

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas - ICEx
LABORATÓRIO DE ESTATÍSTICA I - TTC

Resposta terapêutica na malária por Plasmodium vivax
Variabilidade genética de enzimas metabolizadoras da primaquina e
o clearance de gametócitos

Consultores: Maria Eduarda da Silva Vilas Bôas e Sophia Araújo de Moraes
Orientador: Profa. Lourdes Coral Contreras Montenegro
Clientes: Yanka Evellyn Alves Rodrigues Salazar
Profa. Taís Nóbrega de Sousa

Belo Horizonte
2022

Sumário

| | | |
|----------|--------------------------------------|-----------|
| 1 | Introdução | 3 |
| 1.1 | Objetivos Gerais do estudo | 3 |
| 1.2 | Objetivos Específicos | 3 |
| 2 | Metodologia | 4 |
| 2.1 | Dados | 4 |
| 2.2 | Análise Exploratória | 5 |
| 2.3 | Procedimentos estatísticos | 5 |
| 3 | Resultados | 7 |
| 3.1 | Análise descritiva | 7 |
| 3.2 | Pergunta 1 | 11 |
| 3.3 | Pergunta 2 | 14 |
| 3.4 | Pergunta 3 | 21 |
| 3.5 | Pergunta 4 | 25 |
| 3.6 | Pergunta 5 | 29 |
| 3.7 | Pergunta 6 | 32 |
| 3.8 | Pergunta 7 | 35 |
| 4 | Conclusão | 36 |
| 5 | Referências | 37 |

1 Introdução

A cliente Yanka Evellyn Alves Rodrigues Salazar orientada pela Dra. Taís Nóbrega de Sousa são responsáveis pela pesquisa “RESPOSTA TERAPÊUTICA NA MALÁRIA POR *Plasmodium vivax*: VARIABILIDADE GENÉTICA DE ENZIMAS METABOLIZADORAS DA PRIMAQUINA E O CLEARANCE DE GAMETÓCITOS”, um estudo transversal elaborado juntamente com Programa de Pós-Graduação em Ciências da Saúde. Sendo a fundação responsável o Instituto René Rachou (Fundação Oswaldo Cruz).

A literatura nos traz diversos fatores que podem influenciar a falha terapêutica, em muitos estudos de tratamento da malária (doença infecciosa causada pelo parasito do gênero *Plasmodium*). As mutações no gene POR tem potencial em alterar a atividade metabólica das CYPs, sabe-se também que a variabilidade genética dessas enzimas possui certa influência nos estágios assexuados de *Plasmodium vivax* (*P. vivax*), protozoário mais comum para disseminação de malária. Os gametócitos são as células germinativas eucarióticas, sendo os masculinos chamados de espermátócitos e os femininos de ovócito, na literatura não há muitos estudos sobre o efeito nos gametócitos por isto, nossos clientes sintetizaram um estudo que será pioneiro na compreensão de como polimorfismos em enzimas envolvidas na metabolização de um medicamento para a prevenção de malária conhecido como Primaquina, para verificar se interferem na liberação de gametócitos em infecções causadas por *P. vivax*.

Os polimorfismos têm origem a partir de consequências das mutações, sendo elas variações genéticas. Um ponto que pode influenciar os polimorfismos em CYP2D6 é a enzima que está responsável pela metabolização de Primaquina, e em POR na atividade da Primaquina sobre os gametócitos de *P. vivax*. Neste trabalho, pretende-se analisar a frequência de polimorfismos nos genes CYP2D6 e POR a partir da amostra colhida pelas clientes. Outro ponto é verificar a influência dos polimorfismos em CYP2D6 na atividade de gametócitos da Primaquina em indivíduos infectados por *P. vivax*. Por fim, pretende-se elucidar se há evidências estatísticas que o polimorfismo no gene POR interfere na atividade de gametócitos da Primaquina.

1.1 Objetivos Gerais do estudo

Avaliar a influência de polimorfismos em CYP2D6, principal enzima responsável pela metabolização de PQ, e em POR na atividade da PQ sobre os gametócitos de *Plasmodium vivax*.

1.2 Objetivos Específicos

1. Analisar a frequência de polimorfismos nos genes CYP2D6 e POR na população de estudo;
2. Investigar a influência dos polimorfismos em CYP2D6 na atividade gametocitocida da PQ em indivíduos infectados por *P. vivax*;
3. Determinar se polimorfismos no gene POR interferem na atividade gametocitocida da PQ.

2 Metodologia

2.1 Dados

O banco de dados foi composto por 100 indivíduos e 35 variáveis, totalizando um total de 3.500 observações. Ao total, foram localizados 162 valores ausentes ou não determinados. Esses valores faltantes causaram influências significativas nas análises que serão desenvolvidas posteriormente, tendo em vista que reduziram ainda mais o tamanho amostral, gerando impactos nas análises multivariadas.

Antes do início das análises, vale salientar que conforme orientação das nossas clientes, as variáveis numéricas (Gametócitos femininos no momento inicial (QPVS25D0), Gametócitos femininos no momento 72 horas (QPVS25D72), Razão de Pvs 25 (QPVS25D72/D0) e Concentração de primaquina (ng/mL) em 72 horas (PQCONC)) foram transformadas na escala logarítmo na base 10. Essas transformações foram realizadas com base na literatura trazida pelas pesquisadoras.

Uma breve descrição das variáveis utilizadas e sua faixa de valores são apresentadas abaixo para melhor compreensão dos resultados posteriormente:

A Razão de Pvs 25, foi obtida por meio da divisão da Quantidade de gametócitos femininos maduros no momento de 72 horas, pela Quantidade de gametócitos femininos maduros no momento inicial. O termo Eliminação de gametócitos será usado como sinônimo da variável Razão de Pvs 25. Nos resultados realizou-se teste t pareado de comparação de médias visando averiguar se realmente houve diminuição na quantidade de gametócitos após 72 horas. Os valores da variável Razão de Pvs 25, em escala logarítmo na base 10, vão de -5,000 à 3,848, e quanto mais alto o valor obtido maior a quantidade de gametócitos no momento de 72 horas.

A variável Concentração de primaquina (PQCONC), quantifica a concentração de medicação no momento final da coleta dos dados. Está variável de suma importância pois primaquina foi o fármaco utilizado para o tratamento de malária. A escala de valores para Concentração de primaquina varia de 1,431 à 2,565, sendo que valores mais altos indicam maior concentração deste fármaco no metabolismo do indivíduo no momento 72 horas.

A variável Quantificação de parasitos em 0 horas (PV18SD0NORM), mensura a Quantidade de parasitos de malária no momento inicial. A escala de valores para PV18SD0NORM varia de -5,000 à 7,327, sendo que valores mais altos indicam uma maior concentração de parasitos no organismo do indivíduo. A variável Quantificação de parasitos em 72 horas (PV18SD72NORM), avalia a quantidade de parasitos de malária após 72 horas, e seus valores vão de -5,000 à 5,292.

O Status de metabolização CYP2D6 e POR, foi analisado de forma conjunta e individual. A variável Status de metabolização CYP2D6 possui duas categorias sendo elas: "Normal" e "Alterada". A variável Status de mutação do gene POR possui duas categorias: "Mutada" e "Não-mutada". Já a variável Status de metabolização conjunta (CYP2D6/POR) foi definida com quatro categorias: "Normal/Não-mutada", "Normal/Mutada", "Alterada/Não-mutada" e "Alterada/Mutada". Além disso, a variável CYP2D6 também foi definida quanto ao fenótipo, sendo a classificação feita em termos das categorias: "gPM" - metabolizador nulo, não há metabolização; "gIM" - metabolizador intermediário; "gNM-S" - metabolização reduzida ("gIM" e "gNM-S" apresentam metabolização reduzida), e "gNM-F" - metabolização normal; "gUM" - metabolização ultrarrápida.

Também era de interesse da pesquisa analisar o status dos gametócitos masculinos, por isso utilizou-se as variáveis qualitativas: Indicativo de presença ou ausência de gametócitos masculinos em 0 horas (PVS48D0), e Indicativo de presença ou ausência de gametócitos masculinos em Pvs48/45 em 72 horas (PVS48D72). PVS48D0 e PVS48D72 foram classificadas como: "Positivo" e "Negativo".

2.2 Análise Exploratória

Inicialmente foi sintetizada uma análise exploratória, com objetivo de conhecer as características e identificar padrões nos dados. Parte desse processo foi desenvolvido por meio da obtenção de medidas resumo (média, mediana, desvio padrão, máximo e mínimo), para as variáveis numéricas, e frequência relativa de cada um dos níveis das variáveis categóricas. Além disso, foram elaborados gráficos de dispersão, boxplots e histogramas visando um melhor entendimento das medidas de posição, tendência central e da dispersão dos dados.

Posteriormente, foi realizado o Teste de Shapiro-Wilk, para as variáveis numéricas (Idade, PQCONC, QPVS25D0, QPVS25D72, QPVS25D72/D0, PV18SD0NORM e PV18SD72NORM), com o propósito de verificar quais seguiam a distribuição normal. Por meio desse teste foi possível nortear os testes estatísticos a serem desenvolvidos, visando responder as perguntas formuladas pelas pesquisadoras.

2.3 Procedimentos estatísticos

A estruturação dos resultados foi seguida com base nos seguintes questionamentos:

- Indivíduos com status CYP2D6 alterada/POR mutada possuem menor redução de Pvs25 (72 horas para 0 horas)? (Item 1)
- A Parasitemia inicial (PV18SD0NORM) está associada com a quantidade de gametócitos?
- A Parasitemia inicial (PV18SD0NORM) é um fator de confusão nas análises de associação entre status CYP/POR e Eliminação de gametócitos realizadas no Item 1?
- O Gênero e Idade dos pacientes são fatores de confusão nas análises de associação entre status CYP/POR e Eliminação de gametócitos realizadas no Item 1?
- A Concentração de primaquina em 72 horas (PQCONC) é influenciada pelo fenótipo de CYP2D6 (CYP2D6FENOTIPO), POR (PORMUT) e o status conjunto CYP2D6/POR (CYP2D6POR)?
- Há correlação entre as variáveis: Concentração de primaquina (PQCONC) com a redução de Pvs25 (QPVS25D72/D0)?
- Os Status de metabolização CYP2D6 (CYP2D6CAT) e os Status de mutação do gene POR (PORMUT) influenciam no número de recorrência clínica (NRECORRENCIA)?
- Há associação entre as variáveis: Indicativo de presença ou ausência de gametócitos masculinos em 0 horas (PVS48D0) e Indicativo de presença ou ausência de gametócitos masculinos em 72 horas (PVS48D72) estão associadas com os status de metabolização conjunto e individual CYP2D6/POR?

Os testes de hipóteses e modelos sintetizados para responder as perguntas elaboradas pelas pesquisadoras foram:

- Teste Qui-quadrado;
- Teste t-pareado;
- Teste de Kruskal Wallis;
- Teste de Mann-Whitney;
- Teste de Correlação de Spearman;
- Modelo de regressão linear múltipla.

Para avaliar se há influência das variáveis levantadas pelas pesquisadoras foram ajustados modelos de regressões múltiplas. Os modelos ajustados foram sintetizados para explicar: Razão de Pvs 25 e Concentração de Primaquina em 72 horas.

Vale salientar que em todas as análises foram consideradas um nível de significância de 5% e utilizou-se o software R em sua extensão RStudio para obtenção dos resultados.

3 Resultados

Inicialmente aplicou-se teste de Shapiro Wilk para verificação da normalidade nas seguintes variáveis quantitativas numéricas : “Idade”, “Recorrência de malária”, “Concentração de primaquina (72 horas)”, “Quantificação absoluta de gametócitos femininos (momento inicial)”, “Quantificação absoluta de gametócitos femininos (72 horas)”, “Razão de Pvs25 em 72 horas por 0 horas”, “Quantificação absoluta de parasitos (72 horas)”. Para todas as variáveis não rejeita-se a hipótese de normalidade. (Valor- $p < 0,05$).

Para a verificação da utilização do termo “Eliminação de gametócitos” realizou-se o teste t-pareado unilateral para comparação das médias das variáveis “Quantificação absoluta de gametócitos femininos maduros em 0 horas” e “Quantificação absoluta de gametócitos femininos maduros em 72 horas”.

