
		95% Confidence Interval for Mean	Lower Bound 35.08	Upper Bound 37.10	
		5% Trimmed Mean	35.61		
		Median	36.00		
		Variance	132.565		
		Std. Deviation	11.514		
		Minimum	18		
		Maximum	76		
		Range	58		
		Interquartile Range	18		
		Skewness	.440	.109	
		Kurtosis	-.272	.218	
		Não consumidor de gorduras saturadas	Mean	35.95	1.099
			95% Confidence Interval for Mean	Lower Bound 33.77	Upper Bound 38.13
			5% Trimmed Mean	35.58	
	Median		36.00		
	Variance		120.735		
	Std. Deviation		10.988		
	Minimum		18		
	Maximum		65		
	Range		47		
	Interquartile Range		18		
	Peso	Consumidor de gorduras saturadas	Mean	66.07	.599
			95% Confidence Interval for Mean	Lower Bound 64.90	Upper Bound 67.25
			5% Trimmed Mean	65.56	
Median			65.00		
Variance			179.299		
Std. Deviation			13.390		
Minimum			37		
Maximum			112		
Range			75		
Interquartile Range			17		
Skewness			.607	.109	
Kurtosis			.331	.218	
Or d a			Mean	64.13	1.591

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		95% Confidence Interval	Lower Bound	60.97		
		for Mean	Upper Bound	67.29		
		5% Trimmed Mean		63.23		
		Median		62.00		
		Variance		253.185		
		Std. Deviation		15.912		
		Minimum		39		
		Maximum		110		
		Range		71		
		Interquartile Range		21		
		Skewness		.774 .241		
		Kurtosis		.459 .478		
Estatura	Consumidor de gorduras saturadas	Mean		1.6663 .00355		
		95% Confidence Interval	Lower Bound	1.6593		
		for Mean	Upper Bound	1.6733		
		5% Trimmed Mean		1.6671		
		Median		1.6700		
		Variance		.006		
		Std. Deviation		.07935		
		Minimum		1.45		
		Maximum		1.91		
		Range		.46		
		Interquartile Range		.12		
		Skewness		-.149 .109		
		Kurtosis		-.366 .218		
			Não consumidor de gorduras saturadas	Mean		1.6389 .00933
				95% Confidence Interval	Lower Bound	1.6204
		for Mean		Upper Bound	1.6574	
		5% Trimmed Mean			1.6396	
		Median			1.6350	
		Variance			.009	
		Std. Deviation			.09326	
		Minimum			1.45	
		Maximum			1.82	
		Range			.37	
			Interquartile Range		.14	
			Skewness		-.143 .241	
		Kurtosis		-.787 .478		
Índice de massa corporal	dor de gorduras saturadas	Mean		23.76 .196		
		95% Confidence Interval	Lower Bound	23.37		
		for Mean	Upper Bound	24.15		

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		5% Trimmed Mean	23.56	
		Median	23.00	
		Variance	19.221	
		Std. Deviation	4.384	
		Minimum	15	
		Maximum	45	
		Range	30	
		Interquartile Range	5	
		Skewness	.800	.109
		Kurtosis	1.259	.218
		Mean	23.88	.581
	Não consumidor de gorduras saturadas	95% Confidence Interval for Mean	Lower Bound	22.73
			Upper Bound	25.03
		5% Trimmed Mean	23.41	
		Median	23.00	
		Variance	33.763	
		Std. Deviation	5.811	
		Minimum	14	
		Maximum	46	
		Range	32	
		Interquartile Range	6	
		Skewness	1.403	.241
		Kurtosis	3.239	.478
		Mean	84.98	.567
	Consumidor de gorduras saturadas	95% Confidence Interval for Mean	Lower Bound	83.87
			Upper Bound	86.10
		5% Trimmed Mean	84.53	
		Median	84.00	
		Variance	160.533	
		Std. Deviation	12.670	
		Minimum	50	
		Maximum	132	
		Range	82	
		Interquartile Range	18	
		Skewness	.582	.109
		Kurtosis	.514	.218
Perímetro abdominal	Não consumidor de gorduras saturadas	Mean	83.66	1.323
		95% Confidence Interval for Mean	Lower Bound	81.04
			Upper Bound	86.28
		5% Trimmed Mean	82.80	
		Median	81.00	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Variance	175.015	
		Std. Deviation	13.229	
		Minimum	60	
		Maximum	130	
		Range	70	
		Interquartile Range	14	
		Skewness	1.125	.241
		Kurtosis	1.512	.478
		Mean	117.78	.939
		95% Confidence Interval	Lower Bound	115.93
		for Mean	Upper Bound	119.62
		5% Trimmed Mean	116.71	
		Median	120.00	
		Variance	440.570	
		Std. Deviation	20.990	
		Minimum	80	
		Maximum	200	
		Range	120	
		Interquartile Range	30	
		Skewness	.871	.109
		Kurtosis	1.164	.218
		Mean	120.25	2.146
		95% Confidence Interval	Lower Bound	115.99
		for Mean	Upper Bound	124.51
		5% Trimmed Mean	118.72	
		Median	120.00	
		Variance	460.715	
		Std. Deviation	21.464	
		Minimum	90	
		Maximum	200	
		Range	110	
		Interquartile Range	25	
		Skewness	1.247	.241
		Kurtosis	1.852	.478
		Mean	79.30	.706
		95% Confidence Interval	Lower Bound	77.91
		for Mean	Upper Bound	80.69
		5% Trimmed Mean	78.50	
		Median	80.00	
		Variance	249.300	
		Std. Deviation	15.789	
Pressão arterial sistólica	Consumidor de gorduras saturadas			
	Não consumidor de gorduras saturadas			
Pressão arterial diastólica	Consumidor de gorduras saturadas			

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	Minimum	40	
	Maximum	160	
	Range	120	
	Interquartile Range	20	
	Skewness	1.014	.109
	Kurtosis	3.293	.218
	Mean	77.02	1.532
Não consumidor de gorduras saturadas	95% Confidence Interval for Mean	Lower Bound	73.98
		Upper Bound	80.06
	5% Trimmed Mean		76.36
	Median		80.00
	Variance		234.707
	Std. Deviation		15.320
	Minimum		40
	Maximum		120
	Range		80
	Interquartile Range		10
	Skewness	.644	.241
	Kurtosis	.615	.478

Anexo 11: Tabela de Contingência relativa ao Risco A, por grupos de risco

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Risco A				Total
		Normal	Aumentado	Alto	Muito alto	
velhos	contagem	65	14	16	10	105
	% por faixa etária	61.9%	13.3%	15.2%	9.5%	100.0%
	% por Risco A	15.3%	20.9%	28.1%	19.6%	17.5%
faixa etária novos	contagem	360	53	41	41	495
	% por faixa etária	72.7%	10.7%	8.3%	8.3%	100.0%
	% por Risco A	84.7%	79.1%	71.9%	80.4%	82.5%
total	contagem	425	67	57	51	600
	% por faixa etária	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
masculino	contagem	281	34	20	20	355
	% por género	79.2%	9.6%	5.6%	5.6%	100.0%
	% por Risco A	66.1%	50.7%	35.1%	39.2%	59.2%
género feminino	contagem	144	33	37	31	245
	% por género	58.8%	13.5%	15.1%	12.7%	100.0%
	% por Risco A	33.9%	49.3%	64.9%	60.8%	40.8%
total	contagem	425	67	57	51	600
	% por género	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
sobrepeso com sobrepeso	contagem	53	49	57	51	210
	% por sobrepeso	25.2%	23.3%	27.1%	24.3%	100.0%
	% por Risco A	12.5%	73.1%	100.0%	100.0%	35.0%

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	contagem	372	18	0	0	390
sem sobrepeso	% por sobrepeso	95.4%	4.6%	0.0%	0.0%	100.0%
	% por Risco A	87.5%	26.9%	0.0%	0.0%	65.0%
	contagem	425	67	57	51	600
total	% por sobrepeso	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	65	67	57	51	240
com obesidade abdominal	% por obesidade abdominal	27.1%	27.9%	23.8%	21.3%	100.0%
	% por Risco A	15.3%	100.0%	100.0%	100.0%	40.0%
	contagem	360	0	0	0	360
obesidade abdominal sem obesidade abdominal	% por obesidade abdominal	100.0%	0.0%	0.0%	0.0%	100.0%
	% por Risco A	84.7%	0.0%	0.0%	0.0%	60.0%
	contagem	425	67	57	51	600
total	% por obesidade abdominal	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	103	23	24	24	174
hipertensos	% por hipertensão	59.2%	13.2%	13.8%	13.8%	100.0%
	% por Risco A	24.2%	34.3%	42.1%	47.1%	29.0%
	contagem	322	44	33	27	426
hipertensão normotensos	% por hipertensão	75.6%	10.3%	7.7%	6.3%	100.0%
	% por Risco A	75.8%	65.7%	57.9%	52.9%	71.0%
	contagem	425	67	57	51	600
total	contagem	425	67	57	51	600

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	% por hipertensão	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	42	3	1	1	47
fumadores	% por tabagismo	89.4%	6.4%	2.1%	2.1%	100.0%
	% por Risco A	9.9%	4.5%	1.8%	2.0%	7.8%
	contagem	383	64	56	50	553
tabagismo não fumadores	% por tabagismo	69.3%	11.6%	10.1%	9.0%	100.0%
	% por Risco A	90.1%	95.5%	98.2%	98.0%	92.2%
	contagem	425	67	57	51	600
total	% por tabagismo	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	137	24	20	16	197
consumidores de gorduras saturadas	% por alcoolismo	69.5%	12.2%	10.2%	8.1%	100.0%
	% por Risco A	32.2%	35.8%	35.1%	31.4%	32.8%
	contagem	288	43	37	35	403
alcoolismo não consumidores de gorduras saturadas	% por alcoolismo	71.5%	10.7%	9.2%	8.7%	100.0%
	% por Risco A	67.8%	64.2%	64.9%	68.6%	67.2%
	contagem	425	67	57	51	600
total	% por alcoolismo	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%
gorduras saturadas consumidores de gorduras saturadas	contagem	350	59	49	42	500
	% por gorduras saturadas	70.0%	11.8%	9.8%	8.4%	100.0%

