

Projeto I

Bárbara Oliveira Ribeiro

Nathalia Gabriella Ferreira dos Santos

2023-10-04

1 Introdução

2 Objetivo

3 Metodologia

4 Resultados

4.1 Tamanho Amostral

Para o cálculo do tamanho da amostra onde o objetivo do estudo era comparar dois grupos: grupo tratamento com um controle, no contexto do trabalho o grupo controle seriam as mulheres que tiveram gestações saudáveis (*MN*) e o grupo tratamento seriam mulheres que tiveram pré-eclâmpsia (*MP*). As variâncias para os grupos serão denotados por σ_{MN}^2 e σ_{MP}^2 , respectivamente, e foram obtidas por meio de estudo piloto e repassado pela cliente para o estudo.

As variâncias para a **pressão sistólica** são:

- Grupo MN: 196,90
- Grupo MP: 514,95

Então, escolhendo um nível de significância de $\alpha = 0,05$ e um poder de teste de $\beta = 0,10$, tem-se que:

O *erro* aqui é chamado a diferença das médias do grupo MN e MP. Nota-se que quanto menor o erro, ou seja, quão próximo deseja-se que as médias dos grupos sejam “iguais”, maior será a amostra a ser avaliada para garantir boa estimação.

4.2 Curva de concentração do Sft1

Devido ao não registro de valores de *concentração* de Sft1 (provavelmente devido a baixa acurácia do equipamento na medição para valores pequenos de concentração), e tendo o resultado da *absorbância* observada, foi solicitado a estimação teórica dessas concentrações.

Tabela 1: Cálculo do tamanho amostral para diferentes erros

Erro	Tamanho Amostral
0.1	747.971
0.5	29.919
1.0	7.480
2.0	1.870
5.0	299
10.0	75

Tabela 2: Valores estimados dos parâmetros

	coefs.data
b	-1.3264
c	0.1229
d	10.9819
e	8.6645

Tabela 3: Intervalo de confiança para os parâmetros

	2.5 %	97.5 %
Slope:(Intercept)	-2.198	-0.455
Amin:(Intercept)	-0.054	0.299
Amax:(Intercept)	-41.195	63.159
log2ED50:(Intercept)	3.704	13.625

É conhecido por meio da literatura que essa concentração a ser estimada segue uma equação conhecida como **Four parameter logistic regression** (Modelo de Regressão Logística de 4 Parâmetros) a qual é muito usada para dados de ensaio imunoabsorvente enzimático - ELISA (Enzyme-linked immunosorbent assay). Cabe então, a estimação dos parâmetros da curva por meio da regressão específica do problema.

Esse modelo de regressão possui como variável resposta a **absorbância** e como variável explicativa a **concentração**. Entretanto, o objetivo final da estimação é predizer valores de concentração para uma absorbância conhecida. Assim, a princípio usou-se os dados teóricos disponibilizados para estimar os parâmetros. Foi buscado na literatura ajustes semelhantes e optou-se pela implementação *LL2.4* contida no pacote *drc* do R-Studio. A função *LL2.4* fornece a parametrização da função por meio da log-logística de quatro parâmetros, uma função de auto-inicialização (caso o algoritmo seja iterativo e parta de um “chute inicial”), e os nomes dos parâmetros. Como está sendo usada a transformação \log_2 para estimar os parâmetros, para conversão final da concentração estimada será aplicada a inversa da transformação para achar o valor real da concentração.

- **c** indica a inclinação da curva (Slope)
- **c** indica o limite inferior (Amin)
- **d** indica o limite superior (Amax)
- **e** indica o ponto de inflexão (ED50 que nessa parametrização é $\log_2\text{ED50}$)

O resultado encontrado foi:

Para testar a significância dos parâmetros, um intervalo de confiança foi obtido. Exceto o parâmetro de máximo e mínimo, conclui-se que todos demais são significativos.

Por meio do gráfico é possível notar o quão bem a curva se ajustou aos dados teóricos.

4.2.1 Análise de resíduos

A fim de validar alguns pressupostos da regressão e a partir do gráfico dos resíduos, nota-se que a homocedasticidade dos resíduos é válida.

De mesmo modo, sobre a hipótese nula de que os resíduos possuem distribuição normal, para o mesmo $\alpha = 0,05$, dado o p – valor observado, não rejeita-se a hipótese nula.

Logo conclui-se que a estimação parece se adequar bem aos dados e pode-se seguir para a estimação da concentração por meio desses parâmetros.

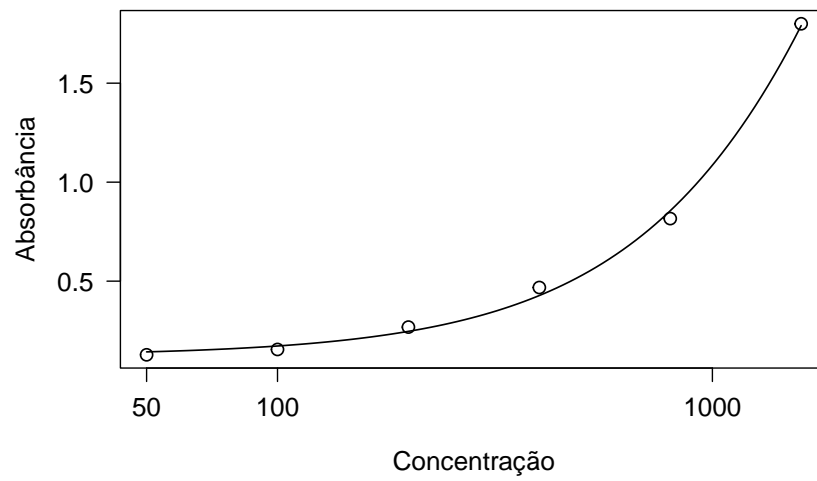


Figura 1: Gráfico da estimação da curva para os dados teóricos

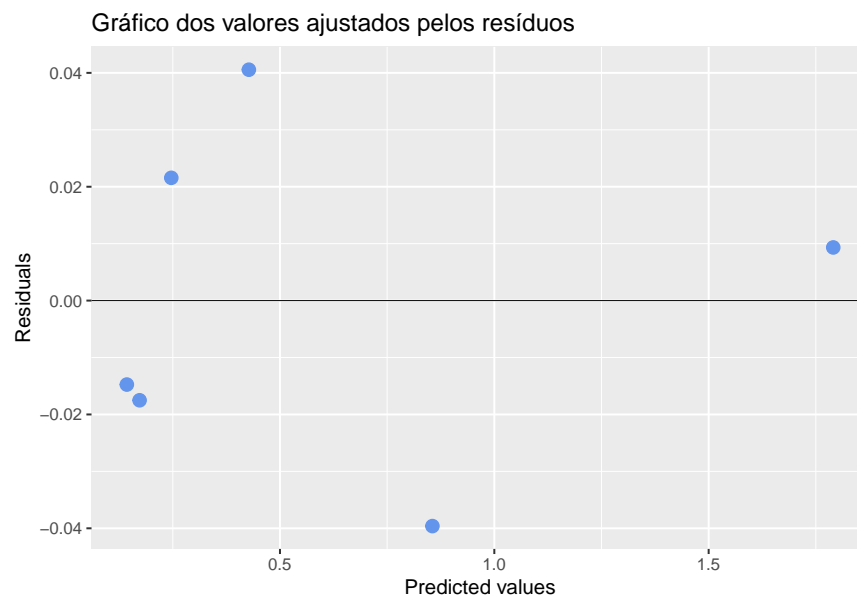


Figura 2: Gráfico dos valores ajustados pelos resíduos do modelo

Tabela 4: Resumo teste de normalidade dos resíduos

Nome	Método	Pvalor	Conclusão
Resíduos	Shapiro-Wilk normality test	0.9177	Não rejeita-se H0

Tabela 5: Comparativo das estimações para concentração de Sft1 abaixo da curva

Absorbância	Concentração Estimada	Ajuste Luiza
0.048	0.000	7.328
0.073	16.934	11.145
0.090	22.226	13.741
0.097	24.098	20.207
0.101	25.112	23.902
0.103	25.605	25.750
0.108	26.803	30.368
0.111	27.499	33.139
0.124	30.350	45.148
0.124	30.350	45.148
0.124	30.350	45.148
0.124	30.350	45.148
0.125	30.560	46.071
0.133	32.192	53.461
0.133	32.192	53.461
0.133	32.192	53.461
0.133	32.192	53.461
0.139	33.370	59.003
0.140	33.563	59.927
0.141	33.754	60.851
0.143	34.135	62.698
0.144	34.324	63.622
0.145	34.512	64.546
0.149	35.256	68.241
0.150	35.439	69.164
0.152	35.804	71.012

4.2.2 Estimação da concentração por imputação do limite inferior

Para estimação da concentração, como foi feita uma parametrização, a *Log2-concentração* será obtida e então convertida para escala padrão. Buscando-se na literatura referências, foi encontrado que além dos valores estimados, como deseja-se estimar concentrações abaixo do limite inferior, um novo candidato a mínimo deveria ser estipulado. Nesse caso, optou-se pelo mínimo ser o menor valor de concentração observado na amostra (0.048). Esse valor está contido no Intervalo de Confiança para o mínimo. Os demais coeficientes serão os mesmos estimados.

As concentrações encontradas podem ser vistas na tabela que segue:

Por fim, nota-se que os valores preditos parecem estar próximos do esperado para o comportamento dos dados teóricos.

4.3 Análises descritivas por grupo

A seguir, será apresentada algumas medidas de resumo para as variáveis de interesse contidas no banco de dados. Essas descritivas foram levantadas para avaliar principalmente o comportamento da variável em cada grupo, *Grupo MN* e *Grupo MP*.

Em síntese, as medidas descritivas escolhidas para os dados contínuos refletem:

- Boxplot: comparativo do comportamento dos dados para cada grupo e se há informações discrepantes;
- Tabela: medidas de resumos de tendência central e variabilidade, indicativo de nulos e número de

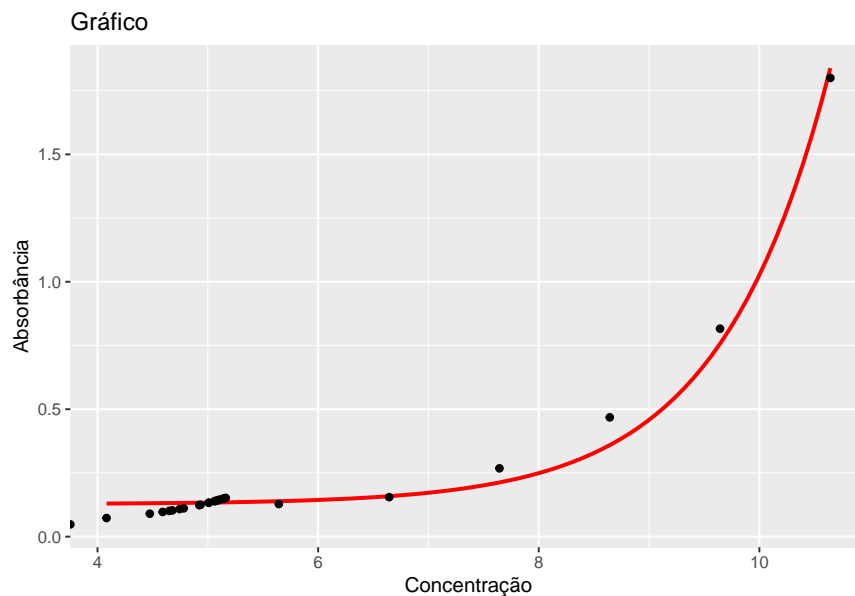


Figura 3: Concentração preditas com dados teóricos

Tabela 6: Tabela da proporção da variavel por grupo

Grupo	Branca	Negra	Parda
Grupo MN	46.53%	13.86%	39.60%
Grupo MP	35.71%	19.05%	45.24%

observações encontradas e

- Histograma: a frequência geral da variável de interesse.

Caso fossem encontrados dados extremos, visando melhores estimativas, eles foram retirados das medidas de resumo.

Para os dados categóricos, a medida descritiva escolhida foi um gráfico de barras para avaliar a proporção de cada grupo (*MN* e *MP*) dentro de cada categoria. Dentre algumas variáveis categóricas, foi gerado uma tabela de frequência para avaliar a proporção da categoria dentro de cada grupo (*MN* e *MP*).

Tabela 7: Tabela da proporção da variavel por grupo

Grupo	4ª série	Ensino Médio	Ensino técnico	Fundamental	Pós-graduação	Superior
Grupo MN	1.98%	14.85%	4.95%	0.00%	34.65%	43.56%
Grupo MP	5.88%	23.53%	5.88%	4.71%	34.12%	25.88%