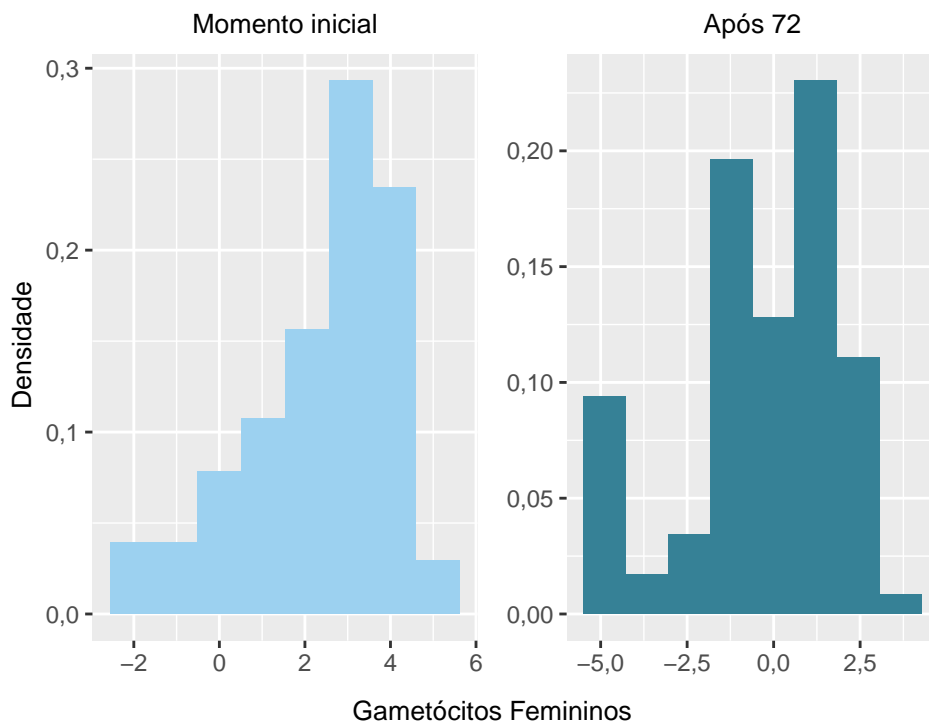
Hipóteses:

H_0 : Não há diferença entre as médias “Quantificação absoluta de gametócitos femininos maduros em 0 horas” e “Quantificação absoluta de gametócitos femininos maduros em 72 horas”;

H_1 : Média inicial de “Quantificação absoluta de gametócitos femininos maduros em 0 horas” > média final e “Quantificação absoluta de gametócitos femininos maduros em 72 horas” (teste unilateral).

P valor < 0,001, logo há evidências de que, após 72 horas, houve uma diminuição estatisticamente significativa da quantidade de gametócitos femininos maduros. Assim, testificando que o termo “eliminação de gametócitos” pode ser usado como sinônimo da variável “Razão de Pvs25”. Na figura 1 observa-se que a média de gametócitos femininos no momento de 72 horas está centrada em valores menor do que a média de gametócitos femininos no momento inicial.

Figura 1 – Análise gametócitos femininos



3.1 Análise descritiva

Foi sintetizada uma análise descritiva para compreender melhor o comportamento das variáveis do banco de dados. Para isto foram realizadas análises gráficas, tabelas com medidas resumo ou teste de hipóteses

cruzando variáveis de interesse, como pode-se ver a seguir.

A Figura 2 e a Tabela 1 apresentam a análise descritiva das variáveis sexo e idade. Desta forma observa-se que a idade média e mediana do grupo feminino é maior que a idade do grupo masculino. A idade média é de 41,90 anos e o desvio padrão 9,49 anos do grupo feminino, quanto ao grupo masculino a idade média é de 31,58 anos e o desvio padrão 10,29 anos.



Table 1: Análise descritiva das variáveis Sexo e Idade

| Sexo | Idade | | | | |
|-----------|--------|--------|---------|---------|---------------|
| | Mínimo | Máximo | Média | Mediana | Desvio Padrão |
| Masculino | 17 | 53 | 31,5814 | 30 | 10,2869 |
| Feminino | 29 | 57 | 40,9091 | 40 | 9,4916 |

Dessa forma, visando responder o seguinte questionamento: “A média das idades, para os níveis da variável sexo são iguais?” foi elaborado um teste de hipótese do teste Mann-Whitney para comparação de medianas de dois grupos. Desse modo assumimos como hipóteses de interesse:

H_0 : A mediana da Idade masculina é igual à mediana da Idade feminina;

H_1 : A mediana da Idade masculina é diferente à mediana da Idade feminina.

O p-valor obtido foi de 0,071, logo há evidências de que a mediana das idades, considerando os níveis do sexo (masculino e feminino), não são estatisticamente diferentes.

De maneira análoga, analisou-se a Recorrência clínica de malária com relação ao Sexo, Local de infecção, Ocupação e Fenótipo CYP. Em todas as análises, realizou-se testes de independência qui-quadrado, assumindo as seguintes hipóteses:

H_0 : A variável analisada não está associada com a Recorrência clínica de malária (são independentes);

H_1 : A variável analisada está associada com a Recorrência clínica de malária (são dependentes).

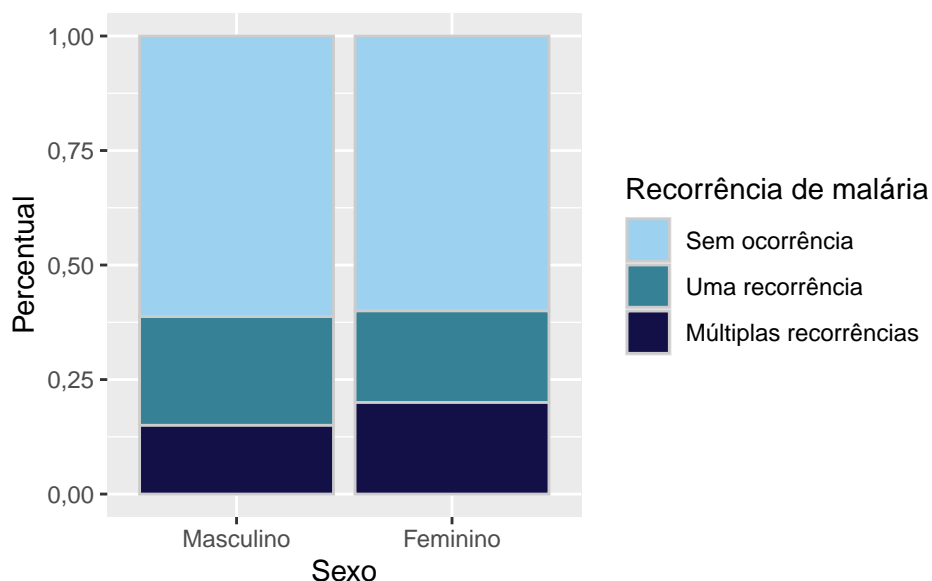
Table 2: Teste de independência qui-quadrado com a variável Recorrência de Malária

| Variáveis | P-Valores |
|----------------------------|-----------|
| Sexo | 0,839 |
| Provável local de infecção | 0,893 |
| Ocupação | 0,557 |
| Fenótipo CYP2D6 | 0,363 |

Analisando-se os p-valores da Tabela 2, é possível concluir que não rejeita-se a hipótese de independência entre as variáveis (Sexo, Local de infecção, ocupação e fenótipo CYP2D6) com recorrência de clínica malária. Para melhor visualização das proporções de Recorrência clínica de malária com relação as demais variáveis foi sintetizado os gráficos a seguir:

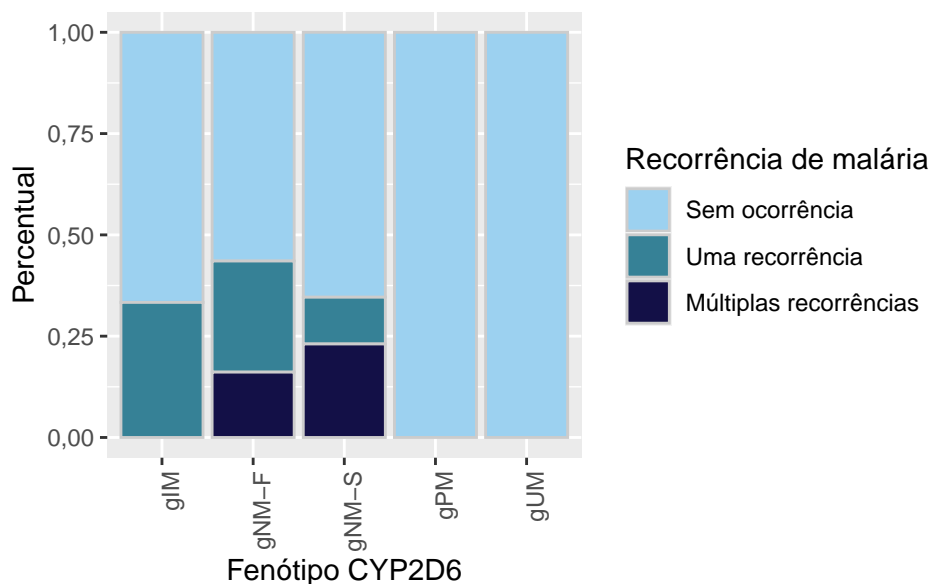
A partir do Figura 3, que relaciona as proporções relativas de recorrência de malária por sexo, pode-se notar que não há grande diferença entre os tipos de recorrência por sexo.

Figura 3 – Gráfico de barra para Recorrência clínica de malária com relação ao Sexo



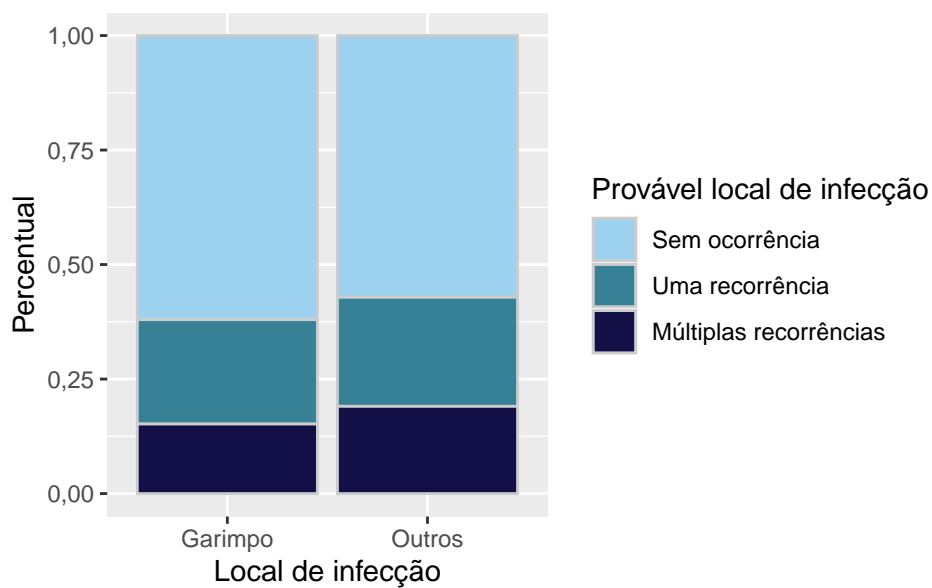
Observando-se o Figura 4 que relaciona Recorrência clínica de malária em relação ao Fenótipo CYP2D6 percebe-se que o Fenótipo gNM-F com maior proporção seria o de pelo menos uma recorrência de malária, já o que possui maior recorrência múltipla é o Fenótipo gNM-S.

Figura 4 – Gráfico de barra para Recorrência clínica de malária com relação o Fenótipo de CYP2D6



Partindo-se do Figura 5 de Recorrência clínica de malária, em relação ao provável local de infecção (Garimpo e outros locais de provável infecção), nota-se que não há muita diferença partindo da comparação da proporção de recorrência clínica com o possível local de infecção.

Figura 5 – Gráfico de barra para Recorrência clínica de malária com relação ao Provável local de infecção



3.2 Pergunta 1

” Indivíduos com status CYP2D6 alterada/POR mutada possuem menor redução de Pvs25 (0 horas para 72 horas; QPVS25D72/D0 ou QPVS25LOGD72/D0; alternativamente pode-se utilizar as variáveis medidas em cada tempo: QPVS25D0LOG e QPVS25D72LOG) quando comparados com indivíduos CYP2D6 normal/POR não-mutada?” - Texto elaborado pelas pesquisadoras.