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	% por Risco A	82.4%	88.1%	86.0%	82.4%	83.3%
não consumidores de gorduras saturadas	contagem	75	8	8	9	100
	% por gorduras saturadas	75.0%	8.0%	8.0%	9.0%	100.0%
	% por Risco A	17.6%	11.9%	14.0%	17.6%	16.7%
total	contagem	425	67	57	51	600
	% por gorduras saturadas	70.8%	11.2%	9.5%	8.5%	100.0%
	% por Risco A	100.0%	100.0%	100.0%	100.0%	100.0%

Anexo 12: Tabela de Contingência relativa ao Risco B, por grupos de risco

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Risco B					Total
		Normal	Acrescido (baixo)	Moderado	Alto	Muito alto	
velhos	contagem	27	18	30	24	6	105
	% por faixa etária	25.7%	17.1%	28.6%	22.9%	5.7%	100.0%
	% por Risco B	9.2%	20.2%	25.0%	28.6%	50.0%	17.5%
faixa etária novos	contagem	268	71	90	60	6	495
	% por faixa etária	54.1%	14.3%	18.2%	12.1%	1.2%	100.0%
	% por Risco B	90.8%	79.8%	75.0%	71.4%	50.0%	82.5%
total	contagem	295	89	120	84	12	600
	% por faixa etária	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
masculino	contagem	153	37	85	68	12	355
	% por género	43.1%	10.4%	23.9%	19.2%	3.4%	100.0%
	% por Risco B	51.9%	41.6%	70.8%	81.0%	100.0%	59.2%
género feminino	contagem	142	52	35	16	0	245
	% por género	58.0%	21.2%	14.3%	6.5%	0.0%	100.0%
	% por Risco B	48.1%	58.4%	29.2%	19.0%	0.0%	40.8%
total	contagem	295	89	120	84	12	600
	% por género	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
sobrepeso com sobrepeso	contagem	84	24	50	45	7	210
	% por sobrepeso	40.0%	11.4%	23.8%	21.4%	3.3%	100.0%

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	% por Risco B	28.5%	27.0%	41.7%	53.6%	58.3%	35.0%
sem sobrepeso	contagem	211	65	70	39	5	390
	% por sobrepeso	54.1%	16.7%	17.9%	10.0%	1.3%	100.0%
	% por Risco B	71.5%	73.0%	58.3%	46.4%	41.7%	65.0%
total	contagem	295	89	120	84	12	600
	% por sobrepeso	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
obesidade abdominal	contagem	103	39	49	44	5	240
	% por obesidade abdominal	42.9%	16.3%	20.4%	18.3%	2.1%	100.0%
	% por Risco B	34.9%	43.8%	40.8%	52.4%	41.7%	40.0%
sem obesidade abdominal	contagem	192	50	71	40	7	360
	% por obesidade abdominal	53.3%	13.9%	19.7%	11.1%	1.9%	100.0%
	% por Risco B	65.1%	56.2%	59.2%	47.6%	58.3%	60.0%
total	contagem	295	89	120	84	12	600
	% por obesidade abdominal	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
hipertensão	contagem	5	38	46	73	12	174
	% por hipertensão	2.9%	21.8%	26.4%	42.0%	6.9%	100.0%
	% por Risco B	1.7%	42.7%	38.3%	86.9%	100.0%	29.0%
normotensos	contagem	290	51	74	11	0	426
	% por hipertensão	68.1%	12.0%	17.4%	2.6%	0.0%	100.0%
	% por Risco B	98.3%	57.3%	61.7%	13.1%	0.0%	71.0%

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	contagem	295	89	120	84	12	600
total	% por hipertensão	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	14	4	15	12	2	47
fumadores	% por tabagismo	29.8%	8.5%	31.9%	25.5%	4.3%	100.0%
	% por Risco B	4.7%	4.5%	12.5%	14.3%	16.7%	7.8%
	contagem	281	85	105	72	10	553
tabagismo não fumadores	% por tabagismo	50.8%	15.4%	19.0%	13.0%	1.8%	100.0%
	% por Risco B	95.3%	95.5%	87.5%	85.7%	83.3%	92.2%
	contagem	295	89	120	84	12	600
total	% por tabagismo	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	81	15	45	48	8	197
consumidores de gorduras saturadas	% por alcoolismo	41.1%	7.6%	22.8%	24.4%	4.1%	100.0%
	% por Risco B	27.5%	16.9%	37.5%	57.1%	66.7%	32.8%
	contagem	214	74	75	36	4	403
alcoolismo não consumidores de gorduras saturadas	% por alcoolismo	53.1%	18.4%	18.6%	8.9%	1.0%	100.0%
	% por Risco B	72.5%	83.1%	62.5%	42.9%	33.3%	67.2%
	contagem	295	89	120	84	12	600
total	% por alcoolismo	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	contagem	247	64	99	78	12	500
gorduras saturadas	consumidores de	contagem	247	64	99	78	500

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

gorduras saturadas	% por gorduras saturadas	49.4%	12.8%	19.8%	15.6%	2.4%	100.0%
	% por Risco B	83.7%	71.9%	82.5%	92.9%	100.0%	83.3%
não consumidores de gorduras saturadas	contagem	48	25	21	6	0	100
	% por gorduras saturadas	48.0%	25.0%	21.0%	6.0%	0.0%	100.0%
	% por Risco B	16.3%	28.1%	17.5%	7.1%	0.0%	16.7%
total	contagem	295	89	120	84	12	600
	% por gorduras saturadas	49.2%	14.8%	20.0%	14.0%	2.0%	100.0%
	% por Risco B	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Anexo 13: Tabela de Contingência relativa ao Risco C, por grupos de risco

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

		Risco C			Total
		Baixo	Moderado	Alto	
velhos	contagem	97	6	2	105
	% por faixa etária	92.4%	5.7%	1.9%	100.0%
	% por Risco C	42.9%	66.7%	100.0%	44.3%
faixa etária novos	contagem	129	3	0	132
	% por faixa etária	97.7%	2.3%	0.0%	100.0%
	% por Risco C	57.1%	33.3%	0.0%	55.7%
total	contagem	226	9	2	237
	% por faixa etária	95.4%	3.8%	.8%	100.0%
	% por Risco C	100.0%	100.0%	100.0%	100.0%
masculino	contagem	145	8	2	155
	% por género	93.5%	5.2%	1.3%	100.0%
	% por Risco C	64.2%	88.9%	100.0%	65.4%
género feminino	contagem	81	1	0	82
	% por género	98.8%	1.2%	0.0%	100.0%
	% por Risco C	35.8%	11.1%	0.0%	34.6%
total	contagem	226	9	2	237
	% por género	95.4%	3.8%	.8%	100.0%
	% por Risco C	100.0%	100.0%	100.0%	100.0%
sobrepeso com sobrepeso	contagem	108	7	1	116
	% por sobrepeso	93.1%	6.0%	.9%	100.0%

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	% por Risco C	47.8%	77.8%	50.0%	48.9%	
	contagem	118	2	1	121	
sem sobrepeso	% por sobrepeso	97.5%	1.7%	.8%	100.0%	
	% por Risco C	52.2%	22.2%	50.0%	51.1%	
	contagem	226	9	2	237	
total	% por sobrepeso	95.4%	3.8%	.8%	100.0%	
	% por Risco C	100.0%	100.0%	100.0%	100.0%	
	contagem	123	6	1	130	
com obesidade abdominal	% por obesidade abdominal	94.6%	4.6%	.8%	100.0%	
	% por Risco C	54.4%	66.7%	50.0%	54.9%	
	contagem	103	3	1	107	
obesidade abdominal	sem obesidade abdominal	% por obesidade abdominal	96.3%	2.8%	.9%	100.0%
	% por Risco C	45.6%	33.3%	50.0%	45.1%	
	contagem	226	9	2	237	
total	% por obesidade abdominal	95.4%	3.8%	.8%	100.0%	
	% por Risco C	100.0%	100.0%	100.0%	100.0%	
	contagem	92	9	2	103	
hipertensos	% por hipertensão	89.3%	8.7%	1.9%	100.0%	
	% por Risco C	40.7%	100.0%	100.0%	43.5%	
hipertensão	contagem	134	0	0	134	
	% por hipertensão	100.0%	0.0%	0.0%	100.0%	
normotensos	% por Risco C	59.3%	0.0%	0.0%	56.5%	

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

	contagem	226	9	2	237	
total	% por hipertensão	95.4%	3.8%	.8%	100.0%	
	% por Risco C	100.0%	100.0%	100.0%	100.0%	
	<hr/>					
	contagem	27	0	0	27	
fumadores	% por tabagismo	100.0%	0.0%	0.0%	100.0%	
	% por Risco C	11.9%	0.0%	0.0%	11.4%	
	<hr/>					
tabagismo não fumadores	contagem	199	9	2	210	
	% por tabagismo	94.8%	4.3%	1.0%	100.0%	
	% por Risco C	88.1%	100.0%	100.0%	88.6%	
<hr/>						
total	contagem	226	9	2	237	
	% por tabagismo	95.4%	3.8%	.8%	100.0%	
	% por Risco C	100.0%	100.0%	100.0%	100.0%	
<hr/>						
consumidores de gorduras saturadas	contagem	90	5	1	96	
	% por alcoolismo	93.8%	5.2%	1.0%	100.0%	
	% por Risco C	39.8%	55.6%	50.0%	40.5%	
<hr/>						
alcoolismo não consumidores de gorduras saturadas	contagem	136	4	1	141	
	% por alcoolismo	96.5%	2.8%	.7%	100.0%	
	% por Risco C	60.2%	44.4%	50.0%	59.5%	
<hr/>						
total	contagem	226	9	2	237	
	% por alcoolismo	95.4%	3.8%	.8%	100.0%	
	% por Risco C	100.0%	100.0%	100.0%	100.0%	
<hr/>						
gorduras saturadas	consumidores de gorduras	contagem	192	7	2	201