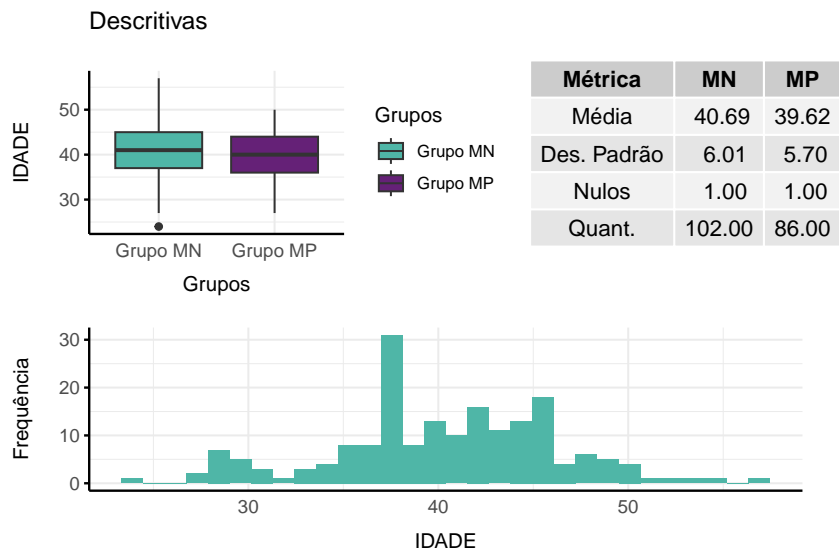


Figura 4: Análise descritiva por grupo para a variável idade

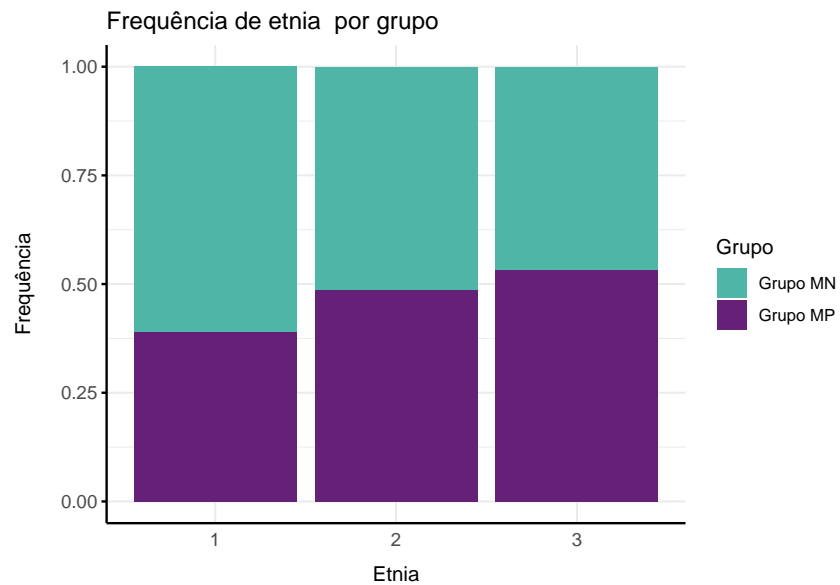


Figura 5: Análise descritiva por grupo para a variável etnia

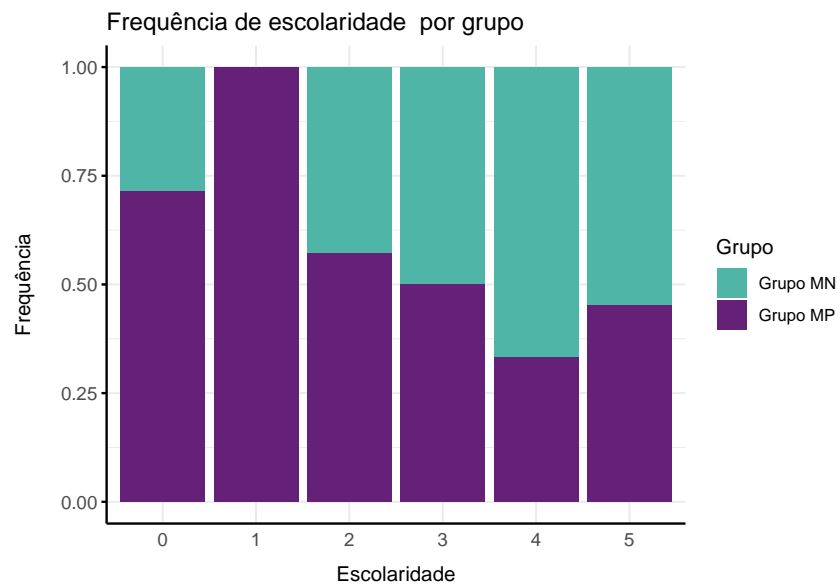


Figura 6: Análise descritiva por grupo para a variável escolaridade

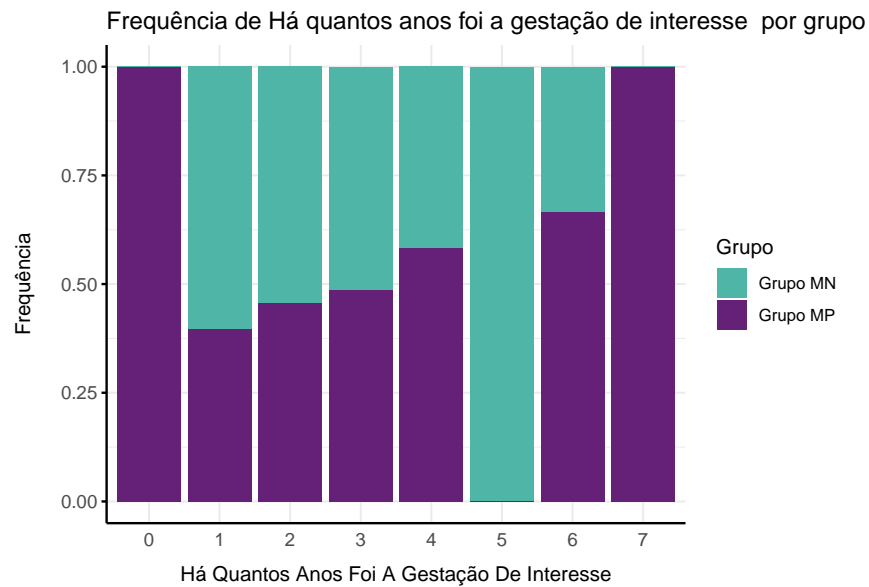


Figura 7: Análise descritiva por grupo para a variável há quantos anos foi a gestação de interesse

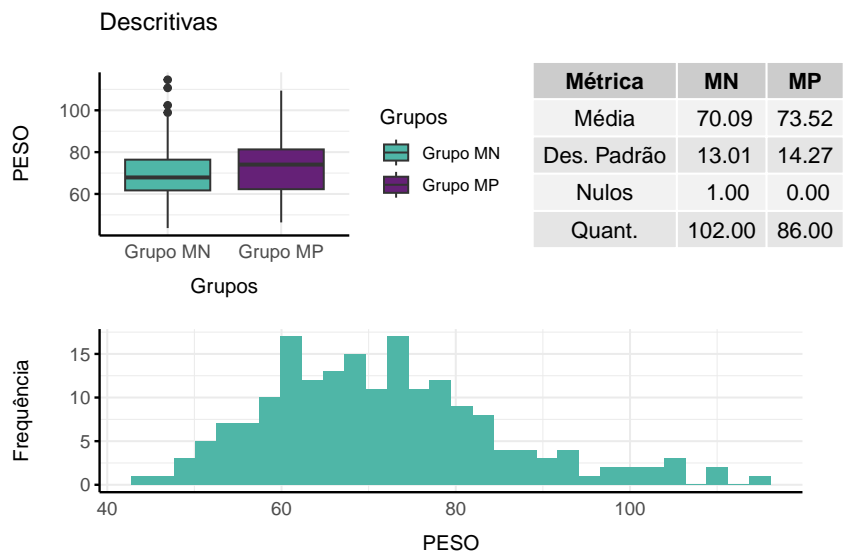


Figura 8: Análise descritiva por grupo para a variável peso

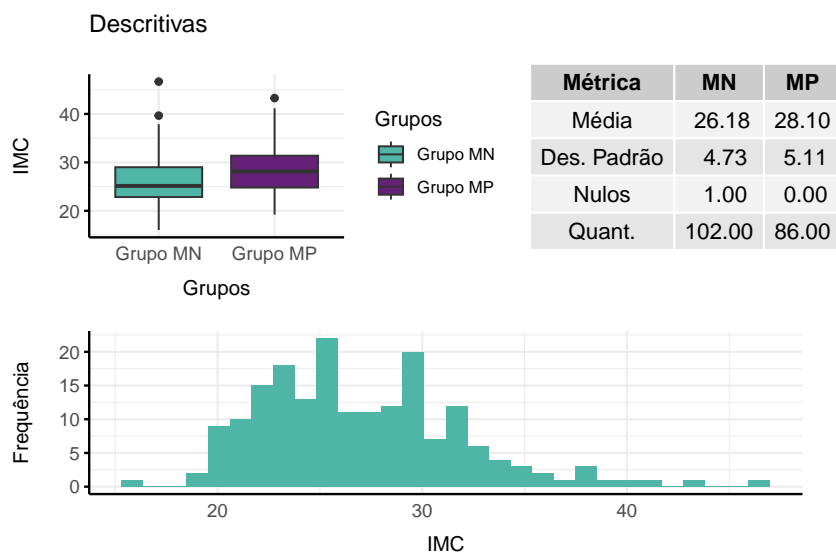


Figura 9: Análise descritiva por grupo para a variável IMC

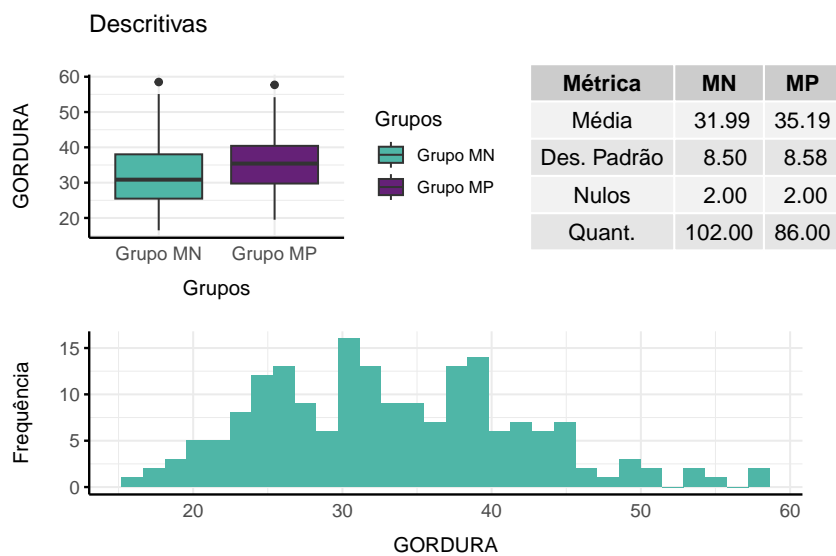


Figura 10: Análise descritiva por grupo para a variável gordura

Tabela 8: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	2.29	1.96
Des. Padrão	3.58	0.71
Nulos	0.00	0.00
Quant.	188.00	182.00

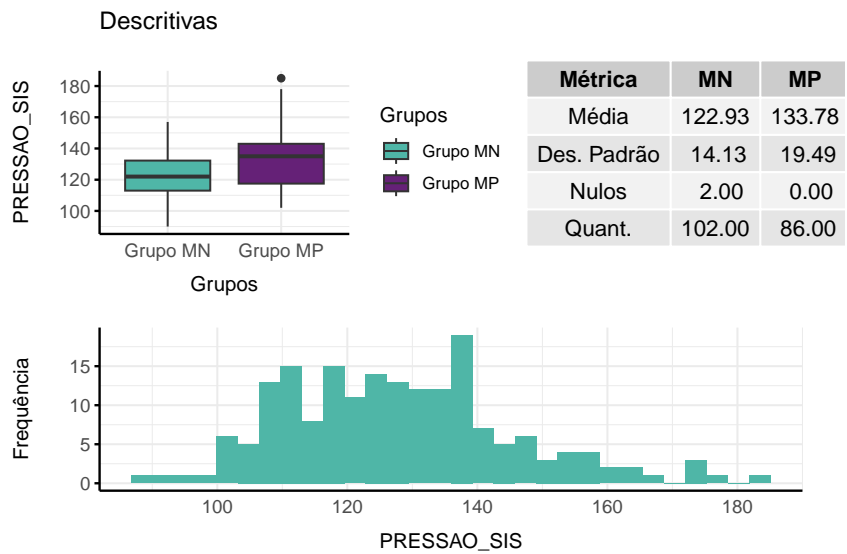


Figura 11: Análise descritiva por grupo para a variável pressão sistólica

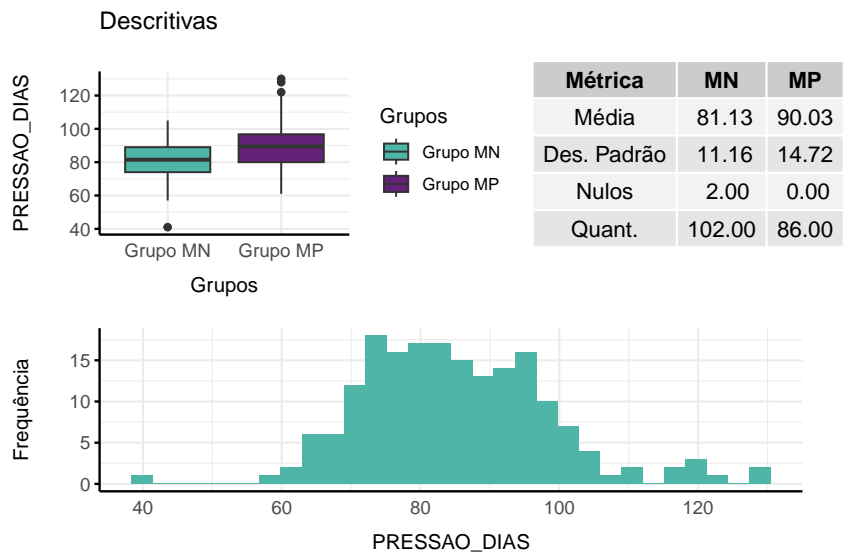


Figura 12: Análise descritiva por grupo para a variável pressão diastólica

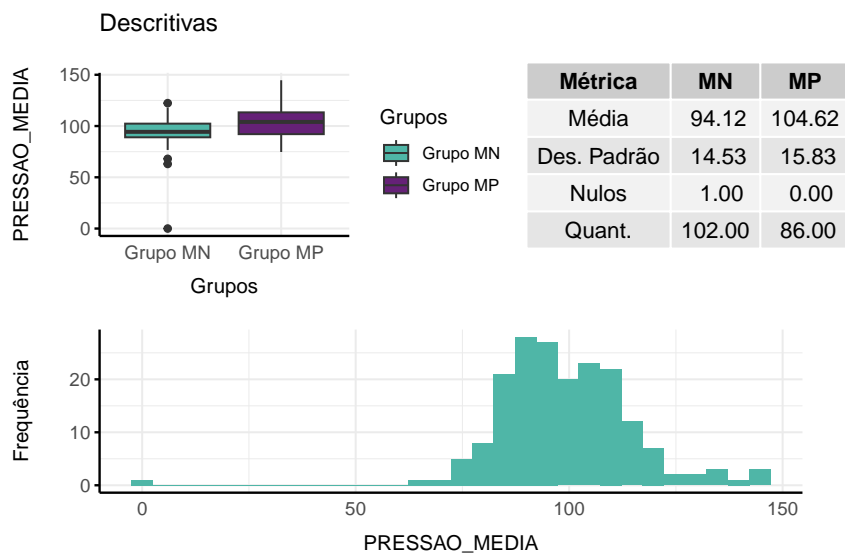


Figura 13: Análise descritiva por grupo para a variável pressão média

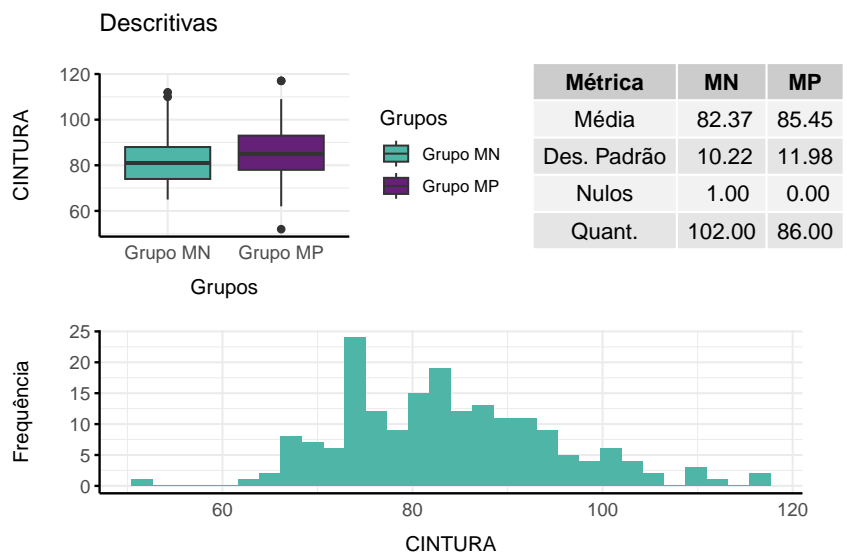


Figura 14: Análise descritiva por grupo para a variável cintura

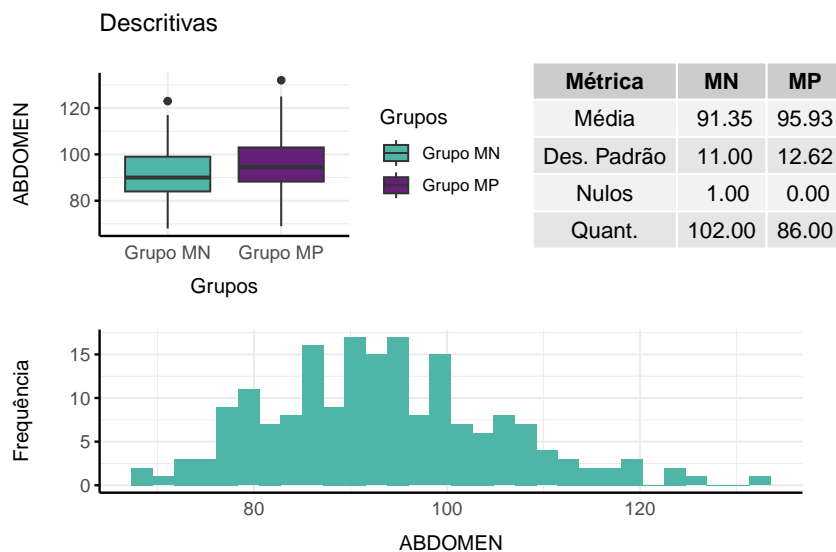


Figura 15: Análise descritiva por grupo para a variável circunferência do abdômen

