

UNIVERSIDADE ABERTA



Análise da relação de supressão viral com medidas clínicas de atendimento aos pacientes adultos vivendo com o HIV/SIDA

Inácio Feliciano Retene

Mestrado em Bioestatística e Biometria

Janeiro de 2025

UNIVERSIDADE ABERTA



Análise da relação de supressão viral com medidas clínicas de atendimento aos pacientes adultos vivendo com o HIV/SIDA

Inácio Feliciano Retene

Mestrado em Bioestatística e Biometria

Dissertação orientada pela Professora Doutora Teresa Paula Costa Azinheira Oliveira e co-orientada pelo Doutor Ramos Bartolomeu José Mboane

Janeiro de 2025

Resumo

A prevalência do Vírus da Imunodeficiência Humana (HIV) em todo o mundo, destaca a importância de adoção de estratégias eficazes para a supressão viral, com foco nos pacientes com comorbidades para países em vias de desenvolvimento. Nessa perspectiva, o objetivo deste estudo, foi de analisar a relação entre a supressão viral com as medidas clínicas de atendimento aos pacientes adultos vivendo com o HIV no Centro de Saúde da Cidade de Lichinga, entre os anos 2020 e 2021. Foram analisadas as seguintes variáveis independentes: Tuberculose (TB), o Peso corporal, Sexo, a Idade, Tempo de tratamento e Apoio Psicossocial. Utilizando o teste de qui-quadrado e a regressão logística múltipla, foram analisados dados de 342 pacientes adultos.

Os principais resultados encontrados são os seguintes: a presença da TB nos pacientes com HIV reduz em 97% a probabilidade de alcançar a supressão viral, dada a Odds Ratio (OR=0.03); Um quilograma de aumento do Peso Corporal, a chance de supressão viral é de 83% (OR=1.83); A cada mês de Tratamento Antirretroviral (TARV) ininterrupto, a probabilidade é muito acima de 100% (OR=5.11); Outra descoberta é o impacto da interação entre a presença da TB e alto Apoio Psicossocial, sugerindo que a prestação de cuidados e Apoio Psicossocial aos pacientes com coinfeção de TB e HIV, aumenta mais rapidamente a chance de alcançar a supressão viral nesses pacientes do que aqueles que não tiveram Apoio Psicossocial dada a OR=48.50 e por fim, houve a descoberta da interação dos fatores Tempo de Tratamento e Peso Corporal.

Com estes resultados, pode se concluir que: aos pacientes em TARV, há necessidade de diagnóstico precoce de outras doenças, como é o caso da TB; por outro lado, o reforço do Apoio Psicossocial e o controlo do Peso é fundamental. Estas conclusões são validadas pelos resultados do Pseudo R² (Cox e Snell= 68.2% e Nagelkerke = 92.6%). O teste de Hosmer-Lemeshow confirma a validade desta teoria (p-valor de 99.56%) e por fim, a curva Receiver Operating Characteristic (ROC= 99.4%).

Palavras-chave: Supressão Viral, HIV/SIDA, Regressão Logística, Terapia Antirretroviral

Abstract

The Prevalence of Human Immunodeficiency Virus (HIV) worldwide highlights the importance of adopting effective strategies for viral suppression, focusing on patients with comorbidities in developing countries. In this context, the objective of this study was to analyze the relationship between viral suppression and clinical measures for the care of adult patients living with HIV at the Health facility at the Lichinga City, between years 2020 and 2021. The following independent variables were analyzed: Tuberculosis (TB), Body Weight, Gender, Age, Treatment Duration and Psychosocial Support. Using the chi-square test and multiple logistic regression, data from 342 adult patients were analyzed.

The main findings are as follows: the presence of TB in an HIV patients reduces the probability of achieving viral suppression by 97%, given the Odds Ratio (OR=0.03); One kilogram of Body Weight increase, the chance of viral suppression is 83% (OR=1.83); For each month of uninterrupted Antiretroviral Treatment (ART), the probability is over than 100% (OR=5.11); Another finding is the impact of the interaction between TB presence and high Psychosocial Support, suggesting that the provision of care and Psychosocial Support to patients with TB and HIV coinfection, increases the chance of achieving viral suppression more quickly than those who did not receive Psychosocial Support, given an OR=48.50. Finally, the interaction between Treatment Duration and body weight was discovered.

Based on these results, it can be concluded that: for patients on ART, there is a need for early diagnosis of other diseases, such as TB as well as a strengthening Psychosocial Support and monitoring the body weight, is essential. These conclusions are validated by the Pseudo R² results (Cox and Snell= 68.2% and Nagelkerke = 92.6%). The Hosmer-Lemeshow test confirms the validity of this theory (p-value of 99.56%), and finally, the Receiver Operating Characteristic curve (ROC=99.4%).

Keywords: Viral Load Suppression, HIV/AIDS, Logistic Regression, Antiretroviral Therapy

Dedicatória

Dedico esta dissertação a minha amada esposa Valdamira José Francisco Retene e os meus filhos Neyonildo, Neylanie, Jonathan, Josivalda e Daniel, que estiveram presentes em todos os momentos, oferecendo apoio incondicional, amor e compreensão. Sem vocês, esta jornada teria sido muito difícil.

Minha gratidão para vocês.

Agradecimentos

Gostaria de expressar a minha profunda gratidão a minha orientadora, Professora Doutora Teresa Paula Costa Azinheira Oliveira, por sua orientação incansável, pela partilha do seu vasto conhecimento e por sempre motivar a alcançar a excelência.

Ao antigo Diretor Provincial de Saúde e Coorientador, Doutor Ramos Bartolomeu José Mboane, por suas valiosas sugestões e críticas construtivas para a materialização desta pesquisa.

Sou grato a Direção Provincial de Saúde, por me proporcionar o campo de pesquisa e coleta de dados que foram usados para análise e conclusão da pesquisa.

Agradeço aos professores do Curso de Mestrado em Bioestatística e Biometria da Universidade Aberta de Lisboa, por terem contribuído para a minha formação académica, em especial o Professor Doutor Pedro Serranho, cujas aulas foram inspiradoras.

Um agradecimento especial aos meus pais Feliciano Retene e Carlota Ginela, pelo seu amor incondicional.

Agradeço também aos meus colegas do serviço que ajudaram no fornecimento de informações úteis e protocolos em uso na área de saúde.

Ao meu amigo Valdemiro Américo, pelos muitos neurónios gastos no incentivo de conclusão deste trabalho.

Índice

Conteúdos	Páginas
Resumo.....	iii
Abstract	iv
Dedicatória	vi
Agradecimentos.....	vii
Capítulo I	1
1. Introdução.....	1
1.1 Justificativa	3
1.2 Objetivos	4
1.2.1 Geral.....	4
1.2.2 Específicos.....	4
1.3 Resultados esperados	4
1.4 Estrutura da dissertação.....	4
Capítulo II.....	6
2. Fundamentação teórica.....	6
2.1 Perfil da epidemia em Moçambique	7
2.1.1 Situação geral de Moçambique.....	7
2.1.2 Situação da epidemia na Província de Niassa.....	9
2.2 Situação da Supressão viral.....	12
2.3 Determinantes da supressão viral.....	13
2.4 Modelos de prestação de cuidados aos pacientes vivendo com HIV	13
Capítulo III.....	14
3. Descrição dos dados	14
3.1 Variáveis de estudo	14
3.2 Critérios de inclusão.....	14
3.3 População e Amostra estudada.....	15
3.4 Metodologia para análise de dados	16
3.4.1 Regressão Logística	16
3.4.1.1 Regressão Logística Binária simples	17
3.4.1.2 Estimação dos coeficientes do modelo de regressão	18

3.4.2	Regressão Logística Múltipla	19
3.4.2.1	Estimação dos coeficientes do modelo de regressão múltipla.....	20
3.4.2.2	Critério de Informação de Akaike	21
3.4.2.3	Razão de chances (odds ratio)	22
3.4.2.4	Teste do Qui-Quadrado para ajustamentos	23
3.4.2.5	Principais pressupostos do teste do Qui-Quadrado.....	24
3.4.3	Métodos de avaliação do modelo logístico.....	25
3.4.3.1	Teste de Razão de Verossimilhança	25
3.4.3.2	Pseudo R² de Cox e Snell	26
3.4.3.3	Nagelkerke's R²	27
3.4.3.4	McFadden's R²	27
3.4.3.5	Tjur's R² (Coeficiente de Discriminação).....	27
3.4.3.6	sqPearson R² (Quadrado da Correlação de Pearson).....	28
3.4.3.7	Teste de Hosmer-Lemeshow	28
3.4.3.8	Curva Receiver Operating Characteristic (ROC)	29
3.5	Materiais e métodos	31
3.5.1	Acesso a base de dados	31
3.5.2	Questões éticas.....	32
3.5.3	Variáveis de estudo.....	32
3.5.4	Definição das variáveis de estudo.....	33
3.5.5	Análise Descritiva (Univariada vs Bivariada)	34
3.5.6	Aplicação do Modelo de Regressão Logística Múltipla.....	34
3.5.7	Definição da qualidade do ajuste do modelo	34
Capítulo IV	35
4.	Aplicação da análise de Regressão Logística Múltipla.....	35
4.1	Apresentação das características demográficas do grupo estudado.....	35
4.2	Análise e verificação dos pressupostos do modelo de Regressão Logística Múltipla.....	37
4.2.1	Análise dos fatores associados à supressão viral	37
4.2.2	Criação do Modelo zero.....	38
4.2.3	Seleção do melhor modelo logístico binário.....	39
4.2.4	Principais diferenças entre os modelos	40

4.2.5	Estimativas dos coeficientes do modelo ajustado.....	41
4.2.6	Definição da equação geral preditiva.....	47
4.2.6.1	Considerando todas as variáveis, independentemente do seu teste de significância na equação final	48
4.2.6.2	Considerando apenas as variáveis significativas na equação final	50
4.2.7	Análise do ajuste geral do modelo.....	52
4.2.8	Análise de ajustamento do modelo	53
4.3	Avaliação do desempenho do modelo.....	53
4.4	Teste do funcionamento da equação preditiva do modelo reduzido	54
Capítulo V.....		57
5.	Conclusões e Sugestões.....	57
5.1	Discussão.....	57
5.2	Conclusão.....	63
5.3	Sugestões.....	65
5.4	Limitações do estudo.....	65
Capítulo VI		67
6.	Referências Bibliográficas	67
WEB-Referências		73
7.	Apêndices	75

Lista de gráficos

Gráfico 2.1 - Cobertura de Unidades Sanitárias TARV ao longo do Tempo	9
Gráfico 2.2 - Tendência de PVHS Adultos.....	10
Gráfico 4.1 - Distribuição dos pacientes por sexo	30
Gráfico 4.2 - Distribuição dos pacientes por faixa-etária	36
Gráfico 4.3 - Análise do modelo pela Curva de ROC	54

Lista de tabelas

Tabela: 2.1 Principais Estimativas do Spectrum 6.06 - 2021	7
Tabela: 2.2 - Distribuição de ativos em TARV em pacientes adultos	11
Tabela 4.1 - Análise do teste do qui-quadrado dos fatores associados à supressão viral em pacientes adultos vivendo com o HIV/SIDA	32
Tabela 4.2: Critério de informação de Akaike para seleção do melhor modelo logístico binário	39
Tabela 4.3 - Estimativas dos coeficientes do Modelo Logístico Ajustado.....	34
Tabela 4.4 - Pseudo R2 para análise do ajuste geral do modelo.....	52
Tabela 4.5 - Teste Hosmer e Lemeshow para análise de ajustamento do modelo	53

Lista de diagramas

Diagrama: 3.1 – Estatísticas de pacientes inscritos ao TARV no período de estudo 15

Lista de figuras

Figura: 2.1 - Distribuição da prevalência do HIV em Moçambique, 2021 8

Lista de abreviaturas, acrónimos, siglas e símbolos matemáticos

AIC -----	(Critério de Informação de Akaike)
APSS -----	(Apoio Psicossocial)
AUC -----	(Area Under Curve)
BIC -----	(Bayesian Information Criterion)
CV -----	(Carga Viral)
HIV -----	(Vírus da Imunodeficiência Humana)
HIV+ -----	(Vírus da Imunodeficiência Humana Positivo)
INS -----	(Instituto Nacional de Saúde)
INSIDA -----	(Inquérito Nacional sobre o Impacto do HIV/AIDS)
MISAU -----	(Ministério da Saúde)
OR -----	(Odds Ratio - Razão de Chances)
OMS -----	(Organização Mundial da Saúde)
PVHIV -----	(Pessoa vivendo com o HIV)
ROC -----	(Receiver Operating Characteristic)
SIDA -----	(Síndrome de Imunodeficiência Adquirida)
TARV -----	(Tratamento Antirretroviral)
TB -----	(Tuberculose)
UNAIDS -----	(Programa Conjunto das Nações Unidas para o HIV)
WHO -----	(World Health Organization)

Capítulo I

1. Introdução

Em todo o mundo, estimava-se que em 2020 havia cerca de 37.6 milhões [30.2 milhões – 45.0 milhões] de pessoas estavam vivendo com o Vírus da Imunodeficiência Humana (HIV) apesar da tendência de redução de novas infecções em cerca de 40%, considerando o pico de 1997, (UNAIDS, 2021). Desta estimativa, cerca de 35.9 milhões [28.9 milhões – 43 milhões] era população adulta. No entanto, a maior parte de pessoas que vivem com o HIV é encontrada em países de baixa renda e cerca de dois terços do total (25.6 milhões) está na África Subariana. Relativamente as mortes, atingem cerca de 1 milhão de pessoas por ano em causas relacionadas ao HIV na África Subariana, (Teixeira, 2021, apud Frank et al., 2019).

Num relatório publicado em 2020 pela Coligação Global para a Prevenção do HIV, refere que o tratamento antirretroviral contribui cada vez mais para a diminuição do número de infetados por HIV, à medida que cada vez mais pessoas recebem e permanecem em tratamento para o HIV o tempo suficiente para suprimir a sua carga viral para níveis não detetáveis.

De acordo com Fisch, 2017, os primeiros impactos do benefício do Tratamento Antirretroviral (TARV), foram vistos nos países desenvolvidos dado que o HIV se tornou numa doença crónica e tratável já no fim da década de 1990. Todavia, nos países desenvolvidos, apesar de grandes esforços, há milhões de pessoas que ainda não têm acesso ao TARV.

No entanto, Moçambique é um País que em 2020 tinha uma população total estimada em cerca de 31.18 milhões de habitantes, segundo os dados o Instituto Nacional de Estatística (INE), 2017. Entretanto, sabe-se que o primeiro caso do Vírus da Imunodeficiência Humana Positivo (HIV+) foi diagnosticado em 1986 (Machel, 2010). Do ponto de vista epidemiológico, apresenta uma situação generalizada de HIV, sendo o país africano de expressão portuguesa e um dos países do mundo em que o número de infetados pelo HIV é maior, Folha, 2021. Face a isso, este país enfrenta uma alta prevalência do HIV, com cerca de 2.2 milhões de pessoas vivendo com esta doença, de acordo com dados de 2020 do Programa Conjunto das Nações Unidas para o HIV (UNAIDS).

Embora tenham sido feitos esforços para melhorar o acesso aos cuidados de saúde para as Pessoas Vivendo com o HIV (PVHIV), ainda há uma lacuna significativa na cobertura do tratamento assim como na supressão viral entre os pacientes. Em 2020, apenas cerca de 64% dos pacientes com HIV em tratamento antirretroviral em Moçambique alcançaram a supressão viral, de acordo com a UNAIDS.

De várias análises feitas, há um consenso científico de que as pessoas HIV positivas que aderem ao tratamento anti- retroviral, apresentam larga probabilidade de alcançar a supressão da carga viral a níveis indetetáveis e não transmitem o HIV sexualmente para outras pessoas pois, o tratamento é um fator que permite à pessoa que vive com HIV, recuperar e manter uma boa saúde. Contudo, é preciso que globalmente se tenha acesso regular aos exames de carga viral, combinando com sistemas de laboratório efetivos, serviços de saúde adequados e recursos humanos qualificados.

Os dados do Inquérito Nacional sobre o Impacto do HIV em Moçambique (INSIDA 2021) revelaram que a prevalência do HIV+ entre adultos em Moçambique foi de 12.5%, o que corresponde a aproximadamente 2,097,000 adultos vivendo com HIV, sendo que é maior entre mulheres (15.0%) do que nos homens (9.5%). Por outro lado, a prevalência de supressão da carga viral entre adultos vivendo com HIV em Moçambique foi de 64.1%, sendo 67.1% entre mulheres e 58.8% entre homens.

No entanto, o Programa Conjunto das Nações Unidas para o HIV estabeleceu metas de 95-95-95%, ou seja, 95% de todas as pessoas vivendo com HIV conheçam o seu estado de HIV; 95% de todas as pessoas diagnosticadas com HIV+ estejam a receber o Tratamento Antirretroviral (TARV) de forma contínua e que 95% das pessoas a receber o TARV alcancem a supressão da carga viral.

Em relação a província de Niassa, os dados epidemiológicos indicaram uma tendência crescente do número de pessoas que estavam em tratamento, pois em menos de 10 anos, o número de adultos que recebia o tratamento antirretroviral passou de 12,282 em 2014 para 41,985 em 2021, ou seja, a epidemia foi mais do que o triplo, (relatório anual da Direção Provincial de Saúde, 2021), dando sinais de que o HIV+ continua ainda sendo um importante problema de saúde pública mundial.

Este problema de saúde pública gera reflexos negativos no estado geral de saúde de pessoas que vivem com o HIV e que não aderem ao tratamento com a terapia antirretroviral de forma regular, (Enferm, 2023).

Na sequência das metas globalmente estabelecidas, esta dissertação traz evidências científicas de análise da correlação entre a supressão da carga viral e a assistência clínica oferecida as PVHIV na Província do Niassa, especificamente no Centro de Saúde da Cidade de Lichinga entre os anos 2020 e 2021.

1.1 Justificativa

A supressão viral é um indicador importante de saúde e bem-estar em pacientes vivendo com HIV. Alcançar a supressão viral pode melhorar significativamente a expectativa de vida, reduzir o risco de transmissão da doença e melhorar a qualidade de vida dos pacientes. No entanto, alcançar e manter a supressão viral é um desafio praticamente inatingível em muitos Países, incluindo Moçambique, UNAIDS, 2020.

O seguimento clínico, adesão ao tratamento e manutenção da terapia, são indicadores importantes da qualidade de atendimento uma vez que garante o controlo da epidemia rumo ao alcance da eliminação até 2030. Esta e outras razões, ditaram a escolha deste tema para pesquisa ao nível do Centro de Saúde da Cidade de Lichinga.

Do ponto de vista de saúde pública, ao estudar a associação entre a supressão viral e as variáveis independentes como a idade, o sexo, peso corporal, tempo de tratamento, presença da Tuberculose (TB) e Apoio Psicossocial (APSS), irá ajudar a orientar várias intervenções em saúde pública, como programas integrados de TB e HIV, políticas para iniciar o tratamento antirretroviral precocemente e intervenções nutricionais. Entender essas relações vai ajudar a otimizar os cuidados e a melhorar os resultados para as pessoas vivendo com o HIV.

1.2 Objetivos

Nesta dissertação foram considerados os seguintes objetivos:

1.2.1 Geral

- Analisar a relação entre a supressão viral com as medidas clínicas de atendimento aos pacientes adultos vivendo com o HIV no Centro de Saúde da Cidade de Lichinga, entre os anos 2020 e 2021.

1.2.2 Específicos

- Fazer a descrição do estado geral da epidemia em Moçambique e em Niassa particularmente;
- Identificar os fatores determinantes para a supressão da carga viral;
- Elaborar um modelo de apoio à tomada de decisão que possa fornecer informações valiosas sobre os fatores que influenciam a supressão viral em pacientes com HIV;
- Contribuir para a identificação de estratégias de melhoria tanto no que diz respeito ao acesso aos cuidados de saúde bem como de procurar melhorar ou otimizar os resultados de diagnóstico e prevenção do HIV.

1.3 Resultados esperados

Espera-se que este estudo forneça informações importantes sobre a relação entre a supressão viral e as medidas clínicas de atendimento em pacientes adultos vivendo com HIV em Moçambique, especificamente a província de Niassa. A análise de regressão logística múltipla permitiu identificar quais as variáveis com uma associação significativa com a supressão viral e quantificar o seu impacto.

1.4 Estrutura da dissertação

Para melhor apresentação dos resultados desta dissertação em consonância com a metodologia traçada, este trabalho obedece a seguinte estrutura:

- **Capítulo I:** é a parte introdutória na qual é descrito de forma sumária, o ponto de situação da epidemia em Moçambique assim como estimativas da taxa de supressão viral. Igualmente estão contemplados os objetivos, a justificativa e os resultados esperados desta dissertação.
- **Capítulo II:** são abordados os principais elementos teóricos que sustentam esta pesquisa, revisão literária e demais conceitos que asseguram a compreensão dos temas abordados.
- **Capítulo III:** este é o capítulo dedicado a descrição dos dados aplicados para o estudo, análise das variáveis e desenho metodológico para a análise de dados.
- **Capítulo IV:** neste capítulo a atenção é dedicada a análise de dados, fazendo o cruzamento e correlação das variáveis de estudo, os modelos explicativos desenhados de supressão da carga viral com base na regressão logística múltipla aplicada no pacote RStudio.
- **Capítulo V:** sendo o último capítulo, as conclusões e sugestões são apresentadas aqui com base nos resultados que o autor encontrou durante a análise das variáveis que impactam sobre a supressão viral nos pacientes vivendo com o HIV.

Capítulo II

2. Fundamentação teórica

A epidemia do Vírus da Imunodeficiência Humana assim como da Síndrome de Imunodeficiência Adquirida (HIV/SIDA) em Moçambique, constitui um desafio que requiere uma resposta combinada e engajamento multisectorial. Nos últimos anos, o Ministério da Saúde (MISAU) de Moçambique tem implementado várias ações com objetivo de alargar as medidas de prevenção, diagnóstico e início atempado dos cuidados e tratamento para esta doença. Face a isso, a expansão dos serviços de saúde para atendimento das pessoas vivendo com o HIV enfrenta algumas barreiras relacionadas com os recursos humanos qualificados, equipamentos laboratoriais, infraestruturas e disponibilidade de medicamentos.

Assim sendo, o tratamento antirretroviral foi introduzido no país de forma gradual em 2003, com linhas terapêuticas limitadas a um regime de primeira linha e esquemas alternativos de segunda linha para casos de falência terapêutica à primeira linha de tratamento.

Nessa abordagem de massificação do tratamento antirretroviral, foi necessário padronizar uma metodologia de seguimento dos pacientes em TARV e por muitos anos Moçambique aplicava os critérios clínicos e imunológicos porém, esta metodologia tinha baixa sensibilidade para identificar indivíduos com falência virológica. Por essa razão, em 2013 a Organização Mundial da Saúde (OMS) recomendou ao país o uso do teste de carga viral para prover a indicação mais precisa da condição de saúde dum paciente em TARV. Foi neste âmbito que se introduziu a primeira fase de uso de Carga Viral (CV) dedicada apenas para mulheres grávidas, lactantes e crianças entre os 2 e 5 anos. Anos mais tarde, foi sendo implementada a colheita de carga viral a todos grupos populacionais na condição de HIV positivas com base em protocolos e algoritmos clínicos, laboratoriais e programáticos.