Modelação Estatística:

Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela

saturadas	% por gorduras saturadas	95.5%	3.5%	1.0%	100.0%
	% por Risco C	85.0%	77.8%	100.0%	84.8%
não consumidores de gorduras saturadas	contagem	34	2	0	36
	% por gorduras saturadas	94.4%	5.6%	0.0%	100.0%
	% por Risco C	15.0%	22.2%	0.0%	15.2%
total	contagem	226	9	2	237
	% por gorduras saturadas	95.4%	3.8%	.8%	100.0%
	% por Risco C	100.0%	100.0%	100.0%	100.0%

Anexo 14: Comandos do *software* R usados no ajuste dos modelos de Regressão Ordinal

```
> require(VGAM)
> dados<-read.table("C:/Users/Ana/Desktop/cardiovascular.txt",header=TRUE)
> risco_a<-dados$risco_a
> x6<-dados$genero
> x4<-dados$idade
> estatura<-dados$estatura
> x3<-dados$peso
> sobrepeso<-dados$sobrepeso
> obesidade<-dados$obesidade
> x2<-dados$imc
> x1<-dados$perimetro
> pas<-dados$pas
> pad<-dados$pad
> x5<-dados$hta
> tabagismo<-dados$tabagismo
> alcoolismo<-dados$alcoolismo
> gorduras<-dados$gorduras
> #AVALIAÇÃO DO AJUSTAMENTO PARA TODOS OS MODELOS#
> sdados<-transform(dados,sx6=-scale(x6),sx5=-scale(x5))
> sx5<-sdados$sx5
> sx6<-sdados$sx6
> ajuste64<-
rrvglm(formula=risco_a~x1+ x2+ x3+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste64)
> AIC(ajuste64)
> ajuste01<-rrvglm(formula=risco_a~x1,family=multinomial,data=sdados)
> deviance(ajuste01)
> AIC(ajuste01)
> pchisq(deviance(ajuste01)-deviance(ajuste64),df=5,lower.tail=FALSE)
> ajuste02<-rrvglm(formula=risco_a~x2,family=multinomial,data=sdados)
> deviance(ajuste02)
> AIC(ajuste02)
> pchisq(deviance(ajuste02)-deviance(ajuste64),df=5,lower.tail=FALSE)
> ajuste03<-rrvglm(formula=risco_a~x3,family=multinomial,data=sdados)
> deviance(ajuste03)
> AIC(ajuste03)
> pchisq(deviance(ajuste03)-deviance(ajuste64),df=5,lower.tail=FALSE)
> ajuste04<-rrvglm(formula=risco_a~x4,family=multinomial,data=sdados)
> deviance(ajuste04)
> AIC(ajuste04)
> pchisq(deviance(ajuste04)-deviance(ajuste64),df=5,lower.tail=FALSE)
> ajuste05<-rrvglm(formula=risco_a~sx5,family=multinomial,data=sdados)
> deviance(ajuste05)
> AIC(ajuste05)
> pchisq(deviance(ajuste05)-deviance(ajuste64),df=5,lower.tail=FALSE)
> ajuste06<-rrvglm(formula=risco_a~sx6,family=multinomial,data=sdados)
> deviance(ajuste06)
> AIC(ajuste06)
> pchisq(deviance(ajuste06)-deviance(ajuste64),df=5,lower.tail=FALSE)
> ajuste07<-rrvglm(formula=risco_a~x1+ x2,family=multinomial,data=sdados)
> deviance(ajuste07)
> AIC(ajuste07)
```

```
> pchisq(deviance(ajuste07)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste08<-rrvglm(formula=risco_a~x1+ x3,family=multinomial,data=sdados)
> deviance(ajuste08)
> AIC(ajuste08)
> pchisq(deviance(ajuste08)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste09<-rrvglm(formula=risco_a~x1+ x4,family=multinomial,data=sdados)
> deviance(ajuste09)
> AIC(ajuste09)
> pchisq(deviance(ajuste09)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste10<-rrvglm(formula=risco_a~x1+ sx5,family=multinomial,data=sdados)
> deviance(ajuste10)
> AIC(ajuste10)
> pchisq(deviance(ajuste10)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste11<-rrvglm(formula=risco_a~x1+ sx6,family=multinomial,data=sdados)
> deviance(ajuste11)
> AIC(ajuste11)
> pchisq(deviance(ajuste11)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste12<-rrvglm(formula=risco_a~x2+ x3,family=multinomial,data=sdados)
> deviance(ajuste12)
> AIC(ajuste12)
> pchisq(deviance(ajuste12)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste13<-rrvglm(formula=risco_a~x2+ x4,family=multinomial,data=sdados)
> deviance(ajuste13)
> AIC(ajuste13)
> pchisq(deviance(ajuste13)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste14<-rrvglm(formula=risco_a~x2+ sx5,family=multinomial,data=sdados)
> deviance(ajuste14)
> AIC(ajuste14)
> pchisq(deviance(ajuste14)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste15<-rrvglm(formula=risco_a~x2+ sx6,family=multinomial,data=sdados)
> deviance(ajuste15)
> AIC(ajuste15)
> pchisq(deviance(ajuste15)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste16<-rrvglm(formula=risco_a~x3+ x4,family=multinomial,data=sdados)
> deviance(ajuste16)
> AIC(ajuste16)
> pchisq(deviance(ajuste16)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste17<-rrvglm(formula=risco_a~x3+ sx5,family=multinomial,data=sdados)
> deviance(ajuste17)
> AIC(ajuste17)
> pchisq(deviance(ajuste17)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste18<-rrvglm(formula=risco_a~x3+ sx6,family=multinomial,data=sdados)
> deviance(ajuste18)
> AIC(ajuste18)
> pchisq(deviance(ajuste18)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste19<-rrvglm(formula=risco_a~x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste19)
> AIC(ajuste19)
> pchisq(deviance(ajuste19)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste20<-rrvglm(formula=risco_a~x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste20)
> AIC(ajuste20)
```

```
> pchisq(deviance(ajuste20)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste21<-rrvglm(formula=risco_a~sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste21)
> AIC(ajuste21)
> pchisq(deviance(ajuste21)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste22<-rrvglm(formula=risco_a~x2+ sx5,family=multinomial,data=sdados)
> deviance(ajuste22)
> AIC(ajuste22)
> pchisq(deviance(ajuste22)-deviance(ajuste64),df=4,lower.tail=FALSE)
> ajuste23<-rrvglm(formula=risco_a~x1+ x2+ x3,family=multinomial,data=sdados)
> deviance(ajuste23)
> AIC(ajuste23)
> pchisq(deviance(ajuste23)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste24<-rrvglm(formula=risco_a~x1+ x2+ x4,family=multinomial,data=sdados)
> deviance(ajuste24)
> AIC(ajuste24)
> pchisq(deviance(ajuste24)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste25<-rrvglm(formula=risco_a~x1+ x2+ sx5,family=multinomial,data=sdados)
> deviance(ajuste25)
> AIC(ajuste25)
> pchisq(deviance(ajuste25)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste26<-rrvglm(formula=risco_a~x1+ x2+ sx6,family=multinomial,data=sdados)
> deviance(ajuste26)
> AIC(ajuste26)
> pchisq(deviance(ajuste26)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste27<-rrvglm(formula=risco_a~x1+ x3+ x4,family=multinomial,data=sdados)
> deviance(ajuste27)
> AIC(ajuste27)
> pchisq(deviance(ajuste27)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste28<-rrvglm(formula=risco_a~x1+ x3+ sx5,family=multinomial,data=sdados)
> deviance(ajuste28)
> AIC(ajuste28)
> pchisq(deviance(ajuste28)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste29<-rrvglm(formula=risco_a~x1+ x3+ sx6,family=multinomial,data=sdados)
> deviance(ajuste29)
> AIC(ajuste29)
> pchisq(deviance(ajuste29)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste30<-rrvglm(formula=risco_a~x1+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste30)
> AIC(ajuste30)
> pchisq(deviance(ajuste30)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste31<-rrvglm(formula=risco_a~x1+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste31)
> AIC(ajuste31)
> pchisq(deviance(ajuste31)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste32<-rrvglm(formula=risco_a~x1+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste32)
> AIC(ajuste32)
> pchisq(deviance(ajuste32)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste33<-rrvglm(formula=risco_a~x2+ x3+ x4,family=multinomial,data=sdados)
> deviance(ajuste33)
> AIC(ajuste33)
```

```
> pchisq(deviance(ajuste33)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste34<-rrvglm(formula=risco_a~x2+ x3+ sx5,family=multinomial,data=sdados)
> deviance(ajuste34)
> AIC(ajuste34)
> pchisq(deviance(ajuste34)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste35<-rrvglm(formula=risco_a~x2+ x3+ sx6,family=multinomial,data=sdados)
> deviance(ajuste35)
> AIC(ajuste35)
> pchisq(deviance(ajuste35)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste36<-rrvglm(formula=risco_a~x2+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste36)
> AIC(ajuste36)
> pchisq(deviance(ajuste36)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste37<-rrvglm(formula=risco_a~x2+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste37)
> AIC(ajuste37)
> pchisq(deviance(ajuste37)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste38<-rrvglm(formula=risco_a~x2+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste38)
> AIC(ajuste38)
> pchisq(deviance(ajuste38)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste39<-rrvglm(formula=risco_a~x3+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste39)
> AIC(ajuste39)
> pchisq(deviance(ajuste39)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste40<-rrvglm(formula=risco_a~x3+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste40)
> AIC(ajuste40)
> pchisq(deviance(ajuste40)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste41<-rrvglm(formula=risco_a~x3+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste41)
> AIC(ajuste41)
> pchisq(deviance(ajuste41)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste42<-rrvglm(formula=risco_a~x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste42)
> AIC(ajuste42)
> pchisq(deviance(ajuste42)-deviance(ajuste64),df=3,lower.tail=FALSE)
> ajuste43<-rrvglm(formula=risco_a~x1+ x2+ x3+ x4,family=multinomial,data=sdados)
> deviance(ajuste43)
> AIC(ajuste43)
> pchisq(deviance(ajuste43)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste44<-rrvglm(formula=risco_a~x1+ x2+ x3+ sx5,family=multinomial,data=sdados)
> deviance(ajuste44)
> AIC(ajuste44)
> pchisq(deviance(ajuste44)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste45<-rrvglm(formula=risco_a~x1+ x2+ x3+ sx6,family=multinomial,data=sdados)
> deviance(ajuste45)
> AIC(ajuste45)
> pchisq(deviance(ajuste45)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste46<-rrvglm(formula=risco_a~x1+ x2+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste46)
> AIC(ajuste46)
```

```
> pchisq(deviance(ajuste46)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste47<-rrvglm(formula=risco_a~x1+ x2+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste47)
> AIC(ajuste47)
> pchisq(deviance(ajuste47)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste48<-rrvglm(formula=risco_a~x1+ x2+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste48)
> AIC(ajuste48)
> pchisq(deviance(ajuste48)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste49<-rrvglm(formula=risco_a~x1+ x3+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste49)
> AIC(ajuste49)
> pchisq(deviance(ajuste49)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste50<-rrvglm(formula=risco_a~x1+ x3+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste50)
> AIC(ajuste50)
> pchisq(deviance(ajuste50)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste51<-rrvglm(formula=risco_a~x1+ x3+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste51)
> AIC(ajuste51)
> pchisq(deviance(ajuste51)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste52<-rrvglm(formula=risco_a~x1+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste52)
> AIC(ajuste52)
> pchisq(deviance(ajuste52)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste53<-rrvglm(formula=risco_a~x2+ x3+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste53)
> AIC(ajuste53)
