

RAE-CEA -11P22

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

Influência dos dados faltantes em análise de sobrevida de pacientes ambulatoriais
portadores de insuficiência cardíaca

Antonio Carlos Pedroso de Lima

Tamy Harumy Moraes Tsujimoto

Tuany de Paula Castro

São Paulo, novembro de 2011

CENTRO DE ESTATÍSTICA APLICADA – CEA – USP**RELATÓRIO DE ANÁLISE ESTATÍSTICA**

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO: “Influência dos dados faltantes em análise de sobrevida de pacientes ambulatoriais portadores de insuficiência cardíaca.”

PESQUISADORES: Prof. Dr. Alfredo José Mansur e Dr. Humberto F.G de Freitas.

INSTITUIÇÃO: Instituto do Coração - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo.

FINALIDADE DO PROJETO: Publicação de artigo.

RESPONSÁVEIS PELA ANÁLISE: Antonio Carlos Pedroso de Lima

Tamy Harumy Moraes Tsujimoto

Tuany de Paula Castro

REFERÊNCIA DESTE TRABALHO: LIMA, A. C. P., TSUJIMOTO, T., CASTRO, T. P. (2011). **Relatório de Análise Estatística sobre o Projeto “Influência dos dados faltantes em análise de sobrevida de pacientes ambulatoriais portadores de insuficiência cardíaca”**. São Paulo, IME-USP, RAE-CEA 11P22.

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

BUSSAB, W. O. e MORETTIN, P. A. (2010). **Estatística Básica**. 6.ed. São Paulo: Saraiva, 540p.

COLOSIMO, E. A. e GIOLO, S. R. (2006). **Análise de Sobrevivência Aplicada**. 1.ed. São Paulo: Edgard Blücher, 367p.

COX, D.R. (1972) Regression models and life-tables. **Journal of the Royal Statistical Society**, **34**, 187-220.

KLEINBAUM, D. G. e KLEIN, M. (2005). **Survival analysis: a self-learning text**. 2nd edition. Springer, 590p.

NETER, J., WASSERMAN, W., KUTNER, M.H. e LI, W., (1996). **Applied linear statistical models**. 4ª edition. Irwin, 1400p.

PAES, A. T., (2007). **Uso de modelos com fração de cura na análise de dados de sobrevivência com omissão nas covariáveis**. - São Paulo. 121 p. Tese (Doutorado). Instituto de Matemática e Estatística – USP.

VAN BUUREN, S., BOSHUIZEN, H. C., e KNOOK, D. L. (1999) Multiple imputation of missing blood pressure covariates in survival analysis. **Statistics in medicine**, **18**, 681-694.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Microsoft Excel *for Windows*®, versão 2007;

Microsoft Word *for Windows*®, versão 2007;

R *for Windows*®, versão 2.12.2

TÉCNICAS ESTATÍSTICAS UTILIZADAS: (entre parênteses a classificação ISI)

Análise Descritiva Unidimensional (03:010);

Análise de Sobrevida (13:070);

ÁREA DE APLICAÇÃO

Ciências Biomédicas (14:990).

Índice

1. Introdução	8
2. Objetivos	8
3. Descrição do estudo	9
4. Descrição das variáveis	10
4.1. Exame físico	10
4.2. Histórico clínico	10
4.3. Exames laboratoriais	11
4.4. Evolução do paciente	12
4.5. Variável resposta	13
4.6. Criação dos grupos	13
5. Análise descritiva	14
5.1. Caracterização da amostra	15
5.1.1. Análise das variáveis qualitativas	15
5.1.2. Análise das variáveis quantitativas	16
5.2. Análise descritiva do tempo de sobrevida	18
5.2.1. Análise da influência dos valores omissos	18
5.2.2. Análise das variáveis qualitativas	19
5.2.3. Análise das variáveis quantitativas	20
5.2.4. Análise dos grupos	22
5.2.5. Análise dos exames clínicos	23
6. Análise inferencial	24
6.1. Estratégia de análise	24
6.1.1. Seleção de variáveis	25

6.2. Modelo 1	27
6.2.1. Construção do modelo	27
6.2.2. Ajuste do modelo.....	27
6.3. Modelo 2	32
6.3.1. Construção do modelo	32
6.3.1.1. Método de imputação.....	32
6.3.2. Ajuste do modelo.....	34
6.4. Comparação dos modelos	38
6.4.1. Variáveis comuns aos dois modelos	39
6.4.2. Variáveis presentes apenas no Modelo 1	41
6.4.3. Variáveis presentes apenas no Modelo 2.....	43
7. Conclusões	44
APÊNDICE A.....	46

RESUMO

Neste relatório, objetiva-se avaliar o efeito de dados incompletos na análise da importância de características demográficas e clínicas no tempo de sobrevivência de pacientes portadores de insuficiência cardíaca. Um total de 1400 pacientes são considerados, atendidos na Unidade Clínica do Ambulatório Geral e Unidade de Insuficiência Cardíaca do Instituto do Coração (InCor) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC – FMUSP). Cada paciente teve 60 características clínicas mensuradas. No entanto, uma vez que o período de acompanhamento superou 18 anos, muitas características não puderam ser avaliadas para todos os pacientes. Uma análise prévia sugeriu que os dados omissos não ocorrem de forma aleatória, pois, de maneira geral, o tempo de sobrevivência do grupo constituído por pacientes com dados faltantes tem comportamento distinto do tempo de sobrevivência dos demais. Foi feita uma análise de sensibilidade utilizando dois modelos, um considerando valores imputados para as observações faltantes e outro em que indivíduos com observações omissas formam um grupo específico. A comparação desses dois modelos não permitiu estabelecer uma regra geral do comportamento das variáveis com valores omissos, sendo necessário analisar cada fator prognóstico individualmente.

1. Introdução

A insuficiência cardíaca é uma condição sistêmica que afeta não só o sistema cardiovascular, mas também os sistemas esquelético, renal, neuroendócrino e imunológico, de forma que tem se tornado importante a análise dos fatores que determinam o prognóstico dessa doença.

Uma das abordagens mais tradicionais é a análise de sobrevivência [ver, por exemplo, Colossimo e Giolo (2006)], em que se estuda o tempo até a ocorrência de óbito para pacientes com diferentes etiologias cardiovasculares.

Recorrente na área médica, a presença de dados faltantes (ou omissos) torna necessário o desenvolvimento de técnicas estatísticas especializadas em contornar esse problema. Alguns exemplos são a imputação múltipla [Van Buuren et al. (1999)] e adaptações aos chamados modelos com fração de curados [Paes (2007)].

Neste estudo, observa-se a ocorrência acentuada de valores omissos, sendo que maiores percentuais nas variáveis avaliadas estão, em geral, associados a piores prognósticos dos pacientes, evidenciando um comportamento não aleatório, isto é, informativo, dos dados faltantes.

O principal objetivo deste estudo é avaliar o tempo de sobrevida de pacientes com insuficiência cardíaca de acordo com as características coletadas, levando em conta o impacto dos valores omissos na análise.

2. Objetivos

Os objetivos da análise estatística dos dados neste estudo são:

- Avaliar o impacto de valores omissos na análise de sobrevida de pacientes portadores de insuficiência cardíaca;

- Verificar a influência de valores omissos em exames invasivos (cateterismo) e não invasivos (ecocardiograma, eletrocardiograma e exames laboratoriais) no tempo de sobrevida de pacientes portadores de insuficiência cardíaca;
- Analisar o comportamento do tempo de sobrevida dos quatro grupos de pacientes do banco de dados (cateterismo e transplante, cateterismo e não cirurgia, cateterismo e outras cirurgias e não cateterismo).

3. Descrição do estudo

Foram coletados os dados de 1400 pacientes com insuficiência cardíaca de diferentes etiologias, atendidos na Unidade Clínica do Ambulatório Geral e Unidade de Insuficiência Cardíaca do Instituto do Coração (InCor) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC – FMUSP).

Incluíram-se no estudo pacientes com idade inferior a 75 anos, portadores de insuficiência cardíaca por diferentes cardiopatias: hipertensiva, idiopática, chagásica e cardiomiopatia isquêmica.

Foram excluídos pacientes submetidos a determinados procedimentos clínicos ou cirúrgicos, portadores de doença pulmonar obstrutiva crônica, com infarto agudo do miocárdio recente, acometidos de angina de peito instável e pacientes com cardiopatias congênitas do adulto. Além disso, descartaram-se do estudo pacientes portadores de insuficiência cardíaca que apresentavam normalidade nas funções sistólica, renal e hepática, portadores de doença arterial periférica, doença cérebro vascular, diabetes melito tipo I, pacientes com infecção recente, neoplasias ou doença peptídica ulcerosa ativa.

O recrutamento dos pacientes foi feito entre abril de 1991 e julho de 2003 e, o acompanhamento, até dezembro de 2009.

4. Descrição das variáveis

Inicialmente, foram coletadas 102 variáveis, submetidas a uma análise preliminar do percentual de valores omissos. A partir desta análise, com a finalidade de evitar possíveis vieses, foram mantidas no estudo apenas as variáveis com, no máximo, 70% de valores faltantes, excetuando-se aquelas cujo pesquisador demonstrou interesse específico. Dessa forma, restaram no banco de dados um total de 60 variáveis, agrupadas conforme o modo de mensuração, ou seja, exame físico, histórico clínico, exames laboratoriais e evolução do paciente:

4.1. Exame físico

- Idade (em anos);
- Gênero (masculino, feminino);
- Etnia (branco, preto, pardo, amarelo);
- Classe funcional NYHA (I – ausência de sintomas durante atividades habituais; II – sintomas durante atividades habituais; III – sintomas durante atividades físicas; IV – sintomas em repouso);
- Peso (em quilogramas);
- Altura (em centímetros);
- Frequência cardíaca máxima (em bpm);
- Frequência cardíaca média (em bpm);
- Frequência cardíaca mínima (em bpm);
- Frequência cardíaca no Holter (em bpm);
- Pressão arterial diastólica (em mmHg);
- Pressão arterial sistólica (em mmHg).

4.2. Histórico clínico

- Diabetes melito (presente, ausente);
- Hipertensão arterial (presente, ausente);

- Tabagismo (até 20 cigarros por dia, mais de 20 cigarros por dia, não fumante, ex-fumante);
- Etilismo (leve, moderado, intenso, ex-etilista, não etilista);
- Diagnóstico etiológico (chagásica, isquêmica, hipertensiva, alcoólica, idiopática, outros);

4.3. Exames laboratoriais

- **Ecocardiograma**

- Espessura do septo interventricular (em milímetros);
- Espessura da parede posterior (em milímetros);
- Diâmetro diastólico do ventrículo esquerdo (em milímetros);
- Diâmetro sistólico do ventrículo esquerdo (em milímetros);
- Fração de ejeção do ventrículo esquerdo (em porcentagem);
- Diâmetro da aorta (em milímetros);
- Diâmetro átrio esquerdo (em milímetros);
- Diâmetro do ventrículo direito (em milímetros);
- Trombo no ventrículo esquerdo (presente, ausente);
- Insuficiência mitral doppler (presente, ausente);
- Insuficiência tricúspide doppler (presente, ausente);
- Massa do ventrículo esquerdo (em gramas);

- **Leucograma**

- Número absoluto de leucócitos (por mm^3);
- Neutrófilos (em percentual);
- Eosinófilos (em percentual);
- Basófilos (em percentual);
- Linfócitos (em percentual);
- Monócitos (em percentual).

- **Eletrocardiograma**

- Ritmo (normal, fibrilação atrial, taquicardia sinusal, outros);

- Frequência cardíaca (em bpm);
- **Perfil lipídico**
 - Colesterol total (em mg/dl);
 - Triglicérides (em mg/dl);
 - Fração HDL do colesterol (em mg/dl);
 - Fração LDL do colesterol (em mg/dl);
 - Fração VLDL do colesterol (em mg/dl).
- **Exames diversos**
 - Sódio (em mEq/dl);
 - Hemoglobina sérica (em g/dl);
 - Hematócrito sérico (em percentual);
 - Número absoluto de plaquetas (por mm³);
 - Glicemia (em mg/dl);
 - Creatinina (em mg/dl);
 - Ácido úrico (em mg/dl);
 - Consumo máximo de oxigênio (em ml/kg/m²);
 - Indicador de cateterismo (sim, não);
 - Número de extrassístoles ventriculares por hora;
 - Número de extrassístoles ventriculares total;
 - Fração de ejeção do ventrículo esquerdo ao Gated (em percentual);
 - Fração de ejeção do ventrículo direito ao Gated (em percentual).

4.4. Evolução do paciente

- Data de inclusão no estudo;
- Data de último retorno;
- Data do óbito;
- Óbito (sim, não);

- Liberação para transplante (liberado, liberado para outras cirurgias, não liberado).

As variáveis em que aparecem os nomes ventrículo direito, ventrículo esquerdo, átrio direito ou átrio esquerdo podem ser mencionadas no relatório com as abreviações VD, VE, AD ou AE, respectivamente.

4.5. Variável resposta

Como o objetivo da análise é avaliar o tempo de sobrevida dos pacientes com insuficiência cardíaca, optou-se por uma abordagem de Análise de Sobrevivência dos Dados [ver Colossimo e Giolo (2006)].

Para a obtenção da variável resposta, foram utilizadas as variáveis *data de inclusão no estudo*, *data do último retorno*, *data do óbito* e *óbito*, de maneira que o tempo de óbito foi definido pelo tempo (em meses) transcorrido entre a *data de inclusão no estudo* e a *data do óbito* de pacientes que apresentem óbito e, o tempo de censura, pelo tempo (em meses) transcorrido entre a *data de inclusão no estudo* e a *data do último retorno* de pacientes que não apresentaram óbito.

4.6. Criação dos grupos

Os pacientes que apresentavam condições mais críticas de insuficiência cardíaca foram encaminhados ao cateterismo e, então, avaliados quanto à necessidade de serem submetidos ao transplante de coração.

Todos os pacientes do banco de dados foram classificados em quatro grupos:

- Grupo 1: pacientes submetidos ao cateterismo e, posteriormente, ao transplante de coração;
- Grupo 2: pacientes submetidos ao cateterismo e, posteriormente, a outras cirurgias do coração;

- Grupo 3: pacientes submetidos ao cateterismo, mas não encaminhados à cirurgia;
- Grupo 4: pacientes não submetidos ao cateterismo.

Para a criação dos grupos citados acima, foram utilizados o *Indicador de cateterismo* e a *Liberção para transplante*, de forma que o Grupo 1 é formado por pacientes classificados na categoria “sim” em *Indicador de cateterismo* e “liberado” em *Liberção para transplante*; o Grupo 2, por pacientes classificados na categoria “sim” em *Indicador de cateterismo* e “liberados para outras cirurgias” em *Liberção para transplante*; o Grupo 3, por sua vez, por pacientes classificados na categoria “sim” em *Indicador de cateterismo* e “não liberado” em *Liberção para transplante*; por último, o Grupo 4, por pacientes classificados na categoria “não” em *Indicador de cateterismo*.

5. Análise descritiva

Nesta seção serão apresentadas as análises descritivas feitas para caracterizar a amostra e avaliar previamente a influência de cada variável no tempo de sobrevida dos pacientes.

Na etapa de caracterização da amostra, foram utilizados gráficos box plot e de percentis para variáveis quantitativas e gráfico de barras para as variáveis qualitativas [para mais detalhes veja, por exemplo, Bussab e Morettin (2010)].

Já numa segunda etapa, avaliou-se a influência individual de cada uma das variáveis no tempo de sobrevida dos pacientes. Para tal, utilizou-se o gráfico de Kaplan-Meier em conjunto com os testes não paramétricos de *logrank* e Wilcoxon. Mais detalhes podem ser vistos, por exemplo, em Colosimo e Giolo (2006).

5.1. Caracterização da amostra

5.1.1. Análise das variáveis qualitativas

Primeiramente, avaliando as variáveis referentes aos exames físicos, verifica-se pelo gráfico de *classe funcional* (Figura A3) que as características mais frequentes na amostra são as classes II (sintomas durante atividades habituais) e III (sintomas durante atividades físicas). Percebe-se ainda que indivíduos com melhores condições sintomáticas (I – ausência de sintomas durante atividades habituais) têm pouca frequência na amostra (cerca de 1%). Quanto ao *gênero* (Figura A6), nota-se predominância de pacientes do sexo masculino (cerca de 78%) e para *etnia* (Figura A7), dentre os casos completos, há maioria de brancos (cerca de 34%), embora a maior parte da amostra (52%) não possua valor para esta variável.

Avaliando as características de histórico clínico, observa-se para o *tabagismo* (Figura A8) uma parcela expressiva de valores omissos (43%) e, dentre os pacientes com valores completos, a predominância de pacientes não fumantes (25%) e tabagista moderado (20%). Para o *etilismo* (Figura A9), verifica-se que a maior parcela da amostra não possui resposta e, dentre os indivíduos com resposta, há uma predominância de não etilistas. Analisando a ocorrência de *diabetes melito* (Figura A10), percebe-se que a maioria dos pacientes (cerca de 62%) não possui a patologia. Quanto à *hipertensão arterial* (Figura A11), nota-se que o percentual de pacientes classificados como “ausente” é ligeiramente maior que o percentual de “presente”. Em relação ao *diagnóstico etiológico* (Figura A12), pode-se observar que não há valores faltantes e que, de maneira geral, os pacientes se distribuem de maneira homogênea entre as categorias.

Para o exame de ecocardiograma, percebe-se que, enquanto a *insuficiência mitral Doppler* (Figura A14) apresenta parcela expressiva de pacientes na categoria “presente” (cerca de 74%), a *insuficiência tricúspide Doppler* (Figura A15) não possui

diferenças expressivas entre os percentuais das categorias “presente” e “ausente”. Quanto à presença de *trombo no ventrículo esquerdo* (Figura A16), verifica-se que a grande maioria se enquadra na categoria “ausente”.

Por último, em relação ao *ritmo* do eletrocardiograma (Figura A13), verifica-se uma parcela expressiva dos pacientes com valores omissos (cerca de 64%) e, dentre os pacientes com resposta, há predominância da categoria “normal”.

5.1.2. Análise das variáveis quantitativas

Inicialmente, analisando os gráficos das variáveis de exame físico, nota-se que a amostra é composta por pacientes com idade mediana de 45 anos, o mais novo com idade igual a 13 anos e, o mais velho, 71 anos (Figura A18). Com relação à *altura* (Figura A19), observa-se simetria, em que metade dos indivíduos está abaixo de 1,65m. Para o *peso* (Figura A20), há certa assimetria dos dados com alguns casos extremos; o máximo observado igual a 172 kg e, o mínimo, 30 kg.

Avaliando os gráficos de frequências cardíacas do exame Holter (Figuras A43, A45, A46 e A47), nota-se certa simetria, com exceção apenas da *frequência cardíaca no Holter*, com mediana igual a 80 bpm. Observa-se que metade dos pacientes apresenta *frequências cardíacas máxima, média e mínima* acima de 128 bpm, 82 bpm e 55 bpm, respectivamente.

Quanto às variáveis de pressão arterial, observa-se maior assimetria para a *pressão arterial sistólica* (Figura A52), que apresenta valor mediano igual a 110 mmHg, já a *pressão arterial diastólica* possui mediana de 80 mmHg (Figura A51).

Seguindo a análise para o ecocardiograma, verificou-se que as variáveis relativas ao *diâmetro de aorta, átrios e ventrículos* (Figuras A36 a A40) possuem certa simetria e pontos extremos. Em relação à *espessura do septo interventricular* e da *parede posterior* (Figuras A41 e A42), nota-se que as distribuições dos dados são

semelhantes, com pouca variabilidade e mediana igual a 8 mm. Para a *fração de ejeção no ecocardiograma* e a *massa*, ambas do ventrículo esquerdo, percebe-se que a mediana da primeira é de 36% (Figura A48) e, a da segunda, de 190 g (Figura A57).

Avaliando as variáveis do leucograma, pode-se observar que o número mediano de *leucócitos* por milímetro cúbico de sangue é 7800 unidades e, ainda para essa variável, nota-se a presença de alguns pontos extremos e valor máximo de 24700 (Figura A26). Quanto aos percentuais de *linfócitos* e *neutrófilos* (Figuras A27 e A30), verifica-se certa simetria dos dados com presença de valores discrepantes acima do terceiro quartil para os *linfócitos* e abaixo do primeiro quartil para os *neutrófilos*. Pode-se observar também assimetria positiva no percentual de *basófilos*, *eosinófilos* e *monócitos*, com medianas de 0%, 2% e 7%, respectivamente (Figuras A28, A29 e A31).

Para o exame de eletrocardiograma, nota-se que a *frequência cardíaca* (Figura A44) apresenta maior variabilidade quando comparada à *frequência cardíaca do exame Holter*. Além disso, não há pontos extremos e a mediana é de 84 bpm.

De acordo com as variáveis do perfil lipídico, nota-se que o *colesterol total* (Figura A21) e suas frações *HDL* (Figura A22) e *LDL* (Figura A23) apresentam simetria, diferentemente da fração *VLDL* (Figura A24) e *triglicérides* (Figura A25), as quais se mostram assimétricas positivas. Todas possuem alguns pontos extremos acima do terceiro quartil e as medianas são 192 mg/dl para o *colesterol total*, 39 mg/dl para *fração HDL*, 125 mg/dl para *fração LDL*, 21 mg/dl para *fração VLDL* e, por último, 105 mg/dl para *triglicérides*.

Finalmente, estudando as variáveis referentes aos exames diversos, observa-se que *número de plaquetas* (Figura A32), *ácido úrico* (Figura A33), *creatinina* (Figura A34) e *sódio* (Figura A35) apresentam medianas 214000 por mm³, 8 mg/dl, 1,2 mg/dl e 139 mEq/dl respectivamente. Além disso, mostram-se simétricas com presença de alguns pontos extremos. Para a *fração de ejeção dos ventrículos direito e esquerdo ao exame Gated* (Figuras A49 e A50), com medidas iguais a 21% e 24,5% respectivamente, nota-se uma leve assimetria positiva com presença de alguns pontos extremos. A

hemoglobina sérica (Figura A53) e o *hematócrito sérico* (Figura A54) possuem medidas iguais a 14,7 g/dl e 44% respectivamente. O *número total e por hora de extrassístoles ventriculares* (Figuras A55 e A56) se mostram assimétricos positivos com medianas de 1285 e 64. Quanto à *glicemia* (Figura A58) e ao *consumo máximo de oxigênio* (Figura A59), nota-se assimetria positiva acentuada para a primeira e uma leve assimetria para a segunda, além de medianas iguais a 98 mg/dl e 14,7 ml/kg/m².

5.2. Análise descritiva do tempo de sobrevida

Primeiramente, com o gráfico de Kaplan-Meier para a amostra total (Figura A60), verifica-se que o tempo mediano de sobrevivência é igual a 56,7 meses, isto é, metade da amostra possui tempo de sobrevida maior ou igual a 56,7 meses.

5.2.1. Análise da influência dos valores omissos

Como o objetivo principal deste estudo é avaliar o impacto dos valores omissos no tempo de sobrevida de pacientes portadores de insuficiência cardíaca, foi criada uma variável denominada percentual de valores faltantes, calculada de acordo com a quantidade de dados omissos em cada um dos pacientes, de maneira que aqueles que possuem todas as observações completas apresentam o percentual de valores faltantes igual a zero e, no caso oposto, quando todos os dados estão omissos, a variável assume o valor 100%.

Inicialmente, foi avaliada a distribuição dessa variável em um gráfico box plot (Figura A17). Por meio desse gráfico, percebe-se que metade da amostra apresenta 40% ou mais de valores faltantes nas variáveis coletadas; além disso, verificam-se alguns casos extremos, com o máximo de 95% de valores omissos. Observa-se também um mínimo de 5%, ou seja, não há nenhum paciente com todas as observações completas. Por conta disso, não foi possível comparar indivíduos com

todas as observações completas (caso completo) com indivíduos com pelo menos uma observação faltante (caso faltante), de forma que a análise do percentual de valores faltantes foi feita em quatro categorias: até 25% (inclusive), de 25% a 50% (inclusive), de 50% a 75% (inclusive) e maior que 75%.

Analisando o gráfico de barras por categoria do percentual de valores faltantes (Figura A2), verifica-se que a maioria dos pacientes (61%) se encontra na categoria “25% a 50%” e apenas uma pequena parcela (5%) apresenta percentual de valores omissos superior a 75%.

Por fim, foram construídas curvas de Kaplan-Meier para as categorias dessa variável (Figura A62), cujo comportamento sugere piores prognósticos associados a maiores percentuais de valores faltantes. De acordo com os testes não paramétricos de *logrank* e Wilcoxon, há pelo menos uma diferença significativa entre as curvas de Kaplan-Meier.

5.2.2. Análise das variáveis qualitativas

Analisando, inicialmente, as variáveis referentes ao exame físico, verifica-se que o tempo de sobrevida de pacientes do sexo feminino parece ser maior que o tempo de sobrevida dos pacientes do sexo masculino, que possui curva semelhante à curva da base geral (Figura A70). Quanto à *classificação funcional NYHA*, nota-se que, apesar dos testes não paramétricos terem se mostrado significativos, aparentemente não há diferenças no tempo de sobrevida para as diferentes categorias da classificação NYHA (Figura A68), exceto pela categoria “I” que parece destoar das demais, possivelmente pela baixa quantidade de pacientes. Para a variável *etnia* (Figura A71), observa-se ausência de diferenças quanto ao tempo de sobrevida.

Quanto às variáveis de histórico clínico, nota-se que a proporção de sobrevida, entre os pacientes classificados na categoria “presente” da variável *hipertensão arterial* (Figura A75), parece ser maior quando comparada com pacientes classificados na

categoria “ausente”. Para o *diagnóstico etiológico* (Figura A78), verifica-se que os piores prognósticos estão associados a pacientes com cardiopatia chagásica e isquêmica e, os melhores prognósticos, a pacientes com etiologias hipertensiva e idiopática. As demais variáveis de histórico clínico não apresentam diferenças expressivas quanto ao tempo de sobrevida.

Para as variáveis do ecocardiograma, percebe-se que pacientes classificados na categoria “ausente” de *insuficiência mitral Doppler* (Figura A76) e *insuficiência tricúspide Doppler* (Figura A77) parecem estar associados a melhores prognósticos, enquanto pacientes com valores omissos nessas variáveis apresentam curva de sobrevida abaixo das demais, sugerindo pior prognóstico. Já para a presença de *trombo no ventrículo esquerdo* (Figura A80), não parece haver diferenças entre as curvas de sobrevida das categorias.

Quanto ao *ritmo* do eletrocardiograma (Figura A79), não se pode afirmar que há diferenças significativas quanto ao tempo de sobrevivência das diferentes categorias.

Por fim, analisando o *indicador de cateterismo* (Figura A69), observa-se que aparentemente não há diferenças entre os indivíduos submetidos ou não ao cateterismo quanto ao tempo de sobrevida.

5.2.3. Análise das variáveis quantitativas

Para avaliar o prognóstico de pacientes portadores de insuficiência cardíaca com relação às variáveis quantitativas, foram construídas, para cada variável, três curvas de Kaplan-Meier para grupos formados a partir de tercís. Também foi construída a curva correspondente aos casos com valores faltantes.

Primeiramente, nota-se que os piores prognósticos estão associados a pacientes com idade acima de 50 anos (Figura A81) e menores pressões arteriais diastólica e sistólica (Figuras A114 e A115). Quanto ao prognóstico de pacientes com valores

omissos nessas variáveis, observa-se que, para a *altura* (Figura A82) e o *peso* (Figura A83), os casos faltantes apresentam menor tempo de sobrevida, enquanto para a *idade* (Figura A81) e para a *frequência cardíaca no Holter* (Figura A106), a curva de valores omissos se mostra acima das demais curvas, indicando um melhor prognóstico. As demais variáveis dos exames físicos não apresentam diferenças expressivas no tempo de sobrevida.

Em seguida, para as variáveis de ecocardiograma, observa-se que quanto maior o *diâmetro do átrio esquerdo* (Figura A100), pior o prognóstico do paciente, comportamento que se repete para o *diâmetro do ventrículo direito* (Figura A101), *diâmetros diastólico e sistólico do ventrículo esquerdo* (Figuras A102 e A103). Diferentemente dessas variáveis, o melhor prognóstico dos pacientes está associado a maiores valores de *espessura do septo interventricular* (Figura A105) e de *fração de ejeção do ventrículo esquerdo* (Figura A111). Para a *massa do ventrículo esquerdo* (Figura A120), nota-se que a proporção de sobreviventes é maior entre os pacientes com menores valores dessa variável. As demais variáveis não apresentam diferenças significativas quanto ao tempo de sobrevida. Além disso, avaliando os valores omissos, observa-se que, com exceção de *diâmetro do ventrículo direito*, todas as variáveis possuem casos faltantes com menor tempo de sobrevida.

Para o exame de leucograma, verifica-se que a curva de Kaplan-Meier dos pacientes com percentual de *eosinófilos* (Figura A92) entre 0% e 1% está abaixo das demais curvas, evidenciando uma menor probabilidade de sobrevida. Já para *neutrófilos* (Figura A93) e *monócitos* (Figura A94), quanto maior o valor da primeira e quanto menor o valor da segunda, pior o prognóstico do paciente. Estudando os casos omissos, apenas o *número absoluto de leucócitos* (Figura A89) apresenta a curva de dados faltantes com diferença expressiva em relação às demais curvas, associada ao melhor prognóstico.

Realizando a análise para o eletrocardiograma, nota-se que, aparentemente, não há diferença entre as curvas de Kaplan-Meier das categorias de *frequência cardíaca* (Figura A107). Pode-se reparar também que a curva de valores omissos se mostra

acima das demais categorias, porém é preciso cuidado com essa interpretação, uma vez que o percentual de valores omissos nesta variável é expressivo.

Avaliando as variáveis de perfil lipídico, observa-se que para o *colesterol total* (Figura A84) e *fração LDL* (Figura A86) a curva da categoria de menor valor se apresenta abaixo das demais curvas, evidenciando um pior prognóstico nesses casos. Para a *fração HDL* (Figura A85), *fração VLDL* (Figura A87) e *triglicérides* (Figura A88), percebe-se que quanto maior o valor da variável, maior proporção de sobreviventes. Além disso, as curvas de valores omissos se mostram associadas a piores prognósticos em todas as variáveis.

Por fim, analisando os exames diversos, verifica-se que quanto maior o valor de *ácido úrico* (Figura A96), *número total de extrassístoles* (Figura A118) e *número de extrassístoles por hora* (Figura A119), pior o prognóstico do paciente. Diferentemente dessas variáveis, quanto maior o valor de *sódio* (Figura A98), *fração de ejeção do ventrículo esquerdo* (Figura A112) e *fração de ejeção do ventrículo direito* (Figura A113), maior o tempo de sobrevida dos pacientes. Além disso, para a *creatinina* (Figura A97) e o *consumo máximo de oxigênio* (Figura A122), maior proporção de sobreviventes está associada à categoria de menores e maiores valores, respectivamente. Para os valores omissos, a *fração de ejeção do ventrículo esquerdo*, a *glicemia* (Figura A121), o *consumo máximo de oxigênio*, a *hemoglobina sérica* (Figura A116) e o *hematócrito sérico* (Figura A117) apresentam curvas de valores faltantes abaixo das demais curvas, indicando menor proporção de sobreviventes.

5.2.4. Análise dos grupos

Observa-se no gráfico de barras por categoria de grupo (Figura A1) que a maioria dos pacientes (64%) não foi submetida ao cateterismo e apenas 8% foram encaminhados ao cateterismo e, posteriormente, para alguma cirurgia (6% de transplante e 2% de outras cirurgias).

De acordo com o gráfico de Kaplan-Meier (Figura A61), nota-se que, aparentemente, não há diferença entre os grupos quanto ao tempo de sobrevida, contrariando a suspeita de que pacientes submetidos ao cateterismo seguido de transplante apresentassem o pior prognóstico.

5.2.5. Análise dos exames clínicos

A fim de avaliar o impacto de valores omissos nos exames clínicos (invasivos e não invasivos), fez-se necessária a definição do critério para identificação do paciente como caso completo ou caso faltante em cada um dos exames e, para tanto, foram utilizadas as suas datas de realização, de forma que pacientes com data de exame foram considerados casos completos e, aqueles que não possuíam data, considerados casos faltantes.

Comparando os exames não invasivos (ecocardiograma, eletrocardiograma, leucograma e perfil lipídico), percebe-se (Figuras B17, B18, B19 e B20) que os menores percentuais de casos faltantes foram no ecocardiograma (13%) e no leucograma (16%), seguidos pelos exames de perfil lipídico e eletrocardiograma (40% e 73%, respectivamente).

Observa-se nos gráficos de Kaplan-Meier que o caso faltante apresentou pior prognóstico nos exames de ecocardiograma e perfil lipídico (Figuras A63 e A66), diferentemente dos exames eletrocardiograma e leucograma (Figura A64 e A65), em que o caso faltante mostra melhor prognóstico.

Por último, no gráfico de Kaplan-Meier (Figura A67), pode-se observar que, aparentemente, não há diferenças significativas quanto ao tempo de sobrevida entre pacientes submetidos e não submetidos ao cateterismo.

6. Análise inferencial

Como o objetivo do projeto é avaliar a influência de dados omissos no tempo de sobrevida de pacientes portadores de insuficiência cardíaca, serão apresentados nesta seção a estratégia adotada na análise e os modelos e técnicas utilizados para mensurar o impacto dos dados faltantes.

6.1. Estratégia de análise

Uma vez que a análise descritiva mostrou que os dados omissos possuem um padrão informativo (ver gráficos de Kaplan-Meier em que o grupo composto por pacientes com dados faltantes tem comportamento distinto dos demais), decidiu-se aplicar uma técnica de imputação de dados e comparar o modelo resultante com um modelo em que indivíduos com observações omissas formavam um grupo específico.

Foi escolhido o modelo de regressão de Cox [Cox (1972)] devido às poucas suposições nas quais se baseia. Dessa forma, modela-se a taxa de óbito, que pode ser interpretada como o *potencial instantâneo de falha* [Kleinbaum e Klein (2005)], de maneira que indivíduos com maiores valores da taxa de óbito estão associados aos piores prognósticos.

Um primeiro modelo, denominado Modelo 1, avaliou o tempo de sobrevida dos pacientes em função de covariáveis resultantes da categorização em termos de tercís, sendo um outro grupo constituído pelos valores faltantes. O segundo modelo, denotado Modelo 2, considerou os tempos de sobrevida em função de covariáveis cujos valores omissos foram imputados por métodos de regressão e, então, categorizados por tercís.

6.1.1. Seleção de variáveis

Como o método de imputação pode trazer viés à análise quando utilizado em variáveis com percentual elevado de valores faltantes, faz-se necessária a seleção prévia do grupo de variáveis que será utilizado na construção dos modelos.

Neste estudo foram adotados os seguintes critérios de escolha das variáveis a serem incluídas na seleção estatística de cada um dos modelos:

- **Significância univariada no teste de Wald:** para cada variável do estudo, previamente categorizada incluindo a categoria *faltante*, construiu-se um modelo de regressão de Cox, controlando por *altura*, *peso* e *gênero*, e avaliou-se o menor nível descritivo das categorias dessa variável, tornando-se possível candidata se apresentasse nível descritivo menor que 5%.
- **Percentual de 40% de valores faltantes:** dentre as variáveis que apresentaram significância no critério anterior, foram selecionadas aquelas com percentual de valores faltantes menor ou igual a 40%. Dessa maneira, selecionaram-se cerca de 50% das variáveis do estudo, incluindo aquelas de interesse do pesquisador.

Utilizando os critérios adotados, foram selecionadas as seguintes variáveis para a construção dos modelos:

- *Diagnóstico etiológico;*
- *Gênero;*
- *Hemoglobina sérica;*
- *Creatinina;*
- *Sódio;*
- *Classe funcional NYHA;*
- *Idade;*

- *Diâmetro diastólico do ventrículo esquerdo;*
- *Fração de ejeção do ventrículo esquerdo no ecocardiograma;*
- *Diâmetro do átrio esquerdo;*
- *Espessura da parede posterior;*
- *Insuficiência mitral Doppler;*
- *Número absoluto de leucócitos;*
- *Insuficiência tricúspide Doppler;*
- *Fração de ejeção do ventrículo esquerdo ao Gated;*
- *Glicemia;*
- *Altura;*
- *Diâmetro do ventrículo direito;*
- *Peso;*
- *Diâmetro sistólico do ventrículo esquerdo;*
- *Basófilos;*
- *Eosinófilos;*
- *Linfócitos;*
- *Neutrófilos;*
- *Monócitos;*
- *Colesterol total;*
- *Triglicérides.*

Para a construção dos modelos, será utilizado o método *Forward-backward* com nível de significância de entrada igual a 10% e, de saída, 5%. Neste método, inicia-se o modelo com apenas uma variável e, a cada passo, uma nova variável é adicionada e avaliada quanto à sua significância. Na primeira etapa, a variável será mantida no

modelo se o nível descritivo do teste da razão de verossimilhanças for inferior a 10% (nível de significância de entrada). Na segunda etapa, será feito o teste da razão de verossimilhanças a fim de verificar se cada uma das variáveis pode ser descartada, dado que as demais já estão presentes no modelo. Nesta etapa, tomou-se o nível descritivo de saída igual a 5%, isto é, se o nível descritivo do teste for superior a 5%, então a variável é descartada do modelo.

6.2. Modelo 1

Nesta seção são apresentados a construção e o ajuste do Modelo 1, que avalia o tempo de sobrevida dos pacientes de acordo com covariáveis previamente categorizadas.

6.2.1. Construção do modelo

Primeiramente, categorizaram-se as variáveis quantitativas de acordo com os tercís, incluindo-se a categoria *faltante*; essa mesma categoria foi criada para as variáveis qualitativas.

Após a categorização, selecionaram-se as variáveis do modelo de acordo com o método *Forward-backward* e agruparam-se as categorias com comportamentos semelhantes de acordo com o teste da razão de verossimilhanças.

6.2.2. Ajuste do modelo

Após o ajuste do modelo, foram significativas as seguintes variáveis: *diagnóstico etiológico*, *hemoglobina sérica*, *diâmetro diastólico do ventrículo esquerdo*, *leucócitos*, *fração de ejeção do ventrículo esquerdo ao Gated*, *glicemia*, *diâmetro do ventrículo*

direito, peso, diâmetro sistólico do ventrículo esquerdo, basófilos, colesterol total e monócitos, as quais apresentaram os seguintes agrupamentos de categorias:

Tabela 1: Agrupamentos realizados no modelo selecionado.

Variável	Níveis iniciais	Níveis finais
Diagnóstico etiológico	Alcoólica, chagásica, isquêmica, hipertensiva, idiopática e outros	Alcoólica e isquêmica; chagásica; hipertensiva, idiopática e outros
Hemoglobina sérica (g/dl)	Menor que 13.9; 13.9 a 15.4; maior que 15.4; faltantes	Completo; faltantes
Diâmetro diastólico do ventrículo esquerdo (mm)	Menor que 70; 70 a 77; maior que 77; faltantes	Completo; faltantes
Leucócitos (número/mm³)	Menor que 6900; 6900 a 8900; maior que 8900; faltantes	Completo; faltantes
Fração de ejeção do ventrículo esquerdo ao Gated (%)	Menor que 19; 19 a 25; maior que 25; faltantes	Faltantes e menor que 19; 19 a 25; maior que 25
Glicemia (mg/dl)	Menor que 93; 93 a 106; maior que 106; faltantes	Faltantes e menor que 106; maior que 106
Diâmetro do ventrículo direito (mm)	Menor que 22; 22 a 29; maior que 29; faltantes	Menor que 22 e maior que 29; 22 a 29; faltantes
Peso (kg)	Menor que 60; 60 a 73; maior que 73; faltantes	Completo; faltantes
Diâmetro sistólico do ventrículo esquerdo (mm)	Menor que 60; 60 a 67; maior que 67; faltantes	Menor que 60; 60 a 67; faltantes e maior que 67
Basófilos (%)	Igual a 0; 0 a 1; maior que 1; faltantes	Faltantes, igual a 0 e maior que 1; 0 a 1
Colesterol total (mg/dl)	Menor que 174; 174 a 213; maior que 213; faltantes	Menor que 174; Maior que 174; faltantes
Monócitos (%)	Menor que 4; 4 a 8; maior que 8; faltantes	Faltantes e menor que 8; maior que 8

As estimativas desse modelo final são apresentadas na Tabela 2.

Tabela 2: Estimativas do modelo de regressão de Cox selecionado.

Variável	Estimativa	Erro padrão	Valor p
Diagnóstico (chagásica)	0.22	0.11	0.05
Diagnóstico (hipertensiva, idiopática e outros)	-0.44	0.64	< 0.01
Hemoglobina sérica (faltante)	0.68	0.22	< 0.01
Diâmetro diastólico do VE (faltantes)	0.44	0.15	< 0.01
Leucócitos (faltantes)	-1.08	0.15	< 0.01
Fração de ejeção do VE (19% a 25%)	-0.30	0.10	< 0.01
Fração de ejeção do VE (maior que 25%)	-0.63	0.10	< 0.01
Glicemia (maior que 106 mg/dl)	0.27	0.09	< 0.01
Diâmetro do VD (22 mm a 29 mm)	0.20	0.10	0.04
Diâmetro do VD (faltantes)	-0.25	0.12	0.03
Peso (faltantes)	0.68	0.08	< 0.01
Diâmetro sistólico do VE (60 mm a 67 mm)	0.21	0.12	0.08
Diâmetro sistólico do VE (faltantes e maior que 67 mm)	0.46	0.10	< 0.01
Basófilos (0% a 1%)	-0.29	0.10	0.01
Colesterol total (maior que 174 mg/dl)	-0.39	0.11	< 0.01
Colesterol total (faltantes)	0.28	0.11	0.01
Monócitos (maior que 8%)	-0.41	0.12	< 0.01

Pode-se observar que todos os coeficientes do modelo foram significativos pelo teste univariado de Wald.

Na Tabela 3 são apresentadas as estimativas pontuais e intervalares (com coeficiente de confiança de 90%) das razões das taxas de óbito, tomando-se como referência as primeiras categorias dos níveis finais de agrupamento (vide Tabela 1) das variáveis estatisticamente significantes na descrição do tempo de sobrevivência.

Tabela 3: Estimativas das taxas de óbitos (TO).

Variável	TO	Limites do IC (90%)	
		Inferior	Superior
Diagnóstico (chagásica)	1.25	1.04	1.51
Diagnóstico (hipertensiva, idiopática e outros)	0.64	0.55	0.75
Hemoglobina sérica (faltante)	1.97	1.38	2.80
Diâmetro diastólico do VE (faltantes)	1.56	1.22	1.99
Leucócitos (faltantes)	0.34	0.27	0.43
Fração de ejeção do VE (19% a 25%)	0.74	0.63	0.87
Fração de ejeção do VE (maior que 25%)	0.53	0.45	0.63
Glicemia (maior que 106 mg/dl)	1.31	1.12	1.52
Diâmetro do VD (22 mm a 29 mm)	1.22	1.04	1.44
Diâmetro do VD (faltantes)	0.78	0.64	0.94
Peso (faltantes)	1.97	1.72	2.26
Diâmetro sistólico do VE (60 mm a 67 mm)	1.24	1.01	1.51
Diâmetro sistólico do VE (faltantes e maior que 67 mm)	1.59	1.34	1.88
Basófilos (0% a 1%)	0.75	0.63	0.89
Colesterol total (maior que 174 mg/dl)	0.68	0.57	0.81
Colesterol total (faltantes)	1.33	1.11	1.58
Monócitos (maior que 8%)	0.66	0.55	0.80

Por meio da Tabela 3, pode-se observar que a maior taxa de óbito está associada a pacientes com insuficiência cardíaca chagásica seguida pela taxa de óbito de pacientes com insuficiência cardíaca alcoólica ou isquêmica e, por último, de pacientes classificados nas etiologias hipertensiva, idiopática e outros.

Para *hemoglobina sérica*, *diâmetro diastólico do ventrículo esquerdo*, *peso* e *leucócitos*, verifica-se que apenas a categoria faltante se mostrou significativa, isto é, não há diferenças entre os tercís. As três primeiras variáveis apresentaram pior prognóstico associado aos pacientes com valores omissos, com um aumento de 97%, 56% e 97%, respectivamente, na taxa de óbito com relação à referência, ao passo que,

para os *leucócitos*, há um decréscimo de 66% na taxa de óbito dos pacientes com valores omissos quando comparados à categoria referência.

Quanto à *fração de ejeção do ventrículo esquerdo ao Gated*, nota-se que pacientes com valores maiores que 19% possuem menores taxas de óbito, sendo tal taxa, em relação aos pacientes com fração de ejeção menor que 19% ou com valor omissos, 26% menor para a categoria de *19% a 25%* e 47% menor para a categoria *maior que 25%* de fração de ejeção.

Analisando o *diâmetro do ventrículo direito*, observa-se que pacientes com diâmetros de *22 mm a 29 mm* possuem menor tempo de sobrevida, uma vez que a taxa de óbito é 22% maior que a taxa da categoria referência (*até 22 mm e maior que 29 mm*). Já os pacientes com valores omissos apresentam maiores tempos de sobrevida, uma vez que a taxa de óbito é 22% menor em relação à referência.

Observa-se, ainda, que a taxa de óbito de pacientes com *diâmetro sistólico do ventrículo esquerdo* acima de 67 mm ou com valor do diâmetro omissos é 59% maior que a taxa de óbito dos pacientes da categoria referência (*diâmetro menor que 60 mm*) e, para pacientes com diâmetro de *60 mm a 67 mm*, a taxa de óbito é 24% maior em relação à categoria referência. Assim, percebe-se que quanto maior o diâmetro sistólico do ventrículo esquerdo, maior a taxa de óbito.

Para as variáveis do leucograma, verifica-se que as taxas de óbitos dos pacientes com menos de 1% de *basófilos* e mais de 8% de *monócitos* são 25% e 34%, respectivamente, menores que as taxas de óbitos das categorias de referência (*faltantes, igual a zero e maior que 1%*, no caso dos *basófilos*; *faltantes e menor que 8%*, no caso de *monócitos*).

Em relação à *glicemia*, pode-se notar que pacientes com mais de 106 mg/dl apresentam pior prognóstico, uma vez que a taxa de óbito é 31% maior que a taxa de óbito daqueles com menos de 106 mg/dl ou com valor omissos.

Por fim, percebe-se que pacientes sem resposta para *colesterol total* possuem pior prognóstico, com taxa de óbito 31% maior em comparação aos pacientes com menos de 174 mg/dl. Já o melhor prognóstico está associado aos pacientes com nível de colesterol total acima de 174 mg/dl, uma vez que a taxa de óbito para esses é 32% menor em relação à referência.

6.3. Modelo 2

Nesta seção são apresentados a construção e o ajuste do Modelo 2, que avalia o tempo de sobrevida dos pacientes de acordo com covariáveis cujos valores faltantes foram imputados.

6.3.1. Construção do modelo

Primeiramente, os valores omissos das variáveis quantitativas e qualitativas foram imputados por meio de técnicas de modelo de regressão [Bussab e Morettin (2010)]. Após a imputação, as variáveis quantitativas foram recategorizadas de acordo com os níveis definidos no Modelo 1, para possibilitar posteriormente a comparação entre os modelos.

Após a imputação e a categorização, selecionaram-se as variáveis do modelo de acordo com o método *Forward-backward* e agruparam-se as categorias com comportamentos semelhantes de acordo com o teste da razão de verossimilhanças.

6.3.1.1. Método de imputação

Os valores omissos das variáveis qualitativas e quantitativas foram imputados por meio de modelos de regressão, sendo, para as qualitativas, regressão logística e, para as quantitativas, regressão linear múltipla. Em ambos os modelos de regressão,

foram consideradas como variáveis explicativas aquelas que não apresentavam nenhum valor omissos, incluindo, portanto, as seguintes variáveis:

- *Gênero;*
- *Diagnóstico etiológico;*
- *Grupo;*
- *Logaritmo do tempo de sobrevivência;*
- *Óbito;*
- *Idade.*

As variáveis *logaritmo do tempo de sobrevivência* e *óbito* foram utilizadas para considerar o padrão informativo dos dados faltantes.

Variáveis qualitativas

As variáveis qualitativas candidatas são:

- *Insuficiência mitral Doppler;*
- *Insuficiência tricúspide Doppler;*
- *Classe funcional NYHA.*

Devido à sugestão do pesquisador, os níveis da variável *classe funcional NYHA* foram agrupados em duas únicas categorias: I – II e III – IV. Assim, como são todas dicotômicas, utilizou-se o modelo de regressão logística [Neter *et al.* (1996)] para a modelagem da probabilidade de um paciente possuir a patologia, no caso de *insuficiência mitral Doppler* e *insuficiência tricúspide Doppler*, ou pertencer à classe sintomática “I e II”, no caso de *classe funcional*.

Para proceder à imputação, foi avaliada a probabilidade estimada pelo modelo de regressão logística (\hat{p}), imputando o valor omissos de acordo com seguinte critério:

$$\text{Valor imputado} = \begin{cases} 0, & \text{se } \hat{p} < p^* \\ 1, & \text{se } \hat{p} \geq p^* \end{cases}$$

em que p^* é dado pela proporção de pacientes classificados na categoria de interesse dentre os pacientes que apresentavam resposta para a variável, ou seja, a proporção de pacientes com a patologia, no caso de *insuficiência mitral Doppler* e *insuficiência tricúspide Doppler*, ou pertencentes à classe sintomática “I e II”, no caso de *classe funcional*, dentre os pacientes sem dados faltantes para essa variáveis.

Variáveis quantitativas

Para as variáveis quantitativas, a imputação foi feita de acordo com o modelo de regressão linear múltipla, tomando-se como explicativas as variáveis sem valores omissos e, como resposta, os valores completos da variável a ser imputada.

Após o ajuste do modelo para os valores completos da variável a ser imputada, procedeu-se a técnica utilizando os parâmetros estimados em conjunto com os valores das variáveis explicativas.

6.3.2. Ajuste do modelo

No ajuste do modelo, foram significativas as seguintes variáveis: *diagnóstico etiológico*, *idade*, *fração de ejeção do ventrículo esquerdo no ecocardiograma*, *diâmetro do átrio esquerdo*, *insuficiência tricúspide Doppler*, *fração de ejeção do ventrículo esquerdo ao Gated*, *altura*, *peso*, *diâmetro sistólico do ventrículo esquerdo*, *eosinófilos*, *neutrófilos*, *monócitos*, e *colesterol total*, as quais apresentaram os seguintes agrupamentos de categorias:

Tabela 4: Agrupamentos realizados no modelo selecionado.