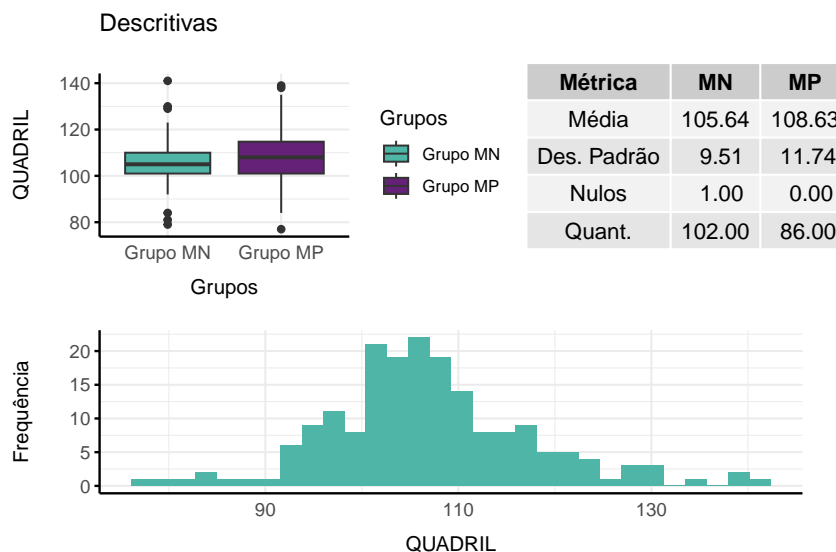


Figura 16: Análise descritiva por grupo para a variável quadril

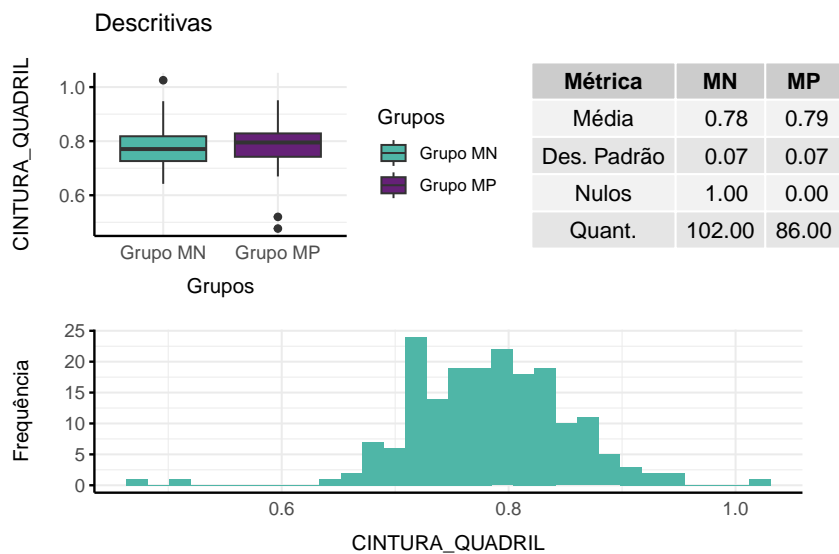


Figura 17: Análise descritiva por grupo para a variável de relação da cintura vs quadril

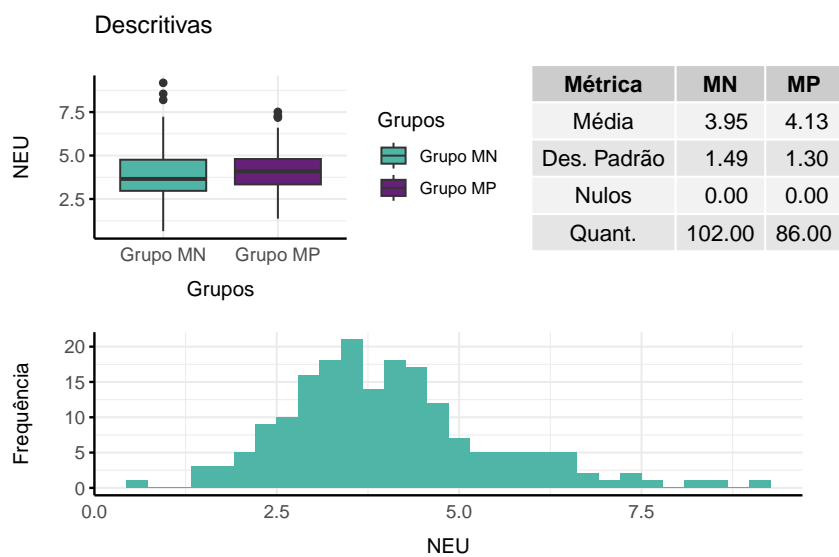


Figura 18: Análise descritiva por grupo para a variável neutrófilos

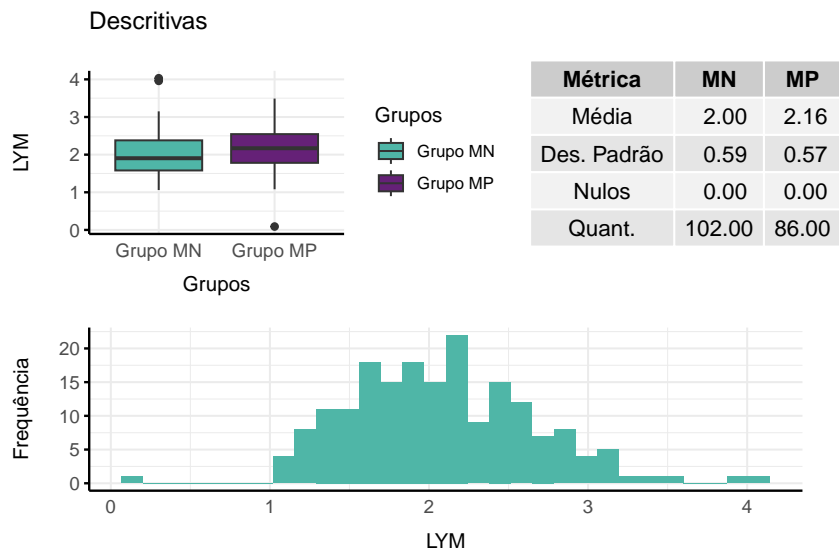


Figura 19: Análise descritiva por grupo para a variável linfócitos

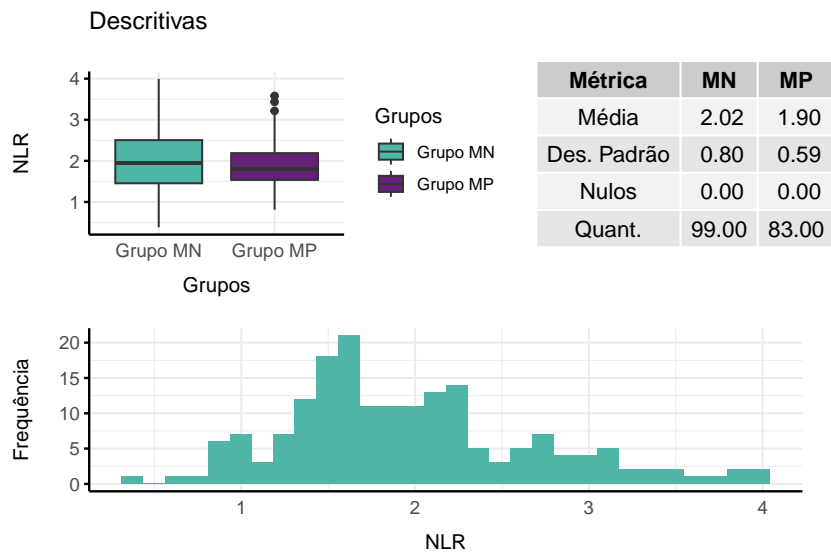


Figura 20: Análise descritiva por grupo para a variável NLR

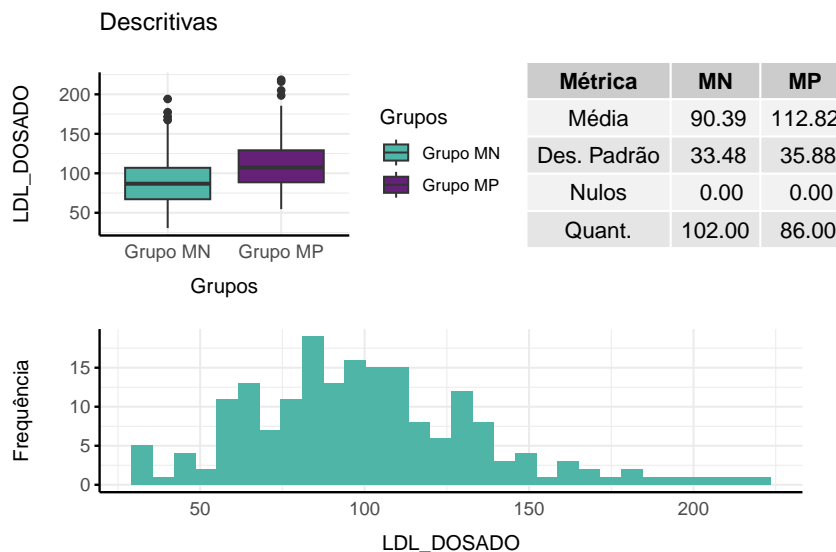


Figura 23: Análise descritiva por grupo para a variável LDL dosado

- 4.3.1 Idade:
- 4.3.2 Etnia:
- 4.3.3 Escolaridade:
- 4.3.4 Há quantos anos foi a gestação de interesse :
- 4.3.5 Peso da mulher:
- 4.3.6 IMC:
- 4.3.7 Percentual de gordura:
- 4.3.8 Pressão Sistólica:
- 4.3.9 Pressão Diastólica:
- 4.3.10 Pressão arterial média:
- 4.3.11 Circunferência de cintura:
- 4.3.12 Circunferência de abdômen:
- 4.3.13 Circunferência de quadril:
- 4.3.14 Relação cintura vs quadril:
- 4.3.15 Neutrófilos:
- 4.3.16 Linfócitos:
- 4.3.17 Razão neutrófilo/linfócito NLR:
- 4.3.18 Colesterol Total:
- 4.3.19 HDL:
- 4.3.20 LDL dosado:
- 4.3.21 VLDL:
- 4.3.22 Não-HDL:
- 4.3.23 Triglicérides:
- 4.3.24 Hemoglobina glicada HbA1c:
- 4.3.25 Creatinina:

Tabela 9: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	21.61	20.46
Des. Padrão	11.20	6.50
Nulos	0.00	0.00
Quant.	188.00	184.00

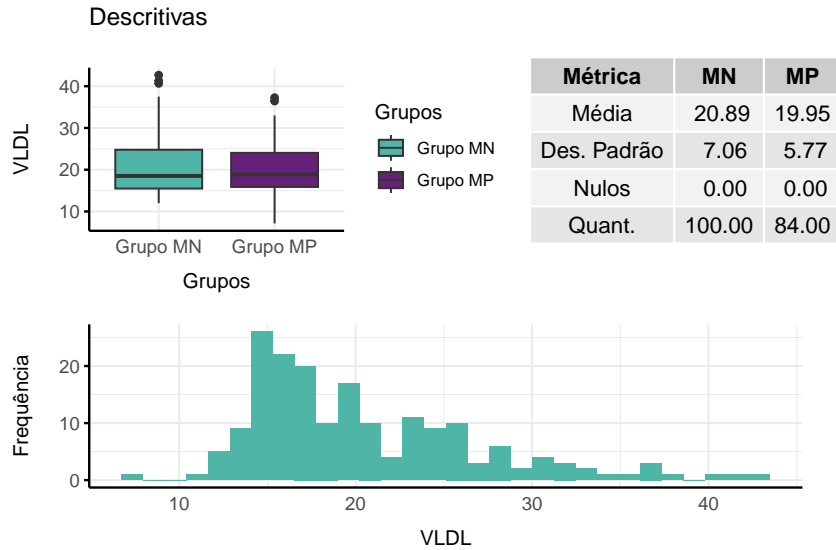


Figura 24: Análise descritiva por grupo para a variável VLDL

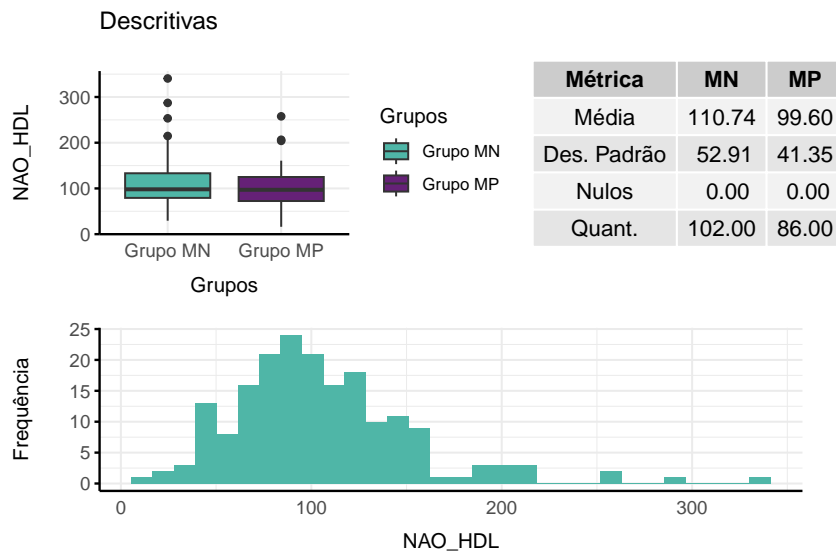


Figura 25: Análise descritiva por grupo para a variável Não-HDL

Tabela 10: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	121.43	112.49
Des. Padrão	84.57	55.37
Nulos	0.00	0.00
Quant.	188.00	183.00