> pchisq(deviance(ajuste53)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste54<-rrvglm(formula=risco_a~x2+ x3+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste54)
> AIC(ajuste54)
> pchisq(deviance(ajuste54)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste55<-rrvglm(formula=risco_a~x2+ x3+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste55)
> AIC(ajuste55)
> pchisq(deviance(ajuste55)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste56<-rrvglm(formula=risco_a~x2+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste56)
> AIC(ajuste56)
> pchisq(deviance(ajuste56)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste57<-rrvglm(formula=risco_a~x3+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste57)
> AIC(ajuste57)
> pchisq(deviance(ajuste57)-deviance(ajuste64),df=2,lower.tail=FALSE)
> ajuste58<-rrvglm(formula=risco_a~x1+ x2+ x3+ x4+ sx5,family=multinomial,data=sdados)
> deviance(ajuste58)
> AIC(ajuste58)
> pchisq(deviance(ajuste58)-deviance(ajuste64),df=1,lower.tail=FALSE)
> ajuste59<-rrvglm(formula=risco_a~x1+ x2+ x3+ x4+ sx6,family=multinomial,data=sdados)
> deviance(ajuste59)
> AIC(ajuste59)
```

```
> pchisq(deviance(ajuste59)-deviance(ajuste64),df=1,lower.tail=FALSE)
> ajuste60<-rrvglm(formula=risco_a~x1+ x2+ x3+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste60)
> AIC(ajuste60)
> pchisq(deviance(ajuste60)-deviance(ajuste64),df=1,lower.tail=FALSE)
> ajuste61<-rrvglm(formula=risco_a~x1+ x2+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste61)
> AIC(ajuste61)
> pchisq(deviance(ajuste61)-deviance(ajuste64),df=1,lower.tail=FALSE)
> ajuste62<-rrvglm(formula=risco_a~x1+ x3+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste62)
> AIC(ajuste62)
> pchisq(deviance(ajuste62)-deviance(ajuste64),df=1,lower.tail=FALSE)
> ajuste63<-rrvglm(formula=risco_a~x2+ x3+ x4+ sx5+ sx6,family=multinomial,data=sdados)
> deviance(ajuste63)
> AIC(ajuste63)
> pchisq(deviance(ajuste63)-deviance(ajuste64),df=1,lower.tail=FALSE)
> #SIGNIFICANCIA DO AJUSTAMENTO (P>=0.05)#
> pchisq(deviance(ajuste61)-deviance(ajuste64),df=1,lower.tail=FALSE)
> pchisq(deviance(ajuste59)-deviance(ajuste64),df=1,lower.tail=FALSE)
> pchisq(deviance(ajuste47)-deviance(ajuste64),df=2,lower.tail=FALSE)
> pchisq(deviance(ajuste26)-deviance(ajuste64),df=3,lower.tail=FALSE)
> #ESTIMAÇÃO DE PARÂMETROS PARA O MODELO 4 (AJUSTE 47)#
> ajuste48<-rrvglm(formula=risco_a~x1+ x2+ sx5+ sx6,multinomial,sdados)
> summary(ajuste48)
> Coef(ajuste48)
> coef(ajuste48,matrix=TRUE)
> residuos<-residuals(ajuste48)
> x<-resíduos[,1]
> y<-resíduos[,2]
> z<-resíduos[,3]
> library(scatterplot3d)
> scatterplot3d(x,y,z, pch=22, highlight.3d=TRUE,type="h",xlim=c(-9,27),ylim=c(-9,50),zlim=c(-8,15))
> #PARA O PERIMETRO
> normal1<-201.252-1.349*x1-2.711*mean(x2)-0.591*mean(sx5)+ 9.284*mean(sx6)
> aumentado1<-105.316-0.671*x1-1.350*mean(x2)-0.294*mean(sx5)+ 4.622*mean(sx6)
> alto1<-36.910-0.223*x1-0.448*mean(x2)-0.098*mean(sx5)+ 1.533*mean(sx6)
> pnormal1<-exp(normal1)/(exp(normal1)+ exp(aumentado1)+ exp(alto1)+ 1)
> paumentado1<-exp(aumentado1)/(exp(normal1)+ exp(aumentado1)+ exp(alto1)+ 1)
> palto1<-exp(alto1)/(exp(normal1)+ exp(aumentado1)+ exp(alto1)+ 1)
> pmtalto1<-1/(exp(normal1)+ exp(aumentado1)+ exp(alto1)+ 1)
> #PARA O PESO
> normal2<-201.252-1.349*mean(x1)-2.711*x2-0.591*mean(sx5)+ 9.284*mean(sx6)
> aumentado2<-105.316-0.671*mean(x1)-1.350*x2-0.294*mean(sx5)+ 4.622*mean(sx6)
> alto2<-36.910-0.223*mean(x1)-0.448*x2-0.098*mean(sx5)+ 1.533*mean(sx6)
> pnormal2<-exp(normal2)/(exp(normal2)+ exp(aumentado2)+ exp(alto2)+ 1)
> paumentado2<-exp(aumentado2)/(exp(normal2)+ exp(aumentado2)+ exp(alto2)+ 1)
> palto2<-exp(alto2)/(exp(normal2)+ exp(aumentado2)+ exp(alto2)+ 1)
> pmtalto2<-1/(exp(normal2)+ exp(aumentado2)+ exp(alto2)+ 1)
> #PARA A HIPERTENSAO
> normal3<-201.252-1.349*mean(x1)-2.711*mean(x2)-0.591*sx5+ 9.284*mean(sx6)
```

```
> aumentado3<-105.316-0.671*mean(x1)-1.350*mean(x2)-0.294*sx5+ 4.622*mean(sx6)
> alto3<-36.910-0.223*mean(x1)-0.448*mean(x2)-0.098*sx5+ 1.533*mean(sx6)
> pnormal3<-exp(normal3)/(exp(normal3)+ exp(aumentado3)+ exp(alto3)+ 1)
> paumentado3<-exp(aumentado3)/(exp(normal3)+ exp(aumentado3)+ exp(alto3)+ 1)
> palto3<-exp(alto3)/(exp(normal3)+ exp(aumentado3)+ exp(alto3)+ 1)
> pmtalto3<-1/(exp(normal3)+ exp(aumentado3)+ exp(alto3)+ 1)
> #PARA O GENERO
> normal4<-201.252-1.349*mean(x1)-2.711*mean(x2)-0.591*mean(sx5)+ 9.284*sx6
> aumentado4<-105.316-0.671*mean(x1)-1.350*mean(x2)-0.294*mean(sx5)+ 4.622*sx6
> alto4<-36.910-0.223*mean(x1)-0.448*mean(x2)-0.098*mean(sx5)+ 1.533*sx6
> pnormal4<-exp(normal4)/(exp(normal4)+ exp(aumentado4)+ exp(alto4)+ 1)
> paumentado4<-exp(aumentado4)/(exp(normal4)+ exp(aumentado4)+ exp(alto4)+ 1)
> palto4<-exp(alto4)/(exp(normal4)+ exp(aumentado4)+ exp(alto4)+ 1)
> pmtalto4<-1/(exp(normal4)+ exp(aumentado4)+ exp(alto4)+ 1)
> #GRAFICOS
> plot(x1,pnormal1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=24)
> plot(x2,pnormal2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=24)
> plot(x1,paumentado1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=3)
> plot(x2,paumentado2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=3)
> plot(x1,palto1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=655)
> plot(x2,palto2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=655)
> plot(x1,pmtalto1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=2)
> plot(x2,pmtalto2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=2)
VALIDAÇÃO
> validacao<-read.table("C:/Users/Ana/Desktop/validacao.txt",header=TRUE)
> risco_a<-validacao$risco_a
> id<-validacao$id
> perimetro<-validacao$perimetro
> imc<-validacao$imc
> hta<-validacao$hta
> genero<-validacao$genero
> srisco_a<--scale(risco_a)
> shta<--scale(hta)
> sgenero<--scale(genero)
> v1<-(exp(201.252-1.349*perimetro-2.711*imc-0.591*shta+ 9.284*sgenero))
> v2<-(exp(105.316-0.671*perimetro-1.350*imc-0.294*shta+ 4.622*sgenero))
> v3<-(exp(36.910-0.223*perimetro-0.448*imc-0.098*shta+ 1.553*sgenero))
> c(v1/(v1+ v2+ v3+ 1), v2/(v1+ v2+ v3+ 1), v3/(v1+ v2+ v3+ 1), 1/(v1+ v2+ v3+ 1))
> pnormal<-v1/(v1+ v2+ v3+ 1)
> pnormal
> paumentado<-v2/(v1+ v2+ v3+ 1)
> paumentado
> palto<-v3/(v1+ v2+ v3+ 1)
> palto
> pmtalto<-1/(v1+ v2+ v3+ 1)
> pmtalto
> c(pnormal,paumentado,palto,pmtalto)
># RISCO B#
> require(VGAM)
> dados<-read.table("C:/Users/Ana/Desktop/cardiovascular.txt",header=TRUE)
> risco_b<-dados$risco_b
> x4<-dados$genero
```

```
> x2<-dados$idade
> x3<-dados$peso
> x1<-dados$hta
> #AVALIAÇÃO DO AJUSTAMENTO PARA TODOS OS MODELOS#
> sdados<-transform(dados,sx4=-scale(x4),sx1=-scale(x1))
> sx1<-sdados$sx1
> sx4<-sdados$sx4
> ajuste15<-rrvglm(formula=risco_b~sx1+ x2+ x3+ sx4,family=multinomial,data=sdados)
> deviance(ajuste15)
> AIC(ajuste15)
> ajuste01<-rrvglm(formula=risco_b~sx1,family=multinomial,data=sdados)
> deviance(ajuste01)
> AIC(ajuste01)
> pchisq(deviance(ajuste01)-deviance(ajuste15),df=5,lower.tail=FALSE)
> ajuste02<-rrvglm(formula=risco_b~x2,family=multinomial,data=sdados)
> deviance(ajuste02)
> AIC(ajuste02)
> pchisq(deviance(ajuste02)-deviance(ajuste15),df=5,lower.tail=FALSE)
> ajuste03<-rrvglm(formula=risco_b~x3,family=multinomial,data=sdados)
> deviance(ajuste03)
> AIC(ajuste03)
> pchisq(deviance(ajuste03)-deviance(ajuste15),df=5,lower.tail=FALSE)
> ajuste04<-rrvglm(formula=risco_b~sx4,family=multinomial,data=sdados)
> deviance(ajuste04)
> AIC(ajuste04)
> pchisq(deviance(ajuste04)-deviance(ajuste15),df=5,lower.tail=FALSE)
> ajuste05<-rrvglm(formula=risco_b~sx1+ x2,family=multinomial,data=sdados)
> deviance(ajuste05)
> AIC(ajuste05)
> pchisq(deviance(ajuste05)-deviance(ajuste15),df=5,lower.tail=FALSE)
> ajuste06<-rrvglm(formula=risco_b~sx1+ x3,family=multinomial,data=sdados)
> deviance(ajuste06)
> AIC(ajuste06)
> pchisq(deviance(ajuste06)-deviance(ajuste15),df=5,lower.tail=FALSE)
> ajuste07<-rrvglm(formula=risco_b~sx1+ sx4,family=multinomial,data=sdados)
> deviance(ajuste07)
> AIC(ajuste07)
> pchisq(deviance(ajuste07)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste08<-rrvglm(formula=risco_b~x2+ x3,family=multinomial,data=sdados)
> deviance(ajuste08)
> AIC(ajuste08)
> pchisq(deviance(ajuste08)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste09<-rrvglm(formula=risco_b~x2+ sx4,family=multinomial,data=sdados)
> deviance(ajuste09)
> AIC(ajuste09)
> pchisq(deviance(ajuste09)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste10<-rrvglm(formula=risco_b~x3+ sx4,family=multinomial,data=sdados)
> deviance(ajuste10)
> AIC(ajuste10)
> pchisq(deviance(ajuste10)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste11<-rrvglm(formula=risco_b~sx1+ x2+ x3,family=multinomial,data=sdados)
> deviance(ajuste11)
```

```

> AIC(ajuste11)
> pchisq(deviance(ajuste11)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste12<-rrvglm(formula=risco_b~sx1+ x2+ sx4,family=multinomial,data=sdados)
> deviance(ajuste12)
> AIC(ajuste12)
> pchisq(deviance(ajuste12)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste13<-rrvglm(formula=risco_b~sx1+ x3+ sx4,family=multinomial,data=sdados)
> deviance(ajuste13)
> AIC(ajuste13)
> pchisq(deviance(ajuste13)-deviance(ajuste15),df=4,lower.tail=FALSE)
> ajuste14<-rrvglm(formula=risco_b~x2+ x3+ sx4,family=multinomial,data=sdados)
> deviance(ajuste14)
> AIC(ajuste14)
> pchisq(deviance(ajuste14)-deviance(ajuste15),df=4,lower.tail=FALSE)
> #ESTIMAÇÃO DE PARÂMETROS PARA O MODELO 2 (AJUSTE 11)#
> ajuste11<-rrvglm(formula=risco_b~sx1+ x2+ x3,multinomial,sdados)
> summary(ajuste11)
> Coef(ajuste11)
> coef(ajuste11,matrix=TRUE)
> residuos<-residuals(ajuste11)
> summary(residuos)
> #PARA A HIPERTENSÃO
> normal1<-8.655-3.258*sx1-0.064*mean(x2)-0.028*mean(x3)
> acrescido1<-5.966-1.604*sx1-0.032*mean(x2)-0.014*mean(x3)
> moderado1<-6.319-1.639*sx1-0.032*mean(x2)-0.014*mean(x3)
> alto1<-3.723-0.620*sx1-0.012*mean(x2)-0.005*mean(x3)
> pnormal1<-
exp(normal1)/(exp(normal1)+ exp(acrescido1)+ exp(moderado1)+ exp(alto1)+ 1)
> pacrescido1<-
exp(acrescido1)/(exp(normal1)+ exp(acrescido1)+ exp(moderado1)+ exp(alto1)+ 1)
> moderado1<-
exp(moderado1)/(exp(normal1)+ exp(acrescido1)+ exp(moderado1)+ exp(alto1)+ 1)
> alto1<-exp(alto1)/(exp(normal1)+ exp(acrescido1)+ exp(moderado1)+ exp(alto1)+ 1)
> malto1<-(1)/(exp(normal1)+ exp(acrescido1)+ exp(moderado1)+ exp(alto1)+ 1)
> #PARA A IDADE
> normal2<-8.655-3.258*mean(sx1)-0.064*x2-0.028*mean(x3)
> acrescido2<-5.966-1.604*mean(sx1)-0.032*x2-0.014*mean(x3)
> moderado2<-6.319-1.639*mean(sx1)-0.032*x2-0.014*mean(x3)
> alto2<-3.723-0.620*mean(sx1)-0.012*x2-0.005*mean(x3)
> pnormal2<-
exp(normal2)/(exp(normal2)+ exp(acrescido2)+ exp(moderado2)+ exp(alto2)+ 1)
> pacrescido2<-
exp(acrescido2)/(exp(normal2)+ exp(acrescido2)+ exp(moderado2)+ exp(alto2)+ 1)
> moderado2<-
exp(moderado2)/(exp(normal2)+ exp(acrescido2)+ exp(moderado2)+ exp(alto2)+ 1)
> alto2<-exp(alto2)/(exp(normal2)+ exp(acrescido2)+ exp(moderado2)+ exp(alto2)+ 1)
> malto2<-(1)/(exp(normal2)+ exp(acrescido2)+ exp(moderado2)+ exp(alto2)+ 1)
> #PARA O PESO
> normal3<-8.655-3.258*mean(sx1)-0.064*mean(x2)-0.028*x3
> acrescido3<-5.966-1.604*mean(sx1)-0.032*mean(x2)-0.014*x3
> moderado3<-6.319-1.639*mean(sx1)-0.032*mean(x2)-0.014*x3
> alto3<-3.723-0.620*mean(sx1)-0.012*mean(x2)-0.005*x3