Nessa perspetiva de prover meios e recursos para a avaliação precisa da condição de saúde da PVHIV, a província de Niassa seguiu as recomendações do MISAU e foi aumentando a capacidade de colheita de amostras para a realização de exames de carga viral pois, os profissionais de saúde de Niassa sabem que a terapia antirretroviral é transformadora para as pessoas nesta condição de doença porque garante a elas uma recuperação da qualidade de vida e transmite algum futuro cheio de esperança. Sabe-se ainda que para além da melhoria do

paciente, há evidências científicas de que as pessoas com carga viral indetetável no sangue não transmitem sexualmente o vírus para outras pessoas.

Análises recentes de maiores estudos com casais sero-discordantes não identificaram um único caso de transmissão de uma pessoa com carga viral indetetável para outra, (UNAIDS, 2017)

2.1 Perfil da epidemia em Moçambique

2.1.1 Situação geral de Moçambique

Moçambique adotou o Spectrum¹ como modelo epidemiológico para estimar os indicadores-chave da epidemia do HIV. Nisso, a projeção feita até Março de 2021 indicava uma redução de mortalidade relacionada ao HIV, redução de novas infeções e da transmissão vertical da mãe para o filho ao longo do tempo, como resultado de diversas intervenções que o serviço nacional de saúde vem realizando para o controlo da epidemia. Entretanto, dos cerca de 2,101,222 PVHIV, 94.05% eram adultos e, no geral, o número de novas infeções por dia era de cerca de 258, como se pode ver na tabela abaixo:

Tabela: 2.1 Principais Estimativas do Spectrum 6.06 - 2021

Moçambique	2021
PVHIV	2,101,222
Adultos vivendo HIV	1,976,250
Dos Adultos, nº de Mulheres VHIV	1,253,516
Crianças vivendo HIV	124,972
Nº de novas infeções	94,053
Nº de novas infeções por dia	258
Nº de mortes relacionadas ao HIV/SIDA	35,463
Nº de mulheres grávidas HIV+	93,925
Taxa de transmissão vertical	12

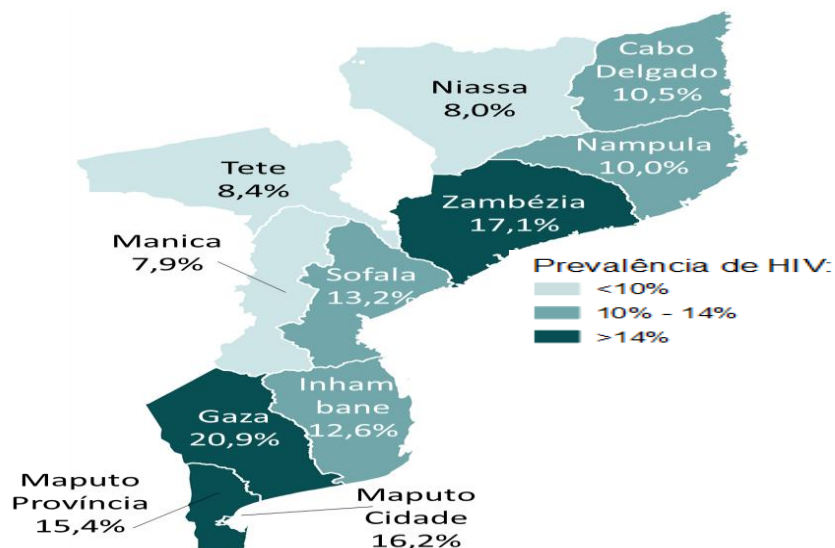
Fonte: Relatório anual do programa nacional de HIV Moçambique, 2021

¹ O Spectrum é um software usado para modelar cenários, estimar o impacto de intervenções e projetar a carga da doença. Ele permite que gestores de saúde pública e tomadores de decisão avaliem a eficácia de políticas e estratégias para responder ao HIV/SIDA.

As estatísticas moçambicanas indicam que de 2,101,222 pessoas vivendo com o HIV, um total de 1,599,317 eram pessoas inscritas nos serviços de cuidados e tratamento do HIV e que estavam ativas até finais de Dezembro de 2021, correspondente a 76.1%. Significa com isso que há uma margem superior a 20% de pessoas que vivem com o HIV mas não estão fazendo o tratamento e isso pode retardar o cumprimento dos Objetivos de Desenvolvimento Sustentável (ODS) até 2030.

Os estudos sobre a prevalência da epidemia em Moçambique, conforme o INSIDA 2021, apontam quatro províncias em situação mais preocupante nomeadamente: Gaza, Zambézia, Maputo Cidade e Maputo Província com 20.9; 17.1; 16.2 e 15.4% respetivamente, pois, somente estas províncias detêm mais de 60% dos pacientes que vivem com a epidemia há mais tempo. O mapa abaixo descreve a situação geral da prevalência da epidemia em Moçambique:

Figura: 2.1 - Distribuição da prevalência do HIV em Moçambique, 2021



Fonte: Instituto Nacional de Saúde (INS), 2021

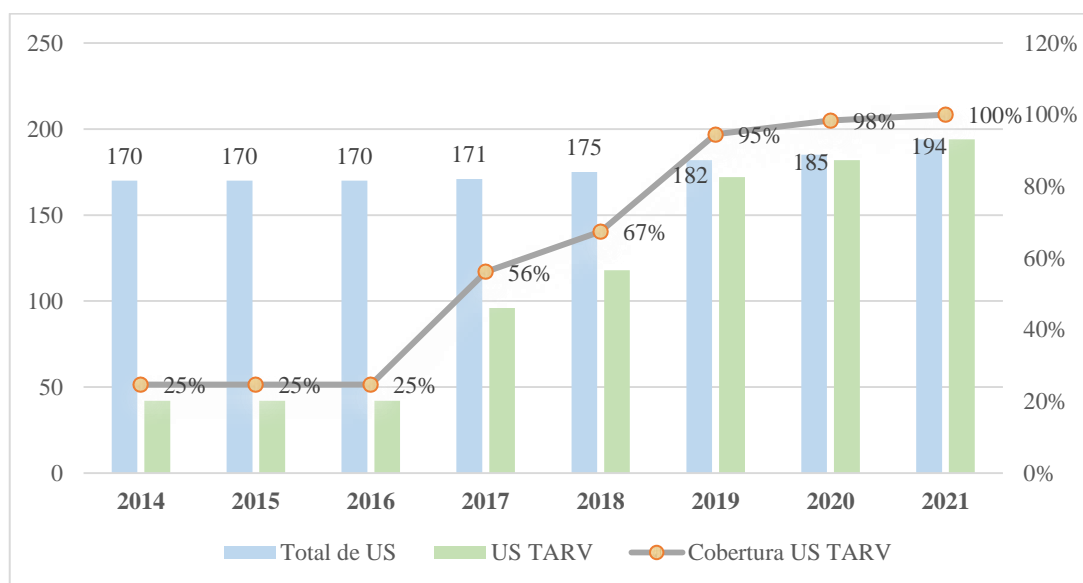
Neste estado da epidemia em Moçambique, estão numa situação relativamente melhor as províncias de Manica, Niassa e Tete pois, apresentam cifras inferiores a 10%.

2.1.2 Situação da epidemia na Província de Niassa

Niassa é uma província que se localiza na região Norte de Moçambique e é constituída por 16 distritos. A situação do HIV na província constitui um grande desafio com destaque para a inclusão dos pacientes aos cuidados e tratamento bem como a oferta de pacotes de melhoria de atendimento e estratégias de retenção das PVHIV.

De acordo o relatório anual da Direção Provincial de Saúde, 2021, a disponibilidade dos serviços de TARV conheceu uma aceleração acentuada a partir de 2017 com a estratégia de oferta massiva em quase todas as unidades sanitárias com intuito de garantir o acesso ao TARV a mais PVHIV, sobretudo em pacientes que percorriam longas distâncias para o acesso ao tratamento antirretroviral somente nas sedes distritais. Houve uma evolução acentuada ao passar de 56% para 100% de unidades sanitárias da província que ofereciam o tratamento até Dezembro de 2021, conforme o gráfico abaixo:

Gráfico: 2.1 - Cobertura de Unidades Sanitárias de TARV ao longo do Tempo



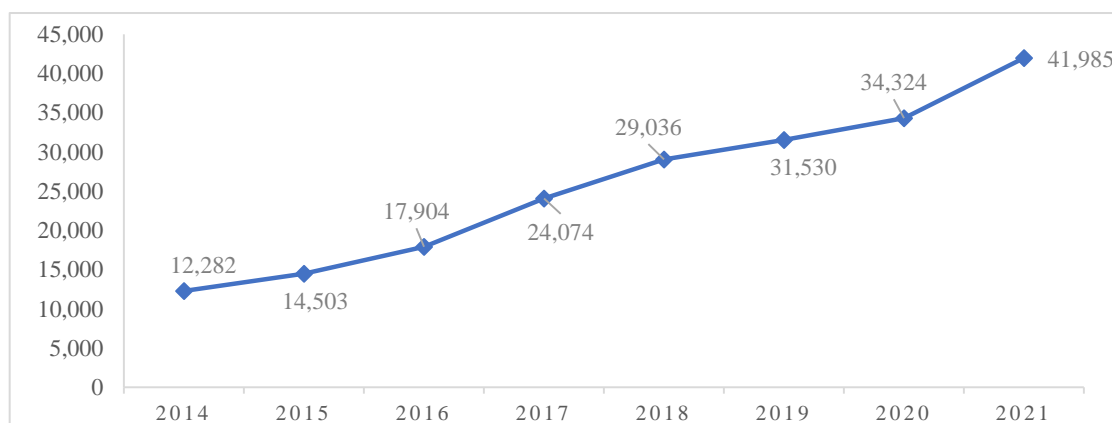
Fonte: Relatório Anual da Direção Provincial de Saúde, 2021

Com esta expansão massiva dos serviços de cuidados e tratamento do HIV para todas as unidades sanitárias da província, a demanda por estes serviços também aumentou

significativamente, sendo que em 2021, a província tinha um total de 44,981 pacientes ativos em TARV e destes, 41,985 eram pacientes adultos, correspondente a 93.3%.

De acordo com relatório anual do programa nacional do HIV em Moçambique, 2021, Niassa é das províncias com prevalência do HIV mais baixa do país assim como apresenta o menor número de pacientes em TARV, em relação ao resto das províncias do país. Contudo, o aumento do número de pacientes adultos em terapia antirretroviral foi crescendo à medida que a província avançava na expansão dos respetivos serviços, conforme nota se no gráfico 2.1.

Gráfico: 2.2 - Tendência de PVHS Adultos



Fonte: Relatórios Anuais da Direção Provincial de Saúde, 2014-2021

Esta tendência ascendente de pessoas adultas vivendo com o HIV/SIDA em Niassa, de acordo com dados do gráfico 2.2, é resultado de diversos esforços do Serviço Nacional de Saúde em colocar o tratamento cada vez mais perto do utente, daí que até 2021, a província possuía um total de 41,985 pacientes adultos ativos em tratamento.

Relativamente a distribuição geográfica da epidemia na província, ela obedece o cenário de distribuição populacional, ou seja, em distritos mais populosos, igualmente se regista o maior número de pacientes em cuidados e tratamento do HIV.

A tabela 2.2 abaixo, descreve o cenário dos dois anos de 2020 e de 2021 em pacientes adultos, considerando todos os distritos:

Tabela: 2.2 - Distribuição de ativos em TARV em pacientes adultos

Distritos	População geral	Adultos ativos em TARV			
	2021	2020	2021	Tendência dos 2 anos	Peso relativo do ano de 2021
Chimbonila	86,137	955	1,336	40%	3%
Cuamba	326,575	7,876	9,888	26%	24%
Lichinga	301,261	9,576	11,406	19%	27%
Lago	128,204	2,192	1,575	-28%	4%
Majune	47526	551	817	48%	2%
Mandimba	254,869	3,560	4,508	27%	11%
Marrupa	93655	902	1,209	34%	3%
Mavago	35,562	242	318	31%	1%
Maúa	81095	678	929	37%	2%
Mecanhelas	321,180	3,730	4,859	30%	12%
Mecula	24815	212	288	36%	1%
Metarica	53,821	689	821	19%	2%
Muembe	52020	485	595	23%	1%
Ngaúma	113,843	1,210	1,513	25%	4%
Nipepe	53564	631	829	31%	2%
Sanga	90,516	835	1,094	31%	3%
Provincia	2,064,643	34,324	41,985	22%	100%

Fonte: Relatório Anual da Direção Provincial de Saúde, 2021

Conforme a tabela 2.2 acima, a distribuição geográfica da epidemia em Niassa é quase que homogênea uma vez que em distritos mais populosos o número de pacientes adultos ativos em TARV é também elevado. Os distritos de Lichinga, Cuamba, Mecanhelas e Mandimba cobrem mais de 70% de pacientes de toda a província por apresentarem cerca de 27%; 24%; 12% e 11%, respetivamente.

Nesta descrição, importa realçar o distrito de Lichinga, sendo este que detém maior número de PVHIV em relação aos demais e, é neste distrito que será a unidade de estudo nesta dissertação. Assim sendo, o distrito de Lichinga, será o foco para análise da relação de supressão viral com medidas clínicas de atendimento aos pacientes adultos vivendo com o HIV/SIDA por se considerar este com maior agregado de pacientes ao nível da província, especificamente a seguinte unidade sanitária: **Centro de Saúde da Cidade de Lichinga.**

2.2 Situação da Supressão viral

A toma dos medicamentos antirretrovirais tem um impacto na redução da morbidade e mortalidade em PVHIV, melhorando desta feita a qualidade de vida. Por outro lado, o seguimento de adesão ao TARV é um dos indicadores relevantes sobre a qualidade de atenção oferecida aos pacientes e permite monitorar o controlo da infeção no indivíduo, De Melo et. al, 2021.

Entende-se por **Supressão Viral** a um resultado dum Carga Viral (CV) dum indivíduo abaixo de 1000 cópias/ml.

Por outro lado, **Indetetável** é definida como uma carga viral abaixo do nível de deteção que varia de (<20 a 75 cópias/ml) dependendo do aparelho usado e o tipo de amostra se é sangue ou plasma. Em Moçambique, uma grande parte de aparelhos utilizados com amostras de sangue o limiar de deteção é de 400 a 550 cópias/ml e para o plasma o limiar é de 50 cópias/ml.

A realização de testes regulares de CV permite monitorar até que ponto as intervenções feitas num determinado paciente impactam sobre ele. Contudo, há que dar muita atenção nas seguintes lacunas identificadas pela UNAIDS: **(i)** lacuna de testagem – pessoas não testadas e elas não conhecem o seu seroestado; **(ii)** lacuna de tratamento – pessoas diagnosticadas com o Vírus de Imunodeficiência Humana Positivas (HIV+) e por várias razões não conseguem iniciar o tratamento; **(iii)** lacuna de supressão viral – pessoas que vivem com o HIV e iniciaram o tratamento todavia, não têm a CV suprimida.

Desta feita, os instrumentos adotados pelo MISAU conseguem, de forma agregada, monitorar a cascata de CV desde a realização, recepção do resultado e supressão viral. O cenário de Niassa mostra que das PVHS em TARV até Dezembro de 2021, apenas 48% tinham pelo menos um teste de carga viral, sendo uma taxa muito baixa para o cumprimento da norma importante para a monitoria de resposta terapêutica das PVHS.

Os dados indicam que dos 48% que foram colhidos CV, cerca de 75% desses pacientes é que receberam seus resultados sendo que destes, 84% é que tinham supressão viral. Significa com isso que a província deve acelerar as suas intervenções com vista a alcançar as metas globais de 95% rumo a eliminação da epidemia até 2030.

2.3 Determinantes da supressão viral

Os determinantes da supressão viral são elementos fundamentais a ter em conta para a contínua melhoria do estado de saúde da PVHS. Neste caso, muitos estudos convergem sobre os mesmos elementos atinentes a supressão viral, sendo os seguintes: *boa adesão, uso correto dos medicamentos prescritos, retenção no programa de cuidados e tratamentos, adoção de estilos de vida saudáveis e monitoria regular da carga viral.*

2.4 Modelos de prestação de cuidados aos pacientes vivendo com HIV

Os modelos diferenciados de serviços são abordagens centradas no utente com vista a simplificar e adaptar os serviços de HIV em todo o ciclo de tratamento. Estes modelos visam melhorar a adesão e retenção dos pacientes com objetivo de reduzir a frequência nas unidades sanitárias e reduzir a pressão sobre os clínicos de modo a melhorar a qualidade da prestação de serviços clínicos aos utentes instáveis.

A tendência de implementação desses modelos foi sendo crescente ao longo do tempo com maior incidência para a Dispensa Trimestral de medicamentos, Grupos de Apoio a Adesão Comunitária e a Dispensa Semestral, impulsionada pelo ajuste dos critérios de elegibilidade para ingresso nestes modelos.

Dentre vários modelos adotados pelo País, Niassa implementa quase todos nas unidades sanitárias que oferecem cuidados e tratamento do HIV sendo, no entanto, a dispensa trimestral o mais incidente com vista a melhorar a prestação de serviços.

Capítulo III

3. Descrição dos dados

Os dados usados para esta dissertação foram de pacientes adultos vivendo com HIV/SIDA e que tinham iniciado o TARV entre os anos 2020 e 2021. O acesso aos dados e variáveis analisadas foi através duma Base de Dados eletrónica usada para gestão de todos os eventos de pacientes em seguimento naquela unidade sanitária, designada de “*OpenMRS*²”.

3.1 Variáveis de estudo

Com vista a compreender o comportamento do fenómeno em estudo, foram consideradas as seguintes variáveis independentes: **idade, sexo, tempo de tratamento, presença ou não da tuberculose, o peso corporal, Apoio Psicossocial e quantidade de Carga Viral**, sendo esta última, a classificada como dependente. As variáveis independentes ajudaram a analisar até que ponto podem influenciar a condição de vida das PVHS, especificamente o alcance da supressão viral dos pacientes em seguimento no Centro de Saúde da Cidade de Lichinga.

3.2 Critérios de inclusão

De acordo com os objetivos do pesquisador, foram incluídos todos os pacientes que iniciaram o TARV no **Centro de Saúde da Cidade de Lichinga** entre Janeiro de 2020 até Dezembro de 2021. Outros critérios de inclusão foram os seguintes:

- Idade superior a 14 anos;
- Ter pelo menos um resultado de carga viral documentado e;
- Paciente no ativo durante o período de estudo.

Todos outros pacientes que não tinham estes critérios naquela unidade de estudo, foram excluídos desta pesquisa. A razão da escolha desta unidade sanitária foi baseada no facto de ser a que detém maior número de pacientes em TARV ao nível da província de Niassa. Importa realçar que esta unidade sanitária está localizada no maior centro urbano da província e é das

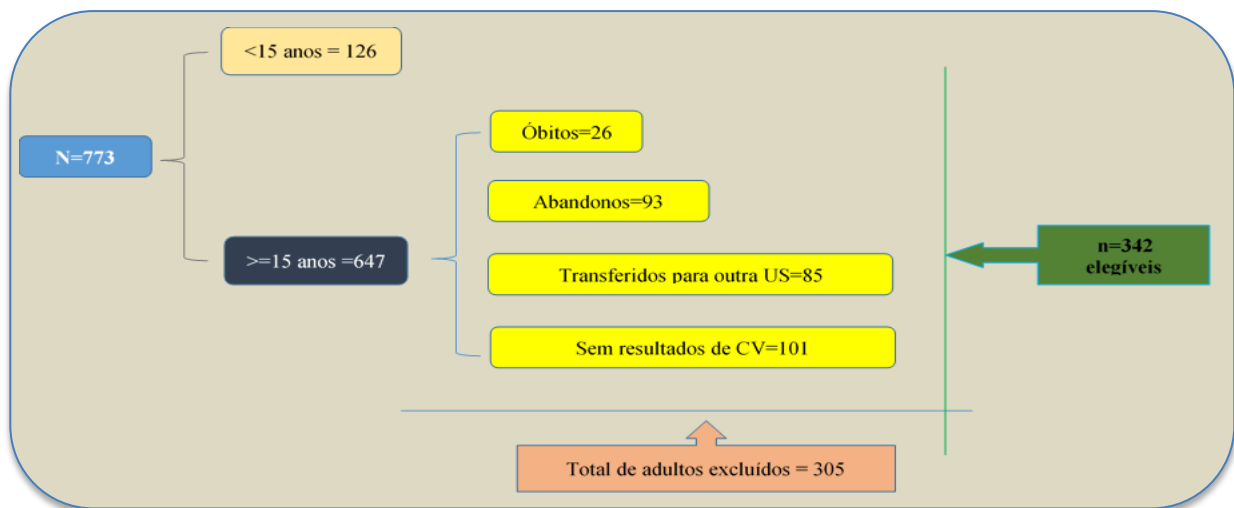
² É um sistema de gestão de registo de pacientes baseado na web. Ele oferece módulos como login baseado em localização, fluxos de trabalho de pacientes incorporados e relatórios flexíveis.

que possui algumas condições infraestruturais, humanas assim como equipamentos laboratoriais para a melhoria da prestação e oferta de cuidados de saúde aos pacientes.

3.3 População e Amostra estudada

Foram parte desta pesquisa todas as pessoas adultas na condição de HIV+ e que iniciaram o TARV no Centro de Saúde da Cidade de Lichinga entre Janeiro de 2020 a Dezembro de 2021. A seleção da amostra foi não probabilística e por conveniência, uma vez que se considerou todos os pacientes elegíveis para esta pesquisa, conforme os critérios de inclusão definidos. O diagrama abaixo retrata as estatísticas de pacientes que iniciaram o TARV no período de estudo:

Diagrama: 3.1 – Estatísticas de pacientes inscritos ao TARV no período de estudo



Fonte: Autor

Como se pode ver no diagrama acima, no período deste estudo, iniciaram o TARV naquela unidade sanitária uma população total de 773 pacientes dos quais, 126 eram pacientes com idade inferior a 15 anos, ou seja, eram crianças. Sendo que o grupo de estudo foi delimitado para adultos, as crianças foram excluídas e ficando um total de 647 pacientes. Ainda sobre os adultos, no período de estudo houve registo de perdas de seguimento por razões de morte, abandono e transferência, estes também foram excluídos da pesquisa. Por fim, quantificou-se o último grupo de 101 pacientes que não tinham nenhum resultado de carga viral documentado no sistema.

Feito isto, duma população total de 773 pacientes, foram excluídas 116 crianças assim como excluídos 305 adultos pelas razões apresentadas acima, e ficando com 342 pacientes elegíveis para análise.

3.4 Metodologia para análise de dados

A metodologia estatística aplicada é de acordo com a descrição teórica no início da dissertação.