Com o propósito de responder à pergunta proposta, construiu-se a Tabela 3, com três vertentes de sumariação dos Status de metabolização conjunto ou individual das variáveis CYP2D6/POR com relação à Razão de Pvs25:

Table 3: Análise descritiva da variável Razão de Pvs25

| Status | Eliminação de gametócitos | | | | |
|-----------------------------------|---------------------------|---------|---------|---------|---------------|
| | Mínimo | Máximo | Média | Mediana | Desvio Padrão |
| Metabolização CYP2D6 e POR | | | | | |
| Normal/Não-mutada | -4,7859 | -0,2127 | -2,5537 | -2,4354 | 1,4322 |
| Alterada/Não-mutada | -4,1551 | 1,5297 | -2,0146 | -2,4667 | 1,7635 |
| Normal/Mutada | -5,0000 | 3,8476 | -1,4474 | -1,5504 | 2,3553 |
| Alterada/Mutada | -4,8144 | 1,0335 | -2,5045 | -2,6610 | 1,9327 |
| Metabolização CYP2D6 | | | | | |
| Atividade alterada | -5,0000 | 1,5297 | -2,3124 | -2,7783 | 1,9041 |
| Atividade normal | -5,0000 | 3,8476 | -2,7880 | -3,0697 | 1,8963 |
| Mutação do gene POR | | | | | |
| Presença de mutação | -5,0000 | 3,8476 | -2,0963 | -2,2196 | 2,0549 |
| Ausência de mutação | -5,0000 | 2,9108 | -2,8644 | -3,0640 | 1,8842 |

Além disso, para melhor visualização dos dados, sintetizou-se os gráficos a seguir:

Figura 6 – Gráfico de boxplot da Razão de Pvs25 em relação ao status de metabolização conjunto CYP e POR

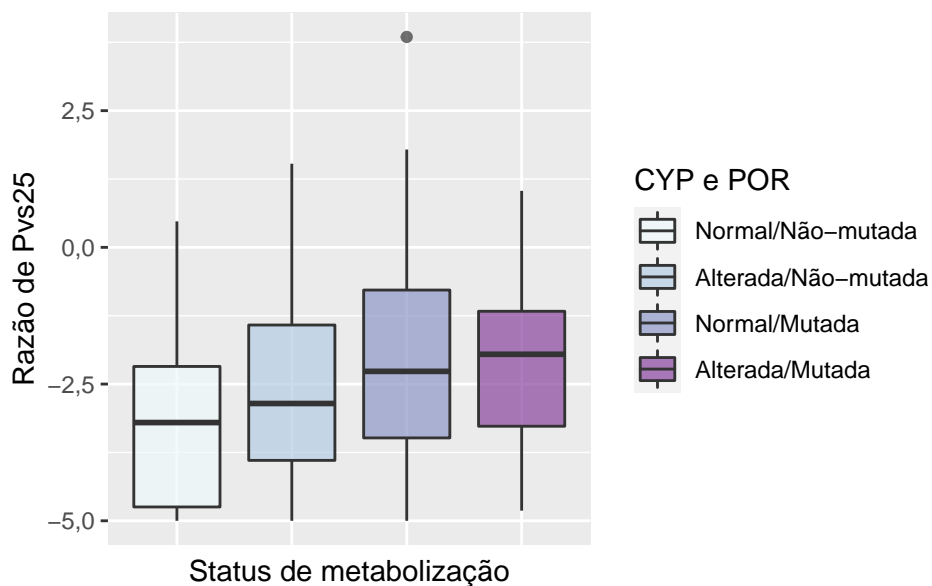


Figura 7 – Gráfico de boxplot da Razão de Pvs25 em relação ao status de metabolização CYP2D6

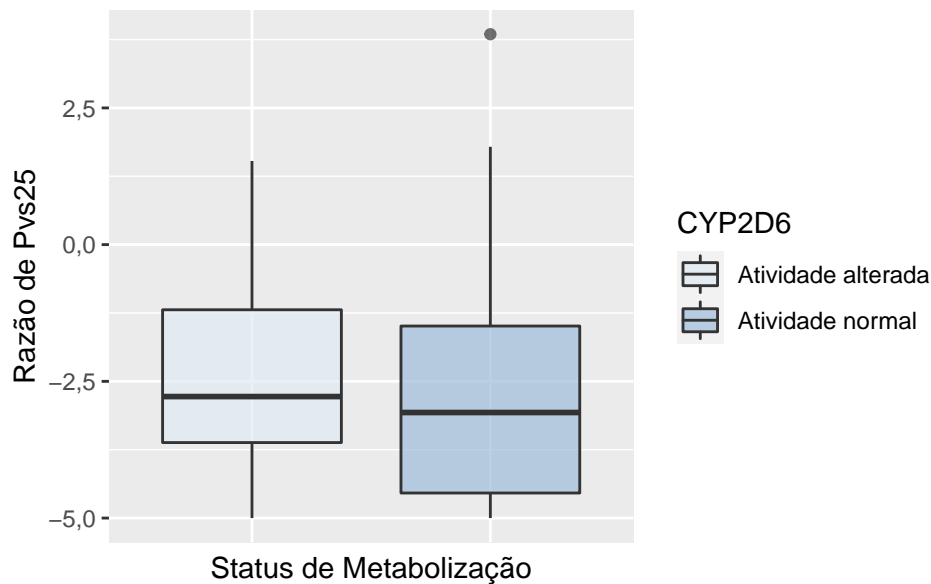
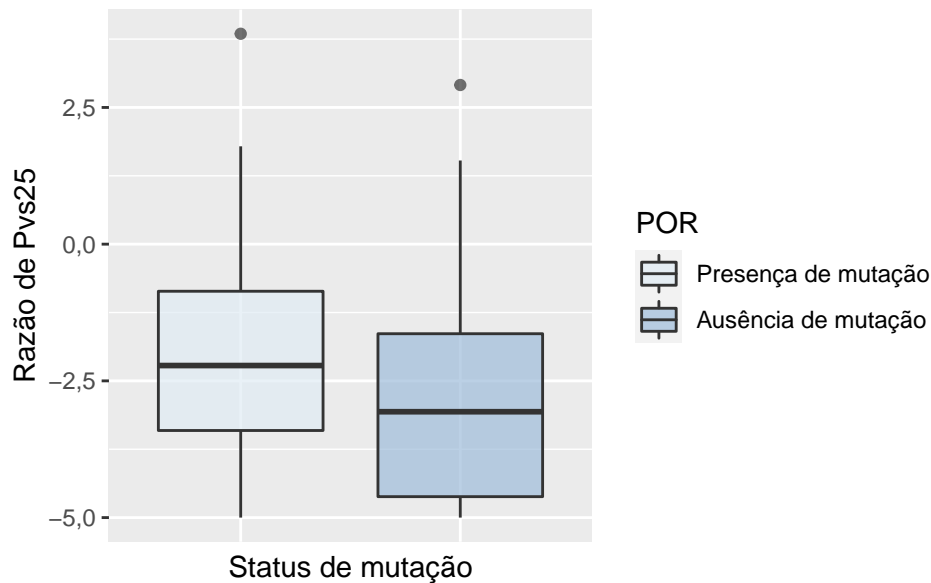


Figura 8 – Gráfico de boxplot da Razão de Pvs25 em relação ao status de mutação do gene POR



Nesse sentido, através da análise construída com base nas Figuras 6, 7 e 8, e na Tabela 3, é possível visualizar que aparentemente não há diferenças significativas quanto ao valor mediano e a variabilidade da Razão de Pvs25, com relação às categorias dos Status de metabolização/mutação conjunto ou individual de CYP2D6/POR. Entretanto, para uma análise mais conclusiva, testes de média e mediana foram realizados, por meio dos métodos não-paramétricos: Teste de Kruskal-Wallis e Mann-Whitney. Em todos os testes realizados buscou-se analisar se existe ou não diferença significativa na média ou na mediana de eliminação de gametócitos.

Desse modo assumimos como hipóteses de interesse:

1º teste) Teste de Kruskal-Wallis:

H_0 : As Razões médias de Pvs25 para cada categoria dos Status de metabolização conjunto são iguais;

H_1 : As Razões médias de Pvs25, para pelo menos uma das categorias dos Status de metabolização conjunto, é diferente;

2º teste) Teste de Mann-Whitney:

H_0 : As Razões medianas de Pvs25 para cada categoria dos Status de metabolização individual de CYP2D6 são iguais;

H_1 : As Razões medianas de Pvs25 para cada categoria dos Status de metabolização individual de CYP2D6 são iguais;

3º teste) Teste de Mann-Whitney:

H_0 : As Razões medianas de Pvs25 para cada categoria dos Status de mutação individual do gene POR são iguais;

H_1 : As Razões medianas de Pvs25 para cada categoria dos Status de mutação individual do gene POR são diferentes;

Table 4: Teste de médias e medianas para Razão de Pvs 25

| Métodos | Variáveis | P-Valores |
|-------------------------|--------------|-----------|
| Teste de Kruskal-Wallis | CYP2D6 e POR | 0,573 |
| Teste de Mann-Whitney | CYP2D6 | 0,216 |
| Teste de Mann-Whitney | POR | 0,0667 |

Os p-valores apresentados na Tabela 4, revelam que em nenhum dos casos a hipótese nula de igualdade das médias ou medianas da Razão de Pvs25, com relação às categorias dos Status de metabolização conjunto/individual de CYP2D6 e POR. Desse modo, pode-se concluir que não há evidências estatísticas de que “Indivíduos com status CYP2D6 alterada/POR mutada possuem menor redução de Pvs 25”. - Texto elaborado pelas pesquisadoras.

3.3 Pergunta 2

“Avaliar se a parasitemia inicial (PV18SD0NORM ou PV18SD0NORMLOG) está associada com a quantidade de gametócitos e se é um fator de confusão nas análises de associação entre status CYP/POR e eliminação de gametócitos realizadas no item 1.” - Texto elaborado pelas pesquisadoras.

Com o intuito de compreender o comportamento das variáveis numéricas, envolvidas na pergunta, foram elaborados histogramas sobre as variáveis Razão de Pvs25 e Quantidade de parasitos no momento inicial:

Figura 9 – Histograma da Razão de Pvs25

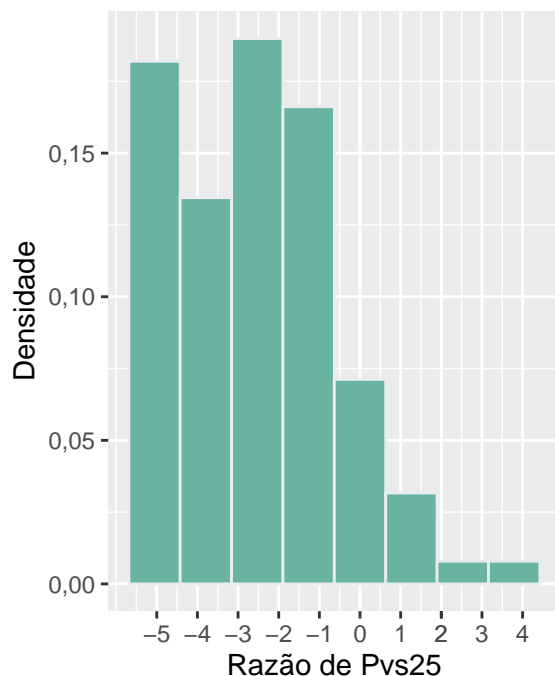
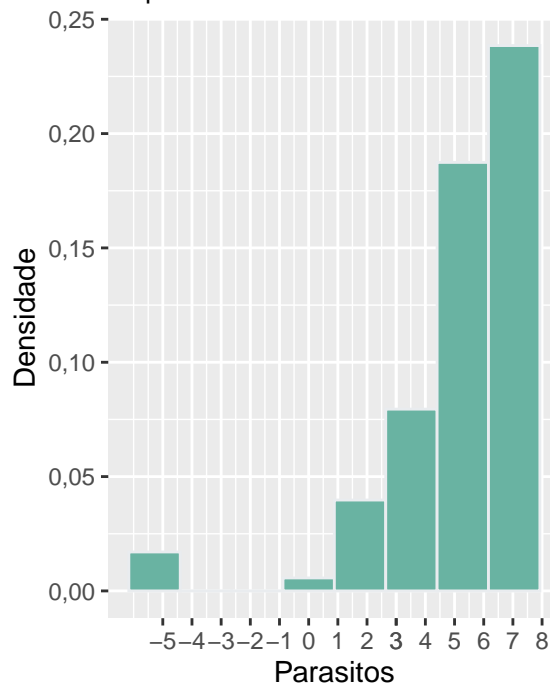


Figura 10 – Histograma da Quantidade de parasitos no momento inicial



Por meio das Figuras 9 e 10, visualiza-se que o possível fator de confusão analisado (Quantidade de parasitos no momento inicial) possui uma distribuição diferente da distribuição da variável resposta analisada no item 1 (Razão de Pvs25). A relação dessas duas variáveis conjuntamente, será analisada por meio de testes estatísticos.