Descritivas

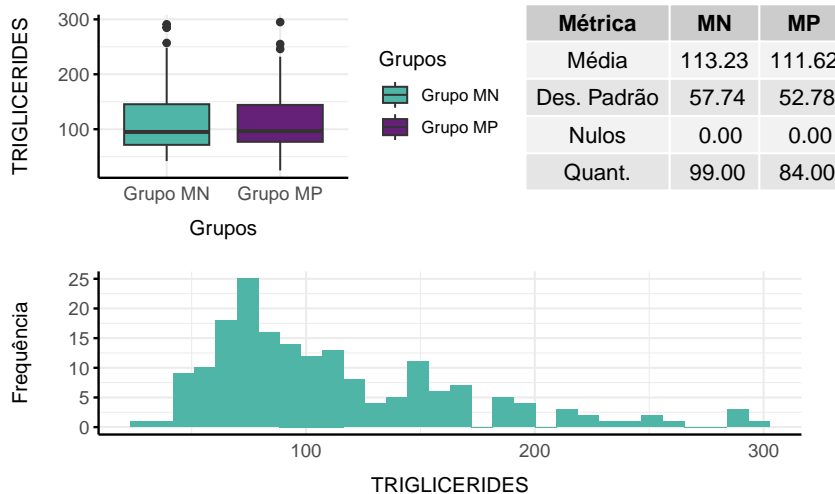


Figura 26: Análise descritiva por grupo para a variável triglicérides

Tabela 11: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	4.78	4.64
Des. Padrão	0.98	0.66
Nulos	44.00	0.00
Quant.	188.00	137.00

Tabela 12: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	4.78	3.48
Des. Padrão	0.98	0.30
Nulos	44.00	0.00
Quant.	188.00	18.00

Tabela 13: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	989.28	121.89
Des. Padrão	11305.86	69.95
Nulos	4.00	0.00
Quant.	188.00	179.00

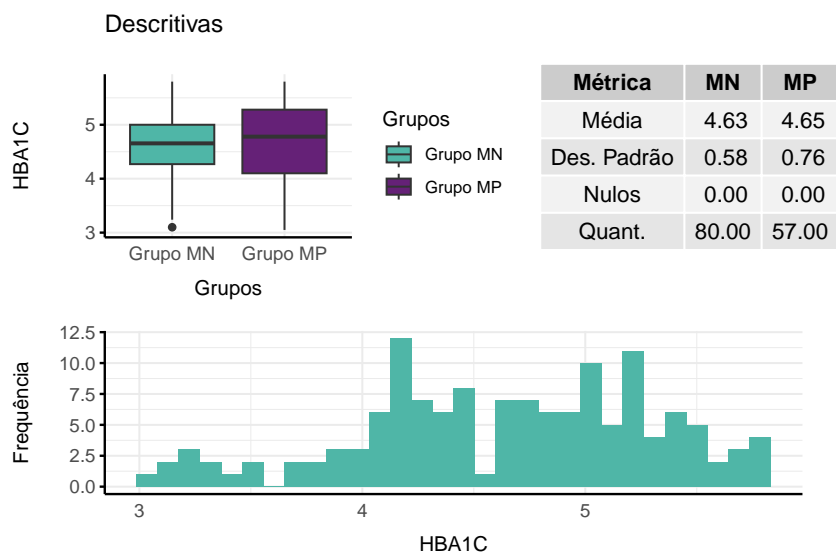


Figura 27: Análise descritiva por grupo para a variável HbA1c

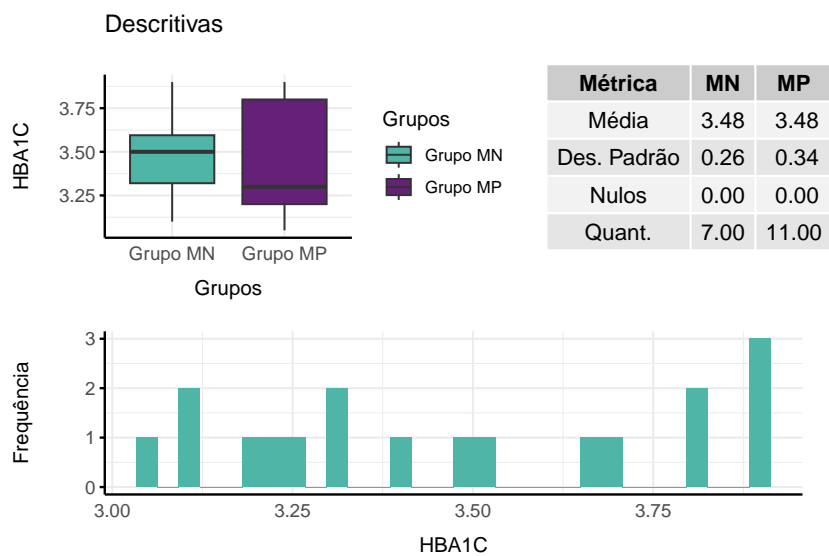


Figura 28: Análise descritiva por grupo para a variável creatinina

Tabela 14: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	4.88	4.66
Des. Padrão	1.62	1.05
Nulos	0.00	0.00
Quant.	188.00	180.00

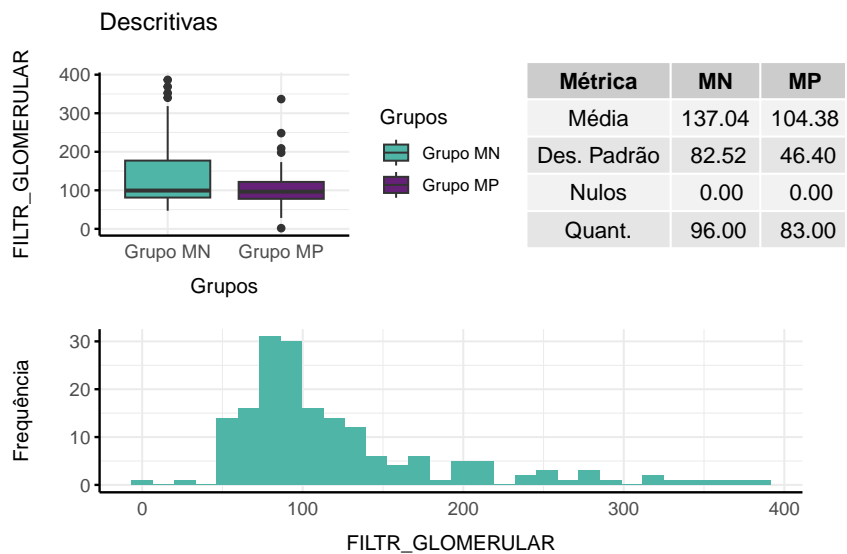


Figura 29: Análise descritiva por grupo para a variável filtração glomerular

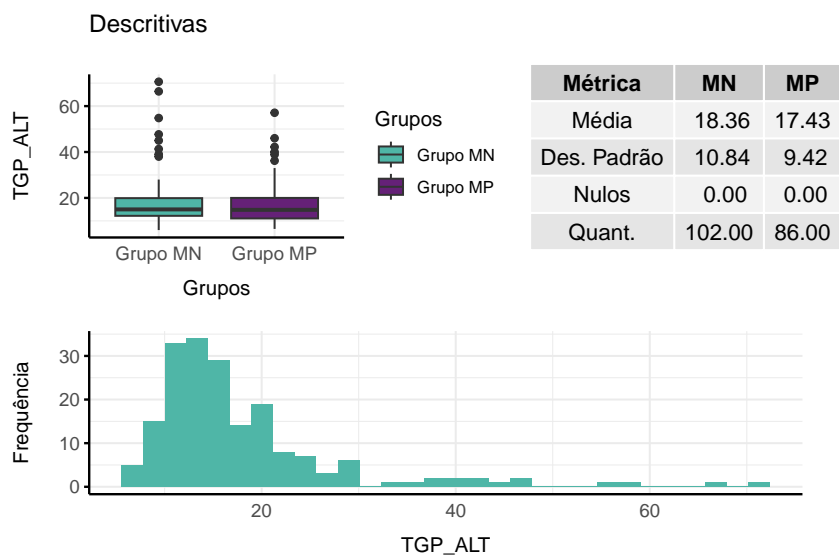


Figura 30: Análise descritiva por grupo para a variável TGP/ALT

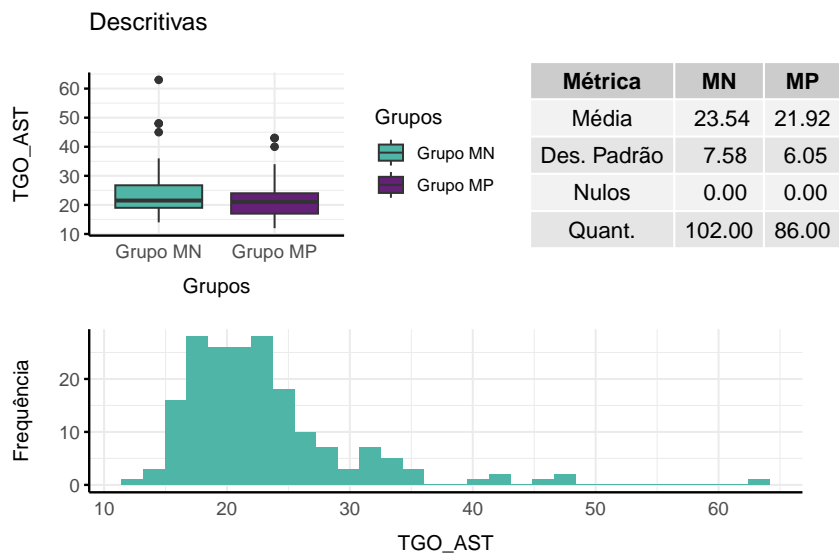


Figura 31: Análise descritiva por grupo para a variável TGO/AST

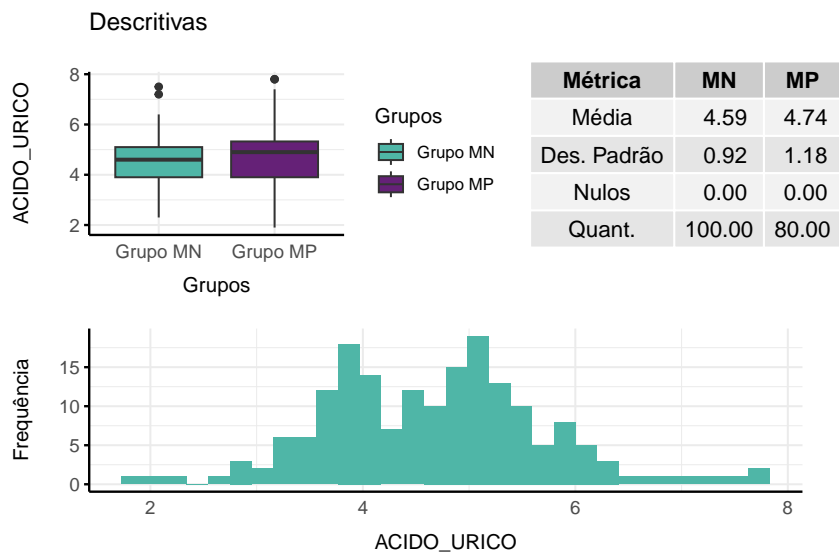


Figura 32: Análise descritiva por grupo para a variável ácido úrico

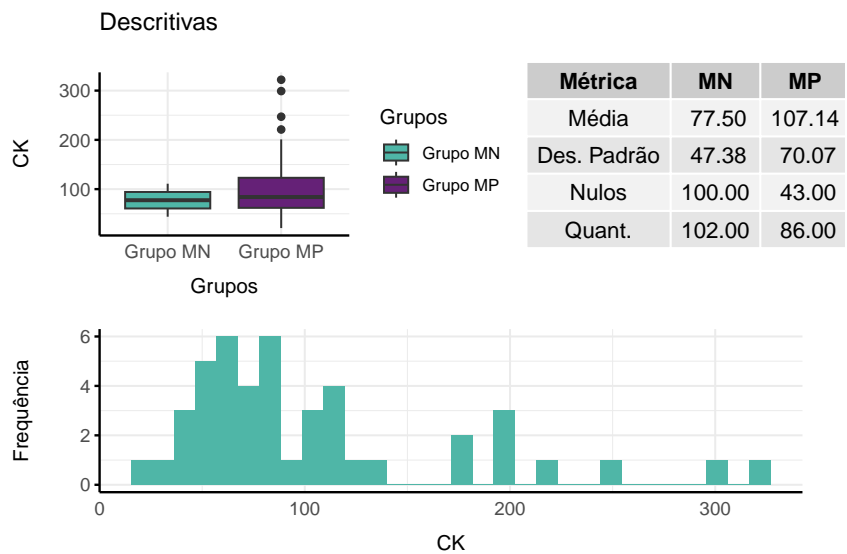


Figura 33: Análise descritiva por grupo para a variável CK

Tabela 15: Quantidade de nulos e não nulos por grupo de CK

grupo	Nulos	Nao nulos
0	100	2
1	43	43

Tabela 16: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	29.79	23.78
Des. Padrão	22.27	8.23
Nulos	0.00	0.00
Quant.	188.00	170.00

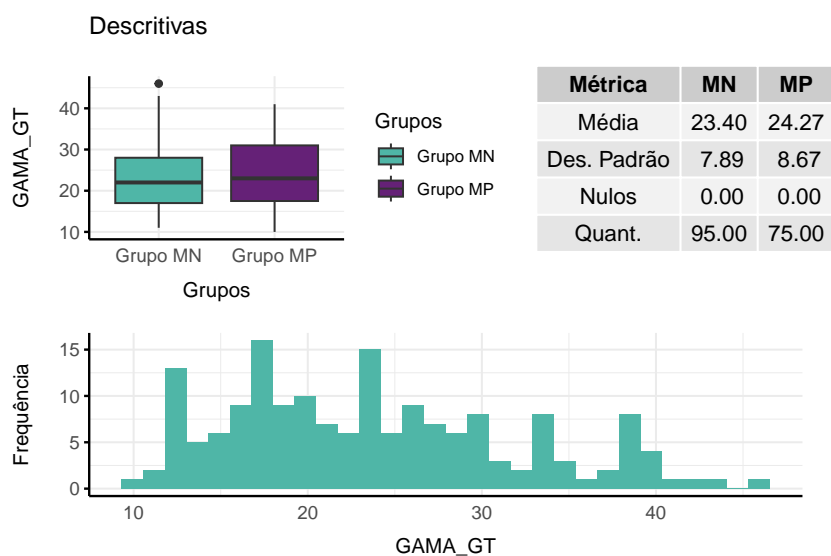


Figura 34: Análise descritiva por grupo para a variável Gama GT

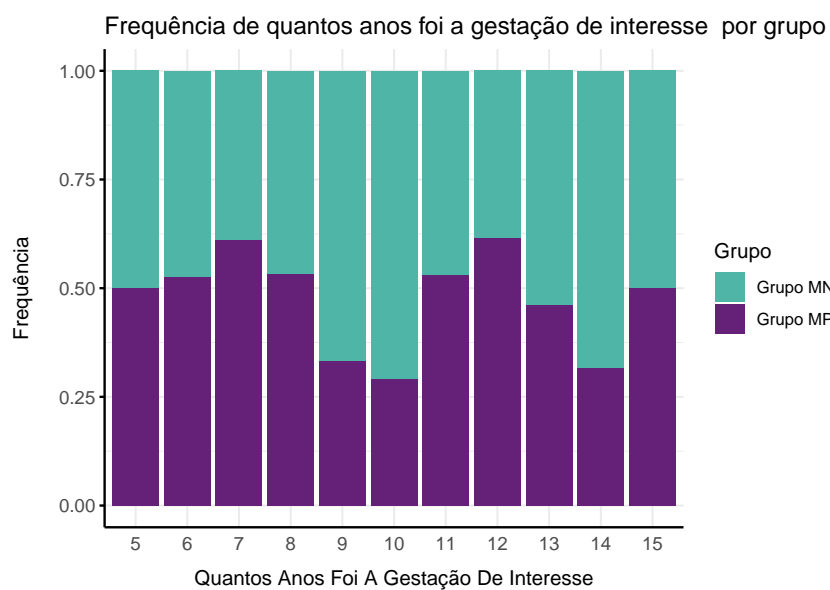


Figura 35: Análise descritiva por grupo para a variável quantos anos passou-se do evento