3.4.1 Regressão Logística

A regressão logística é considerada uma técnica que tem como objetivo modelar o comportamento entre uma variável dependente binária e variáveis independentes métricas ou não métricas, isto é, destina-se a investigar o efeito das variáveis pelas quais os indivíduos, objetos ou sujeitos estão expostos sobre a probabilidade de ocorrência de determinado evento de interesse (Corrar et al. 2007 & Gonzale, 2018).

Vantagens do modelo logístico

Para o Mesquita (2014), o modelo de regressão logística apresenta as seguintes vantagens:

- ✓ Facilidade no uso de variáveis dependentes;
- ✓ Ideal para problemas que requerem estimação de probabilidades;
- ✓ Uso de variáveis métricas e não métricas simultaneamente;
- ✓ Os parâmetros do modelo fornecem um bom conhecimento sobre o fenómeno em estudo;
- ✓ Geralmente fornece um bom ajuste comparativamente aos modelos lineares;
- ✓ Facilidade na interpretação dos resultados;

Segundo (Belfiore, 2015 & Sousa, 2006), a variável dependente Y na regressão logística é geralmente binária, logo, nestes casos ela segue uma distribuição Bernoulli, tendo uma probabilidade desconhecida p . O princípio de Bernoulli determina se os possíveis resultados apresentam ou não certas características, ou seja, no caso de a variável apresentar a característica, ela apresenta valor igual a um, que representa sucesso e no caso contrário, valor igual a zero que representa fracasso. Lembrando que a distribuição de Bernoulli é apenas um

caso especial da distribuição binomial, onde $n = 1$ (que representa a realização de apenas único experimento)

$$Y = \begin{cases} 1, & \text{se ocorrer sucesso} \\ 0, & \text{se ocorrer fracasso} \end{cases}$$

A probabilidade de sucesso é $0 \leq p \leq 1$ e a probabilidade de fracasso é $q = 1 - p$. Portanto a variável aleatória Y tem distribuição Bernoulli, isto é, $Y \sim \text{Bernoulli}(p)$ e sua função de probabilidade é dada por:

$$P(Y = y) = p^y (1 - p)^{1-y}, \text{ com } y = 0, 1 \quad (\text{I})$$

3.4.1.1 Regressão Logística Binária simples

Segundo Cabral (2013), na regressão logística, a variável dependente/explicada é dicotômica, que assume valor 1 ao acontecimento de interesse (sucesso) e 0 ao acontecimento contrário (insucesso). Nas regressões, a quantidade chave é o valor médio da variável resposta dado o valor da variável independente. Esta quantidade é chamada valor médio condicional e é expressa como $E[Y/X]$ onde Y representa a variável dependente e X a variável independente/explicativa. No modelo de regressão linear, admitimos que o valor médio condicional, pode ser expresso como uma equação linear de x :

$$E \left[\frac{Y}{X} = x \right] = \beta_0 + \beta_1 x \quad (\text{II})$$

Portanto, a média pode assumir qualquer valor, quando x varia de $-\infty$ à $+\infty$.

No modelo de regressão logística, no caso onde a variável dependente assume apenas dois valores distintos, assume-se que:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (\text{III})$$

3.4.1.2 Estimação dos coeficientes do modelo de regressão

De acordo com Mesquita (2014), para ajustar um modelo de regressão linear, utiliza-se o método de mínimos quadrados para calcular os parâmetros do modelo, porém este método não é eficaz quando aplicado na regressão logística. No entanto, para regressão logística, utiliza-se o método de máxima verossimilhança.

Segundo Meyer (1980), o método de máxima verossimilhança de β , baseada em uma amostra aleatória x_1, x_2, \dots, x_n é aquele valor de β que torna máxima $L(x_1, x_2, \dots, x_n; \beta)$, considerada como uma função de β para uma dada amostra x_1, x_2, \dots, x_n .

A função de verossimilhança é definida por:

$$L(\beta) = \prod_{i=1}^n [\pi_i^{Y_i} (1 - \pi_i)^{1-Y_i}], \beta \in \mathbf{R}^2 \quad (\text{IV})$$

O princípio da máxima verossimilhança é estimador de valor de β que maximiza $L(\beta)$.

Aplicando o logaritmo:

$$L(\beta) = \ln[\prod_{i=1}^n \pi_i^{Y_i} (1 - \pi_i)^{1-Y_i}] \quad (\text{V})$$

$$L(\beta) = \sum_{i=1}^n [\ln(\pi_i^{Y_i}) + (1 - Y_i)\ln(1 - \pi_i)] \quad (\text{VI})$$

$$L(\beta) = \sum_{i=1}^n [Y_i \ln \frac{\pi_i}{(1-\pi_i)} + \ln(1 - \pi_i)] \quad (\text{VII})$$

$$L(\beta) = \sum_{i=1}^n [Y_i(\beta_0 + \beta_1 x_i) + \ln(\frac{1}{1+e^{-(\beta_0+\beta_1 x_i)}})] \quad (\text{VIII})$$

A expressão é definida como:

$$L(\beta) = \sum_{i=1}^n [Y_i(\beta_0 + \beta_1 x_i) - \ln(1 + e^{-(\beta_0+\beta_1 x_i)})] \quad (\text{IX})$$

Para encontrar o valor de β que maximiza $L(\beta)$, deriva-se $L(\beta)$ em relação a cada parâmetro (β_0, β_1) , obtendo-se duas equações:

$$\frac{\partial(\beta)}{\partial(\beta_1)} = \sum_{i=1}^n [y_i - \frac{1}{1+e^{\beta_0+\beta_1 x_i}} * e^{\beta_0+\beta_1 x_i}] \quad (\text{X})$$

$$\frac{\partial(\beta)}{\partial(\beta_0)} = \sum_{i=1}^n [y_i x_i - \frac{1}{1+e^{\beta_0+\beta_1 x_i}} * x_i e^{\beta_0+\beta_1 x_i}] \quad (\text{XI})$$

Que igualdades a zero, geram o seguinte sistema de equações:

$$\begin{pmatrix} \sum_{i=1}^n (y_i - \pi_i) \\ \sum_{i=1}^n x_i (y_i - \pi_i) \end{pmatrix} = 0 \quad (\text{XII})$$

$$\text{em que } i = 1, 2, \dots, n \text{ e } \pi_i = \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}} \quad (\text{XIII})$$

Agora, vamos colocar essas equações em uma forma matricial

$$X^T (Y - \pi) = 0 \quad (\text{XIV})$$

Onde,

- X é a matriz de variáveis explicativas, com uma coluna de 1s para o intercepto e outra coluna para os valores de x_i
- Y é o vetor das respostas y_1, y_2, \dots, y_n
- π é o vetor das probabilidades preditas $\pi_1, \pi_2, \pi_3 \dots \pi_n$ onde $\pi_i = \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}$

3.4.2 Regressão Logística Múltipla

Conforme Belfiore (2015), a regressão logística múltipla é uma extensão de regressão binária, em que a variável dependente é binária ou dicotômica, ou seja, tem duas categorias e que há

mais de uma variável independente. No modelo anterior, temos uma única variável independente. Para o caso onde temos um conjunto p variáveis independentes expresso pelo vector $x^T = (x_1 x_2 \dots x_p)$. Para o Hosmer-Lemeshow (1989), no modelo de probabilidade múltipla, a probabilidade de sucesso é dada por:

$$\pi_i = P(Y_i = 1 | X_i = x_i) = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}} = \frac{e^{(x_i^T \beta)}}{1 + e^{(x_i^T \beta)}} \quad (\text{XV})$$

3.4.2.1 Estimação dos coeficientes do modelo de regressão múltipla

Para estimar os parâmetros da regressão logística múltipla por máxima verossimilhança encontra-se o valor de $L(\beta)$ que maximiza $L(\beta)$, o que exige um processo iterativo e que faz necessário derivar $L(\beta)$ em relação a cada parâmetro:

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i=1}^n [y_i x_{ij} - \frac{e^{(x_i^T \beta)}}{1 + e^{(x_i^T \beta)}} x_{ij}] \quad (\text{XVI})$$

Onde $L(\beta)$ é a função (vi).

A matriz de covariância dos coeficientes estimados é obtida a partir das derivadas parciais de segunda ordem do logaritmo da função de verossimilhança:

$$\frac{\partial^2 \ln[L(\beta)]}{\partial \beta_j^2} = - \sum_{i=1}^n [x_{ij}^2 \pi_i (1 - \pi_i)] \quad (\text{XVII})$$

$$\frac{\partial^2 \ln[L(\beta)]}{\partial \beta_j \partial \beta_k} = - \sum_{i=1}^n [x_{ij} x_{ik} \pi_i (1 - \pi_i)] \quad (\text{XVIII})$$

Onde: $j, k = 0, 1, 2, \dots, p$ e π_i representa $\pi(x_i)$.

Se for formada uma matriz quadrada de dimensão $(p + 1)$, constituída pelo simétrico dos valores médios dos termos referidos nas equações (10) e (11), obtém-se $I(\beta)$, a chamada matriz de informação.

A matriz de informação de Fisher é dada por:

$$I(\beta) = E \left(- \frac{\partial^2 L(\beta)}{\partial \beta \partial \beta^T} \right) = X^T Q X \quad (\text{XIX})$$

Em que, $Q = \text{diag}[\pi_i(1 - \pi_i)]$, $i = 1, \dots, n$ e X a matriz de dados e sua inversa $[I(\beta)^{-1}]$, é a matriz de variância e covariância das estimativas de máxima verossimilhança dos parâmetros.

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (\text{XX})$$

3.4.2.2 Critério de Informação de Akaike

Segundo Hosmer - Lemeshow (1989), o Critério de Informação de Akaike, conhecido como AIC, e o Critério de Informação Bayesiano, chamado de BIC, não são testes de hipótese; são ferramentas para a seleção de modelos. O AIC é um índice de ajuste geral que pressupõe a comparação de modelos adequados ao mesmo conjunto de dados, utilizando o mesmo método de estimativa.

Os critérios de informação apresentam duas vantagens em relação a outros métodos de seleção. Tanto o AIC quanto o BIC baseiam-se na log-verossimilhança, assim como outros métodos, mas também consideram e penalizam a complexidade do modelo (número de parâmetros), o que busca o alinhamento com o princípio da parcimônia, essencial na modelagem estatística. Além disso, os modelos comparados não precisam ser aninhados, superando a principal limitação da comparação de modelos por meio de testes de hipótese (Araújo, 2017).

Nas equações subsequentes, são apresentados os critérios AIC e BIC, que são expressos, respetivamente, por:

$$AIC = \frac{(-2l)}{n} + \frac{(-2k)}{n} \quad (XXI)$$

$$BIC = \frac{(-2l)}{n} + \frac{(-2\log n)}{n} \quad (XXII)$$

Onde: k é o número de restrições;

n é o tamanho da amostra e l é a função de verossimilhança.

3.4.2.3 Razão de chances (odds ratio)

Razão de Chances, ou Odds Ratio (OR), é uma medida que expressa a mudança nas chances de ocorrência de um determinado evento associado a uma unidade de aumento em uma variável explicativa, mantendo todas as outras variáveis constantes. No modelo logístico, o coeficiente de cada variável independente (X_i) pode ser interpretado como o logaritmo natural da Razão de Chances. Matematicamente, a relação é dada por:

$$OR = e^{\beta_i} \quad (XXIII)$$

Onde:

- ✓ β_i é o coeficiente estimado da variável independente X_i no modelo logístico;
- ✓ e é a base de logaritmo natural.

Assim, se $\beta_i > 0$, um aumento em X_i está associado a um aumento nas chances de evento ocorrer $OR > 1$. Se $\beta_i < 0$, um aumento em X_i está associado a uma diminuição nas chances do evento ocorrer $OR < 1$. Quando $\beta_i = 0$, o OR é igual a 1, indicando que variável X_i não afeta as chances do evento ocorrer.

Assim sendo, em um modelo de **regressão logística**, a probabilidade π_i de um evento ocorrer é dada pela função logística:

$$\pi_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}} \quad (\text{XXIV})$$

Onde, $\pi_i = P(Y_i = 1 | x_i)$ é a probabilidade de Y_i ser igual a 1, dado os valores de x_i

Agora, a **odds** de um evento $Y_i = 1$ é dada pela razão entre a probabilidade do evento ocorrer e a probabilidade de o evento não ocorrer:

$$Odds(Y_i = 1) = \frac{\pi_i}{1 - \pi_i} \quad (\text{XXV})$$

Substituindo a expressão para π_i da regressão logística:

$$Odds(Y_i = 1) = \frac{\frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}}}{1 - \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}}} \quad (\text{XXVI})$$

Isso se simplifica para:

$$Odds(Y_i = 1) = e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \quad (\text{XXVII})$$

3.4.2.4 Teste do Qui-Quadrado para ajustamentos

Segundo Agresti e Finlay (2009), utiliza-se o teste de independência quando os dados da pesquisa se apresentam sob forma de frequências em categorias discretas. Para se aplicar este teste de hipótese, recorre-se a estatística de Qui-quadrado χ^2 para determinar se duas variáveis são independentes, usando a seguinte estatística:

$$\chi^2 = \sum_{i=1}^l \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2 (L - 1)(C - 1) \quad (\text{XXVIII})$$

Onde: O_{ij} – número de casos observados na linha i da coluna j .

E_{ij} – número de casos esperados na linha i da coluna j

As hipóteses a serem testadas são:

H_0 : *As variáveis são independentes, isto é, não há associação entre elas*

H_1 : *As variáveis não são independentes, isto é, há associação entre elas*

3.4.2.5 Principais pressupostos do teste do Qui-Quadrado

Box, Hunter e Hunter (2005), discutem os pressupostos fundamentais para a aplicação do teste de qui-quadrado, enfatizando a importância do tamanho da amostra e da independência das observações, dada a sensibilidade do teste. Estes autores referem que neste teste, as frequências esperadas em cada célula da tabela sejam suficientemente grandes, geralmente superiores a 5.

Em resumo, esses autores apresentam os seguintes principais pressupostos:

- **Independência das observações:** todas as observações devem ser independentes umas das outras, ou seja, o valor de cada observação não deve influenciar o valor da outra;
- **Tamanho da amostra:** para este pressuposto, significa que a amostra deve ser suficientemente grande de modo a garantir que a distribuição do χ^2 se aproxime de uma distribuição qui-quadrado. Em regra, as frequências esperadas em cada célula devem ser pelo menos 5. Se esse critério não for atendido, o teste pode produzir resultados imprecisos;
- **Frequências esperadas:** as frequências esperadas em cada célula da tabela de contingência devem ser suficientemente grandes (maiores ou iguais a 5) para assegurar que a aproximação da distribuição qui-quadrado seja válida;

- **Categorias mutuamente exclusivas:** cada observação deve ser atribuída a apenas uma categoria ou seja, há necessidade de definição clara das categorias e não pode haver sobreposição de modo que não haja ambiguidade na alocação das observações;
- **Distribuição dos dados:** o teste de bondade de ajuste compara as frequências observadas com as esperadas que são calculadas com base numa distribuição teórica (como uma distribuição uniforme, normal, binomial). Para que o teste tenha validade, a distribuição esperada deve ser bem definida;
- **Dados categóricos:** o teste de qui-quadrado é adequado para dados categóricos. As variáveis que estão sendo analisadas precisam ser categóricas (nominais ou ordinais);
- **Não é adequado para pequenos números de classes:** para o teste de bondade de ajuste, o número de categorias deve ser razoável para a distribuição. Se tiver poucas categorias ou uma distribuição de dados muito concentrada em algumas categorias, as frequências esperadas podem ser muito pequenas, o que pode invalidar o teste de qui-quadrado.

3.4.3 Métodos de avaliação do modelo logístico

3.4.3.1 Teste de Razão de Verossimilhança

Como os resíduos do modelo logístico não seguem distribuição normal e nem apresentam variâncias constantes, não é possível obter estimativas ótimas dos coeficientes do modelo pelo método dos mínimos quadrados. O método que melhor se ajusta no modelo logístico é o Método de Máxima Verossimilhança. Este método estima os coeficientes de regressão que maximizam a probabilidade de encontrar as realizações da variável dependente amostrada, isto é, que maximizem a verossimilhança desses valores (Maroco, 2007). No entanto, testa-se as seguintes hipóteses:

$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_p = 0$ (O modelo não é significativo)

$H_1: \text{Existem pelo menos um } \beta_j \neq 0 \text{ } i = 1, 2, 3, \dots, p$ (O modelo é significativo)

A estatística do teste é:

$$G = -2 \ln \frac{L(\text{modelo sem variável})}{L(\text{Modelo com variável})} \quad (\text{XXIX})$$

A estatística G tem como objetivo, comparar o modelo em análise e o modelo saturado, ou seja, a comparação dos valores observados e dos valores esperados usando a função de verossimilhança.

Onde:

- $L(\text{modelo restrito})$ é a **verossimilhança** (likelihood) do modelo **restrito** (geralmente o modelo nulo, ou modelo sem a variável de interesse).
- $L(\text{modelo completo})$ é a verossimilhança do modelo **completo** (ou modelo alternativo, com a variável ou parâmetros de interesse).
- O **fator -2** é incluído para tornar a estatística G mais facilmente comparável com a distribuição qui-quadrado, que é o que torna o teste de razão de verossimilhança útil.

3.4.3.2 Pseudo R^2 de Cox e Snell

O Pseudo R^2 de Cox e Snell foi utilizado para avaliar o ajuste do modelo de regressão logística. Essa medida, embora similar ao R^2 utilizado na regressão linear, possui um limite superior inferior a 1, o que limita sua interpretação direta. Um valor mais próximo de 1 indica um melhor ajuste do modelo, embora essa métrica não atinja o valor máximo de 1, o que a diferencia de outras medidas, como o Pseudo R^2 de Nagelkerke (Cox & Snell, 1989). Pode ser calculado por:

$$R^2_{\text{Loggit}} = \frac{-2LL_{\text{nulo}} - (-2LL_{\text{modelo}})}{-2LL_{\text{nulo}}} \quad (\text{XXX})$$

3.4.3.3 Nagelkerke's R^2

Segundo Nagelkerkes (1991), Nagelkerke's R^2 considera uma versão ajustada do Cox & Snell R^2 , modificado para que seu valor máximo seja 1. Isso facilita a interpretação da proporção de variância explicada pelo modelo. É amplamente usado para avaliar a qualidade do ajuste do modelo logístico. Pode ser calculado por:

$$R^2_{Nagelkerke} = \frac{1 - \frac{L_{Full}}{L_0}}{1 - L_0^{\frac{2}{n}}} \quad (XXXI)$$

Onde:

L_{Full} : é a função de verossimilhança do modelo com os preditores.

L_0 : é a função de verossimilhança do modelo nulo (modelo com apenas o intercepto).

n : é o número de observações

3.4.3.4 McFadden's R^2

É usado para medir a qualidade do ajuste em modelos de regressão logística, comparando a verossimilhança do modelo completo com a do modelo nulo. Valores típicos para bons ajustes de modelos situam-se entre 0,2 e 0,4 enquanto que valores que estes indicam um ajuste mais fraco (McFadden, 1974).

$$R^2_{McFadden} = 1 - \frac{\ln L_{Full}}{\ln L_0} \quad (XXXII)$$

Onde:

$\ln L_{Full}$: é o logaritmo da verossimilhança do modelo completo.

$\ln L_0$: é o logaritmo da verossimilhança do modelo nulo.

3.4.3.5 Tjur's R^2 (Coeficiente de Discriminação)

Segundo Tjur (2009), Tjur's R^2 conhecido como coeficiente de discriminação, é uma medida simples que avalia a capacidade do modelo de distinguir entre os dois grupos de saída (por

exemplo, sucesso vs. fracasso). É a diferença entre as médias das probabilidades previstas para cada grupo. E pode ser calculada por:

$$R_{Tjur}^2 = \text{media}(\hat{y}_1) - \text{media}(\hat{y}_0) \quad (\text{XXXIII})$$

Onde:

\hat{y}_1 : é a média das probabilidades previstas para os casos onde $y = 1$

\hat{y}_0 : é a média das probabilidades previstas para os casos onde $y = 0$

3.4.3.6 sqPearson R^2 (Quadrado da Correlação de Pearson)

De acordo com Menard (2000), sqPearson R^2 é uma medida baseada no quadrado do coeficiente de correlação de Pearson entre as saídas observadas (0 ou 1) e as probabilidades previstas pelo modelo. Ele fornece uma estimativa da força da relação linear entre as saídas reais e as probabilidades previstas.

$$R^2_{sqPearson} = [\text{cor}(y, \hat{y})]^2 \quad (\text{XXXIV})$$

Onde:

y representa as saídas observadas (0 ou 1).

\hat{y} representa as probabilidades previstas pelo modelo.

$\text{cor}(y, \hat{y})$ é o coeficiente de correlação de Pearson entre y e \hat{y}

3.4.3.7 Teste de Hosmer-Lemeshow

Hosmer e Lemeshow (2000), propuseram um teste de ajustamento muito utilizado na regressão logística, com o objetivo de testar a qualidade do ajuste do modelo. Esse teste avalia se o modelo obtido pode explicar adequadamente os dados observados, dividindo os dados em g grupos segundo as probabilidades estimadas.

O teste considera as seguintes hipóteses:

H_0 : Não existem diferenças significativas entre os resultados previstos e observados

H_1 : Existem diferenças significativas entre os resultados previstos e observados.

A estatística do teste (\hat{C}) é dada por:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - E_k)^2}{(1 - E_k/n_k)} \quad (\text{XXXV})$$

Onde:

- O_k : A **frequência observada** (ou número de eventos observados) para o **k-ésimo grupo**.
- E_k : A **frequência esperada** (ou número de eventos esperados) para o **k-ésimo grupo**, com base nas probabilidades previstas pelo modelo.
- n_k : O **número total de observações** no **k-ésimo grupo**.
- g : O número de grupos (geralmente agrupados com base nas probabilidades previstas do modelo, por exemplo, dividindo as observações em decis ou quartis, dependendo da aplicação).

3.4.3.8 Curva Receiver Operating Characteristic (ROC)

De acordo com Hosmer & Lemeshow (2000), a curva ROC (Receiver Operating Characteristic) é uma ferramenta que permite avaliar o desempenho de um modelo de regressão binária (a variável resposta é do tipo 0 e 1). Pode ser feita por meio de um gráfico simples e robusto, que nos permite estudar a variação da sensibilidade e especificidade, para diferentes pontos de quebra. Deveremos considerar um ponto de quebra C e comparar cada probabilidade estimada com o valor de C. O valor mais utilizado para C é 0,5.

A **curva ROC** é um gráfico que representa a **sensibilidade** (ou taxa de verdadeiros positivos) no eixo das ordenadas (eixo y) e a **taxa de falsos positivos** (ou 1 - especificidade) no eixo das abscissas (eixo x). Ela é construída a partir da variação de um valor de corte (limiar de decisão) ao longo do eixo de decisão, mostrando como a sensibilidade e a taxa de falsos positivos mudam

à medida que o ponto de corte é ajustado. Esse gráfico permite visualizar o trade-off entre **sensibilidade** e **especificidade** em diferentes limiares de classificação.