Posteriormente, sintetizou-se gráficos de dispersão para visualização da relação entre as variáveis Razão de Pvs25 (Eliminação de gametócitos) e Quantidade de Parasitos (0 horas), com relação aos status de metabolização conjunto e individuais:

Figura 11 – Gráfico de dispersão de Parasitos no momento inicial e da Razão de Pvs25, com relação ao status de metabolização conjunc

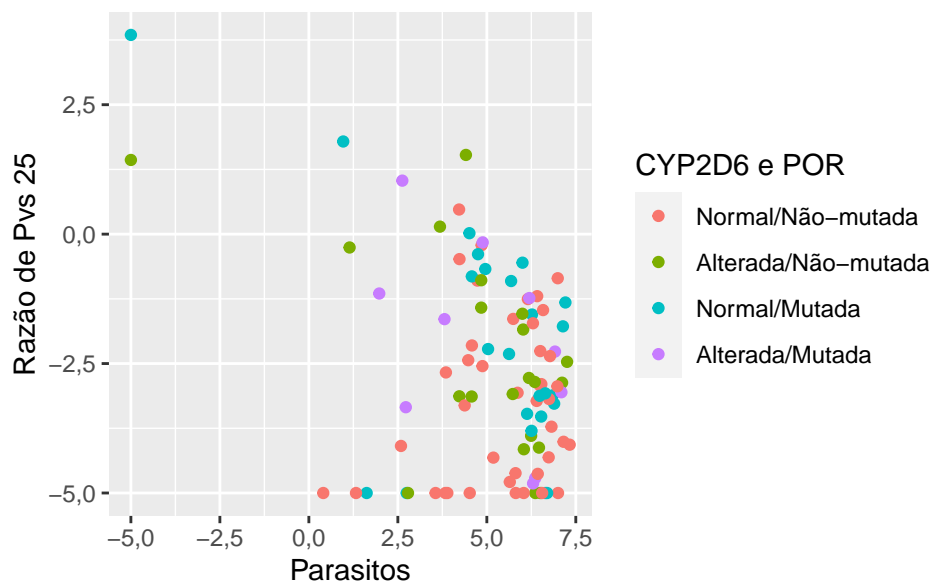


Figura 12 – Gráfico de dispersão de Parasitos no momento inicial e da Razão de Pvs25, com relação ao status CYP2D6

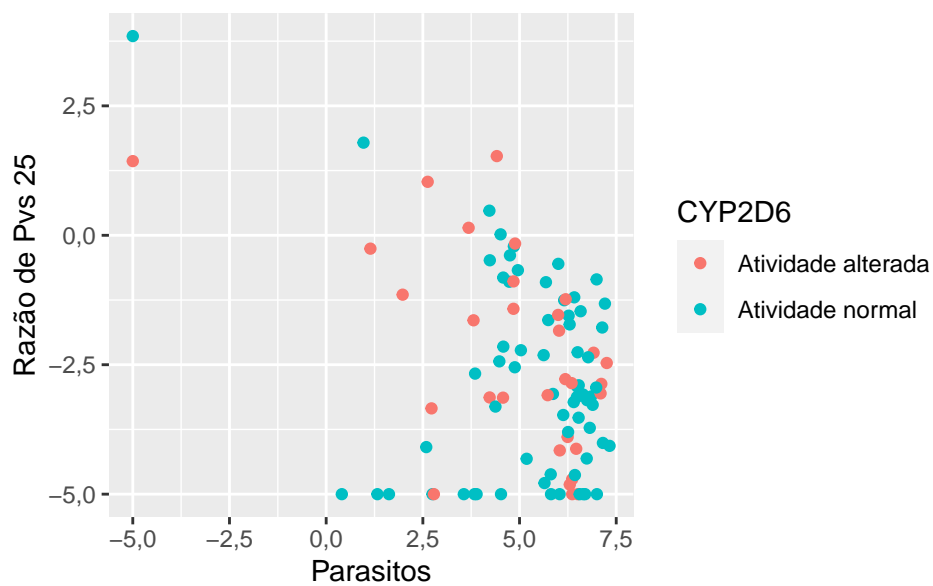
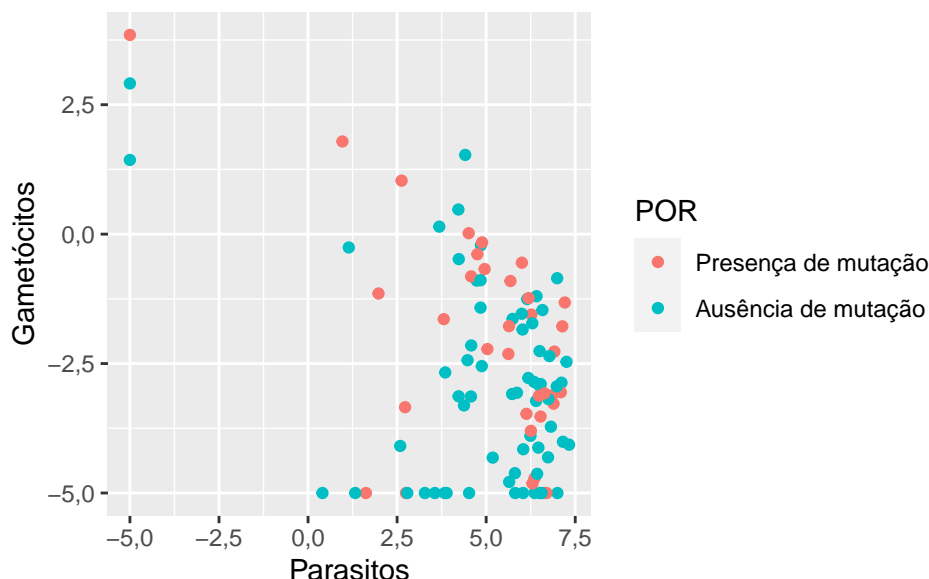


Figura 13 – Gráfico de dispersão de Parasitos no momento inicial e da Razão de Pvs25, com relação ao status POR



Analisando as Figuras 11, 12 e 13, não é possível observar uma tendência linear na dispersão dos Status de metabolização CYP2D6/POR conjunto/individual com relação à Quantidade de parasitos no momento inicial e a Razão de Pvs25.

Como comentado anteriormente, visando testar estatisticamente se existe uma associação entre as variáveis de interesse, postulou-se testes de independencia qui-quadrado, teste de correlação e testes de médias/medanas, os quais possuem as seguintes hipótese de interesse:

1º teste) Teste qui-quadrado de independência

Esse teste é utilizado para verificar se existe associação entre duas variáveis qualitativas:

Hipóteses:

H_0 : As variáveis Status de metabolização CYP2D6 e Status de mutação do gene POR (PORMUT) são independentes

H_1 : As variáveis Status de metabolização CYP2D6 e Status de mutação do gene POR (PORMUT) estão associadas

2º teste) Teste de Correlação de Spearman

Esse teste é utilizado quando não existe normalidade e/ou não existe relação linear, deve ser usado quando não se deseja utilizar nenhuma suposição de normalidade ou da presença de qualquer outra distribuição para a variável ou para a estatística de teste

Hipóteses:

H_0 : A correlação entre as variáveis, Eliminação de gametócitos e quantificação de parasitos no momento inicial, é zero/nula;

H_1 : A correlação entre as variáveis, Eliminação de gametócitos e quantificação de parasitos no momento inicial, não é nula.

3º teste) Teste de Kruskal Wallis

Hipóteses:

H_0 : Se em todos os níveis CYP2D6 e POR conjunto as médias de Gametócitos são iguais;

H_1 : Se em pelo menos um nível CYP2D6 e POR conjunto as médias de Gametócitos é diferente.

4º teste) Teste de Kruskal Wallis

Hipóteses:

H_0 : Se em todos os níveis CYP2D6 e POR conjunto as médias de Parasitos são iguais;

H_1 : Se em pelo menos um nível CYP2D6 e POR conjunto as médias de Parasitos é diferente.

5º teste) Teste de Mann-Whitney

Hipóteses:

H_0 : A mediana de Gametócitos é igual em CYP2D6 (Alterada e não alterada);

H_1 : A mediana de Gametócitos é diferente em CYP2D6 (Alterada e não alterada).

6º teste) Teste de Mann-Whitney

Hipóteses:

H_0 : A mediana de Parasitos é igual em CYP2D6 (Alterada e não alterada);

H_1 : A mediana de Parasitos é diferente em CYP2D6 (Alterada e não alterada).

7º teste) Teste de Mann-Whitney

Hipóteses:

H_0 : A mediana de Gametócitos é igual em POR (Mutada e não mutada);

H_1 : A mediana de Gametócitos é diferente em POR (Mutada e não mutada).

8º teste) Teste de Mann-Whitney

Hipóteses:

H_0 : A mediana de Parasitos é igual em POR (Mutada e não mutada);

H_1 : A mediana de Parasitos é diferente em POR (Mutada e não mutada).

Table 5: Análise da associação das variáveis principais

| Métodos | Variáveis | P-Valores |
|-------------------------|--------------------------|-----------|
| Teste de independência | CYP2D6 e POR | 0,867 |
| Teste de Spearman | Gametócitos e Parasitos | 0,0454 |
| Teste de Kruskal-Wallis | Gametócitos e CYP2D6/POR | 0,238 |
| Teste de Kruskal-Wallis | Parasitos e CYP2D6/POR | 0,0816 |
| Teste de Mann-Whitney | Gametócitos e CYP2D6 | 0,216 |
| Teste de Mann-Whitney | Parasitos e CYP2D6 | 0,307 |
| Teste de Mann-Whitney | Gametócitos e POR | 0,0667 |
| Teste de Mann-Whitney | Parasitos e POR | 0,806 |

Note:

Gametócitos= Razão de Pvs25

Parasitos= Quantidade de parasitos no momento inicial

CYP2D6/POR= Status de metabolização conjunto de CYP2D6 e POR

Pela Tabela 5, observa-se que o único teste que resultou um p-valor significativo foi o teste de Spearman sobre Razão de Pvs25 (Gametócitos) e Quantidade de Parasitos no momento inicial (Parasitos), assim é possível concluir que há evidências de que essas variáveis são correlacionadas (associadas entre si). Desse modo, esse quesito indica que a variável Quantidade de parasitos (0 horas) é um possível fator

de confusão para as análises construídas no item 1. Contudo, outras análises serão realizadas visando verificar a relação dessa variável com Status de metabolização CYP2D6/POR individual/conjunto.

A seguir se encontra a Tabela 6 com as medidas resumos da quantidade de parasitos no momento inicial com relação ao status de metabolização conjunto (CYP2D6 e Por); status de metabolização CYP2D6; e status de mutação do gene POR.

Table 6: Análise descritiva da variável Quantidade de parasitos (momento inicial)

| Status | Quantidade de parasitos | | | | |
|-----------------------------------|-------------------------|--------|--------|---------|---------------|
| | Mínimo | Máximo | Média | Mediana | Desvio Padrão |
| Metabolização CYP2D6 e POR | | | | | |
| Normal/Não-mutada | 3,8509 | 7,3274 | 5,8991 | 6,2959 | 1,0361 |
| Alterada/Não-mutada | 1,1441 | 7,2548 | 5,3957 | 6,0217 | 1,6502 |
| Normal/Mutada | -5,0000 | 7,2076 | 4,7823 | 5,6454 | 3,2574 |
| Alterada/Mutada | 2,6256 | 7,0954 | 5,2529 | 6,2468 | 1,8800 |
| Metabolização CYP2D6 | | | | | |
| Atividade alterada | -5,0000 | 7,2548 | 4,8602 | 5,9987 | 2,4543 |
| Atividade normal | -5,0000 | 7,3274 | 5,2808 | 5,9335 | 2,0448 |
| Mutação do gene POR | | | | | |
| Presença de mutação | -5,0000 | 7,2076 | 5,0214 | 6,0026 | 2,4471 |
| Ausência de mutação | -5,0000 | 7,3274 | 5,0366 | 5,8137 | 2,3650 |

3.3.1 Regressão múltipla

A ideia dos modelos de regressão construídos é averiguar a contribuição da adição da variável “Quantidade de parasitos (momento inicial)” nos modelos, antes explicados apenas pelas variáveis status de metabolização/mutação CYP2D6 e POR.

