



Modelo matemático para prever probabilidade de quilombolas desenvolverem síndrome metabólica com fluxograma de atendimento de saúde

Mathematical model to predict the probability of quilombolas developing metabolic syndrome with a health care flowchart

R. C. Torres^{1*}; M. A. P. Nunes²; M. M. Mota³; T. L. T. B. da Silva³; C. C. da C. Oliveira⁴; C. M. de Melo⁵

¹Laboratório de Planejamento e Promoção em Saúde (LPPS), Universidade Tiradentes, 49032-490, Aracaju-Sergipe, Brasil

²Departamento de Medicina, Universidade Federal de Sergipe, Aracaju, 49032-490, Aracaju-Sergipe, Brasil

³Curso de Educação Física, Centro Universitário Estácio de Sergipe, 49020-530, Aracaju-Sergipe, Brasil

⁴Instituto de Hematologia e Hemoterapia de Sergipe – IHHS, 49020-270, Aracaju-Sergipe, Brasil

⁵Laboratório de Biologia Molecular do LPPS-ITP, Universidade Tiradentes, 49032-490, Aracaju-Sergipe, Brasil

*ruthcristini@gmail.com

(Recebido em 20 de novembro de 2022, aceito em 31 de agosto de 2023)

Desenvolveu-se um modelo matemático utilizando algoritmos de *Machine learning* para prever probabilidade de quilombolas desenvolverem síndrome metabólica, bem como propor um fluxograma de atendimento de saúde. Este foi um estudo transversal com uso de inteligência artificial. Adotou-se como variável dependente ter ou não síndrome metabólica. Foi realizada a análise bivariada comparando-se as variáveis independentes, os indicadores antropométricos e bioquímicos em relação a presença de síndrome metabólica e as variáveis categóricas que foram avaliadas pelo teste do Qui-quadrado ($p < 0,05$). Utilizou-se o teste de Análise de Variância ou Kruskal-Wallis conforme a tendência de normalidade avaliada pelo teste de Shapiro-Wilk e a ferramenta de análise de dados *machine learning*, por meio do método de Árvore de Decisão. A árvore de decisão para predição de síndrome metabólica em quilombolas apresentou acurácia de 75%, gerando um gráfico em relação ao processo, ilustrado por meio de fluxograma para orientar a tomada de decisão em relação à saúde e prevenção de síndrome metabólica. O modelo preditivo permitiu identificar os pontos de corte específicos dos indicadores antropométricos mais importantes a serem mensurados no primeiro atendimento de saúde dos quilombolas. A acurácia do modelo preditivo permite a aplicação do fluxograma em outras comunidades quilombolas, apresentando-se como uma ferramenta tecnológica facilitadora para tomada de decisão em saúde.

Palavras-chave: grupo com ancestrais do continente africano, síndrome metabólica, aprendizado de máquina.

To develop a mathematical model using Machine learning algorithms to predict the probability of quilombolas developing metabolic syndrome, as well as to propose a health care flowchart. Cross-sectional study using artificial intelligence. Having or not metabolic syndrome was adopted as the dependent variable. Bivariate analysis was performed comparing independent variables, anthropometric and biochemical indicators in relation to the presence of metabolic syndrome and categorical variables that were evaluated using the chi-square test ($p < 0.05$). The Analysis of Variance or Kruskal-Wallis test was used according to the normality trend evaluated by the Shapiro-Wilk test. The machine learning data analysis tool was used, through the Decision Tree method. The decision tree for predicting metabolic syndrome in quilombolas was 75% accurate, generating a graph in relation to the decision process illustrated by means of a flowchart to guide decision-making in relation to health and prevention of metabolic syndrome. The predictive model allowed identifying the specific cutoff points of the most important anthropometric indicators to be measured in the first health care of the quilombolas. The accuracy of the predictive model allows the application of the flowchart in other quilombola communities, presenting itself as a technological tool that facilitates decision-making in health.

Keyword: group with ancestors of the african continent, metabolic syndrome, machine learning.