```

```

> pnormal3<-
exp(normal3)/(exp(normal3)+ exp(acrescido3)+ exp(moderado3)+ exp(alto3)+ 1)
> pacrescido3<-
exp(acrescido3)/(exp(normal3)+ exp(acrescido3)+ exp(moderado3)+ exp(alto3)+ 1)
> moderado3<-
exp(moderado3)/(exp(normal3)+ exp(acrescido3)+ exp(moderado3)+ exp(alto3)+ 1)
> alto3<-exp(alto3)/(exp(normal3)+ exp(acrescido3)+ exp(moderado3)+ exp(alto3)+ 1)
> mtalto3<-(1)/(exp(normal3)+ exp(acrescido3)+ exp(moderado3)+ exp(alto3)+ 1)
> #GRAFICOS
> plot(x1,pnormal1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=24)
> plot(x2,pnormal2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=24)
> plot(x1,paumentado1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=3)
> plot(x2,paumentado2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=3)
> plot(x1,palto1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=655)
> plot(x2,palto2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=655)
> plot(x1,pmtalto1, pch=8, xaxp=c(40,140,100), yaxp=c(0,1,10), col=2)
> plot(x2,pmtalto2, pch=8, xaxp=c(10,50,40), yaxp=c(0,1,10), col=2)
VALIDAÇÃO
> validacao<-read.table("C:/Users/Ana/Desktop/validacao2.txt",header=TRUE)
> risco_b<-validacao$risco_b
> id<-validacao$id
> hta<-validacao$hta
> idade<-validacao$idade
> peso<-validacao$peso
> srisco_b<--scale(risco_b)
> shta<--scale(hta)
> v1<-(exp(8.655-3.258*shta-0.064*idade-0.028*peso))
> v2<-(exp(5.966-1.604*shta-0.032*idade-0.014*peso))
> v3<-(exp(6.319-1.639*shta-0.032*idade-0.014*peso))
> v4<-(exp(3.723-0.620*shta-0.012*idade-0.005*peso))
> c(v1/(v1+ v2+ v3+ v4+ 1), v2/(v1+ v2+ v3+ v4+ 1), v3/(v1+ v2+ v3+ v4+ 1),
1/(v1+ v2+ v3+ v4+ 1))
> pnormal<-v1/(v1+ v2+ v3+ v4+ 1)
> pnormal
> pacrescido<-v2/(v1+ v2+ v3+ v4+ 1)
> pacrescido
> pmoderado<-v3/(v1+ v2+ v3+ v4+ 1)
> pmoderado
> palto<-v4/(v1+ v2+ v3+ v4+ 1)
> palto
> pmtalto<-1/(v1+ v2+ v3+ v4+ 1)
> pmtalto
> c(pnormal,pacrescido,pmoderado,palto,pmtalto)