Na abordagem de Hosmer & Lemeshow (2000) descreve a sensibilidade e especificidade da seguinte maneira:

A sensibilidade (também chamada de taxa de verdadeiros positivos ou recall) é uma medida crucial para avaliar a capacidade de um modelo em identificar corretamente os casos positivos. Em termos da curva ROC, a sensibilidade é representada pela taxa de verdadeiros positivos (TPR) no eixo vertical. Ela é calculada como a razão entre o número de casos positivos corretamente classificados (verdadeiros positivos) e o total de casos reais positivos (soma de verdadeiros positivos e falsos negativos). Uma alta sensibilidade indica que o modelo é eficaz em identificar a maioria dos casos positivos, sendo particularmente importante em contextos onde é crucial não perder muitos casos positivos, como em diagnósticos de doenças graves.

A especificidade, por outro lado, mede a capacidade do modelo de identificar corretamente os casos negativos, sendo a taxa de falsos negativos o principal parâmetro relacionado à sua interpretação. Na curva ROC, a especificidade é representada pela taxa de falsos positivos (FPR) no eixo horizontal, sendo calculada como a razão entre o número de negativos incorretamente classificados como positivos (falsos positivos) e o total de casos reais negativos (soma de falsos positivos e verdadeiros negativos). Quanto maior a especificidade, menor a probabilidade de que os casos negativos sejam incorretamente classificados como positivos, o que é importante em cenários onde o custo de classificar um caso negativo como positivo é alto. A curva ROC permite visualizar o trade-off entre sensibilidade e especificidade, pois à medida que se ajusta o limiar de classificação para aumentar a sensibilidade, pode-se acabar sacrificando a especificidade, e vice-versa. Esse equilíbrio é essencial para escolher o ponto de corte ideal que atenda às necessidades específicas de cada aplicação.

Assim, a representação da curva ROC permite evidenciar os valores para os quais existe otimização da Sensibilidade em função da Especificidade, correspondente ao ponto que se encontra mais próximo do canto superior esquerdo do diagrama, uma vez que o índice de verdadeiro positivo é 1 e o de falso positivo é 0.

A área abaixo da curva de ROC, fornece-nos uma medida de discriminação, que nos indica a possibilidade de um individuo não doente ter uma probabilidade estimada associada mais elevada do que um individuo doente.

Seja R , o valor que corresponde à área abaixo da curva de ROC, como regra geral de acordo com Hosmer & Lemeshow (2000), temos as seguintes linhas de orientação:

1. Se $R = 0.5$ não há discriminação
2. Se $0.5 < R \leq 0.6$ a discriminação é muito ruim
3. Se $0.6 < R \leq 0.7$ a discriminação é aceitável
4. Se $0.7 < R \leq 0.8$ a discriminação é boa
5. Se $0.8 < R \leq 0.9$ a discriminação é muito boa
6. Se $0.9 < R \leq 1.0$ a discriminação é excelente
7. Se $R = 1.0$ a discriminação é perfeita

3.5 Materiais e métodos

3.5.1 Acesso a base de dados

A base de dados usada na aplicação desta dissertação, obedeceu a um desenho de coorte retrospectivo, sendo que os dados foram recolhidos a partir duma base de dados secundária de pacientes adultos vivendo com HIV/SIDA no Centro de Saúde de Lichinga entre os anos 2020 e 2021. Os pacientes incluídos no estudo obedeceram aos critérios de inclusão descritos no subcapítulo anterior.

Os dados constam nos registos médicos dos pacientes com informações relevantes como: idade, sexo, quantidade de carga viral, duração do tratamento, histórico de comorbidades e outras medidas clínicas relevantes. Os dados foram acessados numa base de dados de armazenamento eletrónico com denominação de *OpenMRS*, a qual é usada para registo de todos os eventos dos pacientes em cuidados e tratamento do HIV. Importa referir que a base de dados é de acesso restrito pois, apenas os utilizadores credenciados tem acesso e o pesquisador solicitou

autorização para acessar e explorar a base de dados. Ademais, os dados já introduzidos nela, regra geral, permanecem inalterados, excepto por razões que justifiquem.

A análise estatística foi realizada usando o teste de qui-quadrado e Regressão Logística Múltipla no RStudio, ajustando o modelo com as variáveis mais relevantes. Igualmente foram apresentadas técnicas de Análise Exploratória de dados, Estatística Descritiva, Regressão Logística Múltipla incluindo as Técnicas de Inferência usualmente aplicadas aos parâmetros de regressão logística.

3.5.2 Questões éticas

Neste estudo, foram usados dados secundários coletados de pacientes que, a priori, foram anonimizados para garantir a privacidade e confidencialidade dos participantes. Não foram coletados dados diretamente dos pacientes e todas as informações utilizadas não contêm identificadores pessoais como nomes e outros dados que possam revelar a identidade dos participantes.

O uso dos dados foi realizado em conformidade com as diretrizes éticas aplicáveis, respeitando-se os princípios de confidencialidade e integridade. Como o estudo não envolveu interação direta com os pacientes, não foi necessário obter consentimento informado adicional.

3.5.3 Variáveis de estudo

✓ **Variável dependente:** Supressão viral

$$f(y) = \begin{cases} 1, & y < 1000 \\ 0, & y \geq 1000 \end{cases}$$

✓ **Variáveis independentes:** Idade, Sexo, Peso, Tempo, TB e APSS.

Do ponto de vista de saúde pública, o estudo da relação entre estas variáveis nesta pesquisa, irá ajudar a desenvolver intervenções e estratégias de tratamento mais eficazes e integradas, abordando não apenas o HIV mas também comorbidades como a Tuberculose e questões

nutricionais. Para além disso, vai assegurar o aprimoramento de políticas de saúde através de fornecimento de evidências para programas de saúde, integração dos serviços de TB e HIV e a promoção de adesão prolongada do tratamento antirretroviral.

3.5.4 Definição das variáveis de estudo

De modo a assegurar maior clareza das análises que serão feitas, há necessidade de clarificar a descrição de cada variável, conforme consta na tabela abaixo:

Tabela 3.1 Descrição das variáveis estudadas

Nome da variável	Classificação	Descrição
Supressão viral	Dependente	É uma variável categórica binária que expressa o desfecho da condição de saúde do paciente. Entende-se por Supressão Viral a um resultado dum Carga Viral (CV) dum paciente abaixo de 1000 cópias/ml. Foi classificada como: 1=sucesso para pacientes abaixo 1000 cópias/ml e 0=insucesso para o restante grupo
Idade	Independentes	É uma variável quantitativa contínua que expressa a idade dos pacientes em anos. Esta variável foi integrada para análise no modelo logístico ajustado considerando-a em idades simples. Para além de ser analisada no formato contínuo, a mesma variável foi convertida para categórica através de agrupamento de idades para faixas etárias. Com esta conversão, permitiu ser analisada como categórica no teste de qui-quadrado
Sexo		É uma variável categórica binária que expressa o sexo do paciente classificada em 0=masculino e 1=feminino
Peso		É uma variável contínua referente ao peso corporal dos pacientes medido em quilogramas
Tuberculose		Variável categórica binária que classifica a presença ou não da Tuberculose nos pacientes sendo: 0=Não presença e 1=Presença da tuberculose no paciente
Apoio Psicossocial		É uma variável categórica binária que classifica o nível de apoio Psicossocial sendo 1=alto e 0=baixo . Esta classificação é baseada no número de pacotes de apoio psicossocial oferecidos ao doente. No padrão clínico moçambicano, os pacientes HIV+ devem ser oferecidos 7 pacotes de apoio psicossocial e o pesquisador classificou como baixo aos pacientes com acesso até 4 pacotes de apoio psicossocial e classificou como alto para o restante grupo

Fonte: Autor

3.5.5 Análise Descritiva (Univariada vs Bivariada)

- ✓ Apresentação das características demográficas e clínicas da amostra, como idade, sexo, distribuição dos valores de supressão viral;
- ✓ Análise da associação entre supressão viral com outras características.

3.5.6 Aplicação do Modelo de Regressão Logística Múltipla

- ✓ Ajuste do modelo de regressão binária, utilizando a variável de supressão viral como desfecho e as variáveis independentes como: Idade, Sexo, Peso, Tempo, TB e APSS;
- ✓ Interpretação dos coeficientes estimados e as razões de chance (odds ratio) para cada variável independente;
- ✓ Avaliação da significância estatística dos coeficientes usando a distribuição qui-quadrado;
- ✓ Identificação de quais variáveis independentes estão significativamente associadas à supressão viral com base no modelo logístico.

3.5.7 Definição da qualidade do ajuste do modelo

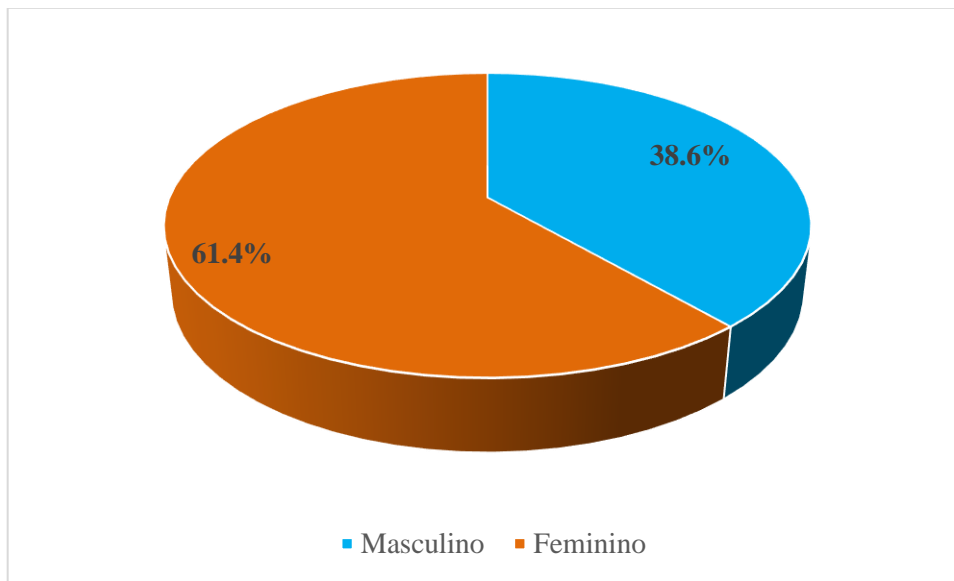
- ✓ Avaliada a qualidade do ajuste do modelo, utilizando medidas como AIC, BIC e teste de Hosmer - Lemeshow;
- ✓ Através do Pseudo R^2 , verificado se os pressupostos do modelo, são atendidos. E por último,
- ✓ Validação do modelo usando a Curva ROC.

Capítulo IV

4. Aplicação da análise de Regressão Logística Múltipla

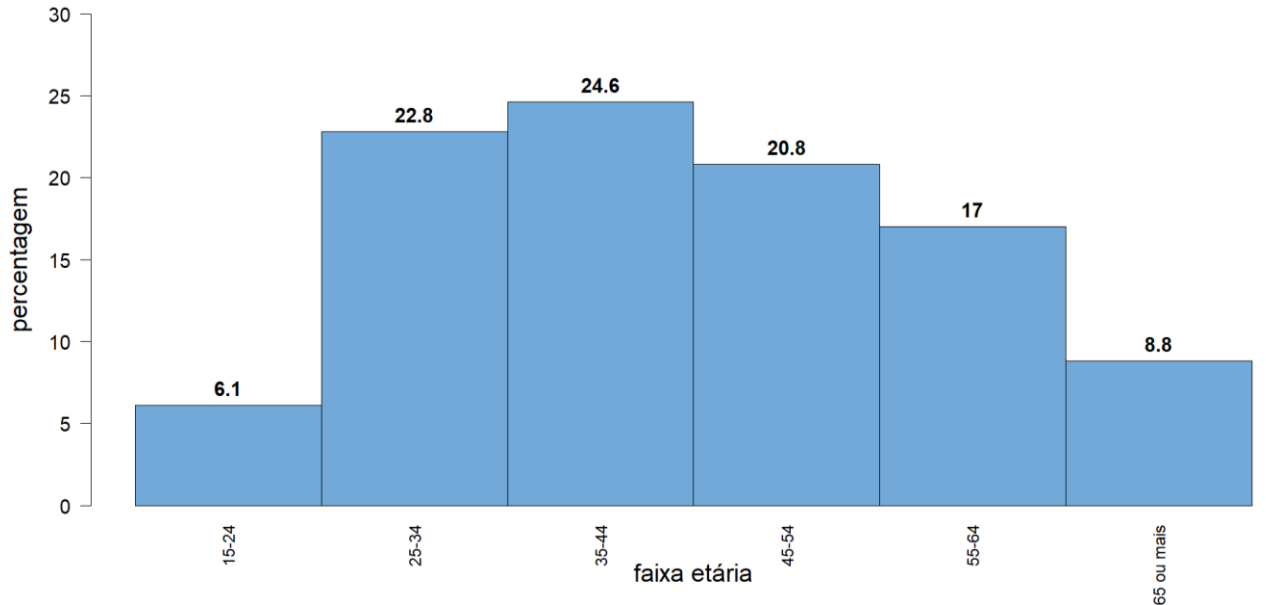
4.1 Apresentação das características demográficas do grupo estudado

Gráfico 4.1 Distribuição dos pacientes por sexo



O gráfico 4.1, revela uma distribuição desigual dos pacientes por sexo, com uma maior predominância do sexo feminino (61.4%) em comparação com pacientes do sexo masculino (38.6%). Esta maior proporção pode refletir fatores como maior prevalência de determinadas condições de saúde entre mulheres, maior tendência das mulheres em buscar cuidados médicos, ou uma combinação destes e outros fatores. Relativamente a esta menor proporção dos homens, pode indicar que poucos buscam cuidados médicos, ou que certas condições são menos prevalentes entre homens na população estudada.

Gráfico 4.2: Distribuição dos pacientes por faixa-etária



Fonte: Autor

A análise da distribuição dos pacientes por faixa etária revela que a maioria está concentrada nas faixas etárias entre os 25 a 54 anos, perfazendo cerca de 68.1% do total. As faixas de 65 anos ou mais (8.8%) e 15-24 anos (6.1%) representam proporções menores do grupo estudado.

1. Faixa Etária 15-24 anos (6.1%):

- Representa uma das menores proporções de pacientes. Isso pode indicar que os jovens são menos afetados pela condição estudada ou têm menor probabilidade de buscar atendimento médico.

2. Faixa Etária 25-34 anos (22.8%), 35-44 (24.6%) e 45-54 anos (20.8%):

- Estas três faixas etárias têm as maiores proporções de pacientes, indicando que a população adulta está mais vulnerável e há um aumento na prevalência da condição ou na procura por atendimento médico nestas faixas etárias, pois é a idade sexualmente mais ativa. Contudo, nota se pequenas diferenças entre estas faixas sendo que a faixa dos 35-44 é a que apresenta maior proporção de todas.

3. Faixa Etária 55-64 anos (17.0%):

- A proporção é ligeiramente menor do que a faixa anterior, mas ainda considerável, refletindo um aumento contínuo na necessidade de cuidados de saúde com o aumento da idade.

4. Faixa Etária 65 anos ou mais (8.8%):

- Esta é também das faixas etárias com a menor proporção de pacientes. A baixa prevalência pode ser devido ao diagnóstico tardio do HIV entre idosos ou baixa procura e utilização dos serviços de saúde a partir desta faixa etária.

4.2 Análise e verificação dos pressupostos do modelo de Regressão Logística Múltipla

4.2.1 Análise dos fatores associados à supressão viral

Tabela 4.1: Análise do teste do qui-quadrado dos fatores associados à supressão viral em pacientes adultos vivendo com o HIV/SIDA

	Sem supressão viral	Supressão viral	P-valor
Sexo			
Masculino	42	67	1
Feminino	66	106	
Faixa-etária			
15-24	6	10	
25-34	26	35	
35-44	25	45	0.82
45-54	21	37	
55-64	17	32	
65 ou mais	13	14	
Tuberculose			
Não	74	154	<0.000038
Sim	34	19	
APSS			
Alto	64	115	0.27
Baixo	44	58	

A Tabela 4.1 apresenta os resultados do teste do qui-quadrado para vários fatores associados à supressão viral em pacientes adultos vivendo com HIV/SIDA. Os fatores analisados incluem sexo, faixa etária, presença ou não da tuberculose e nível de apoio psicossocial (APSS).

Analisando cada fator com base nos resultados fornecidos pelo teste qui-quadrado, temos o seguinte:

No fator sexo, notou-se que não há diferença significativa na supressão viral entre homens e mulheres (p-valor = 1). Isso sugere que o sexo dos pacientes não influencia significativamente a probabilidade de alcançar a supressão viral. Segundo a mesma tabela, não há uma associação estatisticamente significativa entre a faixa etária e a supressão viral (p-valor = 0.82). Portanto, a idade dos pacientes não parece ter um impacto significativo na probabilidade de alcançar a supressão viral.

Em relação ao fator tuberculose, notou-se com os resultados que existe uma associação estatisticamente significativa entre a presença de tuberculose e a supressão viral (p-valor < 0.000038). Significa com isso que pacientes com tuberculose têm uma probabilidade significativamente diferente de alcançar a supressão viral em comparação com aqueles sem tuberculose.

Quanto ao apoio psicossocial, não mostrou-se um fator estatisticamente significativo para a supressão viral (p-valor = 0.27). Portanto, o nível de apoio psicossocial não parece influenciar significativamente a probabilidade de alcançar a supressão viral.

De todos os fatores analisados, a presença de tuberculose é o único fator significativamente associado à supressão viral em pacientes adultos vivendo com HIV/SIDA. Fatores como sexo, faixa etária e nível de apoio psicossocial não mostraram uma associação estatisticamente significativa com a supressão viral.

4.2.2 Criação do Modelo zero

Com o objetivo de assegurar uma sequência adequada na análise dos dados, foi desenvolvido o **Modelo Zero**. A partir desse modelo inicial, foi utilizado o método “*stepwise*”³, aliado à

³ O método *stepwise* é uma abordagem utilizada em regressões para selecionar as variáveis preditoras mais significativas para incluir no modelo, de forma iterativa. O objetivo é construir um modelo estatístico mais simples, sem sacrificar a precisão das previsões.

ferramenta de AIC (Critério de Informação de Akaike), para identificar o melhor modelo entre aqueles avaliados. O **Modelo Zero** foi formulado com base na seguinte equação:

$$\text{Modelo0} = \text{glm}(CV_class \sim (as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo + as.factor(APSS_Class))^2, data = train, family = binomial(link = "logit")) \quad (\text{XXXVI})$$

4.2.3 Seleção do melhor modelo logístico binário

A definição do melhor modelo com base no Critério de Informação de Akaike (AIC⁴) tem sido uma das etapas cruciais na modelagem dos dados e a cautela é um dos princípios que deve ser priorizado, tendo em conta que o AIC continua a ser a ferramenta mais usada pelos profissionais para a seleção do modelo, (Lucambio, 2020).

Por esse motivo, com objetivo de encontrar o modelo que explica melhor os dados, o pesquisador adotou o AIC para encontrar o modelo que ajustou melhor os seus dados. Na tabela abaixo, são apresentados os resultados:

Tabela 4.2: Critério de informação de Akaike para seleção do melhor modelo logístico binário

Modelos	Critério de informação de Akaike (AIC)
1	90.48
2	88.49
3	86.50
4	84.53
5	82.62
6	80.74
7	78.91
8	77.38
9	77.20
10	77.01
11	76.67

A Tabela 4.2 apresenta os valores do AIC para diversos modelos logísticos binários. O AIC é uma medida usada para comparar a qualidade de diferentes modelos estatísticos. Em geral,

⁴ Akaike Information Criterion – ferramenta inventada em 1973 pelo estatístico japonês Hirotugu Akaike

modelos com valores de AIC mais baixos são preferidos, pois indicam um melhor equilíbrio entre a qualidade do ajuste do modelo e a complexidade do modelo (número de parâmetros). Os valores do AIC diminuem à medida que avançamos do Modelo 1 para o Modelo 11. Isso indica que cada modelo sucessivo ajusta melhor os dados do que o anterior, levando em conta a complexidade do modelo. O Modelo 11 tem o menor valor de AIC (76.67), indicando que é o modelo preferido entre os avaliados, pois oferece o melhor equilíbrio entre o ajuste do modelo e a penalização pela complexidade. Isso sugere que este modelo fornece a melhor explicação dos dados, considerando tanto a qualidade do ajuste quanto a complexidade do modelo.

4.2.4 Principais diferenças entre os modelos

Durante o processo de seleção de variáveis utilizando o método *stepwise* com AIC, foram avaliados diversos modelos. A cada iteração, variáveis e interações foram removidas com base na redução do AIC, que reflete a melhoria do ajuste do modelo e a diminuição da sua complexidade. O modelo inicial (Start) apresentou um AIC de 90.48, incluindo uma grande variedade de variáveis e interações.

A partir do modelo inicial com AIC=90.48, foram progressivamente removidas interações e variáveis menos significativas. Por exemplo, as interações *as.factor(Sexo):as.factor(TB)* e *Tempo:as.factor(APSS_class)* foram descartadas nas primeiras iterações, com o AIC caindo para 88.49. Modelos subsequentes, com AIC em torno de 84-86, continuaram ajustando o conjunto de variáveis, eliminando termos como *Idade:Peso* e *Peso:as.factor(APSS_class)*.

No estágio final (AIC=76.67), o modelo inclui as variáveis principais, como Sexo, Idade, TB, Peso e Tempo, e algumas interações chave como *Idade:as.factor(APSS_class)* e *Peso:Tempo*, ao mesmo tempo em que exclui outras interações complexas e redundantes.

O modelo final tem o melhor ajuste e complexidade, com o menor valor de AIC, o que indica a sua maior eficácia preditiva.

Diante destas principais diferenças entre o Modelo 1 e o 11, importa referir que os demais detalhes e diferenças entre os modelos, constam no anexo desta dissertação em formato de *output* do RStudio.

Assim sendo, o modelo final tomou a seguinte equação: (XXXVII)

```
ModeloFinal <- glm(CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso + Idade:as.factor(APSS_class) +
as.factor(TB):as.factor(APSS_class) + Peso:Tempo, family = binomial(link = "logit"), data = train)
```

A partir deste modelo final, gerou se as estimativas finais dos coeficientes do modelo logístico ajustado.