Tabela 17: Tabela da proporção da variavel por grupo

Grupo	Cesariana	Fórceps	Induzido	Normal
Grupo MN	57.58%	3.03%	0.00%	39.39%
Grupo MP	76.19%	1.19%	1.19%	21.43%

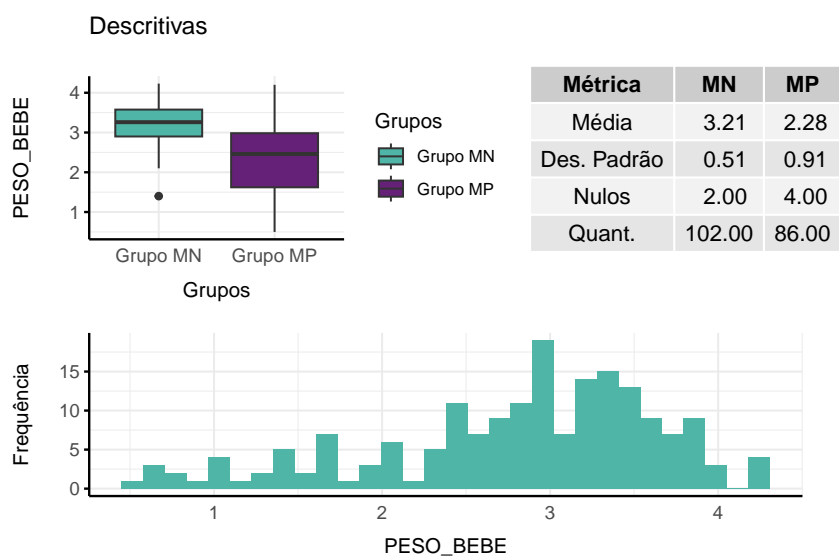


Figura 36: Análise descritiva por grupo para a variável peso do bebê

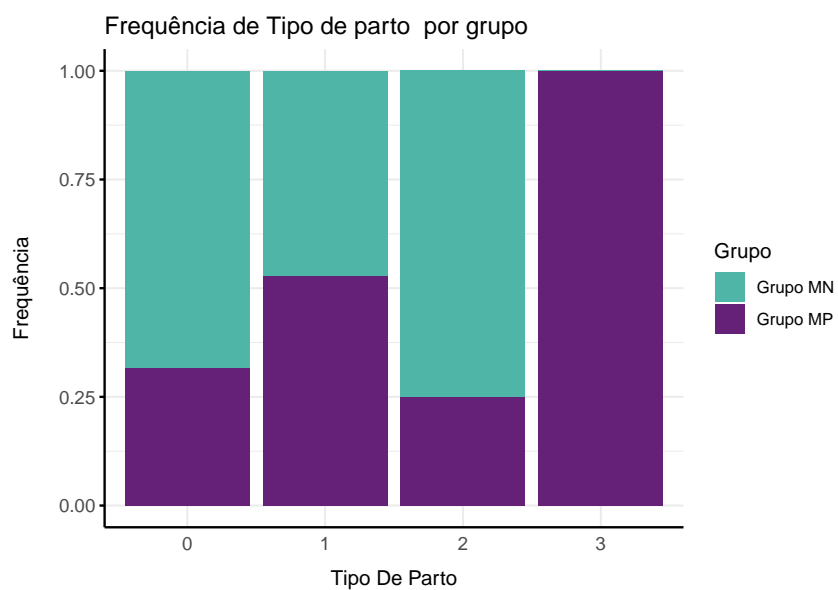


Figura 37: Análise descritiva por grupo para a variável tipo de parto

Tabela 18: Tabela da proporção da variável por grupo

Grupo	Não	Sim
Grupo MN	90.10%	9.90%
Grupo MP	49.41%	50.59%

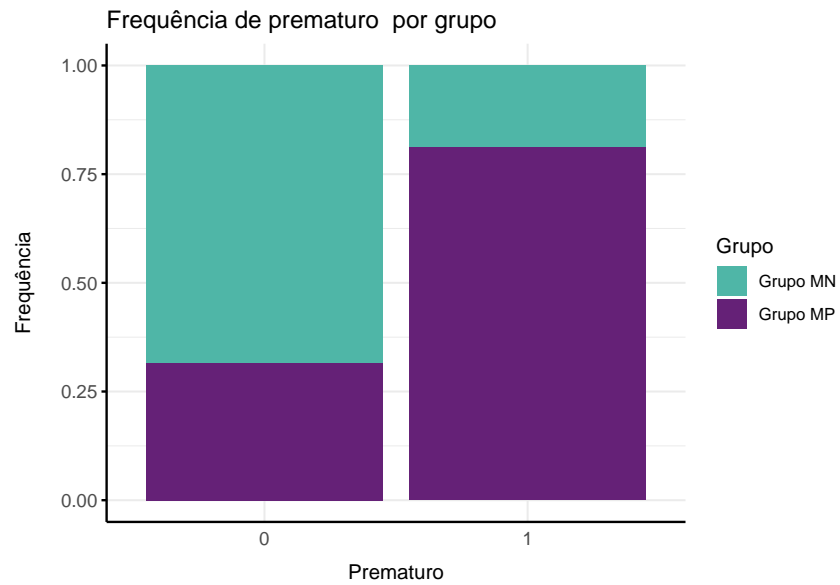


Figura 38: Análise descritiva por grupo para a variável prematuro

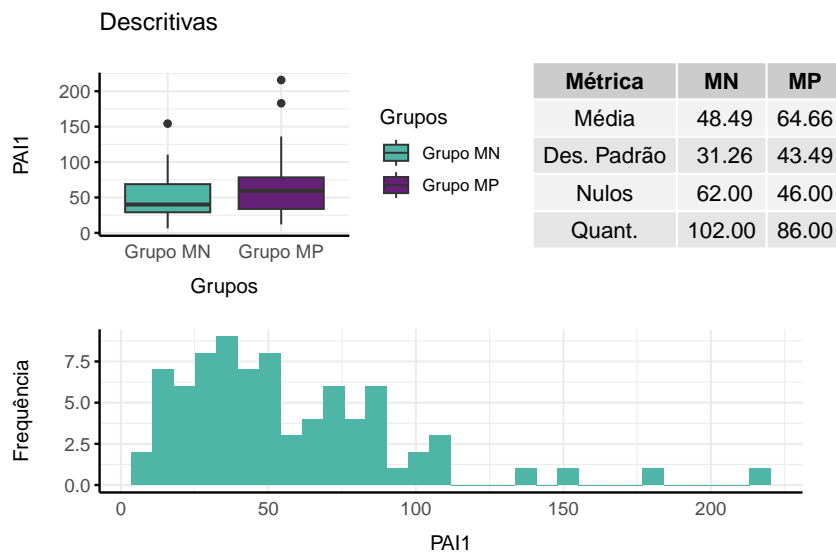


Figura 39: Análise descritiva por grupo para a variável PAI-1

Tabela 19: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	103.86	79.88
Des. Padrão	122.70	35.44
Nulos	111.00	0.00
Quant.	188.00	73.00

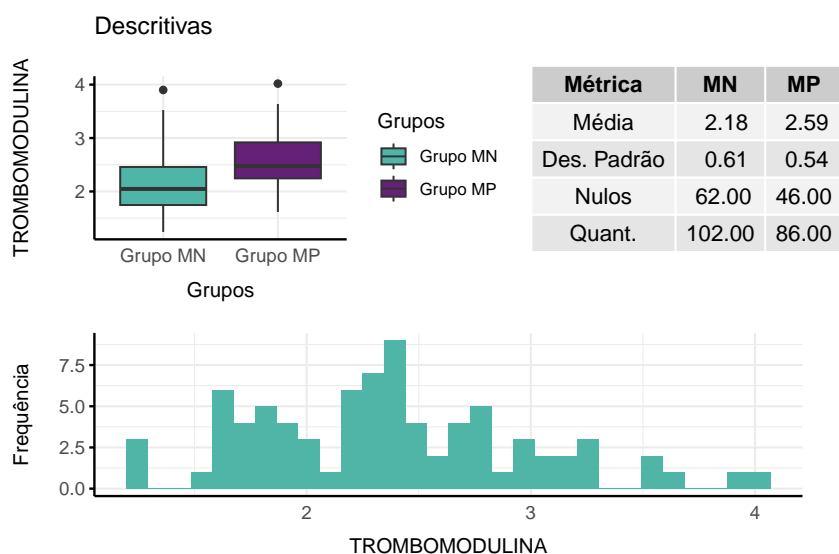


Figura 40: Análise descritiva por grupo para a variável trombomodulina

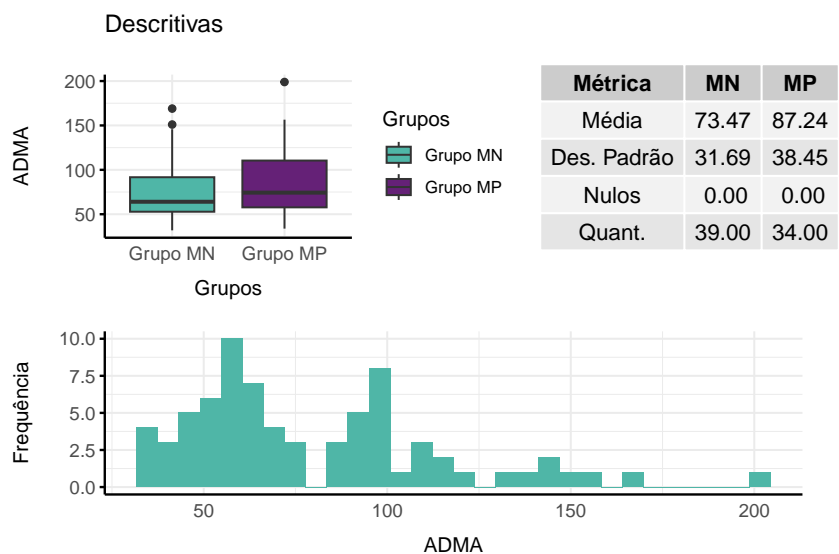


Figura 41: Análise descritiva por grupo para a variável ADMA

Tabela 20: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	172.02	62.47
Des. Padrão	877.84	37.94
Nulos	121.00	0.00
Quant.	188.00	65.00

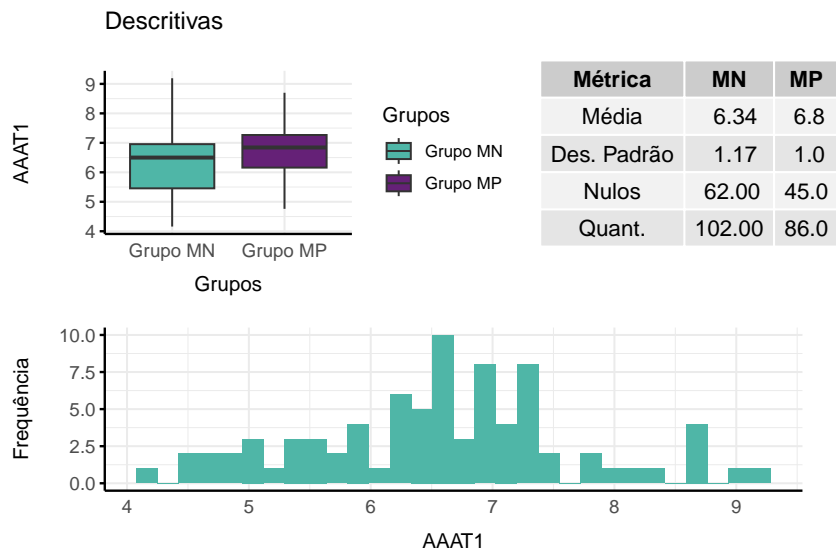


Figura 42: Análise descritiva por grupo para a variável AA-AT1

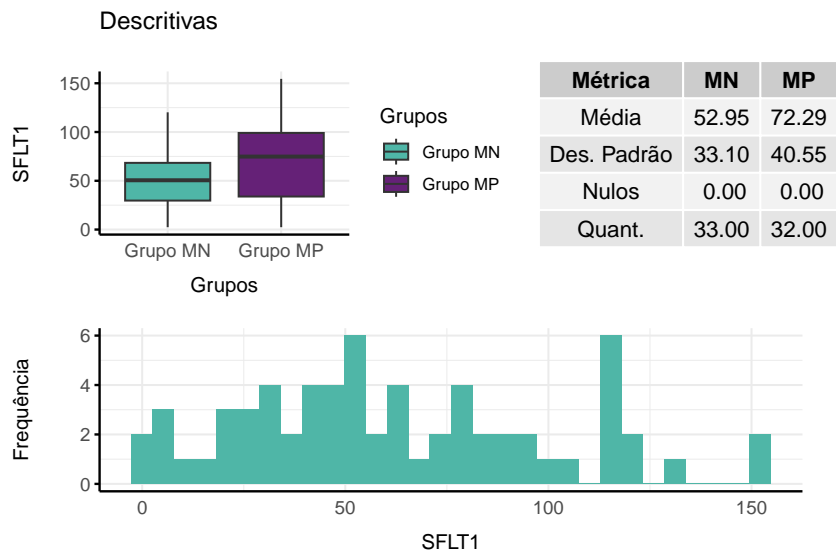


Figura 43: Análise descritiva por grupo para a variável sflt1 sem previsões

Tabela 21: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	132.09	52.97
Des. Padrão	746.32	35.60
Nulos	99.00	0.00
Quant.	192.00	91.00