```

Anexo 15: Apresentação em Powerpoint, usada na defesa pública da dissertação (14 de Fevereiro de 2014)

MODELAÇÃO ESTATÍSTICA: Estudo do perfil cardiovascular de adultos assintomáticos, da cidade de Benguela



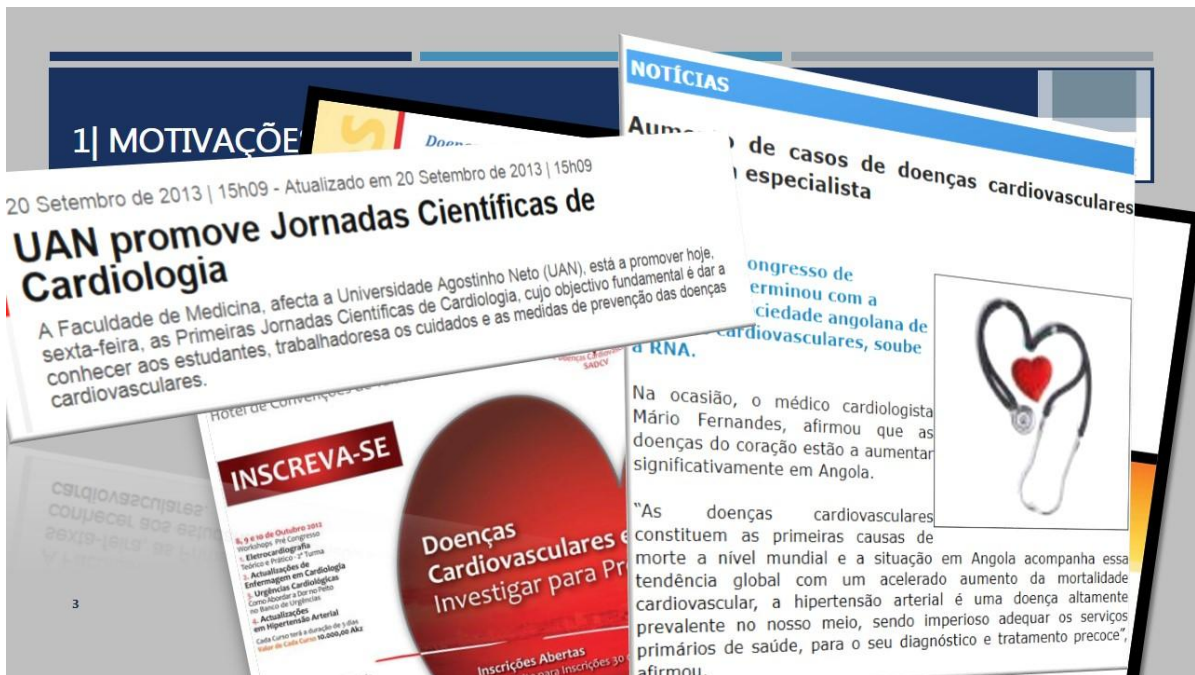
Por Ana Isabel Cunha, sob orientação da Professora
Doutora Teresa Oliveira

Defesa Pública da Dissertação para obtenção do grau
de Mestre em Bioestatística e Biometria

ESTRUTURA DA APRESENTAÇÃO



- 1| MOTIVAÇÕES
- 2| OBJETIVOS GERAIS E ESPECÍFICOS
- 3| ENQUADRAMENTO
- 4| CASO PRÁTICO
- 5| CONSIDERAÇÕES FINAIS
- 6| PERSPETIVAS FUTURAS DE INVESTIGAÇÃO



1| MOTIVAÇÃO



- Tema extremamente relevante e atual, no panorama epidemiológico angolano
- A modelação estatística e, em particular, do modelo estereótipo para análise de risco
- Perspetivas de aprofundamento da investigação desta temática

2| OBJETIVOS GERAIS



- Contribuir para a produção de conhecimento sobre o perfil epidemiológico das doenças cardiovasculares, em Angola
- Chamar a atenção para a relevância do modelo estereótipo em análise de risco e para a importância de se desenvolver *software* de apoio ao mesmo

2| OBJETIVOS ESPECÍFICOS



- Descrever a amostra quanto a alguns indicadores de risco e ao risco cardiovascular
- Comparar o grau de risco cardiovascular entre
- Construir um modelo para análise de risco

- Homens e mulheres
- Indivíduos mais velhos e mais novos
- Indivíduos com sobrepeso e sem sobrepeso
- Indivíduos com obesidade abdominal e sem obesidade abdominal
- Hipertensos e não hipertensos
- Fumadores e não fumadores
- Indivíduos que ingerem álcool e que não ingerem álcool
- Consumidores de gorduras saturadas e não consumidores

3| ENQUADRAMENTO (PARTE I)



1. Técnicas de Inferência Estatística não Paramétrica

- Teste de ajustamento à normalidade de Kolmogorov-Smirnov
- Teste da Binomial

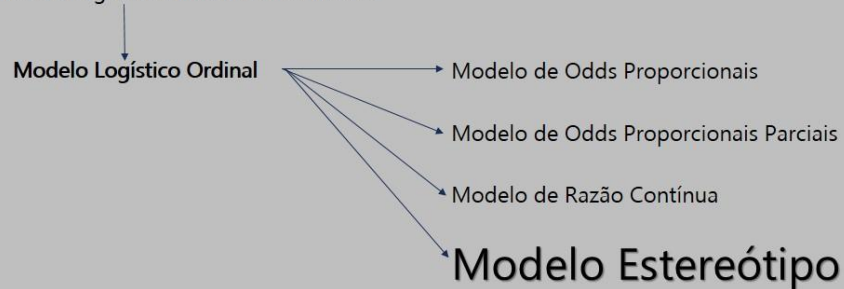
- Teste de independência de Mann-Whitney/ Coeficiente Eta
- Teste de independência de Kolmogorov-Smirnov/ Coeficiente V de Cramer
- Odds Ratio

3| ENQUADRAMENTO (PARTE I)



2. Modelação Estatística

- Modelos de Regressão Linear Generalizados

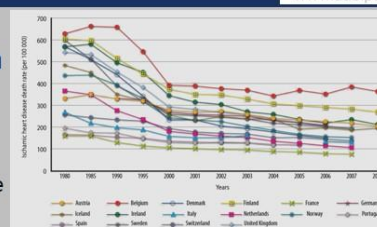


3| ENQUADRAMENTO (PARTE I)

AbERTA
www.uab.pt

3. Doenças Cardiovasculares- Dimensão do problema

- As doenças cardiovasculares lideram o ranking das doenças mais mortíferas do mundo
- Os países desenvolvidos, apostando em estratégias de prevenção e controlo, têm conseguido baixar as suas prevalências

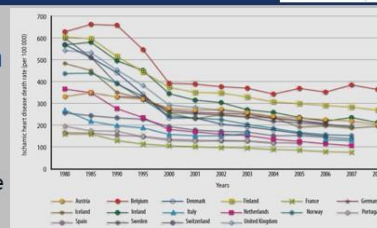


3| ENQUADRAMENTO (PARTE I)

AbERTA
www.uab.pt

3. Doenças Cardiovasculares- Dimensão do problema

- As doenças cardiovasculares lideram o ranking das doenças mais mortíferas do mundo
- Os países desenvolvidos, apostando em estratégias de prevenção e controlo, têm conseguido baixar as suas prevalências



País	Angola	Cabo Verde	Guiné-Bissau	Moçambique	São Tomé e Príncipe	Portugal
% de mortes por doenças cardiovasculares (2008)	27.2	30.4	26.0	23.0	28.0	22.8

3| ENQUADRAMENTO (PARTE I)



3. Doenças Cardiovasculares- Fatores de risco

- O **género** masculino é universalmente apontado como grupo de risco
- A OMS aponta a **idade** como um fator de risco bastante poderoso
- Em 2008, 34% da população mundial estava com **sobrepeso**
- Em 2008, 48% da população mundial apresentava **hipertensão**
- 9% das mortes mundiais devem-se ao **tabagismo**. África tem as prevalências mais baixas (10% de fumadores)

3| ENQUADRAMENTO (PARTE I)



3. Doenças Cardiovasculares- Análise de risco

- Análise do risco cardiovascular tendo em conta o género, o índice de massa corporal e o perímetro abdominal (Risco A)

Índice de massa corporal (kg/m ²)	Perímetro abdominal (cm)	
	Homem: 94-102 Mulher: 80-88	Homem: >102 Mulher: >88
Baixo peso <18.5	Normal	Normal
Peso saudável 18.5-24.9	Normal	Aumentado
Sobrepeso 25.0-29.9	Aumentado	Alto
Obesidade >30	Alto	Muito alto

Fonte: OMS

3| ENQUADRAMENTO (PARTE I)