4.2.5 Estimativas dos coeficientes do modelo ajustado

Tabela 4.3: Estimativas dos coeficientes do Modelo Logístico Ajustado

Variáveis	Estimativas	OR	Limite Inferior (IC 2.5%)	Limite Superior (IC 97.5%)	Erro padrão	P-valor
Intercepto	-41.21	-	2.17×10^{-29}	5.03×10^{-11}	10.47	0.00008
SexoFeminino	9.65	15589.1	17.3	9.44×10^8	5.02	0.05
Idade	0.06	1.06	0.96	1.2	0.06	0.29
TBSim	-3.58	0.03	0.0015	0.39	1.36	0.009
Peso	0.6	1.83	1.36	2.78	0.17	0.0005
Tempo	1.63	5.11	2.45	13.97	0.43	0.0001
APSSAlto	5.72	306.54	0.46	461151.2	3.44	0.09
SexoFeminino:Idade	-0.08	0.92	0.81	1.02	0.05	0.15
SexoFeminino:Peso	-0.15	0.88	0.67	1.05	0.11	0.17
Idade:APSSAlto	-0.12	0.89	0.77	1.02	0.07	0.09
TBSim:APSSAlto	3.88	48.5	1.28	3184.61	1.94	0.04
Peso:Tempo	-0.02	0.98	0.96	0.99	0.01	0.008

Conforme os resultados da tabela 4.3 acima, o **intercepto** é estatisticamente significativo (p-valor <0.05), indicando que a função de ligação logit do modelo ajustado não é zero quando todas as outras variáveis são zero, ou seja, com 95% de confiança pode-se rejeitar a hipótese nula de que o intercepto é igual a zero.

A estimativa negativa sugere uma baixa probabilidade inicial de supressão viral na ausência de outras variáveis.

Quanto ao **Sexo**, o coeficiente não é estatisticamente significativo (p -valor = 0.05). A razão de chances (OR = 15589.10) sugere que as mulheres têm mais probabilidade de alcançar a supressão viral em comparação com homens. Apesar da falta de significância estatística para considerar real este efeito, este resultado sugere a uma associação extremamente forte entre ser do sexo feminino e a probabilidade de alcançar a supressão viral. O intervalo de confiança de **17.3 a 9.44×10^8** também sugere grande incerteza, apesar do valor muito elevado.

Embora nesta pesquisa mostra que as mulheres têm mais probabilidade de alcançar a supressão viral em comparação aos homens, uma pesquisa realizada pelo Myers et al., 2017, relativa a diferenças de gênero na supressão viral, teve resultados diferentes ao revelar que as mulheres podem enfrentar desafios adicionais relacionados à adesão ao tratamento e à supressão viral em comparação com os homens, possivelmente devido a fatores sociais, económicos e hormonais, podendo com isso ter a menor chance de alcançar a supressão viral quando comparadas aos homens.

A **idade** não é estatisticamente significativa conforme o (p -valor = 0.29). Por outro lado, o coeficiente positivo e a (OR = 1.06) sugere que para cada unidade adicional da idade, as chances de alcançar a supressão viral aumentam em cerca de 6%. Apesar de não ser significativo este efeito, contudo, pode haver uma associação entre a idade e um nível de aumento na probabilidade de alcance da supressão viral, apesar de que seja uma mudança relativamente ínfima.

Um estudo feito para analisar a relação entre faixas etárias e a supressão viral de Reiss et al. (2019), revelaram que a resposta do tratamento antirretroviral pode variar com a idade. Em alguns casos, pessoas mais velhas podem ter uma resposta menos eficaz ao tratamento devido a comorbidades ou efeitos colaterais de medicamentos. Pessoas mais velhas, geralmente enfrentam comorbidades como doenças cardiovasculares e diabetes o que pode impactar a eficácia do tratamento e a supressão viral.

A presença de **tuberculose** (TBSim) é estatisticamente significativa ao nível de 5% pois apresentou um (p-valor = 0.009). Esta estatística de teste sugere uma forte evidência de que a presença da TB nos pacientes HIV+ tem um impacto significativamente negativo para alcançar a supressão viral em relação aos pacientes que não tenham essa doença. Por outro lado, a (OR = 0.03), significa que as chances de alcançar a supressão viral em pacientes com TB, tem uma redução de cerca de 97%. Entretanto, o coeficiente negativo de -3.58 sugere que a presença da TB está associado a uma de redução da probabilidade de alcançar a supressão viral. O intervalo de confiança (**0.0015 a 0.39**) confirma essa associação com um risco muito baixo, mas com um intervalo relativamente largo.

Um estudo recentemente realizado no Centro de Tratamento Integrado do Hospital Militar de Maputo, revelou que pacientes com TB ativa apresentavam taxas altas de não alcançar a supressão viral (23.1%) em comparação àqueles sem TB, sugerindo que a TB pode comprometer a eficácia do TARV, Chicanequisso, et al.,(2024)

O peso é estatisticamente significativo (p-valor = 0.0005), neste modelo, a variável peso foi identificada como um preditor significativo para um paciente alcançar a supressão viral. Relativamente a razão de chances (OR de 1.83) sugere que para cada unidade adicional de peso, resulta em 83% de mais chances para alcançar a supressão viral. Além disso, o coeficiente de 0.60 evidencia que há um aumento de 0.60 unidades no logaritmo das odds de supressão viral, à medida que o peso for aumentando. Igualmente, o intervalo de confiança (1.36 a 2.78) é estreito, o que sugere um efeito confiável.

Gomes e Lourival (2016) afirmam que conhecer o perfil nutricional dos pacientes é o primeiro passo para uma nutrição adequada e um tratamento bem-sucedido, garantindo ao paciente uma melhora na qualidade de vida. Esta teoria reforça a necessidade de que no contexto de prestação de cuidados aos pacientes HIV+, focalizar outros elementos que ajudem a melhorar a condição de vida do paciente é muito prioritário.

Isso é caso para alertar os profissionais da área de nutrição que devem focalizar também as suas atenções aos pacientes HIV+ com vista a diagnosticar casos de pacientes com subnutrição e orientar intervenções de forma particularizada para o sucesso da terapia.

Duran (2009), no seu trabalho de dissertação, refere que a existência da relação entre o HIV e nutrição é evidente, de forma que a assistência nutricional pode ajudar as PVHS a lidar com as complicações relacionadas ao HIV, promover boas respostas ao tratamento e melhorar a qualidade de vida do indivíduo pela manutenção de funcionalidade e dignidade humana.

Estes fundamentos reforçam a necessidade do controlo de peso pois, sabe-se que o peso muito baixo como resultado de subnutrição assim como a obesidade relacionada a práticas nutricionais inadequadas, têm uma associação a probabilidade menor de alcançar a supressão viral em relação aos pacientes com peso normal.

Quanto ao **tempo de tratamento**, é um preditor estatisticamente significativo (p-valor = 0.0001) e isso sugere que o tempo em que o paciente está em tratamento antirretroviral, influencia a probabilidade de alcançar a supressão viral. Uma OR de 5.11 indica que a probabilidade de alcançar a supressão viral aumenta com o tempo, ou seja, a cada mês adicional do tempo de tratamento, a chance alcançar a supressão aumenta em **411%**. Considerando o coeficiente de 1.63 evidencia igualmente que há um aumento na log-odds da supressão viral à medida que o tempo de tratamento do paciente aumenta. O intervalo (2.45 a 13.97) também é relativamente largo, mas sugere uma associação forte e significativa. Quando o paciente com HIV/SIDA está em tratamento antirretroviral, o resultado final esperando ao longo do tempo é que alcance a supressão viral, ou seja, resultado duma Carga Viral abaixo de 1000 cópias/ml e chegando a alcançar níveis indetectáveis no sangue.

De acordo com Deeks, S.G., et al (2015) a maioria dos pacientes que aderem ao tratamento consistentemente ao TARV, alcançam a supressão viral de 6 a 12 meses de tratamento. Esta teoria de que o tempo longo no tratamento influencia na supressão viral, é sustentada pelo Ford, N., et al (2014) pois na sua pesquisa, concluíram que pacientes com alta adesão ao tratamento, a mediana de tempo para alcançar supressão viral foi aproximadamente de 6 meses, variando conforme os tipos de TARV usados e adesão dos pacientes.

Face ao exposto acima, o seguimento de adesão é crucial para a melhoria da qualidade de vida dos pacientes em TARV e é necessário que se cumpra com o calendário mensal nos primeiros

meses após o início do TARV, seguido de seguimento trimestral aliado a oferta do pacote de prevenção positiva.

Relativamente ao nível de **apoio psicossocial (APSS Alto)** aos pacientes em TARV naquela unidade de estudo, mostra que não é estatisticamente significativo considerando (p-valor = 0.09), porém, **OR de 306.54** sugere que a oferta de **APSSAlto** está associada a um aumento significativo nas chances de alcançar a supressão viral, evidenciando o impacto potencialmente muito grande. Contudo, dada a ausência da estatística de teste para provar, este efeito pode não ser real e, o intervalo de confiança muito amplo (0.46 a 461151.2) indica que há grande incerteza quanto ao efeito exato.

Vários estudos revelam que oferecer pacotes de apoio psicossocial aos pacientes em TARV tem uma influência em saber lidar com os desafios emocionais e práticas de viver com o HIV, aumentando a chance de alcançar a supressão viral. A contínua oferta de apoio psicossocial ajuda na redução do estresse e da depressão e concorre para a adoção de estilos de vida saudáveis.

Evidências científicas trazidas por Safren, S.A., et al (2014) demonstraram que intervenções psicossociais, como terapia cognitivo-comportamental e aconselhamento, melhoraram significativamente a adesão ao TARV e os níveis de supressão viral em pacientes com HIV. Por outro lado, Sherr, L., et al (2011) encontraram evidências consistentes de que o apoio psicossocial melhora a adesão ao tratamento, reduz o estresse e a depressão e consequentemente melhora os resultados da saúde do paciente incluindo a supressão viral.

No concernente a **interação entre sexo feminino e idade** não é estatisticamente significativa (p-valor = 0.15). A OR de 0.92 sugere que para cada unidade adicional na interação entre sexo feminino e idade, as chances de alcançar a supressão viral diminuem em cerca de 8%, embora essa associação não seja estatisticamente significativa. O intervalo de confiança (0.81 a 1.02) sugere que essa interação não tem um efeito muito claro.

A **interação entre sexo feminino e peso** também não é estatisticamente significativa (p-valor = 0.17). A **OR de 0.88** sugere que o efeito do aumento do peso nas mulheres pode ser um pouco menor, mas o intervalo (0.67 a 1.05) indica incerteza.

Relativamente a interação entre a **idade e o APSS alto**, a estatística de teste com (p-valor = 0.09), mostra que não há evidência suficiente para afirmar que essa interação tenha um efeito significativo na supressão viral. A OR=0.89 indica que para cada unidade adicional na interação entre idade e alto APSS, as chances de alcançar a supressão viral diminuem em cerca de 11%. Por outro lado, o coeficiente -0.12 evidencia que à medida que a idade aumenta em contexto de alto apoio psicossocial, o logaritmo da odds de supressão viral tende a reduzir. Contudo, este efeito pode não ser verdadeiro devido a ausência do teste estatístico para validar. Por outro lado, o intervalo (0.77 a 1.02) novamente sugere incerteza sobre esse efeito.

A interação entre a presença de **tuberculose e alto apoio psicossocial** é estatisticamente significativa (p-valor = 0.04), sugerindo uma relação importante entre a presença da TB e a eficácia do apoio psicossocial na supressão viral. A OR de 48.50 indica que para cada unidade adicional na interação “TB:APSS alto” as chances de alcançar a supressão viral aumentam substancialmente. Este elevado aumento sugere que os pacientes que têm TB e que recebem um alto nível de apoio psicossocial, têm significativamente mais probabilidade de alcançar a supressão viral em comparação àqueles que não têm esse apoio. O intervalo de confiança (1.28 a 3184.61) confirma a grande incerteza, mas ainda assim é um efeito forte.

Sabe-se no entanto que o apoio psicossocial pode desempenhar um papel crucial no aumento da adesão ao tratamento e consequentemente, melhores resultados no tratamento.

No estudo feito pelo Kipp, W., et al (2011) concluiu que as intervenções de apoio psicossocial em pacientes co-infetados com TB e HIV, melhoraram significativamente a adesão ao tratamento e os resultados de supressão viral nesse grupo de pacientes.

O mesmo resultado foi apresentado pelo estudo realizado por Pachi, A., et al (2013) segundo o qual, numa pesquisa longitudinal investigou o impacto do apoio psicossocial na adesão ao tratamento da TB e supressão viral em pacientes com HIV. Nessa pesquisa concluíram que o

apoio psicossocial foi associado a uma maior adesão ao tratamento e melhores taxas de supressão viral.

Com essas evidências, é caso para afirmar que realizar sessões de aconselhamento psicossocial para ajudar os pacientes a lidar com a sua condição de saúde, ansiedade e estresse, é fundamental e benéfico para a maioria deles. Ademais, o envolvimento de equipas multidisciplinares de profissionais de saúde (enfermeiros, médicos, psicólogos e outras áreas) pode ajudar na prestação de cuidados de forma integrada aos pacientes.

Por fim, a interação entre o **peso e tempo** de tratamento é estatisticamente significativa (p-valor = 0.008), sugerindo que há influência do peso na probabilidade de alcançar a supressão viral com aumento do tempo de tratamento. Por outro lado, a razão de chances (OR = 0.98) sugere que para cada unidade adicional na interação entre o peso e tempo de tratamento, as chances de alcançar a supressão viral diminuem ligeiramente em cerca de 2%. Por outro lado, o coeficiente de -0.02 confirma que a interação entre o peso e tempo de tratamento está associada a uma diminuição da log-odds de supressão viral. O intervalo de confiança (0.96 a 0.99) confirma que o efeito é pequeno e negativo (as chances diminuem ligeiramente). A ilação que se pode tirar desta interação entre o peso e tempo de tratamento, é avaliar a eficácia do tratamento antirretroviral, sugerindo que ajustes no tratamento em função do peso, podem ser necessários para otimizar a supressão viral em diferentes perfis de pacientes.

4.2.6 Definição da equação geral preditiva

Depois da análise de todos os resultados gerados pelo modelo final, será desenhada a equação geral considerando todos os parâmetros de cada variável, seja ela significativa ou não, a ser dada pelo modelo geral:

$$\text{logit}(P(Y = 1)) = \ln \left(\frac{P(Y=1)}{1-P(Y=1)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (\text{XXXVIII})$$

Sendo que,

- $\text{logit}(P(Y = 1))$ é o logaritmo das odds de Y ser igual a 1, ou seja é o logaritmo da razão de chance para que seja alcançada a supressão viral;
- $P(Y = 1)$ é a probabilidade de Y seja igual a 1, ou seja, a probabilidade de que a supressão viral seja alcançada a 100%;
- β_0 é o intercepto (constante), é um ponto de referência inicial do modelo;
- $\beta_1, \beta_2, \dots, \beta_n$ são os coeficientes das variáveis preditoras X_1, X_2, \dots, X_n

4.2.6.1 Considerando todas as variáveis, independentemente do seu teste de significância na equação final

Numa primeira fase, todas as variáveis do Modelo de Regressão Logística Múltipla foram mantidas na equação, independentemente de sua significância estatística dada pelo p-valor. Essa abordagem é baseada em um modelo “full model” ou modelo completo, onde todas as variáveis independentes (preditoras) são incluídas na equação de regressão, independentemente do valor do p-valor de cada variável.

Em Modelos de Regressão Logística Múltipla, essa abordagem é sustentada pelo Kutner et al. (2005), segundo o qual sugere que a inclusão de variáveis relevantes, mesmo aquelas com p-valores maiores que o limiar convencional de 0.05, é importante para evitar o viés de omissão e garantir que o modelo reflita adequadamente a complexidade e as relações entre as variáveis.

Além disso, a inclusão de todas as variáveis permite uma melhor interpretação dos efeitos marginais e a avaliação do impacto de variáveis potencialmente importantes, que podem influenciar o resultado de forma indireta ou em interação com outras variáveis. Assim, mesmo que algumas variáveis individuais não apresentem significância estatística, elas foram mantidas para preservar a consistência e a integridade do modelo ajustado.

Por via disso, para calcular a probabilidade $P(Y = 1)$ a equação é desdobrada da seguinte forma:

(XXXIX)

$$P(Y = 1) = \frac{e^{\beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + \beta_6 * X_6 + \beta_7 * X_7 + \beta_8 * X_8 + \beta_9 * X_9 + \beta_{10} * X_{10} + \beta_{11} * X_{11}}}{1 + e^{\beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + \beta_6 * X_6 + \beta_7 * X_7 + \beta_8 * X_8 + \beta_9 * X_9 + \beta_{10} * X_{10} + \beta_{11} * X_{11}}}$$

No entanto, considerando os valores de todos os coeficientes de cada variável tem-se o seguinte:

- $\beta_0 = -42.21$; $\beta_1 = 9.65$; $\beta_2 = 0.06$; $\beta_3 = -3.58$; $\beta_4 = 0.6$; $\beta_5 = 1.63$;
- $\beta_6 = 5.72$; $\beta_7 = -0.08$; $\beta_8 = -0.15$; $\beta_9 = -0.12$; $\beta_{10} = 3.88$;
- $\beta_{11} = -0.02$

Assim sendo, a equação do modelo substituindo todos os parâmetros pelas respectivas estimativas encontradas pelo método de máxima verossimilhança é:

(XL)

$$\begin{aligned} \text{logit}(P(Y = 1)) = & -41.21 + 9.65 * \text{SexoFeminino} + 0.06 * \text{Idade} - 3.58 * \text{TBSim} + 0.6 * \text{Peso} + 1.63 \\ & * \text{Tempo} + 5.72 * \text{APSSAlto} - 0.08 * \text{SexoFeminino:Idade} - 0.15 * \text{SexoFeminino:Peso} \\ & - 0.12 * \text{Idade:APSSAlto} + 3.88 * \text{TBSim:APSSAlto} - 0.02 * \text{Peso:Tempo} \end{aligned}$$

Esta fórmula representa a logaritmo de razão de chances (odds ratio) para que seja alcançada a supressão viral pelos pacientes em tratamento antirretroviral no Centro de Saúde da Cidade de Lichinga ($Y = 1$), considerando todas as variáveis do modelo, independentemente do seu teste de significância, sendo que:

- **SexoFeminino** – é uma variável categoria binária classificada em (0=Masculino e 1=Feminino);
- **Idade** - é uma variável contínua que expressa a idade dos pacientes em anos completos.
- **TBSim** - é uma variável binária indicando a presença ou não da tuberculose (0 = Não; 1 = Sim);
- **Peso** - é uma variável quantitativa contínua e representa o peso corporal do paciente em quilogramas;
- **Tempo** - é também uma variável quantitativa contínua que representa o período de tratamento do paciente em meses;

- **APSSAlto**: É uma variável categórica binária que classifica o nível de Apoio Psicossocial sendo 1=Alto e 0=Baixo.
- **SexoFeminino:Idade** – representa a interação da variável categórica Sexo Feminino com a variável contínua Idade;
- **SexoFeminino:Peso** - representa a interação da variável categórica Sexo Feminino com a variável contínua Peso;
- **Idade:APSSAlto** - representa a interação da variável contínua Idade com a variável categórica Apoio Psicossocial Alto;
- **TBSim:APSSAlto** – representa a interação entre a presença da Tuberculose e o nível Alto de acesso aos pacotes de Apoio Psicossocial;
- **Peso:Tempo** – é uma variável quantitativa contínua que representa a interação entre o peso do paciente e seu tempo de tratamento (é determinada pelo produto do peso e tempo).

Entretanto, para calcular a probabilidade $P(Y = 1)$, a fórmula toma o seguinte formato:

$$\begin{aligned}
 & \text{(XLI)} \\
 P(Y = 1) = & \frac{e^{-41.21+9.65*\text{SexoFeminino}+0.06*\text{Idade}-3.58*\text{TBSim}+0.6*\text{Peso}+1.63*\text{Tempo}+5.72*\text{APSSAlto}-0.08*\text{SexoFeminino:Idade}-0.15*\text{SexoFeminino:Peso}-0.12*\text{Idade:APSSAlto}+3.88*\text{TBSim:APSSAlto}-0.02*\text{Peso:Tempo}}{1+e^{-41.21+9.65*\text{SexoFeminino}+0.06*\text{Idade}-3.58*\text{TBSim}+0.6*\text{Peso}+1.63*\text{Tempo}+5.72*\text{APSSAlto}-0.08*\text{SexoFeminino:Idade}-0.15*\text{SexoFeminino:Peso}-0.12*\text{Idade:APSSAlto}+3.88*\text{TBSim:APSSAlto}-0.02*\text{Peso:Tempo}}
 \end{aligned}$$

Nesta equação de regressão logística múltipla, a constante e , de acordo com Euler, L. (1707 – 1783), representa a base do logaritmo natural, que é aproximadamente igual a 2.71828.

4.2.6.2 Considerando apenas as variáveis significativas na equação final

Nesta pesquisa, objetivamente, o interesse final é construir um modelo mais **parcimonioso**, ou seja, escolher apenas as variáveis que têm uma associação estatisticamente significativa com a variável dependente. O critério mais comum para determinar a significância é o **p-valor**, e normalmente se usa um limiar de 0,05, o que significa que as variáveis que têm p-valor menor do que esse são mantidas no modelo, enquanto as que têm p-valor maior são removidas.

O Modelo reduzido proposto pelo pesquisador (somente variáveis significativas) assenta no facto de favorecer um modelo mais simples e interpretável, eliminando variáveis com p-valor maior que um limiar pré-estabelecido.

Esta preferência do pesquisador tem sustento com base no livro de Hosmer & Lemeshow (2000), pois os autores discutem a construção de modelos de regressão logística e a escolha das variáveis predictoras. No seu livro, “*Applied Logistic Regression*” abordam a importância de realizar a seleção de variáveis de maneira cuidadosa, considerando testes de significância (como o p-valor) e o risco de incluir variáveis irrelevantes (o que pode resultar em modelos mais complexos e com menor capacidade de generalização).

Esta abordagem do modelo reduzido, também é sustentada pelo Agresti, A. (2018) uma vez que ele discute como construir modelos de regressão logística, incluindo a abordagem de seleção de variáveis, e a escolha de manter apenas aquelas que são significativas em termos de p-valor.

Dentre vários autores que defendem a consideração do modelo reduzido, há convergência das conclusões relativas as suas vantagens como é o caso de: **Simplicidade e Interpretabilidade, Redução da multicolinearidade, Prevenção do overfitting, Eficiência computacional, Teoria e relevância prática.**

Diante dos pressupostos apresentados acima, o pesquisador **adoptou o uso do modelo reduzido**, por ser o mais confiável e por ele proporcionar um melhor equilíbrio entre a simplicidade, precisão e generalização, alinhando com as melhores práticas de análise estatística.

Assim sendo, a **equação final do modelo reduzido**, substituindo os parâmetros pelas respectivas estimativas encontradas pelo método de máxima verossimilhança é:

(XLII)

$$P(Y = 1) = \frac{e^{-41.21-3.58*TBSim+0.60*Peso+1.63*Tempo+3.88*TBSim:APSSAlto-0.02*Peso:Tempo}}{1 + e^{-41.21-3.58*TBSim+0.60*Peso+1.63*Tempo+3.88*TBSim:APSSAlto-0.02*Peso:Tempo}}$$

De forma conclusiva, esta é a **equação preditiva definitiva** produzida pelo autor, baseada na análise dos dados e nos critérios estatísticos aplicados, que oferece a melhor estimativa para o cálculo da probabilidade de alcance da supressa viral em pacientes adultos vivendo com o HIV/SIDA no Centro de Saúde da Cidade de Lichinga.