Nesse sentido, foram sintetizados seis modelos de Regressão linear múltipla. Em dois desses modelos não foi obtido nenhuma variável explicativa significativa no modelo. Vale salientar que os dois modelos, construídos sem nenhuma variável explicativa significativa, eram explicados apenas pelas variáveis explicativas: Status de Metabolização CYP2D6; e Status de Mutação do gene POR.

A seguir segue os quatro modelos que possuem pelo menos uma variável explicativa significativa para explicar a média da Eliminação de gametócitos:

Modelo 1: Status de metabolização conjunto

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i$

Modelo ajustado: $Y_i = -3,1846 + 0,7875x_{1i} + 1,0909x_{2i} + 1,0503x_{3i} + e_i$

Legenda:

β_0 : Intercepto; *significativa*

$\beta_1 x_{1i}$: CYP2D6PORAAlterada/Não-mutada;

$\beta_2 x_{2i}$: CYP2D6PORNNormal/Mutada; *significativa*

$\beta_3 x_{3i}$: CYP2D6PORAAlterada/Mutada;

ϵ_i : Erro aleatório.

Modelo 2: Status de Metabolização conjunto + parasitos

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i$

Modelo ajustado: $Y_i = -1,5040 + 0,6118x_{1i} + 0,9787x_{2i} + 0,8877x_{3i} - 0,3105x_{4i} + e_i$

Legenda:

β_0 : Intercepto; *significativa*

β_1 : CYP2D6PORAAlterada/Não-mutada;

β_2 : CYP2D6PORNORMAL/Mutada; *significativa*

β_3 : CYP2D6PORAAlterada/Mutada;

β_4 : PV18SD0NORM; *significativa*

ϵ_i : Erro aleatório.

Modelo 3: Status de Metabolização CYP + parasitos

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$ Modelo ajustado: $Y_i = -0,7459 - 0,3400x_{1i} - 0,3223x_{2i} + e_i$

Legenda:

β_0 : Intercepto;

β_1 : Atividade CYP normal;

β_2 : PV18SD0NORM; *significativa*

ϵ_i : Erro aleatório.

Modelo 4: Status de Mutação POR + parasitas

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$

Modelo ajustado: $Y_i = -0,3011 - 0,7627x_{1i} - 0,3575x_{2i} + e_i$

Legenda:

β_0 : Intercepto;

β_1 : Ausência de mutação POR; *significativa*

β_2 : PV18SD0NORM; *significativa*

ϵ_i : Erro aleatório.

Visando analisar se a variável quantificação de parasitas no momento inicial (0 horas) é fator de confusão para as análises os modelos de regressão construídos, analisou-se os valores dos coeficientes de determinação ajustados:

Table 7: Análise da contribuição da variável 'Parasitas' para os modelos de regressão

| <i>Variáveis explicativas</i> | <i>Coefficiente de determinação ajustado</i> |
|---|---|
| Modelo 1 | |
| Status de metabolização conjunto | 0,0373 |
| Modelo 2 | |
| Status de metabolização conjunto + Parasita | 0,157 |
| Modelo 3 | |
| Status de metabolização CYP2D6 + Parasita | 0,131 |
| Modelo 4 | |
| Status de mutação POR + Parasita | 0,206 |

Note:

Parasitas= Quantidade de parasitas no momento inicial

A partir da Tabela 7 conseguimos analisar a contribuição da variável Quantidade de Parasitas no momento inicial (Parasita) para os modelos de regressão com algum coeficiente significativo. À vista disso, é possível concluir que os modelos que contavam com a adição dessa variável explicativa, obtiveram valores maiores do coeficiente de determinação ajustado. Testamos também o pressuposto de normalidade dos resíduos dos modelos construídos, e quanto a análise dos p-valores, observou-se que o Modelo 1 foi o único que teve o pressuposto de normalidade dos resíduos rejeitado (p valor $< 0,05$).

Portanto, resumizando os resultados obtidos na Tabela 7, é possível concluir que a variável Quantidade de parasitas no momento inicial gera uma melhora na capacidade explicativa dos modelos de regressão. No entanto, importa salientar que essa melhora não é tão significativa em termos estatísticos, haja vista que o coeficiente de correlação da relação das variáveis numéricas Quantidade de parasitas no momento inicial e Razão de Pvs25 é -0,433, indicando uma correlação moderada entre as variáveis.

3.4 Pergunta 3

“Avaliar se sexo e idade dos pacientes são fatores de confusão nas análises de associação entre status CYP/POR e eliminação de gametócitos realizadas no item 1.” - Texto elaborado pelas pesquisadoras.

Com o propósito de visualizar se as variáveis Sexo e Idade são fatores de confusão nos resultados obtidos no item 1, primeiramente foram construídos boxplots (Figura 14 e 15) com as variáveis Sexo e Idade:

Figura 14 – Gráfico de boxplot da Razão de Pvs25 em relação ao Sexo

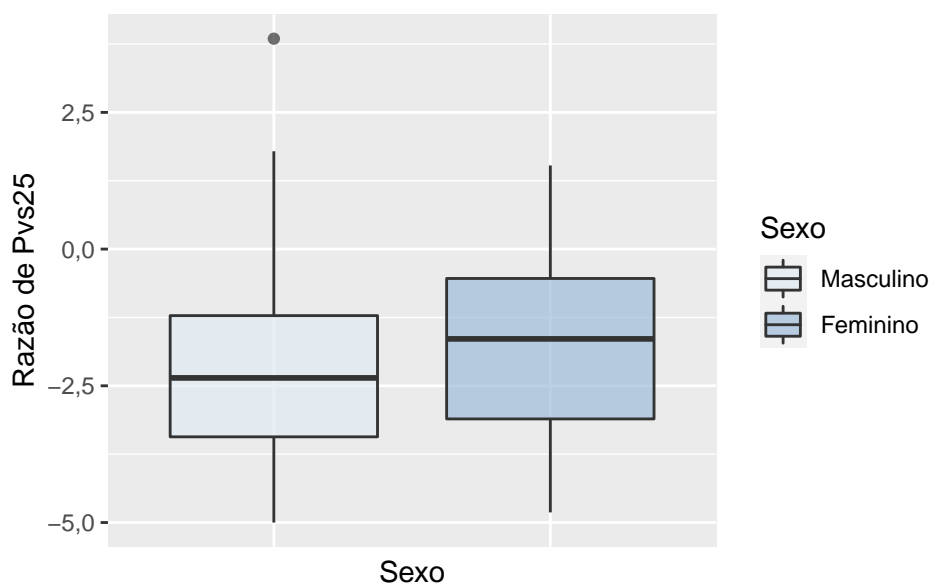
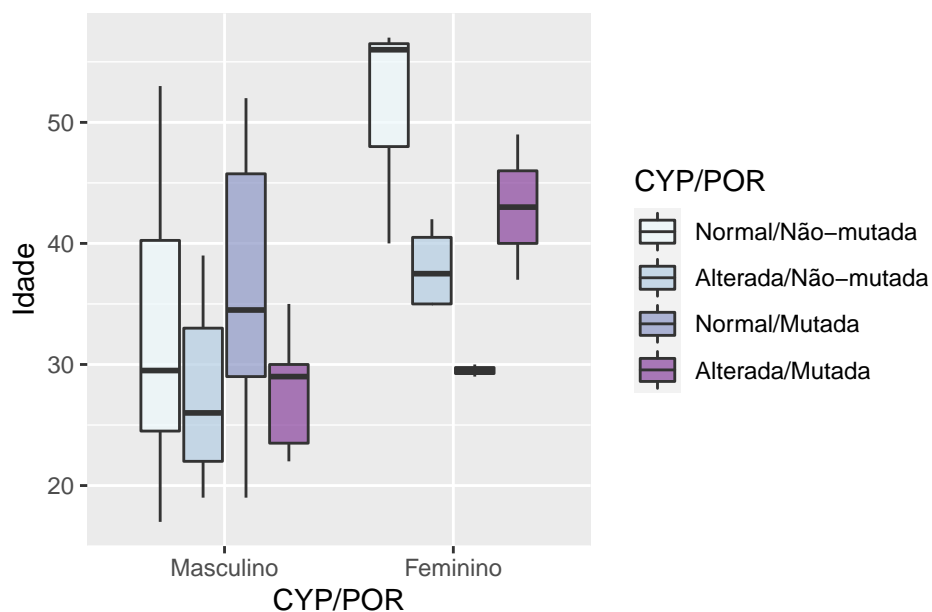


Figura 15 – Gráfico de boxplot Idade por CYP/POR conjunto



Analisando a Figura 14, visualiza-se que aparentemente a mediana da Razão de Pvs25 do grupo feminino não parece ser diferente da Razão de Pvs25 do grupo masculino. Além disso, observa-se uma certa

similaridade na variabilidade dos grupos. Vale salientar que testes estatísticos serão desenvolvidos posteriormente, visando verificar se existem diferenças significativas entre os grupos de análise.

Quanto à Figura 15, como a variável Idade foi plotada em termos de uma variável categórica de quatro níveis será realizado nessa sessão, um teste de Kruskal-Wallis com o propósito de proporcionar uma análise mais acertiva da relação dessas duas variáveis.

Como comentado anteriormente, visando testar estatisticamente se existe uma associação entre as variáveis, postulou-se testes estatísticos, cujos resultados foram apresentados na Tabela 8. As hipóteses de interesse dos testes realizados foram:

1º teste) Teste qui-quadrado de independência:

Esse teste é utilizado para verificar se existe associação entre as duas variáveis qualitativas:

Hipóteses:

H_0 : As variáveis sexo e status de metabolização conjunto são independentes;

H_1 : As variáveis sexo e status de metabolização conjunto são dependentes.

2º teste) Teste de Correlação de Spearman

Hipóteses:

H_0 : A correlação entre as variáveis Idade e Razão de Pvs25 é nula;

H_1 : A correlação entre as variáveis idade e razão de Pvs25 não é nula.

3º teste) Teste de Mann-Whitney

Hipóteses:

H_0 : A mediana da Idade masculina é igual à mediana da Idade feminina;

H_1 : A mediana da Idade masculina é diferente à mediana da Idade feminina.

4º teste) Teste de Mann-Whitney

Hipóteses:

H_0 : A mediana da Razão de Pvs25 do grupo masculina é igual à mediana da Razão de Pvs25 no grupo feminino;

H_1 : A mediana da Razão de Pvs25 do grupo masculina é diferente à mediana da Razão de Pvs25 no grupo feminino

5º teste) Teste de Kruskal-Wallis:

H_0 : As Idades médias para cada categoria dos Status de metabolização conjunto são iguais;

H_1 : As Idades médias, para pelo menos um dos Status de metabolização conjunto, é diferente;

Table 8: Análise da associação das variáveis principais

| Métodos | Variáveis | P-Valores |
|-------------------------|-----------------------------|-----------|
| Teste de independência | Sexo e CYP2D6/POR conjunto | 0,76 |
| Teste de Spearman | Idade e Razão de Pvs25 | 0,981 |
| Teste de Mann-Whitney | Idade e sexo | 0,0707 |
| Teste de Mann-Whitney | Sexo e Razão de Pvs25 | 0,567 |
| Teste de Kruskal-Wallis | Idade e CYP2D6/POR conjunto | 0,82 |

Os p-valores apresentados na Tabela 8, revelam que em nenhum dos casos há uma associação significativa entre as variáveis. Apesar dos resultados dos testes de hipótese, ainda não é possível concluir que

as variáveis Idade e Sexo não são fatores de confusão dos resultados obtidos no item 1. À vista disso, construiu-se modelos de regressão múltipla, com o propósito de averiguar a contribuição das variáveis Idade e Sexo ao modelo explicado apenas pela variável Status de metabolização CYP2D6/POR.

3.4.1 Regressão múltipla

A pergunta 1 foi desenvolvida tendo como base a análise da relação entre as variáveis: Razão de Pvs 25 e Status de metabolização conjunto CYP2D6/POR. Como exposto o modelo de regressão múltipla que rege a relação dessas variáveis é:

Modelo item 1

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i$

Modelo ajustado:

$$Y_i = -3,1846 + 0,7875x_{1i} + 1,0909x_{2i} + 1,0503x_{3i} + e_i$$

Legenda:

Y_i : Razão de Pvs25

β_0 : Intercepto; *significativa*

β_1 : CYP2D6PORAAlterada/Não-mutada;

β_2 : CYP2D6PORNNormal/Mutada; *significativa*

β_3 : CYP2D6PORAAlterada/Mutada;

ϵ_i : Erro aleatório.