3. Doenças Cardiovasculares- Análise de risco

- Análise do risco cardiovascular tendo em conta a condição de hipertensão e o número de fatores de risco adicionais (Risco B)

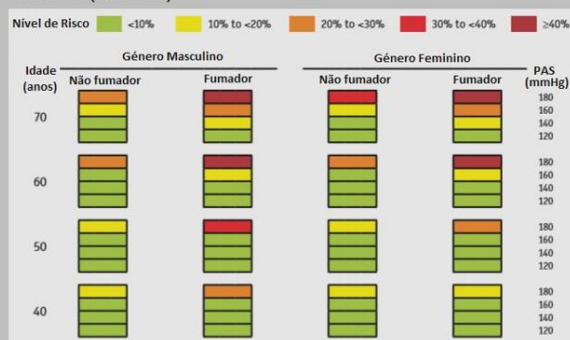
Hipertensão	Número de fatores de risco adicionais		
	0	1-2	>3
Normal	Normal	Baixo acrescido	Moderado acrescido
Normal alta	Normal	Baixo acrescido	Alto acrescido
Hipertensão de grau I	Baixo acrescido	Moderado acrescido	Alto acrescido
Hipertensão de grau II	Moderado acrescido	Moderado acrescido	Alto acrescido
Hipertensão de grau III	Alto acrescido	Muito alto acrescido	Muito alto acrescido

3| ENQUADRAMENTO (PARTE I)



3. Doenças Cardiovasculares- Análise de risco

- Análise do risco cardiovascular tendo em conta o género, os hábitos tabágicos, a idade e a pressão arterial sistólica (Risco C)

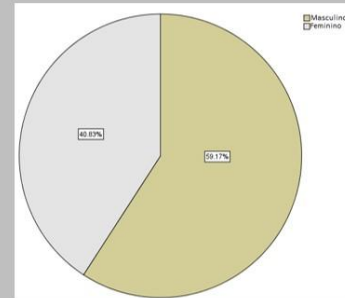


4| CASO PRÁTICO (PARTE II)



Comparação do risco cardiovascular entre homens e mulheres

- As mulheres apresentaram cinco vezes mais obesidade abdominal
- O consumo de tabaco era cinco vezes superior nos homens
- O consumo de álcool era quatro vezes superior nos homens
- A prevalência de hipertensão nos homens era o triplo da das mulheres

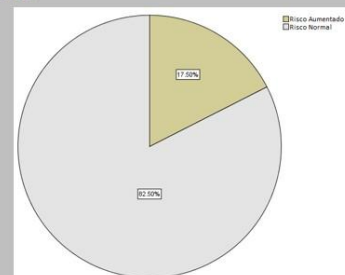


4| CASO PRÁTICO (PARTE II)



Comparação do risco cardiovascular entre indivíduos mais velhos e mais novos

- A prevalência de sobrepeso nos mais velhos era o dobro da dos mais novos
- A prevalência de hipertensão nos mais velhos era quatro vezes superior
- Os hábitos alcoólicos eram três vezes superior nos mais velhos
- Dos indivíduos que ingeriam álcool regularmente, 72.6% eram mais novos

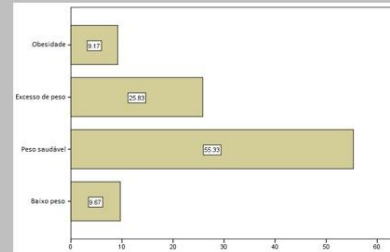


4| CASO PRÁTICO (PARTE II)



Comparação do risco cardiovascular entre indivíduos com sobrepeso e sem sobrepeso

- 35.0% dos benguelenses estava com sobrepeso
- 83.3% afirmaram consumir regularmente gorduras saturadas
- A prevalência de obesidade abdominal no grupo com sobrepeso era onze vezes superior
- A prevalência de hipertensão era duas vezes superior no grupo com sobrepeso
- Os indivíduos com sobrepeso apresentaram maior risco cardiovascular

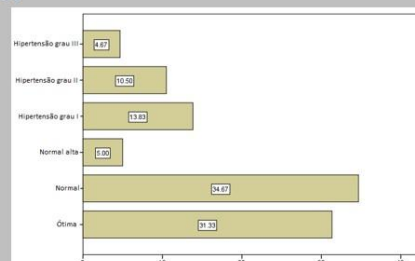


4| CASO PRÁTICO (PARTE II)



Comparação do risco cardiovascular entre hipertensos e normotensos

- 29% de hipertensos
- Como já foi referido anteriormente, os homens, os mais velhos, os indivíduos com sobrepeso e/ou obesidade abdominal e os consumidores de álcool foram associados à hipertensão



4| CASO PRÁTICO (PARTE II)



Comparação do risco cardiovascular entre fumadores e não fumadores

- 7.8% de fumadores
- A probabilidade de um fumador ingerir álcool é cinco vezes maior
- O tabagismo funciona como um fator de proteção em relação ao desenvolvimento de sobrepeso e/ou obesidade abdominal (um terço de probabilidades)

4| CASO PRÁTICO (PARTE II)

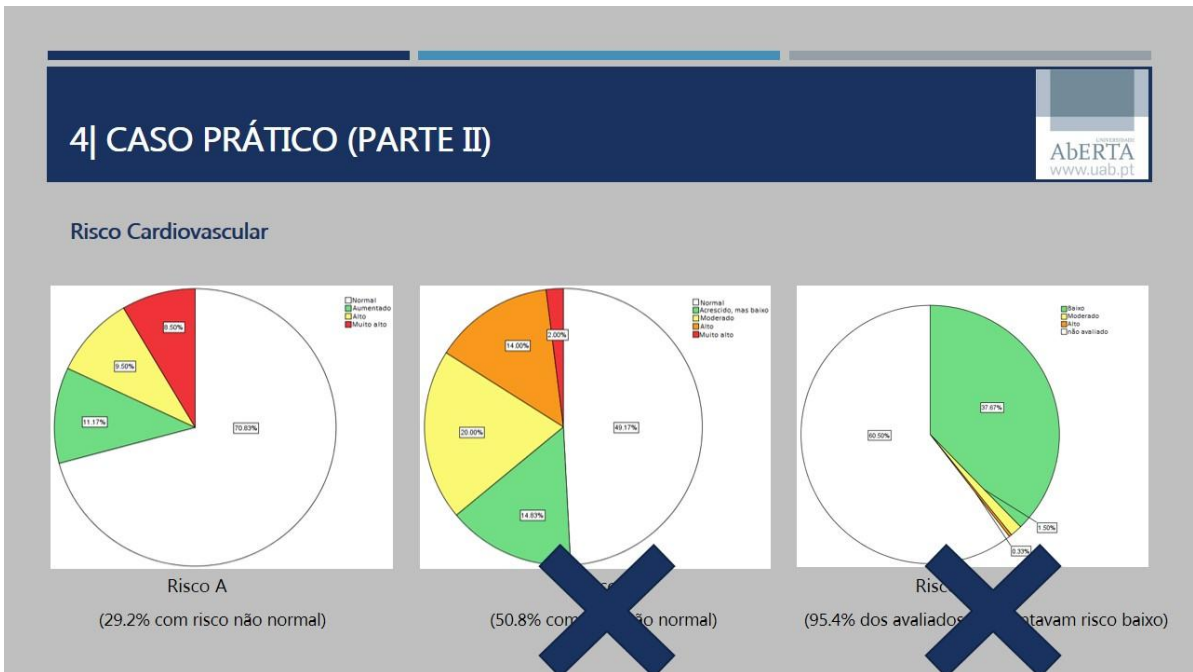
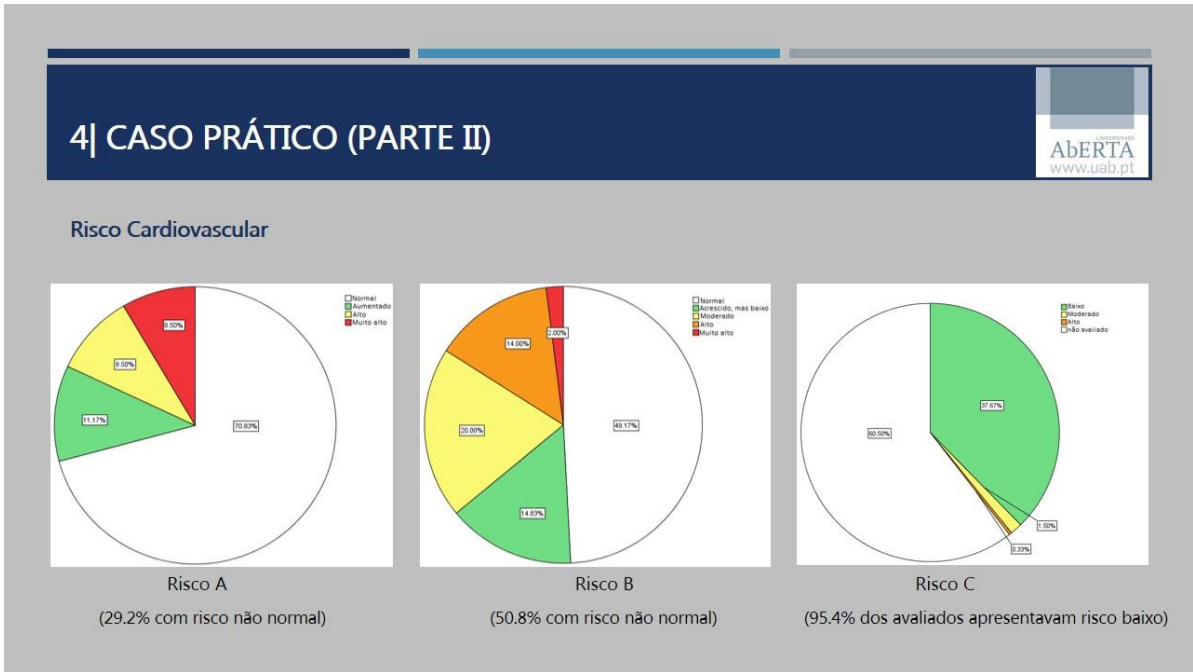


Comparação do risco cardiovascular entre fumadores e não fumadores

- 7.8% de fumadores
- A probabilidade de um fumador ingerir álcool é cinco vezes maior do que a de um não fumador
- O tabagismo funciona como um fator de proteção em relação ao desenvolvimento de sobrepeso e/ou obesidade abdominal (um terço de probabilidades)

Comparação do risco cardiovascular entre indivíduos que ingerem álcool regularmente e que não ingerem

- 32.8% ingerem álcool regularmente
- Os indivíduos que ingeriam álcool consumiam o dobro de gorduras saturadas



4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

- Modelo Estereótipo = escassez de técnicas e de software de apoio
- O software R, graças ao pacote VGAM e à função `rrvglm`, revelou-se como o melhor aliado, porém é ainda muito incompleto e a função referida revelou-se incompatível com outras que são compatíveis com muitos outros pacotes.