1. INTRODUÇÃO

As comunidades quilombolas (remanescentes de quilombos) tiveram origem a partir da fuga de escravos que formavam grupos socialmente organizados contra os abusos sofridos durante o

período da escravidão [1]. Após a abolição da escravatura no Brasil, não foram traçadas políticas públicas imediatas de atenção à população negra, que não tiveram assistência social ou jurídica para que possibilitasse a inserção dos negros de forma igualitária na sociedade brasileira [2], o que contribuiu para as inequidades e desigualdades sociais e de saúde que influenciam nos modos de vida, mortalidade e morbidade deste grupo populacional [3, 4].

No Brasil existem cerca de 1,3 milhão de pessoas que se identificam como quilombolas, tendo a maior concentração na região Nordeste, com 905.415 quilombolas, correspondendo a 68,2% da população quilombola do país [5]. Já foram reconhecidas 3.502 comunidades quilombolas, dentre as quais, 44 (1,25%) estão localizadas no Estado de Sergipe [6], correspondendo a um total de 28.124 pessoas quilombolas [5].

A Agenda Nacional de Prioridades de Pesquisa em Saúde, destacou a saúde da população negra e das comunidades tradicionais no Brasil e as Doenças Crônicas não Transmissíveis (DCNT) como 2 (dois) dos 14 (quatorze) eixos temáticos de saúde prioritários para investigações científicas, tecnológicas e de inovação. A agenda possibilita o direcionamento de investimentos em temas de pesquisa estratégicos para o Sistema Único de Saúde (SUS), visando contribuir para a implementação de políticas públicas de saúde mais efetivas [7].

Neste contexto, encontra-se a Síndrome Metabólica (SM) caracterizada por um conjunto de fatores de risco cardiovasculares e desequilíbrios metabólicos que podem contribuir para desenvolvimento de DCNT [8].

Para que a SM seja caracterizada, o sujeito deve apresentar 3 (três) dos fatores de risco a seguir: Circunferência Abdominal (CA) >90 cm para homens e >80 cm para mulheres [9], triglicérides ≥ 150 mg/dL ou uso de medicamentos para tratamento de hipertrigliceridemia, glicemia em jejum ≥ 100 mg/dL ou uso de medicamentos para diabetes; High Density Lipoproteins (HDL) <40 mg/dL (homem) ou <50 mg/dL (mulher) ou uso de medicamentos para tratamento de HDL baixo e valores de pressão arterial sistólica ≥ 130 mmHg e/ou diastólica ≥ 85 mmHg, ou o uso de medicamento anti-hipertensivo [10].

A população negra apresenta suscetibilidade aos riscos cardiometabólicos, sendo necessárias mais pesquisas com metodologias robustas para compreender este cenário [11]. Destaca-se que condições socioeconômicas desfavoráveis, acesso limitado a cuidados de saúde, e menor disponibilidade de alimentos saudáveis podem contribuir para maior prevalência da SM na população negra [3]. Diversos índices antropométricos como Índice de Massa Corporal (IMC), Circunferência de Cintura (CC), Relação Cintura Quadril (RCQ), Relação Cintura-Estatura (RCEst) e CA, têm sido postos em prática para determinar a relação entre sobrepeso e fatores de risco cardiovascular [9, 12], entretanto os estudos anteriores não tem organizado estes fatores em sistemas de predição de risco ao serem avaliados em conjunto como foi apresentado no presente artigo.

A circunferência do pescoço (CP) auxilia na detecção precoce de adiposidade corporal, bem como na avaliação de risco cardiovascular e correlaciona-se positivamente os componentes da SM [13]. A circunferência de quadril (CQ) é considerada um dos indicadores de obesidade associados à mortalidade em pessoas negras, brancas não-hispânicas e mexicanas-americanas [14].

A CC apresenta score de risco alto para desenvolvimento da SM em pessoas com baixos níveis socioeconômico e de educação, contribuindo para o desfecho do fenômeno [15]. O Índice de Adiposidade Corporal (IAC) permite a classificação do peso dos sujeitos em três classes: obesidade, peso em excesso e peso saudável [16]. O IAC aliado à CC e ao IMC pode contribuir para a predição de risco cardiometabólico, pois estes três indicadores de forma semelhante e consistente estimam adiposidade e o risco de Doenças Cardiovasculares (DCV) [17].