4.3.31 Gama GT:

4.3.32 Há quantos anos foi a gestação de interesse

4.3.33 Peso do bebê:

4.3.34 Tipo Parto

4.3.35 Prematuro

4.3.36 PAI-1 ng/mL:

4.3.37 Trombomodulina ng/mL:

4.3.38 ADMA ng/mL:

4.3.39 AA-AT1 ng/mL:

4.3.40 Sflt1 sem previsões

4.3.41 Sflt1 com previsões feitas pelo ajuste da curva

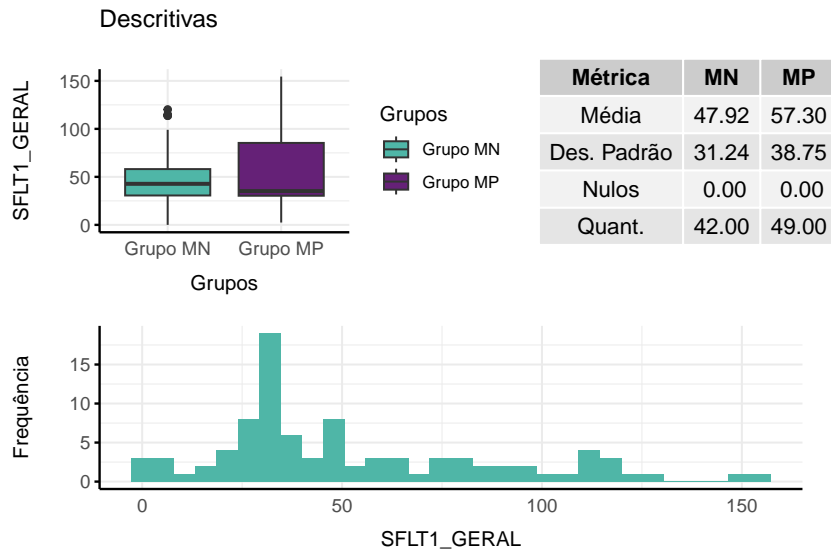


Figura 44: Análise descritiva por grupo para a variável Sflt1 com previsões feitas pelo ajuste da curva

Tabela 22: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	129.75	56.97
Des. Padrão	715.71	33.07
Nulos	99.00	0.00
Quant.	200.00	99.00

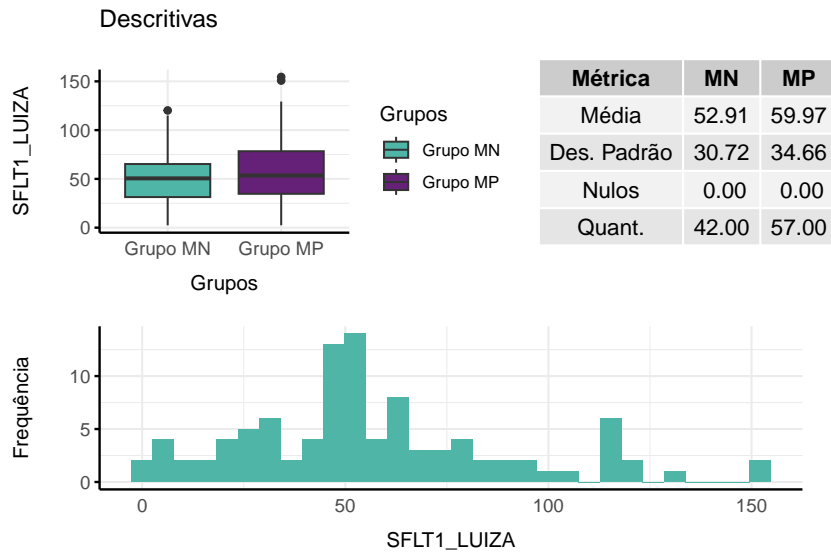


Figura 45: Análise descritiva por grupo para a variável Sflt1 com previsões feitas pelo método da Luiza

Tabela 23: Resumo da retirada

Métrica	Antes.da.retirada	Após.a.retirada
Média	334.60	308.53
Des. Padrão	301.24	195.52
Nulos	337.00	0.00
Quant.	415.00	77.00

Tabela 24: Tabela da proporção da variavel por grupo

Grupo	Não	Sim
Grupo MN	97.03%	2.97%
Grupo MP	74.42%	25.58%

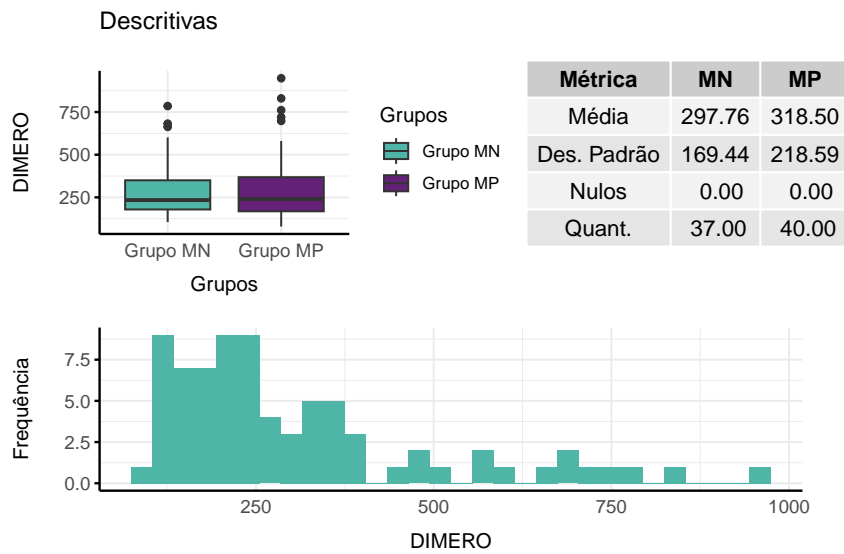


Figura 46: Análise descritiva por grupo para a variável dimero

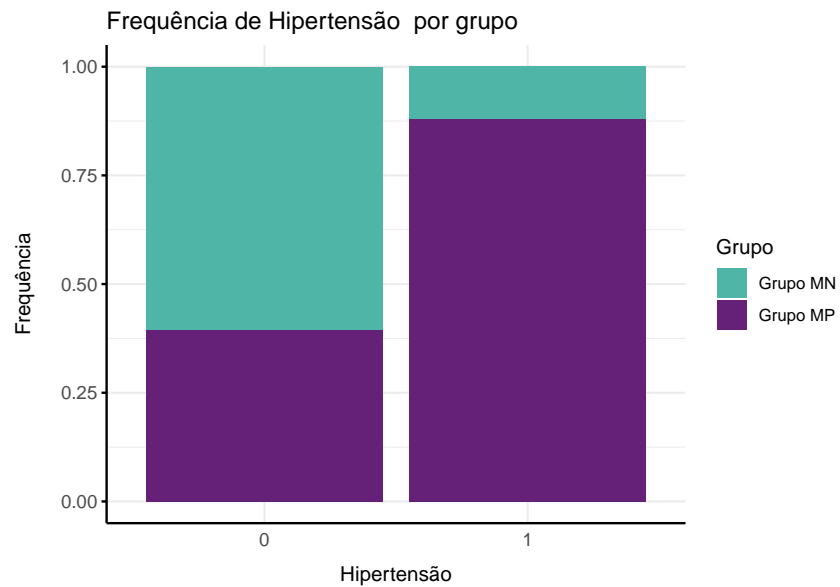


Figura 47: Análise descritiva por grupo para a variável hipertensão

Tabela 25: Tabela da proporção da variavel por grupo

Grupo	Não	Sim
Grupo MN	96.04%	3.96%
Grupo MP	70.93%	29.07%

Tabela 26: Tabela da proporção da variavel por grupo

Grupo	Não	Sim
Grupo MN	99.01%	0.99%
Grupo MP	96.51%	3.49%

4.3.42 Sft1 com previsões feitas pelo método da Luiza

4.3.43 Dímero-D ng/mL:

4.3.44 Hipertensão

4.3.45 Anti-hipertensivo

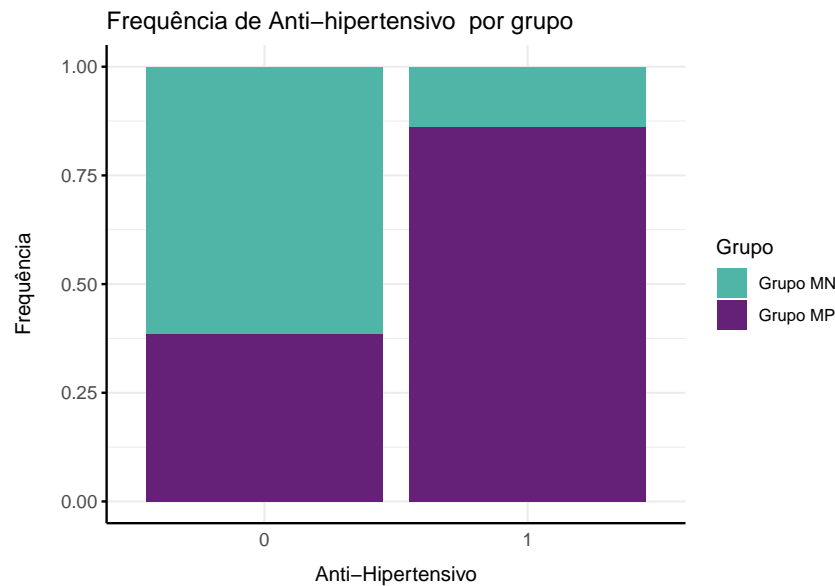


Figura 48: Análise descritiva por grupo para a variável anti-hipertensivo

4.3.46 Diabetes

4.3.47 Medicamento para diabetes

4.3.48 Hipercolesterolemia

4.3.49 Medicamento para hipercolesterolemia

4.3.50 Fumantes

4.4 Resumo testes de hipóteses

Para os dados numéricos, os testes tiveram por hipótese geral:

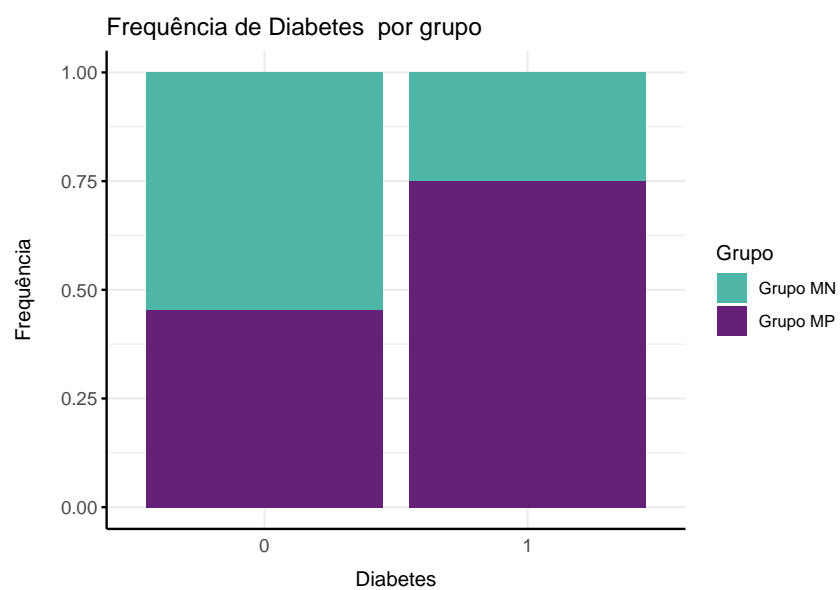


Figura 49: Análise descritiva por grupo para a variável diabetes

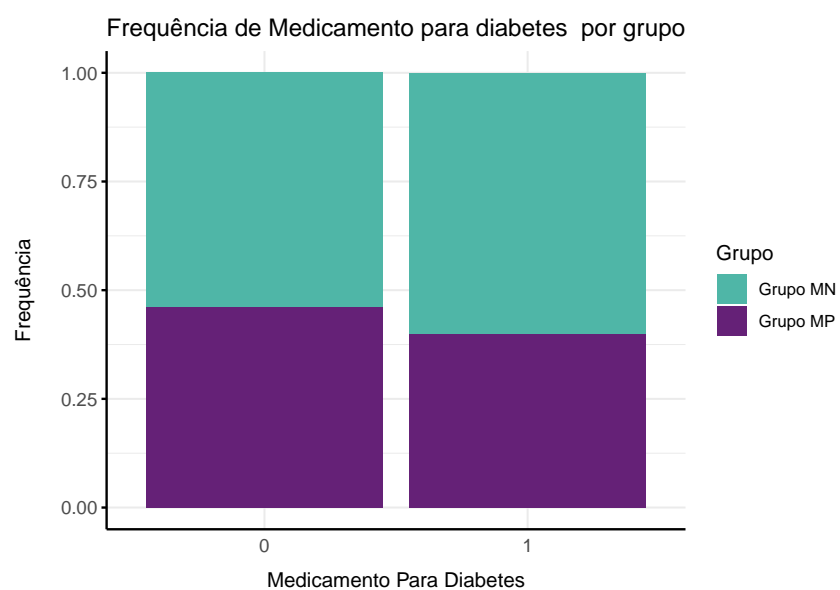


Figura 50: Análise descritiva por grupo para a variável medicamento para diabetes

Tabela 27: Tabela da proporção da variável por grupo

Grupo	Não	Sim
Grupo MN	97.03%	2.97%
Grupo MP	97.67%	2.33%

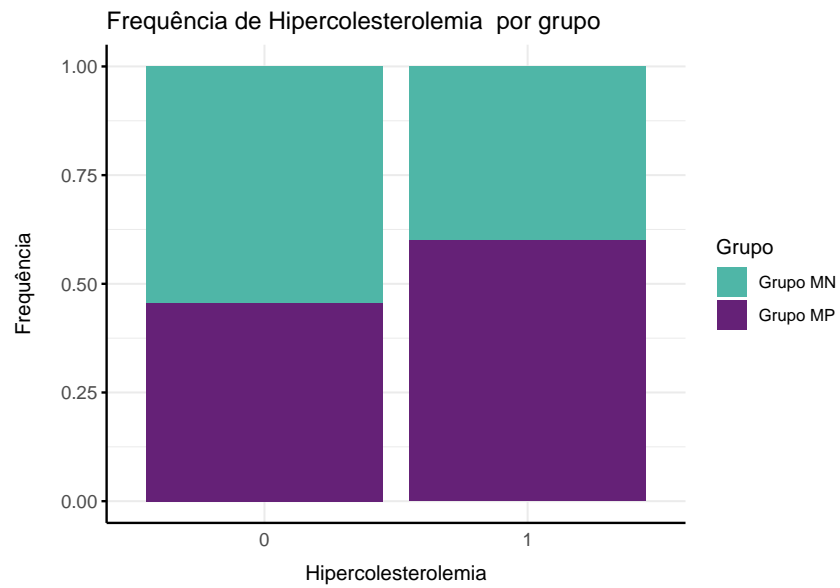


Figura 51: Análise descritiva por grupo para a variável hipercolesterolemia

Tabela 28: Tabela da proporção da variavel por grupo

Grupo	Não	Sim
Grupo MN	98.02%	1.98%
Grupo MP	96.51%	3.49%

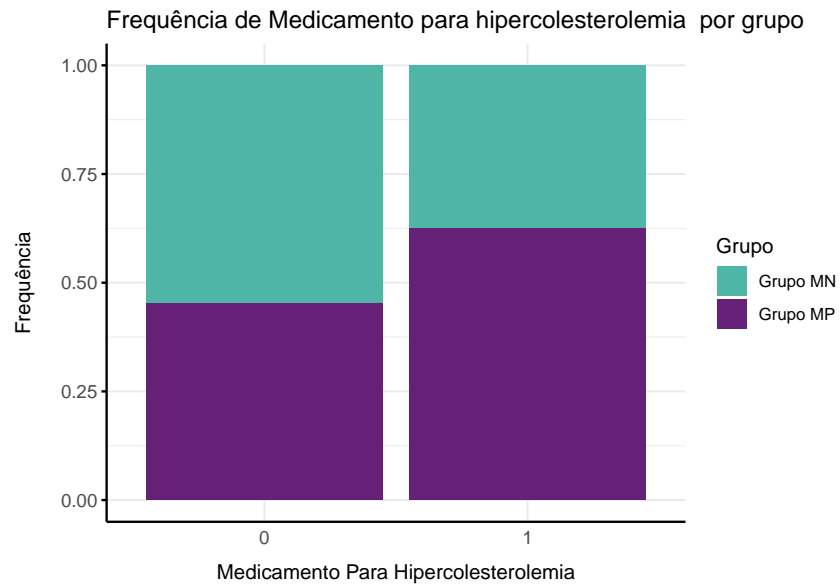


Figura 52: Análise descritiva por grupo para a variável medicamento para hipercolesterolemia