4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

- 1- Pré- seleção das possíveis variáveis explicativas, que foram significativamente correlacionadas com o risco
- 2- Seleção das variáveis explicativas mais relevantes para o objetivo do estudo (perímetro abdominal; índice de massa corporal; peso; idade; hipertensão arterial; género)
- 3- Construção dos 63 modelos possíveis, recorrendo à função `rrvglm` e à standardização das variáveis binárias, recorrendo às funções `transform` e `scale`
- 4- Cálculo da estatística do teste da razão de verosimilhança e da sua significância, tendo-se eliminado todos os modelos que apresentaram $p\text{-value} \leq .05$. Para complementar, calculou-se também a estatística BIC, para aplicação do critério de informação Bayesiano (função `stepwise` incompatível com a função `rrvglm`)

4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

5- Escolha do modelo

Dos 63 possíveis modelos, apenas 7 ajustavam os dados de forma significativa

Modelo	Covariáveis	Estatística	p-value	BIC
M1	X1+X2+X4+X5+X6	.763	.382	299.604
M2	X1+X2+X3+X5+X6	2.205	.138	301.046
M3	X1+X2+X3+X4+X6	1.128	.288	299.970
M4	X1+X2+X5+X6	3.088	.214	295.532
M5	X1+X2+X4+X6	1.836	.399	294.280
M6	X1+X2+X3+X6	3.890	.143	296.335
M7	X1+X2+X3	4.691	.196	290.738

x1- perímetro abdominal; x2-índice de massa corporal; x3- peso; x4- idade; x5- hipertensão arterial; x6- género

Parcimoniosamente, escolheu-se o modelo M4

4| CASO PRÁTICO (PARTE II)



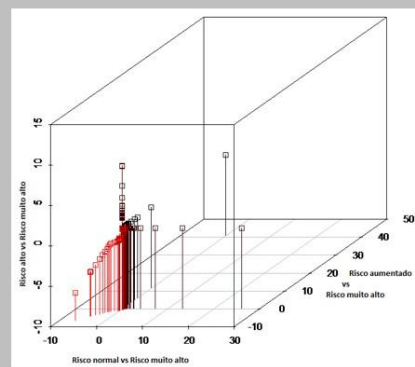
Modelação do Risco Cardiovascular (risco A)

6- Construção de um gráfico com os resíduos para cada um

dos três contrastes, recorrendo à função `residuals` e à função `scatterplot3d`, para analisar a qualidade do ajuste

7- Cálculo das estimativas dos coeficientes e do erro padrão, recorrendo à função `summary`

8- Cálculo dos intervalos de confiança para os coeficientes (a função `confint` não era compatível com a função `rrvglm`)



4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

9- Cálculo do peso dos efeitos das quatro covariáveis do modelo, pela função `coef`, e dos respetivos produtos pelas estimativas dos parâmetros ($w_j\beta$), recorrendo à função `coef(...,matrix=T)`

10- Cálculo dos *Odds Ratio*: $OR = \exp(w_j\beta)$

	Condição de Risco A					
	Normal vs Muito alto		Aumentado vs Muito alto		Alto vs Muito alto	
	$w_j\beta$	OR	$w_j\beta$	OR	$w_j\beta$	OR
Perímetro abdominal	-1.349	.259	-671	.511	-.223	.800
Índice de massa corporal	-2.711	.066	-1.350	.259	-.448	.639
Hipertensão (hipertenso)	-.591	.553	-.294	.745	-.098	.907
Género (masculino)	9.284	10764	4.622	101.697	1.533	4.632

4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

$$\frac{P(\text{Risco normal})}{P(\text{Risco muito alto})} = .553 \Rightarrow P(\text{Risco muito alto}) = 1.808P(\text{Risco normal})$$

	Condição de Risco A					
	Normal vs Muito alto		Aumentado vs Muito alto		Alto vs Muito alto	
	$w_j\beta$	OR	$w_j\beta$	OR	$w_j\beta$	OR
Perímetro abdominal	-1.349	.259	-671	.511	-.223	.800
Índice de massa corporal	-2.711	.066	-1.350	.259	-.448	.639
Hipertensão (hipertenso)	-.591	.553	-.294	.745	-.098	.907
Género (masculino)	9.284	10764	4.622	101.697	1.533	4.632

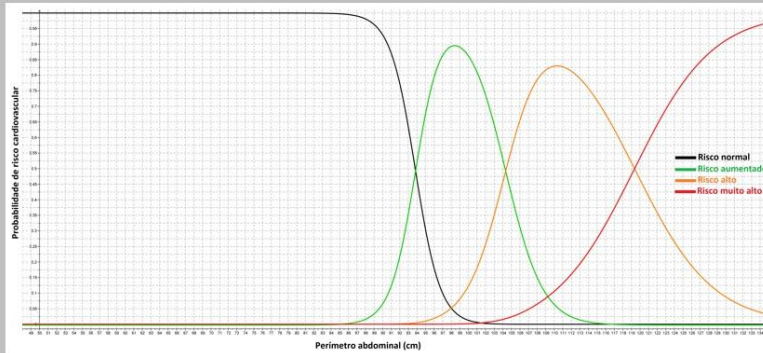
Isto é, um indivíduo hipertenso tem duas vezes mais probabilidades de estar com risco muito alto do que com risco normal!

4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

13- Construção dos gráficos de probabilidades de níveis de risco para cada uma das variáveis contínuas, recorrendo à forma funcional do modelo estereótipo e ao software Geogebra

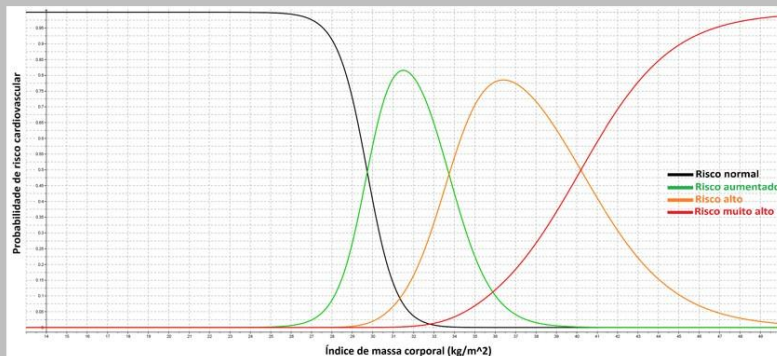


4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

13- Construção dos gráficos de probabilidades de níveis de risco para cada uma das variáveis contínuas, recorrendo à forma funcional do modelo estereótipo e ao software Geogebra

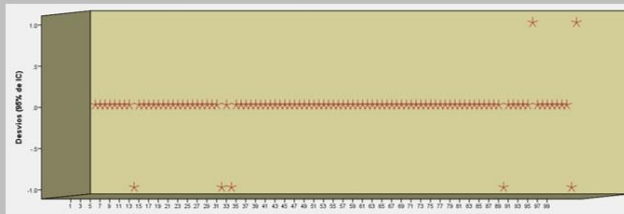


4| CASO PRÁTICO (PARTE II)



Modelação do Risco Cardiovascular (risco A)

11- Validação do modelo- Construção de um gráfico de resíduos



2.0% de riscos sobrevalorizados

93.0% de riscos bem ajustados

5.0% de riscos subvalorizados

Cálculo do coeficiente de correlação de Spearman: $\rho = .945$ → Correlação forte

O MODELO É VÁLIDO E PODE SER USADO PARA ANÁLISE DE RISCO CARDIOVASCULAR!

5| CONSIDERAÇÕES FINAIS



- É importante que em Angola as pessoas sejam informadas acerca da gravidade das doenças cardiovasculares e das consequências de alguns comportamentos de risco
- Recomenda-se, também, a aposta numa estratégia individual de análise de risco, para que cada pessoa conheça o seu risco cardiovascular e seja responsabilizado pela sua própria saúde
- A investigação em doenças cardiovasculares tem que ser prioritária, para que se definam estratégias de prevenção realmente adaptadas à realidade angolana
- O modelo analisado anteriormente é simples e relevante no panorama atual do sistema de saúde angolano, pois é baseado em indicadores antropométricos, sendo, por isso, de baixo custo, no entanto...

6| PERSPETIVAS FUTURAS DE INVESTIGAÇÃO



- É urgente que os estudos se tornem mais complexos (e, inevitavelmente, mais caros!), considerando amostras maiores e indicadores clínicos importantes, como os níveis de glicémia, de colesterol e/ou de triglicéridos
- Sugere-se o início de um estudo de coorte, na província de Benguela, selecionando um conjunto de indivíduos assintomáticos, expostos a alguns fatores de risco, e acompanhando-os para comparar a ocorrência de eventos cardiovasculares
- Do ponto de vista estatístico, é urgente que mais atenções se voltem para a construção de técnicas e de *software* de apoio ao modelo estereótipo, pois é extremamente útil em análise de risco



GRATA PELA VOSSA ATENÇÃO!

♥ Feliz dia do... coração! ♥

anagoncunha@gmail.com