Dados sobre as condições de saúde da população negra ainda são insuficientes, o que dificulta a implementação de intervenções preventivas e assistenciais rápidas e assertivas neste grupo populacional [18]. Com os ascendentes índices de prevalência da SM e sua associação ao aumento do risco de diabetes mellitus tipo II e DCV, é de suma importância o diagnóstico precoce [19].

A inteligência artificial vem sendo utilizada com o intuito de melhorar a tomada de decisão em saúde em relação ao diagnóstico de doenças, tratamento e manutenção da saúde [2, 20]. Modelos preditivos podem ser aplicados para estimar o risco de determinado desfecho ocorrer, com base em características socioeconômicas, demográficas, relacionadas ao hábito de vida e às condições de saúde, perfil clínico, entre outras.

O presente estudo é pioneiro na utilização de todos os indicadores antropométricos e bioquímicos citados em conjunto, em modelo de predição de SM em quilombolas, possibilitando tanto uma avaliação completa e robusta da composição corporal desta população, como a identificação de quais indicadores analisados podem prever a SM.

Até o momento não foi identificado na literatura científica estudos que utilizem inteligência artificial para a predição de SM em quilombolas, o que poderia auxiliar na efetividade de intervenções com baixo custo, tratamentos e estratégias preventivas. Deste modo, objetivou-se desenvolver um modelo matemático utilizando algoritmos de *Machine learning* para prever probabilidade de quilombolas desenvolverem síndrome metabólica, bem como propor um fluxograma de atendimento de saúde.

2. MATERIAL E MÉTODOS

Trata-se de um estudo transversal com uso de inteligência artificial. A coleta de dados foi realizada no período de novembro 2017 a julho de 2019 nas comunidades quilombolas Canta Galo, Pirangy e Terra Dura e Coqueiral, localizadas no município de Capela, na região leste do Estado de Sergipe, que abrangem 386 famílias e aproximadamente 1.158 indivíduos em uma área territorial contígua. A comunidade Pirangi tem como base econômica a agricultura de subsistência. As comunidades Terra Dura e Coqueiral e Canta Galo possuem predomínio de atividades de agricultura de subsistência e a produção artesanal.

Foram incluídos no estudo os sujeitos com idade \geq a 18 anos, de ambos os sexos, moradores cadastrados como quilombola na associação de moradores local e usuários da Estratégia de Saúde da Família (ESF) das comunidades em questão. Os critérios de exclusão foram: indivíduos que tiveram incapacidade de compreender os questionamentos e os que não estavam em jejum de, no mínimo, 12 horas no dia da realização dos exames sanguíneos.

Os participantes foram convidados para estarem presentes nos locais da pesquisa (escolas das comunidades) em dias previamente marcados com a liderança quilombola. Apesar da seleção ser por conveniência, respeitou-se os critérios de inclusão e exclusão previstos para a amostra. Foram identificados 465 indivíduos maiores de 18 anos, de ambos os sexos, moradores das comunidades em questão e excluídos 2 que apresentaram incapacidade de compreender os questionamentos.

A coleta de dados foi realizada por meio de um questionário estruturado para a identificação das características sociodemográficas (sexo, idade, escolaridade, renda familiar, cor da pele), perfil clínico, uso de tabaco, hábitos alimentares.

Foi utilizada uma ficha clínica para registro da mensuração dos índices antropométricos, da aferição de pressão arterial, dos exames laboratoriais de lipidograma e glicemia de jejum [21], do Escore de Risco de Framingham (ERF) [22], da classificação do nível de atividade física por meio do IPAQ - International Physical Activity Questionnaire [23] e aplicada a Escala Brasileira de Insegurança Alimentar 24. Para a definição do ERF foram consideradas a faixa etária, o sexo, valores de pressão arterial sistólica, valores da razão entre o colesterol total e a dosagem de HDL, presença de uso de tabaco e diagnóstico de diabetes.