Tabela 29: Tabela da proporção da variavel por grupo

Grupo	Não	Sim
Grupo MN	97.03%	2.97%
Grupo MP	94.19%	5.81%

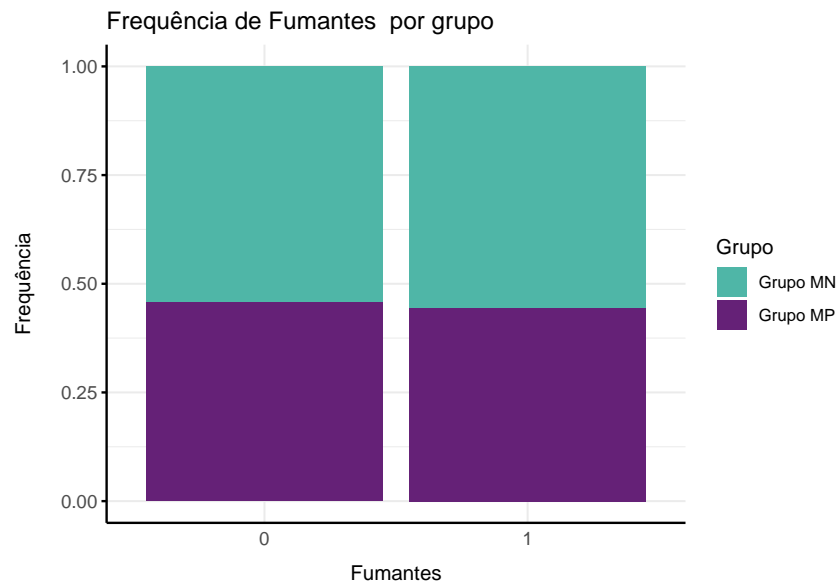


Figura 53: Análise descritiva por grupo para a variável fumantes

Tabela 30: Tabela da proporção da variavel por grupo

Grupo	Não	Sim
Grupo MN	95.05%	4.95%
Grupo MP	95.29%	4.71%

$$\begin{cases} H_0 : \mu_{Grupo\ MN} = \mu_{Grupo\ MP} \\ H_1 : \mu_{Grupo\ MN} \neq \mu_{Grupo\ MP} \end{cases}$$

Caso os dados atendessem a suposição de normalidade para os grupos, foi aplicado o *Teste T* para comparação da média de dois grupos. Caso contrário, o *Teste de Permutação* foi aplicado como teste não paramétrico. Em suma, deseja-se avaliar se a média entre os grupos é igual para uma dada variável ou se há diferença.

Para as variáveis qualitativas, o Teste Exato de Fisher para tabelas de contingência $2 \cdot m$ foi aplicado sobre a hipótese nula de que a proporção das categorias fossem iguais para ambos grupos (*MN* e *MP*).

Em seguida, será apresentado o resumo de todos testes de hipóteses aplicados as variáveis.

Resumo Descritivas favoráveis para o estudo sobre as variáveis que impactam na média ou frequência dentro do grupo *Grupo MN* e *Grupo MP* para um nível de significância de $\alpha = 0,05$.

`\begin{table}`

`\caption{Resumo de todos os resultados favoráveis dos testes para 5% de significância}`

Nome	Método	Pvalor	Conclusão
pressao_sis	Teste permutação	0.0000	Rejeita-se H0
pressao_dias	Teste permutação	0.0000	Rejeita-se H0
pressao_media	Teste permutação	0.0000	Rejeita-se H0
ldl_dosado	Teste permutação	0.0000	Rejeita-se H0
peso_bebe	Welch Two Sample t-test	0.0000	Rejeita-se H0
prematureo	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
hipertensao	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
anti_hiper	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
filtr_glomerular	Teste permutação	0.0014	Rejeita-se H0
trombomodulina	Welch Two Sample t-test	0.0021	Rejeita-se H0
imc	Teste permutação	0.0082	Rejeita-se H0
abdomen	Welch Two Sample t-test	0.0095	Rejeita-se H0
gordura	Teste permutação	0.0108	Rejeita-se H0
tipo_parto	Fisher's Exact Test for Count Data	0.0108	Rejeita-se H0
escolaridade	Fisher's Exact Test for Count Data	0.0239	Rejeita-se H0
sft1	Welch Two Sample t-test	0.0397	Rejeita-se H0

`\end{table}`

Caso fosse assumido um nível de significância de $\alpha = 0,10$, as seguintes variáveis seriam candidatas a impactar no grupo.

`\begin{table}`

`\caption{Resumo de todos os resultados favoráveis dos testes para 10% de significância}`

Tabela 31: Resumo de todos os resultados dos testes

Nome	Método	Pvalor	Conclusão
pressao_sis	Teste permutação	0.0000	Rejeita-se H0
pressao_dias	Teste permutação	0.0000	Rejeita-se H0
pressao_media	Teste permutação	0.0000	Rejeita-se H0
ldl_dosado	Teste permutação	0.0000	Rejeita-se H0
peso_bebe	Welch Two Sample t-test	0.0000	Rejeita-se H0
prematureo	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
hipertensao	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
anti_hiper	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
filtr_glomerular	Teste permutação	0.0014	Rejeita-se H0
trombomodulina	Welch Two Sample t-test	0.0021	Rejeita-se H0
imc	Teste permutação	0.0082	Rejeita-se H0
abdomen	Welch Two Sample t-test	0.0095	Rejeita-se H0
gordura	Teste permutação	0.0108	Rejeita-se H0
tipo_parto	Fisher's Exact Test for Count Data	0.0108	Rejeita-se H0
escolaridade	Fisher's Exact Test for Count Data	0.0239	Rejeita-se H0
sft1	Welch Two Sample t-test	0.0397	Rejeita-se H0
quadril	Teste permutação	0.0580	Não rejeita-se H0
hdl	Teste permutação	0.0598	Não rejeita-se H0
aat1	Welch Two Sample t-test	0.0602	Não rejeita-se H0
cintura	Teste permutação	0.0616	Não rejeita-se H0
pail	Teste permutação	0.0640	Não rejeita-se H0
lym	Teste permutação	0.0718	Não rejeita-se H0
adma	Teste permutação	0.0900	Não rejeita-se H0
peso	Teste permutação	0.0918	Não rejeita-se H0
tgo_ast	Teste permutação	0.1062	Não rejeita-se H0
nao_hdl	Teste permutação	0.1104	Não rejeita-se H0
colesterol_total	Teste permutação	0.1604	Não rejeita-se H0
sft1_geral	Teste permutação	0.2078	Não rejeita-se H0
nlr	Teste permutação	0.2346	Não rejeita-se H0
sft1_luiza	Teste permutação	0.2780	Não rejeita-se H0
etnia	Fisher's Exact Test for Count Data	0.2943	Não rejeita-se H0
acido_urico	Welch Two Sample t-test	0.3302	Não rejeita-se H0
diabetes	Fisher's Exact Test for Count Data	0.3353	Não rejeita-se H0
vldl	Teste permutação	0.3424	Não rejeita-se H0
neu	Teste permutação	0.3572	Não rejeita-se H0
quantos_anos_evento	Fisher's Exact Test for Count Data simulated	0.3643	Não rejeita-se H0
medic_hipercol	Fisher's Exact Test for Count Data	0.4735	Não rejeita-se H0
cintura_quadril	Teste permutação	0.4746	Não rejeita-se H0
gama_gt	Teste permutação	0.4840	Não rejeita-se H0
tgp_alt	Teste permutação	0.5284	Não rejeita-se H0
num_gestacoes	Fisher's Exact Test for Count Data	0.5611	Não rejeita-se H0
dimero	Teste permutação	0.6510	Não rejeita-se H0
hipercolesterolemia	Fisher's Exact Test for Count Data	0.6627	Não rejeita-se H0
triglicerides	Teste permutação	0.8522	Não rejeita-se H0
hba1c	Teste permutação	0.8942	Não rejeita-se H0
hba1c	Teste permutação	0.9864	Não rejeita-se H0
medic_diabetes	Fisher's Exact Test for Count Data	1.0000	Não rejeita-se H0
tabagista	Fisher's Exact Test for Count Data	1.0000	Não rejeita-se H0

Nome	Método	Pvalor	Conclusão
pressao_sis	Teste permutação	0.0000	Rejeita-se H0
pressao_dias	Teste permutação	0.0000	Rejeita-se H0
pressao_media	Teste permutação	0.0000	Rejeita-se H0
ldl_dosado	Teste permutação	0.0000	Rejeita-se H0
peso_bebe	Welch Two Sample t-test	0.0000	Rejeita-se H0
prematureo	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
hipertensao	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
anti_hiper	Fisher's Exact Test for Count Data	0.0000	Rejeita-se H0
filtr_glomerular	Teste permutação	0.0014	Rejeita-se H0
trombomodulina	Welch Two Sample t-test	0.0021	Rejeita-se H0
imc	Teste permutação	0.0082	Rejeita-se H0
abdomen	Welch Two Sample t-test	0.0095	Rejeita-se H0
gordura	Teste permutação	0.0108	Rejeita-se H0
tipo_parto	Fisher's Exact Test for Count Data	0.0108	Rejeita-se H0
escolaridade	Fisher's Exact Test for Count Data	0.0239	Rejeita-se H0
sft1	Welch Two Sample t-test	0.0397	Rejeita-se H0
quadril	Teste permutação	0.0580	Rejeita-se H0
hdl	Teste permutação	0.0598	Rejeita-se H0
aat1	Welch Two Sample t-test	0.0602	Rejeita-se H0
cintura	Teste permutação	0.0616	Rejeita-se H0
pail	Teste permutação	0.0640	Rejeita-se H0
lym	Teste permutação	0.0718	Rejeita-se H0
adma	Teste permutação	0.0900	Rejeita-se H0
peso	Teste permutação	0.0918	Rejeita-se H0