Visando averiguar se as variáveis Idade e Sexo contribuem para o modelo, foi sintetizado outro modelo de regressão múltipla, o qual pode ser representado pela equação:

Modelo com adição das variáveis explicativas: Idade e Sexo

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \epsilon_i$

Modelo ajustado:

$$Y_i = -3,228471 + 0,728832x_{1i} + 1,205164x_{2i} + 1,022845x_{3i} + 0,001203x_{4i} + 0,099869x_{5i} + e_i$$

Legenda:

Y_i : Razão de Pvs25

β_0 : Intercepto; *significativa*

β_1 : CYP2D6PORAAlterada/Não-mutada;

β_2 : CYP2D6PORNNormal/Mutada; *significativa*

β_3 : CYP2D6PORAAlterada/Mutada;

β_4 : IDADE

β_5 : SEXOFeminino

ϵ_i : Erro aleatório.

Ressalta-se que a normalidade dos resíduos, desse modelo de regressão múltipla, foi testada. E obteve-se um p-valor pequeno (0,0185), o que indica que o modelo construído não está bem ajustado.

Além disso, com o propósito de avaliar a contribuição das variáveis Sexo e Idade para o modelo de regressão múltipla do item 1, a Tabela 9 foi construída, com os coeficientes de determinação ajustado dos dois modelos:

Table 9: Análise da contribuição das variáveis Idade e Sexo para o modelo de regressão

| <i>Coefficiente de determinação ajustado</i> |
|---|
| Modelo item 1 |
| 0,0373 |
| Modelo com adição de Idade e Sexo |
| 0,0216 |

Portanto, é possível concluir que além da falta de ajuste do modelo, a adição das variáveis Sexo e Idade é responsável pela diminuição da explicação de Razão de Pvs25, tendo em vista que o coeficiente de determinação do modelo construído com as novas variáveis, é menor do que o coeficiente de determinação construído apenas com a variável CYP/POR conjunto.

3.5 Pergunta 4

“Identificar se a concentração de primaquina em 72 horas (PQCONC ou PQCONCLOG) é influenciada pelo fenótipo de CYP2D6 (CYP2D6FENOTIPO), POR (PORMUT) e o status conjunto CYP2D6/POR (CYP2D6POR). Exemplo: Indivíduos com fenótipo gPM possuem maior concentração de primaquina do que indivíduos gNM-F?” - Texto elaborado pelas pesquisadoras.

Haja vista a variável Concentração de primaquina em 72 horas (PQCON) construiu-se boxplots, com o propósito de identificar se essa variável é influenciada pelas variáveis: Fenótipo de CYP2D6 (CYP2D6FENOTIPO), Status de mutação do gene POR (PORMUT), e Status de metabolização conjunto CYP2D6/POR (CYP2D6POR). Nesse sentido as Figuras 16, 17 e 18 foram construídas:

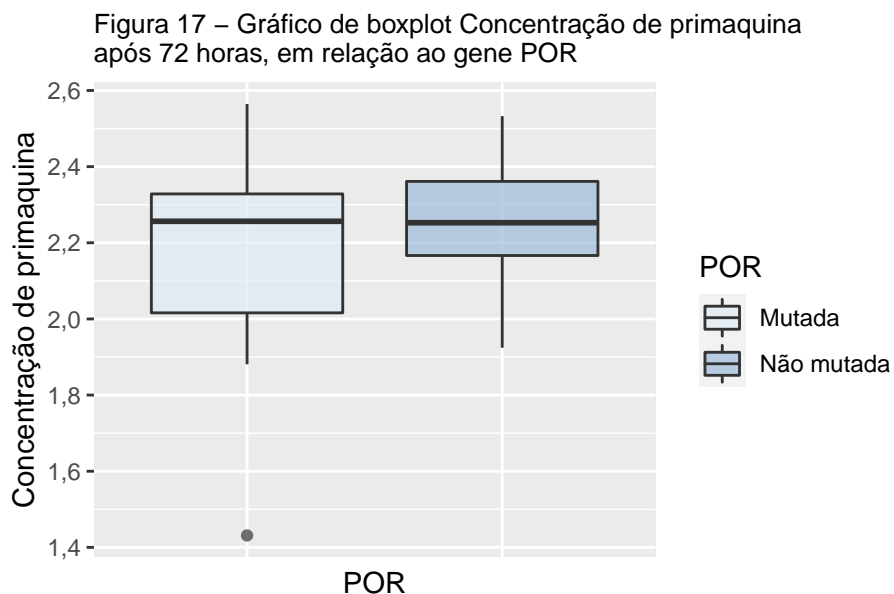
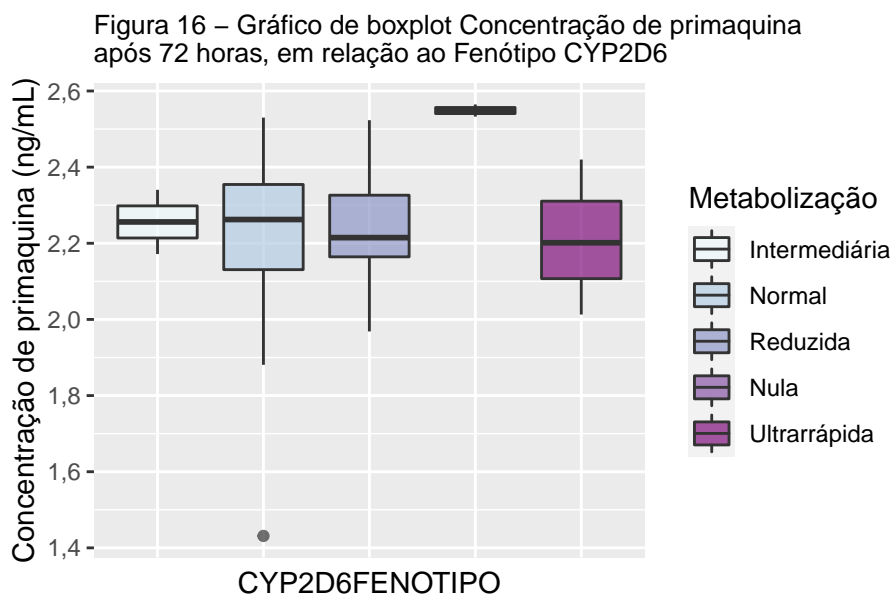
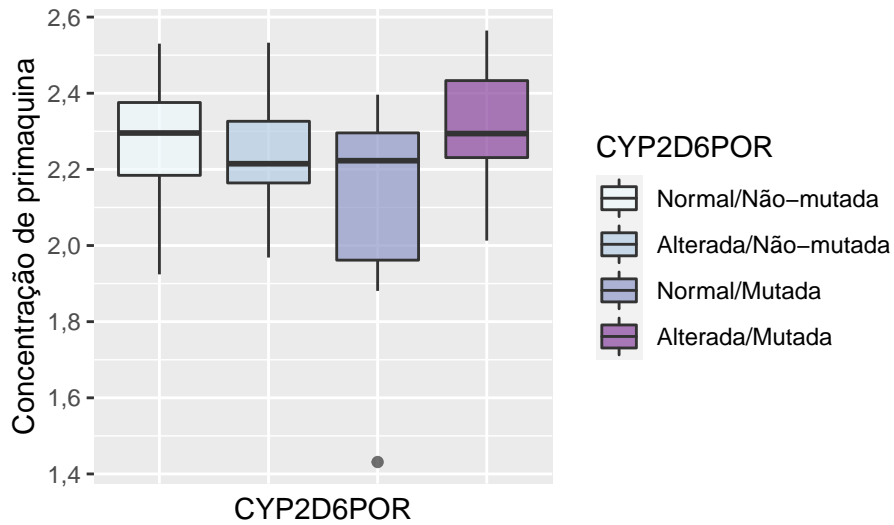


Figura 18 – Gráfico de boxplot Concentração de primaquina após 72 horas, em relação à CYP2D6



Nos gráficos 16, 17 e 18 é possível visualizar que aparentemente não há diferenças significativas quanto ao valor mediano e a variabilidade da Concentração de primaquina após 72 horas com relação às categorias das variáveis de interesse (CYP2D6FENOTIPO, PORMUT e CYP2D6POR). Entretanto, para uma análise mais conclusiva, testes de média e mediana foram realizados, por meio dos métodos não-paramétricos: Teste de Kruskal-Wallis e Mann-Whitney. Em todos os testes realizados buscou-se analisar se existe ou não diferença significativa na média ou na mediana da concentração de primaquina.

Desse modo, assumimos como hipóteses de interesse:

1º teste) Teste de Kruskal-Wallis:

H_0 : A Concentração média de primaquina, após 72 horas, para cada categoria dos Status de metabolização conjunto são iguais;

H_1 : A Concentração média de primaquina, após 72 horas, para pelo menos uma das categorias dos Status de metabolização conjunto, é diferente;

2º teste) Teste de Kruskal-Wallis:

H_0 : A Concentração média de primaquina, após 72 horas, para cada categoria do fenótipo CYP2D6 são iguais;

H_1 : A Concentração média de primaquina, após 72 horas, para pelo menos uma das categoria do fenótipo CYP2D6, é diferente;

3º teste) Teste de Mann-Whitney:

H_0 : A Concentração mediana de primaquina, após 72 horas, para cada categoria dos Status de mutação Status de mutação do gene POR são iguais;

H_1 : A Concentração mediana de primaquina, após 72 horas, para cada categoria dos Status de mutação do gene POR são diferentes;

Table 10: Teste de médias e medianas para Concentração de Primaquina após 72 horas

| Métodos | Variáveis | P-Valores |
|-------------------------|---------------------|-----------|
| Teste de Kruskal-Wallis | CYP2D6/POR conjunto | 0,148 |
| Teste de Kruskal-Wallis | Fenótipo CYP2D6 | 0,155 |
| Teste de Mann-Whitney | Metabolização POR | 0,68 |

Os p-valores apresentados na Tabela 10, revelam que em nenhum dos casos houve rejeição da hipótese nula de igualdade das médias ou medianas de Concentração de primaquina 72 horas. Essa análise é um indício de que talvez as variáveis analisadas (Status de metabolização conjunto CYP2D6/POR, Fenótipo CYP2D6, Status de mutação do ge POR) não exercem influência sobre a variável PQCONC. Entretanto, para uma análise mais precisa, foi postulados modelos de regressão múltipla com o propósito de verificar a contribuição de cada variável na explicação de um modelo cuja variável resposta é Concentração de primaquina após 72 horas.

3.5.1 Regressão Múltipla

Sobre demanda das pesquisadoras, modelos para explicação da concentração de primaquina em 72 horas (PQCONC) foram desenvolvidos. Nesse sentido, foram sintetizados seis modelos de Regressão linear múltipla. Em três desses modelos não foi obtido nenhuma variável explicativa significativa no modelo. Vale salientar que os três modelos, construídos sem nenhuma variável explicativa significativa, eram explicados pelas variáveis explicativas: CYP2D6FENOTIPO; CYP2D6FENOTIPO + idade + sexo; PORMUT. Além disto, é importante resaltar que rejeita-se a normalidade dos erros em todos os modelos de regressão, logo esses modelos não estão bem ajustados.

Modelo 1: CYP2D6FENOTIPO + Idade + Sexo

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \epsilon_i$

Modelo ajustado:

$$Y_i = 2,100350 - 0,058641x_{1i} - 0,022574x_{2i} + 0,316897x_{3i} - 0,118646x_{4i} + 0,005663x_{5i} - 0,105175x_{6i} + \epsilon_i$$

Legenda:

Y_i : concentração de primaquina em 72 horas (PQCONC);

β_0 : Intercepto; *significativa*

β_1 : CYP2D6FENOTIPONormal;

β_2 : CYP2D6FENOTIPOReducada;

β_3 : CYP2D6FENOTIPONula;

β_4 : CYP2D6FENOTIPOUltrarrápida;

β_5 : IDADE; *significativa*

β_6 : SEXOFeminino.

ϵ_i : Erro aleatório.

Modelo 2: Status de metabolização conjunto CYP/POR

Modelo de referência : $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i$

Modelo ajustado:

$$Y_i = 2,31564 - 0,08454x_{1i} - 0,09679x_{2i} - 0,20614x_{3i} - 0,118646x_{4i} + \epsilon_i$$

Legenda:

Y_i : concentração de primaquina em 72 horas (PQCONC);

β_0 : Intercepto; *significativa*

β_1 : CYP2D6PORAN;

β_2 : CYP2D6PORND;

β_3 : CYP2D6PORNM;*significativa*

β_4 : CYP2D6PORN;

ϵ_i : Erro aleatório.