Variável	Níveis iniciais	Níveis finais
Diagnóstico etiológico	Alcoólica, chagásica, isquêmica, hipertensiva, idiopática e outros	Alcoólica e isquêmica; chagásica; hipertensiva, idiopática e outros
Idade (anos)	Menor que 40; 40 a 50; maior que 50	Menor que 40; maior que 40
Diâmetro do átrio esquerdo (mm)	Menor que 43; 43 a 50; maior que 50	Menor que 43; maior que 43
Insuficiência tricúspide Doppler	Presente; ausente	Presente; ausente
Fração de ejeção do ventrículo esquerdo ao Gated (%)	Menor que 19; 19 a 25; maior que 25	Menor que 25; maior que 25
Altura (cm)	Menor que 161; 161 a 169; maior que 169	Menor que 161 e maior que 169; 161 a 169
Peso (kg)	Menor que 60; 60 a 73; maior que 73	Menor que 60; 60 a 73; maior que 73
Diâmetro sistólico do ventrículo esquerdo (mm)	Menor que 60; 60 a 67; maior que 67	Menor que 60; 60 a 67; maior que 67
Eosinófilos (%)	Menor que 1; 1 a 4; maior que 4	Menor que 4; maior que 4
Neutrófilos (%)	Menor que 59; 59 a 68; maior que 68	Menor que 59; maior que 59
Monócitos (%)	Menor que 4; 4 a 8; maior que 8	Menor que 4; 4 a 8; maior que 8
Colesterol total (mg/dl)	Menor que 174; 174 a 213; maior que 213	Menor que 174; 174 a 213; maior que 213

As estimativas desse modelo final são apresentadas na Tabela 5.

Tabela 5: Estimativas do modelo de regressão de Cox selecionado.

Variável	Estimativa	Erro padrão	Valor p
Diagnóstico (chagásica)	0.23	0.12	0.06
Diagnóstico (hipertensiva, idiopática e outros)	-0.41	0.10	< 0.01
Idade (maior que 40 anos)	0.24	0.08	< 0.01
Fração de ejeção do VE (maior que 25%)	-0.53	0.10	< 0.01
Diâmetro do AE (maior que 43 mm)	0.35	0.09	< 0.01
Insuficiência tricúspide Doppler (ausente)	-0.22	0.08	< 0.01
Peso (60 kg a 73 kg)	0.20	0.10	0.06
Peso (maior que 73 kg)	-0.27	0.12	0.01
Diâmetro sistólico do VE (60 mm a 67 mm)	0.50	0.10	< 0.01
Diâmetro sistólico do VE (maior que 67 mm)	0.72	0.11	< 0.01
Eosinófilos (maior que 4%)	-0.76	0.10	< 0.01
Neutrófilos (maior que 59%)	0.22	0.10	0.04
Monócitos (4% a 8%)	-0.21	0.09	0.05
Monócitos (maior que 8%)	-0.89	0.13	< 0.01
Colesterol total (174 mg/dl a 213 mg/dl)	-0.36	0.09	< 0.01
Colesterol total (maior que 213 mg/dl)	-0.65	0.12	< 0.01
Altura (161 cm a 169 cm)	0.32	0.08	< 0.01

Pode-se observar que todos os coeficientes do modelo foram significativos pelo teste univariado de Wald.

Na Tabela 6 são apresentadas as estimativas pontuais e intervalares (com coeficiente de confiança de 90%) das razões das taxas de óbito, tomando-se como referência as primeiras categorias dos níveis finais de agrupamento (vide Tabela 4) das variáveis estatisticamente significantes na descrição do tempo de sobrevivência.

Tabela 6: Estimativas das taxas de óbitos (TO).

Variável	TO	Limites do IC (90%)	
		Inferior	Superior
Diagnóstico (chagásica)	1.25	1.03	1.51
Diagnóstico (hipertensiva, idiopática e outros)	0.65	0.56	0.77
Idade (maior que 40 anos)	1.27	1.11	1.46

Fração de ejeção do VE (maior que 25%)	0.57	0.49	0.67
Diâmetro do AE (maior que 43 mm)	1.43	1.23	1.67
Insuficiência tricúspide Doppler (ausente)	0.79	0.68	0.90
Peso (60 kg a 73 kg)	1.21	1.03	1.43
Peso (maior que 73 kg)	0.73	0.60	0.89
Diâmetro sistólico do VE (60 mm a 67 mm)	1.70	1.44	2.00
Diâmetro sistólico do VE (maior que 67 mm)	2.40	1.97	2.92
Eosinófilos (maior que 4%)	0.46	0.39	0.55
Neutrófilos (maior que 59%)	1.23	1.04	1.45
Monócitos (4% a 8%)	0.83	0.72	0.97
Monócitos (maior que 8%)	0.42	0.34	0.52
Colesterol total (174 mg/dl a 213 mg/dl)	0.67	0.58	0.78
Colesterol total (maior que 213 mg/dl)	0.50	0.41	0.61
Altura (161 cm a 169 cm)	1.36	1.19	1.56

De acordo com a Tabela 6, verifica-se que, assim como no Modelo 1, os pacientes com etiologia chagásica possuem pior prognóstico, seguido pelos pacientes com etiologias alcoólicas e isquêmicas, e, por fim, pelos pacientes com etiologias hipertensiva, idiopática e outros.

Avaliando a variável *fração de ejeção do ventrículo esquerdo ao Gated*, nota-se que, assim como no Modelo 1, os melhores prognósticos estão associados a pacientes com maiores valores de fração de ejeção, uma vez que a taxa de óbito de pacientes classificados na categoria *maior que 25%* é 43% menor que a taxa de óbito dos pacientes classificados na categoria referência.

Quanto ao *diâmetro do átrio esquerdo* e ao *diâmetro sistólico do ventrículo esquerdo*, verifica-se que as maiores taxas de óbito estão presentes em pacientes com maiores diâmetros do átrio esquerdo e diâmetros sistólicos do ventrículo esquerdo. Para a primeira variável, a taxa de óbito daqueles com mais de 43 mm é 43% maior que a taxa da referência (*menor que 43 mm*). Já para a segunda variável, a taxa de óbito de pacientes classificados na categoria *maior que 67 mm* é mais que o dobro da taxa de óbito de pacientes classificados na primeira categoria.

Para a variável *idade*, observa-se que a taxa de óbito de pacientes mais velhos (idade superior a 40 anos) é 27% maior que a taxa de óbito de pacientes mais novos (idade inferior a 40 anos). Na variável *peso*, nota-se que pacientes com 73 kg ou mais possuem maior probabilidade de sobrevivência que os demais pacientes, uma vez que a taxa de óbito é 27% menor que a característica referência (*menor que 60 kg*). Quanto à *altura*, verifica-se que pacientes com altura entre 161 cm e 169 cm possuem menor probabilidade de sobrevivência.

Analisando as variáveis do leucograma, percebe-se que os melhores prognósticos estão associados a pacientes com percentual de *eosinófilos* maior que 4%; percentual de *neutrófilos* menor que 59% e percentual de *monócitos* maior que 8%.

De acordo com a variável *insuficiência tricúspide Doppler*, nota-se que a taxa de óbito de pacientes sem insuficiência tricúspide Doppler é 21% menor que a taxa de óbito de pacientes que possuem essa patologia.

Por fim, observa-se que quanto maior a quantidade de *colesterol total*, melhor o prognóstico do paciente, uma vez que há uma redução de 33% e 50% da taxa de óbito para as categorias *174 mg/dl a 213 mg/dl* e *maior que 213 mg/dl* com relação à característica referência, respectivamente.

6.4. Comparação dos modelos

Com os ajustes dos modelos 1 e 2, verificam-se variáveis significativas em ambos ou significativas em apenas um deles. A relação de variáveis significativas nos modelos segue na Tabela 7.

Tabela 7: Relação de variáveis significativas nos modelos 1 e 2.

Modelos 1 e 2	Modelo 1	Modelo 2
Diagnóstico etiológico	Hemoglobina sérica	Idade
Fração de ejeção do VE ao Gated	Diâmetro diastólico do VE	Diâmetro do AE
Peso	Diâmetro do VD	Insuficiência tricúspide Doppler

Diâmetro sistólico do VE
 Colesterol total
 Monócitos

Leucócitos
 Basófilos
 Glicemia

Altura
 Eosinófilos
 Neutrófilos

6.4.1. Variáveis comuns aos dois modelos

Inicialmente, os modelos serão comparados quanto às estimativas e aos erros padrões das variáveis que apresentaram significância em ambos (Tabela 8).

Tabela 8: Estimativas pontual e intervalar das taxas de óbito das variáveis comuns aos modelos 1 e 2.

Variável	Modelo	TO	Limites do IC (90%)	
			Inferior	Superior
Diagnóstico (chagásica)	1	1.25	1.04	1.51
	2	1.25	1.03	1.51
Diagnóstico (hipertensiva, idiopática e outros)	1	0.64	0.55	0.75
	2	0.65	0.56	0.77
Fração de ejeção do VE (19% a 25%)	1	0.74	0.63	0.87
Fração de ejeção do VE (maior que 25%)	1	0.53	0.45	0.63
	2	0.57	0.49	0.67
Peso (faltantes)	1	1.97	1.72	2.26
Peso (60 kg a 73 kg)	2	1.21	1.03	1.43
Peso (maior que 73 kg)	2	0.73	0.60	0.89
Diâmetro sistólico do VE (60 mm a 67 mm)	1	1.24	1.01	1.51
	2	1.70	1.44	2.00
Diâmetro sistólico do VE (faltantes e maior que 67 mm)	1	1.59	1.34	1.88
Diâmetro sistólico do VE (maior que 67 mm)	2	2.40	1.97	2.92
Colesterol total (maior que 174 mg/dl)	1	0.68	0.57	0.81
Colesterol total (faltantes)	1	1.33	1.11	1.58
Colesterol total (174 mg/dl a 213 mg/dl)	2	0.67	0.58	0.78
Colesterol total (maior que 213 mg/dl)	2	0.50	0.41	0.61
Monócitos (4% a 8%)	2	1.33	1.11	1.58
Monócitos (maior que 8%)	1	0.66	0.55	0.80
	2	0.42	0.34	0.52

Por meio da Tabela 8 verifica-se que as variáveis *fração de ejeção do ventrículo esquerdo ao Gated*, *diâmetro sistólico do ventrículo esquerdo* e *monócitos* apresentam categorias comuns aos dois modelos, possibilitando a comparação das estimativas pontuais e intervalares. Para a *fração de ejeção do ventrículo esquerdo ao Gated (maior que 25%)*, nota-se que, aparentemente, não houve mudanças expressivas nas estimativas. Já para o *diâmetro sistólico do ventrículo esquerdo (60mm a 67mm)* observa-se aumento nas estimativas pontuais e intervalares da taxa de óbito e um ligeiro aumento na amplitude do intervalo de confiança. Por último, avaliando a variável *monócitos (maior que 8%)*, observa-se diminuição das estimativas, assim como da amplitude do intervalo de confiança.

Pode-se verificar também que algumas categorias deixam de ser significativas ou passam a ser significativas de um modelo para o outro. Esse é o caso de todas as variáveis listadas na Tabela 8, exceto apenas o *diagnóstico etiológico*, uma vez que esta variável se mostrou completa.

Avaliando a *fração de ejeção do ventrículo esquerdo ao Gated*, nota-se que a categoria 19% a 25% foi significativa apenas no Modelo 1, isto é, após a imputação, esta categoria passou a ter o mesmo comportamento que a categoria referência (menor que 19%). Isso pode ser evidenciado também pelos gráficos de Kaplan-Meier da fração de ejeção do ventrículo esquerdo ao Gated antes e depois da imputação (Figura A112 e A123).

Quanto ao *peso*, no Modelo 1, apenas os pacientes com dados omissos apresentaram taxa de óbito diferente dos demais, o que pode ser observado pelas curvas de Kaplan-Meier dos tercís aparentemente confundidas na Figura A83. Após a imputação, passaram a ser significativas as categorias *60 kg a 73 kg* e *maior que 73 kg*, que também pode ser evidenciado pela Figura A137, em que, aparentemente, as curvas de Kaplan-Meier do segundo e terceiro tercil passaram a ficar acima e abaixo da curva de referência, respectivamente.

De acordo com o *diâmetro sistólico do ventrículo esquerdo* verifica-se que em ambos os modelos a categoria *maior que 67%* se mostra significativa, entretanto, no Modelo 1, o comportamento dessa categoria foi considerado igual ao comportamento dos tempos de sobrevida de pacientes com valores omissos. Esse resultado pode ser observado nas Figuras A103 e A126, sendo que na primeira, a curva de valores faltantes está mais próxima da curva da categoria *maior que 67%* e, na segunda, houve um decaimento das curvas de Kaplan-Meier do segundo e terceiro tercil.

Para o *colesterol total*, nota-se na Figura A84 que as curvas de Kaplan-Meier dos pacientes classificados nas categorias *174 mg/dl a 213 mg/dl* e *maior que 213 mg/dl* apresentam comportamentos semelhantes, resultado este obtido no ajuste do Modelo 1, que apresentou também significância da categoria *faltantes*. Já na Figura A135, observa-se que, aparentemente, quanto maior o colesterol total, melhor o prognóstico do paciente, que resultou na significância das categorias *174 mg/dl a 213 mg/dl* e *maior que 213 mg/dl* no Modelo 2.

Por último, percebe-se que a categoria *4% a 8% de monócitos*, que apresentou o mesmo comportamento da categoria *menor que 4%* e *faltantes* no Modelo 1, passou a ser significativa no Modelo 2, resultados que podem ser verificados pelas Figuras A94 e A129.

6.4.2. Variáveis presentes apenas no Modelo 1

Nesta etapa da análise são avaliados os gráficos de Kaplan-Meier das variáveis presentes apenas no Modelo 1, antes e depois da imputação dos dados, a fim de compreender, descritivamente, o motivo pelo qual a variável deixou de ser significativa no Modelo 2.

Analisando inicialmente a variável *hemoglobina sérica*, pode-se verificar que apenas a categoria *faltante* foi significativa no Modelo 1, resultado que pode ser observado nas Figuras A116 e A128 em que as curvas de Kaplan-Meier dos tercís não

parecem diferir muito, havendo apenas uma ligeira diferença da curva de valores faltantes com relação às demais no primeiro gráfico.

Quanto ao *diâmetro diastólico do ventrículo esquerdo* nota-se, pelas Figuras A102 e A125, que aparentemente as curvas de Kaplan-Meier do segundo e terceiro tercil sofreram um decaimento após a imputação dos dados, não havendo, portanto, um motivo aparente para a não significância dessa variável no Modelo 2.

Para o *leucócito* pode-se observar, pelo Modelo 1, que apenas a categoria *faltante* foi significativa, resultado que foi visto anteriormente na Figura A89, em que a curva de valores faltantes se mostra acima das demais curvas. A partir da imputação dos dados, nota-se que a curva do segundo tercil sofre um leve deslocamento para cima (Figura A130), entretanto, essa diferença entre as curvas não é detectada no Modelo 2.

De acordo com a variável *basófilos*, percebe-se que apenas a curva do segundo tercil parece destoar das demais na Figura A91, comportamento este que foi confirmado no Modelo 1, em que apenas a categoria *0% a 1%* se mostrou significativa. Com a imputação dos dados, essa curva parece ter sofrido um decaimento, fazendo com que as curvas dos tercils passassem a ser mais próximas e, possivelmente, dificultando a detecção da significância dessa variável no Modelo 2.

Com relação à *glicemia*, pode-se verificar que apenas a categoria *maior que 106 mg/dl* se mostrou significativa no Modelo 1, apesar de todas as curvas terem se mostrado próximas na Figura A121. Após a imputação dos dados (Figura A134) nota-se que, aparentemente, as curvas não diferem, resultado confirmado pela ausência de significância desta variável no Modelo 2.

Por fim, avaliando a variável *diâmetro do ventrículo direito*, nota-se pelas Figuras A101 e A127 que a curva do segundo tercil desloca-se para cima após a imputação dos dados e não há um motivo aparente para a não significância dessa variável no segundo modelo.

6.4.3. Variáveis presentes apenas no Modelo 2

Analisando, primeiramente, a *idade*, sabe-se que essa variável não apresenta dados omissos, de maneira que não há um motivo aparente para justificar o fato de ela ser significativa apenas no segundo modelo. Uma possível explicação para isso pode ser o ganho de significância da *idade* na presença de alguma outra variável do Modelo 2.

Avaliando o *diâmetro do átrio esquerdo*, pode-se perceber pelas Figuras A100 e A124 que houve um decaimento da curva de Kaplan-Meier do segundo tercil, que ficou mais próxima da curva do terceiro tercil, após a imputação dos dados. Além disso, não há um motivo evidente da não significância dessa variável no Modelo 1.

De acordo com a variável *insuficiência tricúspide Doppler* verifica-se, por meio das Figuras A77 e A138, que após a imputação dos dados, houve um decaimento da curva da categoria *presente*, fazendo com que a diferença entre as curvas de Kaplan-Meier ficasse mais acentuada, que pode ser o motivo pelo qual essa variável se mostrou significativa no Modelo 2.

Para a *altura*, nota-se que as curvas de Kaplan-Meier dos tercís parecem próximas na Figura A82. Após a imputação dos valores, percebe-se que a curva da categoria *161 cm a 169 cm* parece ter se deslocado para baixo, destoando-se das demais, resultado este que se confirmou com a significância dessa categoria no Modelo 2.

Quanto aos *eosinófilos*, percebe-se que, na Figura A92, as curvas dos tercís estão próximas, justificando a não significância no Modelo 1. Após a imputação (Figura A132), pode-se notar que houve um decaimento das curvas de Kaplan-Meier das duas primeiras categorias, as quais ficaram com o comportamento semelhante, e um deslocamento para cima da terceira categoria, confirmando os resultados do segundo modelo.

Finalmente, analisando os *neutrófilos*, verifica-se que, com a imputação dos dados, houve um deslocamento para cima das curvas das duas primeiras categorias (Figuras A93 e A133), acentuando a diferença entre as curvas de Kaplan-Meier e, assim, possivelmente tornando significativa a variável *neutrófilos* no Modelo 2.

7. Conclusões

De acordo com a análise descritiva, concluiu-se, de maneira geral, que o percentual de valores faltantes parece estar associado ao tempo de sobrevida dos pacientes estudados, em que maiores percentuais apresentam os piores prognósticos. Esse mesmo resultado foi confirmado na análise inferencial por meio do Modelo 1, em que a categoria *faltante*, quando significativa, se mostrou ligada a maiores taxas de óbito, com exceção da variável *leucócitos*, que apresentou menor taxa de óbito para pacientes com valores omissos.

Após a imputação dos dados e a construção do Modelo 2, pôde-se observar que metade das variáveis que se mostraram significativas no Modelo 1 também se mostraram significativas no Modelo 2, possibilitando a comparação das estimativas pontuais e intervalares dos parâmetros das categorias em comum. Por meio dessa comparação, verificaram-se mudanças nas estimativas pontuais e intervalares, assim como na amplitude dos intervalos de confiança, como é o caso das variáveis *diâmetro sistólico do ventrículo esquerdo* e *monócitos*. Houve ainda o caso em que as estimativas não sofreram alterações expressivas (*fração de ejeção do ventrículo esquerdo ao Gated*).

Além das comparações das estimativas, também puderam ser verificados, por meios descritivos, os motivos pelos quais algumas categorias ou variáveis perderam ou ganharam significância de um modelo para o outro, como por exemplo, a *hemoglobina sérica* e os *leucócitos* que, aparentemente, se mostraram significativos apenas no Modelo 1 por causa da categoria *faltante*, a qual apresentou comportamento distinto

das demais categorias, que puderam ser consideradas iguais quanto à contribuição na taxa de óbito.

Finalmente, é importante ressaltar que não houve, dentre os dois modelos avaliados nesta análise, um que possa ser considerado mais adequado ou superior, cabendo ao pesquisador a escolha daquele que traga as interpretações mais adequadas no contexto médico.

APÊNDICE A

Figuras

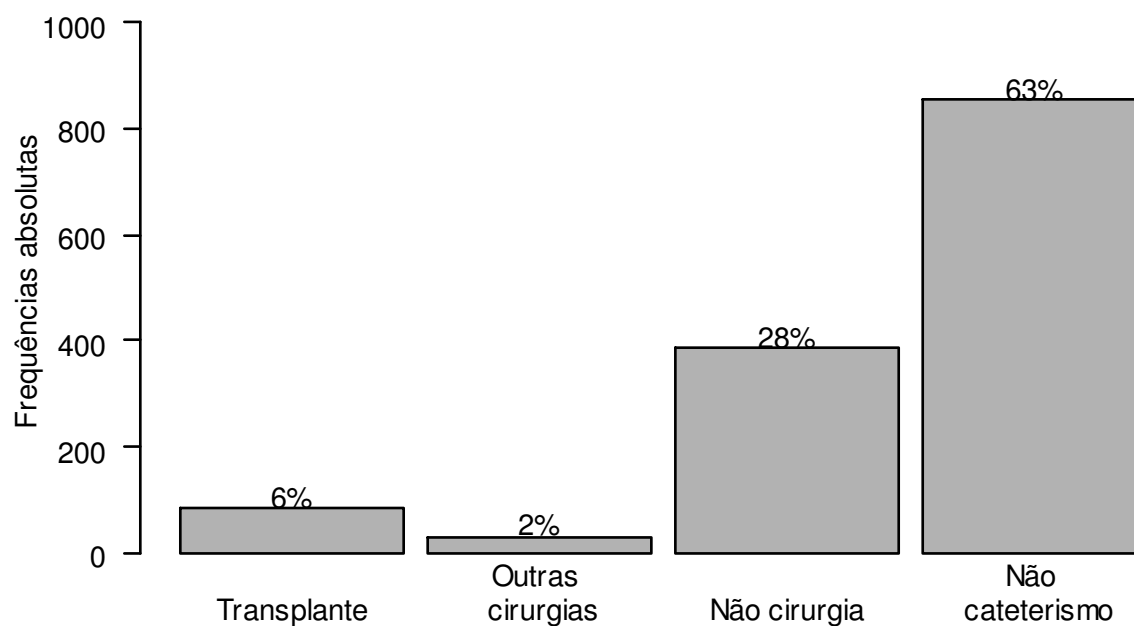
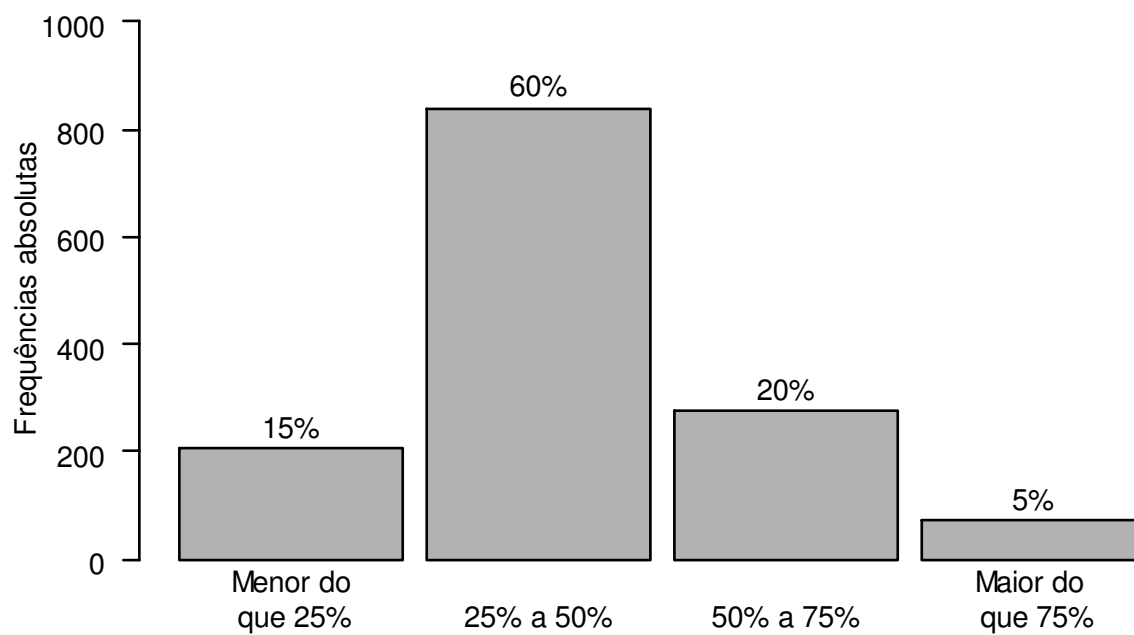
Figura A1: Gráfico de barras por categoria de grupo.**Figura A2:** Gráfico de barras por categoria do percentual de valores faltantes.

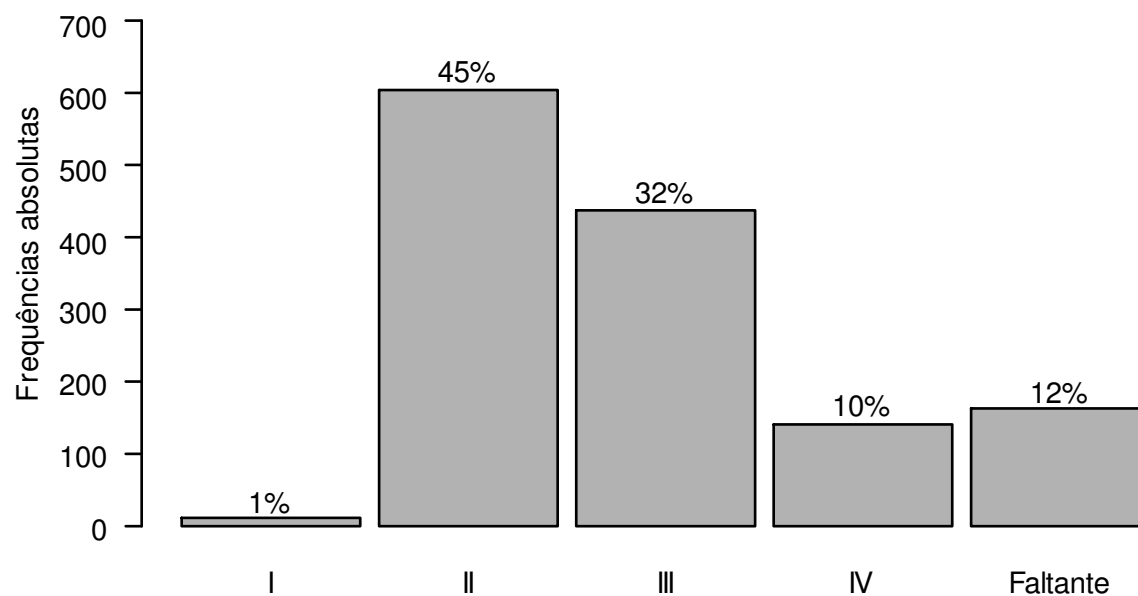
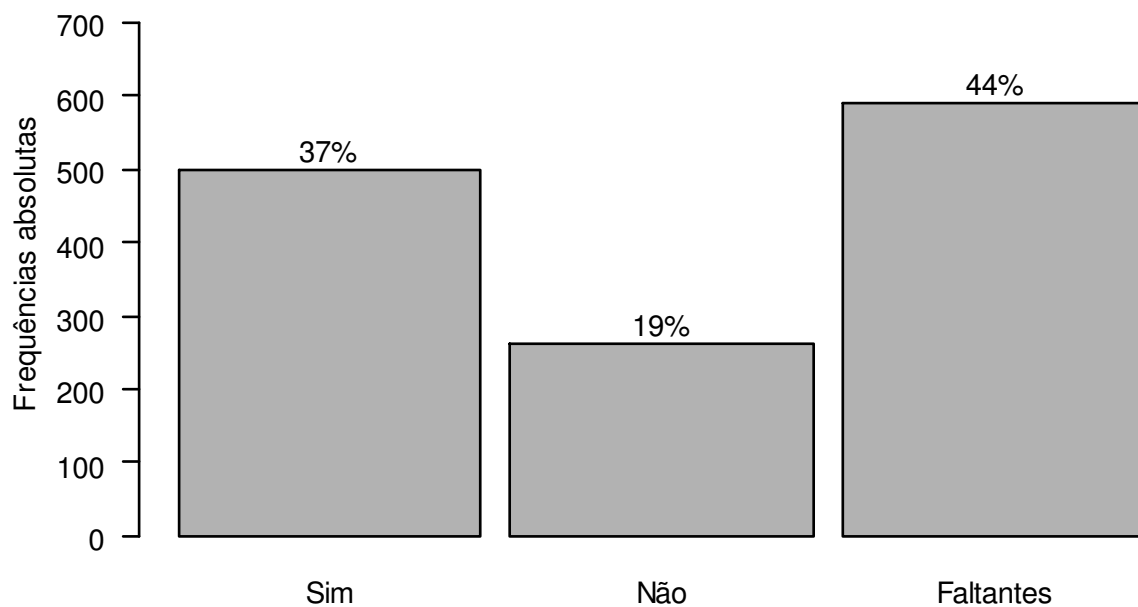
Figura A3: Gráfico de barras por categoria de classe funcional.**Figura A4:** Gráfico de barras por categoria de indicador de cateterismo.

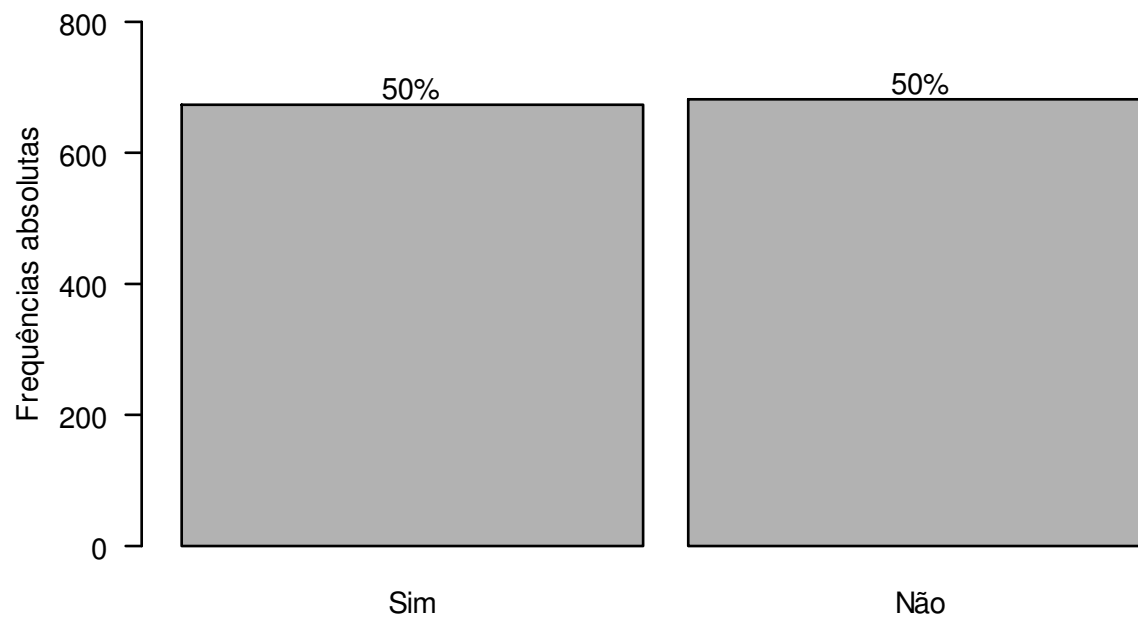
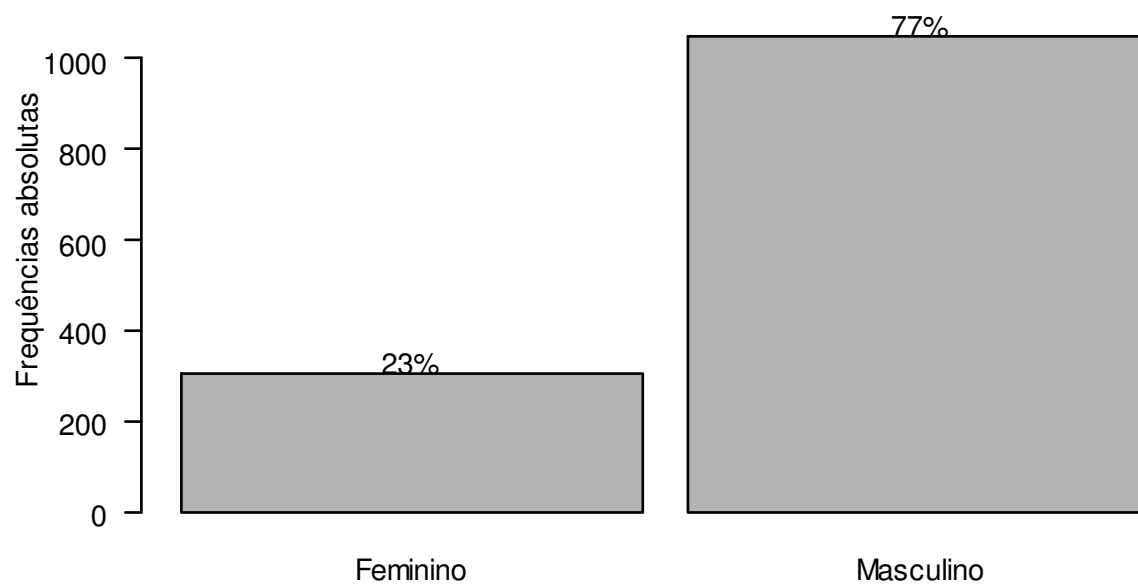
Figura A5: Gráfico de barras por categoria de óbito.**Figura A6:** Gráfico de barras por categoria de gênero.

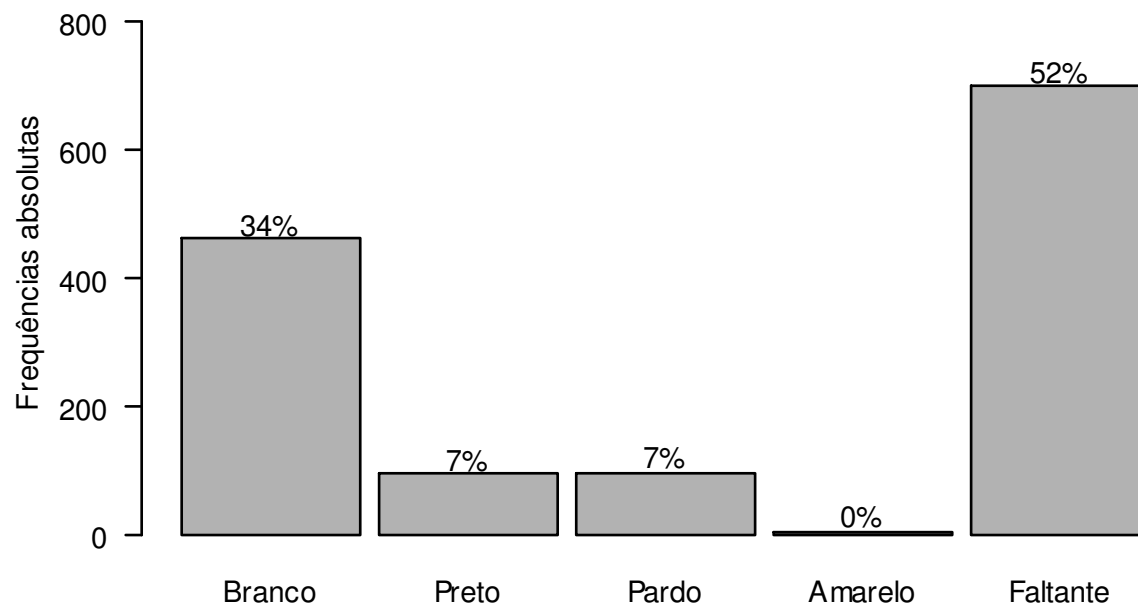
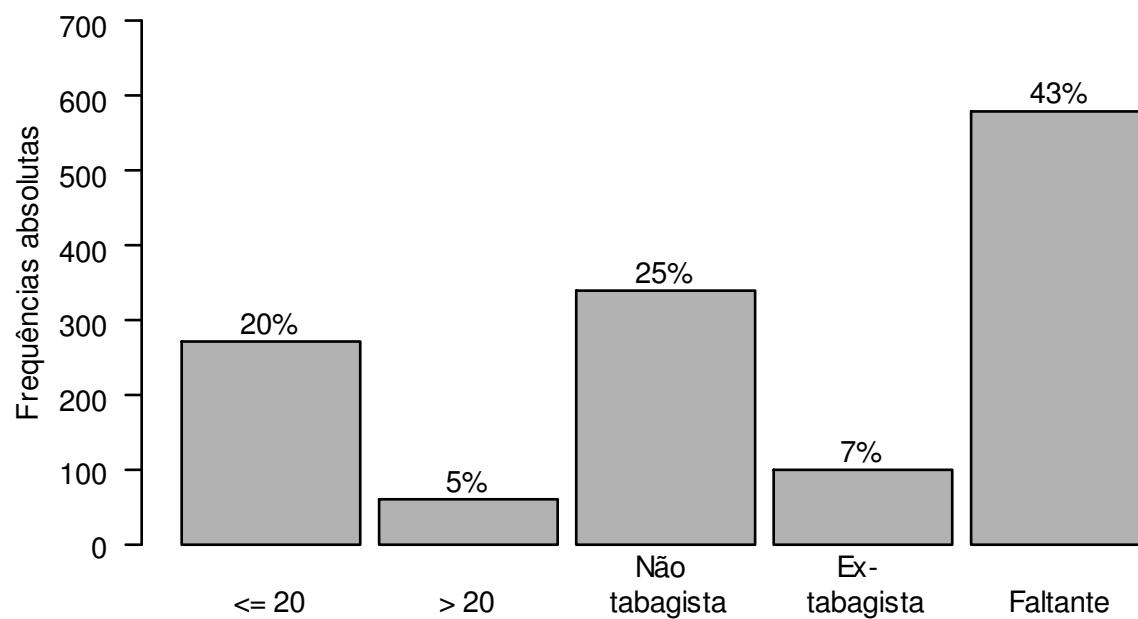
Figura A7: Gráfico de barras por categoria de etnia.**Figura A8:** Gráfico de barras por categoria de tabagismo.

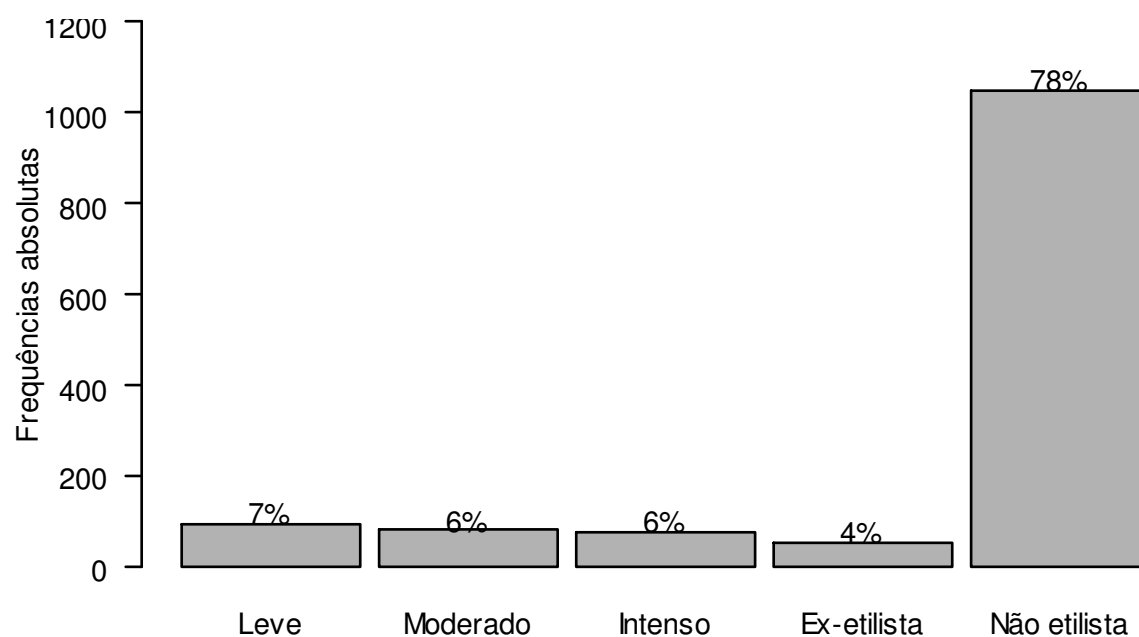
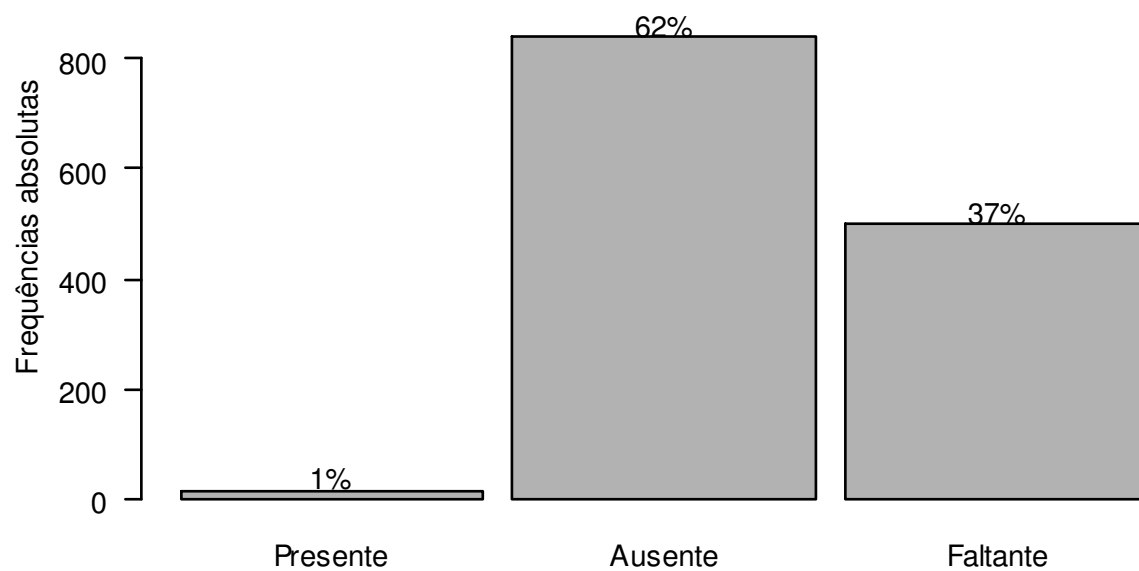
Figura A9: Gráfico de barras por categoria de etilismo.**Figura A10:** Gráfico de barras por categoria de diabetes melito.

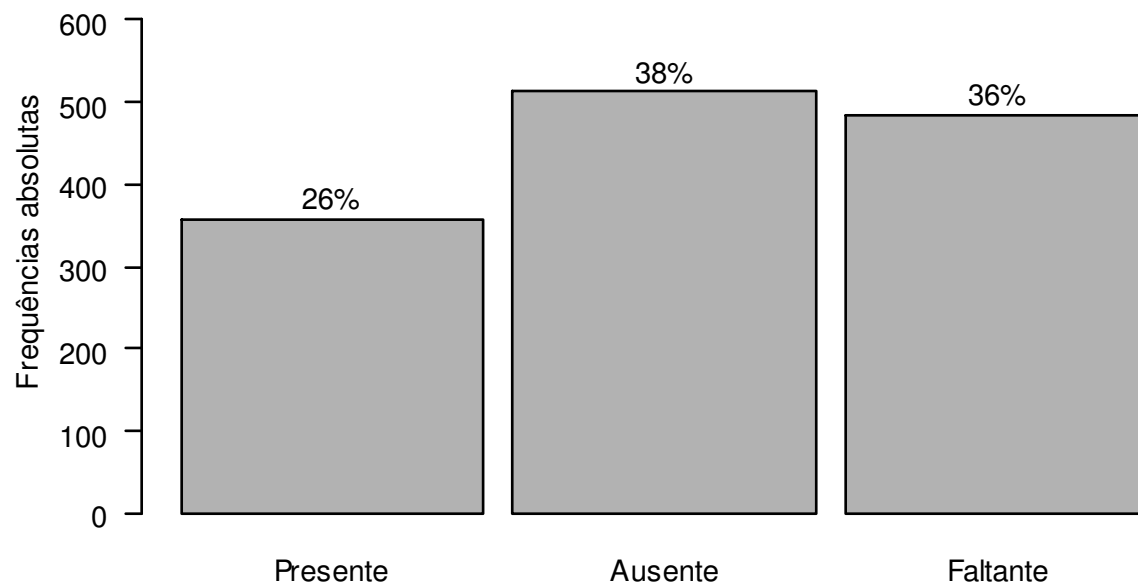
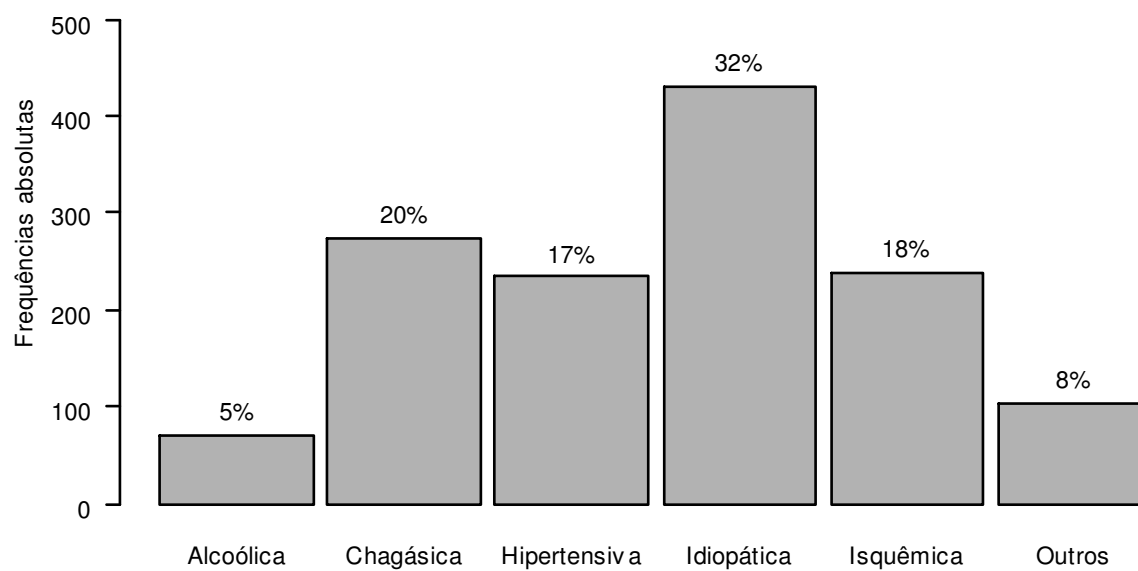
Figura A11: Gráfico de barras por categoria de hipertensão arterial.**Figura A12:** Gráfico de barras por categoria de diagnóstico etiológico.

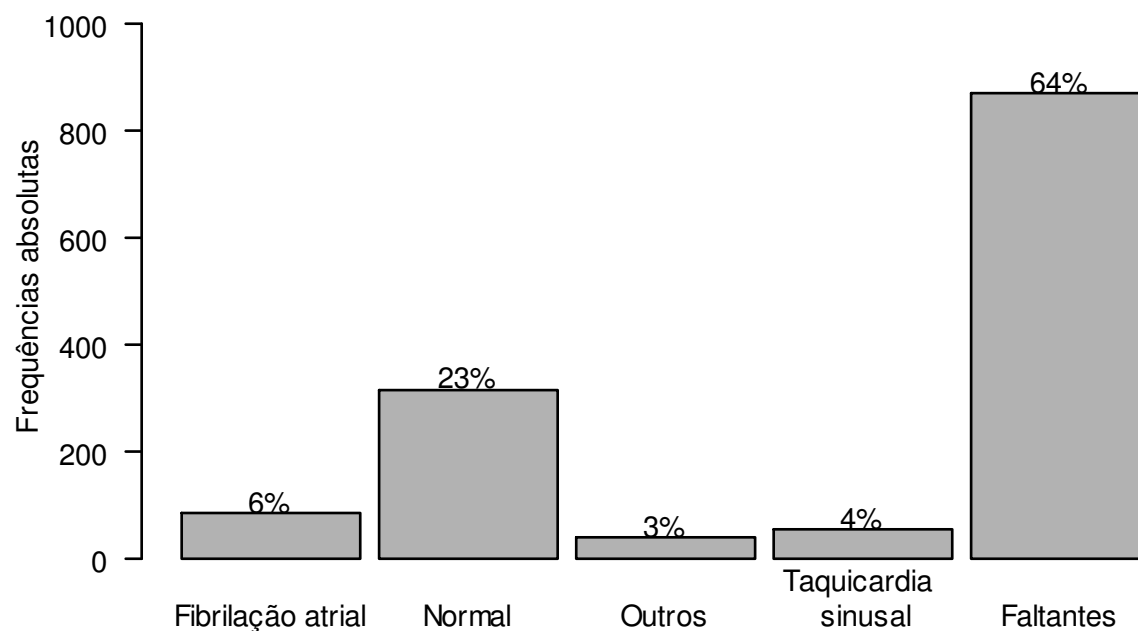
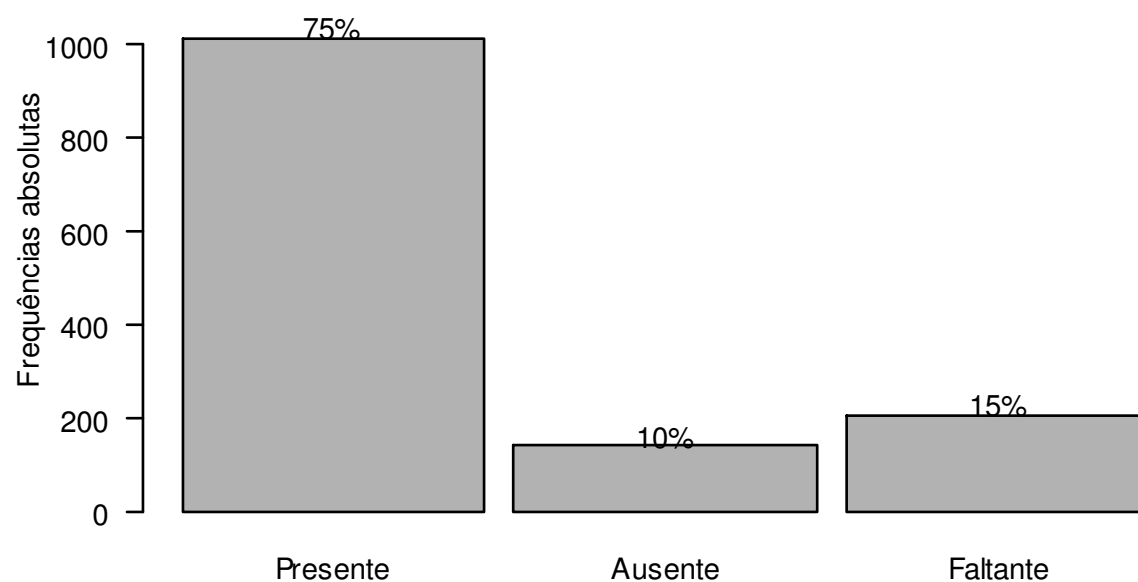
Figura A13: Gráfico de barras por categoria de ritmo do eletrocardiograma.**Figura A14:** Gráfico de barras por categoria de insuficiência mitral Doppler.