4.2.7 Análise do ajuste geral do modelo

A Tabela 4.4 apresenta as diferentes medidas de Pseudo R^2 , que são usadas para avaliar o ajuste geral do modelo logístico. Essas medidas são análogas ao R^2 nos modelos de regressão linear, mas adaptadas para modelos de regressão logística. Cada uma das medidas fornecidas tem uma interpretação ligeiramente diferente.

Tabela 4. 4: Pseudo R^2 para análise do ajuste geral do modelo

Cox-Snell	Nagelkerke	McFadden	Tjur	sqPearson
0.682	0.926	0.859	0.889	0.889

O valor de 0.682 indica que o modelo explica aproximadamente 68.2% da variação na variável dependente. Esta medida é limitada a valores menores que 1, portanto, mesmo um valor relativamente alto como 0.682 pode indicar um bom ajuste do modelo. O Nagelkerke R^2 é uma versão ajustada do Cox-Snell R^2 , recalibrada para variar entre 0 e 1. Um valor de 0.926 sugere que o modelo explica 92.6% da variação na variável dependente, indicando um excelente ajuste do modelo.

O McFadden R^2 é baseado na razão de verossimilhança do modelo. Um valor de 0.859 é considerado muito bom, indicando que o modelo tem um excelente poder preditivo e explica bem a variância observada nos dados. O Tjur R^2 , também conhecido como coeficiente de discriminação, é uma medida simples que representa a diferença média entre as probabilidades preditas para os casos e os controlos. Um valor de 0.889 indica uma excelente discriminação do modelo entre os dois grupos, sugerindo um ajuste muito bom. O R^2 de Pearson é baseado na

correlação entre os valores observados e preditos. Um valor de 0.889 indica uma correlação muito forte, sugerindo que 88.9% da variabilidade na variável dependente pode ser explicada pelo modelo, indicando um ajuste excelente.

4.2.8 Análise de ajustamento do modelo

O valor do qui-quadrado (1.299) com 8 graus de liberdade é uma medida da discrepância entre as frequências observadas e esperadas. O p-valor (0.9956) é maior do que o nível de significância comumente utilizado (0.05). Isso indica que não há evidência estatisticamente significativa para rejeitar a hipótese nula de que o modelo ajusta bem os dados. Não há diferença significativa entre as frequências observadas e esperadas, sugerindo que o modelo se ajusta bem aos dados.

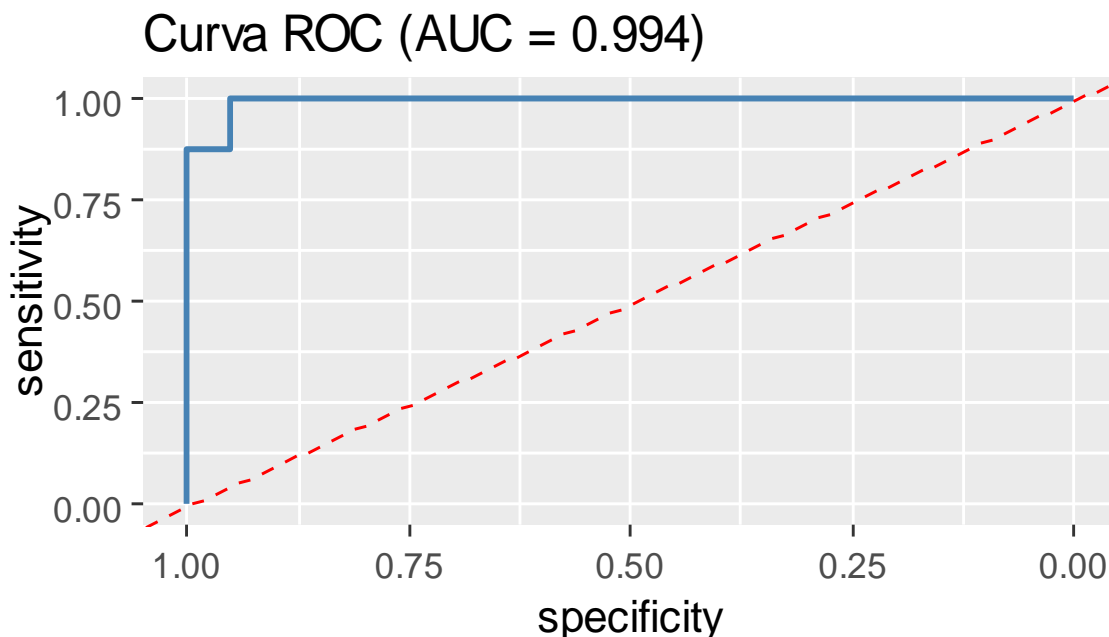
Tabela 4.5: Teste Hosmer e Lemeshow para análise de ajustamento do modelo

Chi-quadrado	Grau de liberdade	P-valor
1.299	8	0.9956

4.3 Avaliação do desempenho do modelo

A curva ROC (**R**eceiver **O**perating **C**haracteristic) é uma ferramenta gráfica usada para avaliar o desempenho de um modelo de classificação binária. A Área Sob a Curva (AUC - Area Under Curve) é um valor numérico que resume a performance da Curva ROC. O valor de AUC varia de 0 a 1, onde 1 indica um modelo perfeito e 0.5 indica um modelo que não tem poder de discriminação (equivalente a uma classificação aleatória).

Gráfico 4.3: Análise do modelo pela Curva de ROC



Com base no gráfico acima, verifica-se que o valor de (AUC) é de 0.994, o que significa que em média há uma probabilidade de 99.4% de que o modelo classifica corretamente uma observação positiva como mais provável de ser positiva do que uma observação negativa. A curva ROC está muito próxima do canto superior esquerdo do gráfico. Isso implica uma alta sensibilidade e especificidade, ou seja, o modelo tem uma alta taxa de verdadeiros positivos e uma baixa taxa de falsos positivos. O valor alto do AUC sugere que o modelo é robusto e confiável para a classificação dos dados. A probabilidade de erro de classificação é baixa.

4.4 Teste do funcionamento da equação preditiva do modelo reduzido

De modo a testar a robustez e funcionamento do modelo, serão colocados alguns exercícios práticos como teste.

Para este ensaio, foi escolhido o paciente na ordem número 76 do ficheiro de dados em Excel.

Este paciente tem os seguintes dados:

TB	Peso	Tempo de Tratamento	APSSAlto
1	62kg	8 meses	1

Tendo em conta a seguinte fórmula do autor:

(XLIII)

$$P(Y = 1) = \frac{e^{-41.21-3.58*TB\text{Sim}+0.60*Peso+1.63*Tempo+3.88*TB\text{Sim}:APSS\text{Alto}-0.02*Peso:Tempo}}{1 + e^{-41.21-3.58*TB\text{Sim}+0.60*Peso+1.63*Tempo+3.88*TB\text{Sim}:APSS\text{Alto}-0.02*Peso:Tempo}}$$

Primeiro serão calculadas as interações de:

$$TB\text{Sim}: APSS\text{Alto} = 1 * 1 = 1$$

$$Peso: Tempo = 62 * 8 = 496$$

Substituindo os fatores na fórmula, teremos:

$$P(Y = 1) = \frac{2.71828^{-41.21-3.58*1+0.60*62+1.63*8+3.88*1-0.02*496}}{1 + 2.71828^{-41.21-3.58*1+0.60*62+1.63*8+3.88*1-0.02*496}}$$

$$P(Y = 1) = \frac{2.71828^{-0.59}}{1 + 2.71828^{-0.59}} = 0.3566$$

Calculada a probabilidade deste paciente alcançar a supressão viral, considerando que tem 62kg de peso, tempo de tratamento de 8 meses, tem tuberculose e teve acesso ao nível alto de apoio psicossocial, tem a chance estimada $P(Y = 1)$ de **35,66%**.

Teste número 2

Feito o cálculo da primeira situação do paciente, vamos imaginar que o mesmo paciente continuou em tratamento por mais seis meses, passando dos 8 para 14 meses e teve um ganho de mais 1kg de peso, passando dos 62 para 63kg, teremos os seguintes dados:

TB	Peso	Tempo de Tratamento	APSSAlto
1	63kg	14 meses	1

Primeiro serão calculadas as interações de:

$$TB\text{Sim}: APSS\text{Alto} = 1 * 1 = 1$$

$$Peso: Tempo = 63 * 14 = 882$$

Substituindo os fatores na fórmula, teremos:

$$P(Y = 1) = \frac{2.71828^{-41.21-3.58*1+0.60*63+1.63*8+3.88*1-0.02*882}}{1 + 2.71828^{-41.21-3.58*1+0.60*63+1.63*8+3.88*1-0.02*882}}$$

$$P(Y = 1) = \frac{2.71828^{2.07}}{1 + 2.71828^{2.07}} = \mathbf{0.8879}$$

O resultado da segunda simulação mostra que o mesmo paciente, passados mais 6 meses de tratamento antirretroviral, a sua probabilidade de alcançar a supressão viral passou dos 35,66% para cerca de **88.79%**, significando com isso que a equação geral desta pesquisa é um excelente preditor da chance de alcance da supressão viral aos pacientes adultos em tratamento antirretroviral no Centro de Saúde da Cidade de Lichinga.

Capítulo V

5. Conclusões e Sugestões

5.1 Discussão

O presente estudo teve como objetivo analisar a relação entre a supressão viral e as medidas clínicas de atendimento aos pacientes adultos vivendo com HIV/SIDA no Centro de Saúde da Cidade de Lichinga, entre os anos de 2020 e 2021. A análise baseou-se numa série de testes estatísticos e modelos para identificar fatores associados à supressão viral e avaliar o desempenho desses modelos, tendo em conta as variáveis independentes: Sexo, Idade, Peso, Tempo de tratamento, Tuberculose e Apoio Psicossocial. Com base nos resultados obtidos, podemos determinar o seguinte:

- ✓ Utilizando o teste do qui-quadrado, os fatores como sexo, faixa etária e APSS não apresentaram associações estatisticamente significativas com a supressão viral ($p > 0.05$). Isso sugere que, enquanto a tuberculose pode desempenhar um papel crucial na supressão viral, outros fatores demográficos e psicossociais não mostraram uma relação direta e significativa no contexto deste estudo. Entretanto, apesar de que tais fatores não tenham sido significativos conforme a estatística deste teste, no contexto de saúde pública é importante considera-los fatores porque geralmente pessoas mais velhas (especialmente de terceira idade) requerem maior atenção nos cuidados de saúde dada a debilidade imunológica desgastada com o tempo de vida. A eficácia de tratamento é geralmente mais lenta em pessoas de idade avançada do que na idade jovem e adulta, sobre tudo em grupos populacionais desfavorecidas economicamente. Outro dado importante no contexto de saúde pública moçambicano, é oferecer maior atenção e prestação de cuidados de saúde diferenciados para os homens, uma vez que grande parte de programas e serviços de saúde nas unidades sanitárias, estão dedicados a mulheres e crianças
- ✓ Relativamente a TB, verificou-se que a supressão viral está significativamente associada à presença de tuberculose ($p < 0.000$). Sabe-se no entanto que a coinfeção por TB e HIV pode trazer várias complicações aos pacientes para que possam alcançar a supressão

viral devido a razões relacionadas com a interação de vários medicamentos, podendo criar certos efeitos adversos.

Estudos científicos como é o caso de Lawn, S.D., et al. (2006), concluiu se que pacientes com coinfeção por TB e HIV tinham menor probabilidade de alcançar a supressão viral em comparação aos que somente tinham HIV.

O mesmo resultado foi encontrado por Manosuthi, W., et al. (2006) em seu estudo concluiu que pacientes com coinfeção TB/HIV apresentavam uma resposta virológica mais lenta ao TARV, devido a interações medicamentosas e complicações associadas ao tratamento de TB. Por fim, evidências publicadas em 2008 no artigo *The Lancet Infectious Diseases* também concluiu que o tratamento simultâneo da TB e HIV é desafiador contudo, a integração de serviços de tratamento pode melhorar os resultados de supressão viral.

Com esses fundamentos, conclui se que um individuo com HIV, em princípio apresenta uma imunidade baixa e suscetível a qualquer patologia, portanto, a presença da TB cria um impacto negativo na evolução da doença uma vez que o paciente terá dois pesos no mesmo organismo. Clinicamente, recomenda-se que na presença da TB há que dar prioridade ao tratamento desta doença, de modo que mais tarde, a administração do TARV possa ter o efeito desejado.

- ✓ Para além do uso do teste de qui-quadrado, foram realizados outros cruzamentos em conformidade ao modelo de regressão logística múltipla, ao qual evidenciou no Modelo Final que o **peso do paciente** tem uma significância estatística no alcance da supressão viral.

Conforme provado por Gomes e Lourival (2016), que conhecer o estado nutricional do paciente com HIV é um dos primeiros passos para orientar uma nutrição adequada e obter um ganho de peso ideal e garantir que o tratamento seja bem-sucedido no paciente.

Por outro lado, Duran AC (2009), refere que a existência da relação entre AIDS e nutrição é evidente de tal forma que em presença de deficiências nutricionais que

influenciam no peso do paciente, pode implicar na eficácia do TARV e consequentemente, a capacidade de alcançar e manter a supressão viral.

No estudo provado por Koethe, J.R., et al. (2016), na sua meta-análise, provaram que a obesidade (excesso de peso) pode estar associada a menor probabilidade de alcançar a supressão viral, possivelmente devido a complicações metabólicas e inflamatórias associadas.

Com essas evidências, surge a necessidade de rastreio regular das comorbidades aos pacientes com vista a identificar aqueles com baixo peso e outros com sobre peso ou obesidade de modo a orientar ou aconselhar melhor sobre boas práticas nutricionais, dietética e estilo de vida. Ademais, a monitoria regular do estado nutricional e a carga viral dos pacientes pode ajudar na identificação de problemas de adesão e eficácia do TARV para posterior intervenção.

- ✓ No Modelo Final, foi também significativa a variável **tempo de tratamento** como um elemento que tem uma influência forte na supressão viral. Na pesquisa de Deeks, S.G., et al. (2015) refere que a maioria dos pacientes que aderem ao tratamento consistentemente ao TARV, alcançam a supressão viral no intervalo de 6 a 12 meses de tratamento. Para além deste autor, Gomilia. E. S, et al (2023) concluiu que pacientes que aderem regularmente ao tratamento, o tempo estimado para alcançar a supressão viral tem uma mediana de 6 meses.

Em outro estudo longitudinal realizado por Bangsberg, D.R., et al, (2001), os resultados encontrados mostraram que a adesão ao TARV é o fator mais importante para alcançar a supressão viral ou seja, pacientes com alta adesão (>95%) tiveram uma taxa significativamente maior de supressão viral.

Em torno dessas abordagens de outros autores, a ilação que se pode tirar é que a manutenção no programa de cuidados e tratamento do HIV com seguimento rigoroso da terapia antirretroviral são fatores determinantes para alcançar e manter a supressão viral. Por outro lado, o início do tratamento logo após o diagnóstico da doença associado a

avaliação de possíveis riscos de resistência a medicamentos, são fatores fundamentais para o sucesso do tratamento.

- ✓ Ainda sobre os resultados de significância no Modelo Final, a interação entre a **tuberculose e o apoio psicossocial** sugeriu ser de grande relevância na medida em que as chances de supressão viral dum paciente com TB, aumentam na medida em que vai tendo apoio psicossocial para melhoria da adesão e manutenção no programa de cuidados e tratamento do HIV. Esta conclusão neste estudo coincide com os resultados encontrados pelo Kipp, W., et al (2011), segundo o qual as intervenções de apoio psicossocial em pacientes co-infetados com TB e HIV, ajudaram a melhorar de forma significativa na adesão e os seus resultados de supressão viral foram positivos.

Um outro estudo realizado por Pachi, A., et al (2013), confirmou o impacto do apoio psicossocial na adesão ao tratamento da TB e conseqüentemente, melhorias nos resultados de supressão viral.

Considerando que a TB é uma doença que afeta o sistema imunológico do paciente, é preciso dar atenção ao seu tratamento para eficácia do TARV, tendo em conta que a TB tem cura. Entretanto, o seu processo de tratamento requer um acompanhamento e apoio psicossocial, reduzir o estigma e isolamento para permitir o sucesso do tratamento sem resistência ou complicações relacionadas a medicação. Considerando essas variantes, o envolvimento das equipas multidisciplinares de saúde como enfermeiros, médicos, psicólogos e demais carreiras, pode ajudar na melhoria da prestação de cuidados ao paciente.

- ✓ O último resultado que demonstrou significância nos testes realizados no Modelo Final tem a ver com a interação entre o **peso e o tempo de tratamento**. Tal como referido anteriormente na discussão individualizada de cada uma destas duas variáveis interativas, a sua influência no alcance da supressão viral é alta.

Uma pesquisa realizada por Lake, J.E., et al. (2017) referente ao impacto do Peso e Tempo de tratamento na resposta ao TARV, mostrou resultados diferentes, ao ter concluído que o ganho de peso durante o tempo de tratamento pode estar associado a

uma resposta virológica menos eficaz, particularmente em pacientes com longa duração de tratamento.

Num outro estudo de Meta-análise sobre o peso, tempo de tratamento e supressão viral realizado por Schouten, J., et al, (2014) mostrou que o Índice de Massa Corporal (IMC) inicial quanto as mudanças no peso durante o tempo de tratamento, eram fatores preditivos para a supressão viral. Nisto tudo, a duração do tratamento teve um papel modulador nesses efeitos, com o tempo de tratamento prolongado associado a melhores resultados em pacientes com peso estável.

A conclusão que se pode ter é que é crucial o controlo da evolução da doença dos pacientes de modo a permitir possíveis ajustamentos da dosagem do medicamento e assegurar sua eficácia. Por outro lado, a monitoria do peso dos pacientes, especificamente com sobre peso e obesidade, precisam dum seguimento mais intensivo com vista a recomendar melhor sobre práticas nutricionais e estilos de vida saudáveis, cujo fim último é o alcance e manutenção da supressão viral.

Assim sendo, a análise do Critério de Informação de Akaike (AIC) ajudou a selecionar o melhor modelo logístico binário. Observou-se uma diminuição contínua nos valores de AIC à medida que mais variáveis foram incluídas entre o modelo 1 até modelo 11, com o menor valor de AIC (76.67) indicando o melhor modelo. Este processo reforça a importância de selecionar um modelo que balanceie a complexidade e o ajuste, evitando o *overfitting*⁵.

Face a isso, o autor determinou a seguinte equação preditiva do modelo final: (XLIV)

$$\text{logit}(P(Y = 1)) = -41.21 - 3.58 * TBSim + 0.60 * Peso + 1.63 * Tempo + 3.88 * TBSim: APSSAlto - 0.02 * Peso: Tempo$$

Esta é a equação preditora final para calcular a probabilidade de alcance da supressão viral dos pacientes adultos em atendimento no Centro de Saúde da Cidade de Lichinga.

⁵ O *overfitting* ocorre quando um modelo se ajusta excessivamente aos dados de treinamento, perdendo a capacidade de generalizar para novos dados.

Em relação aos valores de Pseudo R^2 na Tabela 4.4, indicam que o modelo possui um excelente ajuste geral, especialmente com valores elevados em Nagelkerke (0.926) e McFadden (0.859). Isso demonstra que o modelo explica uma grande parte da variabilidade nos dados. O teste de Hosmer e Lemeshow (Tabela 4.5) confirmou que o modelo ajusta bem os dados observados ($p = 0.9956$), indicando que não há evidência significativa de falta de ajuste ou seja, **não há evidência estatisticamente significativa para rejeitar a hipótese nula de que o modelo ajusta bem os dados.**

5.2 Conclusão

De acordo com a WHO, o HIV é um problema de saúde pública global ainda não ultrapassado, responsável por cerca de 42.3 milhões de vidas perdidas e a sua transmissão continua existir em todos os países do mundo. Foi estimado que, no final de 2023, havia cerca de 39.9 milhões de pessoas a viver com o HIV, das quais 65% se encontravam na Região Africana da OMS. Em 2023 estimou-se que 630,000 pessoas morreram por causas relacionadas com HIV e que 1.3 milhões de pessoas adquiriram o HIV. Não existindo até a data a cura para a doença, esta tornou-se uma condição crónica de saúde capaz de gerir e para a redução da sua ocorrência tem vindo a apostar se fortemente em medidas de prevenção.

Este estudo revelou importantes resultados sobre os fatores que influenciam a supressão viral em pacientes adultos vivendo com HIV/SIDA, como uma das medidas de contenção da propagação do vírus, por via sexual.

A presença de tuberculose mostrou ser um fator crítico associado à supressão viral. Outros fatores como tempo de tratamento, peso e apoio psicossocial também desempenham papéis significativos. Especificamente, descobriu-se que a presença da TB foi identificada como fator de risco, complicando o manejo da terapia e potencialmente reduzindo a probabilidade de alcançar a supressão viral.

Por outro lado, o tempo de tratamento prolongando demonstrou ser fundamental para alcançar níveis indetectáveis de carga viral e desta feita, assegurar a melhor qualidade de vida do paciente.

Uma outra descoberta, é relativa ao peso do paciente pois, desempenha um papel crucial, porque um paciente com peso inferior ao recomendado, a medicação pode trazer várias complicações de saúde, daí que há necessidade de monitoria regular e recomendar boas práticas nutricionais e estilos de vida saudáveis.

Finalmente, o apoio psicossocial mostrou-se vital para adesão ao tratamento e manejo de estresse e estigma, facilitando assim a supressão viral.

A robustez do modelo logístico ajustado, evidenciada pelos altos valores de Pseudo R² e pelo adequado ajuste no teste de Hosmer e Lemeshow, reforça a validade dos achados.

As implicações práticas dos achados neste estudo são vastas. Políticas de saúde e práticas clínicas devem considerar a integração de diagnóstico precoce e tratamento da tuberculose, especialmente dada a interação significativa entre a tuberculose e o apoio psicossocial na supressão viral assim como a interação entre o peso do paciente e o tempo em que se encontra em tratamento. Por outro lado, a monitoria regular do peso e ajuste adequado da dosagem dos medicamentos antirretrovirais é necessário para otimizar o tratamento. Para além destas intervenções, o apoio psicossocial deve ser considerado como parte essencial no programa de tratamento do HIV, pois melhora a saúde e bem estar do paciente.

Este estudo contribui significativamente para a compreensão dos determinantes da supressão viral em pacientes vivendo com HIV/SIDA, fornecendo uma base sólida para intervenções clínicas e políticas de saúde pública mais eficazes.

Resumidamente, o estudo contribuiu para entendimento dos fatores que influenciam a supressão viral em pacientes adultos vivendo com HIV no Centro de Saúde da Cidade de Lichinga entre os anos 2020 e 2021, destacando se o papel crucial da TB, Peso corporal, Tempo de tratamento e APSS. Esses *insights* são fundamentais para adoção de estratégias de tratamento mais eficazes e personalizadas, com objetivo de melhorar a qualidade de vida dos pacientes e alcançar melhores resultados de saúde e evitar novas infeções.