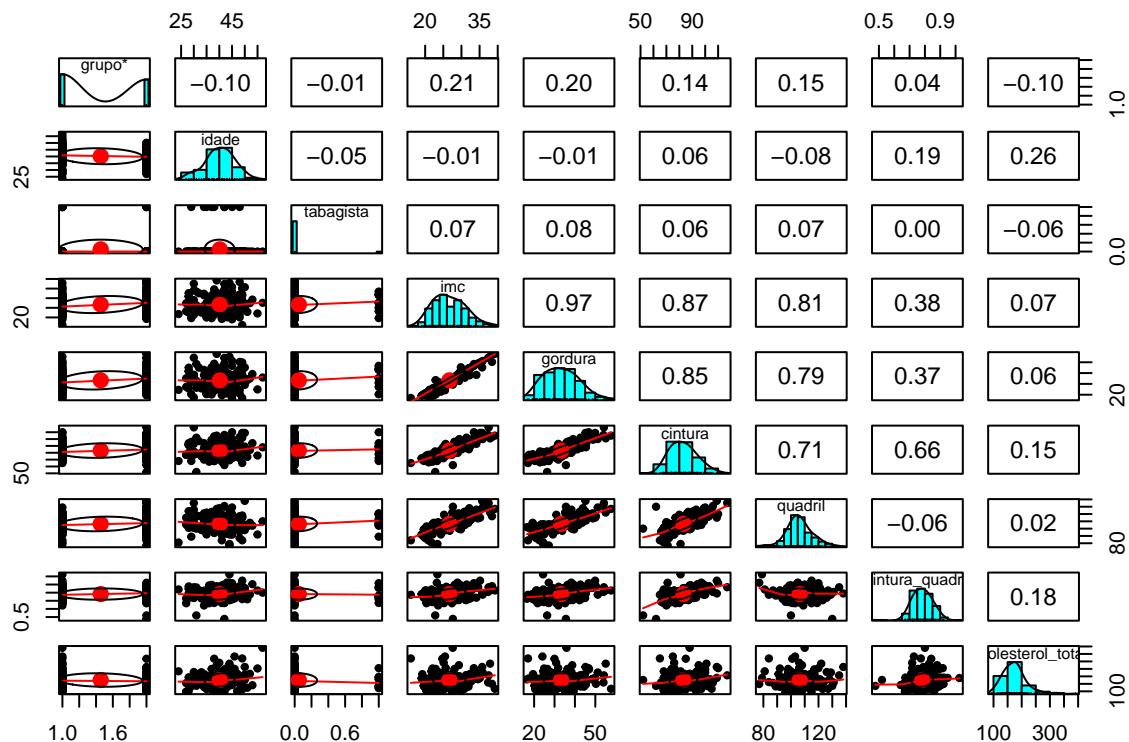
\end{table}

4.5 Modelo Linear Generalizado para hipertensão

```
## # A tibble: 6 x 53
##   iniciais grupo.x etnia idade escolaridade num_gestacoes peso imc gordura
##   <chr>      <fct>  <dbl> <dbl> <fct>                <dbl> <dbl> <dbl>  <dbl>
## 1 LDS        1        2   39 1                    4  68.4  29.6  39.1
## 2 MRS        0        2   45 0                    3 102.  35.9  51.2
## 3 ZCAS        1        1   31 2                    2  67.3  26.3  32.9
## 4 AAFR        1        2   45 2                    3  69.8  25.6  31.9
## 5 AMAA        1        1   30 2                    2  71.5  29.0  38.3
## 6 DBGA        1        2   36 2                    3  79.9  29.0   38
## # i 44 more variables: pressao_sis <dbl>, pressao_dias <dbl>,
## #   pressao_media <dbl>, cintura <dbl>, abdomen <dbl>, quadril <dbl>,
## #   cintura_quadril <dbl>, neu <dbl>, lym <dbl>, nlr <dbl>,
## #   colesterol_total <dbl>, hdl <dbl>, ldl_dosado <dbl>, vldl <dbl>,
## #   nao_hdl <dbl>, triglicerides <dbl>, hba1c <dbl>, creatinina <dbl>,
## #   filtr_glomerular <dbl>, tgp_alt <dbl>, tgo_ast <dbl>, acido_urico <dbl>,
## #   ck <dbl>, gama_gt <dbl>, anti_hiper <dbl>, medic_diabetes <dbl>, ...
##
##           iniciais      grupo colesterol_total      hdl
##           0.000000000      0.000000000      0.000000000      0.000000000
##           ldl_dosado  triglicerides      imc      cintura
##           0.000000000      0.000000000      0.005319149      0.005319149
##           quadril  cintura_quadril  pressao_sis  pressao_dias
##           0.005319149      0.005319149      0.010638298      0.010638298
##           idade      tabagista      gordura      hba1c
```

```
##      0.010638298      0.010638298      0.021276596      0.234042553
##      aaat1          pai1          dimero
##      0.569148936      0.574468085      0.585106383
```

```
## # A tibble: 10 x 4
##   Nome          Método          Pvalor Conclusão
##   <chr>         <chr>         <dbl> <chr>
## 1 imc           Teste permutação      0.0014 Rejeita-se H0
## 2 ldl_dosado    Teste permutação      0.0018 Rejeita-se H0
## 3 cintura      Welch Two Sample t-test 0.0025 Rejeita-se H0
## 4 gordura      Teste permutação      0.0028 Rejeita-se H0
## 5 cintura_quadril Teste permutação      0.0124 Rejeita-se H0
## 6 quadril       Teste permutação      0.0606 Não rejeita-se H0
## 7 idade        Welch Two Sample t-test 0.230  Não rejeita-se H0
## 8 hdl          Teste permutação      0.623  Não rejeita-se H0
## 9 colesterol_total Teste permutação      0.634  Não rejeita-se H0
## 10 triglicerides Teste permutação      0.716  Não rejeita-se H0
```



```
##
## Call:
## glm(formula = alta_pressao ~ grupo + imc + cintura + cintura_quadril +
##      idade + quadril + gordura + ldl_dosado, family = binomial(link = "logit"),
##      data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8947  -0.8883  -0.6075   1.0715   1.9733
```

```
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)  31.105454  23.054389   1.349   0.1773
## grupo2       0.875230   0.364709   2.400   0.0164 *
## imc          0.317023   0.184817   1.715   0.0863 .
## cintura     0.424896   0.267903   1.586   0.1127
## cintura_quadril -44.711264  28.624103  -1.562   0.1183
## idade       0.035508   0.030367   1.169   0.2423
## quadril     -0.380309   0.221018  -1.721   0.0853 .
## gordura     -0.080569   0.083715  -0.962   0.3358
## ldl_dosado   0.006558   0.005151   1.273   0.2030
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 207.41  on 173  degrees of freedom
## AIC: 225.41
##
## Number of Fisher Scoring iterations: 4
##
##      grupo      imc      cintura cintura_quadril      idade
##      1.168764    23.271350    285.881923    118.081902    1.127745
##      quadril      gordura      ldl_dosado
##      162.878990    17.676427    1.179525
##
## Call:
## glm(formula = alta_pressao ~ grupo + idade + ldl_dosado, family = binomial(link = "logit"),
##      data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5392  -0.9073  -0.6792   1.1403   1.9245
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.383135   1.216383  -2.781   0.00541 **
## grupo2       0.959907   0.345115   2.781   0.00541 **
## idade       0.036610   0.028486   1.285   0.19872
## ldl_dosado  0.007961   0.004825   1.650   0.09894 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 219.18  on 178  degrees of freedom
## AIC: 227.18
##
## Number of Fisher Scoring iterations: 4
##
##      grupo      idade ldl_dosado
```

```

##    1.118696    1.059741    1.103919
##
## Call:
## glm(formula = alta_pressao ~ grupo + idade + imc + ldl_dosado,
##      family = binomial(link = "logit"), data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5705  -0.8860  -0.6386   1.1197   2.0676
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.878246   1.636090  -3.593 0.000327 ***
## grupo2       0.848765   0.353921   2.398 0.016477 *
## idade       0.039990   0.029137   1.373 0.169903
## imc         0.093538   0.038252   2.445 0.014474 *
## ldl_dosado  0.006895   0.004935   1.397 0.162369
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 212.96  on 177  degrees of freedom
## AIC: 222.96
##
## Number of Fisher Scoring iterations: 4
##
## Call:
## glm(formula = alta_pressao ~ grupo + idade + gordura + ldl_dosado,
##      family = binomial(link = "logit"), data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5578  -0.9016  -0.6500   1.1486   2.0349
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.876953   1.430422  -3.409 0.000651 ***
## grupo2       0.858950   0.352453   2.437 0.014807 *
## idade       0.039476   0.028959   1.363 0.172830
## gordura     0.043939   0.019834   2.215 0.026732 *
## ldl_dosado  0.007390   0.004897   1.509 0.131282
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 214.13  on 177  degrees of freedom
## AIC: 224.13
##
## Number of Fisher Scoring iterations: 4

```



```
##
## Call:
## glm(formula = alta_pressao ~ grupo + idade + cintura + ldl_dosado,
##      family = binomial(link = "logit"), data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5824  -0.8695  -0.6532   1.0740   2.0273
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -6.247899   1.775274  -3.519 0.000433 ***
## grupo2       0.913133   0.351632   2.597 0.009409 **
## idade        0.035300   0.028911   1.221 0.222097
## cintura      0.037187   0.016006   2.323 0.020160 *
## ldl_dosado   0.006235   0.004950   1.260 0.207845
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 213.56  on 177  degrees of freedom
## AIC: 223.56
##
## Number of Fisher Scoring iterations: 4
##
## Call:
## glm(formula = alta_pressao ~ grupo + idade + quadril + ldl_dosado,
##      family = binomial(link = "logit"), data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5761  -0.9061  -0.6824   1.1419   1.9605
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.798008   2.263386  -2.562 0.01042 *
## grupo2       0.925435   0.348328   2.657 0.00789 **
## idade        0.041632   0.029083   1.431 0.15229
## quadril      0.021549   0.016723   1.289 0.19754
## ldl_dosado   0.007279   0.004888   1.489 0.13647
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 217.48  on 177  degrees of freedom
## AIC: 227.48
##
## Number of Fisher Scoring iterations: 4
##
```

```
## Call:
## glm(formula = alta_pressao ~ grupo + idade + cintura_quadril +
##      ldl_dosado, family = binomial(link = "logit"), data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5466  -0.9104  -0.6354   1.1421   1.9950
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -6.420382   2.123925  -3.023  0.00250 **
## grupo2         0.949895   0.348100   2.729  0.00636 **
## idade         0.027041   0.028880   0.936  0.34912
## cintura_quadril 4.449829   2.502679   1.778  0.07540 .
## ldl_dosado     0.007256   0.004875   1.489  0.13659
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 215.87  on 177  degrees of freedom
## AIC: 225.87
##
## Number of Fisher Scoring iterations: 4
##
## Start:  AIC=254.25
## alta_pressao ~ grupo + imc + cintura + cintura_quadril + idade +
##      quadril + gordura + ldl_dosado
##
##              Df Deviance    AIC
## - gordura      1   208.34 249.97
## - idade         1   208.80 250.43
## - ldl_dosado    1   209.04 250.67
## - imc           1   210.41 252.04
## - cintura_quadril 1   210.59 252.22
## - cintura       1   210.76 252.39
## - quadril       1   211.32 252.96
## <none>          1   207.41 254.25
## - grupo         1   213.26 254.89
##
## Step:  AIC=249.97
## alta_pressao ~ grupo + imc + cintura + cintura_quadril + idade +
##      quadril + ldl_dosado
##
##              Df Deviance    AIC
## - idade         1   209.74 246.17
## - ldl_dosado    1   210.31 246.74
## - cintura_quadril 1   211.16 247.59
## - imc           1   211.29 247.72
## - cintura       1   211.30 247.73
## - quadril       1   211.87 248.30
## <none>          1   208.34 249.97
## - grupo         1   214.04 250.46
```

```

## + gordura          1    207.41 254.25
##
## Step: AIC=246.17
## alta_pressao ~ grupo + imc + cintura + cintura_quadril + quadril +
##      ldl_dosado
##
##              Df Deviance    AIC
## - cintura_quadril  1    212.40 243.62
## - ldl_dosado       1    212.40 243.62
## - imc              1    212.54 243.76
## - cintura          1    212.62 243.84
## - quadril          1    213.18 244.41
## - grupo            1    214.63 245.86
## <none>              209.74 246.17
## + idade            1    208.34 249.97
## + gordura          1    208.80 250.43
##
## Step: AIC=243.62
## alta_pressao ~ grupo + imc + cintura + quadril + ldl_dosado
##
##              Df Deviance    AIC
## - cintura          1    212.69 238.71
## - imc              1    214.39 240.41
## - quadril          1    214.57 240.59
## - ldl_dosado       1    214.98 241.00
## - grupo            1    217.36 243.38
## <none>              212.40 243.62
## + cintura_quadril  1    209.74 246.17
## + idade            1    211.16 247.59
## + gordura          1    211.80 248.22
##
## Step: AIC=238.71
## alta_pressao ~ grupo + imc + quadril + ldl_dosado
##
##              Df Deviance    AIC
## - quadril          1    214.88 235.69
## - ldl_dosado       1    215.68 236.49
## - grupo            1    217.42 238.24
## <none>              212.69 238.71
## - imc              1    219.57 240.39
## + idade            1    211.31 242.53
## + gordura          1    212.15 243.37
## + cintura          1    212.40 243.62
## + cintura_quadril  1    212.62 243.84
##
## Step: AIC=235.69
## alta_pressao ~ grupo + imc + ldl_dosado
##
##              Df Deviance    AIC
## - ldl_dosado       1    217.64 233.26
## - grupo            1    219.76 235.37
## <none>              214.88 235.69
## - imc              1    220.85 236.47
## + quadril          1    212.69 238.71

```

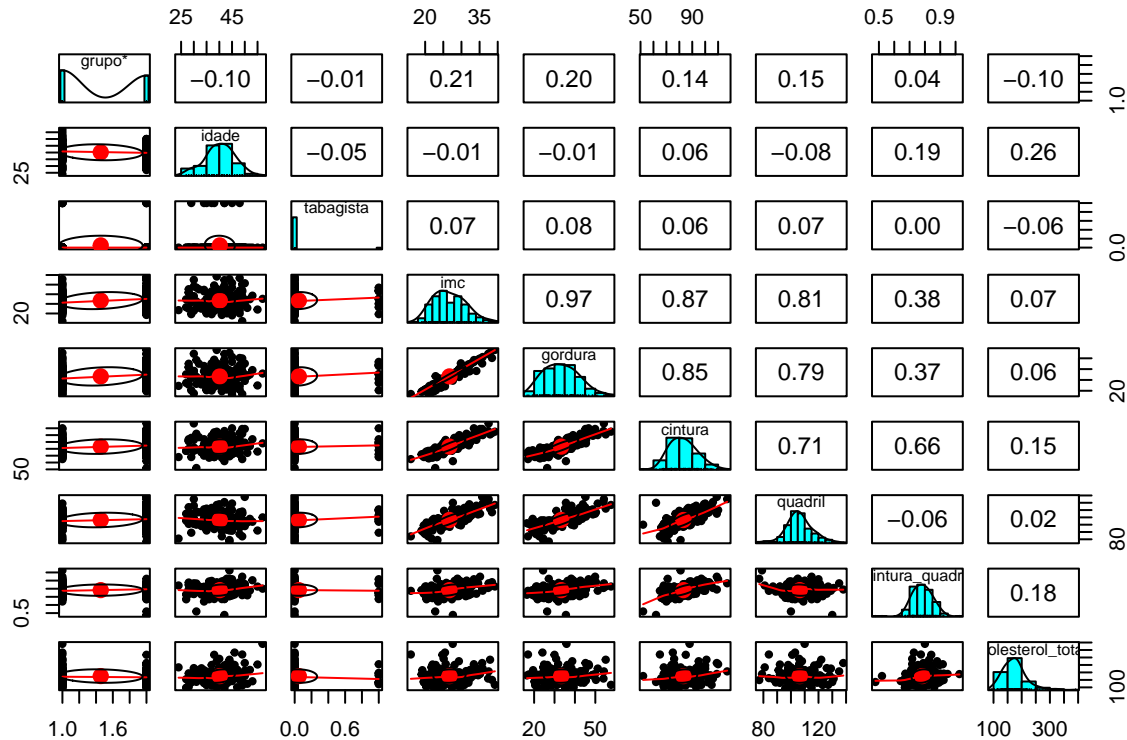
```

## + idade          1    212.96 238.98
## + cintura_quadril 1    213.40 239.42
## + gordura         1    214.32 240.34
## + cintura         1    214.57 240.59
##
## Step:  AIC=233.26
## alta_pressao ~ grupo + imc
##
##              Df Deviance   AIC
## <none>              217.64 233.26
## - imc                1    224.62 235.03
## + ldl_dosado         1    214.88 235.69
## + idade              1    214.94 235.75
## - grupo              1    225.54 235.95
## + quadril            1    215.68 236.49
## + cintura_quadril    1    215.76 236.58
## + gordura            1    216.73 237.55
## + cintura            1    216.94 237.75
##
## Call:
## glm(formula = alta_pressao ~ grupo + imc, family = binomial(link = "logit"),
##      data = dfmodelo)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6454  -0.8795  -0.6531   1.1501   1.9564
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.71339    1.03720  -3.580 0.000343 ***
## grupo2       0.92084    0.33053   2.786 0.005336 **
## imc          0.09810    0.03794   2.586 0.009717 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 236.04  on 181  degrees of freedom
## Residual deviance: 217.64  on 179  degrees of freedom
## AIC: 223.64
##
## Number of Fisher Scoring iterations: 4
##
## z test of coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  31.1054539  23.0543890   1.3492 0.17727
## grupo2       0.8752297   0.3647095   2.3998 0.01640 *
## imc          0.3170233   0.1848167   1.7153 0.08628 .
## cintura     0.4248956   0.2679030   1.5860 0.11274
## cintura_quadril -44.7112642  28.6241029  -1.5620 0.11828
## idade        0.0355083   0.0303671   1.1693 0.24228
## quadril      -0.3803090   0.2210181  -1.7207 0.08530 .

```

```
## gordura          -0.0805694    0.0837154 -0.9624    0.33584
## ldl_dosado       0.0065575     0.0051511  1.2730    0.20300
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## (Intercept)      grupo2          imc          cintura cintura_quadril
## 3.227952e+13    2.399426e+00    1.373035e+00    1.529431e+00    3.820715e-20
## idade          quadril          gordura          ldl_dosado
## 1.036146e+00    6.836501e-01    9.225908e-01    1.006579e+00
```



A análise da matriz de dispersão e da correlação entre as variáveis pode ajudar a identificar a multicolinearidade, que ocorre quando duas ou mais variáveis independentes estão altamente correlacionadas entre si. Isso pode ser observado em gráficos de dispersão onde as variáveis estão fortemente alinhadas. É possível identificar uma relação linear positiva entre as seguintes variáveis: imc, gordura, cintura e quadril.

Characteristic	**OR**	**95% CI**	**p-value**
grupo			
1	—	—	
2	2.14	1.09, 4.23	0.028
imc	1.10	1.02, 1.18	0.016
ldl_dosado	1.01	1.00, 1.02	0.10

5 Referências

<https://dl.acm.org/doi/pdf/10.1145/503506.503509>

<https://est.ufba.br/sites/est.ufba.br/files/kim/matd49-aula04-fisher.pdf>

<https://www.jstor.org/stable/2527460?seq=1>
<http://ndl.ethernet.edu.et/bitstream/123456789/15304/1/9103.pdf>
<https://janalin.github.io/analyse-ELISA/results.html>
<https://cran.r-project.org/web/packages/drc/drc.pdf>
<https://nomato.files.wordpress.com/2015/03/curvas-de-dose-resposta-no-software-r.pdf>