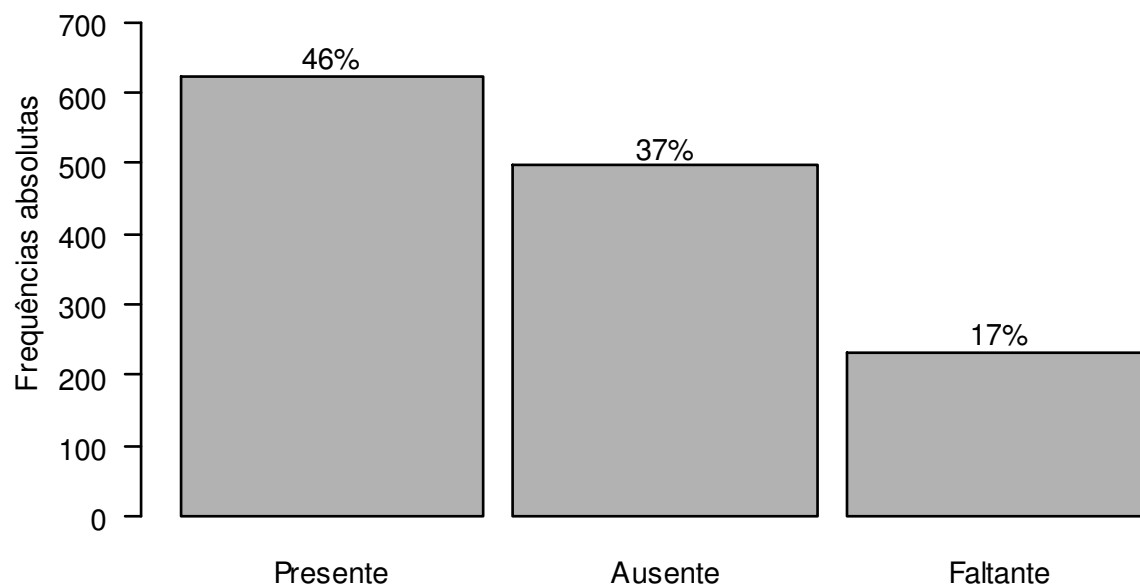
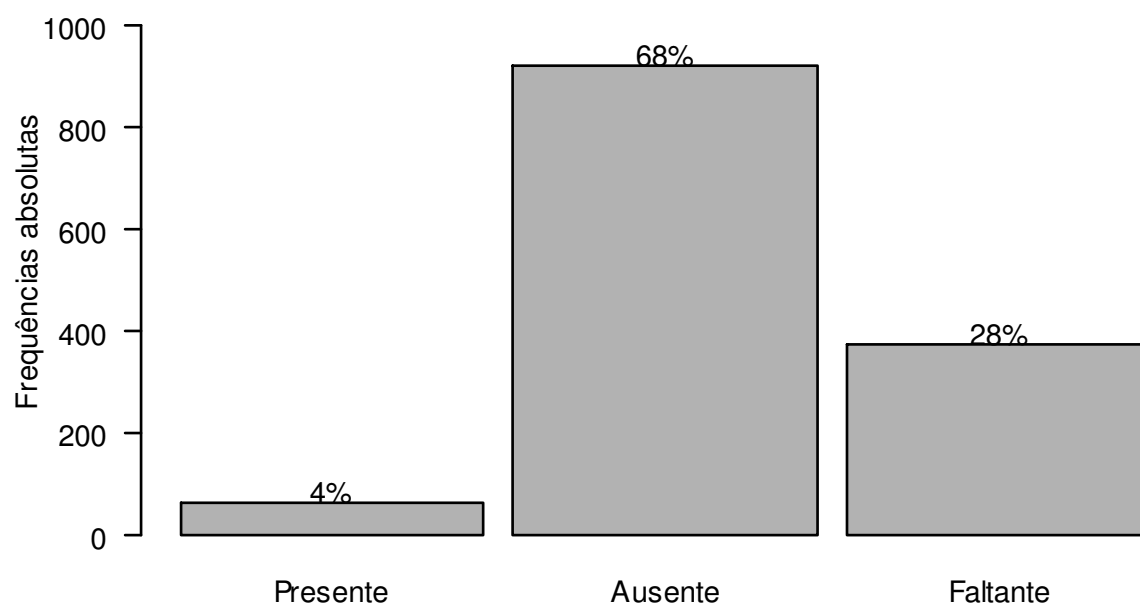
Figura A15: Gráfico de barras por categoria de insuficiência tricúspide Doppler.**Figura A16:** Gráfico de barras por categoria de trombo no ventrículo esquerdo.

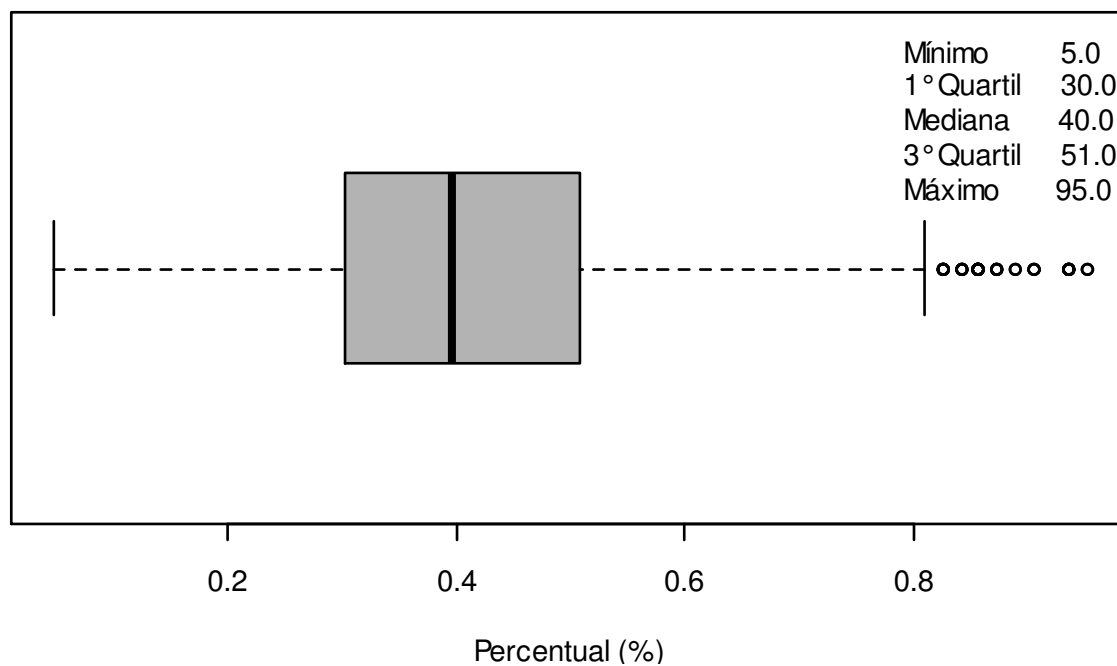
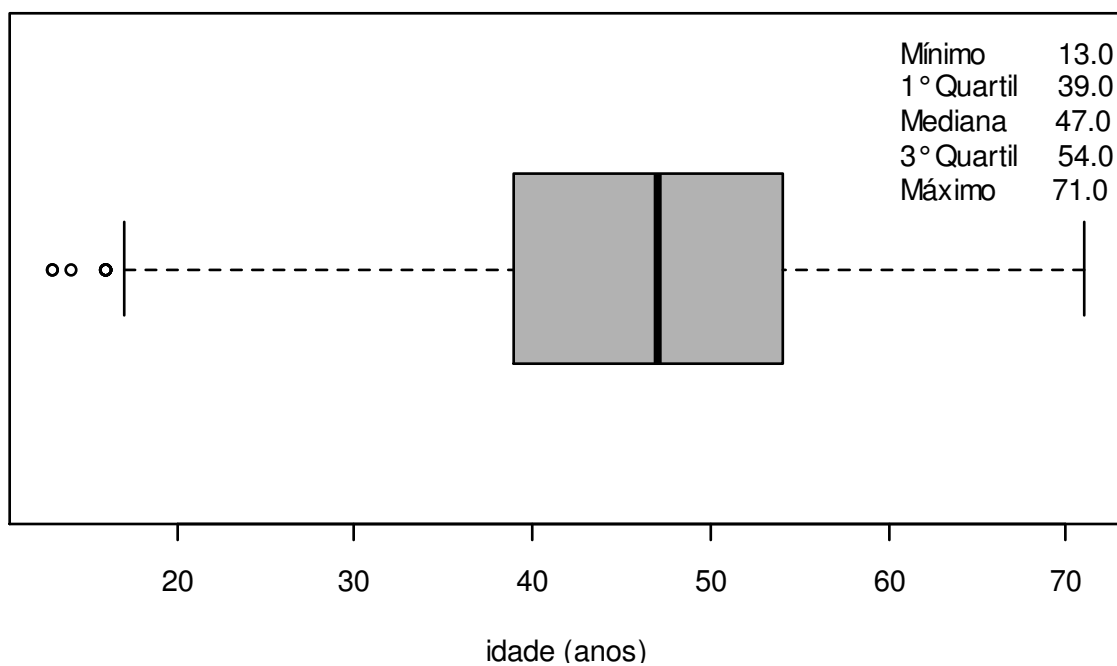
Figura A17: Box plot para percentual de dados faltantes.**Figura A18:** Box plot para idade (anos).

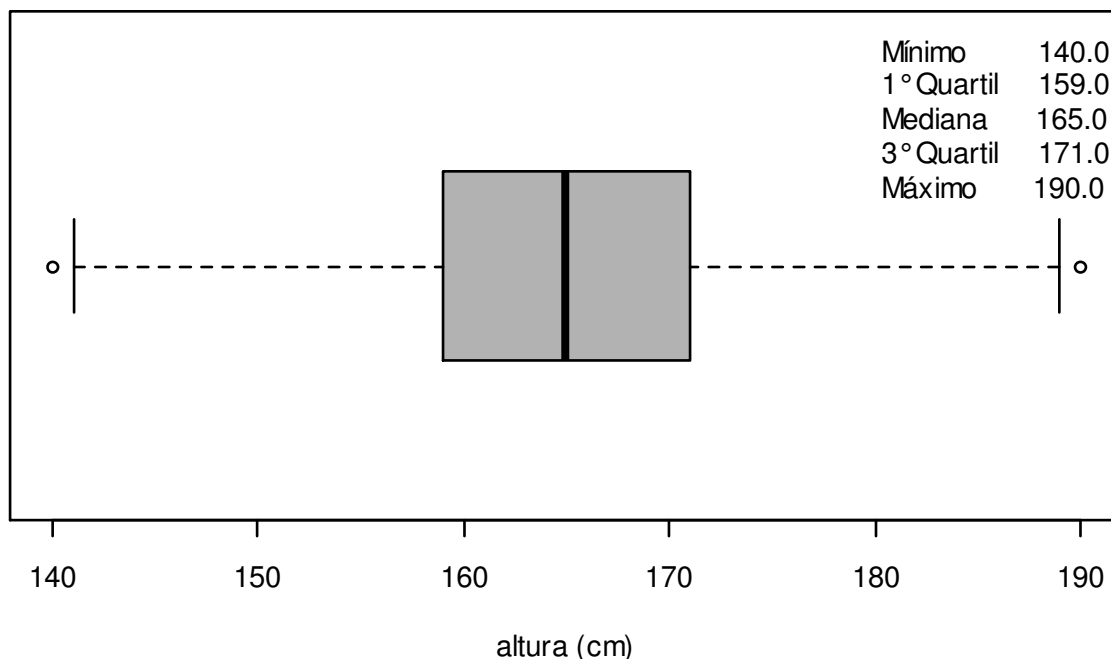
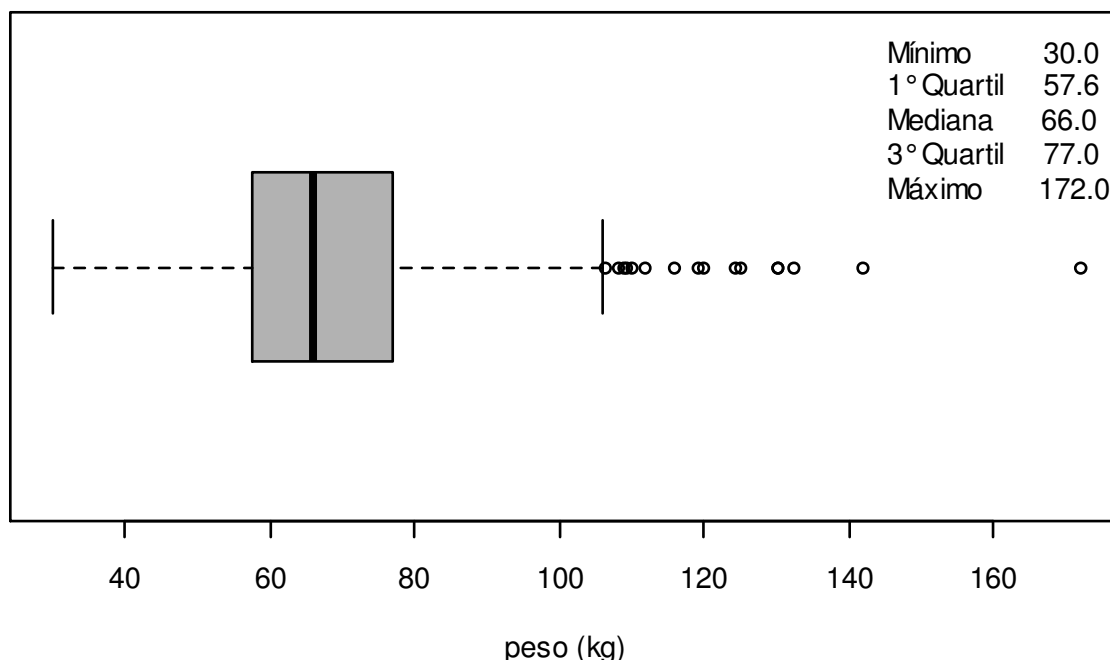
Figura A19: Box plot para altura (cm).**Figura A20:** Box plot para peso (kg).

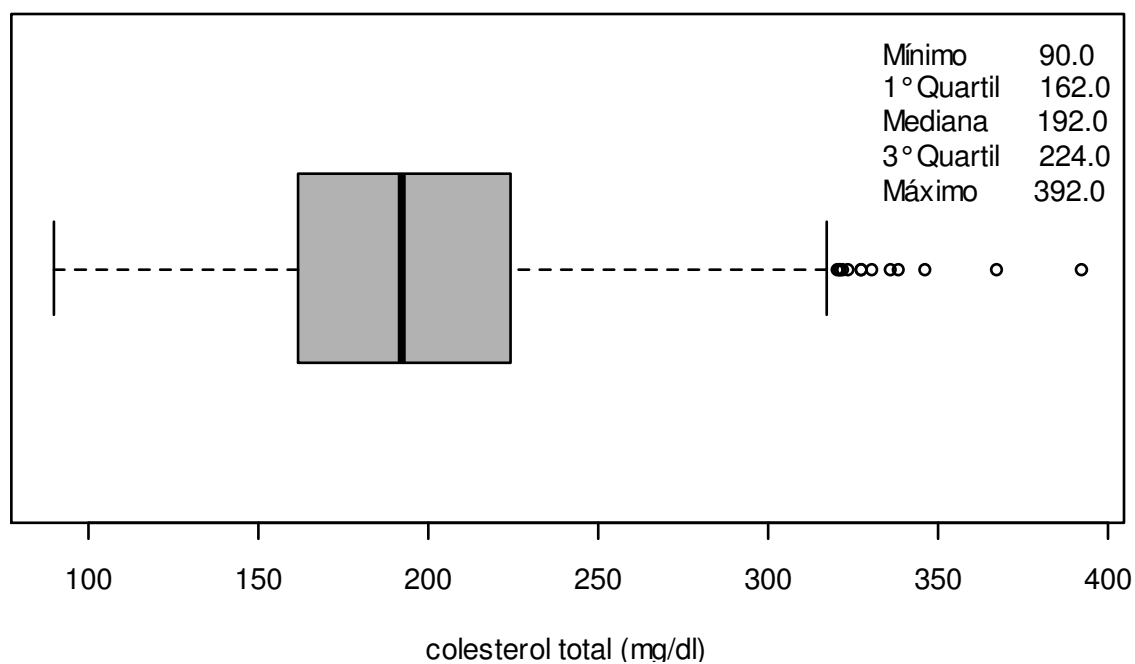
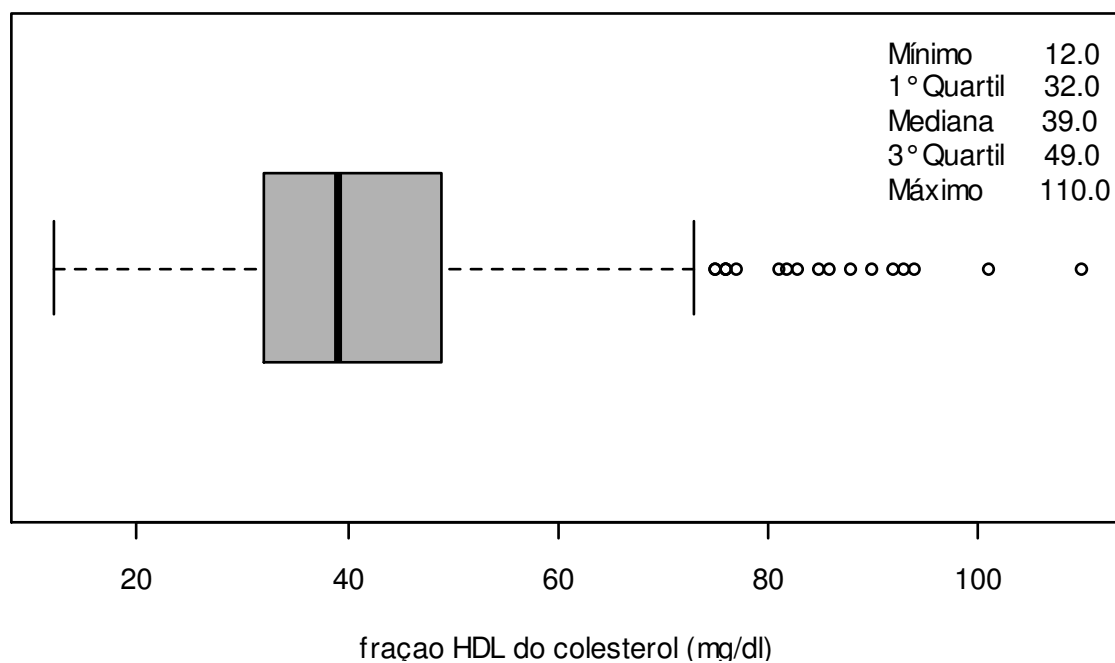
Figura A21: Box plot para colesterol total (mg/dl).**Figura A22:** Box plot para fração HDL do colesterol (mg/dl).

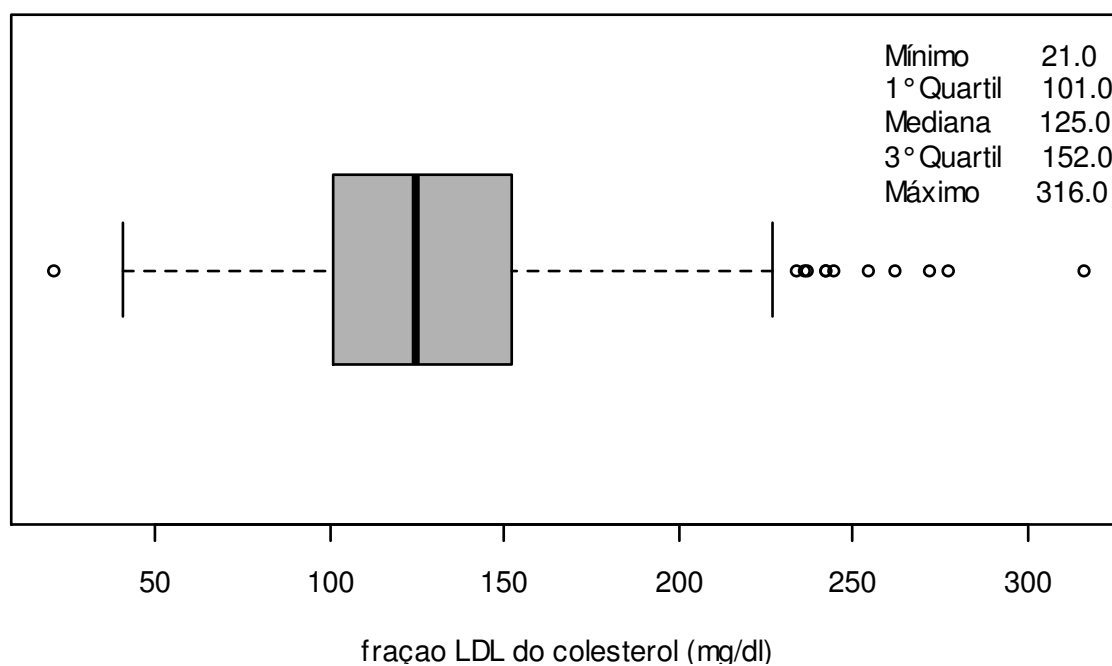
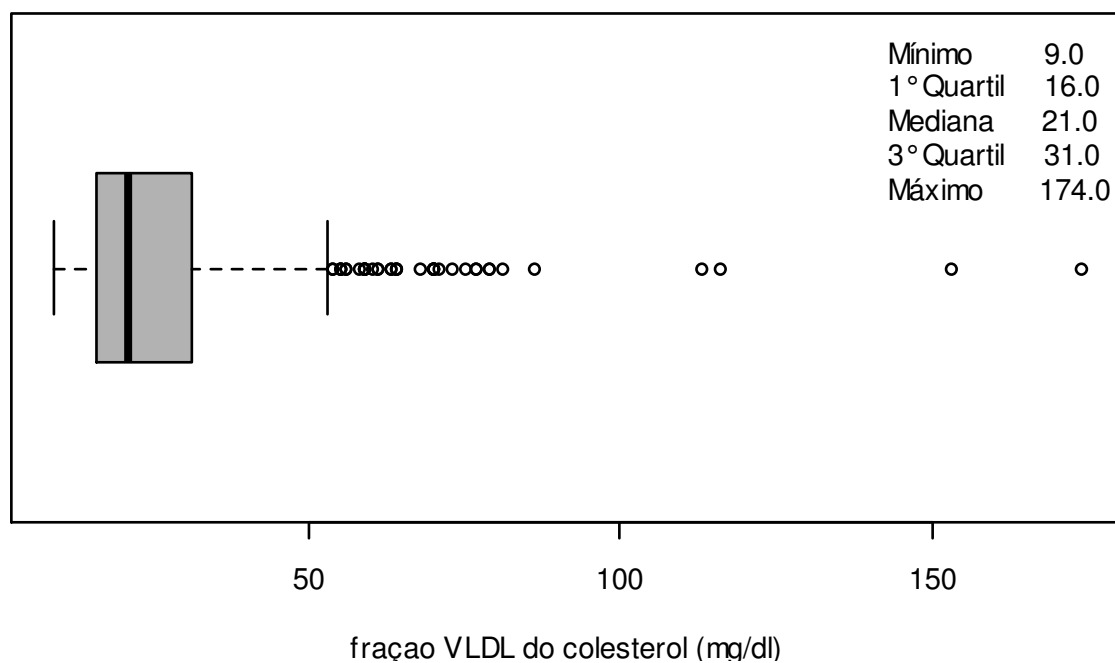
Figura A23: Box plot para fração LDL do colesterol (mg/dl).**Figura A24:** Box plot para fração VLDL do colesterol (mg/dl).

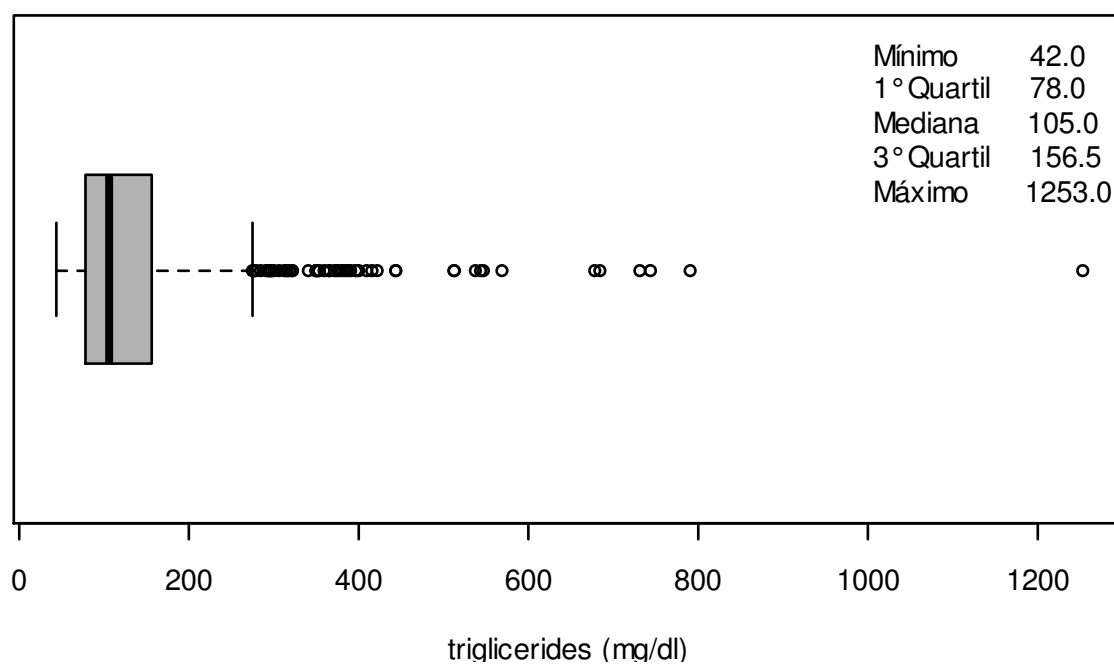
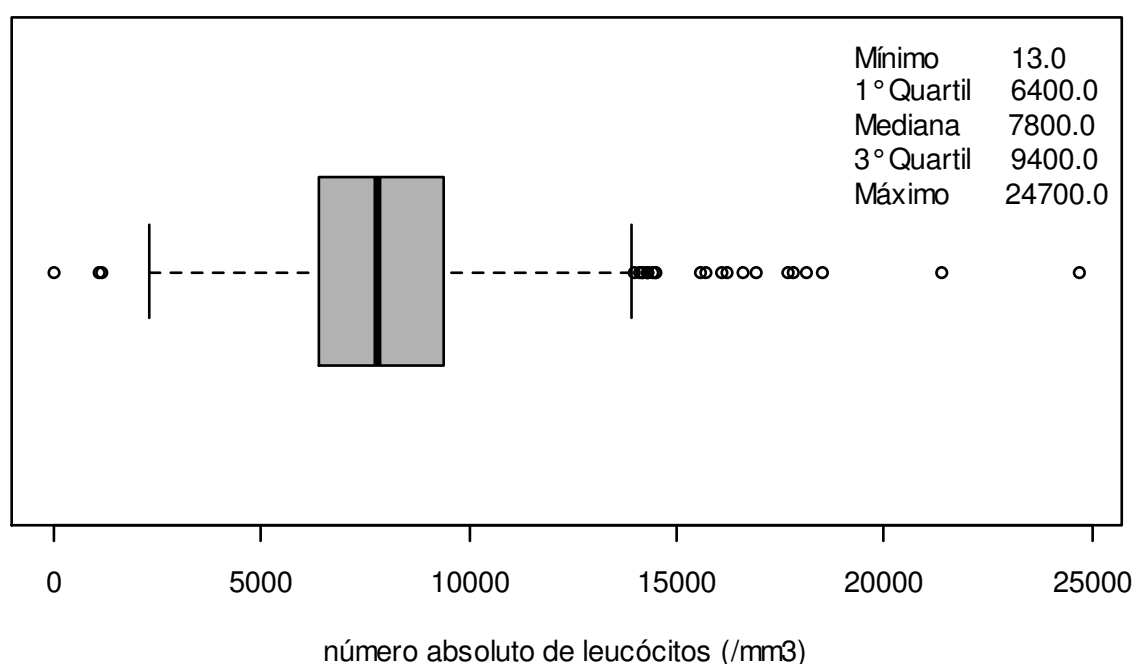
Figura A25: Box plot para triglicérides(mg/dl).**Figura A26:** Box plot para número absoluto de leucócitos (/mm³).

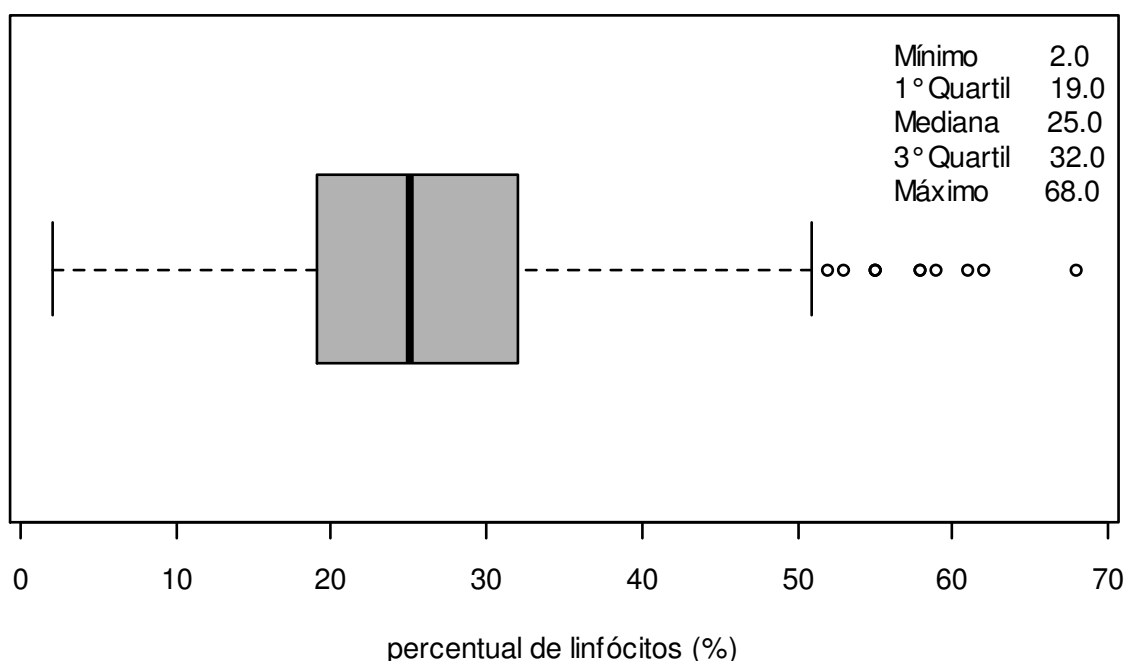
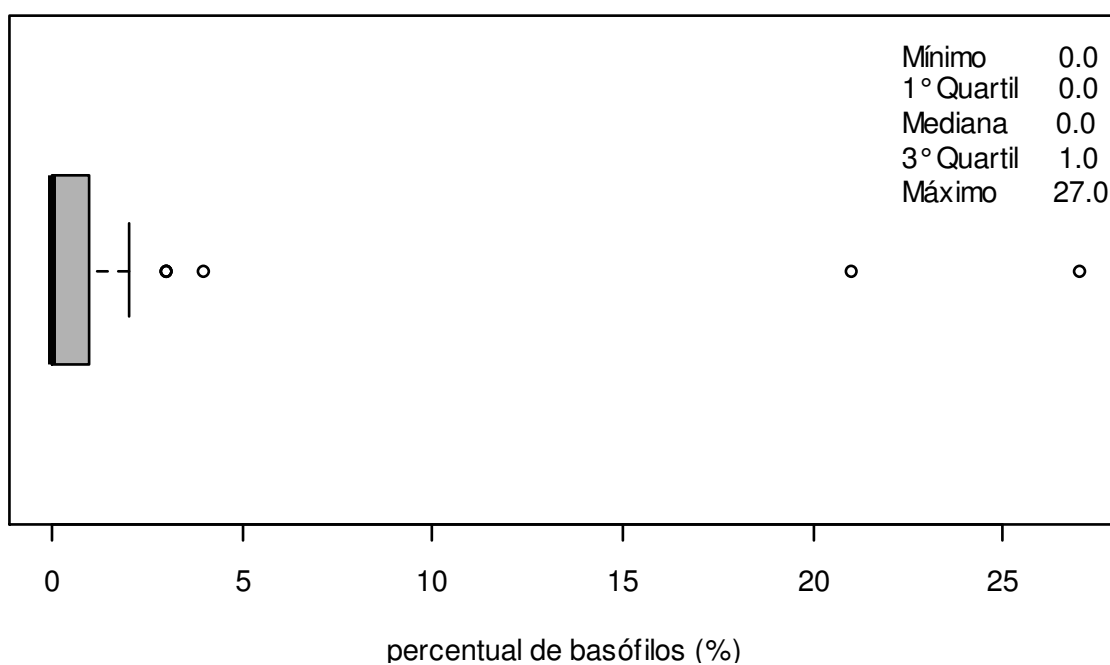
Figura A27: Box plot para percentual de linfócitos (%).**Figura A28:** Box plot para percentual de basófilos (%).

Figura A29: Box plot para percentual de eosinófilos (%).

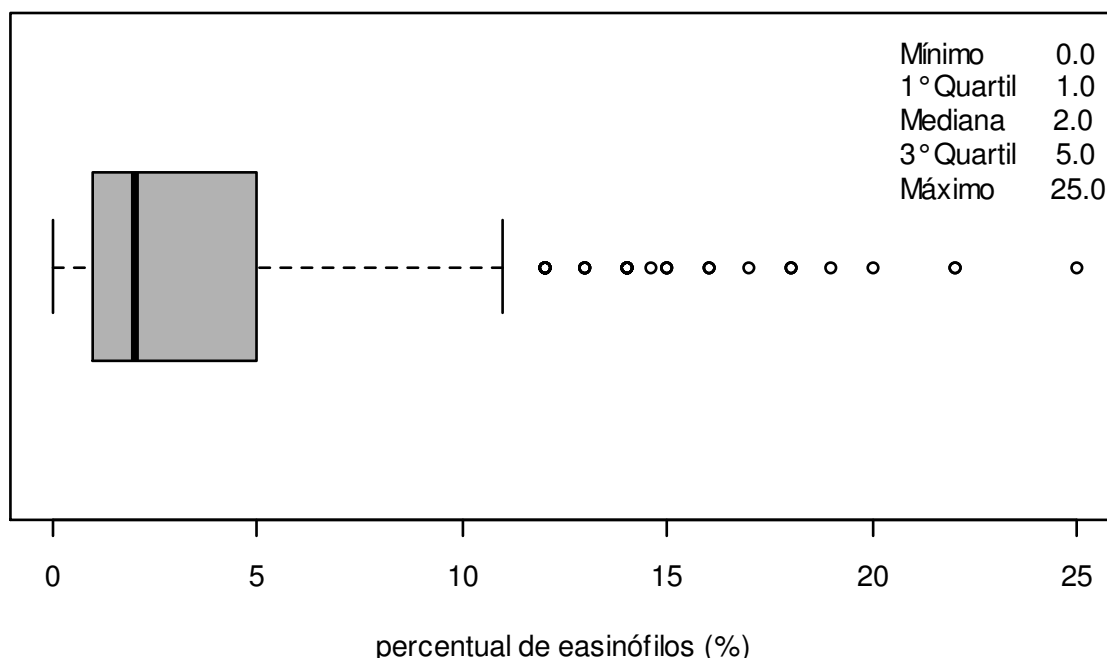


Figura A30: Box plot para percentual de neutrófilos (%).

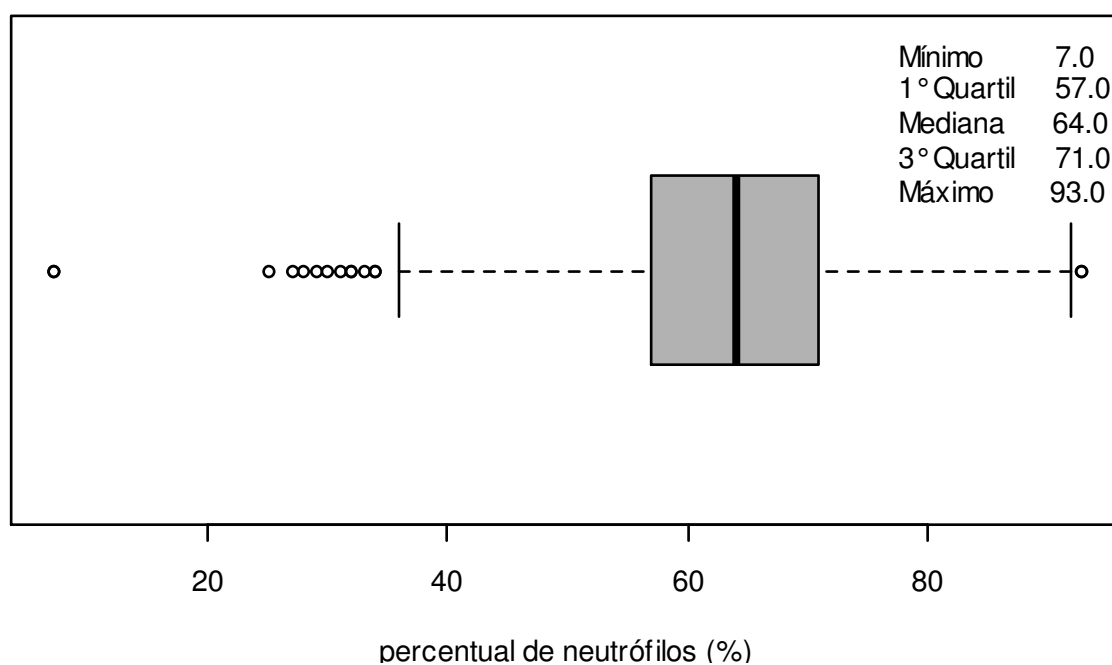


Figura A31: Box plot para percentual de monócitos (%).

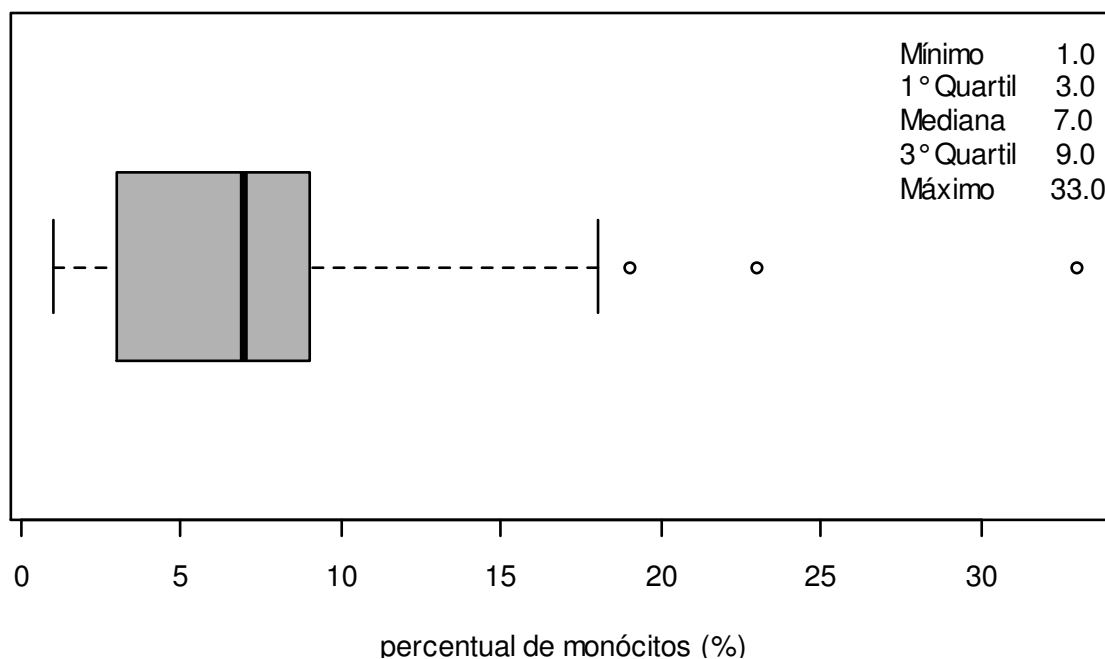


Figura A32: Box plot para número absoluto de plaquetas (/mm³).

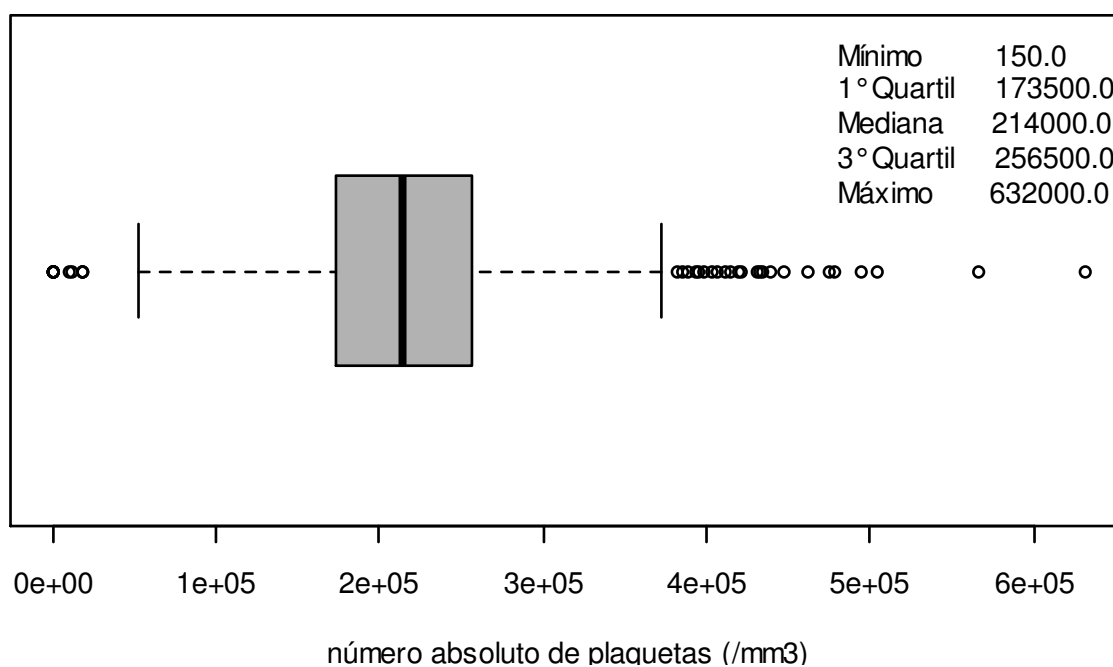


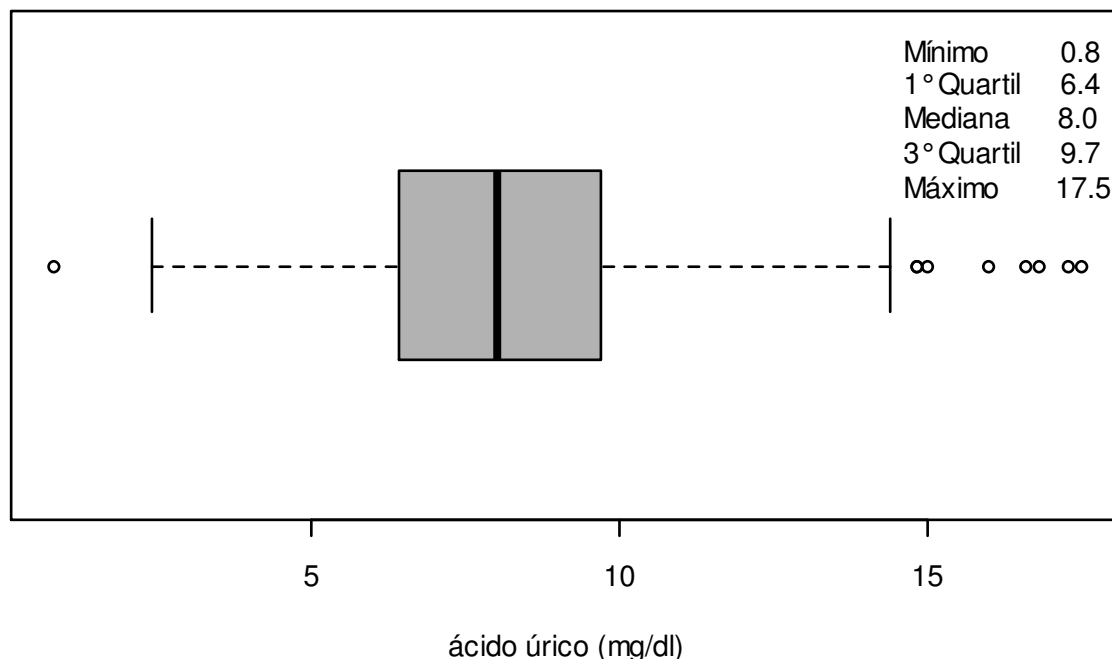
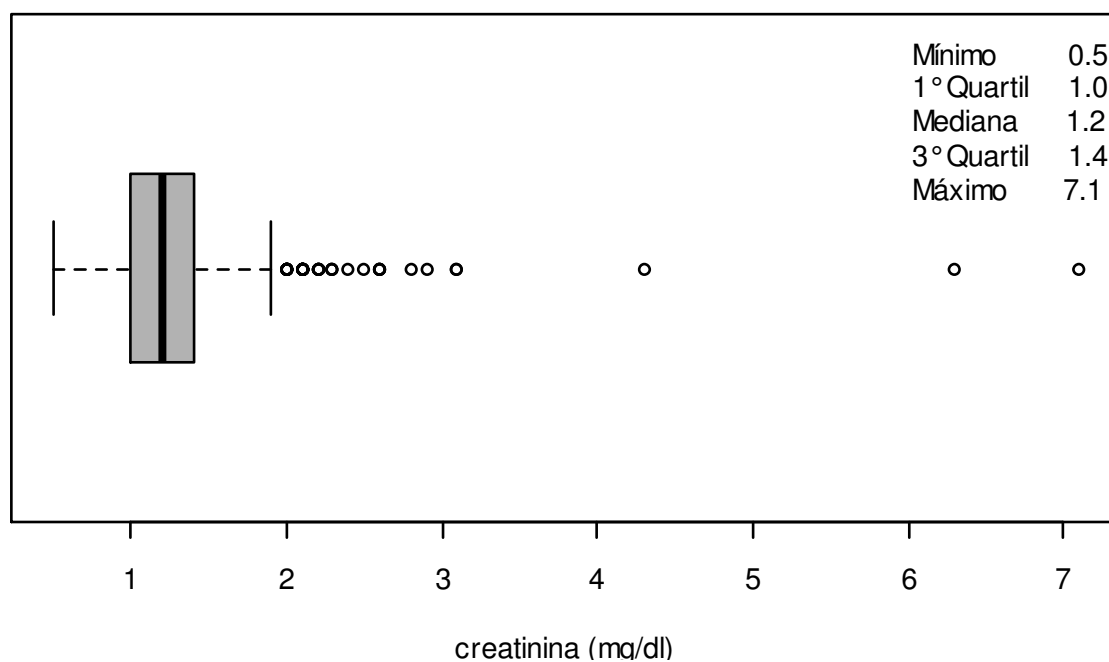
Figura A33: Box plot para ácido úrico (mg/dl).**Figura A34:** Box plot para creatinina (mg/dl).

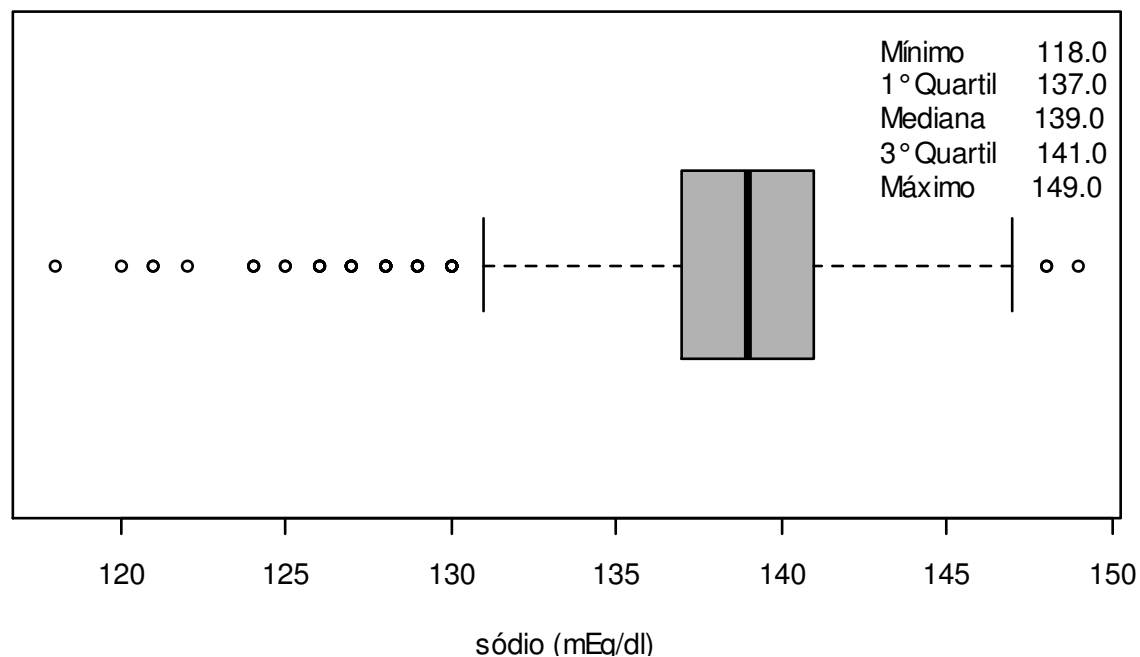
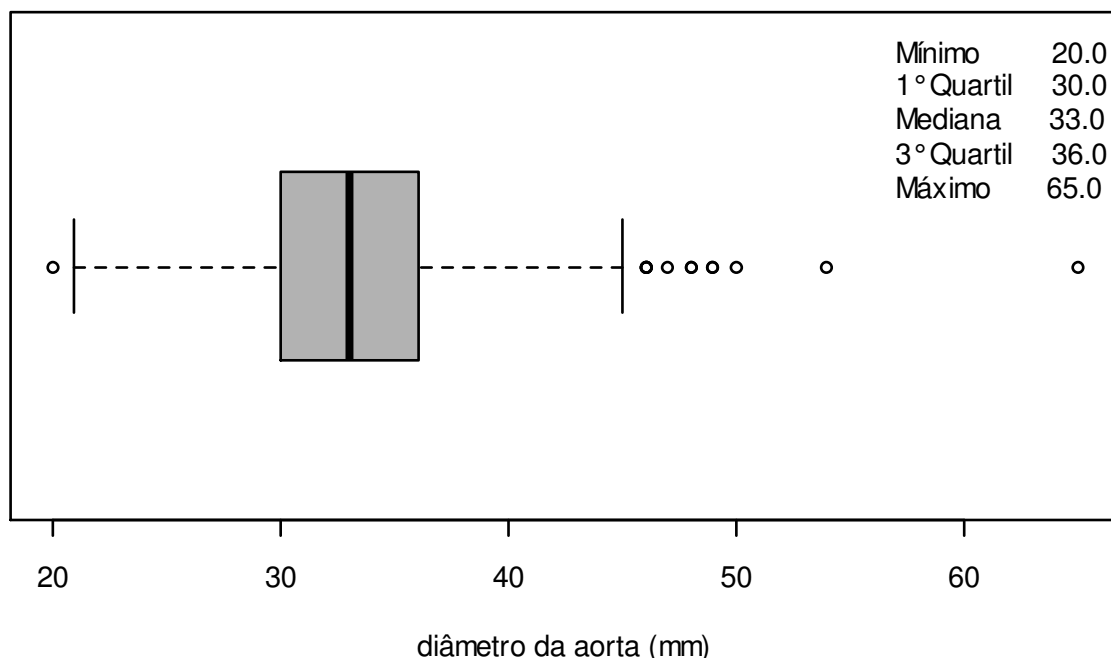
Figura A35: Box plot para sódio (mEq/dl).**Figura A36:** Box plot para diâmetro da aorta (mm).