No contexto de combate da epidemia, a OMS, o Fundo Global e a UNAIDS têm estratégias globais sobre o HIV que estão alinhadas com o objetivo 3.3 dos Objetivos de Desenvolvimento Sustentável de acabar com a epidemia do HIV até 2030. Prevê-se que até 2025, 95% de todas as pessoas a viver com o HIV terão um diagnóstico, 95% das quais deverão estar a receber o TARV, e 95% das pessoas a viver com o HIV em tratamento deverão alcançar uma carga viral suprimida para benefício da saúde da pessoa e para a redução da transmissão do HIV. Contudo, globalmente em 2023, esses percentuais foram de 86%, 89% e 93% respetivamente.

5.3 Sugestões

Considerando os resultados encontrados neste estudo, o pesquisador sugere a realização de outras pesquisas em outros grupos populacionais da Província do Niassa e sempre que possível, explorar a interação de outros fatores, comorbidades adicionais assim como noutra grupo etário (crianças). É no entender também do pesquisador que estudos longitudinais com amostras maiores também poderiam ser úteis para confirmar estes principais achados do presente trabalho e explorar profundamente as interações encontradas. Igualmente o pesquisador sugere que para trabalhos futuros, a análise possa ser feita segregando os grupos etários com vista a verificar se o efeito identificado num único conjunto, ele se mantém quando a análise é também feita em grupos etários e/ou separados por sexo.

5.4 Limitações do estudo

Algumas limitações do estudo devem ser consideradas na interpretação dos resultados deste estudo.

Todos os dados analisados nesta pesquisa são secundários provenientes duma base de dados eletrónica que faz a gestão de dados individuais e todo o histórico de tratamento antirretroviral dos pacientes com HIV. Por esse motivo, a ausência de alguma informação do paciente, com destaque para falta de colheita ou digitação dos resultados da carga viral na base de dados, fez com que o pesquisador excluísse um número significativo de pacientes para estudo.

Por outro lado, apesar de que os resultados da pesquisa têm uma contribuição significativa, é importante referir que esta pesquisa se limita apenas a uma unidade sanitária, podendo desta forma a amostra não ser representativa para toda a população adulta vivendo com HIV na província do Niassa e, variáveis não estudadas pelo pesquisador podem ter outra influência.

Capítulo VI

6. Referências Bibliográficas

1. Acharya S, Palkar A, Sayed AP, Setia MS.(2021).Retrospective cohort analysis of survival of children living with HIV/AIDS in Mumbai, India. *BMJ Open*. 2021 Sep 13;11(9):e050534. doi: 10.1136/bmjopen-2021-050534.
2. Agresti, A. e Finlay, B. (2009). *Statistical Methods for the Social Sciences*. 4th ed. Prentice Hall.
3. Araújo, P. (2017). *Regressão Logística Multinível: um estudo sobre o consumo de álcool e inalantes entre crianças e adolescentes*. Brasília, Brasil.
4. Bangsberg, D.R., et al. (2006). Adherence to antiretroviral therapy in sub-Saharan Africa and North America: A meta-analysis. *JAMA*, 296(6), 679-690. <https://doi.org/10.1001/jama.296.6.679>
5. Belfiore, P. (2015). *Estatística aplicada a administração, contabilidade e economia com excel em SPSS. Exemplo de aplicação ao Resolving Credit*. Porto. Vida Económica.
6. Box, G. E. P., Hunter, J. S., & Hunter, W. G.(2005). *Statistics for experimenters: Design, innovation, and discovery: (2nd ed.)*
7. Cabral, C. S. (2013). *Aplicação do Modelo de Regressão Logística num Estudo de Mercado*. Universidade de Lisboa. Dissertação de mestrado em Matemática aplicada à Economia e à Gestão.
8. Cardona-Arias J, Higueta-Gutiérrez L. Impact of HIV/AIDS on quality of life: meta-analysis 2002–2012; 2014.
9. Chein, Flávia *Introdução aos modelos de regressão linear: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas* (2019). Enap Brasília.
10. Chicanequisso, et al., Impact of active tuberculosis on viral suppression in patients with HIV at the Military Hospital of Maputo, *PLOS Global Public Health*, 2024

11. Corrar, S.R., Natário, I. e Martins, M.R. (2007). Modelos de regressão logística com apoio computacional. IME-USP, São Paulo
12. Da W, Li X, Qiao S, Zhou Y, Shen Z. Antiretroviral therapy and mental health among people living with HIV/AIDS in China. *Psychol Health Med.* 2020;25:45–52. - PMC – PubMed
13. De Melo et al., (2021). Adesão ao tratamento antirretroviral e supressão viral entre pacientes do Hospital de Clínicas da Unicamp: buscando a meta UNAIDS 90-90-90.2021. *Research, Society and Development*, v. 10, n. 13, e395101321428.
14. Deeks, S. F., et al. The end of AIDS:HIV infection as a chronic disease, 2013, *The Lancet* , [https://doi.org/10.1016/S0140-6736\(13\)618](https://doi.org/10.1016/S0140-6736(13)618)
15. Dessie, Z.G., Zewotir, T., Mwambi, H. et al. Multivariate multilevel modeling of quality of life dynamics of HIV infected patients. *Health Qual Life Outcomes* 18, 80 (2020). <https://doi.org/10.1186/s12955-020-01330-2>
16. Dong et al.,(2022) Prediction Model for the Risk of HIV Infection among MSM in China: Validation and Stability. *Int. J. Environ. Res. Public Health*, 19, 1010. <https://doi.org/10.3390/ijerph19021010>
17. Drewes et al.,(2021). Associations Between Experienced and Internalized HIV Stigma, Adversarial Growth, and Health Outcomes in a Nationwide Sample of People Aging with HIV in Germany. *AIDS Behav* 25, 1037–1046 <https://doi.org/10.1007/s10461-020-03061-3>
18. Duran, A. C. da F. Leitão, qualidade da dieta dos adultos vivendo com HIV/AIDS e seus fatores associados 2009, *apud* World Bank, HIV/Aids, Nutrition, and Food Security: What we can do. A Synthesis of international guidance. AIDS: Washington DC: The World Bank, 2007
19. Enferm, A. P (2023) Tratamento como prevenção na perspectiva de pessoas vivendo com HIV/AIDS <https://doi.org/10.37689/acta-ape/2023AO012323>
20. Euler, L. (1748). *Introductio in analysin infinitorum*. Ex officina Kummelii.

21. Fisch, P. (2017) Qualificando a cascata do HIV: Diferentes tempos até a supressão viral a partir dos dados da coorte de Alvorada -Tese de Doutorado, Universidade Federal do Rio Grande do Sul, Porto Alegre, BR-RS, 28 - 29
22. Folha, Yoann Rodrigues (2021) Tendências e determinantes acerca do conhecimento do HIV/SIDA em Moçambique, Dissertação de Mestrado em Estatística, Universidade do Minho, Portugal
23. Fox, J. (2015). *Applied regression analysis and generalized linear models* (3rd ed.). Sage
24. Gomes, T. B., & Lourival, N. B. dos S. (2022). Perfil nutricional e pacientes HIV positivo do município de Apucarana. *Revista Brasileira de Nutrição Clínica*, 37(1), 15-24
25. Gomilia. E. S, et al. (2023). Rapid antiretroviral initiation at a community-based clinic in Jackson. *Journal of Community Health*, 48(2), 290-298. <https://aidsrestherapy.biomedcentral.com/articles/10.1186/s12981-020-00319-7>
26. Gonzale, L. A. (2018). Regressão Logística e suas aplicações. Universidade Federal de Maranhão. Dissertação de mestrado
27. Harrell, F. E. Jr. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis*. Springer
28. Hosmer, D. e Lemeshow, S. (1989). *Applied Logistic Regression*. New Work, John Wiley
29. Hosmer, J. e Lemeshow, S. (2000). *Applied Logistic Regression*. 2nd ed. New York: John Wiley e Sons
30. Kinuthia, Rachael N., Thigiti, Joseph M., & Gakinya, Benson N.. (2016). Relationship between HIV stage and psychomotor speed neurocognitive score at a Kenyan sub-county hospital. *African Journal of Primary Health Care & Family Medicine* , 8(1), 1-8. <https://dx.doi.org/10.4102/phcfm.v8i1.1061>
31. Kipp, W.,et al. (2009). Challenges in tuberculosis care in Western Uganda: Health care worker and patient perspectives. *International Journal of Tuberculosis and Lung Disease*, 13(8), 956-962

32. Koethe, J.R., et al. Obesity and Weight Gain in Persons with HIV Pathogenesis and Treatment, AL Landay and NS Uday, Section Editors, 2020, Volume 17, pages 138–150
33. Kutner, M., Nachtsheim, C., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). McGraw-Hill/Irwin
34. Lake, J.E., et al. (2017). Weight gain following initiation of antiretroviral therapy: Risk factors in randomized comparative clinical trials. *Journal of Acquired Immune Deficiency Syndromes*, 74(4), 435-444. <https://doi.org/10.1097/QAI.0000000000001311>
35. Li W, Fazeli PL, Wang G, Maqbool M, Del Bene VA, Triebel K, Martin KM, Vance D. (2023). Predictors of an Active Lifestyle in Middle-Aged and Older Adults with HIV in the United States Deep South. *HIV AIDS*. P.15:63-70. <https://doi.org/10.2147/HIV.S399842>
36. Lucambio, F. Seleção de modelos: Critério de Informação de Akaike, Novembro de 2020
37. Machel, J. (2010). O HIV/SIDA em Moçambique: Reflexões sobre a epidemia e o papel da mulher. Editora
38. Manosuthi, W., et al. (2006). Integrated therapy for HIV and tuberculosis. *AIDS Research and Therapy*, 3, 30. <https://doi.org/10.1186/1742-6405-3-30>
39. Maroco, J. (2007). *Análise estatística Com Utilização do SPSS*, 3ª Edição. Lisboa: Edição Silabo
40. McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press
41. Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, 54(1), 17–24
42. Mesquita, P.S.B. (2014). Um modelo de regressão logística para avaliação do nível de conhecimento em relação ao uso de métodos de planeamento familiar (Dissertação de Mestrado). Universidade Estadual do Norte Fluminense

43. Meyer, A.N (1980). Análise da qualidade de vida no trabalho utilizando um modelo de regressão logística. Dissertação de Mestrado em Engenharia de produção. Ponto Grossa, Brasil
44. Montgomery, D. C., Peck, E. A. (2006). Introduction to Linear Regression Analysis, 3rd Ed. Índia: Wiley India Pvt. Limited
45. Myers, J. P., Haas, J. M., & R., P. M. (2017). Gender differences in HIV outcomes among adults with HIV infection in the United States. *Journal of Acquired Immune Deficiency Syndromes*, 75(4), 457-465. <https://doi.org/10.1097/QAI.0000000000001476>
46. Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), 691–692
47. Nunn, P., et al. (2008). The impact of HIV on resource utilization by patients with tuberculosis in a tertiary referral hospital, Nairobi, Kenya. *International Journal of Tuberculosis and Lung Disease*, 12(2), 219-224
48. Oliveira, T.A. (2004). Estatística Aplicada. Edições Universidade Aberta.
49. Pachi, A., Bratis, D., Moussas, G., & Tselebis, A. (2013). Psychiatric morbidity and other factors affecting treatment adherence in pulmonary tuberculosis patients. *Tuberculosis Research and Treatment*, 2013, 489865. <https://doi.org/10.1155/2013/489865>
50. Qiao, S.; Li, X.; Stanton, B. Social support and HIV-related risk behaviors: A systematic review of the global literature. *AIDS Behav.* 2014, 18, 419–441.
51. Reis, Elizabeth; Melo, Paulo; Andrade, Rosa; e Calapez, Tereza (2016). Estatística Aplicada 2. Lisboa – Portugal. Edições Sílabo, Ltda
52. Reiss, P., Carr, A., Currier, J., et al. (2019). Impact of aging on HIV disease progression and antiretroviral therapy response. *Clinical Infectious Diseases*, 69(1), 91-98. <https://doi.org/10.1093/cid/ciy978>

53. Safren, S. A., et al (2009). A randomized controlled trial of cognitive behavioral therapy for adherence and depression (CBT-AD) in HIV-infected individuals. *Health Psychology*, 28 (1)<https://doi.org/10.1037/a0012715>
54. Schouten, J., et al. (2014). Current evidence on hospital antimicrobial stewardship objectives: A systematic review and meta-analysis. *Clinical Infectious Diseases*, 59(1), 51-60. <https://doi.org/10.1093/cid/ciu240>
55. Sherr, L., et al. (2011). HIV and depression – a systematic review of interventions. *AIDS Care*, 23(9), 1147-1162. <https://doi.org/10.1080/09540121.2011.554557>
56. Shriharsha, C. and Rentala, S. (2019) Quality of life among people living with HIV/AIDS and its predictors: A cross-sectional study at ART center, Bagalkot, Karnataka. *J Family Med Prim Care*. 2019 Mar; 8(3): 1011–1016. doi: 10.4103/jfmpe.jfmpe_411_18
57. Sousa, E.C. (2006). Análise da influência local no modelo de regressão logística. Universidade de São Paulo. Dissertação de Mestrado em Estatística e Experiência Agronômica
58. Teixeira, A. B. B. (2021). Métodos de amostragem para atingir populações difíceis: avaliando e comparando diferentes métodos estatísticos com aplicação a um Estudo de prevalência do VIH, 2021, Dissertação de Doutor em Saúde Internacional
59. Tjur, T. (2009). Coefficients of determination in logistic regression models-A new proposal: The coefficient of discrimination." *The American Statistician*, 63(4), 366–372.
60. Wu AW (2000). Quality of life assessment comes of age in the era of highly active antiretroviral therapy. *Aids*.14:1449–51
61. Xi Xia et al., (2013) "HIV drug resistance prediction using multiple regression," IEEE 3rd International Conference on Computational Advances in Bio and medical Sciences (ICCABS), New Orleans, LA, USA, 2013, pp. 1-2, doi: 10.1109/ICCABS.2013.6629203

WEB-Referências

1. Coligação Global para a Prevenção do HIV, implementação do Roteiro de Prevenção do HIV, 2020, quarto relatório intercalar, Novembro de 2020, https://www.unaids.org/sites/default/files/media/documents/fourth-annual-progress-report-global-hiv-prevention-coalition_pt.pdf
2. Ending AIDS: progress towards the 90–90–90 targets. Geneva: UNAIDS; 2016.
3. http://erepository.uonbi.ac.ke/bitstream/handle/11295/93597/Ali_Logistic_Regression_to_Determine_theRelationship_Between_Hiv_Testing,_Hiv_Knowledge_and_Attitude_Among_Adults_in_Kenya.pdf?sequence=3
4. https://repositorio.enap.gov.br/bitstream/1/4788/1/Livro_Regress%C3%A3o%20Linear.pdf
5. <https://ubrisa.ub.bw/handle/10311/1498>
6. <https://www.datacamp.com/tutorial/multiple-linear-regression-r-tutorial?>
7. https://www.estrategiaods.org.br/wp-content/uploads/2015/09/ODS_oficial.pdf
8. [https://www.gnpplus.net/resources/positive-health-dignity-and-prevention-operational-guidelines/.](https://www.gnpplus.net/resources/positive-health-dignity-and-prevention-operational-guidelines/)
9. <https://www.msmanuals.com/pt-pt/casa/infec%C3%A7%C3%B5es/infec%C3%A7%C3%A3o-pelo-v%C3%ADrus-da-imunodefici%C3%A4ncia-humana-hiv/infec%C3%A7%C3%A3o-pelo-v%C3%ADrus-da-imunodefici%C3%A4ncia-humana-hiv>
10. <https://www.scielo.br/j/ape/a/JNfJPWLRyB6ZwWYtFyJXg4c/>
11. https://www.who.int/news-room/fact-sheets/detail/hiv-aids?gad_source=1&gclid=Cj0KCQjwn9y1BhC2ARIsAG5IY-

6w6gSuf3Z3FnNMOClwqnuTSw4Q4fsrWs9CLbMWupkJx065dOs40IIaAs2pEALw
wcB

12. Ministério da Saúde (2020), Relatório final da Avaliação do Plano de Aceleração da Resposta ao HIV e SIDA em Moçambique, 2013 – 2017. <https://www.misau.gov.mz/index.php/planos-estrategicos-do-hiv?download=528:avaliacao-do-plano-de-aceleracao-da-resposta-ao-hiv-e-sida-em-mocambique-2013-2017-relatorio-final>
13. Ministério da Saúde de Moçambique. (2020). Relatório do Inquérito Nacional de Prevalência, Riscos Comportamentais e Informação sobre o HIV e Sífilis em Moçambique. Disponível em: https://www.medbox.org/document/inquerito-nacional-de-prevalencia-de-hiv-e-sifilis-2015_mozambique-pdf-5/download
14. Multiple Linear Regression in R: Tutorial With Examples. DATA CAMP.
15. PNUD, Objetivos do Desenvolvimento Sustentável, 2030
16. UNAIDS. (2020). Country Factsheets: Mozambique. Disponível em: <https://www.unaids.org/en/regionscountries/countries/mozambique>.
17. https://unaids.org.br/wp-content/uploads/2021/06/2020_11_19_UNAIDS_FactSheet_PORT_Revisada-Final.pdf

7. Apêndices

Output do R

```
R version 4.4.0 (2024-04-24 ucrt) -- "Puppy Cup"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

> setwd("C:/Users/inaci/Documents/Dissertacao/Final")
> library("rstatix")

Attaching package: `rstatix`

The following object is masked from `package:stats`:

  filter

Warning message:
package `rstatix` was built under R version 4.4.1
> library("lmtest")
Loading required package: zoo

Attaching package: `zoo`

The following objects are masked from `package:base`:

  as.Date, as.Date.numeric

Warning message:
package `lmtest` was built under R version 4.4.1
> library("ggpubr")
Loading required package: ggplot2
Warning message:
package `ggpubr` was built under R version 4.4.1
> library("readxl")
```

```
Warning message:
package 'readxl' was built under R version 4.4.2
> library("car")
Loading required package: carData
Warning message:
package 'car' was built under R version 4.4.1
> library("DescTools")

Attaching package: 'DescTools'

The following object is masked from 'package:car':

  Recode

> library("gtsummary")
> library("psych")

Attaching package: 'psych'

The following objects are masked from 'package:DescTools':

  AUC, ICC, SD

The following object is masked from 'package:car':

  logit

The following objects are masked from 'package:ggplot2':

  %+%, alpha

Warning message:
package 'psych' was built under R version 4.4.1
> library("QuantPsyc")
Loading required package: boot

Attaching package: 'boot'

The following object is masked from 'package:psych':

  logit

The following object is masked from 'package:car':

  logit

Loading required package: dplyr

Attaching package: 'dplyr'

The following object is masked from 'package:car':

  recode
```

```

The following objects are masked from `package:stats`:

  filter, lag

The following objects are masked from `package:base`:

  intersect, setdiff, setequal, union

Loading required package: purrr
Attaching package: `purrr`

The following object is masked from `package:car`:

  some

Loading required package: MASS
Attaching package: `MASS`

The following object is masked from `package:dplyr`:

  select

The following object is masked from `package:gtsummary`:

  select

The following object is masked from `package:rstatix`:

  select

Attaching package: `QuantPsyc`

The following objects are masked from `package:DescTools`:

  Kurt, Skew

The following object is masked from `package:base`:

  norm

Warning message:
package `QuantPsyc` was built under R version 4.4.1
> library("dplyr")
> library("scatterplot3d")
> library("ggplot2")
> library("pacman")
> library("devtools")
Loading required package: usethis
> library("bptest")
Error in library("bptest") : there is no package called `bptest`
> #Import data set

```

```

> library(readxl)
> Dados <- read_excel("C:/Users/inaci/Documents/Dissertacao/Final/Dados.xlsx")
> #Calcular as frequencias da Supressao viral
> freq <- table(Dados$CV_class)
> #Calcular as frequencias absolutas da Faixa Etária
> Total_Faixa <- table(Dados$Faixa_etaria)
> #Exibir o gráfico de frequencias das Faixas
> barplot(Total_Faixa, main="Distribuição de Faixas Etárias",
+         xlab="Faixa Etária", ylab="Total", col="blue")
> hist(Total_Faixa, main="Distribuição de Faixas Etárias",
+       xlab="Faixa Etária", ylab="Total", col="blue")
> # Dados das frequências absolutas
> Total_Faixa <- c(`15-24` = 21, `25-34` = 78, `35-44` = 84, `45-54` = 71,
+                 `55-64` = 58,
+                 `65 ou mais` = 30)
> # Criar um vetor de nomes para as faixas etárias
> faixas <- names(Total_Faixa)
> # Gerar o histograma (gráfico de barras neste caso)
> barplot(Total_Faixa,
+         main = "Distribuição de Faixas Etárias",
+         xlab = "Faixa Etária",
+         ylab = "Frequência",
+         col = "blue",
+         names.arg = faixas,
+         las = 2) # Rotacionar os rótulos do eixo x
> # Dados das frequências absolutas
> percent_faixa <- c(`15-24` = 6.1, `25-34` = 22.8, `35-44` = 24.6, `45-54`
+                  = 20.8,
+                  `55-64` = 17.0, `65 ou mais` = 8.8)
> # Gerar o gráfico de barras com menos espaço entre as barras
> barplot_heights <- barplot(percent_faixa,
+                             xlab = "faixa etária",
+                             ylab = "percentagem",
+                             col = "#72A9D8", #Um azul com um pouco mais de
intensidade, mas ainda claro e agradável
+                             names.arg = faixas,
+                             las = 2, # Rotacionar os rótulos do eixo X
+                             space = 0, # Ajustar o espaço entre as barras
+                             ylim = c(0, 30),# Ajuste o limite superior do
eixo Y
+                             family = "Times New Roman",# Fonte definida
+                             cex.lab = 1.5, # Aumentar o tamanho da fonte
dos rótulos dos eixos
+                             cex.axis = 1.2) # Aumentar o tamanho da fonte
dos números no eixo,
There were 15 warnings (use warnings() to see them)
> freq_rel <- prop.table(freq)
There were 30 warnings (use warnings() to see them)
> print(freq_rel)
      0      1
0.377193 0.622807
> barplot(frequencias_supressao, main="Frequências de Supressão Viral",