Modelo 3: Status de metabolização CYP/POR + Idade + Sexo

Modelo de referência :

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \epsilon_i$$

Modelo ajustado:

$$Y_i = 2,193303 - 0,082986x_{1i} - 0,109217x_{2i} - 0,226303x_{3i} - 0,073692x_{4i} + 0,004362x_{5i} - 0,062416x_{6i} + \epsilon_i$$

Legenda:

Y_i : concentração de primaquina em 72 horas (PQCONC);

β_0 : Intercepto; *significativa*

β_1 : CYP2D6PORAN;

β_2 : CYP2D6PORND;

β_3 : CYP2D6PORNM;*significativa*

β_4 : CYP2D6PORN;

β_5 : IDADE;*significativa*

β_6 : SEXOFeminino.

ϵ_i : Erro aleatório.

Visando analisar a capacidade explicativa de cada modelo, calculou-se os respectivos coeficientes de determinação ajustados:

Table 11: Análise da contribuição das variáveis 'Idade e Sexo' para os modelos de regressão

| <i>Variáveis explicativas</i> | <i>Coeficiente de determinação ajustado</i> |
|--------------------------------------|--|
| Modelo 1 | |
| CYP2D6FENOTIPO + Idade + Sexo | 0,0934 |
| Modelo 2 | |
| Status conjunto CYP/POR | 0,0531 |
| Modelo 3 | |
| CYP/POR + Idade + Sexo | 0,0867 |

Note:

Variável resposta= Concentração de primaquina em 72 horas

CYP2D6FENOTIPO = Metabolização do gene CYP2D6

Status de metabolização conjunto

Status de mutação do gene POR

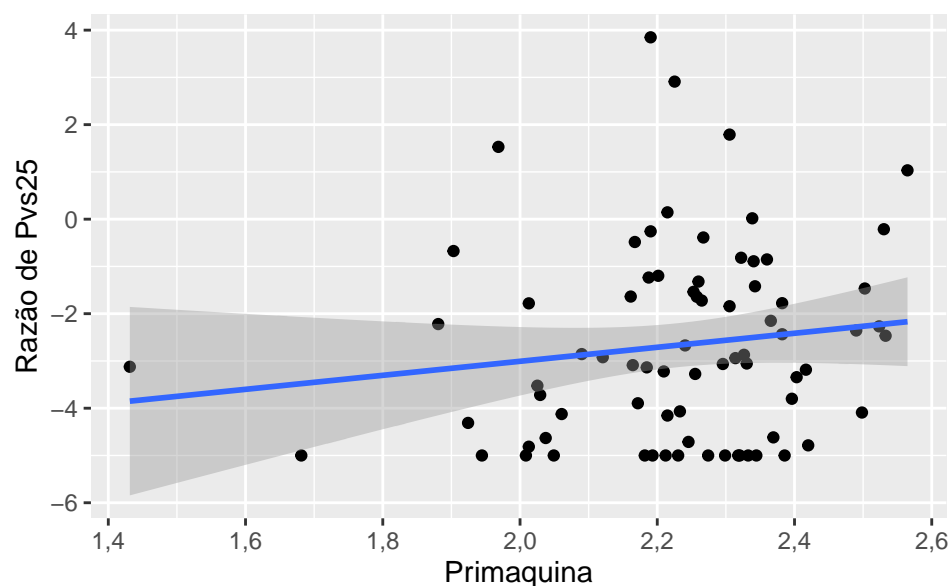
A partir da Tabela 11 pode-se analisar se as variáveis CYP2D6FENOTIPO e CYP/POR contribuem para um melhor ajuste dos modelos. Além disso, é possível avaliar a contribuição da adição das variáveis Sexo e Idade nos modelos. Nesse sentido, como os coeficientes de determinação ajustados são relativamente muito pequenos, é possível concluir que aparentemente essas não são variáveis influentes para explicação da variável Concentração de primaquina em 72 horas.

3.6 Pergunta 5

“Correlacionar a concentração de primaquina (PQCONC ou PQCONCLOG) com a redução de Pvs25 (QPVS25D72/D0 ou QPVS25LOGD72/D0; alternativamente pode-se utilizar as variáveis medidas em cada tempo: QPVS25D0LOG e QPVS25D72LOG). Exemplo: Aqueles indivíduos que possuem maior concentração de primaquina tiveram maior redução de Pvs25? Essas análises também podem ser realizadas considerando apenas presença/ausência de gametócitos em D0 e D72 (PVS25D0 e PVS25D72)”- Texto elaborado pelas pesquisadoras.

Primeiramente foi verificado, através de um diagrama de dispersão, se as variáveis Concentração de primaquina (ng/mL; fármaco estudado) em 72 horas e Razão de Pvs 25 apresentam algum padrão de linearidade, pois caso se confirme pode-se utilizar o coeficiente de correlação linear de Pearson para mensurar tal correlação. Nesse sentido, por meio da Figura 19, é possível verificar que as variáveis aparentemente não mostram padrão linear e sugere que não há correlação.

Figura 19 – Gráfico de dispersão Concentração de primaquina após 72 horas, em relação à Razão de Pvs25



À vista disso, prosseguimos a análise da relação das variáveis, utilizando o método de correlação de Spearman, o qual assume como hipóteses de interesse:

H_0 : A correlação entre as variáveis, Concentração de primaquina (ng/mL; fármaco estudado) em 72 horas e Razão de Pvs 25 (72 horas - 0 horas), é zero/nula;

H_1 : A correlação entre as variáveis, Concentração de primaquina (ng/mL; fármaco estudado) em 72 horas e Razão de Pvs 25 (72 horas - 0 horas), não é nula.

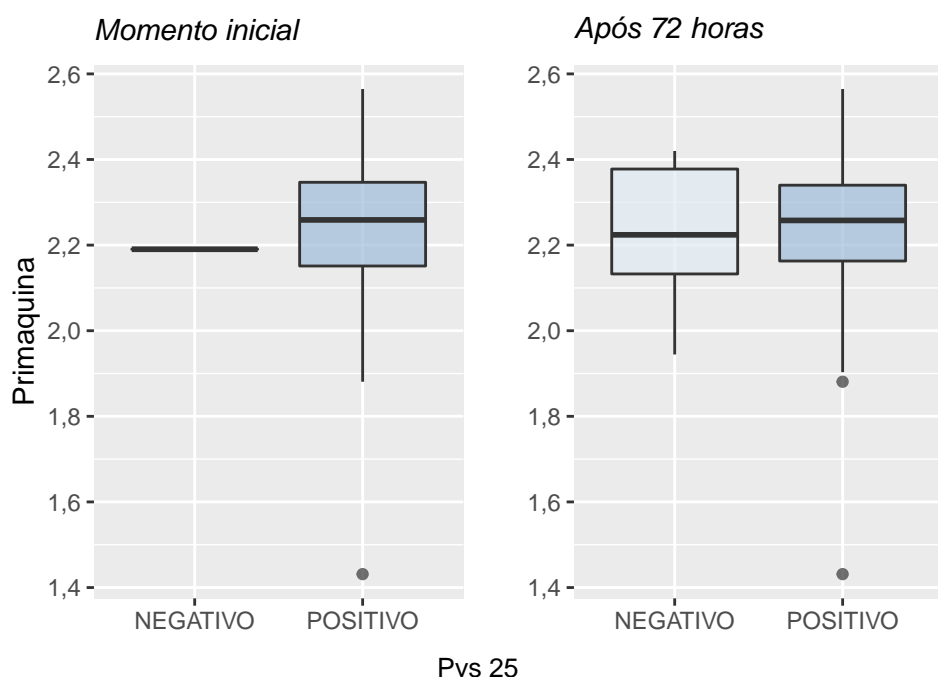
Analisando o p-valor obtido no teste (0,2485) não rejeita-se a hipótese nula do teste de Correlação de Spearman. Assim, não podemos concluir que indivíduos que apresentam uma maior razão de Pvs 25, possuem uma maior concentração de primaquina em 72 horas, tendo em vista que as variáveis não são variáveis correlacionadas.

Foi orientado também conduzir a análise tendo em vista a relação das variáveis: Concentração de primaquina (ng/mL; fármaco estudado) em 72 horas e Presença/ausência de gametócitos nos momentos 0 horas e 72 horas. Nesse sentido, a Tabela 12 foi construída visando obter uma descrição de como a variável Concentração de primaquina- 72 horas se distribui, tendo em vista os níveis das variáveis: Status de Pvs 25 no momento inicial e Status de Pvs 25 após 72 horas. O resultado de maior destaque na tabela seria o aumento da variabilidade do status Negativo de gametócitos, considerando os momentos 0 horas e 72 horas; tal variação é nítida na Figura 20.

Table 12: Análise descritiva da variável Concentração de primaquina- 72 horas, com relação ao Status de Pvs25

| Variável | Concentração de primaquina | | | | |
|-----------------|----------------------------|--------|--------|---------|---------------|
| | Mínimo | Máximo | Média | Mediana | Desvio Padrão |
| PVS25D0 | | | | | |
| NEGATIVO | 2,1903 | 2,1903 | 2,1903 | 2,1903 | 0,0000 |
| POSITIVO | 1,4314 | 2,5647 | 2,2304 | 2,2589 | 0,2052 |
| PVS25D72 | | | | | |
| NEGATIVO | 1,9445 | 2,4200 | 2,2233 | 2,2240 | 0,1733 |
| POSITIVO | 1,4314 | 2,5647 | 2,2299 | 2,2577 | 0,2077 |

Figura 20– Gráfico boxplot Concentração de primaquina com relação ao Status de Pvs25



Nesse sentido, com o propósito de responder o questionamento levantado no item 5, “Aqueles indivíduos que possuem maior concentração de primaquina tiveram maior redução de Pvs25?”, utilizamos o teste Mann-Whitney (unilateral), para comparação de médias de dois grupos independentes. Desse modo assumimos como hipóteses de interesse:

H_0 : A mediana da Concentração de primaquina 72 horas no Status de Pvs 25 é igual à mediana da Concentração de primaquina 72 horas no Status de Pvs 25 negativo;

H_1 : A mediana da Concentração de primaquina 72 horas no Status de Pvs 25 positivo é maior do que a mediana da Concentração de primaquina 72 horas no Status de Pvs 25 negativo.

Table 13: Teste Mann-Whitney para Concentração de primaquina- 72 horas

| Métodos | Variáveis | P-Valores |
|--------------------|--------------------------------|-----------|
| Teste Mann-Whitney | Status Pvs 25- momento inicial | 0,803 |
| Teste Mann-Whitney | Status Pvs 25- após 72 horas | 0,741 |

Analisando a Tabela 13, é possível concluir que a Concentração mediana de primaquina (72 horas) é a mesma para status negativo e positivo de Pvs 25, tanto no momento inicial quanto no final. Dessa forma, não pode-se afirmar que indivíduos que possuem maior concentração de primaquina tiveram maior redução de Pvs25.

3.7 Pergunta 6

“Verificar se os status de CYP2D6 (CYP2D6CAT) e POR (PORMUT) influenciam no número de recorrência clínica (RECIDIVACAT ou NRECORRENCIA). Exemplo: Indivíduos com CYP2D6 alterada possuem mais recorrências do que indivíduos com CYP2D6 normal? Indivíduos com POR mutada possuem mais recorrências do que POR nao-mutada?” - Texto elaborado pelas pesquisadoras.

Com o propósito de responder a pergunta, realizou-se primeiramente testes qui-quadrados de independência. Desse modo, analisamos se as variáveis (Recorrência de malária e Status de metabolização conjunto; Recorrência de malária e Status de metabolização CYP2D6; Recorrência de malária e Status de mutação do gene POR) são dependentes ou independentes. Vale salientar que duas variáveis são independentes se a ocorrência de uma não afeta a ocorrência da outra. Nesse sentido, vale salientar as hipóteses consideradas nos testes realizados:

H_0 : A variável Recorrência de malária não está associado com a variável Status de metabolização/Status de mutação (independentes);

H_1 : A variável Recorrência de malária está associado com a variável Status de metabolização/Status de mutação (dependentes).