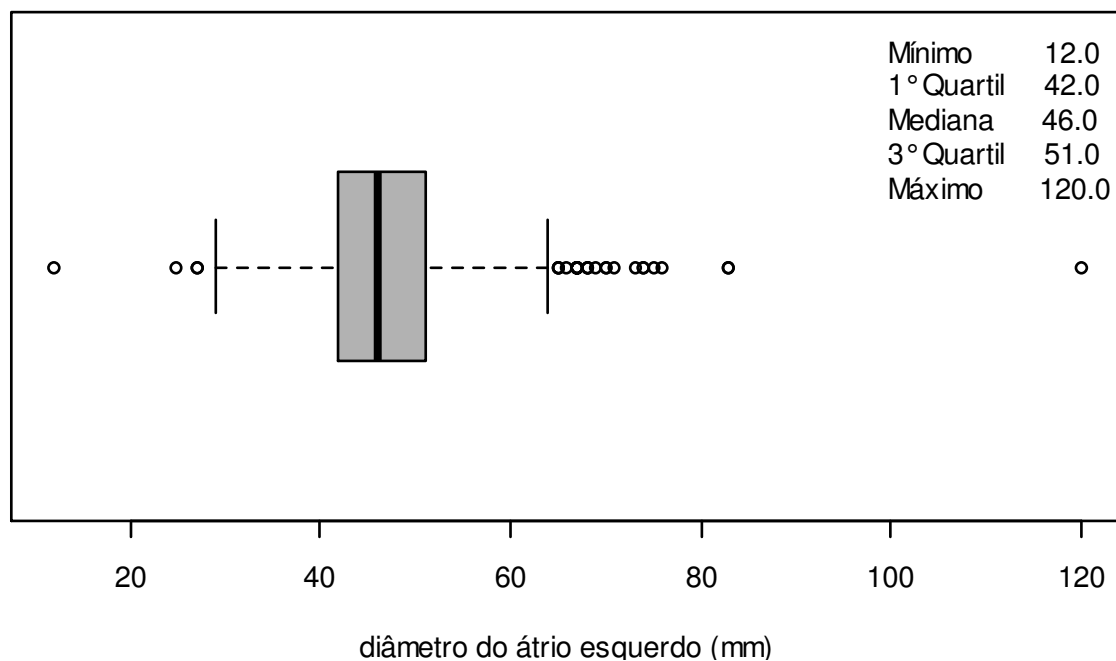
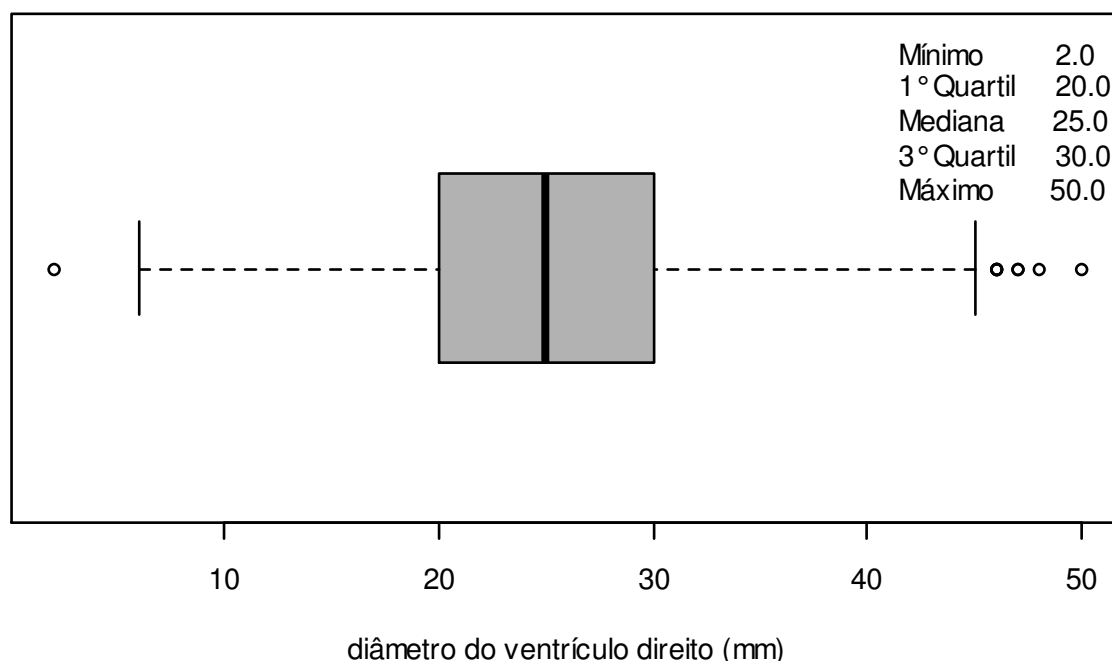
Figura A37: Box plot para diâmetro do átrio esquerdo (mm).**Figura A38:** Box plot para diâmetro do ventrículo direito (mm).

Figura A39: Box plot para diâmetro diastólico do ventrículo esquerdo (mm).

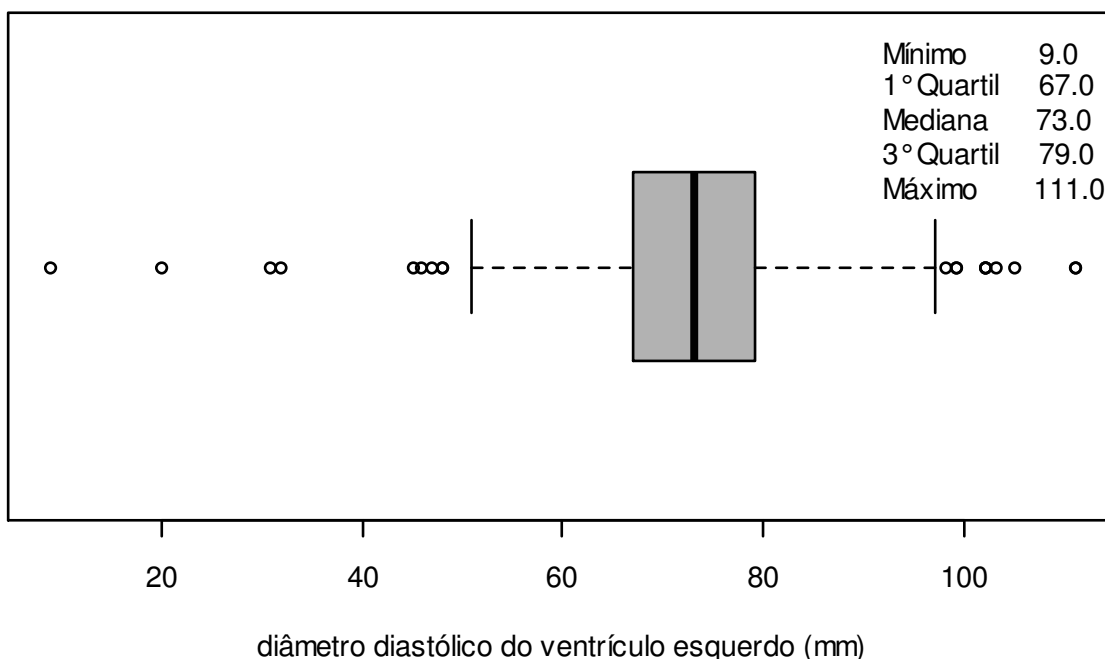


Figura A40: Box plot para diâmetro sistólico do ventrículo esquerdo (mm).

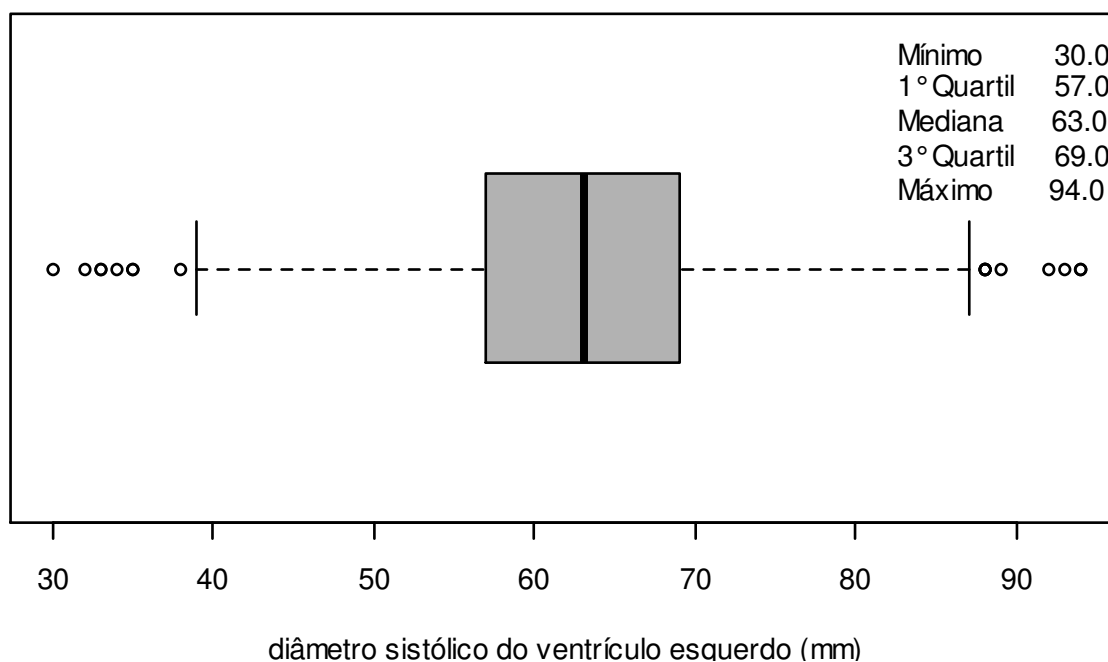


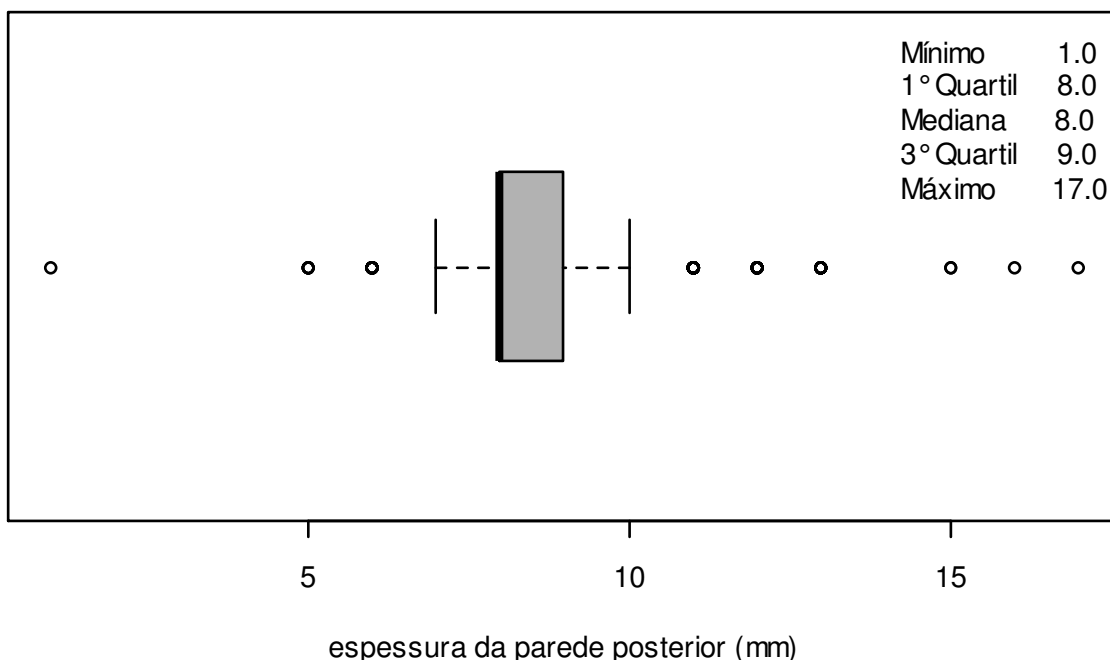
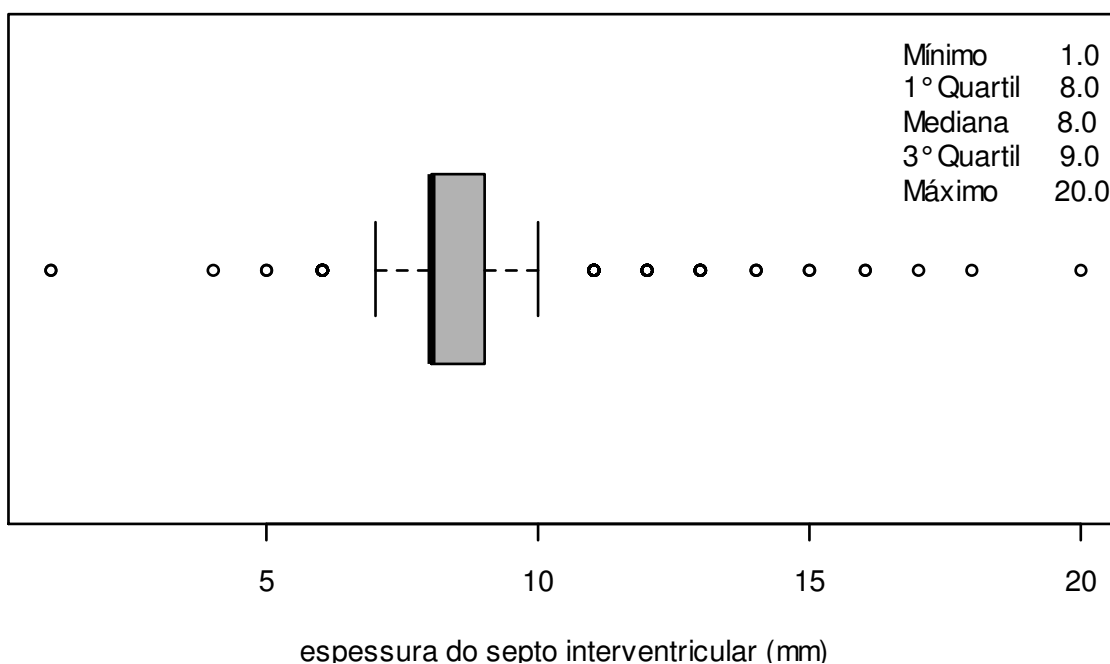
Figura A41: Box plot para espessura da parede posterior (mm).**Figura A42:** Box plot para espessura do septo interventricular (mm).

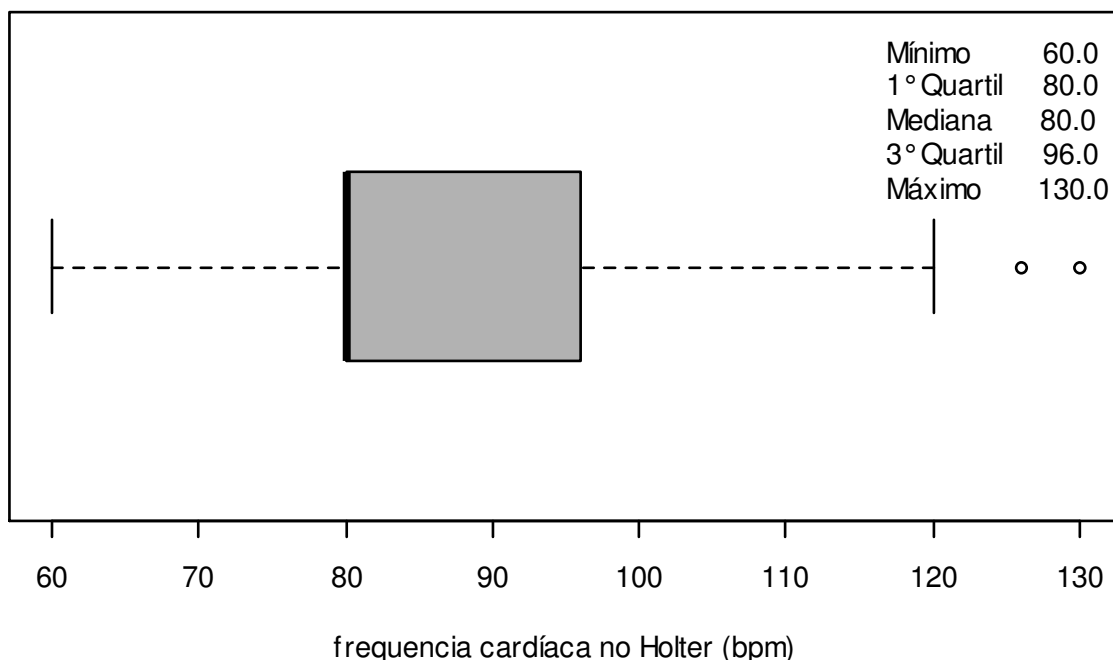
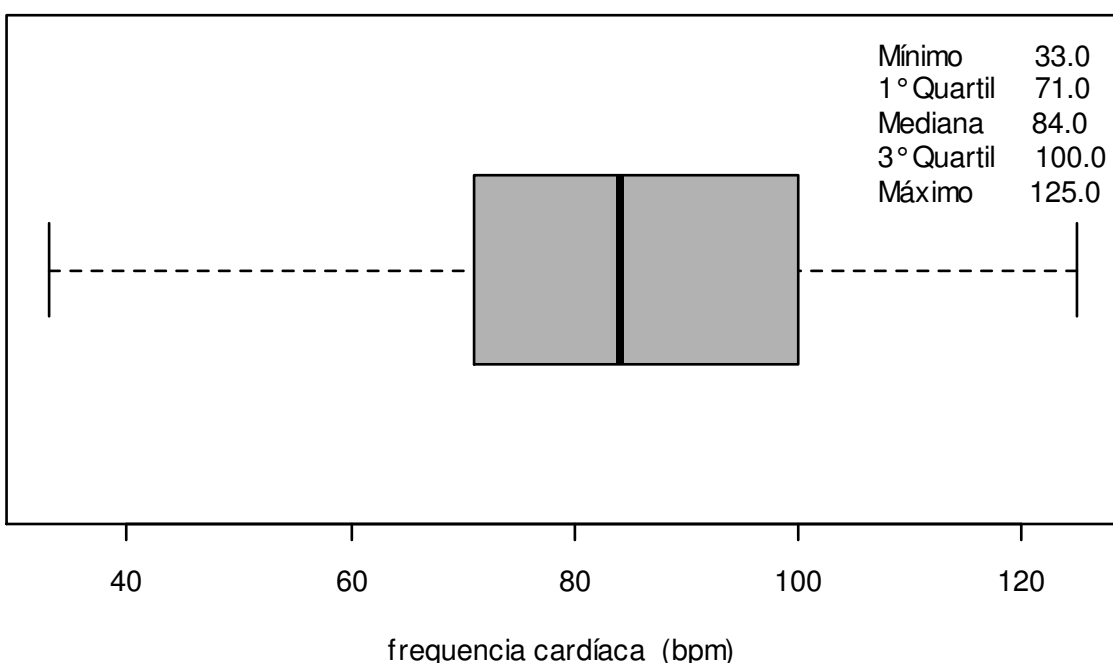
Figura A43: Box plot para frequência cardíaca no Holter (bpm).**Figura A44:** Box plot para frequência cardíaca (bpm).

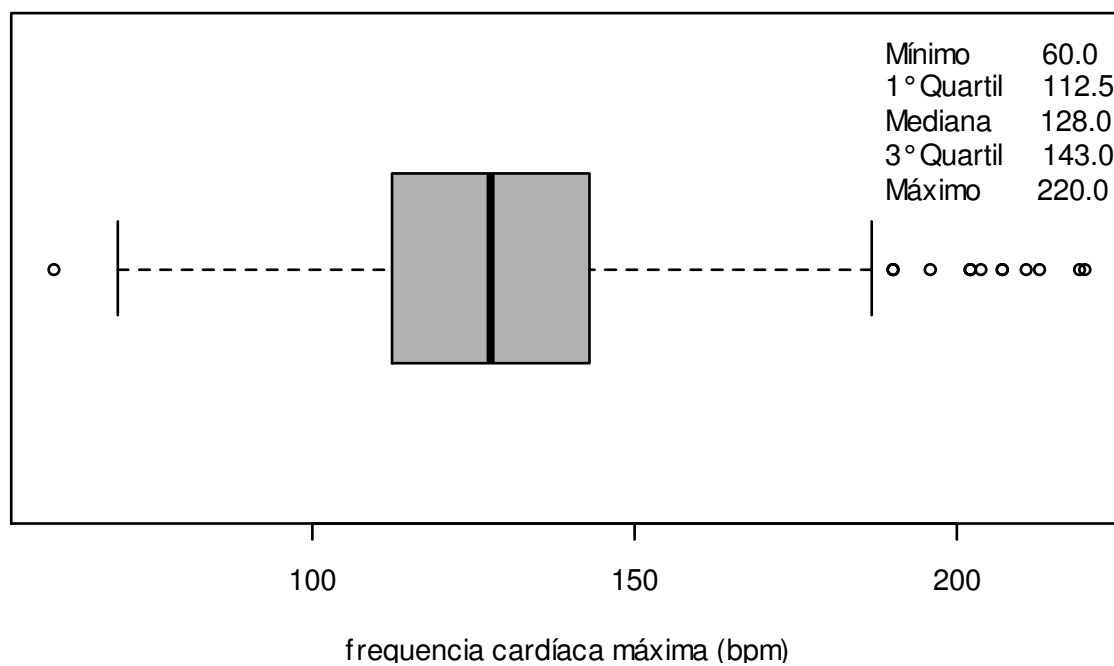
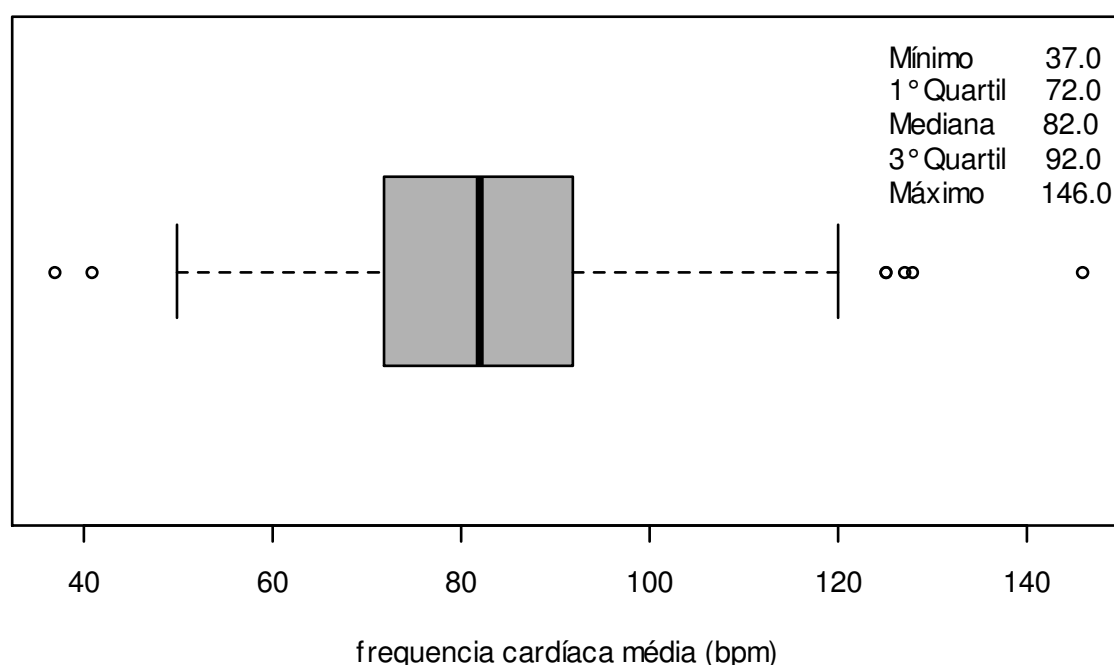
Figura A45: Box plot para frequência cardíaca máxima (bpm).**Figura A46:** Box plot para frequência cardíaca média (bpm).

Figura A47: Box plot para frequência cardíaca mínima (bpm).

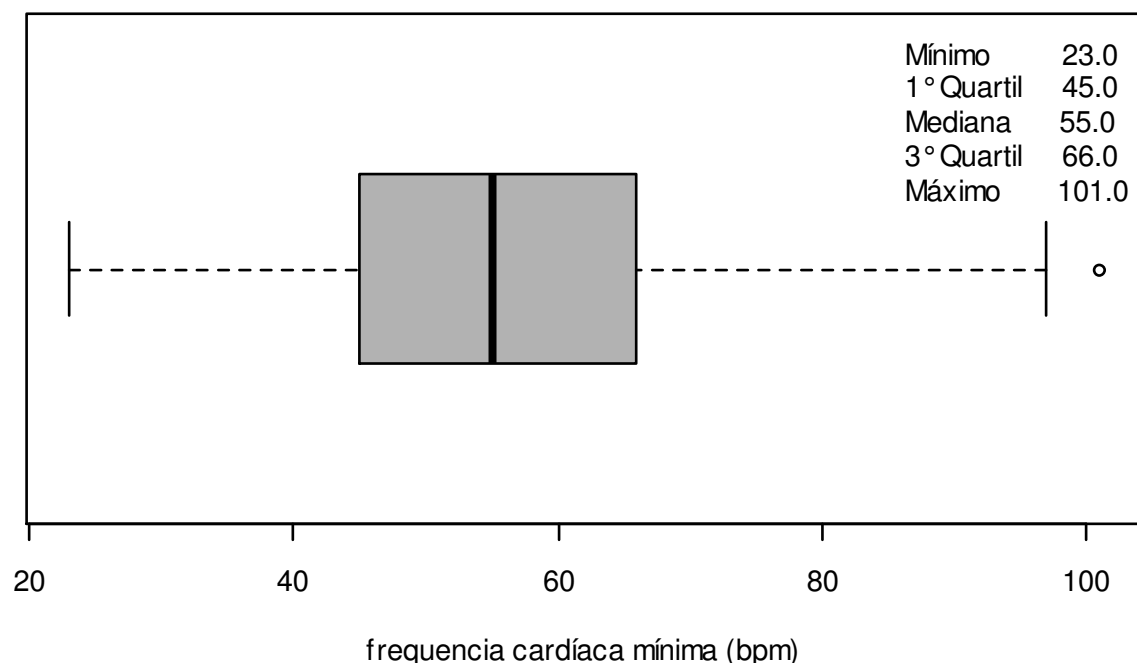


Figura A48: Box plot para fração de ejeção do ventrículo esquerdo no ecocardiograma (%).

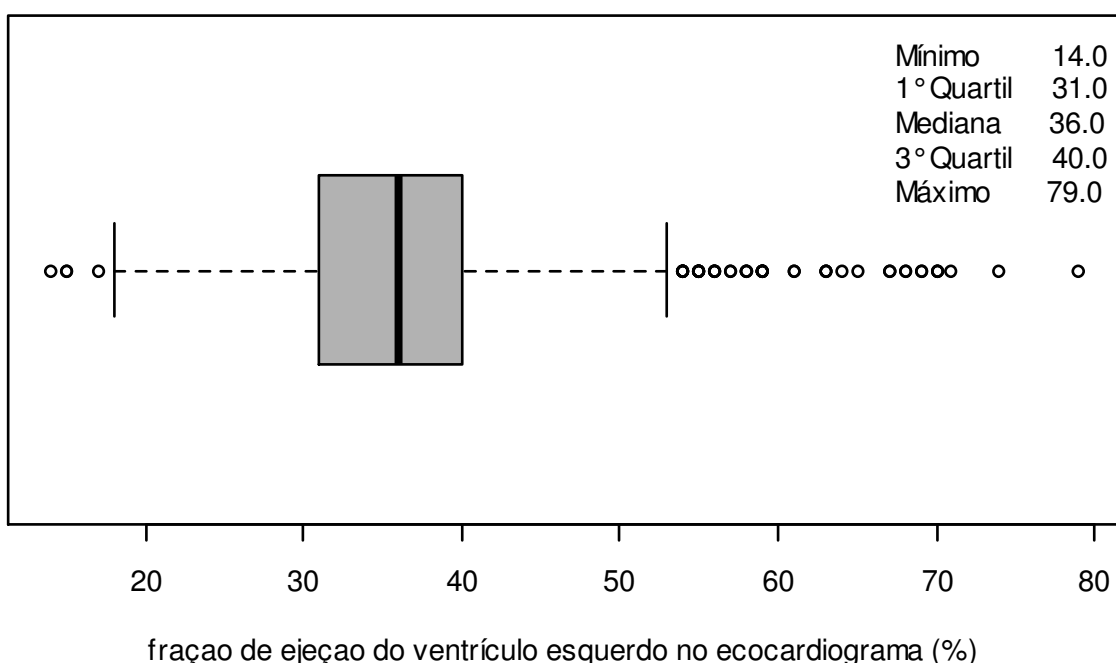


Figura A49: Box plot para fração de ejeção do ventrículo esquerdo (%).

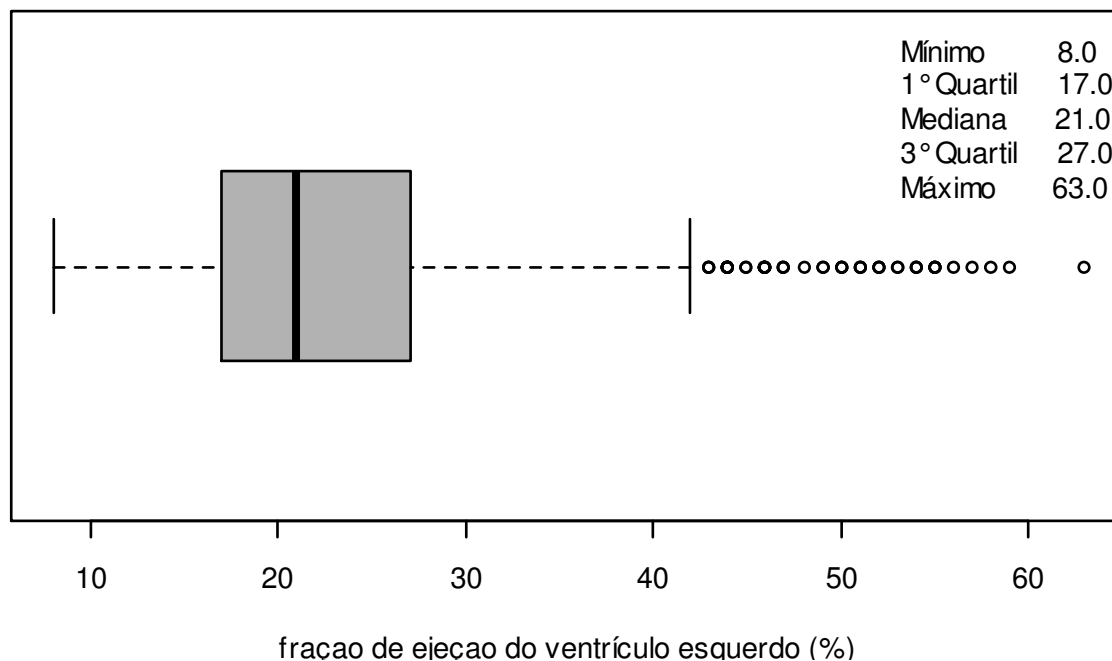


Figura A50: Box plot para fração de ejeção do ventrículo direito (%).

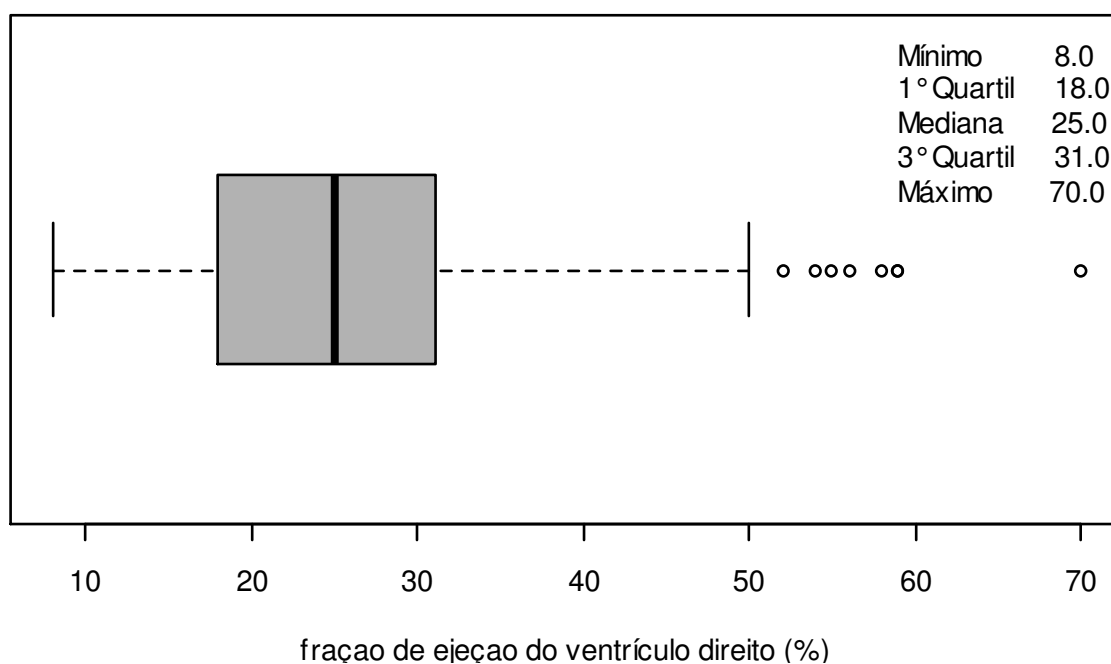


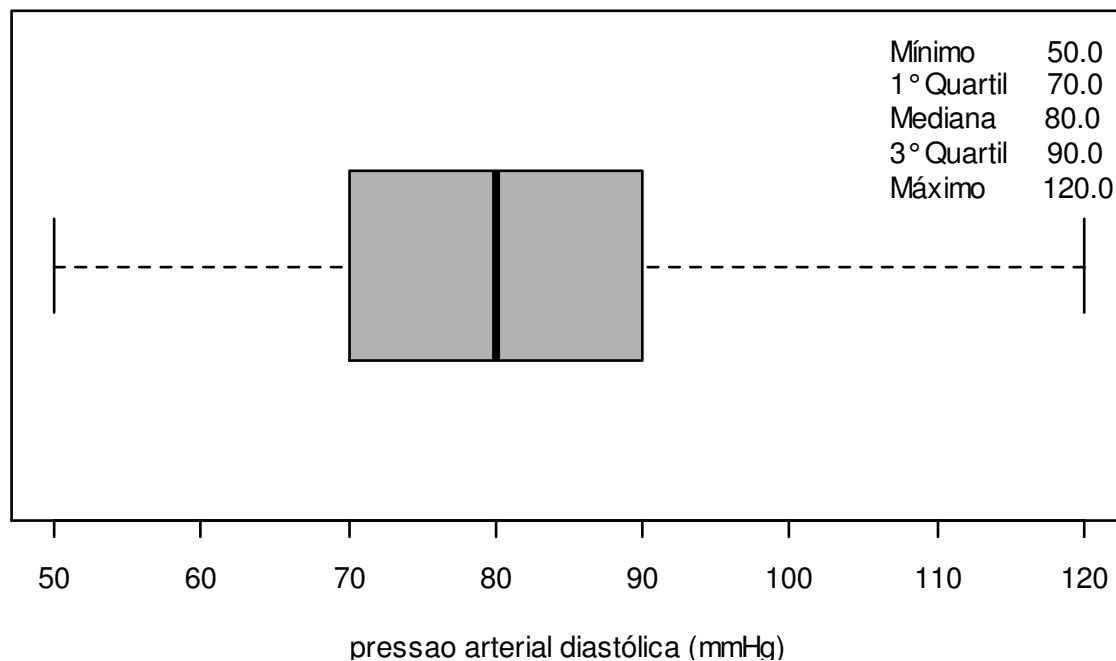
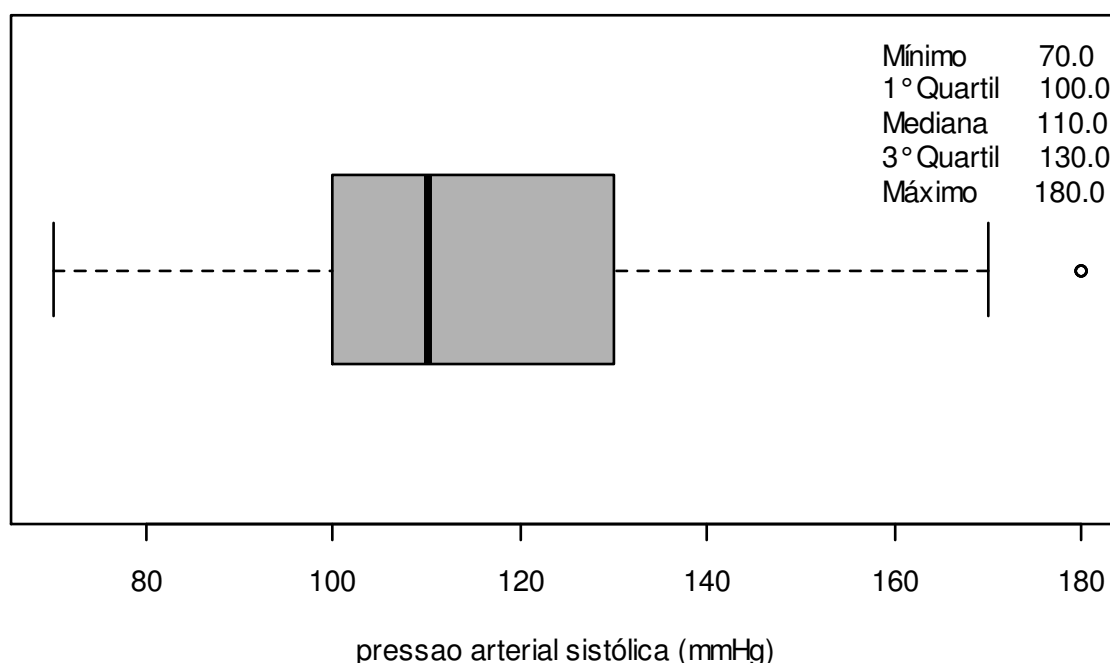
Figura A51: Box plot para pressão arterial diastólica (mmHg).**Figura A52:** Box plot para pressão arterial sistólica (mmHg).

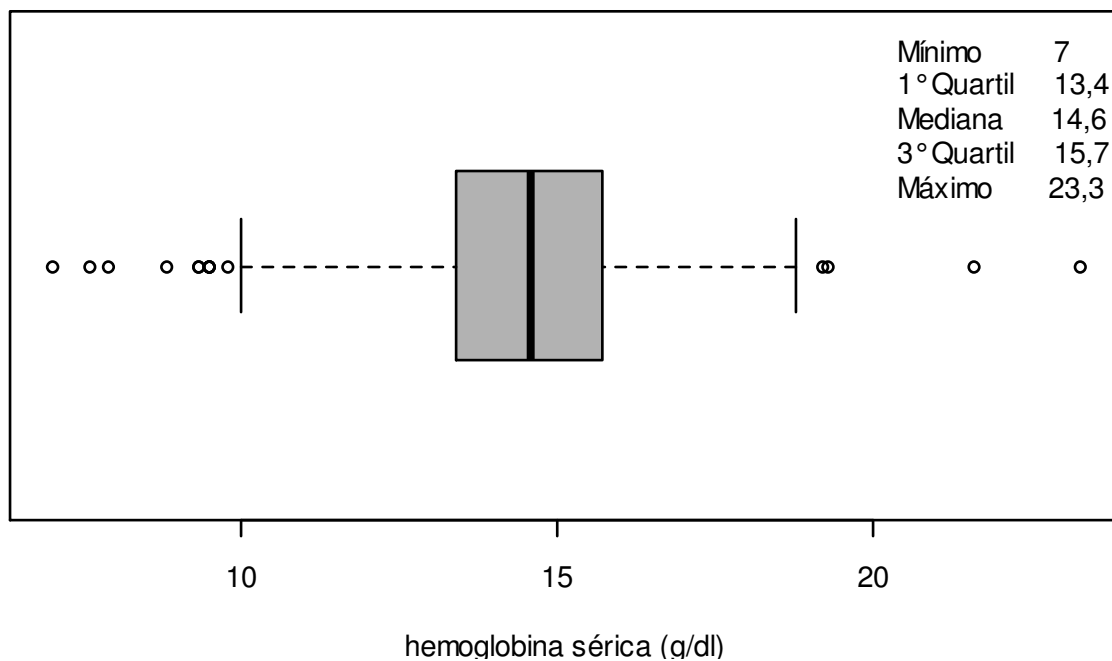
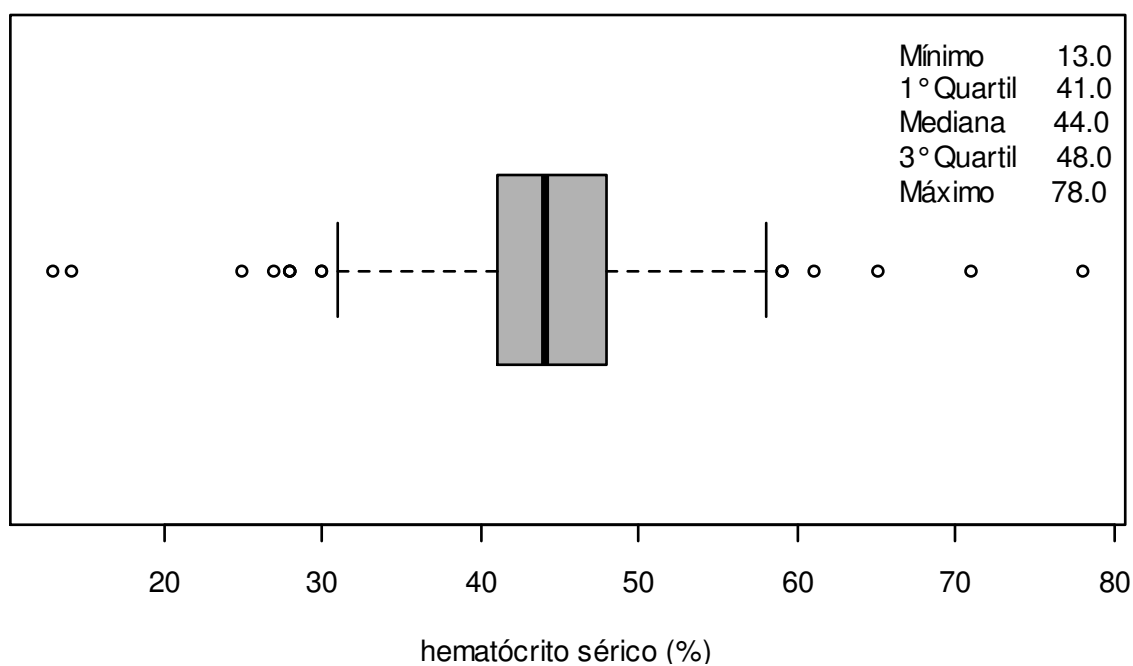
Figura A53: Box plot para hemoglobina sérica (g/dl).**Figura A54:** Box plot para hematócrito sérico (%).

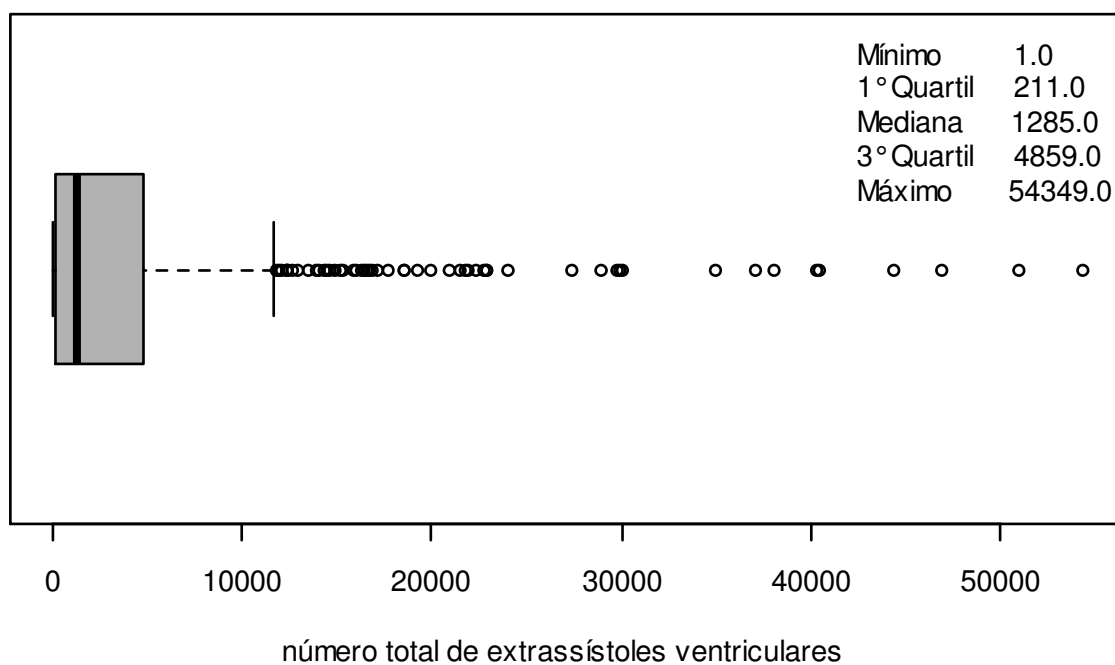
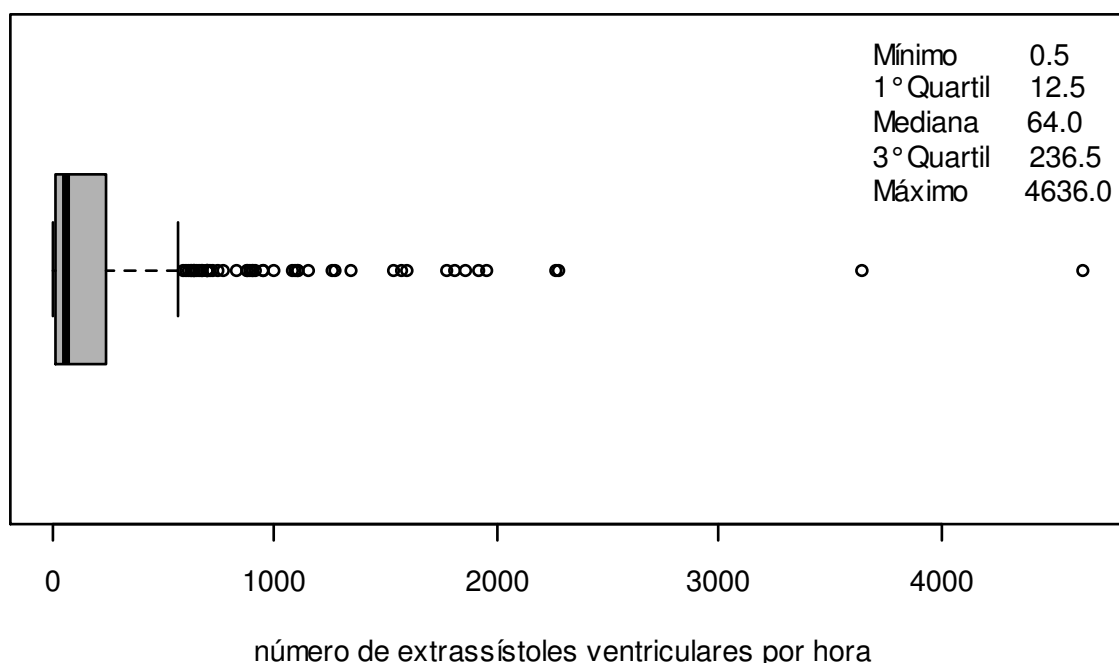
Figura A55: Box plot para número total de extrassístoles ventriculares.**Figura A56:** Box plot para número de extrassístoles ventriculares por hora.

Figura A57: Box plot para massa do ventrículo esquerdo (g).

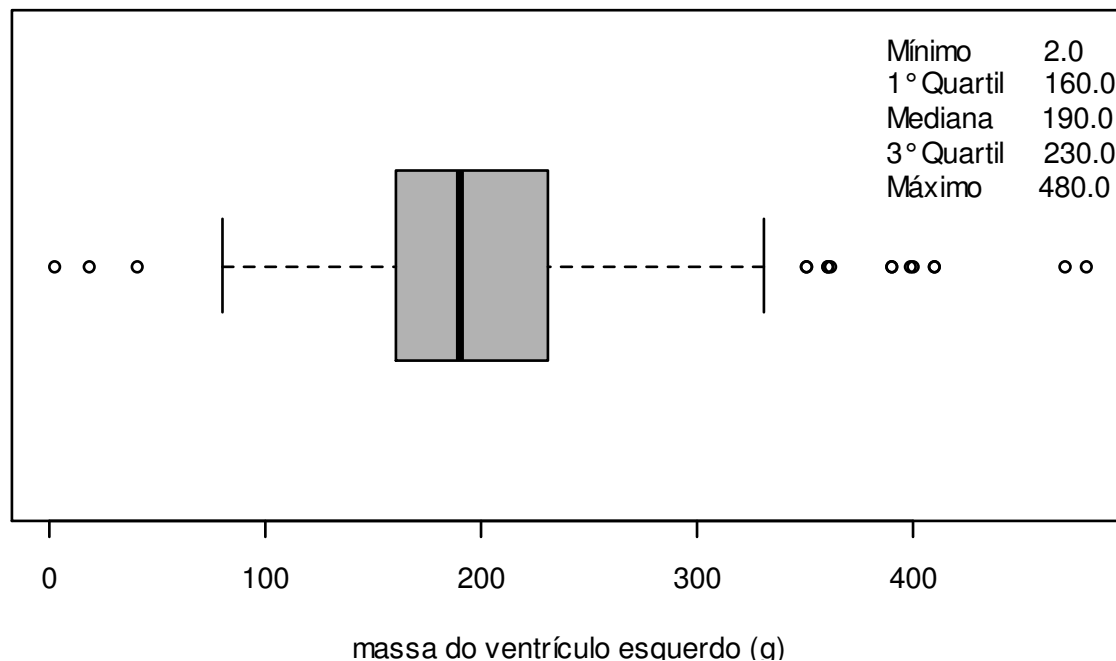


Figura A58: Box plot para glicemia (mg/dl).