```

```

+       xlab="Supressão Viral", ylab="Frequência")
Error: object 'frequencias_supressao' not found
> barplot(freq_rel, main="Frequências de Supressão Viral",
+       xlab="Supressão Viral", ylab="Frequência")
> barplot(freq, main = "Proporção", xlab = "Supressão", ylab = "frequencias
",
+       col="purple")
> #Gráfico circular
> pie(freq, main = "Distribuição de Supressão Viral", labels = c ("Não", "
Sim"),
+       col= c("red", "green"), int.angle=90, clockwise = TRUE)
Warning messages:
1: In text.default(1.1 * P$x, 1.1 * P$y, labels[i], xpd = TRUE, adj = ifels
e(P$x < :
  "int.angle" is not a graphical parameter
2: In text.default(1.1 * P$x, 1.1 * P$y, labels[i], xpd = TRUE, adj = ifels
e(P$x < :
  "int.angle" is not a graphical parameter
3: In title(main = main, ...) : "int.angle" is not a graphical parameter
> #Divisão de base de dados (Treinamento vs Teste)
Warning messages:
1: In doTryCatch(return(expr), name, parentenv, handler) :
  "int.angle" is not a graphical parameter
2: In doTryCatch(return(expr), name, parentenv, handler) :
  "int.angle" is not a graphical parameter
3: In doTryCatch(return(expr), name, parentenv, handler) :
  "int.angle" is not a graphical parameter
4: In doTryCatch(return(expr), name, parentenv, handler) :
  "int.angle" is not a graphical parameter
5: In doTryCatch(return(expr), name, parentenv, handler) :
  "int.angle" is not a graphical parameter
6: In doTryCatch(return(expr), name, parentenv, handler) :
  "int.angle" is not a graphical parameter
> set.seed(1)
> amostra <- sample(c(TRUE, FALSE), nrow(Dados), replace=TRUE, prob=c(0.8,0
.2))
> train <- Dados[amostra, ]
> test <- Dados[!amostra, ]
> chisq.test(train$CV_class, train$Sexo)

      Pearson's Chi-squared test with Yates' continuity correction

data:  train$CV class and train$Sexo
X-squared = 1.2051e-30, df = 1, p-value = 1

> tab1 <- table(train$Sexo, train$CV_class)
> chisq.test(train$CV_class, train$Faixa_etaria)

      Pearson's Chi-squared test

data:  train$CV_class and train$Faixa_etaria
X-squared = 2.1651, df = 5, p-value = 0.8259

> tab2 <- table(train$Faixa_etaria, train$CV_class)

```

```

> chisq.test(train$CV_class, train$TB)

        Pearson's Chi-squared test with Yates' continuity correction

data:  train$CV_class and train$TB
X-squared = 16.942, df = 1, p-value = 3.854e-05

> tab3 <- table(train$TB, train$CV_class)
> chisq.test(train$CV_class, train$APSS_class)

        Pearson's Chi-squared test with Yates' continuity correction

data:  train$CV_class and train$APSS_class
X-squared = 1.201, df = 1, p-value = 0.2731

> tab4 <- table(train$APSS_class, train$CV_class)
> attach(train)
> Modelo0=glm(CV_class~(as.factor(Sexo)+Idade+as.factor(TB)+Peso+Tempo+as.f
actor(APSS_class))^2,
+             data=train, family = binomial(link = "logit"))
Warning message:
glm.fit: fitted probabilities numerically 0 or 1 occurred
> #Stepwise
> step(Modelo0, direction = 'both')
Start:  AIC=90.48
CV_class ~ (as.factor(Sexo) + Idade + as.factor(TB) + Peso +
Tempo + as.factor(APSS_class))^2

              Df Deviance    AIC
- as.factor(TB):Tempo      1  46.485  88.485
- as.factor(Sexo):as.factor(TB)  1  46.490  88.490
- Tempo:as.factor(APSS_class)  1  46.528  88.528
- Idade:Peso                1  46.540  88.540
- as.factor(TB):Peso        1  46.621  88.621
- as.factor(Sexo):as.factor(APSS_class)  1  46.632  88.632
- as.factor(Sexo):Tempo     1  47.096  89.096
- as.factor(Sexo):Peso      1  47.103  89.103
- Idade:as.factor(TB)      1  47.314  89.314
- Idade:Tempo              1  47.710  89.710
- Peso:as.factor(APSS_class)  1  47.732  89.732
- as.factor(Sexo):Idade    1  48.325  90.325
<none>                      1  46.483  90.483
- as.factor(TB):as.factor(APSS class)  1  49.605  91.605
- Idade:as.factor(APSS_class)  1  50.554  92.554
- Peso:Tempo               1  55.123  97.123

Step:  AIC=88.49
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):as.fact
or(TB) +
as.factor(Sexo):Peso + as.factor(Sexo):Tempo + as.factor(Sexo):as.facto
r(APSS_class) +
Idade:as.factor(TB) + Idade:Peso + Idade:Tempo + Idade:as.factor(APSS_c
lass) +

```

```

as.factor(TB):Peso + as.factor(TB):as.factor(APSS_class) +
Peso:Tempo + Peso:as.factor(APSS_class) + Tempo:as.factor(APSS_class)

Df Deviance AIC
- as.factor(Sexo):as.factor(TB) 1 46.498 86.498
- Tempo:as.factor(APSS_class) 1 46.528 86.528
- Idade:Peso 1 46.549 86.549
- as.factor(Sexo):as.factor(APSS_class) 1 46.634 86.634
- as.factor(TB):Peso 1 46.656 86.656
- as.factor(Sexo):Peso 1 47.106 87.106
- as.factor(Sexo):Tempo 1 47.125 87.125
- Idade:as.factor(TB) 1 47.349 87.349
- Idade:Tempo 1 47.727 87.727
- Peso:as.factor(APSS_class) 1 47.749 87.749
- as.factor(Sexo):Idade 1 48.332 88.332
<none> 46.485 88.485
+ as.factor(TB):Tempo 1 46.483 90.483
- Idade:as.factor(APSS_class) 1 50.576 90.576
- as.factor(TB):as.factor(APSS_class) 1 51.086 91.086
- Peso:Tempo 1 55.201 95.201

Step: AIC=86.5
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
as.factor(Sexo):Tempo + as.factor(Sexo):as.factor(APSS_class) +
Idade:as.factor(TB) + Idade:Peso + Idade:Tempo + Idade:as.factor(APSS_c
lass) +
as.factor(TB):Peso + as.factor(TB):as.factor(APSS_class) +
Peso:Tempo + Peso:as.factor(APSS_class) + Tempo:as.factor(APSS_class)

Df Deviance AIC
- Tempo:as.factor(APSS_class) 1 46.533 84.533
- Idade:Peso 1 46.554 84.554
- as.factor(Sexo):as.factor(APSS_class) 1 46.634 84.634
- as.factor(TB):Peso 1 46.657 84.657
- as.factor(Sexo):Peso 1 47.111 85.111
- as.factor(Sexo):Tempo 1 47.129 85.129
- Idade:as.factor(TB) 1 47.512 85.512
- Idade:Tempo 1 47.732 85.732
- Peso:as.factor(APSS_class) 1 47.752 85.752
- as.factor(Sexo):Idade 1 48.457 86.457
<none> 46.498 86.498
+ as.factor(Sexo):as.factor(TB) 1 46.485 88.485
+ as.factor(TB):Tempo 1 46.490 88.490
- Idade:as.factor(APSS_class) 1 50.724 88.724
- as.factor(TB):as.factor(APSS_class) 1 51.277 89.277
- Peso:Tempo 1 55.380 93.380

Step: AIC=84.53
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
as.factor(Sexo):Tempo + as.factor(Sexo):as.factor(APSS_class) +
Idade:as.factor(TB) + Idade:Peso + Idade:Tempo + Idade:as.factor(APSS_c
lass) +

```

```

as.factor(TB):Peso + as.factor(TB):as.factor(APSS_class) +
Peso:Tempo + Peso:as.factor(APSS_class)

Df Deviance AIC
- Idade:Peso 1 46.623 82.623
- as.factor(Sexo):as.factor(APSS_class) 1 46.647 82.647
- as.factor(TB):Peso 1 46.680 82.680
- as.factor(Sexo):Tempo 1 47.134 83.134
- as.factor(Sexo):Peso 1 47.138 83.138
- Idade:as.factor(TB) 1 47.605 83.605
- Idade:Tempo 1 47.737 83.737
- as.factor(Sexo):Idade 1 48.458 84.458
<none> 46.533 84.533
- Peso:as.factor(APSS_class) 1 48.586 84.586
+ Tempo:as.factor(APSS_class) 1 46.498 86.498
+ as.factor(Sexo):as.factor(TB) 1 46.528 86.528
+ as.factor(TB):Tempo 1 46.531 86.531
- Idade:as.factor(APSS_class) 1 50.907 86.907
- as.factor(TB):as.factor(APSS_class) 1 51.282 87.282
- Peso:Tempo 1 55.844 91.844

Step: AIC=82.62
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
as.factor(Sexo):Tempo + as.factor(Sexo):as.factor(APSS_class) +
Idade:as.factor(TB) + Idade:Tempo + Idade:as.factor(APSS_class) +
as.factor(TB):Peso + as.factor(TB):as.factor(APSS_class) +
Peso:Tempo + Peso:as.factor(APSS_class)

Df Deviance AIC
- as.factor(Sexo):as.factor(APSS_class) 1 46.737 80.737
- as.factor(TB):Peso 1 46.793 80.793
- as.factor(Sexo):Tempo 1 47.174 81.174
- as.factor(Sexo):Peso 1 47.486 81.486
- Idade:as.factor(TB) 1 47.915 81.915
<none> 46.623 82.623
- Peso:as.factor(APSS_class) 1 48.838 82.838
- as.factor(Sexo):Idade 1 49.067 83.067
- Idade:Tempo 1 49.076 83.076
+ Idade:Peso 1 46.533 84.533
+ Tempo:as.factor(APSS_class) 1 46.554 84.554
+ as.factor(TB):Tempo 1 46.619 84.619
+ as.factor(Sexo):as.factor(TB) 1 46.623 84.623
- Idade:as.factor(APSS_class) 1 50.908 84.908
- as.factor(TB):as.factor(APSS_class) 1 52.119 86.119
- Peso:Tempo 1 56.105 90.105

Step: AIC=80.74
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
as.factor(Sexo):Tempo + Idade:as.factor(TB) + Idade:Tempo +
Idade:as.factor(APSS_class) + as.factor(TB):Peso + as.factor(TB):as.factor(APSS_class) +
Peso:Tempo + Peso:as.factor(APSS_class)

```

	Df	Deviance	AIC
- as.factor(TB):Peso	1	46.913	78.913
- as.factor(Sexo):Tempo	1	47.183	79.183
- Idade:as.factor(TB)	1	47.933	79.933
- as.factor(Sexo):Peso	1	48.072	80.072
<none>		46.737	80.737
- Peso:as.factor(APSS_class)	1	48.860	80.860
- Idade:Tempo	1	49.109	81.109
- as.factor(Sexo):Idade	1	49.870	81.870
+ as.factor(Sexo):as.factor(APSS_class)	1	46.623	82.623
+ Idade:Peso	1	46.647	82.647
+ Tempo:as.factor(APSS_class)	1	46.699	82.699
+ as.factor(Sexo):as.factor(TB)	1	46.732	82.732
+ as.factor(TB):Tempo	1	46.737	82.737
- Idade:as.factor(APSS_class)	1	50.986	82.986
- as.factor(TB):as.factor(APSS_class)	1	52.154	84.154
- Peso:Tempo	1	56.304	88.304

Step: AIC=78.91

```
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
  as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
  as.factor(Sexo):Tempo + Idade:as.factor(TB) + Idade:Tempo +
  Idade:as.factor(APSS_class) + as.factor(TB):as.factor(APSS_class) +
  Peso:Tempo + Peso:as.factor(APSS_class)
```

	Df	Deviance	AIC
- as.factor(Sexo):Tempo	1	47.384	77.384
- Idade:as.factor(TB)	1	47.987	77.987
- as.factor(Sexo):Peso	1	48.103	78.103
<none>		46.913	78.913
- Idade:Tempo	1	49.156	79.156
- Peso:as.factor(APSS_class)	1	49.170	79.170
- as.factor(Sexo):Idade	1	49.929	79.929
+ as.factor(TB):Peso	1	46.737	80.737
+ as.factor(Sexo):as.factor(APSS_class)	1	46.793	80.793
+ Idade:Peso	1	46.801	80.801
+ as.factor(TB):Tempo	1	46.867	80.867
+ Tempo:as.factor(APSS_class)	1	46.887	80.887
+ as.factor(Sexo):as.factor(TB)	1	46.888	80.888
- Idade:as.factor(APSS_class)	1	51.385	81.385
- as.factor(TB):as.factor(APSS_class)	1	52.171	82.171
- Peso:Tempo	1	56.305	86.305

Step: AIC=77.38

```
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
  as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
  Idade:as.factor(TB) + Idade:Tempo + Idade:as.factor(APSS_class) +
  as.factor(TB):as.factor(APSS_class) + Peso:Tempo + Peso:as.factor(APSS_
class)
```

	Df	Deviance	AIC
- Idade:as.factor(TB)	1	49.198	77.198
<none>		47.384	77.384

```

- Peso:as.factor(APSS_class)          1    49.558  77.558
- Idade:Tempo                          1    49.561  77.561
- as.factor(Sexo):Peso                 1    49.698  77.698
- as.factor(Sexo):Idade                1    49.938  77.938
+ as.factor(Sexo):Tempo                1    46.913  78.913
+ as.factor(TB):Peso                   1    47.183  79.183
+ as.factor(TB):Tempo                  1    47.255  79.255
+ Idade:Peso                            1    47.321  79.321
+ as.factor(Sexo):as.factor(APSS_class) 1    47.373  79.373
+ Tempo:as.factor(APSS_class)          1    47.373  79.373
+ as.factor(Sexo):as.factor(TB)        1    47.380  79.380
- Idade:as.factor(APSS_class)          1    52.597  80.597
- as.factor(TB):as.factor(APSS_class)  1    54.050  82.050
- Peso:Tempo                            1    58.274  86.274

Step:   AIC=77.2
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
           as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
           Idade:Tempo + Idade:as.factor(APSS_class) + as.factor(TB):as.factor(APSS_class) +
           Peso:Tempo + Peso:as.factor(APSS_class)

           Df Deviance   AIC
- Peso:as.factor(APSS_class)          1    51.010  77.010
<none>                                49.198  77.198
+ Idade:as.factor(TB)                  1    47.384  77.384
- Idade:Tempo                          1    51.495  77.495
- as.factor(Sexo):Peso                 1    51.803  77.803
+ as.factor(Sexo):Tempo                1    47.987  77.987
+ as.factor(TB):Tempo                  1    48.519  78.519
- as.factor(Sexo):Idade                1    52.594  78.594
+ as.factor(Sexo):as.factor(TB)        1    48.917  78.917
+ as.factor(Sexo):as.factor(APSS_class) 1    48.952  78.952
- Idade:as.factor(APSS_class)          1    53.024  79.024
+ Idade:Peso                            1    49.067  79.067
+ Tempo:as.factor(APSS_class)          1    49.126  79.126
+ as.factor(TB):Peso                   1    49.160  79.160
- as.factor(TB):as.factor(APSS_class)  1    54.051  80.051
- Peso:Tempo                            1    58.890  84.890

Step:   AIC=77.01
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
           as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
           Idade:Tempo + Idade:as.factor(APSS_class) + as.factor(TB):as.factor(APSS_class) +
           Peso:Tempo

           Df Deviance   AIC
- Idade:Tempo                          1    52.669  76.669
- as.factor(Sexo):Peso                 1    52.936  76.936
<none>                                51.010  77.010
+ Peso:as.factor(APSS_class)          1    49.198  77.198
+ Idade:as.factor(TB)                  1    49.558  77.558
- Idade:as.factor(APSS_class)          1    53.643  77.643

```

```

+ Tempo:as.factor(APSS_class)          1  50.173 78.173
+ as.factor(Sexo):Tempo                 1  50.227 78.227
- as.factor(Sexo):Idade                 1  54.299 78.299
+ as.factor(TB):Tempo                   1  50.551 78.551
+ as.factor(Sexo):as.factor(TB)        1  50.682 78.682
+ as.factor(Sexo):as.factor(APSS_class) 1  50.776 78.776
+ Idade:Peso                             1  50.870 78.870
+ as.factor(TB):Peso                    1  50.959 78.959
- as.factor(TB):as.factor(APSS_class)   1  56.028 80.028
- Peso:Tempo                             1  59.111 83.111

Step:  AIC=76.67
CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Peso + Tempo +
  as.factor(APSS_class) + as.factor(Sexo):Idade + as.factor(Sexo):Peso +
  Idade:as.factor(APSS_class) + as.factor(TB):as.factor(APSS_class) +
  Peso:Tempo

              Df Deviance   AIC
<none>              52.669 76.669
- as.factor(Sexo):Peso          1  54.756 76.756
+ Idade:as.factor(TB)           1  50.951 76.951
+ Idade:Tempo                   1  51.010 77.010
- as.factor(Sexo):Idade         1  55.226 77.226
+ Peso:as.factor(APSS_class)    1  51.495 77.495
- Idade:as.factor(APSS_class)   1  55.658 77.658
+ as.factor(Sexo):Tempo        1  51.911 77.911
+ as.factor(Sexo):as.factor(TB) 1  52.135 78.135
+ Tempo:as.factor(APSS_class)   1  52.241 78.241
+ as.factor(TB):Tempo          1  52.347 78.347
+ Idade:Peso                    1  52.499 78.499
+ as.factor(Sexo):as.factor(APSS_class) 1  52.553 78.553
+ as.factor(TB):Peso            1  52.669 78.669
- as.factor(TB):as.factor(APSS class) 1  57.072 79.072
- Peso:Tempo                    1  61.074 83.074

Call:  glm(formula = CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) +
  Peso + Tempo + as.factor(APSS_class) + as.factor(Sexo):Idade +
  as.factor(Sexo):Peso + Idade:as.factor(APSS class) + as.factor(TB):as.f
actor(APSS_class) +
  Peso:Tempo, family = binomial(link = "logit"), data = train)

Coefficients:
              (Intercept)              as.factor(Sexo)
1              -41.20715              9.6543
3
              Idade              as.factor(TB)
1              0.06361              -3.5765
0
              Peso              Tempo
0              0.60258              1.6311
2

```

```

          as.factor(APSS_class)1          as.factor(Sexo)1:Idade
e
          5.72535                      -0.0800
2
          as.factor(Sexo)1:Peso          Idade:as.factor(APSS_class)
1
          -0.15217                      -0.1170
2
as.factor(TB)1:as.factor(APSS_class)1          Peso:Tempo
o
          3.88148                      -0.0207
4

Degrees of Freedom: 280 Total (i.e. Null); 269 Residual
Null Deviance:      374.4
Residual Deviance: 52.67      AIC: 76.67
There were 50 or more warnings (use warnings() to see the first 50)
> ModeloFinal<-glm(CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) + Pes
o + Tempo + as.factor(APSS_class) +
+          as.factor(Sexo):Idade +
+ as.factor(Sexo):Peso + Idade:as.factor(APSS_class) + as.factor(TB):as.fac
tor(APSS_class) + Peso:Tempo,
+ family = binomial(link = "logit"), data = train)
> summary(ModeloFinal)

Call:
glm(formula = CV_class ~ as.factor(Sexo) + Idade + as.factor(TB) +
    Peso + Tempo + as.factor(APSS_class) + as.factor(Sexo):Idade +
    as.factor(Sexo):Peso + Idade:as.factor(APSS_class) + as.factor(TB):as.f
actor(APSS_class) +
    Peso:Tempo, family = binomial(link = "logit"), data = train)

Coefficients:
                    Estimate Std. Error z value Pr(>|z|
)
(Intercept)          -41.207147   10.477810  -3.933  8.4e-0
5 ***
as.factor(Sexo)1         9.654327    5.016245   1.925  0.05427
8 .
Idade                   0.063606    0.059770   1.064  0.28724
8
as.factor(TB)1         -3.576502    1.362885  -2.624  0.00868
5 **
Peso                    0.602579    0.173054   3.482  0.00049
8 ***
Tempo                  1.631121    0.430404   3.790  0.00015
1 ***
as.factor(APSS_class)1  5.725349    3.438137   1.665  0.09586
3 .
as.factor(Sexo)1:Idade -0.080016    0.055888  -1.432  0.15222
3
as.factor(Sexo)1:Peso  -0.152172    0.111936  -1.359  0.17400
2

```

```

Idade:as.factor(APSS_class)1      -0.117019   0.070313  -1.664  0.09606
0 .
as.factor(TB)1:as.factor(APSS_class)1  3.881482   1.943471   1.997  0.04580
4 *
Peso:Tempo                          -0.020739   0.007792  -2.662  0.00777
5 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 374.376 on 280 degrees of freedom
Residual deviance: 52.669 on 269 degrees of freedom
AIC: 76.669

Number of Fisher Scoring iterations: 9

> #Odds ratio vs Intervalo de Confiança
> exp(cbind(OR=coef(ModeloFinal), confint(ModeloFinal)))
Waiting for profiling to be done...

```

	OR	2.5 %	97.5 %
(Intercept)	1.270467e-18	2.169641e-29	5.027021e-1
as.factor(Sexo)1	1.558910e+04	1.725127e+00	9.435868e+0
Idade	1.065672e+00	9.567993e-01	1.204862e+0
as.factor(TB)1	2.797337e-02	1.471349e-03	3.890004e-0
Peso	1.826825e+00	1.360726e+00	2.775977e+0
Tempo	5.109597e+00	2.447778e+00	1.397122e+0
as.factor(APSS_class)1	3.065402e+02	4.583660e-01	4.611512e+0
as.factor(Sexo)1:Idade	9.231016e-01	8.085571e-01	1.016716e+0
as.factor(Sexo)1:Peso	8.588404e-01	6.729925e-01	1.053171e+0
Idade:as.factor(APSS_class)1	8.895685e-01	7.678245e-01	1.015184e+0
as.factor(TB)1:as.factor(APSS_class)1	4.849604e+01	1.284280e+00	3.184606e+0
Peso:Tempo	9.794749e-01	9.622554e-01	9.933986e-0

```

There were 26 warnings (use warnings() to see them)
> #Curva ROC
> library(ggplot2)
> library(pROC)
Type 'citation("pROC")' for a citation.

Attaching package: 'pROC'

```

```

The following objects are masked from 'package:stats':

  cov, smooth, var

Warning message:
package 'pROC' was built under R version 4.4.1
> predicted <- predict(ModeloFinal, test, type="response")
> roc <- roc(test$CV_class, predicted)
Setting levels: control = 0, case = 1
Setting direction: controls < cases
> auc <- round(auc(test$CV_class, predicted),4)
Setting levels: control = 0, case = 1
Setting direction: controls < cases
> ggroc(roc, colour = 'steelblue', size = 1) +
+   ggtitle(paste0('Curva ROC ', '(AUC = ', auc, ')'))+
+   geom_abline(intercept = 1, slope = 1, colour = 'red', linetype = 'dashed'
) +
+   ggtitle(paste0('Curva ROC ', '(AUC = ', auc, ')'))
> library(ResourceSelection)
ResourceSelection 0.3-6      2023-06-27
Warning message:
package 'ResourceSelection' was built under R version 4.4.1
> hoslem.test(CV_class,fitted(ModeloFinal),g=10)

      Hosmer and Lemeshow goodness of fit (GOF) test

data:  CV_class, fitted(ModeloFinal)
X-squared = 1.2986, df = 8, p-value = 0.9956

> library("modEvA")

Attaching package: 'modEvA'

The following object is masked from 'package:psych':

  AUC

The following objects are masked from 'package:DescTools':

  AUC, RMSE

Warning message:
package 'modEvA' was built under R version 4.4.1
> RsqGLM(ModeloFinal)
$CoxSnell
[1] 0.6817328

$Nagelkerke
[1] 0.9261037

$McFadden
[1] 0.8593145

$Tjur

```

```
[1] 0.8889199
```

```
$sqPearson
```

```
[1] 0.8889746
```