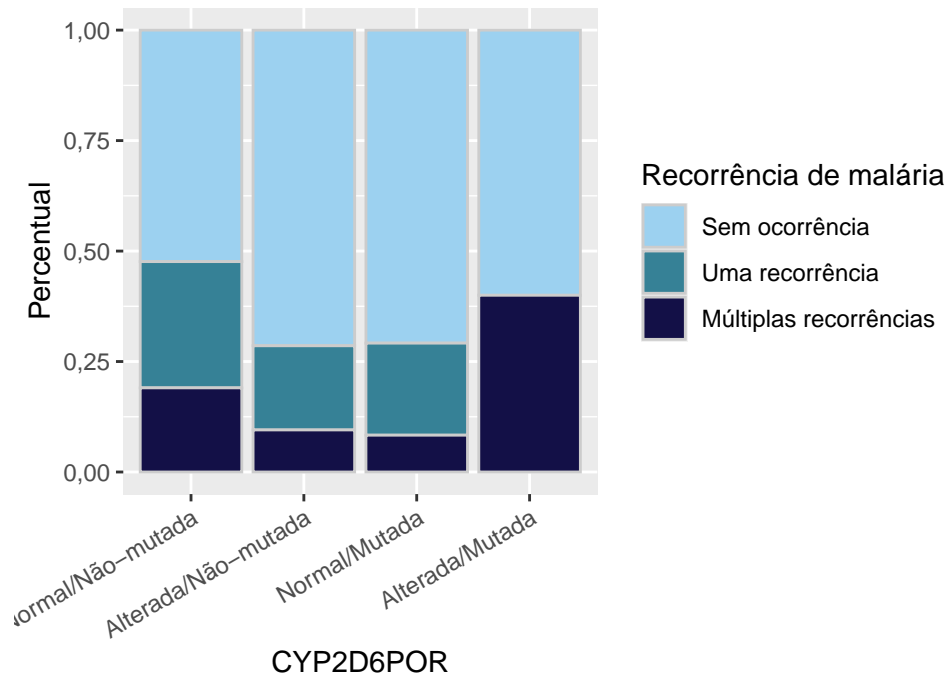
Table 14: Teste de independencia qui-quadrado com a variável Recorrência de Malária

| Variáveis | P-Valores |
|---|-----------|
| Status de metabolização conjunta (CYP2D6/POR) | 0,147 |
| Status de metabolização CYP2D6 | 0,352 |
| Status de mutação POR | 0,594 |

Analisando a Tabela 14, há evidências para não rejeitar a hipótese nula em todos os testes realizados. Desse modo, pode-se afirmar que há evidências de que as variáveis (Status de metabolização conjunta, Status de metabolização CYP2D6 e Status de mutação POR) são independentes da variável Recorrência clínica de malária. Assim, não foi constatado que indivíduos com CYP2D6 alterada possuem mais recorrências do que indivíduos com CYP2D6 normal. Também não é possível afirmar que indivíduos com POR mutada possuem mais recorrências do que POR nao-mutada.

Com o propósito de complementar a análise, três gráficos foram construídos (Figura 21, Figura 22 e Figura 23), visando visualizar as proporções relativas da variável Recorrência Clínica de malária com relação ao Status de metabolização conjunto e individual CYP2D6/POR.

Figura 21 – Gráfico de barras de Recorrência de malária com relação a CYP/POR conjunto



Recorrência clínica de malária, em relação ao Status de Metabolização CYP2D6:

A partir do gráfico de que relaciona as proporções relativas de recorrência de malária por status de metabolização CYP2D6.

Figura 22 – Gráfico de barras de Recorrência de malária com relação a CYP2D6

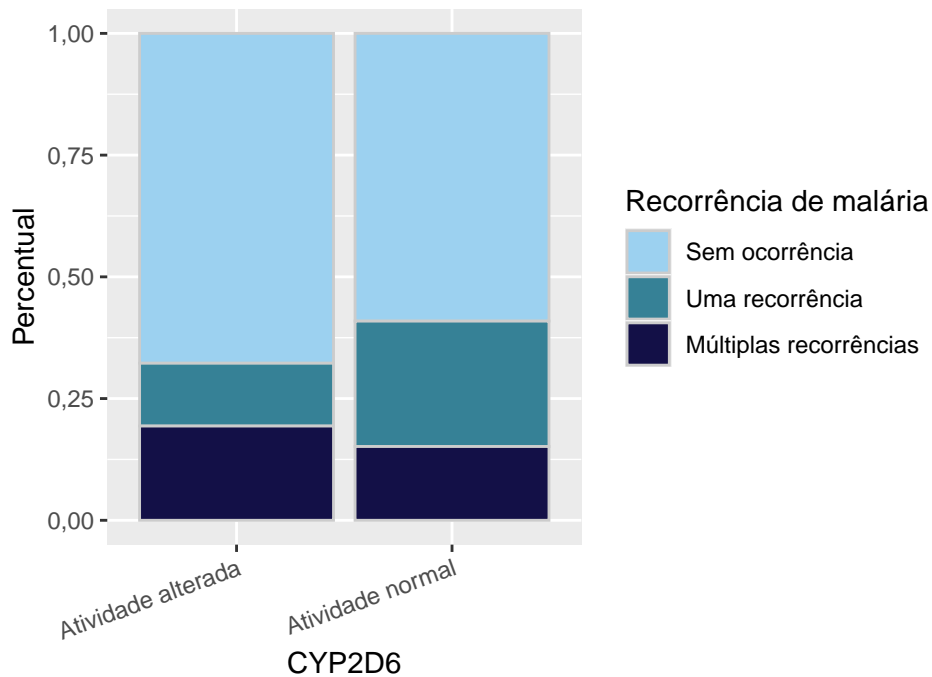
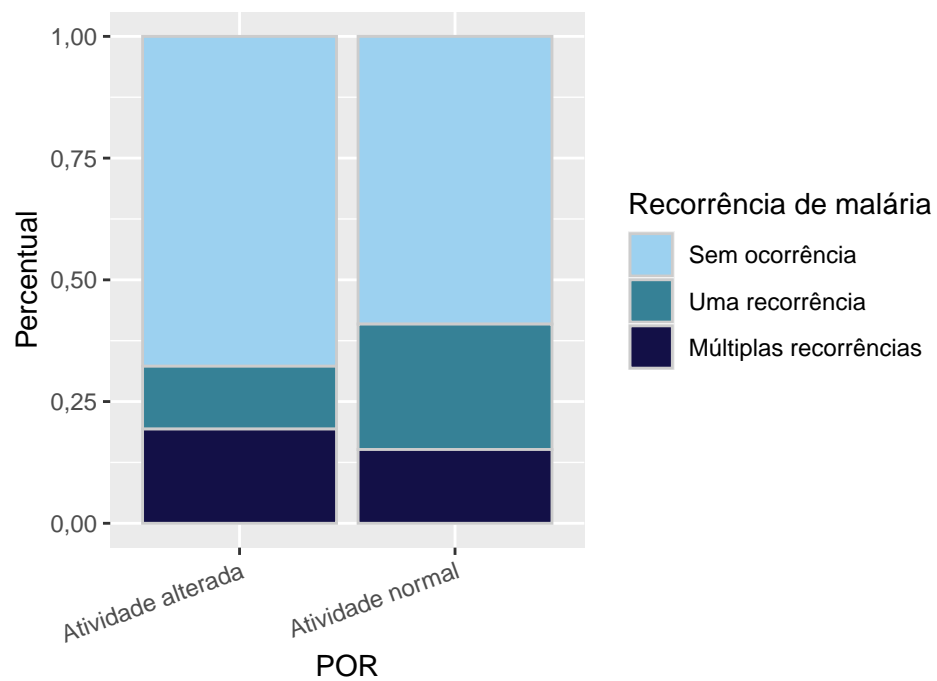


Figura 23 – Gráfico de barras de Recorrência de malária com relação a POR



3.8 Pergunta 7

“A eliminação dos gametócitos masculinos Pvs48 (PVS48D0 e PVS48D72) está associada com o status CYP2D6/POR (CYP2D6POR)?”- Texto elaborado pelas pesquisadoras.

Com o propósito de responder a pergunta realizamos testes de independência qui-quadrado, visando verificar se existe associação entre as variáveis: Status dos gametócitos masculinos (Pvs48) no momento inicial e Status de metabolização conjunta/individual CYP2D6 e gene POR; Status dos gametócitos masculinos (Pvs48) após 72 horas e Status de metabolização conjunta/individual CYP2D6 e gene POR. À vista disso, foi construído a Tabela 15 com os p-valores desses testes:

Table 15: Análise dos status dos gametócitos masculinos com relação à CYP2D6/POR

| Variáveis | P-valores |
|-----------------------------|------------------|
| CYP/POR conjunto e PVS48D0 | 0,177 |
| CYP/POR conjunto e PVS48D72 | 0,804 |
| CYP2D6 e PVS48D0 | 0,923 |
| CYP2D6 e PVS48D72 | 0,887 |
| POR e PVS48D0 | 0,216 |
| POR e PVS48D72 | 1 |

Note:

PVS48D0 = Gametócitos masculinos em 0 horas

PVS48D72 = Gametócitos masculinos em 72 horas

Analisando os p-valores da Tabela 15 é possível concluir que em todos os casos, não rejeita-se a hipótese nula do teste qui-quadrado. Desse modo, não é possível realizar afirmações do tipo “A eliminação dos gametócitos masculinos Pvs48 (PVS48D0 e PVS48D72) está associada com o status CYP2D6/POR (CYP2D6POR)”, tendo em vista que não há evidências de que as variáveis são dependentes/associadas.

4 Conclusão

Um ponto importante para ressaltar é que para todos os modelos de regressão múltipla foi sintetizado outros métodos como transformação de box-cox, e também modelos lineares generalizados em alguns casos com a distribuição Poisson, Gamma e Gamma inversa e os resultados obtidos traziam uma explicabilidade muito próxima do modelo de regressão linear múltiplo. De modo, como todos os ajustes de modelos não ficaram bem ajustados. Então decidiu-se manter os modelos de mais sucinto entendimento, sendo estes os modelos de regressão linear múltipla.

Quanto aos resultados obtidos, vale salientar que o obtivemos apenas um resultado significativo:

- Pergunta 2: Essa pergunta tinha o intuito de avaliar se a parasitemia inicial (PV18SD0NORM) está associada com a quantidade de gametócitos, e se essa variável é um fator de confusão nas análises de associação entre Status CYP/POR e Eliminação de gametócitos realizadas no item 1. Nesse sentido, a associação entre as variáveis Razão de Pvs25 (Gametócitos) e Quantidade de Parasitos no momento inicial (Parasitos), foi testada através do teste de correlação de Spearman, sendo obtido evidências de uma correlação não nula entre as variáveis. Posteriormente essa correlação foi mensurada: -0,433, o que indica uma correlação moderada e negativa entre as variáveis Razão de Pvs25 e Quantidade de Parasitos no momento inicial.

Outro resultado significativo obtido nesse item foi quanto aos modelos de regressão múltipla desenvolvidos com a variável resposta Razão de Pvs25, haja vista que os modelos explicados com as variáveis: Status de metabolização CYP2D6/POR conjunto e individual + variável Quantidade de parasitas no momento inicial, obtiveram maiores valores no coeficiente de determinação ajustado.

Nas demais perguntas não foram obtidos resultados significativos. Um fato que prejudicou as análises foi o tamanho amostral, pois ao se realizar as análises múltiplas rejeitamos em todos os casos o pressuposto de normalidade. Desse modo, tivemos que proceder as análises por meio de testes não paramétricos, que são menos poderosos do que os testes paramétricos. À vista disso muitos resultados que poderiam ser significativos, foram refutados

5 Referências

- Ministério da Saúde. Guia de tratamento da malária no Brasil. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/guia_tratamento_malaria_brasil.pdf. Acesso em 18 de maio de 2022.
- PIRES, Magda Carvalho et al. Estatística não paramétrica básica no software R: uma abordagem por resolução de problemas. Departamento de estatística. Universidade Federal de Minas Gerais, 2018.
- SALLES, Paula Ferreira. Variabilidade Do CYP2D6 em cinco municípios endêmicos de malária por Plasmodium vivax na Amazônia brasileira. Universidade Federal Fluminense, Niterói, 2020.
- SILVINO, Ana Carolina Rios. Recorrências na malária por Plasmodium vivax: variabilidade na enzima do complexo citocromo P450 2D6 (CYP2D6) e sua influência na falha terapêutica por primaquina. Belo Horizonte, 2019.
- Suporte ao Minitab. Compreendendo os métodos não paramétricos. Disponível em : <https://support.minitab.com/pt-br/minitab/19/help-and-how-to/statistics/nonparametrics/supporting-topics/understanding-nonparametric-methods/>. Acessado em 18 de maio de 2022.