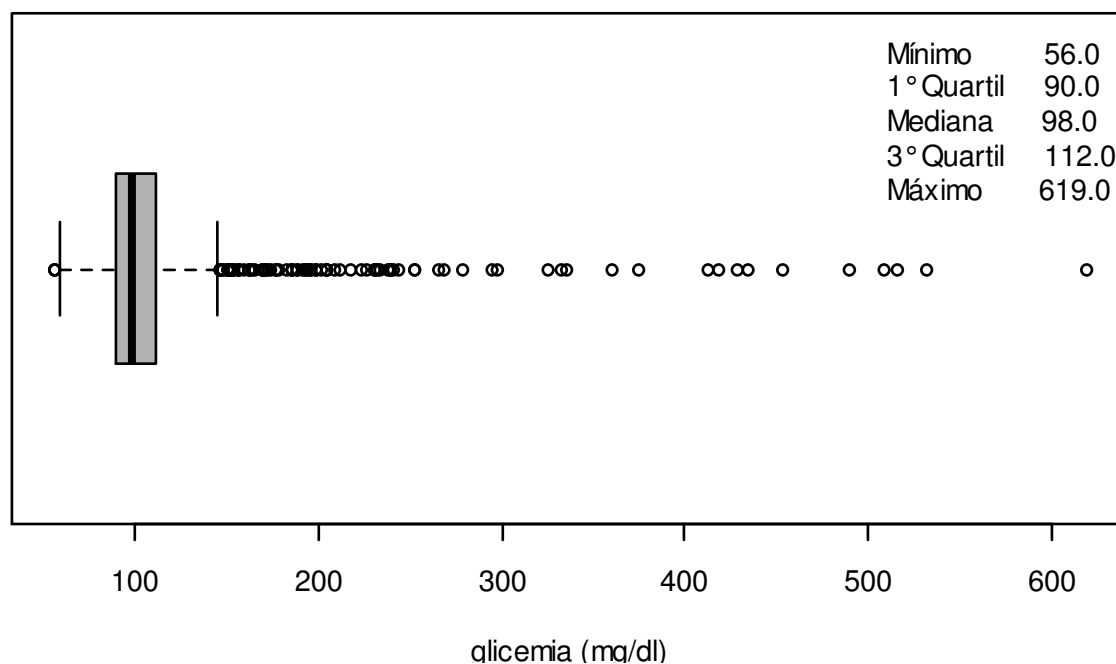


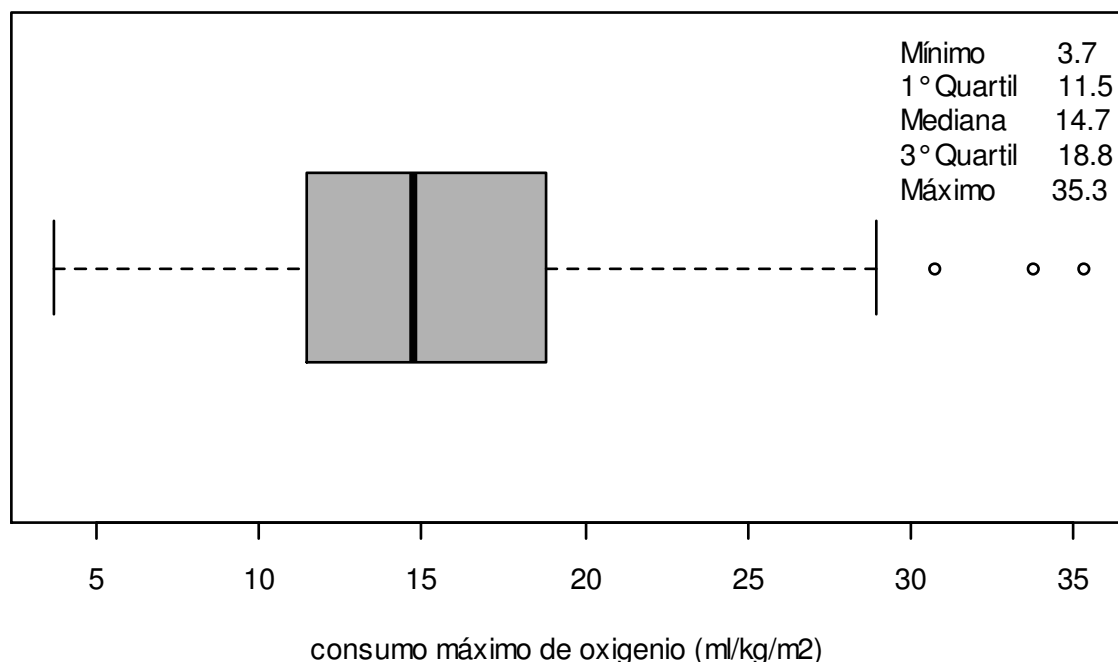
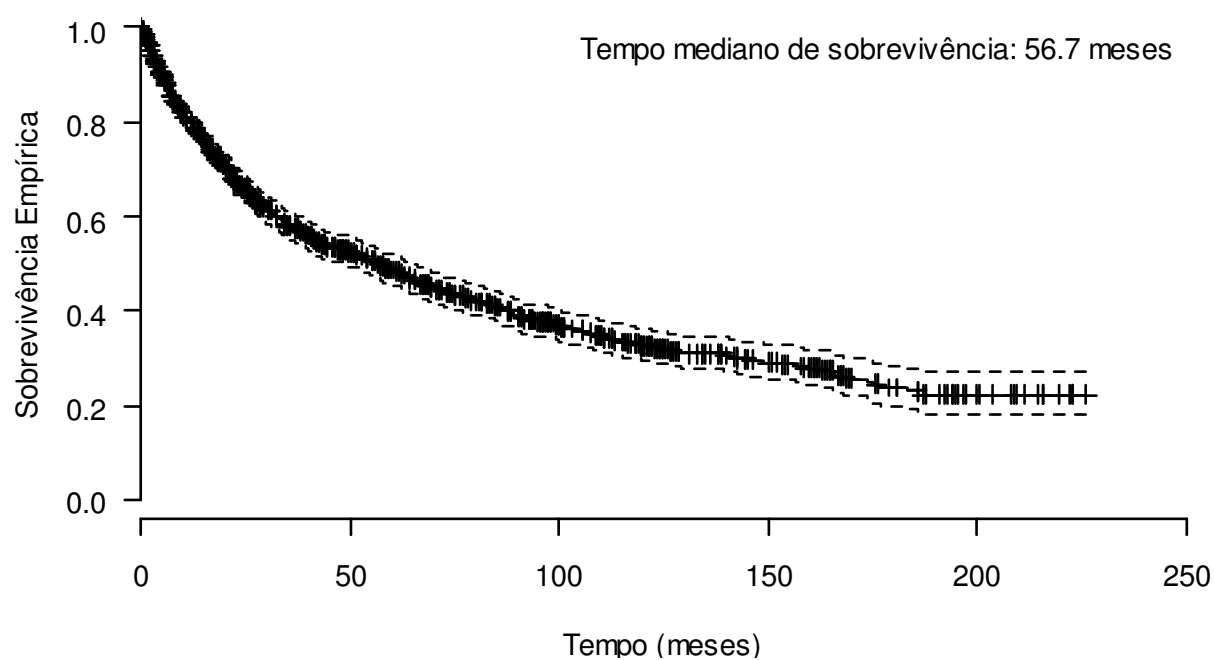
Figura A59: Box plot para consumo máximo de oxigênio (ml/kg/m^2).**Figura A60:** Curva de Kaplan-Meier da amostra.

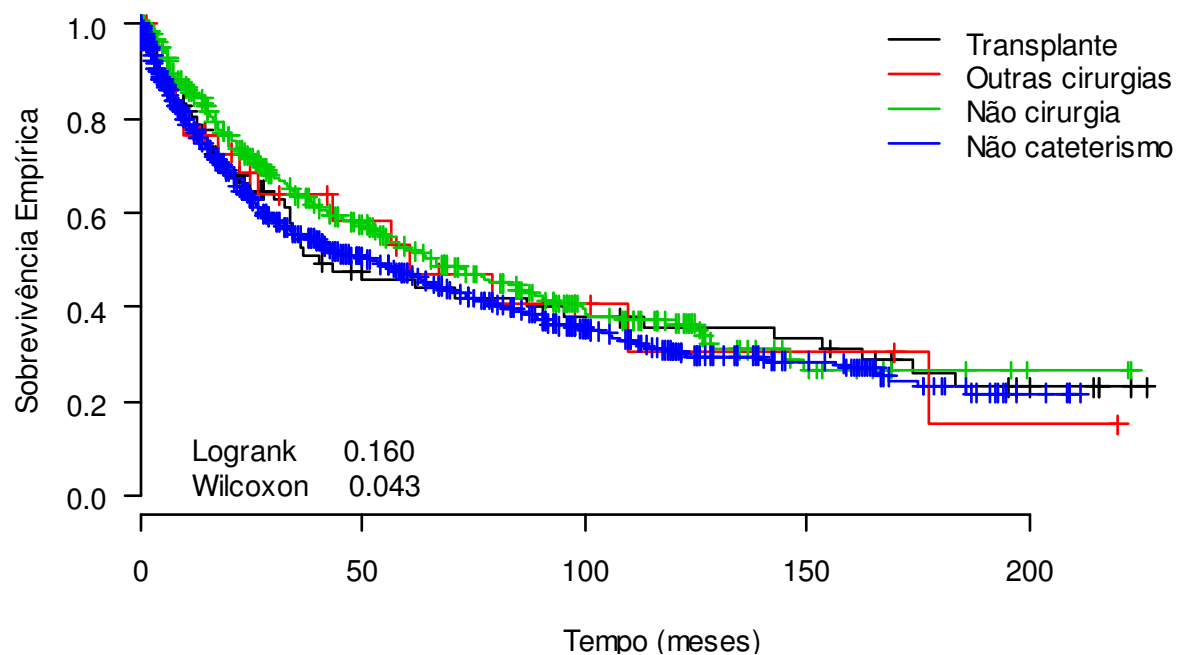
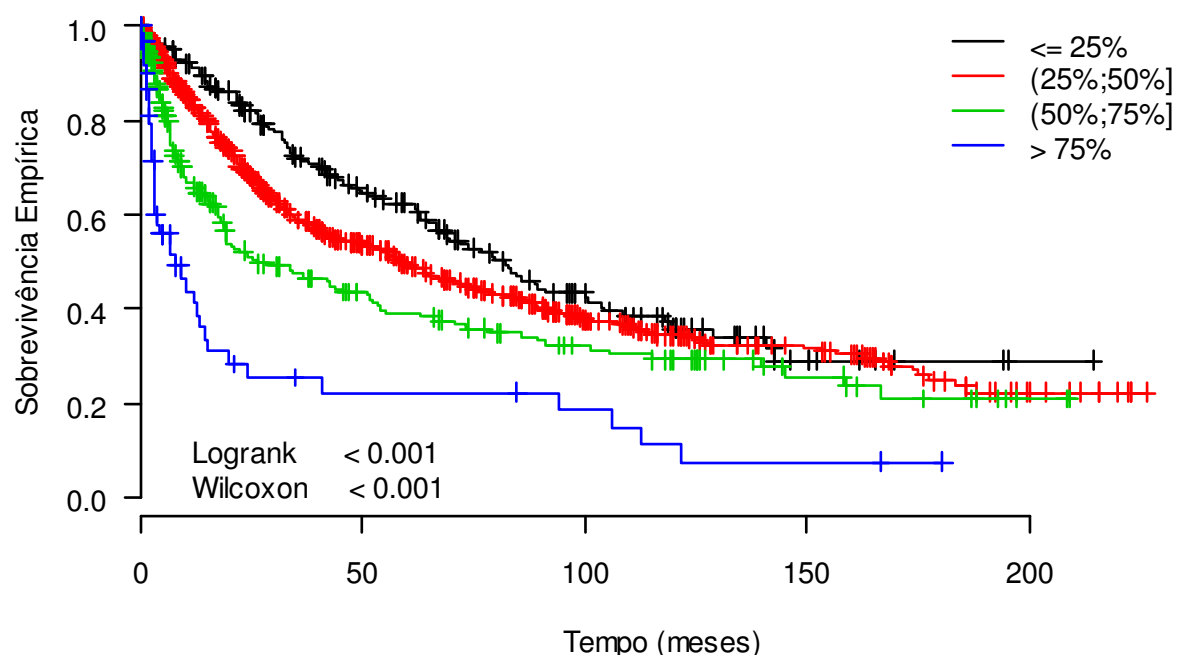
Figura A61: Curva de Kaplan-Meier por categoria de grupo.**Figura A62:** Curva de Kaplan-Meier por categoria do percentual de valores faltantes.

Figura A63: Curva de Kaplan-Meier por categoria de valores faltantes no ecocardiograma.

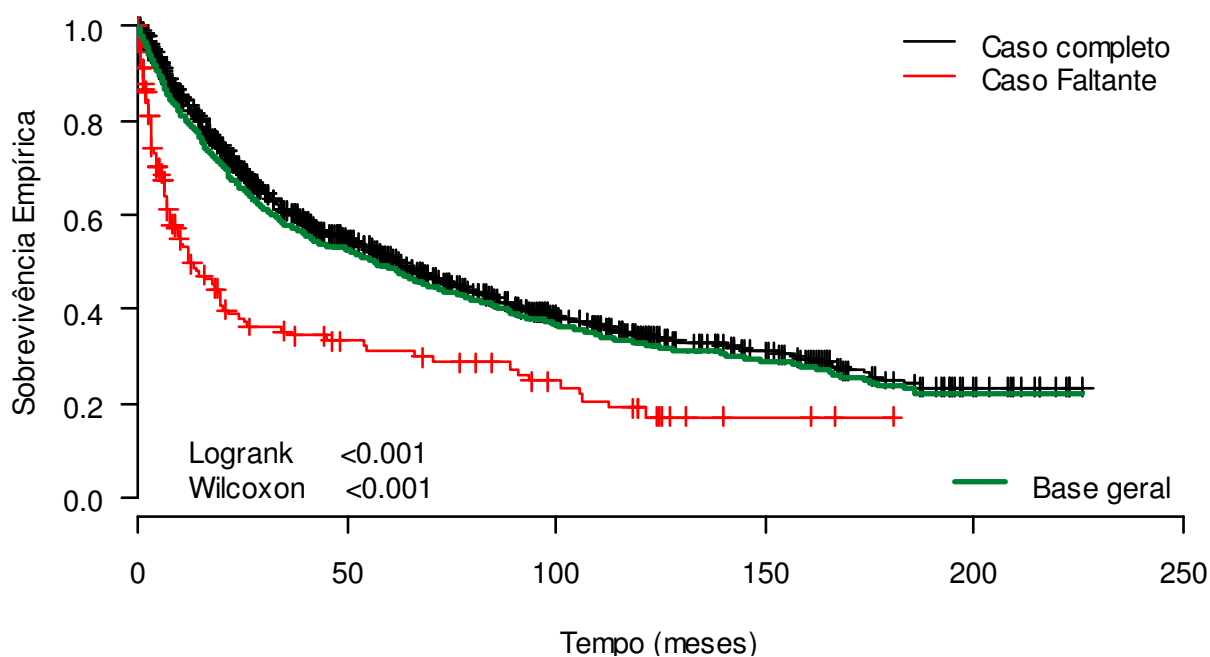


Figura A64: Curva de Kaplan-Meier por categoria de valores faltantes no eletrocardiograma.

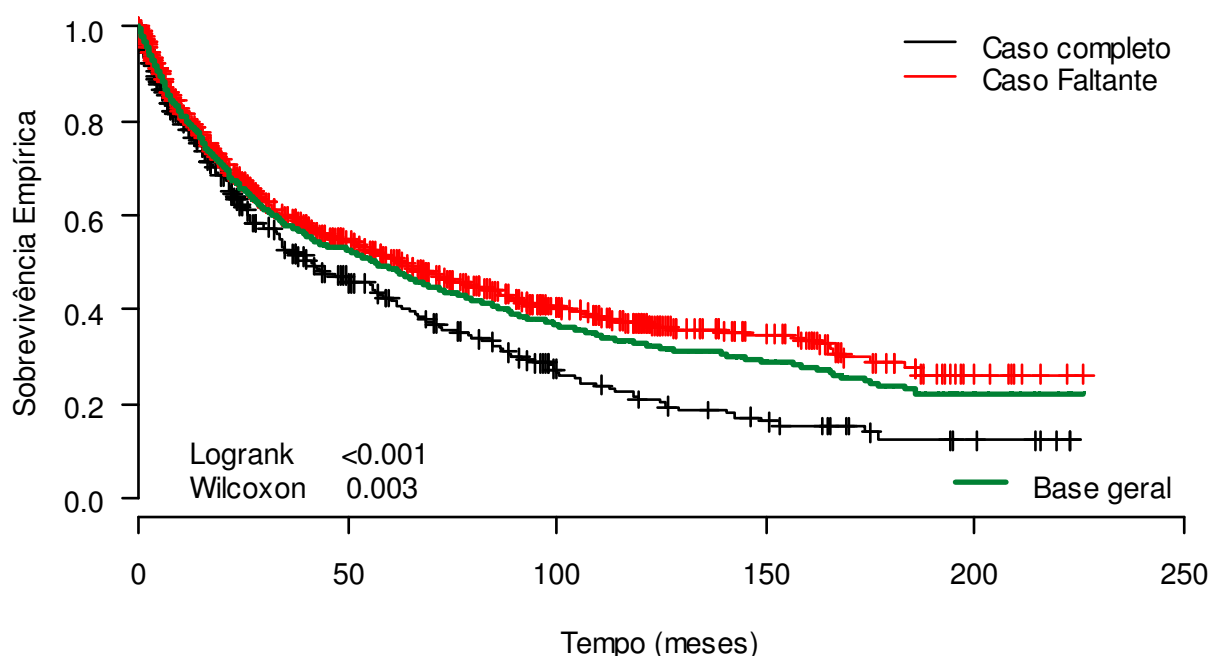


Figura A65: Curva de Kaplan-Meier por categoria de valores faltantes no leucograma.

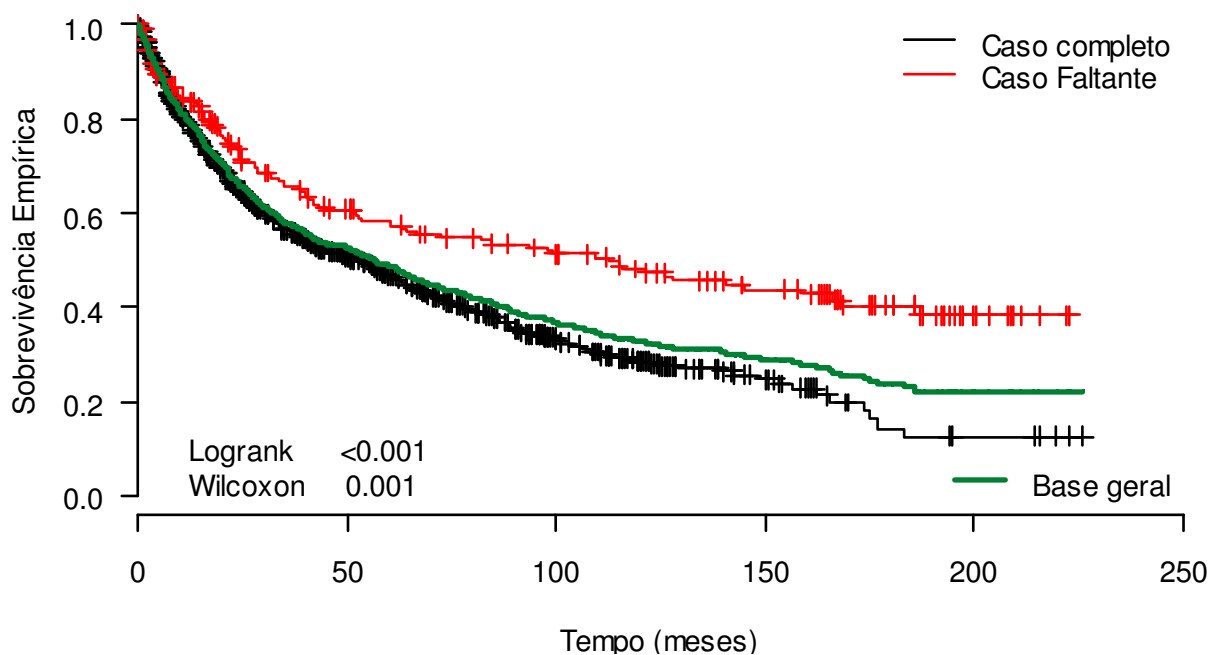


Figura A66: Curva de Kaplan-Meier por categoria de valores faltantes no perfil lipídico.

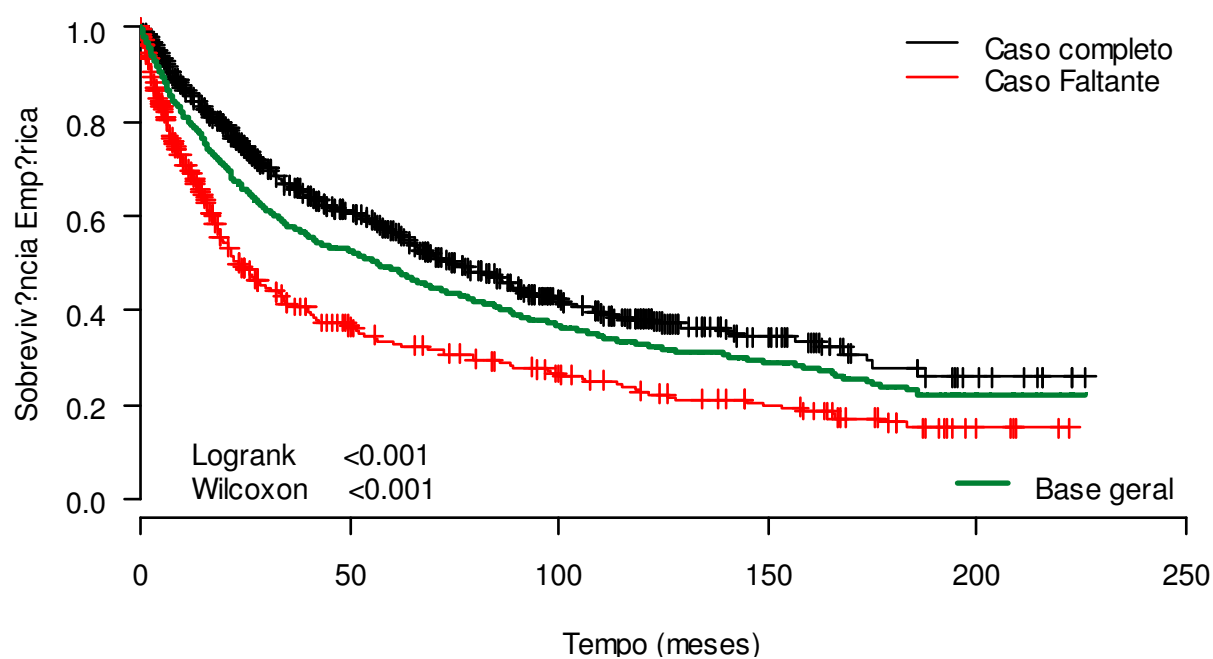


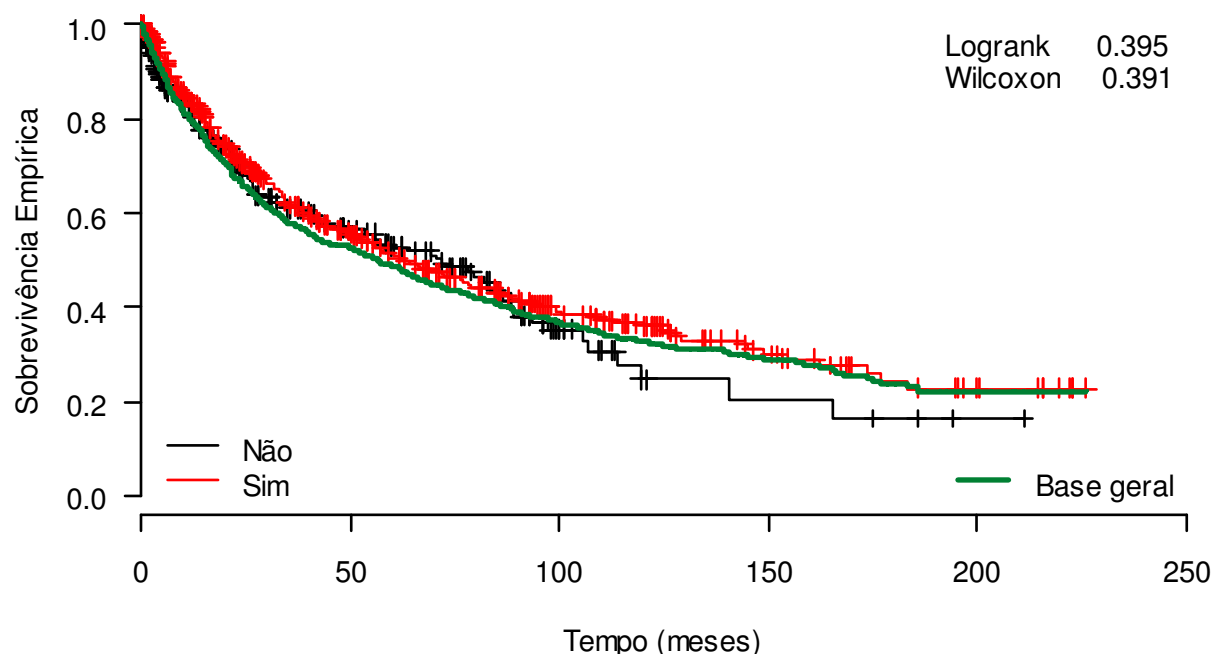
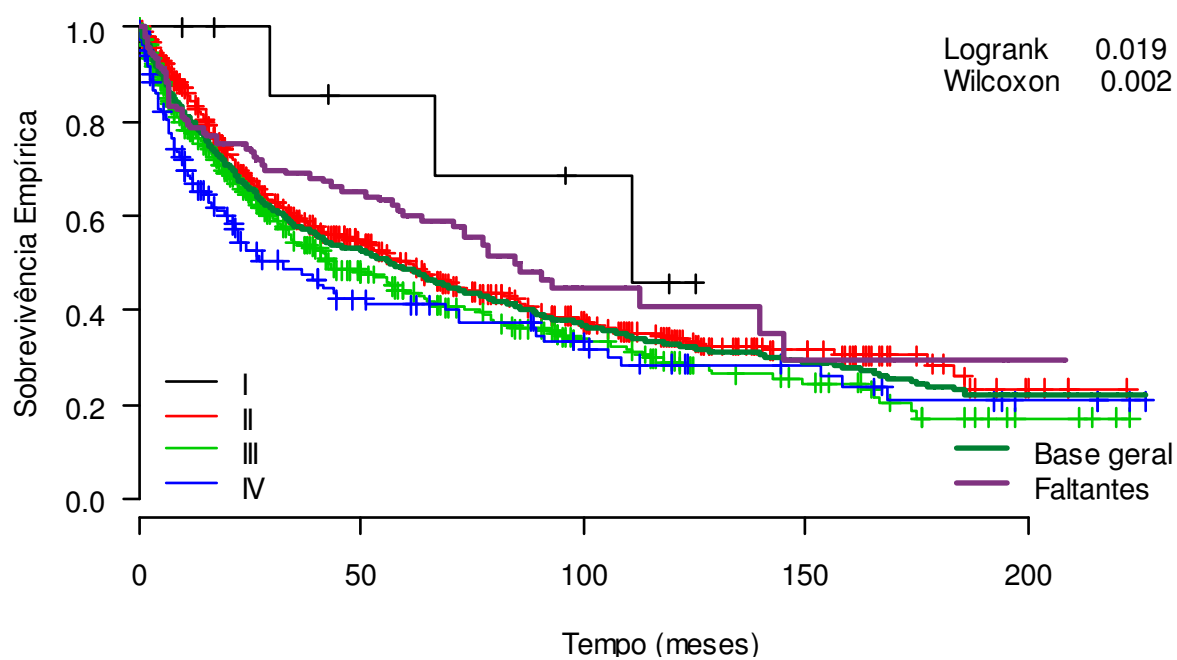
Figura A67: Curva de Kaplan-Meier por categoria de valores faltantes no cateterismo.**Figura A68:** Curva de Kaplan-Meier por categoria de classe funcional.

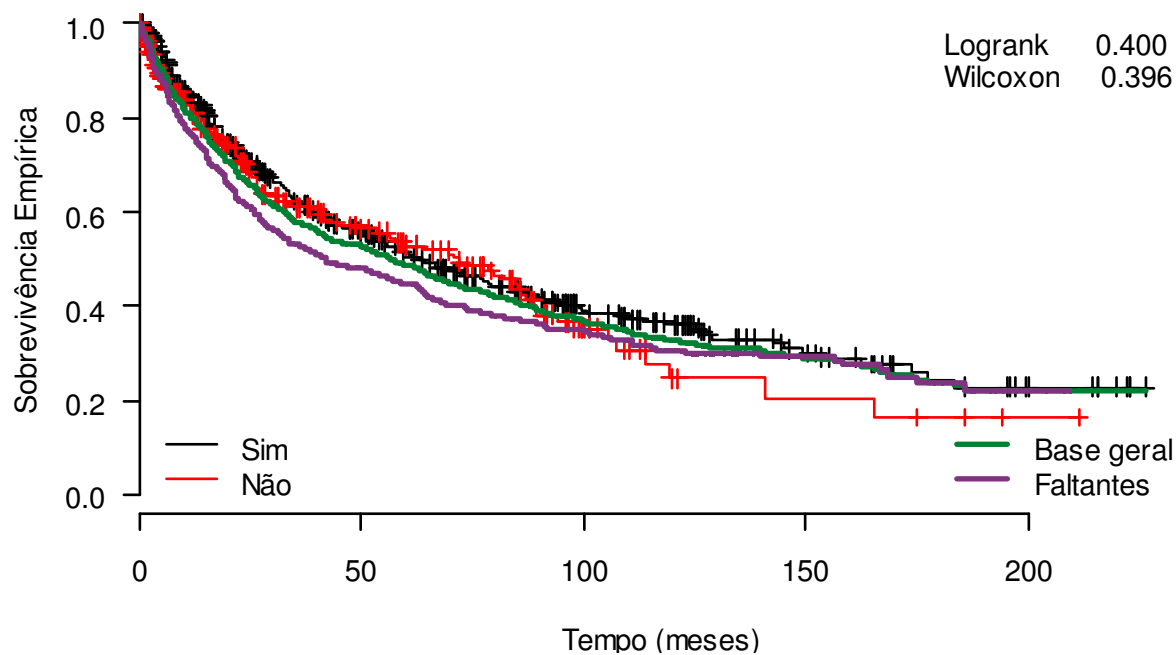
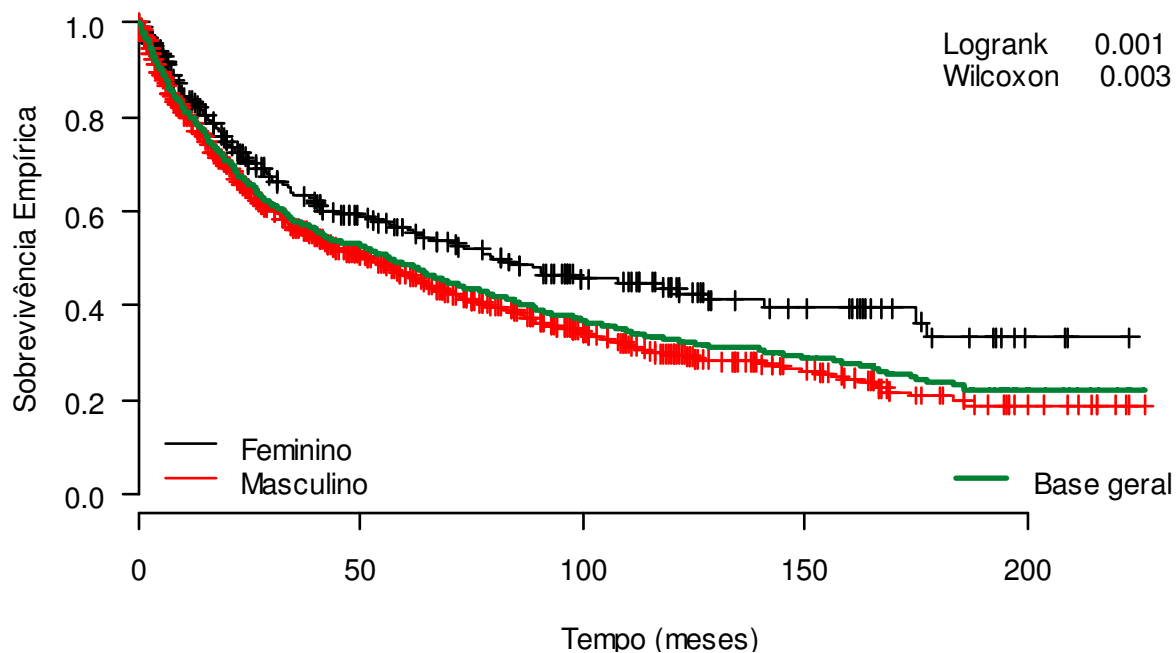
Figura A69: Curva de Kaplan-Meier por categoria de indicador de cateterismo.**Figura A70:** Curva de Kaplan-Meier por categoria de gênero.

Figura A71: Curva de Kaplan-Meier por categoria de etnia.

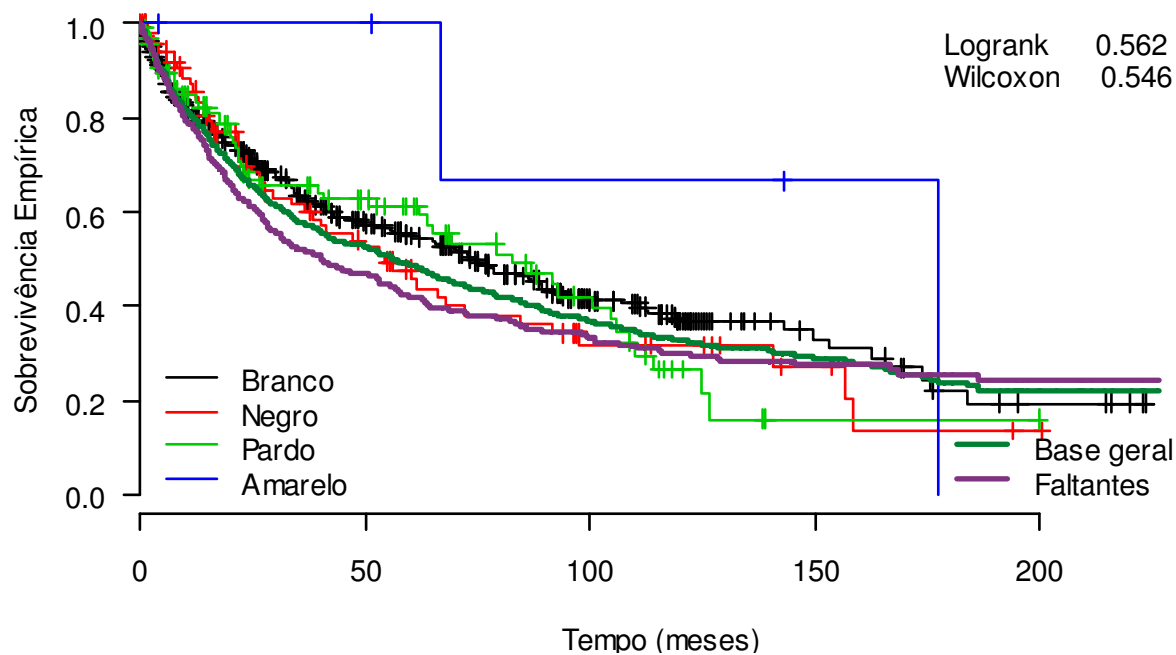


Figura A72: Curva de Kaplan-Meier por categoria de tabagismo.

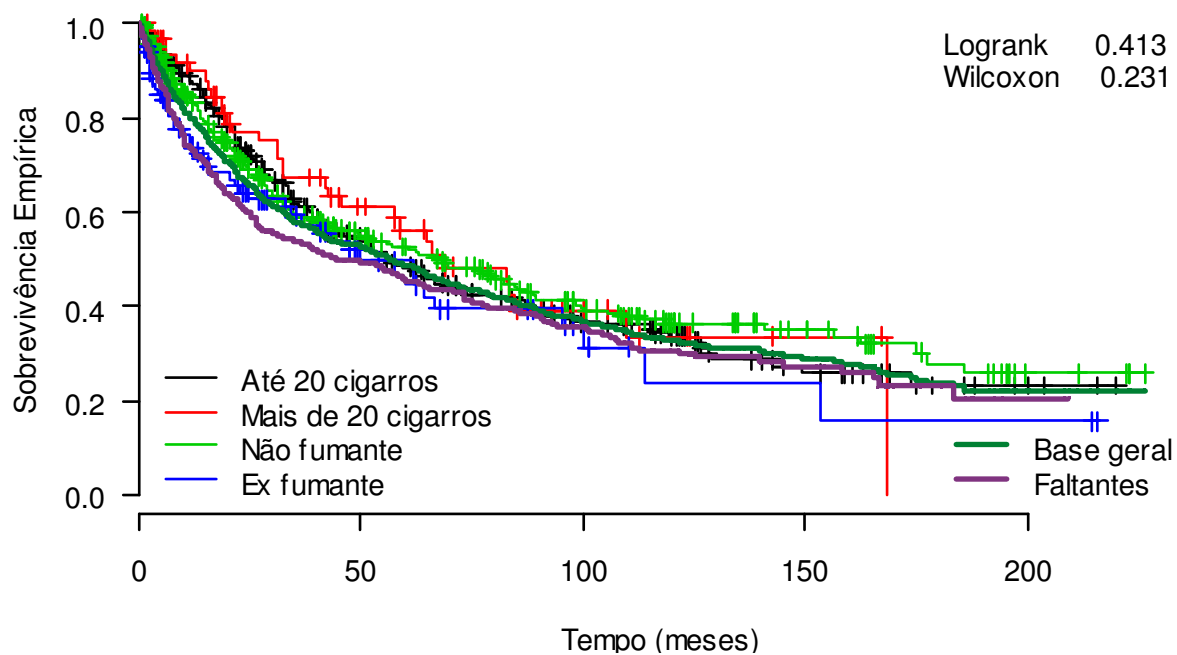


Figura A73: Curva de Kaplan-Meier por categoria de etilismo.

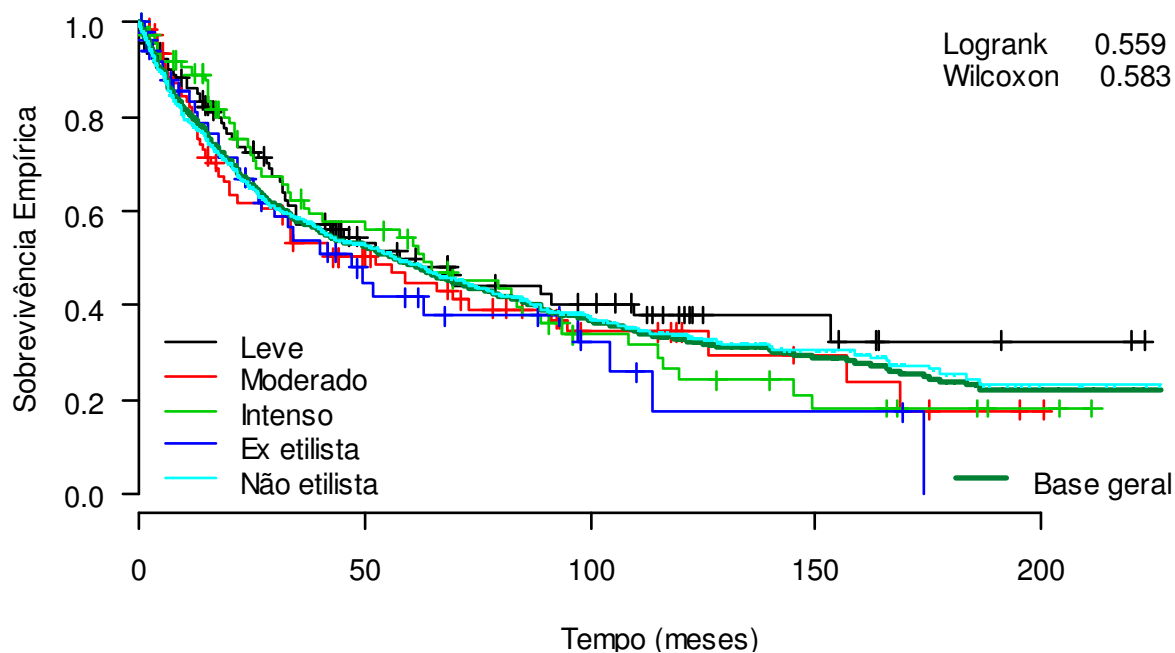


Figura A74: Curva de Kaplan-Meier por categoria de diabetes melito.

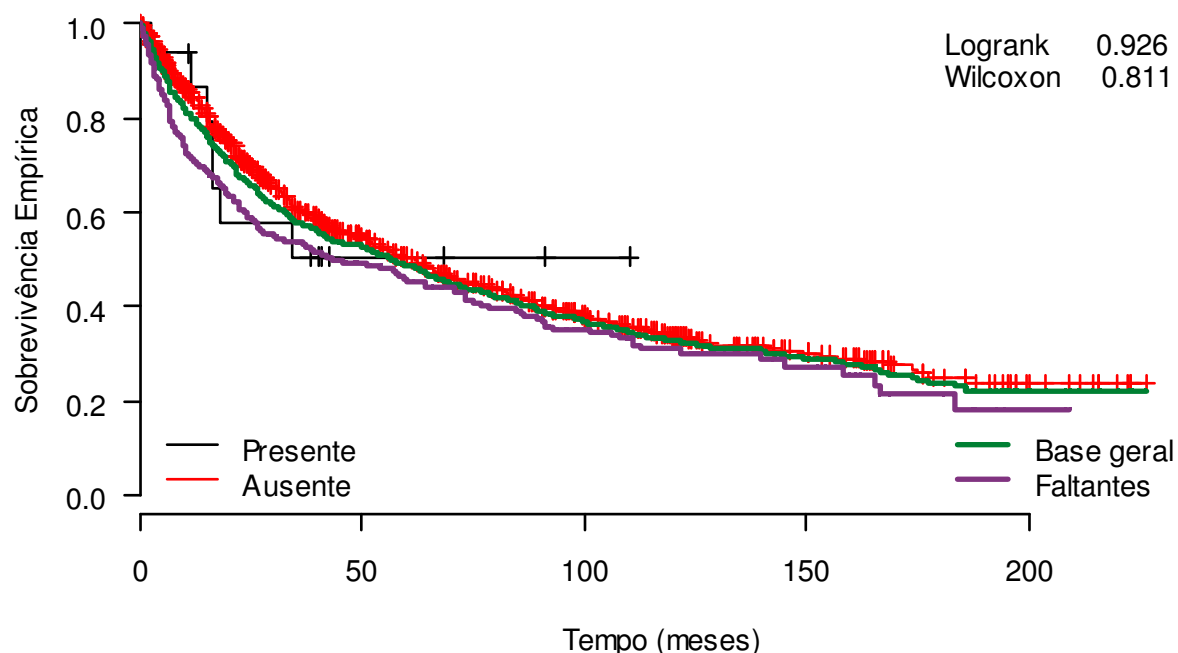


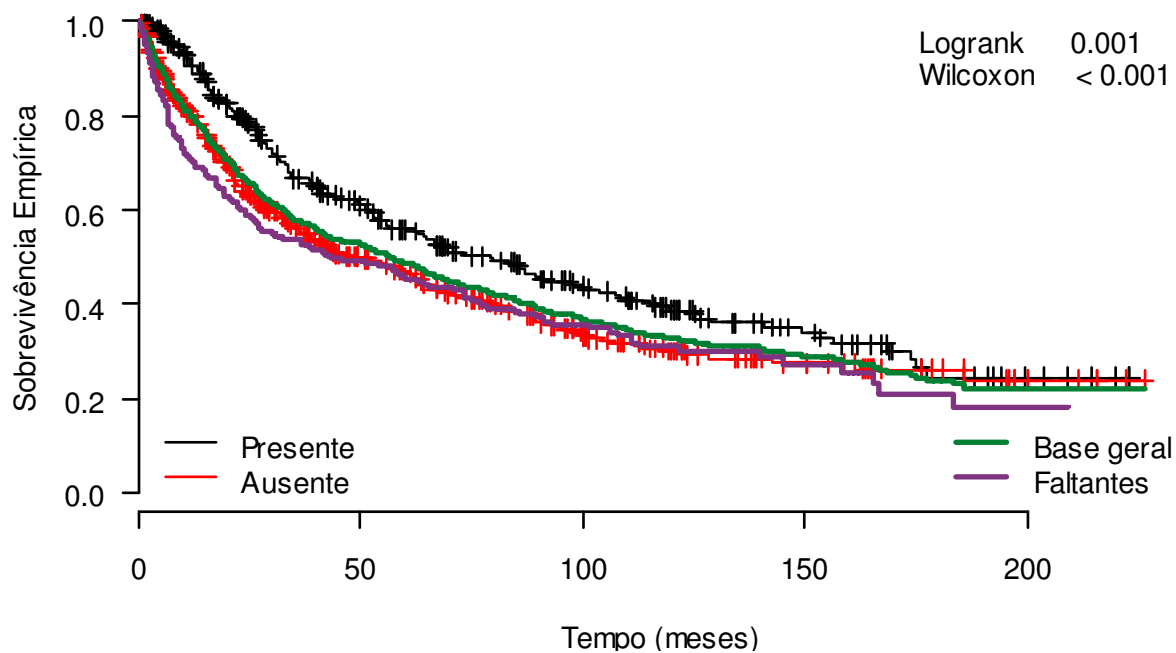
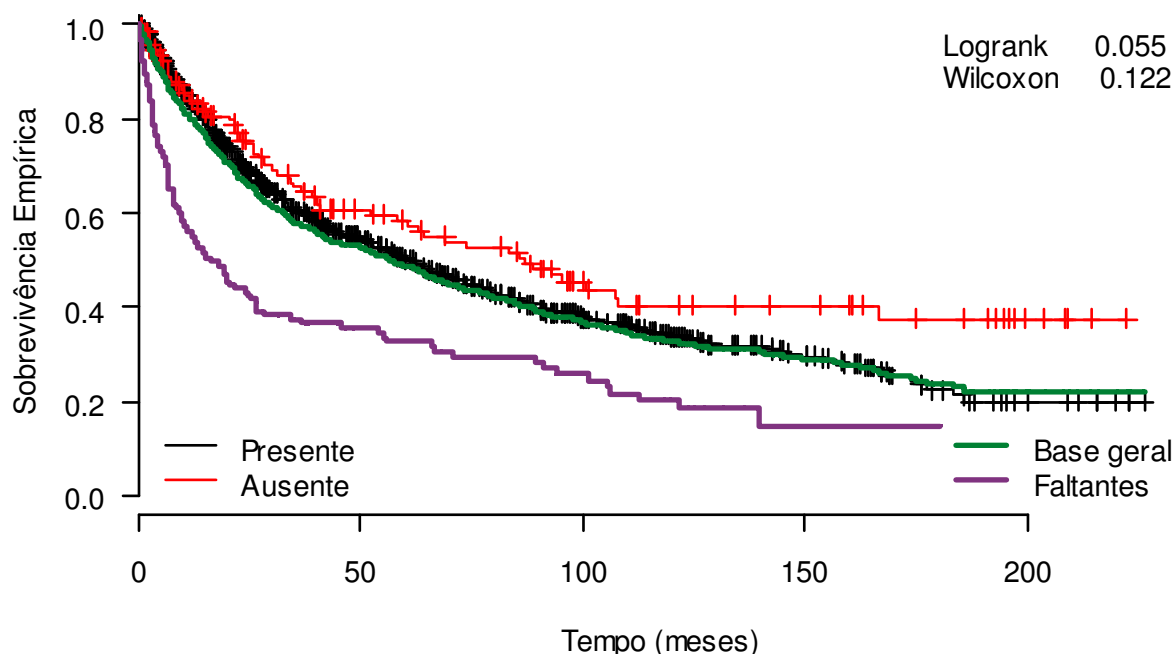
Figura A75: Curva de Kaplan-Meier por categoria de hipertensão arterial.**Figura A76:** Curva de Kaplan-Meier por categoria de insuficiência mitral Doppler.

Figura A77: Curva de Kaplan-Meier por categoria de insuficiência tricúspide Doppler.

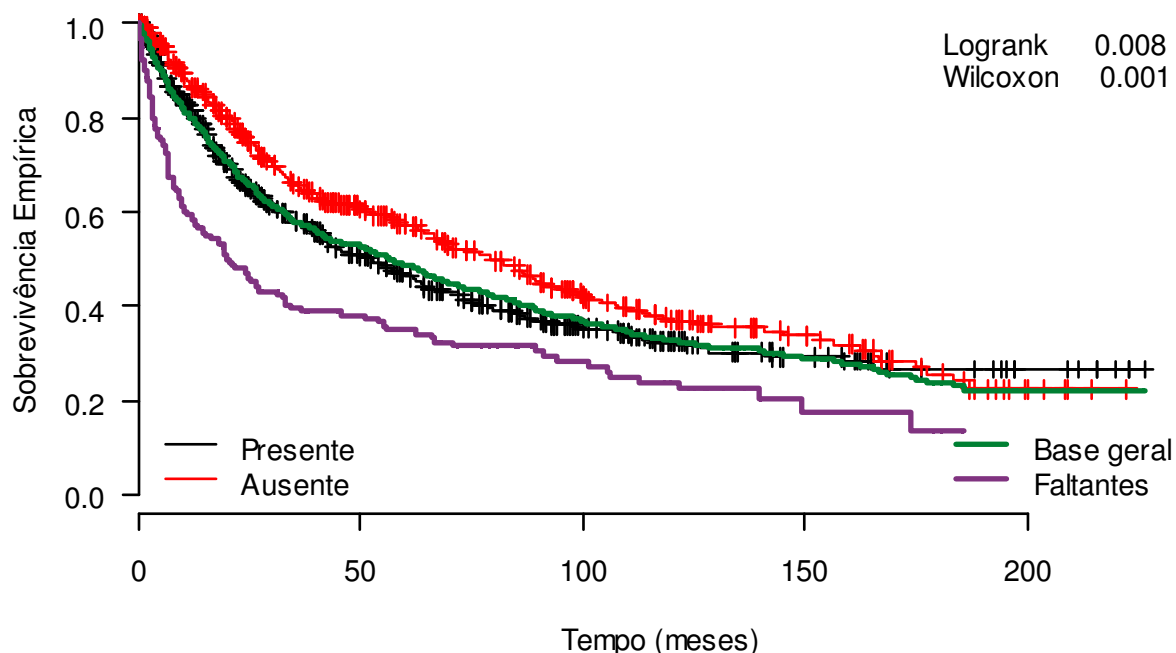


Figura A78: Curva de Kaplan-Meier por categoria de diagnóstico etiológico.

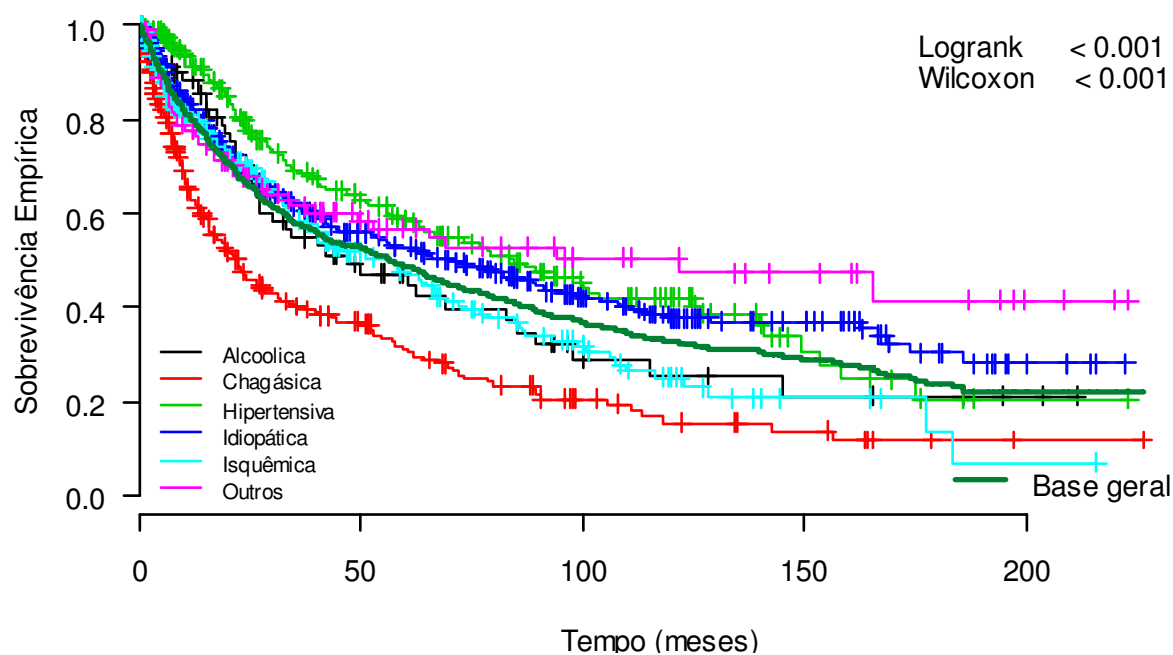


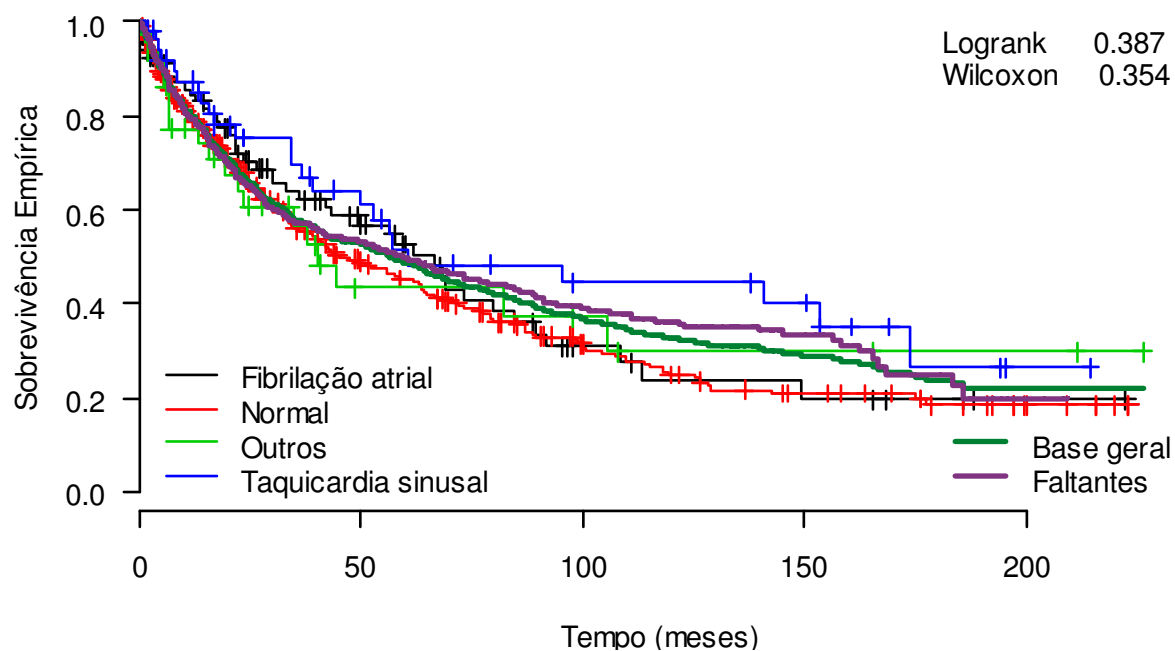
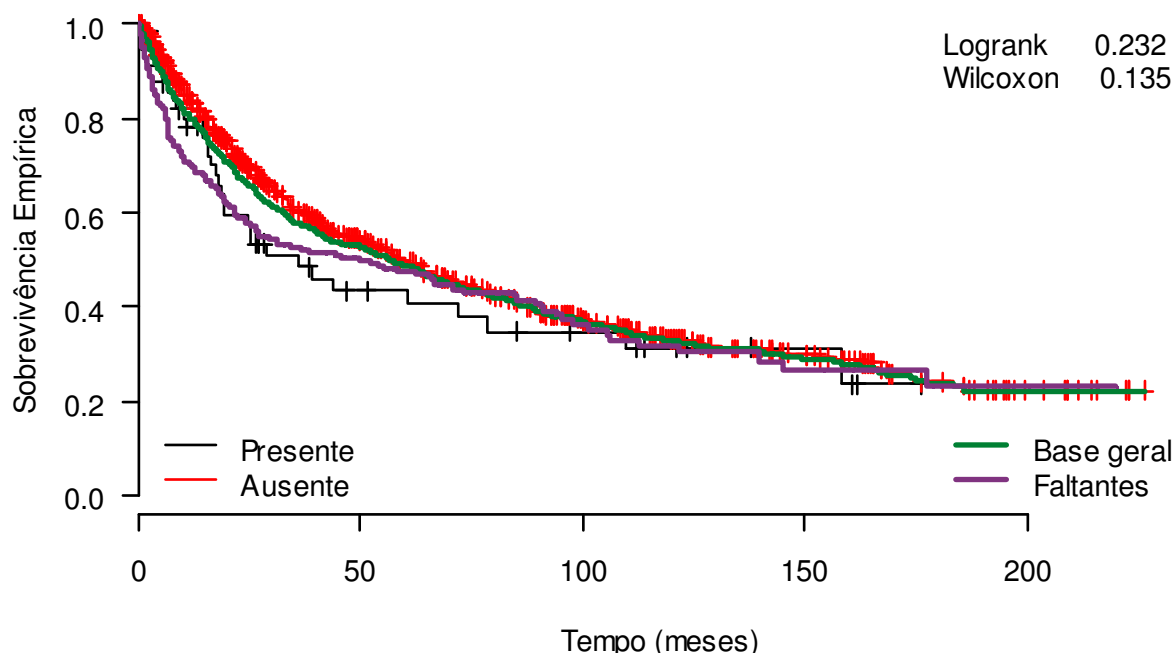
Figura A79: Curva de Kaplan-Meier por categoria de ritmo.**Figura A80:** Curva de Kaplan-Meier por categoria de trombo no ventrículo esquerdo.

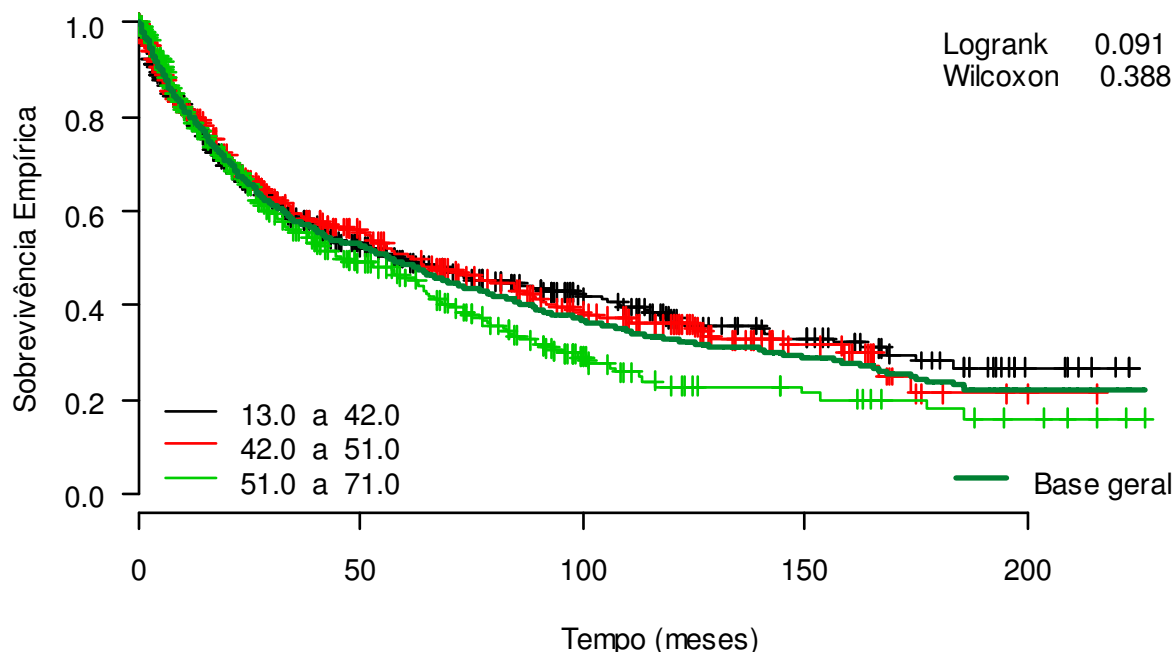
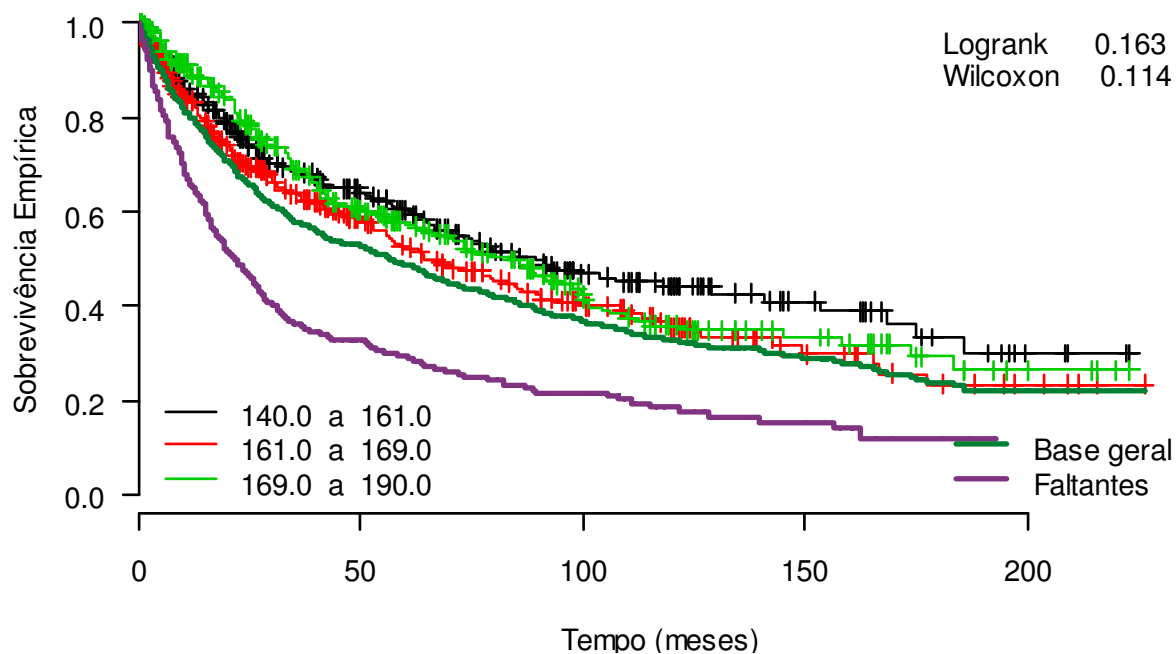
Figura A81: Curva de Kaplan-Meier por categoria de idade (anos).**Figura A82:** Curva de Kaplan-Meier por categoria de altura (cm).

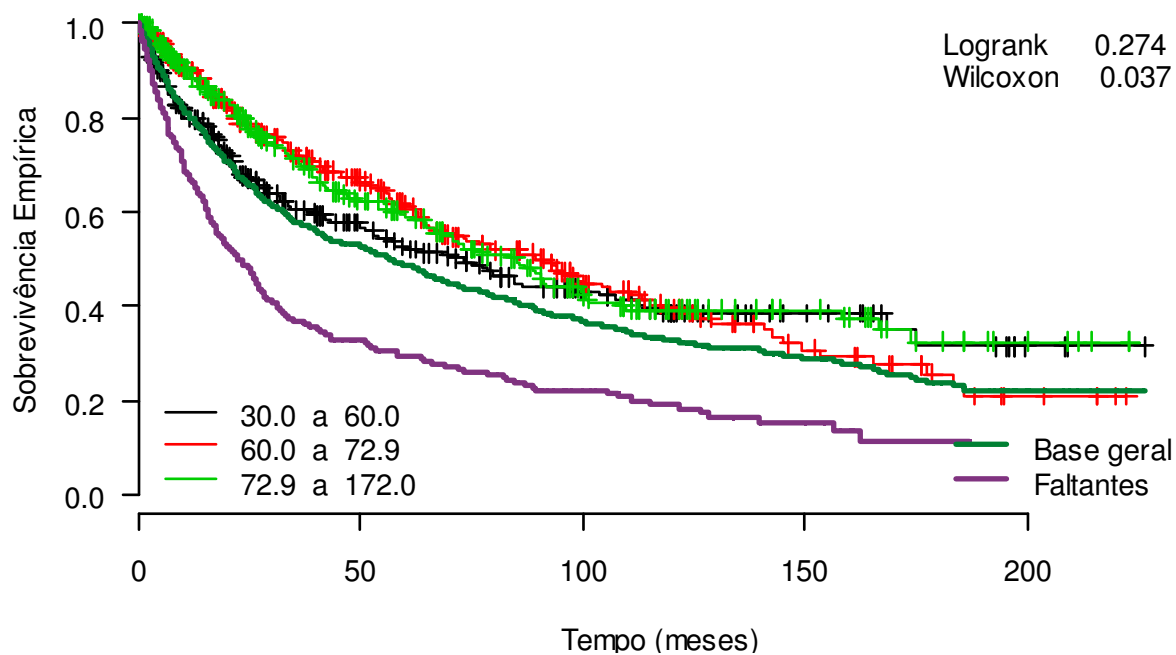
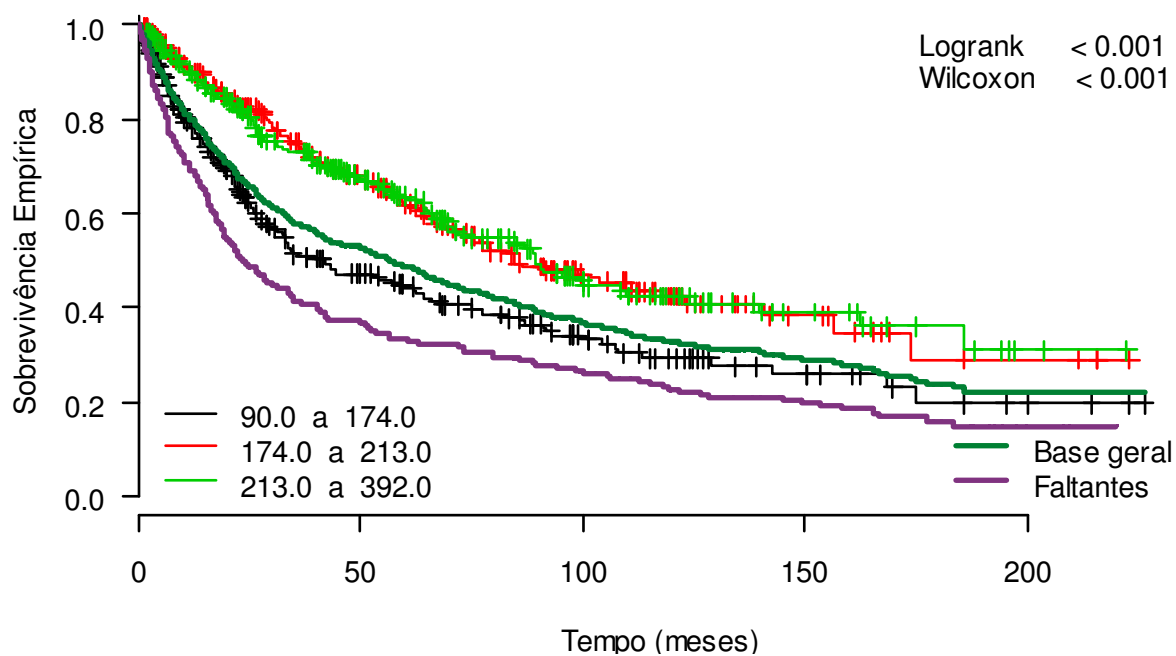
Figura A83: Curva de Kaplan-Meier por categoria de peso (kg).**Figura A84:** Curva de Kaplan-Meier por categoria de colesterol total (mg/dl).

Figura A85: Curva de Kaplan-Meier por categoria de fração HDL do colesterol (mg/dl).

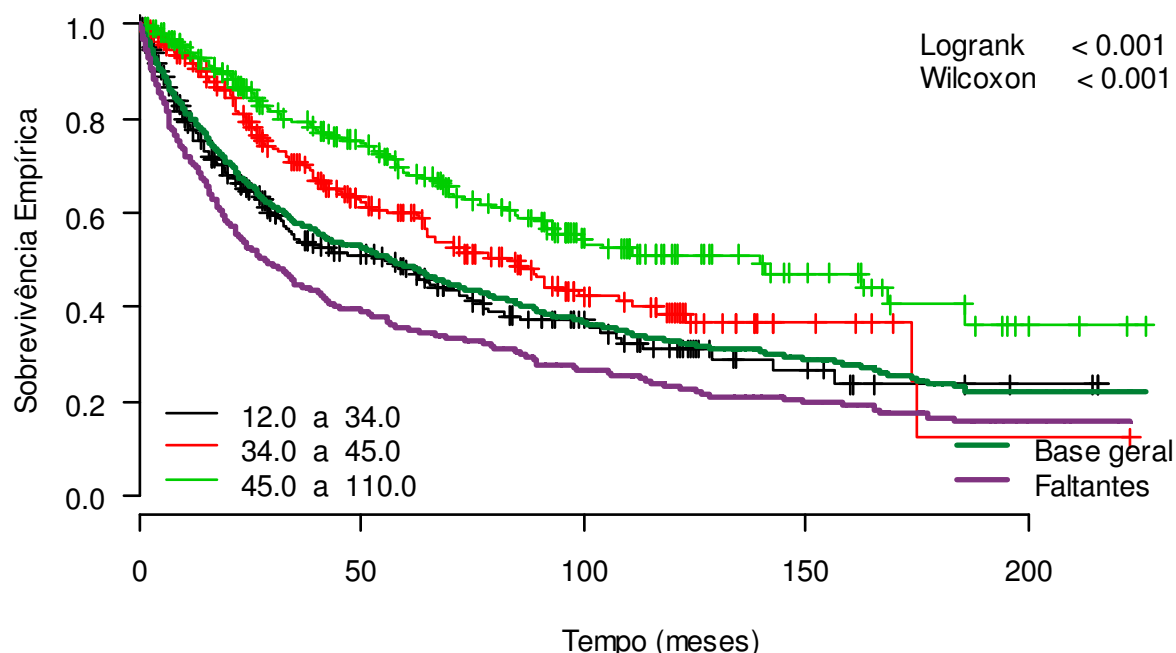


Figura A86: Curva de Kaplan-Meier por categoria de fração LDL do colesterol (mg/dl).

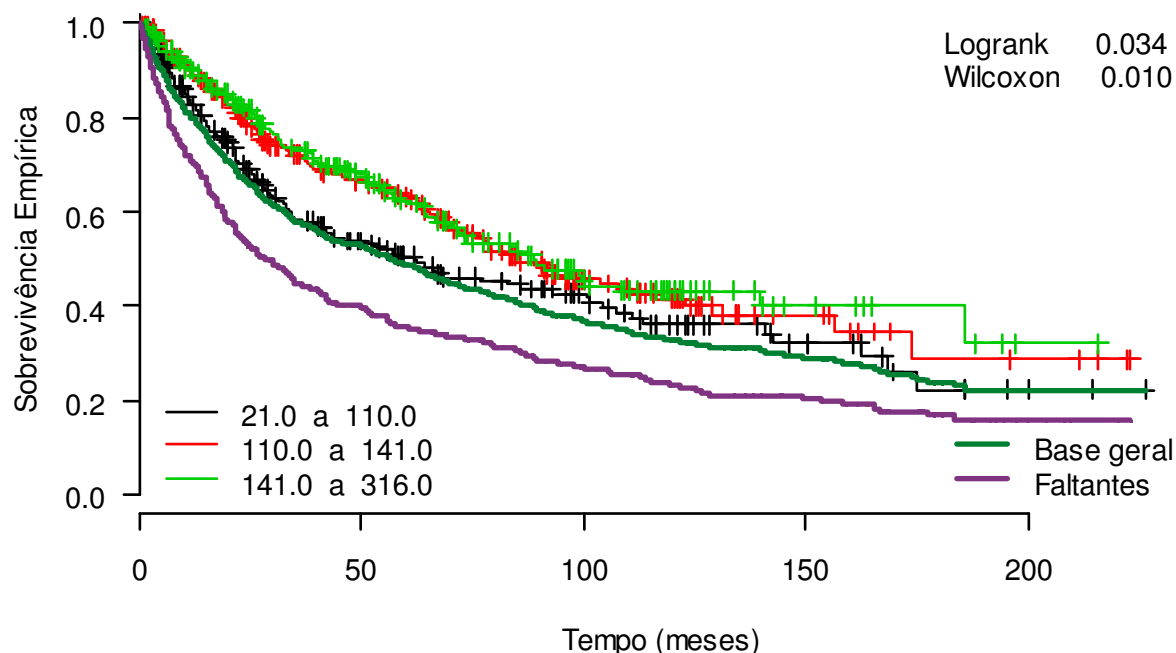


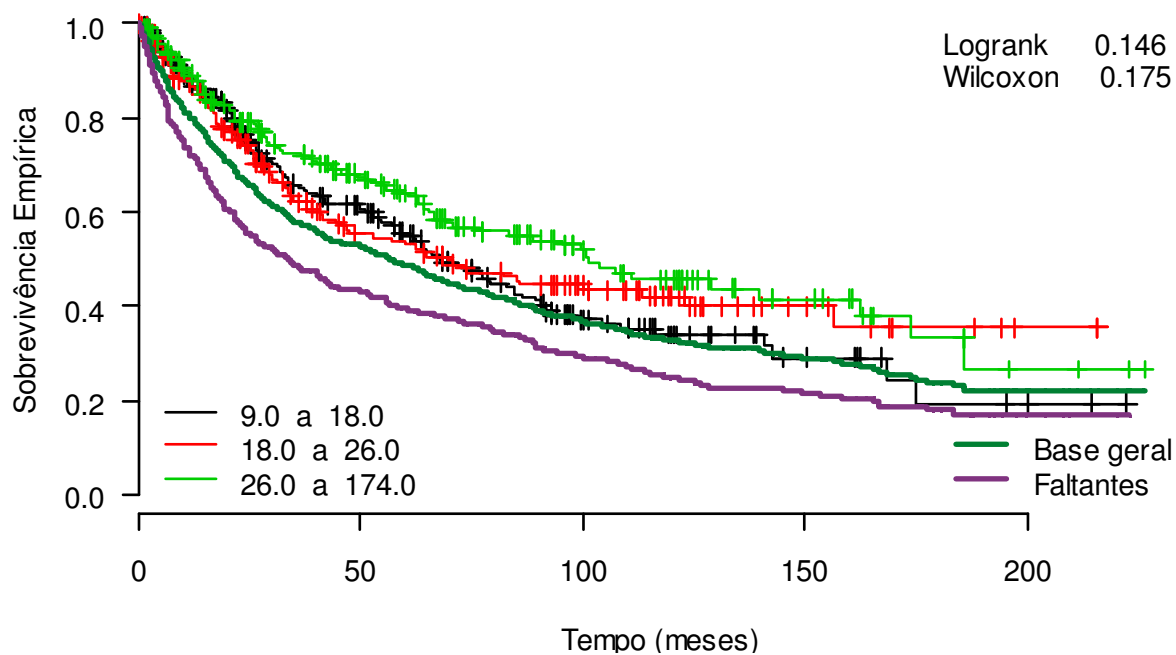
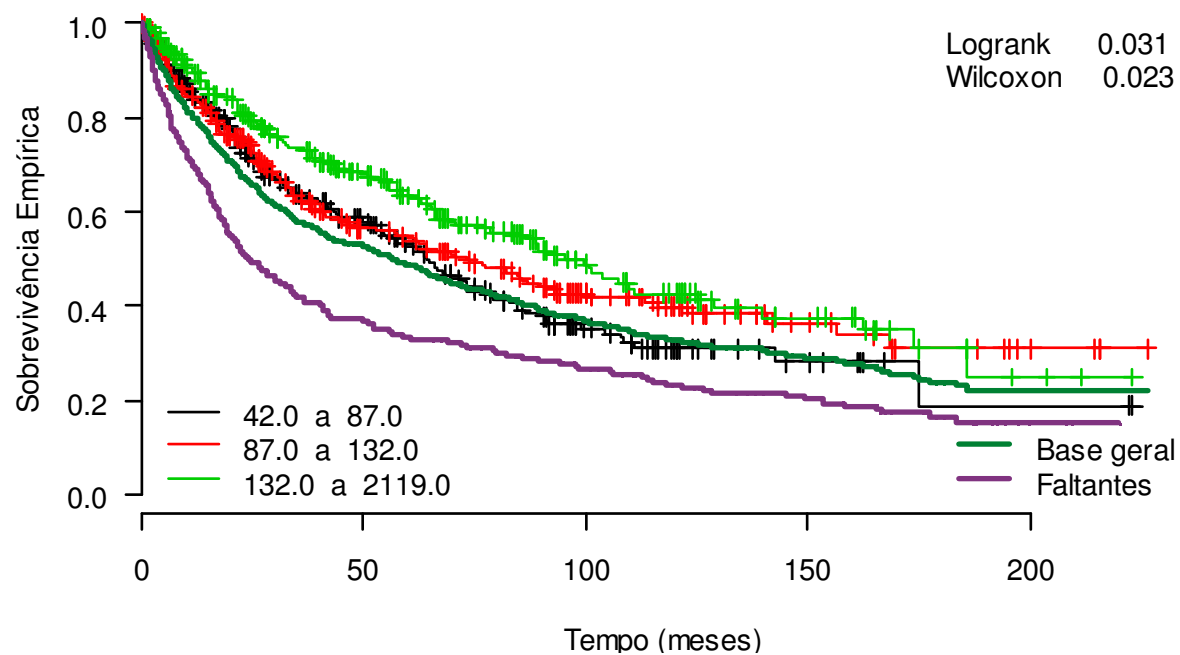
Figura A87: Curva de Kaplan-Meier por categoria de fração VLDL do colesterol (mg/dl).**Figura A88:** Curva de Kaplan-Meier por categoria de triglicérides (mg/dl).

Figura A89: Curva de Kaplan-Meier por categoria do número absoluto de leucócitos (/mm³).

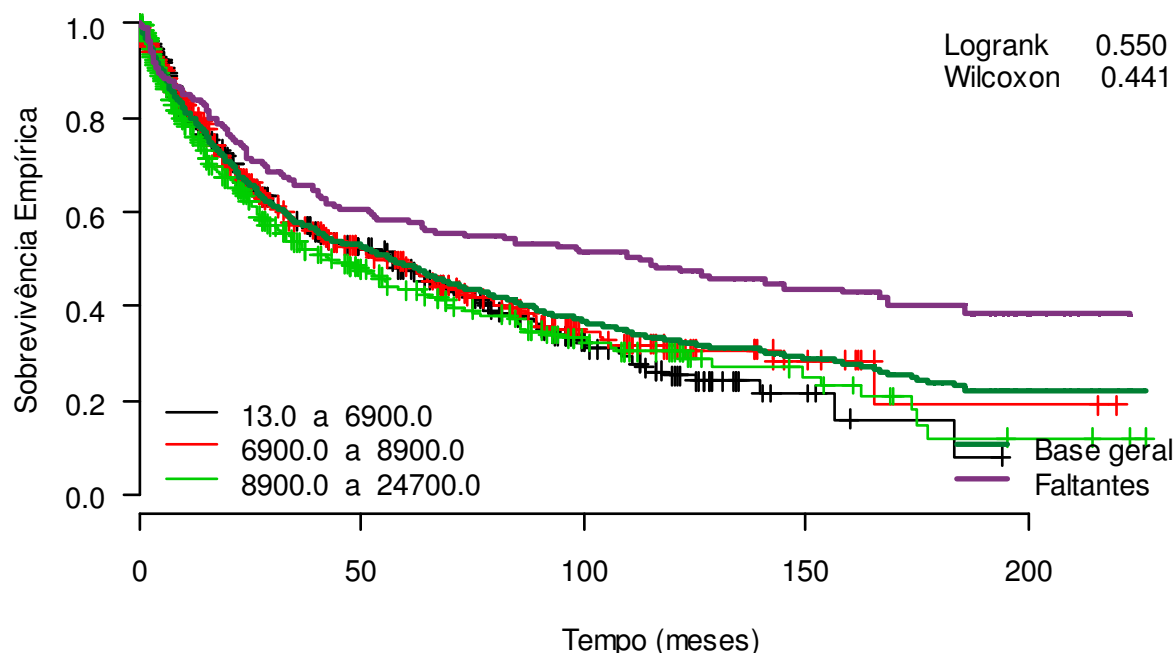


Figura A90: Curva de Kaplan-Meier por categoria de linfócitos (%).

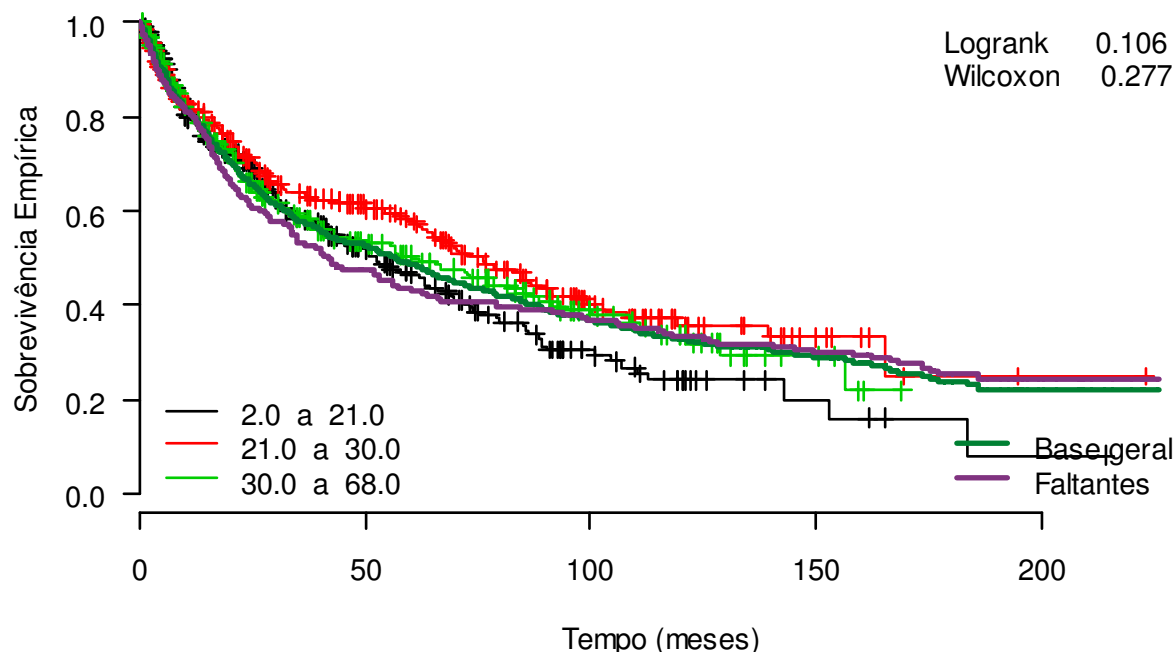


Figura A91: Curva de Kaplan-Meier por categoria de basófilos (%).

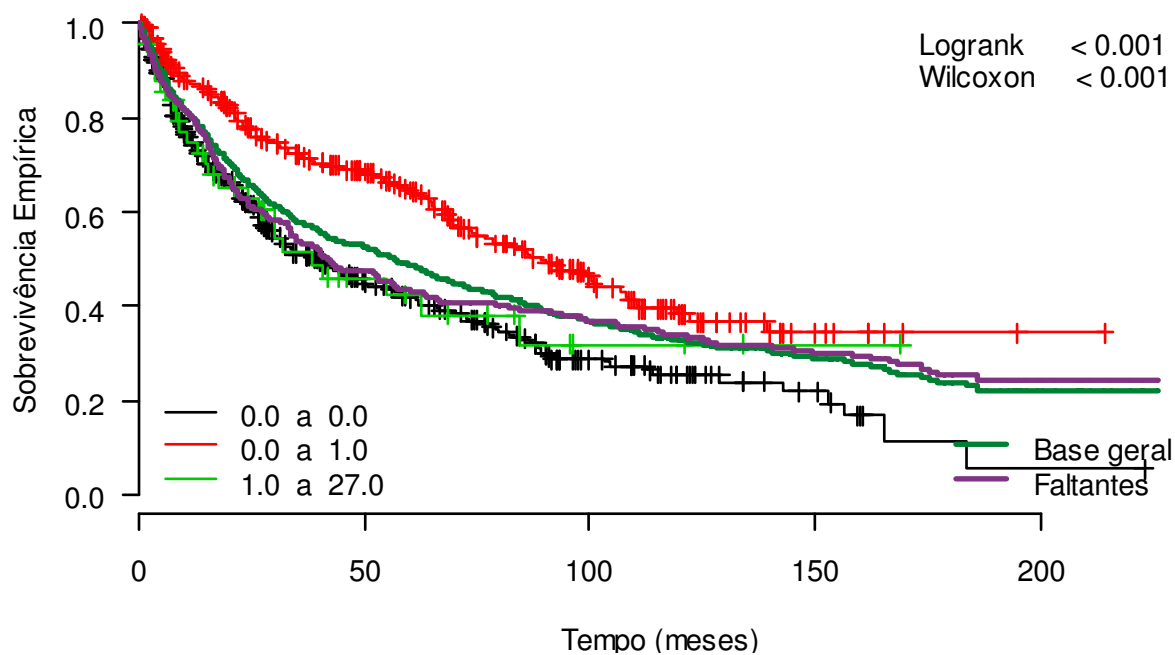


Figura A92: Curva de Kaplan-Meier por categoria de eosinófilos (%).

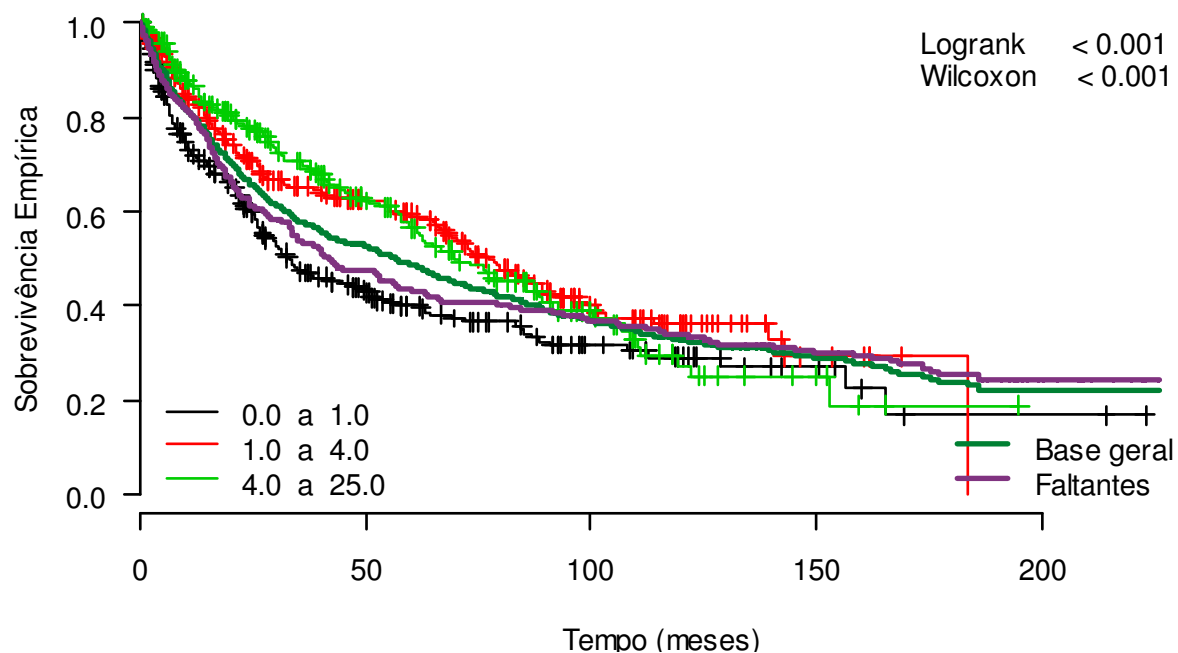


Figura A93: Curva de Kaplan-Meier por categoria de neutrófilos (%).

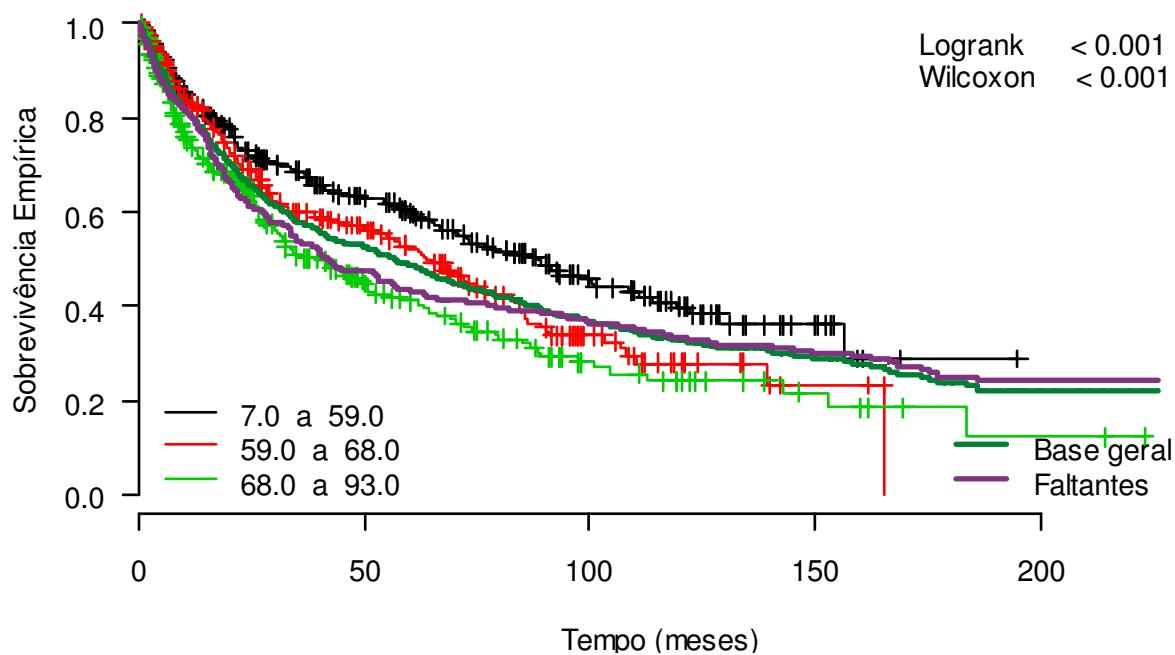


Figura A94: Curva de Kaplan-Meier por categoria de monócitos (%).

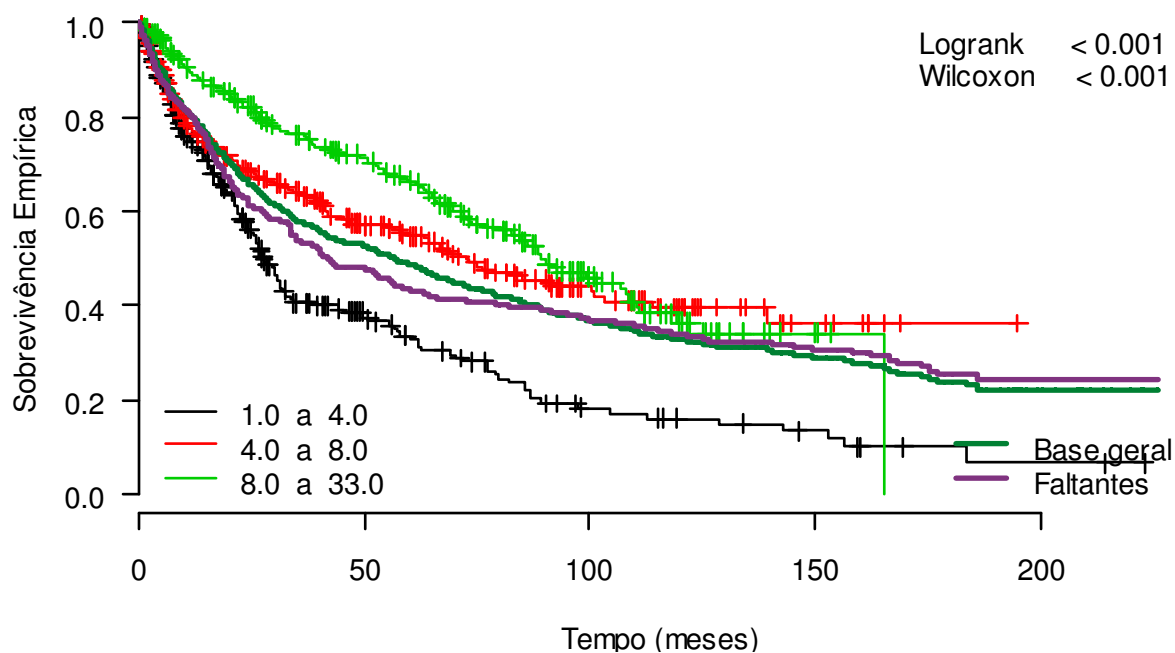


Figura A95: Curva de Kaplan-Meier por categoria do número absoluto de plaquetas (/mm³).

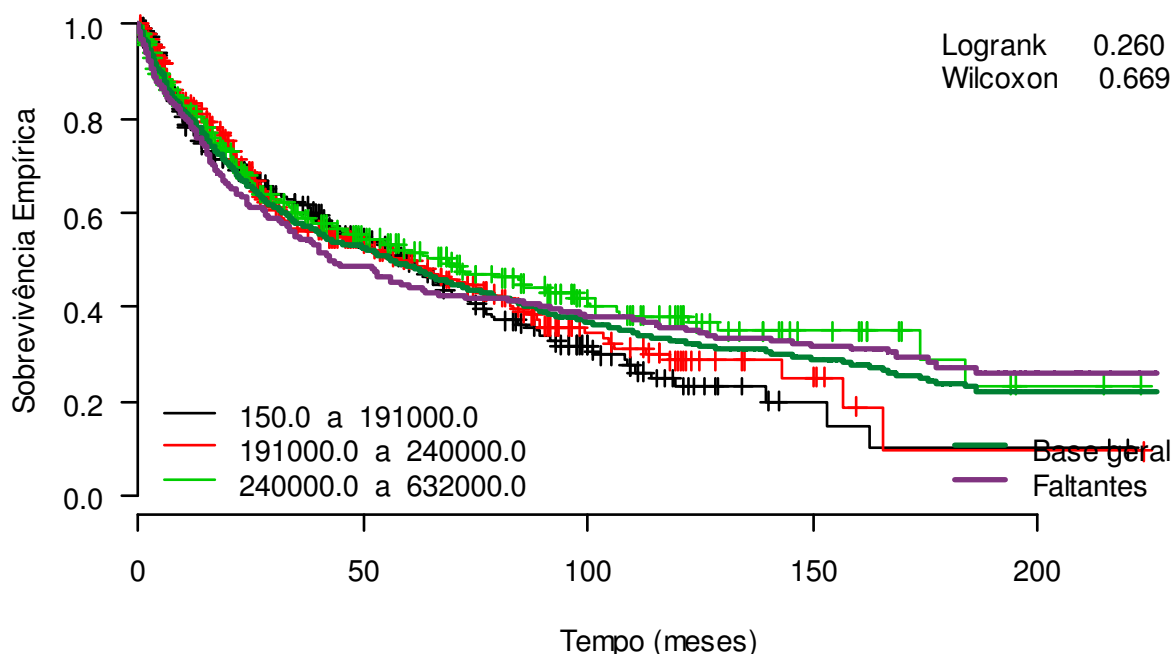


Figura A96: Curva de Kaplan-Meier por categoria de ácido úrico (mg/dl).

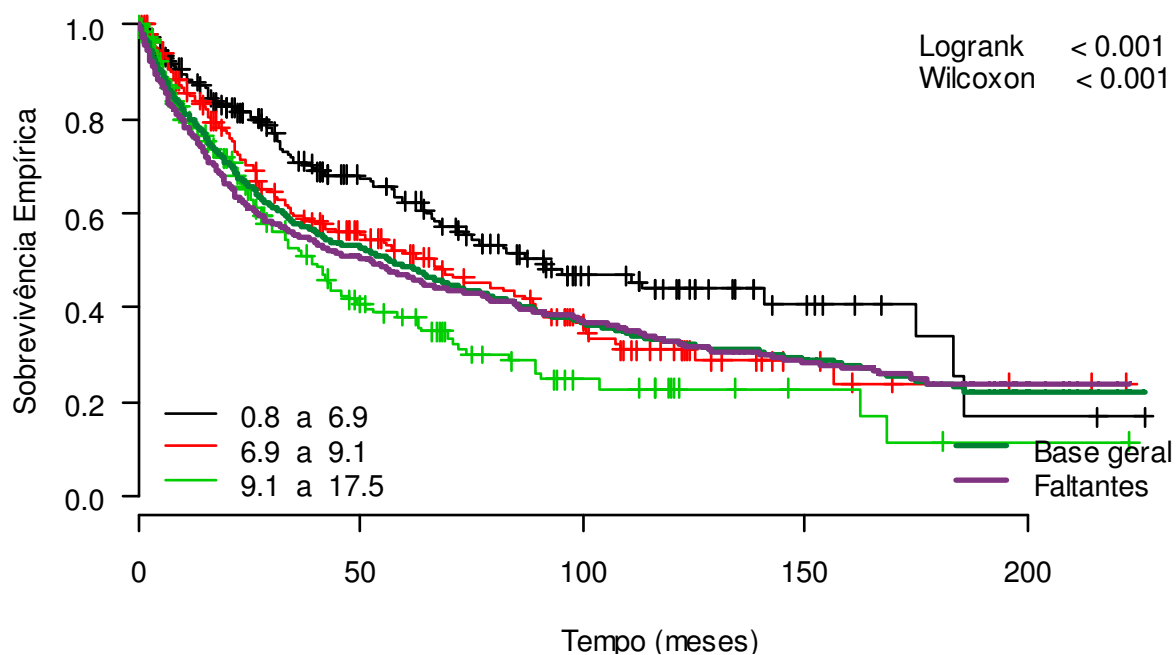


Figura A97: Curva de Kaplan-Meier por categoria de creatinina (mg/dl).

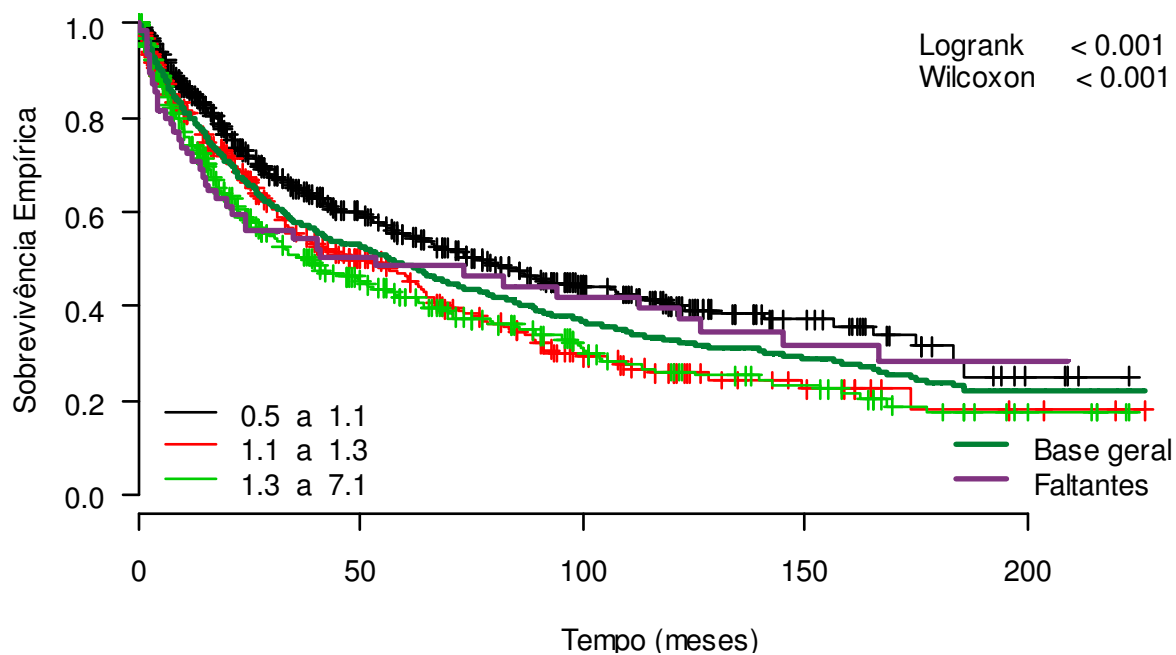


Figura A98: Curva de Kaplan-Meier por categoria de sódio (mEq/dl).

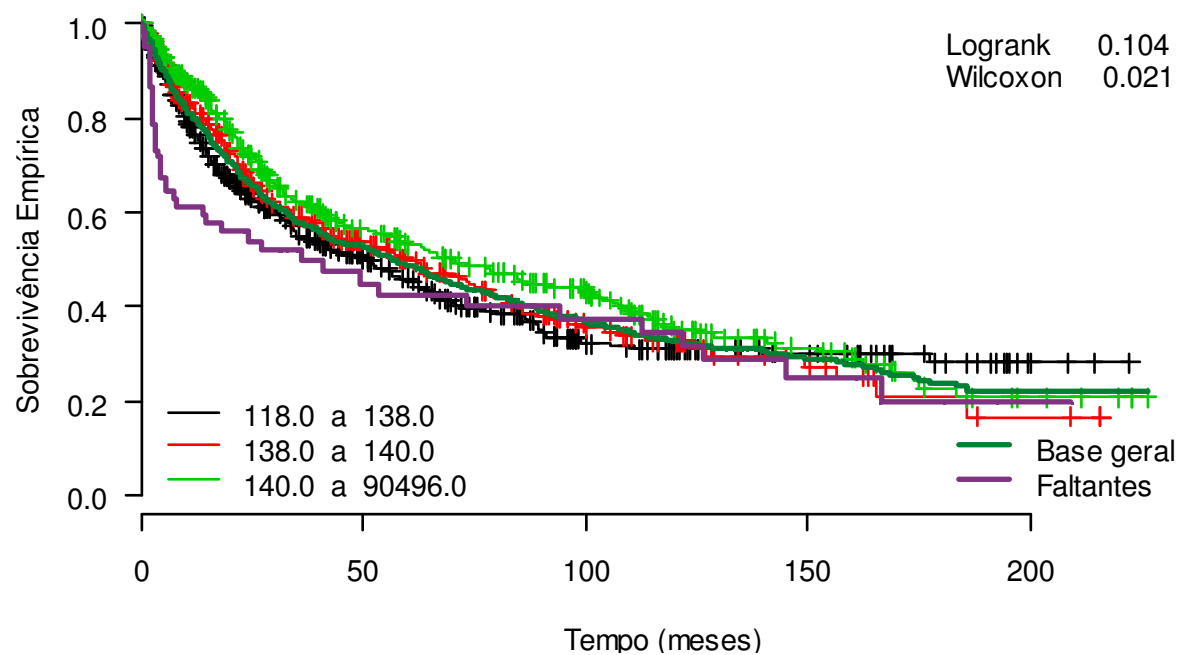


Figura A99: Curva de Kaplan-Meier por categoria de diâmetro da aorta (mm).

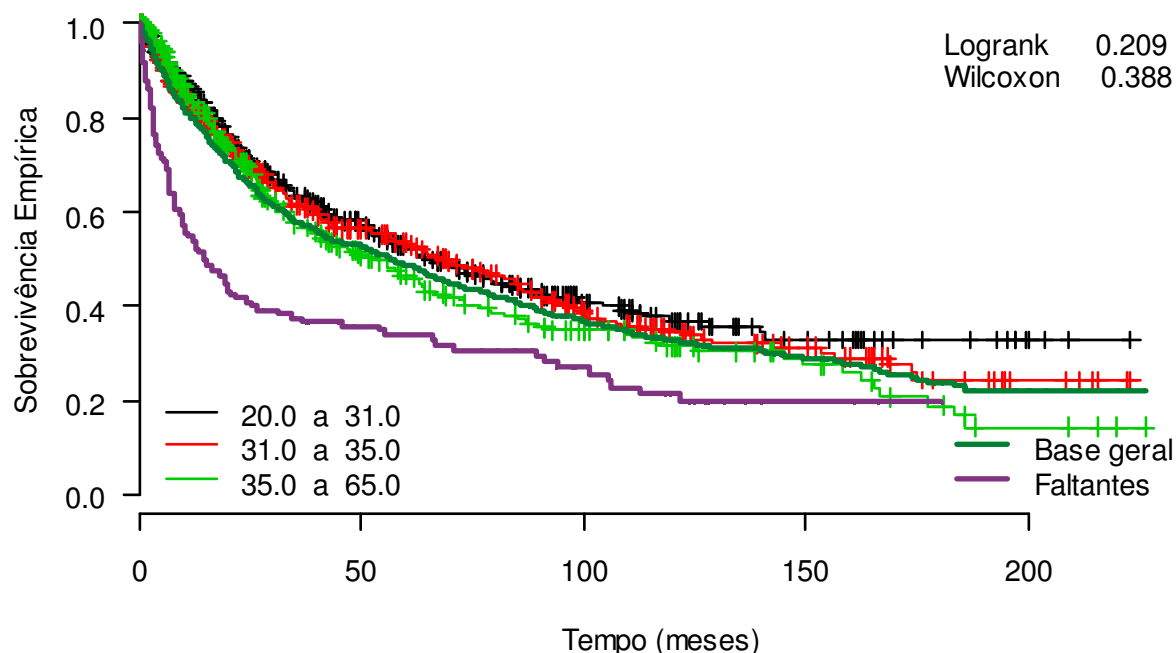


Figura A100: Curva de Kaplan-Meier por categoria de diâmetro do átrio esquerdo (mm).

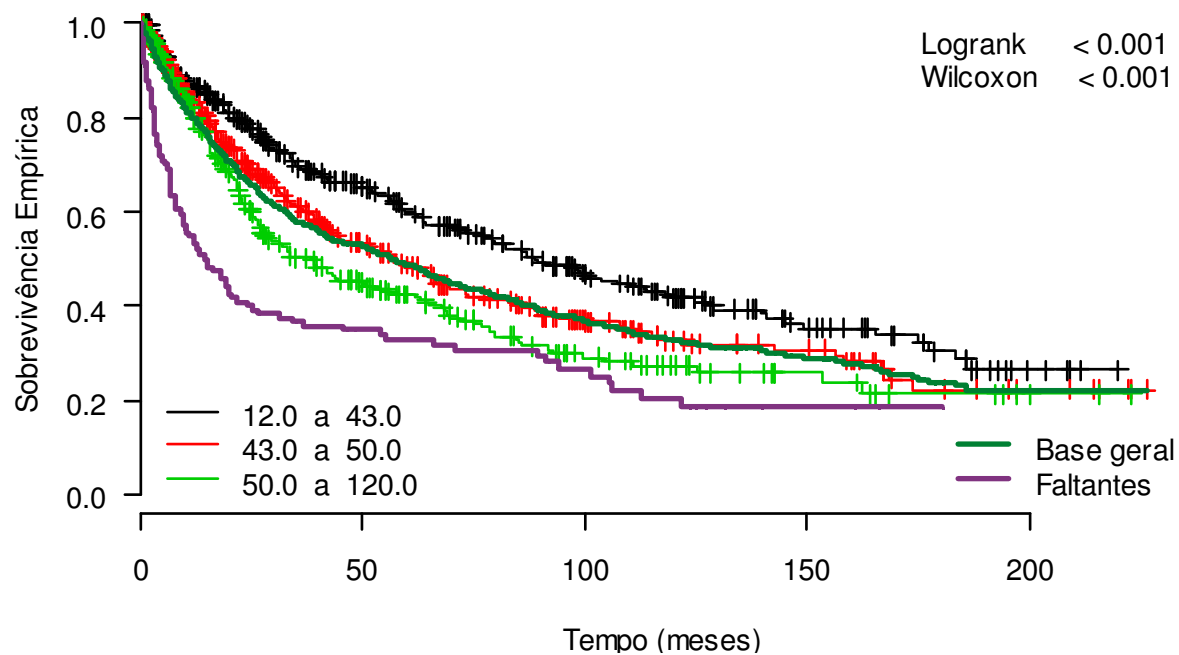


Figura A101: Curva de Kaplan-Meier por categoria de diâmetro do ventrículo direito (mm).

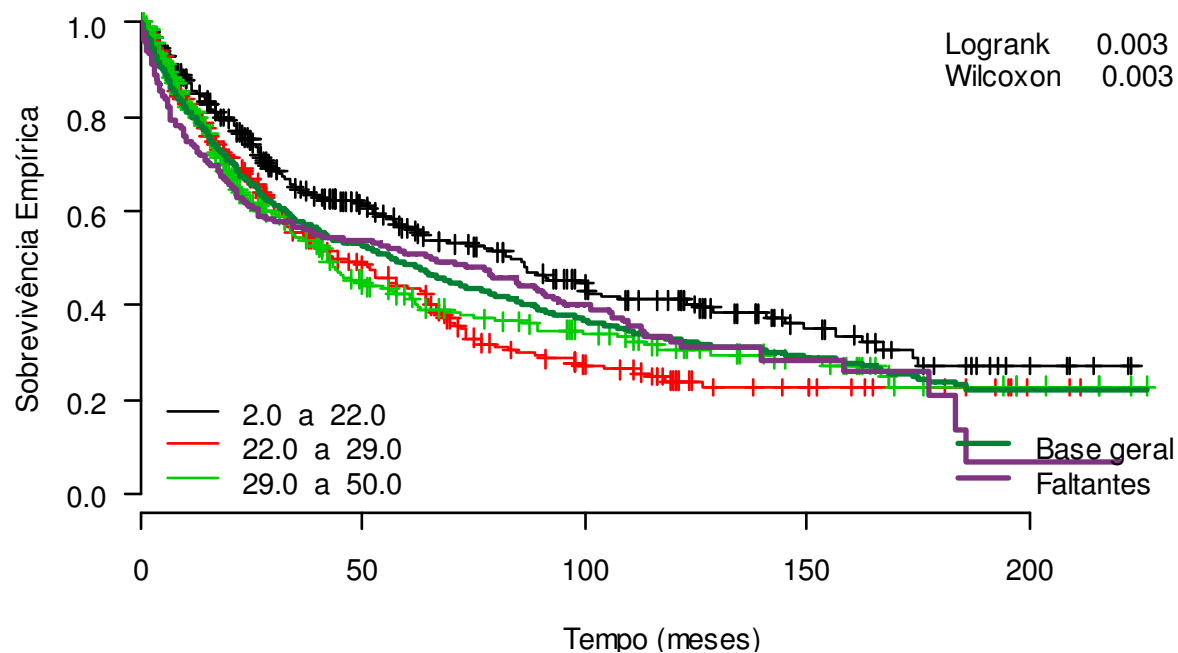


Figura A102: Curva de Kaplan-Meier por categoria de diâmetro diastólico do ventrículo esquerdo (mm).

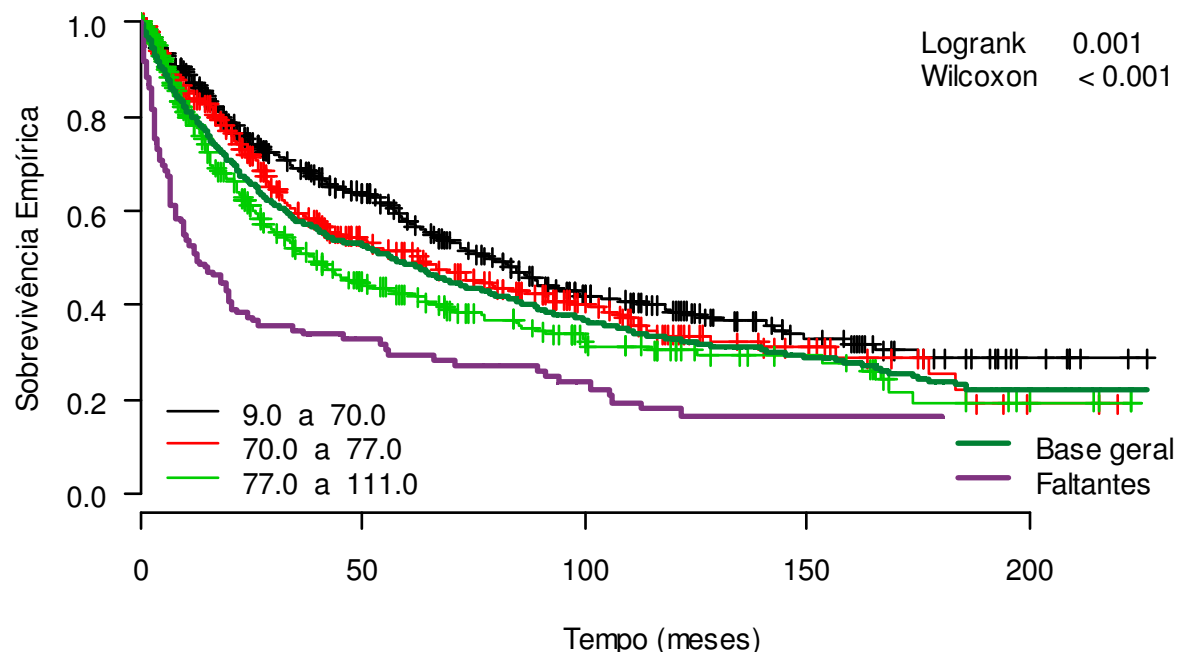


Figura A103: Curva de Kaplan-Meier por categoria de diâmetro sistólico do ventrículo esquerdo (mm).

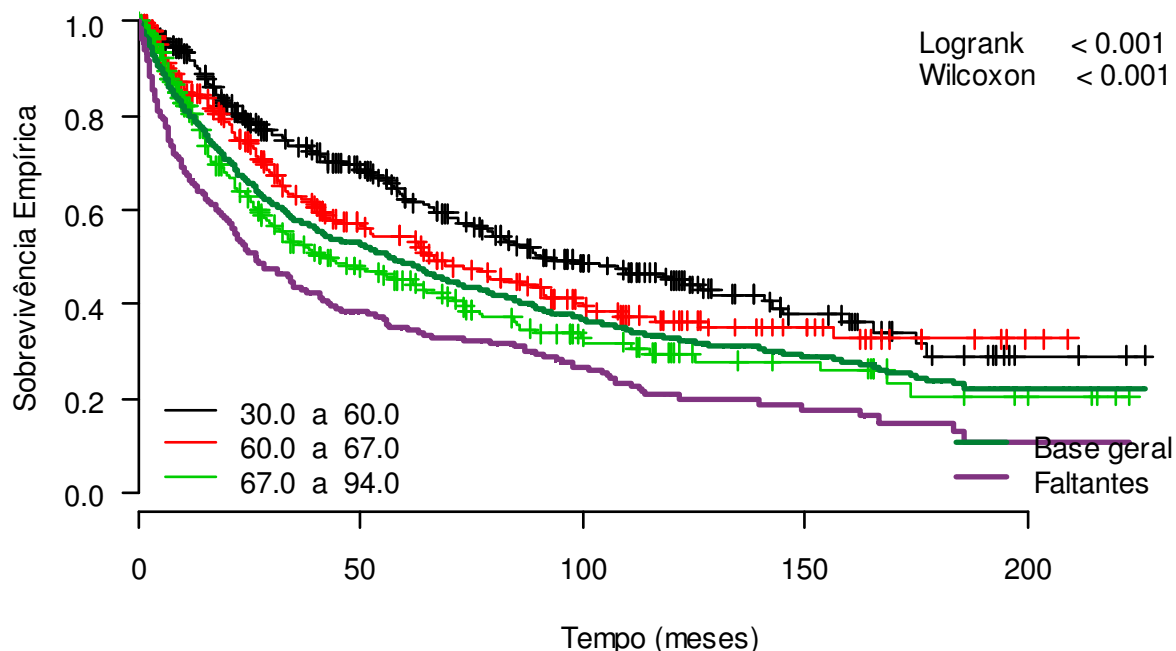


Figura A104: Curva de Kaplan-Meier por categoria de espessura da parede posterior (mm).

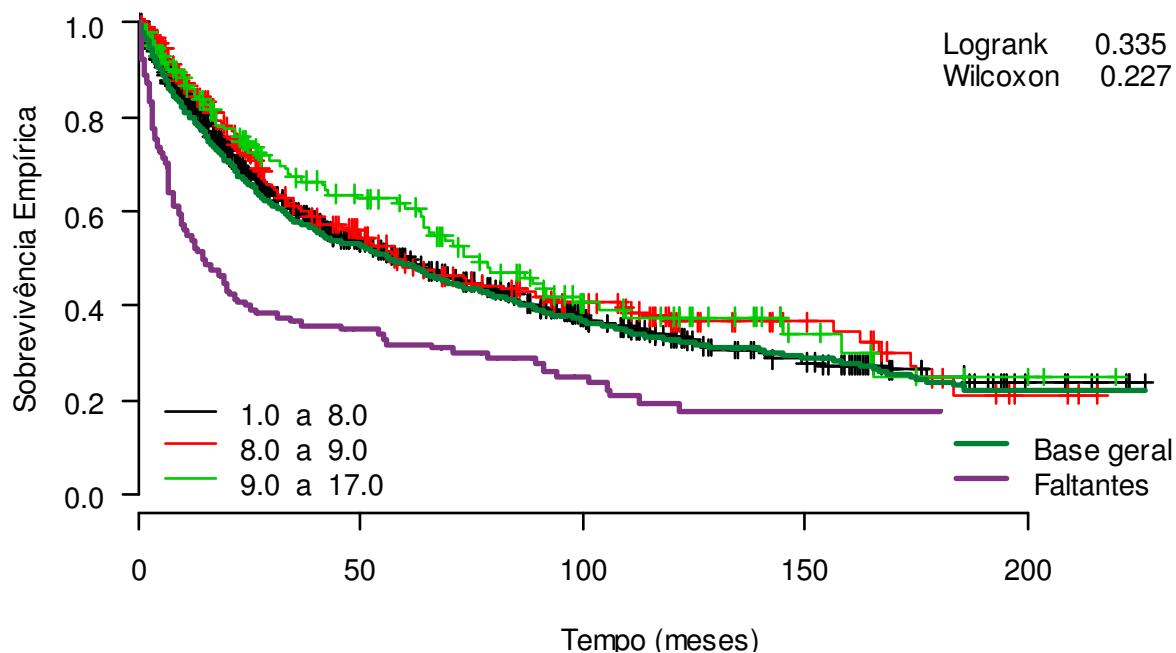


Figura A105: Curva de Kaplan-Meier por categoria de espessura do septo interventricular (mm).

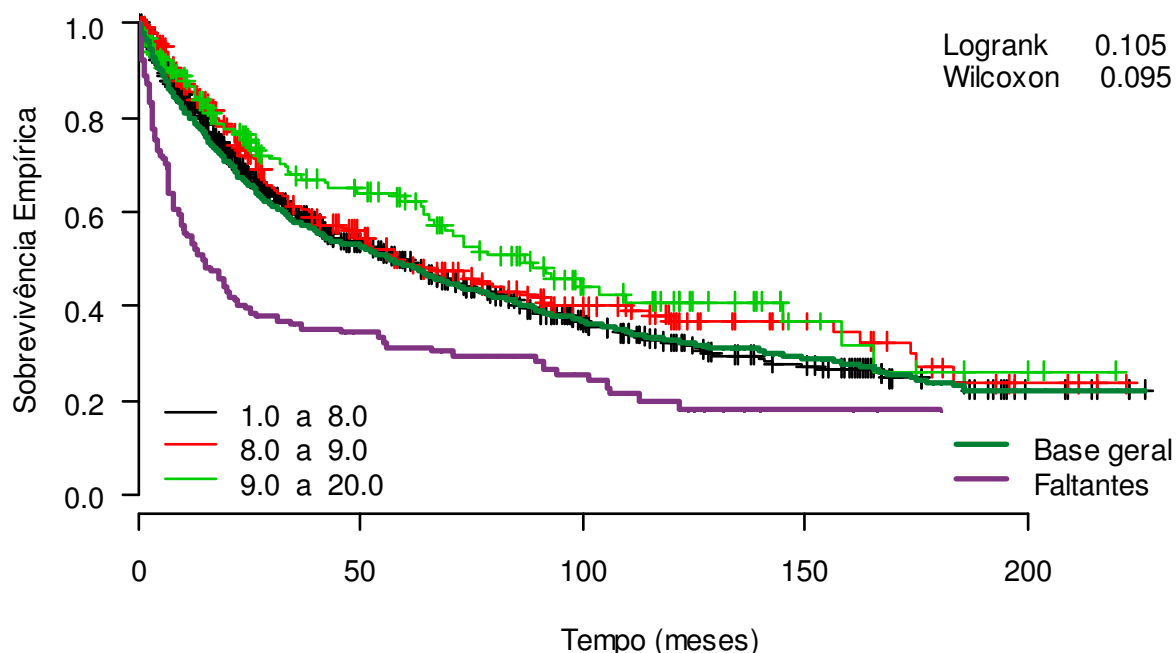


Figura A106: Curva de Kaplan-Meier por categoria de frequência cardíaca no Holter (bpm).

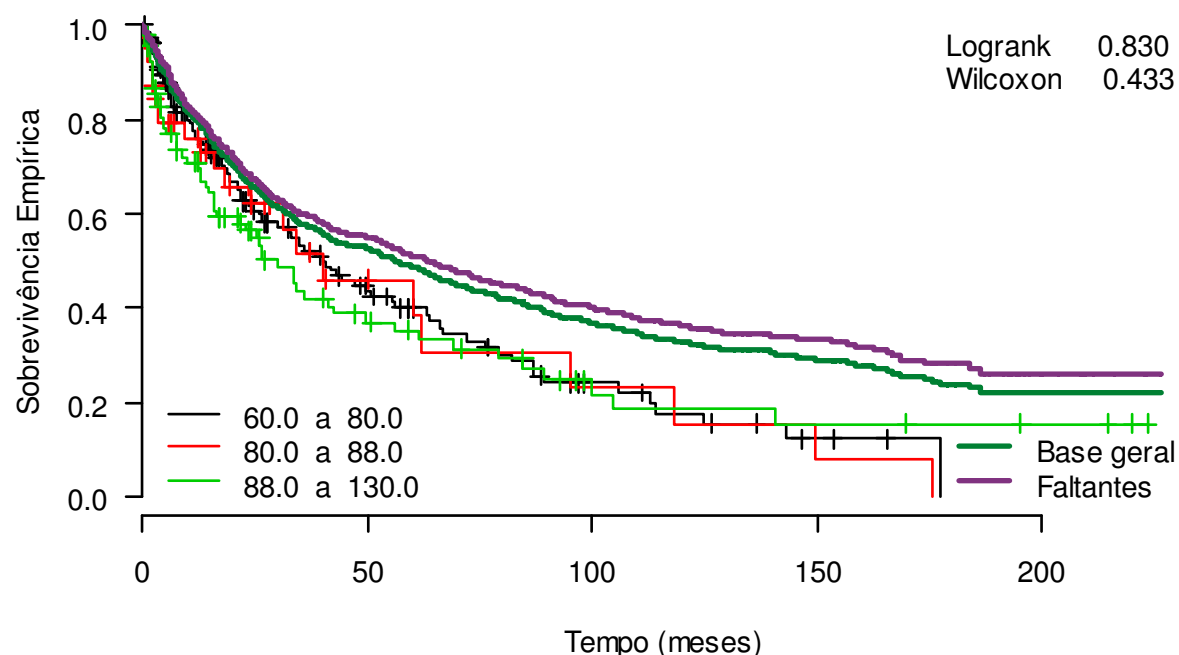


Figura A107: Curva de Kaplan-Meier por categoria de frequência cardíaca ao Gated (bpm).

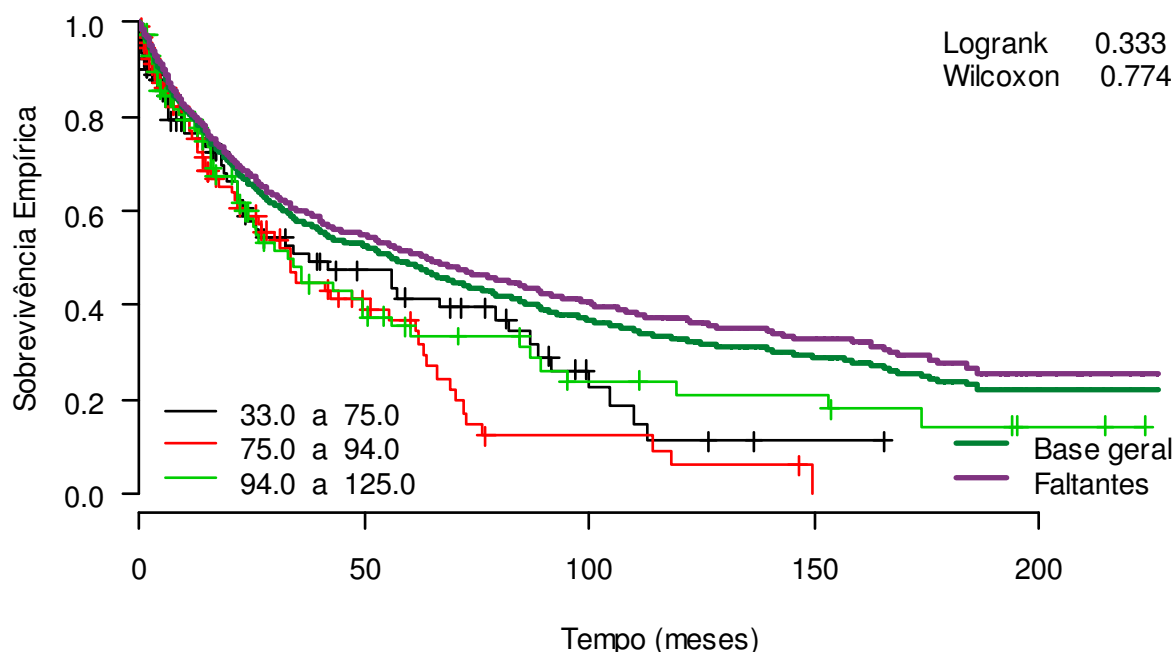


Figura A108: Curva de Kaplan-Meier por categoria de frequência cardíaca máxima (bpm).

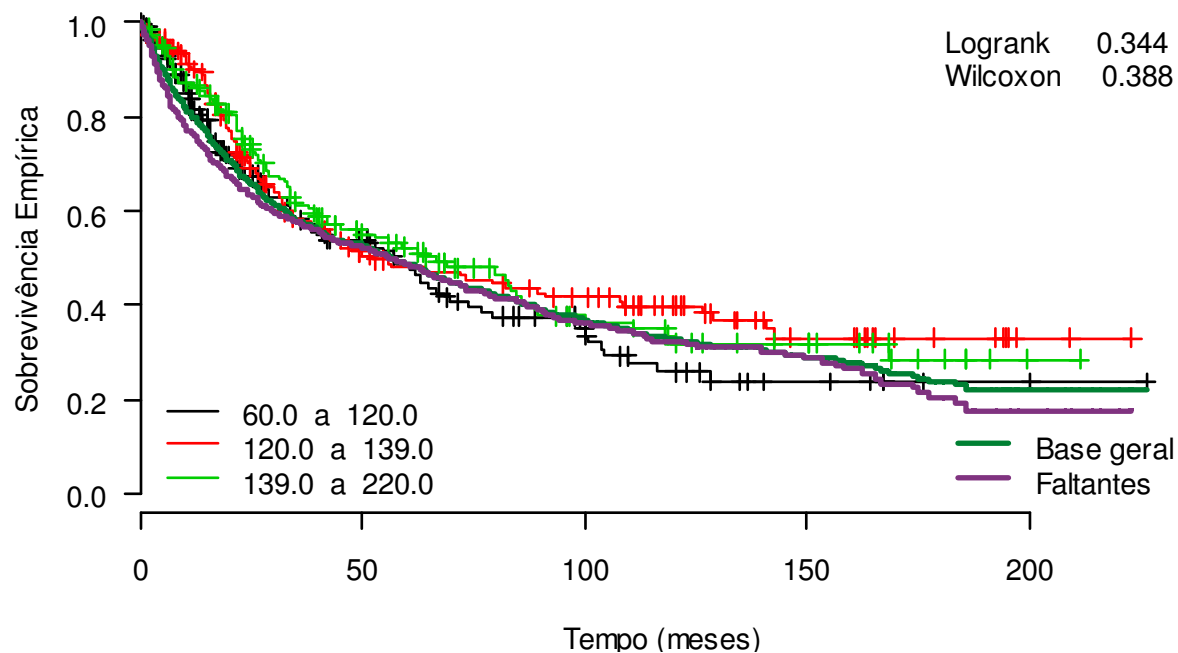


Figura A109: Curva de Kaplan-Meier por categoria de frequência cardíaca média (bpm).

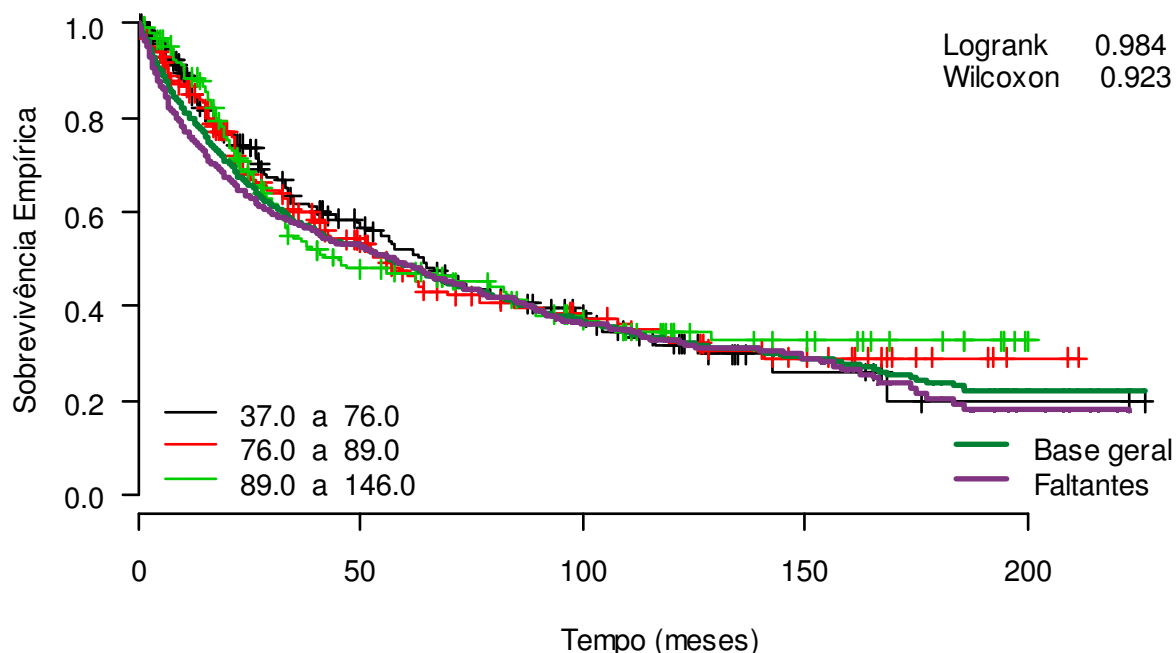


Figura A110: Curva de Kaplan-Meier por categoria de frequência cardíaca mínima (bpm).

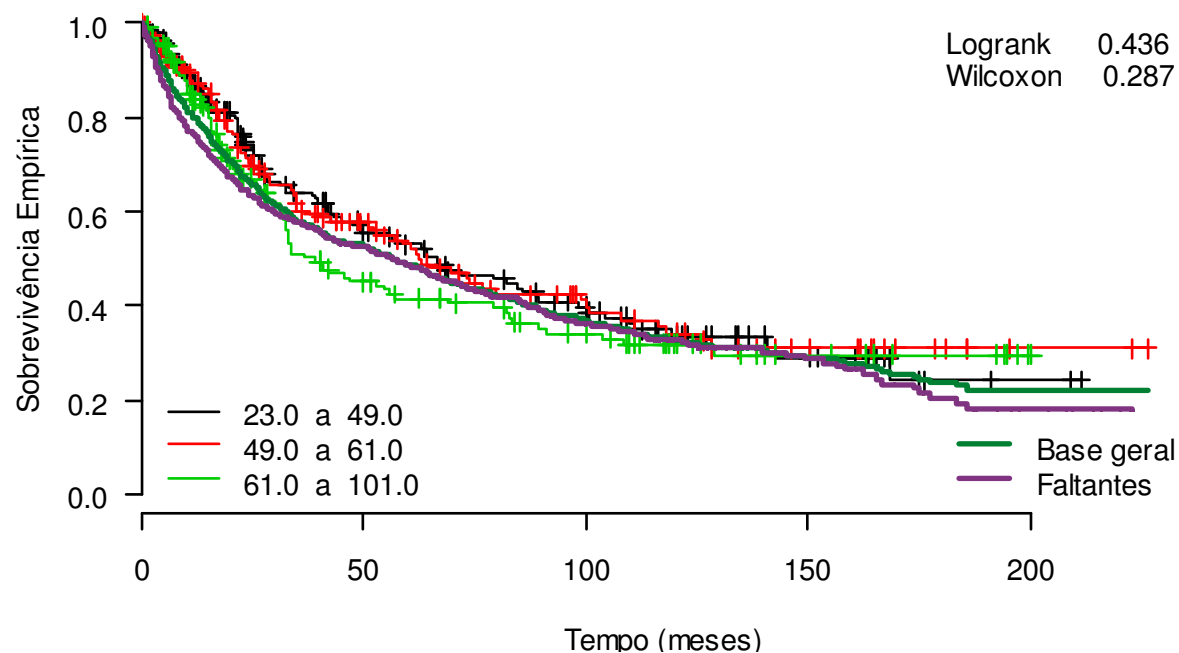


Figura A111: Curva de Kaplan-Meier por categoria de fração de ejeção do ventrículo esquerdo no ecocardiograma (%).

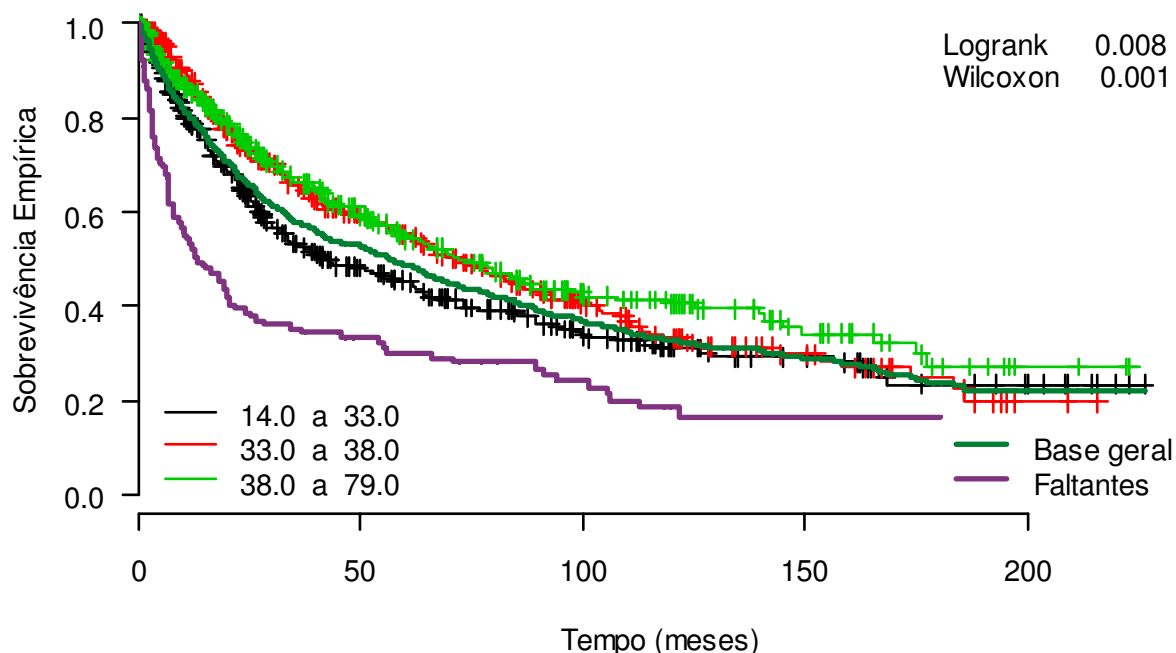


Figura A112: Curva de Kaplan-Meier por categoria de fração de ejeção do ventrículo esquerdo (%).

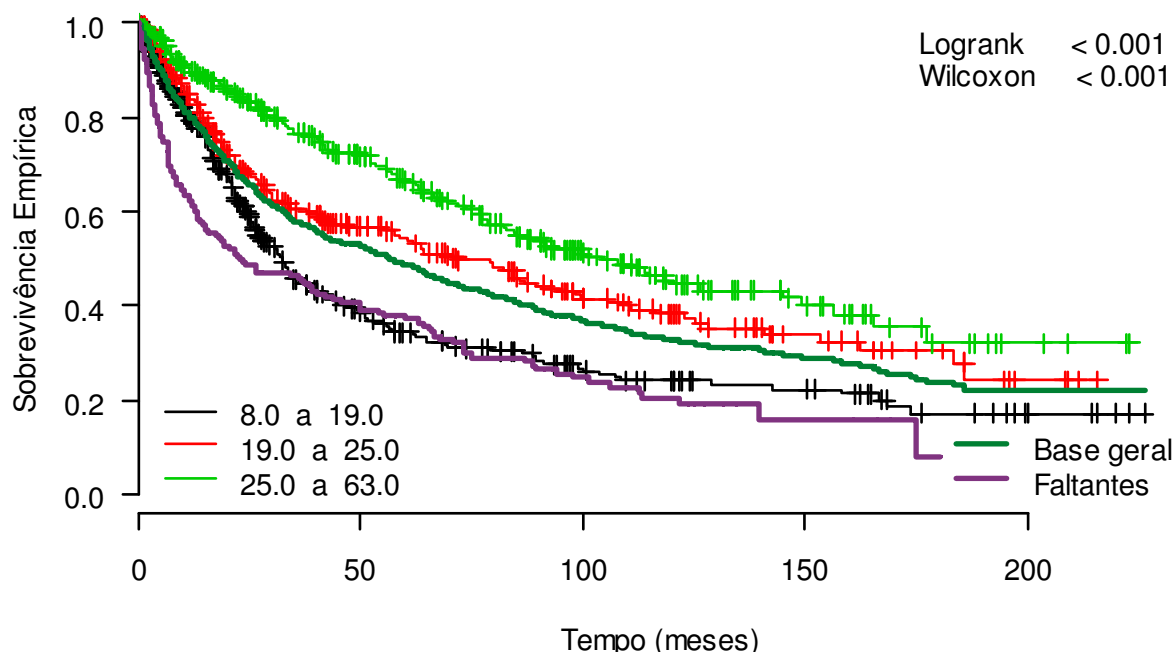


Figura A113: Curva de Kaplan-Meier por categoria de fração de ejeção do ventrículo direito (%).

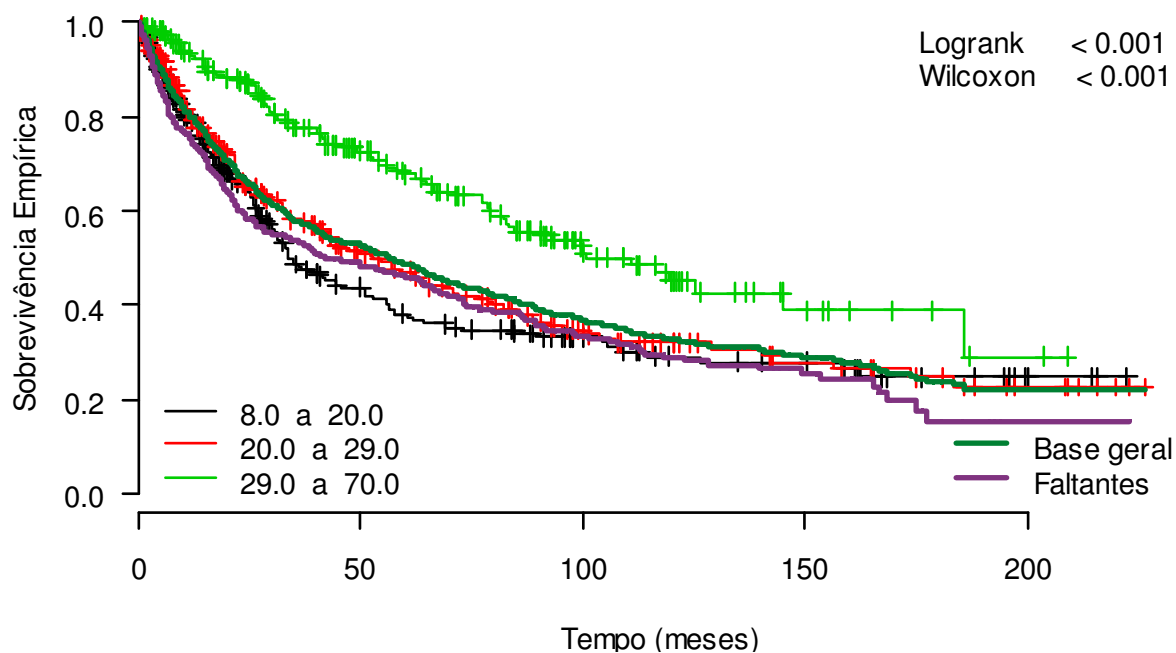


Figura A114: Curva de Kaplan-Meier por categoria de pressão arterial diastólica (mmHg).

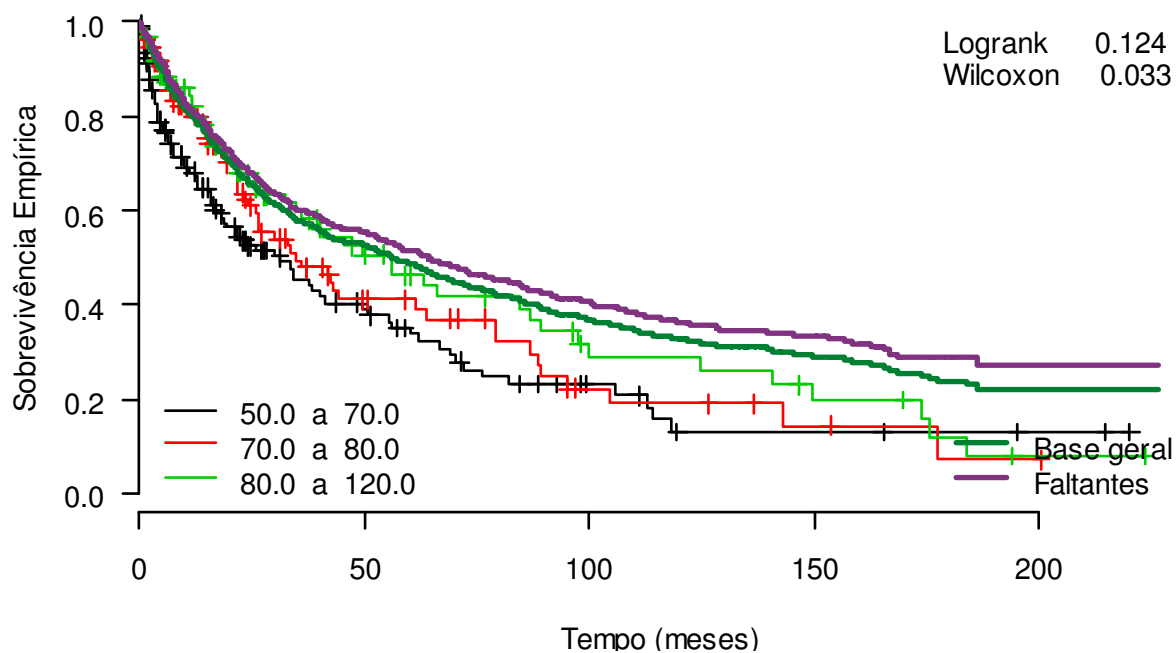


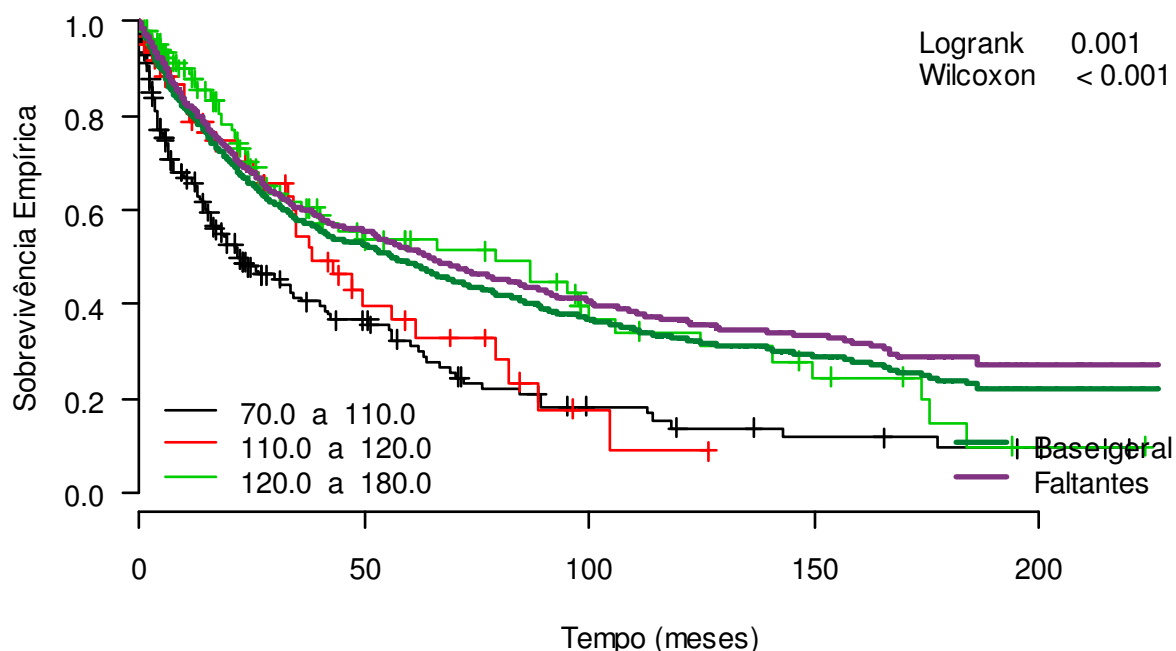
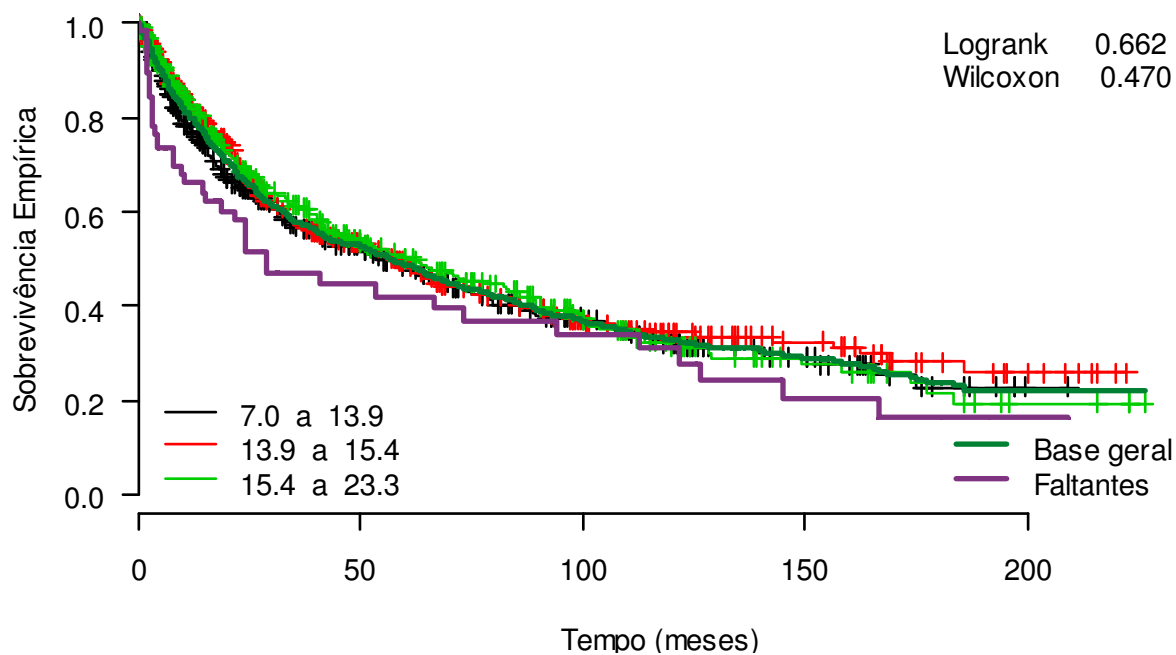
Figura A115: Curva de Kaplan-Meier por categoria de pressão arterial sistólica (mmHg).**Figura A116:** Curva de Kaplan-Meier por categoria de hemoglobina sérica (g/dl).

Figura A117: Curva de Kaplan-Meier por categoria de hematócrito sérico (%).

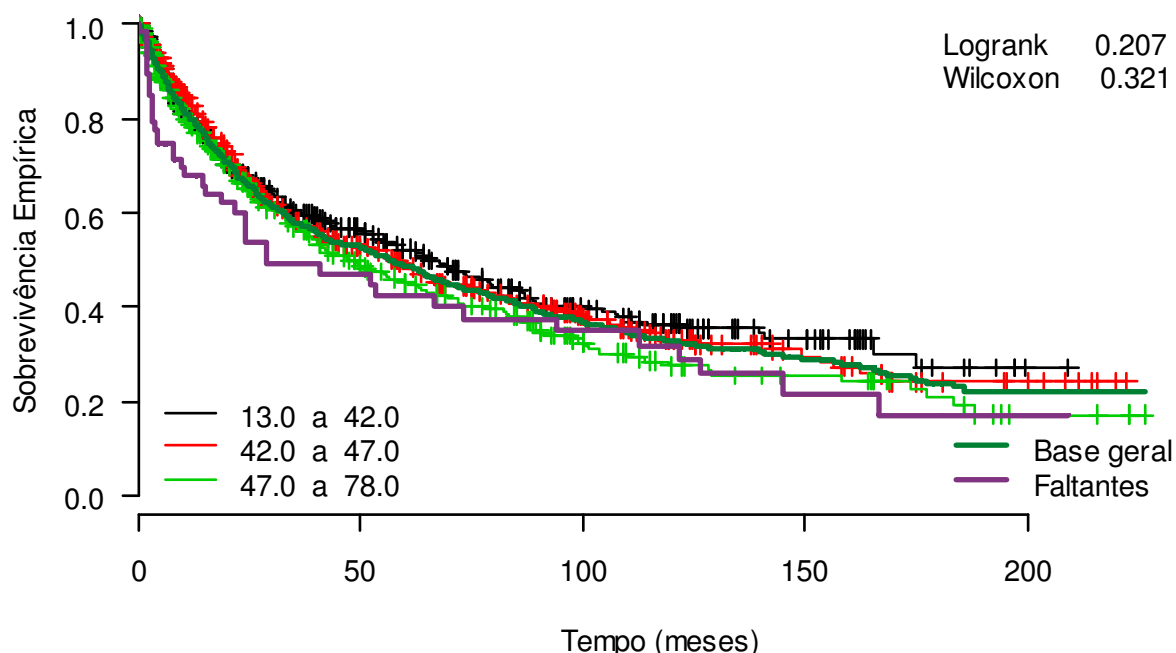


Figura A118: Curva de Kaplan-Meier por categoria de número total de extrassístoles ventriculares.

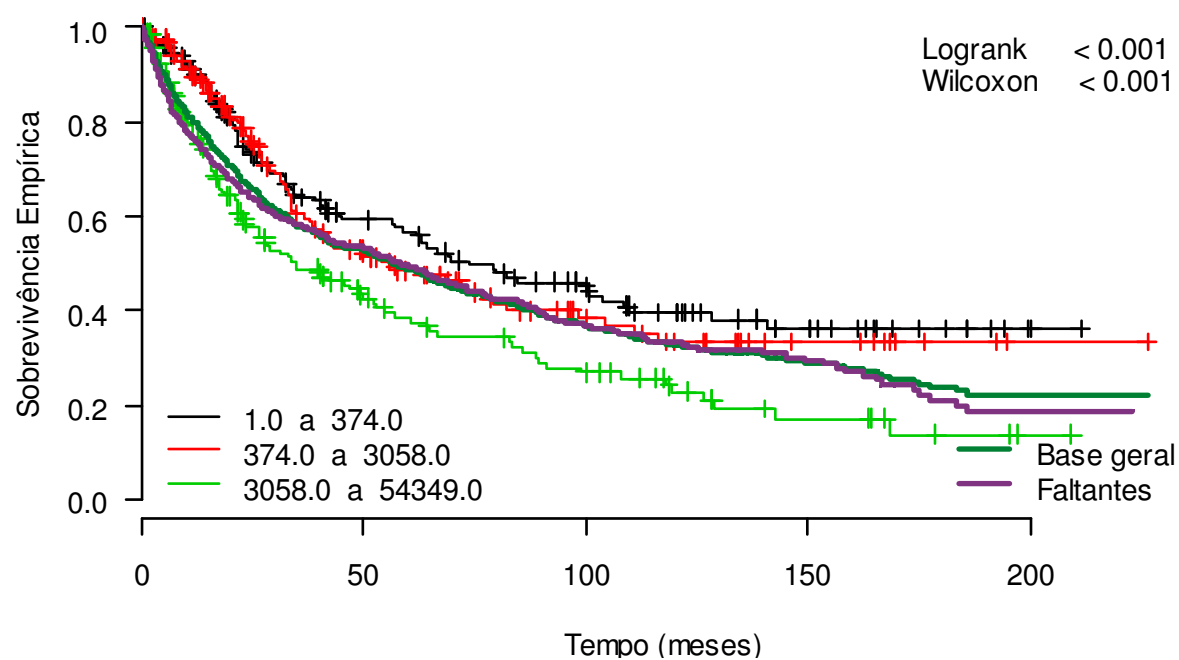


Figura A119: Curva de Kaplan-Meier por categoria de número de extrassístoles ventriculares por hora.

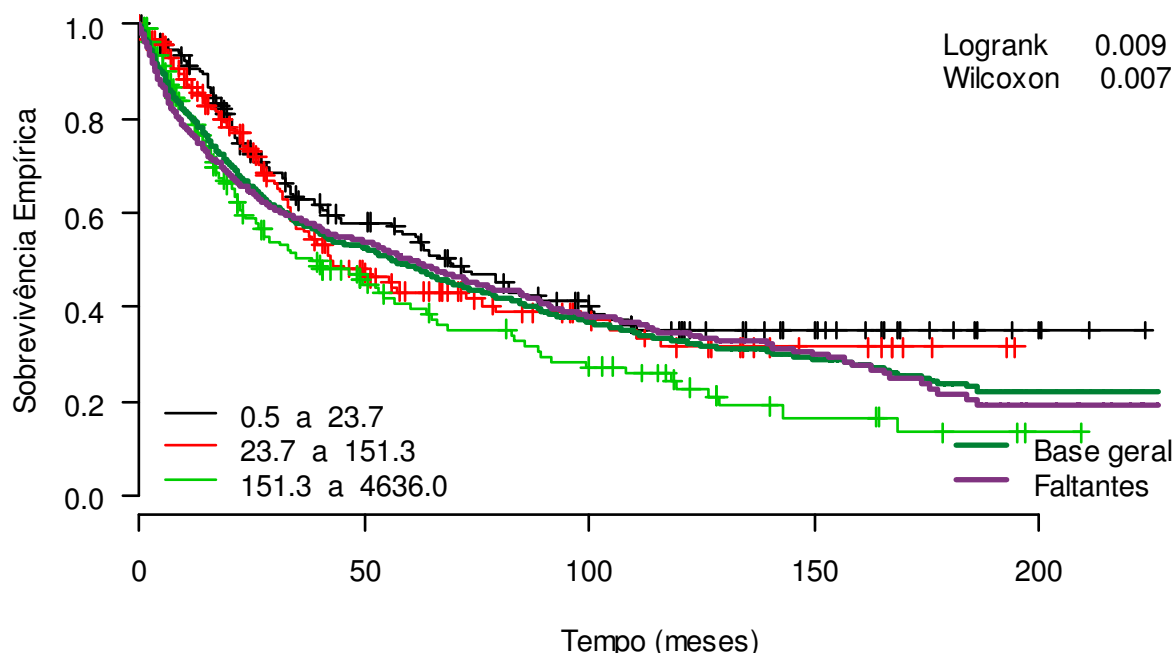


Figura A120: Curva de Kaplan-Meier por categoria de massa do ventrículo esquerdo (g).

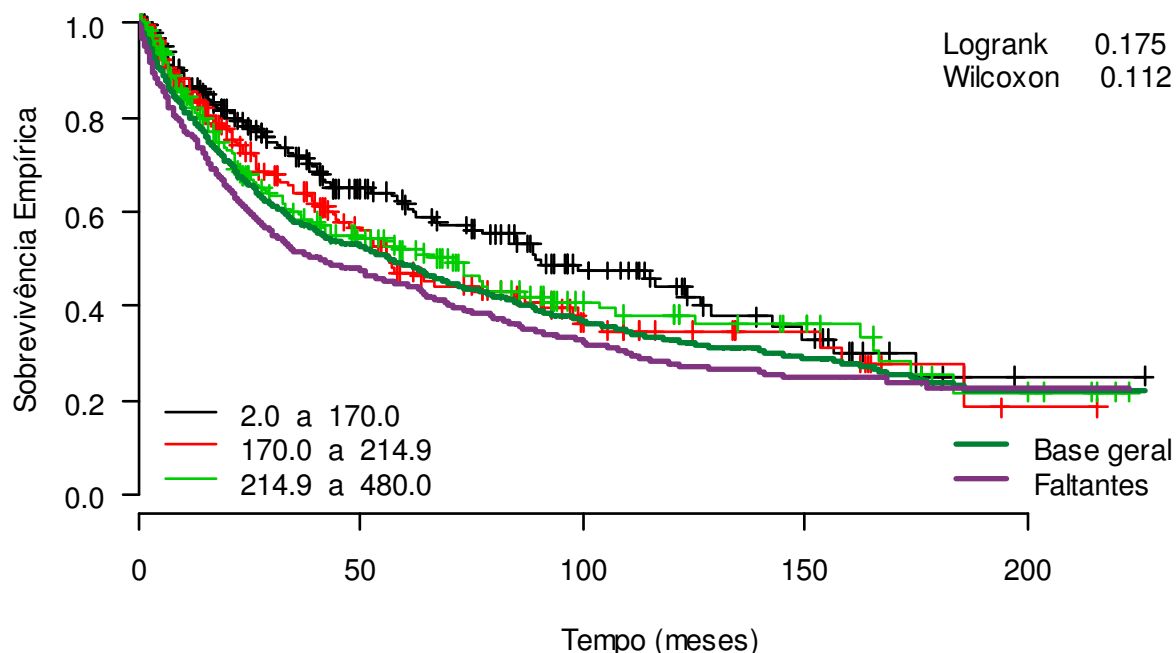


Figura A121: Curva de Kaplan-Meier por categoria de glicemia (mg/dl).

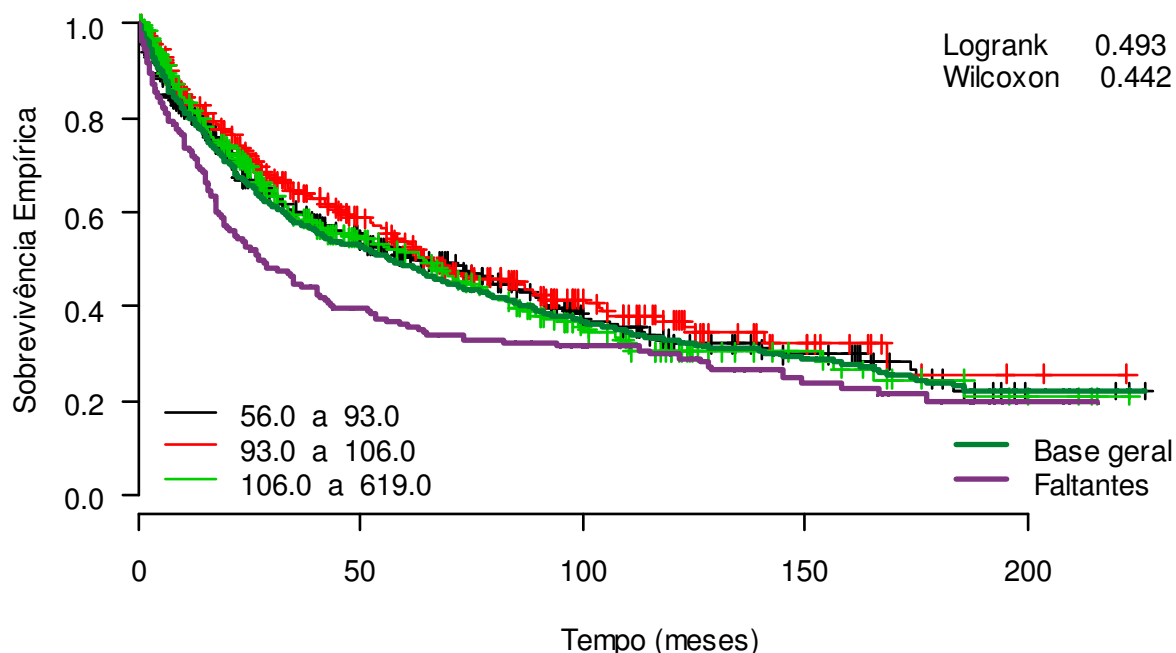


Figura A122: Curva de Kaplan-Meier por categoria de consumo máximo de oxigênio (ml/kg/m^2).

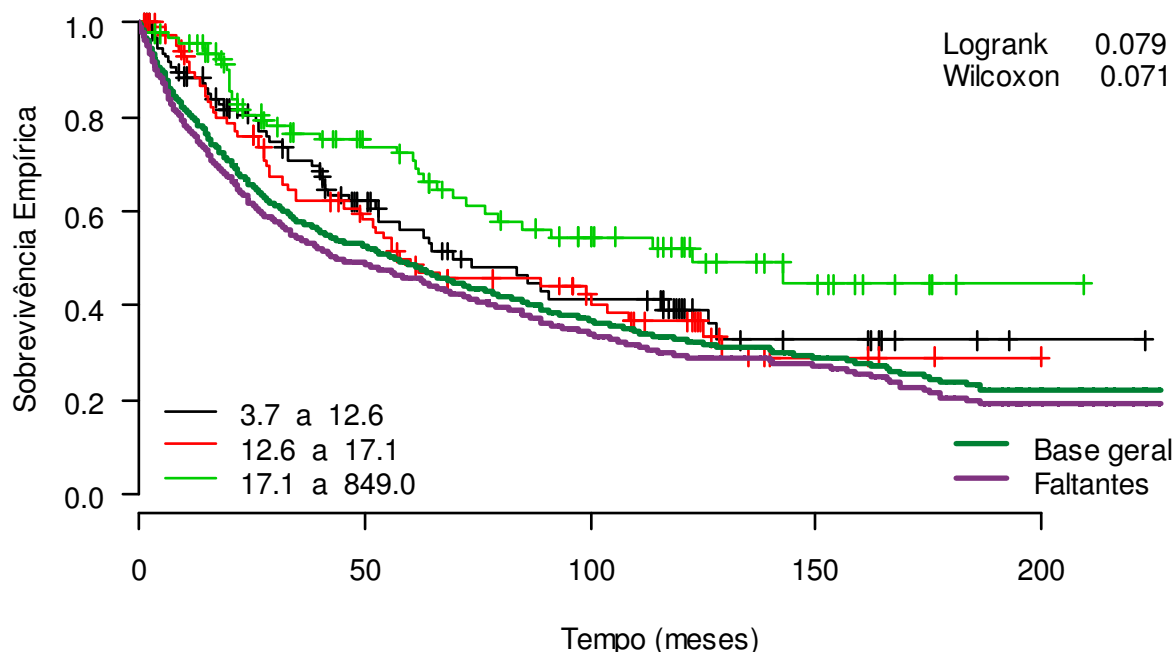


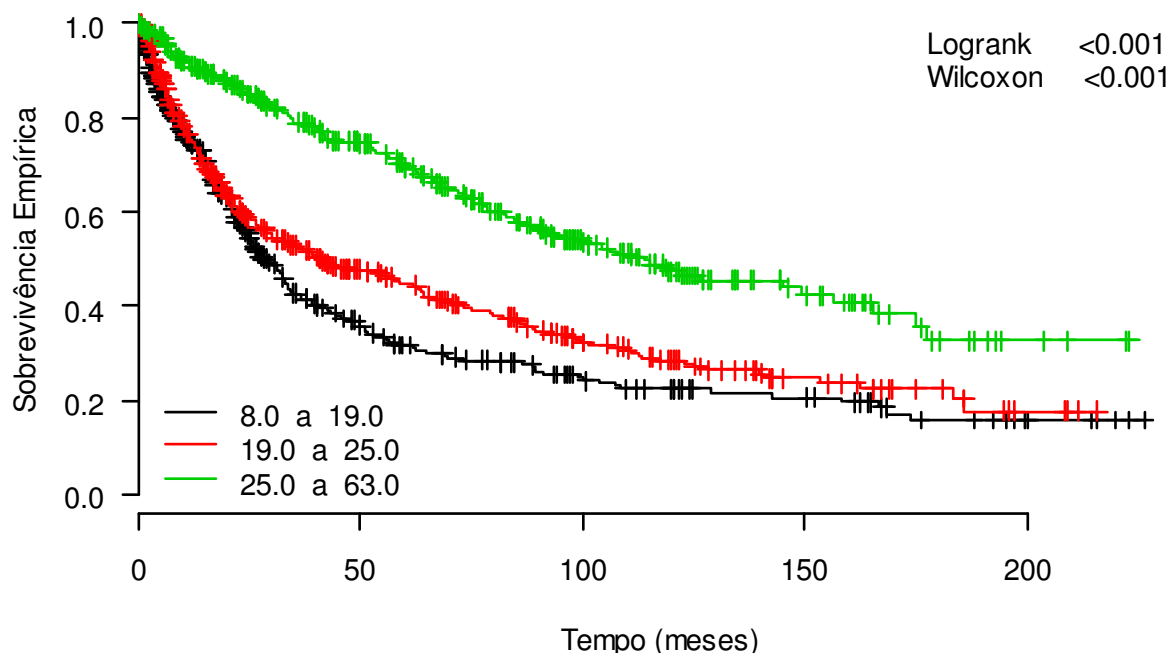
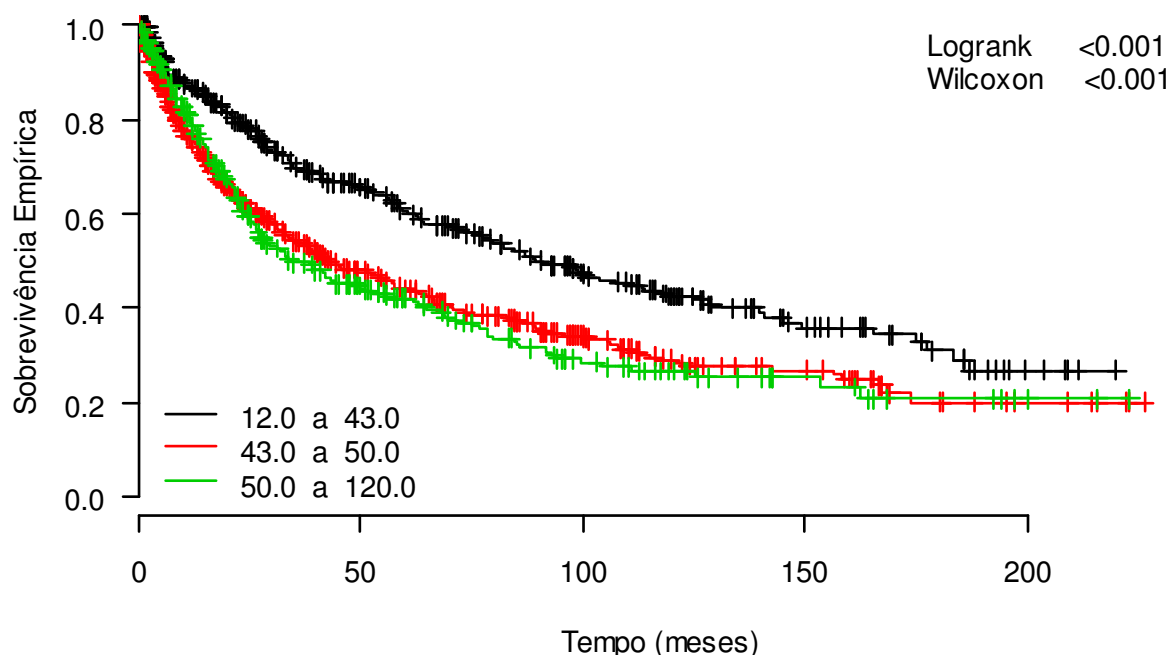
Figura A123: Curva de Kaplan-Meier por categoria de fração de ejeção ao Gated imputada (%).**Figura A124:** Curva de Kaplan-Meier por categoria de diâmetro do átrio esquerdo imputado (mm).

Figura A125: Curva de Kaplan-Meier por categoria de diâmetro diastólico do VE imputado (mm).

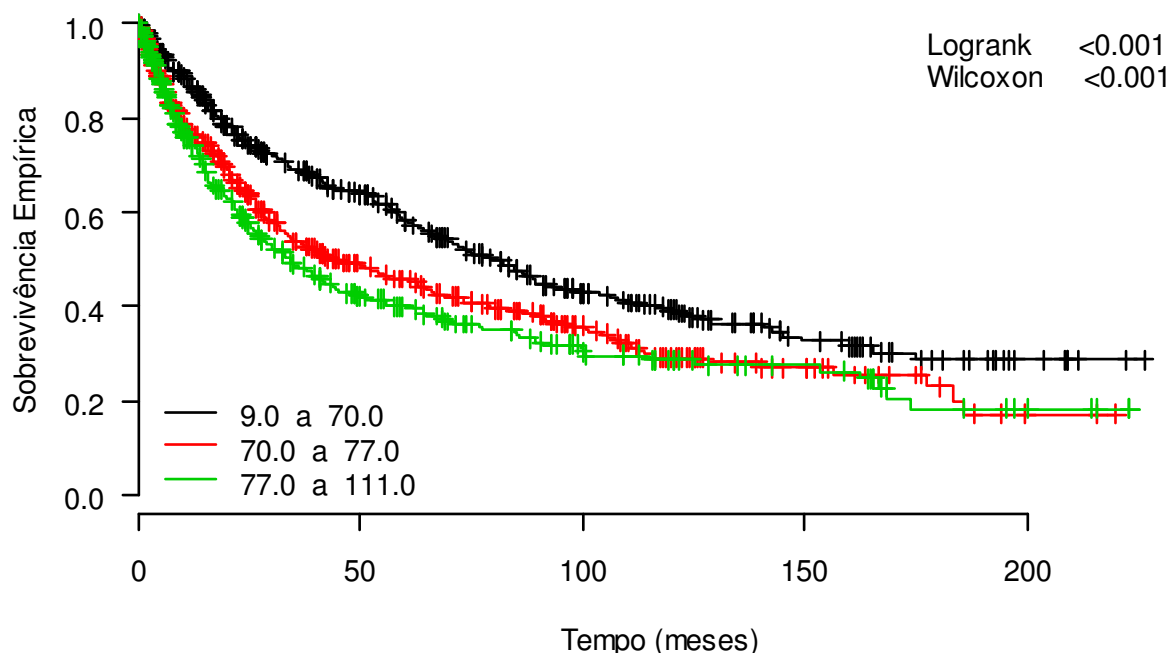


Figura A126: Curva de Kaplan-Meier por categoria de diâmetro sistólico do VE imputado (mm).

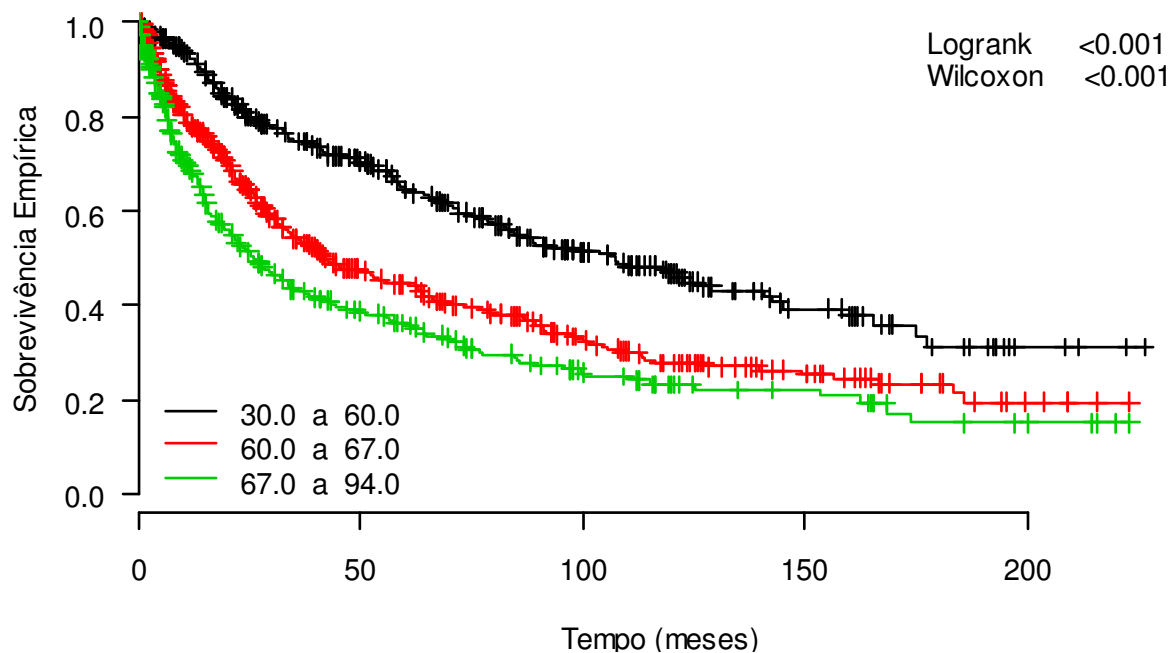


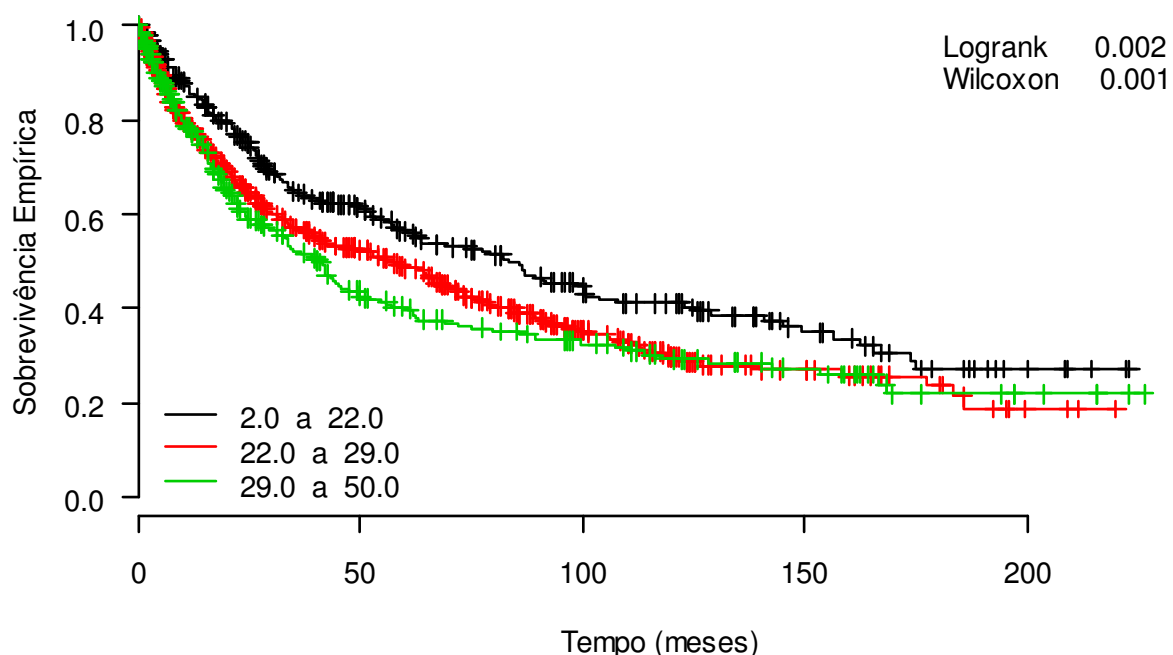
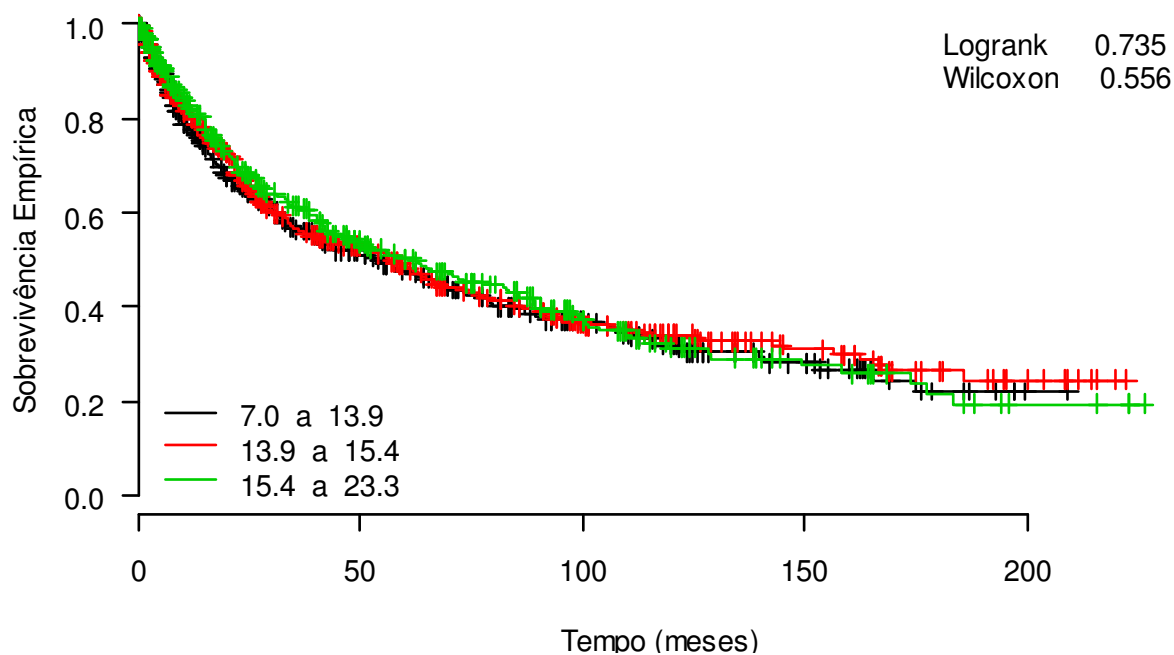
Figura A127: Curva de Kaplan-Meier por categoria de diâmetro do VD imputado (mm).**Figura A128:** Curva de Kaplan-Meier por categoria de hemoglobina sérica imputada (mg/dl).

Figura A129: Curva de Kaplan-Meier por categoria de monócitos imputados (%).

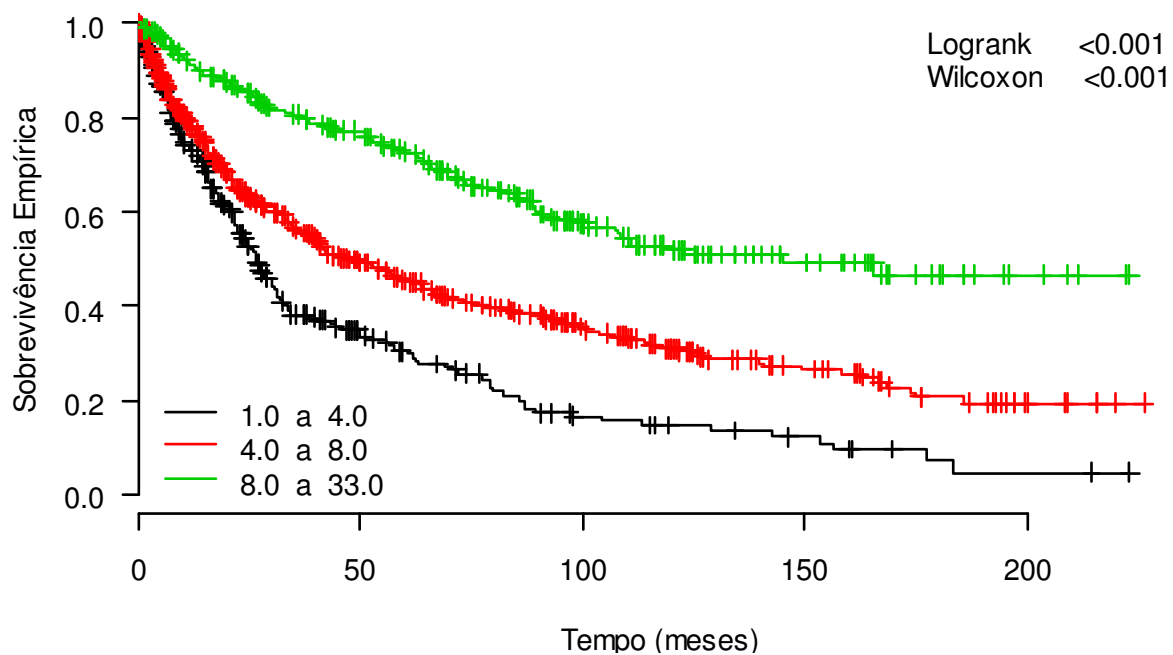


Figura A130: Curva de Kaplan-Meier por categoria do número absoluto de leucócitos imputados (/mm³).

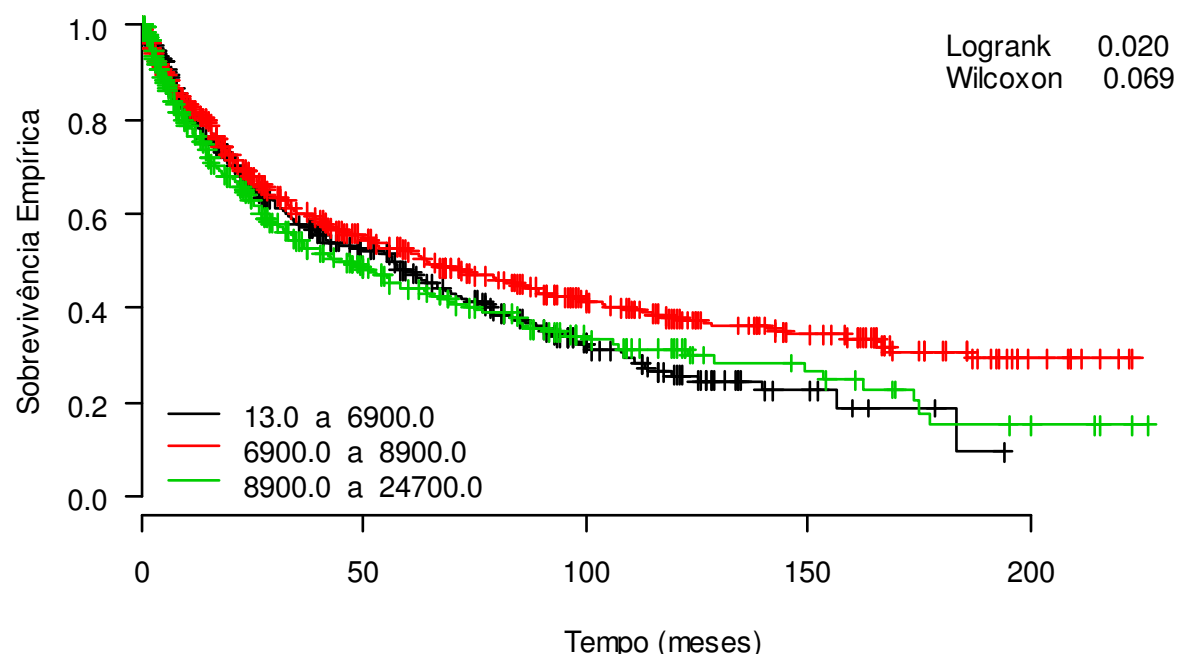


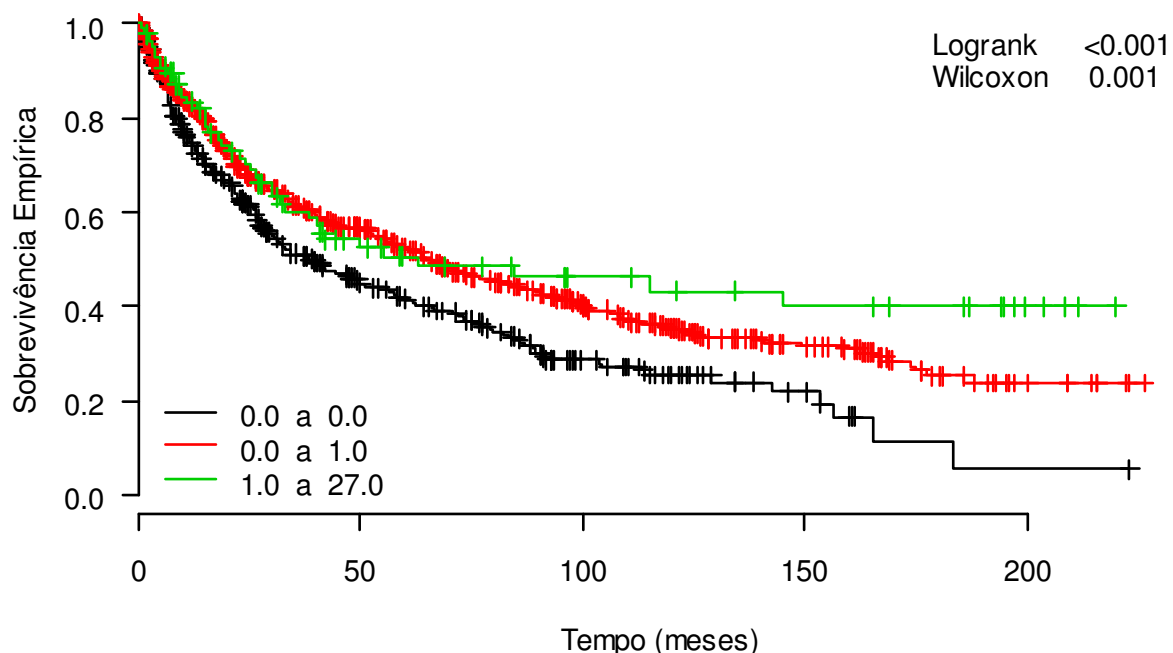
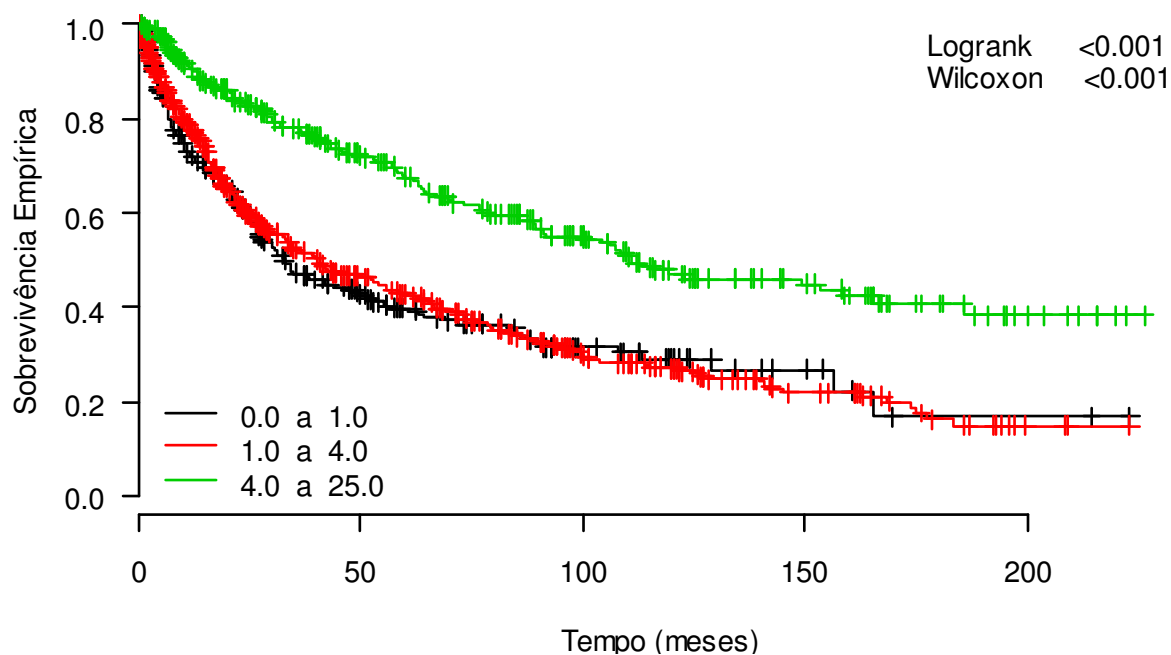
Figura A131: Curva de Kaplan-Meier por categoria de basófilos imputados (%).**Figura A132:** Curva de Kaplan-Meier por categoria de eosinófilos imputados (%).

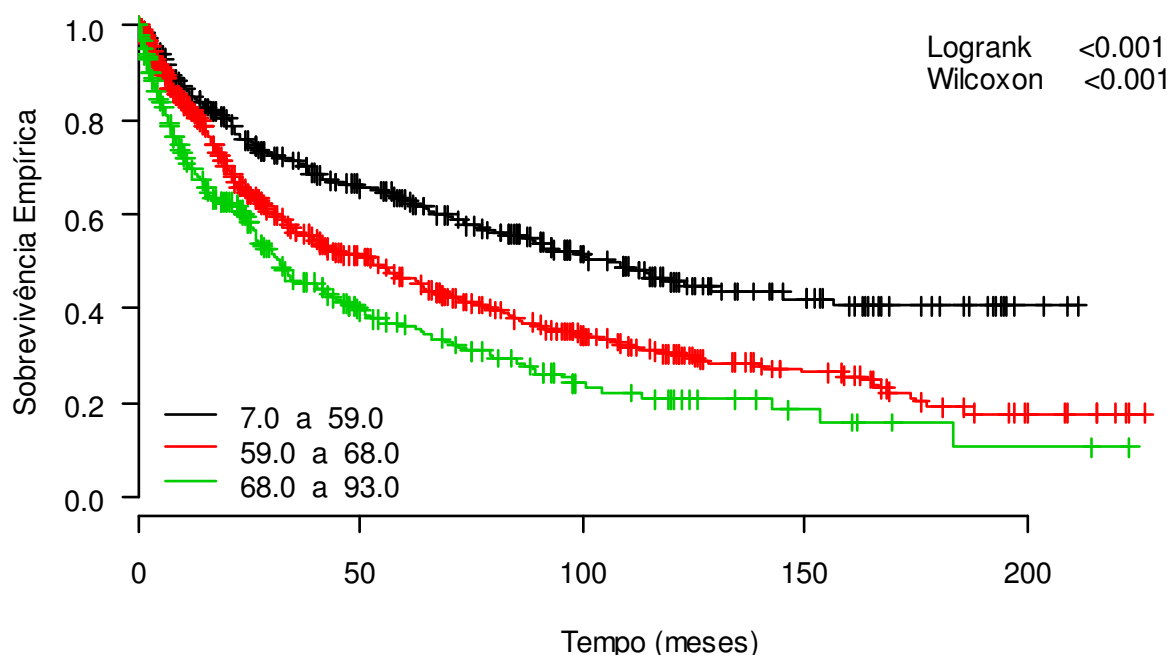
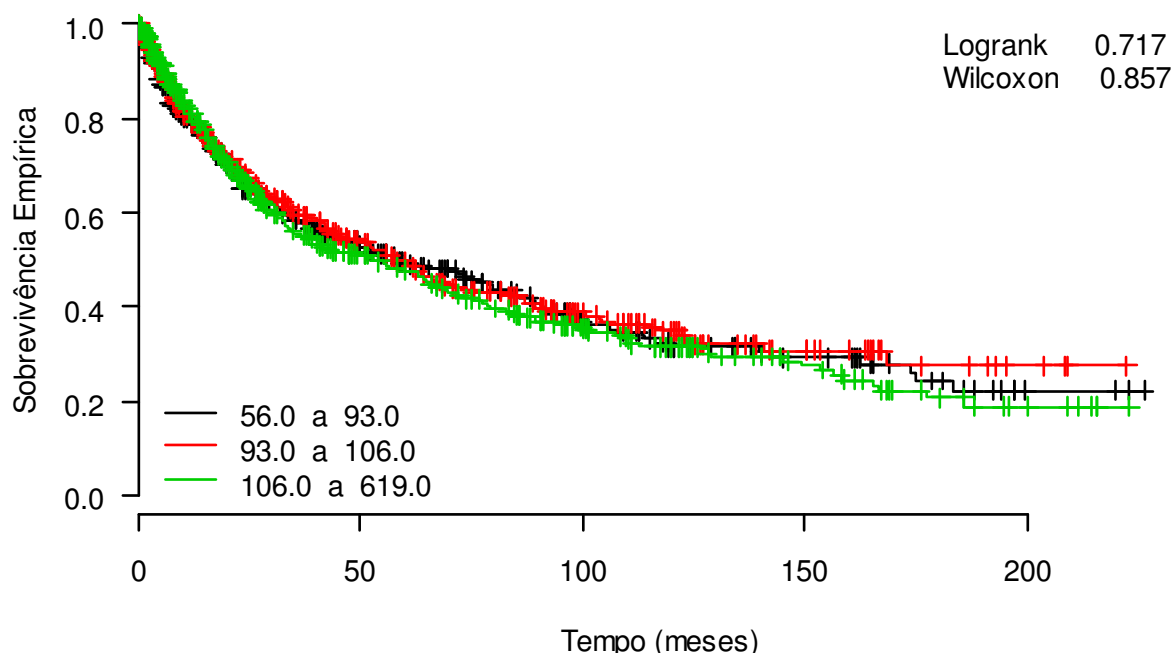
Figura A133: Curva de Kaplan-Meier por categoria de neutrófilos imputados (%).**Figura A134:** Curva de Kaplan-Meier por categoria de glicemia imputada (mg/dl).

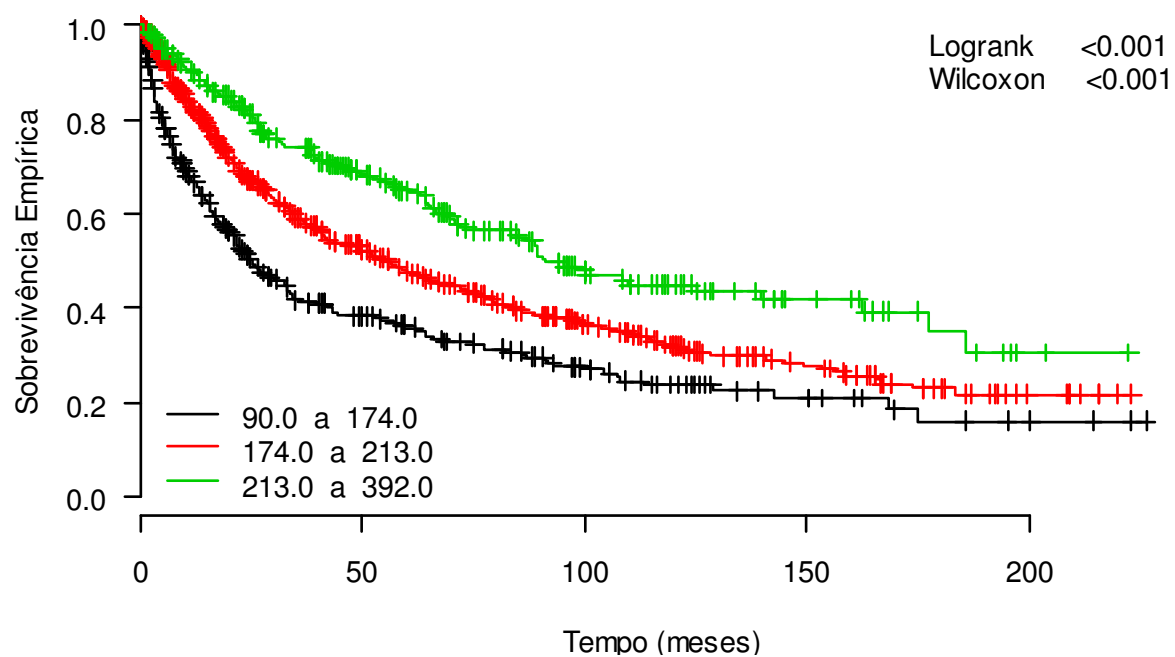
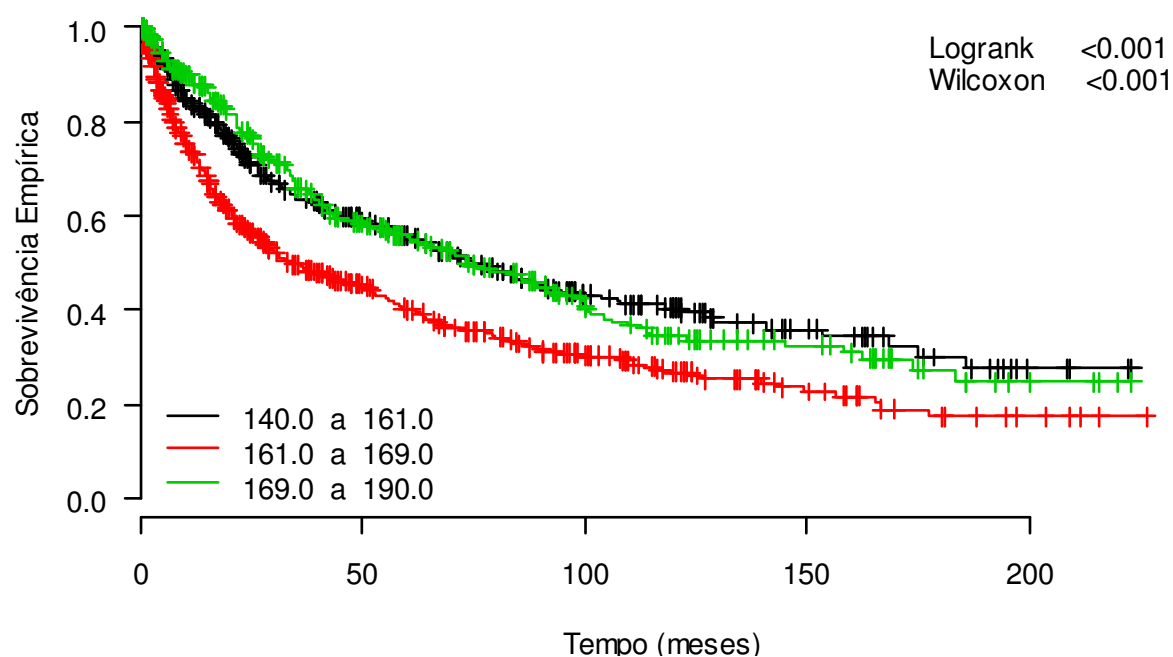
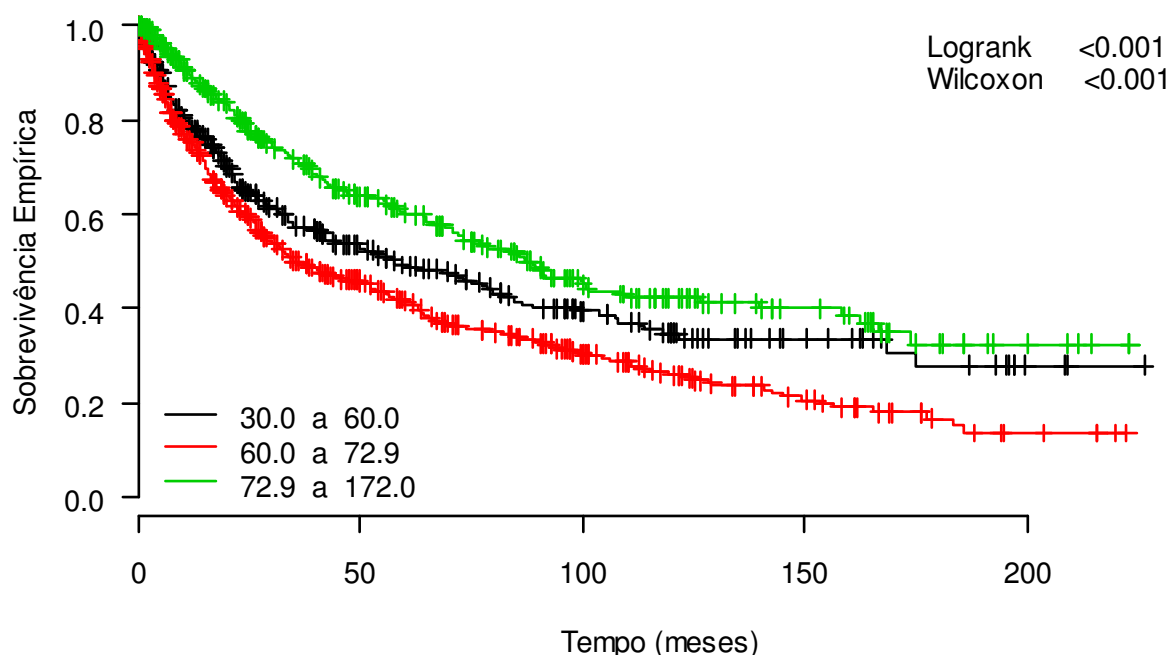
Figura A135: Curva de Kaplan-Meier por categoria de colesterol total imputado (mg/dl).**Figura A136:** Curva de Kaplan-Meier por categoria de altura imputada (cm).

Figura A137: Curva de Kaplan-Meier por categoria de peso imputado (kg).**Figura A138:** Curva de Kaplan-Meier por categoria de insuficiência tricúspide Doppler imputada.