A avaliação antropométrica foi realizada por meio da mensuração dos valores das circunferências de cintura (CC), quadril (CQ), pescoço (CP), cálculo do Índice de Massa Corporal (IMC), Índice de Adiposidade Corporal (IAC), Relação Cintura Estatura (RCEst), Relação Cintura Quadril (RCQ) [25], Circunferência abdominal (CA) [11], e índice de Conicidade (IC) [26].

Utilizou-se a ferramenta de análise de dados *Machine learning*, por meio do método de Árvore de Decisão (técnica de mineração de dados), com vistas a construção de um modelo matemático que possibilite a predição das probabilidades de desenvolvimento de SM na população quilombola e oriente a tomada de decisão em relação à saúde e prevenção de SM. Entre as principais vantagens das árvores de decisão estão a fácil interpretação dos dados, rapidez na geração dos resultados, baixo custo computacional e admissão de variáveis categóricas e nominais.

O algoritmo Árvore de Decisão pertence à família de algoritmos de aprendizado supervisionado e é usado para resolver problemas de regressão e classificação. O objetivo de usar esta técnica é criar um modelo de treinamento que possa ser usado para prever a classe ou valor da variável de referência (target) aprendendo regras de decisão simples inferidas de dados

anteriores (dados de treinamento). O processo ocorre em duas etapas, uma seria a etapa de aprendizado e a outra seria a etapa de previsão. A Árvore de Decisão utiliza uma função que apresenta os dados como um vetor de atributos e uma decisão como valor de saída. Seu funcionamento se dá pela divisão subsequente do conjunto de dados, de forma a separá-lo em um subconjunto homogêneo, com casos de única classe [27].

Uma árvore de decisão cresce por particionamento recursivo binário. A variável resposta foi usada na fórmula especificada e foram escolhidas as divisões dos termos do lado direito. As variáveis numéricas foram divididas em $X < a$ e $X > a$; e as variáveis categóricas foram divididas em dois grupos não vazios. Foi escolhida a divisão que maximizou a redução da impureza e o processo foi repetido. A divisão continuou até que os nós terminais fossem muito pequenos ou poucos para serem divididos.

Algumas das suposições foram realizadas ao usar a árvore de decisão, conforme exposto a seguir:

- Inicialmente, todo o conjunto de treinamento é considerado como a raiz;
- Preferencialmente as variáveis devem ser categóricas. Quando forem contínuas, deverão discretizados antes da construção do modelo;
- Os registros são distribuídos recursivamente com base nos valores dos atributos;
- A ordem de colocação de atributos como raiz ou nó interno da árvore é feita usando alguma abordagem estatística;
- A classificação das variáveis e a identificação dos subconjuntos ocorrem a partir de cálculos de probabilidade das decisões no conjunto do banco de dados e em todos os subconjuntos de variáveis independentes relacionadas à decisão;
- O cálculo de heterogeneidade nos subconjuntos gerados permite avaliar que variáveis têm maior Ganho de Informação (GI) e participam da formação da árvore no apoio à tomada de decisão.

A técnica de *Machine learning* utilizada parte do pressuposto de que os algoritmos que têm a capacidade de aprenderem sozinhos a partir de dados. Desta forma, para medir o desempenho real do modelo criado, foi necessário realizar testes utilizando dados diferentes dos que foram apresentados em sua criação. Com esta finalidade, o banco de dados foi subdividido em dados de treinamento (80% dos dados) e dados de teste (20% dos dados), este último usado para testar o modelo e calcular a acurácia por meio do Programa R version 3.6.1. Dessa forma os dados de treino foram apresentados ao algoritmo de *Machine learning* para criação do modelo.

Utilizou-se o seguinte algoritmo de árvore de decisão utilizado para criar o modelo:

```
# decision tree
library(tree)
library(caTools)
sample <- sample.split(ruth, SplitRatio = 0.8)
train <- subset(ruth, sample == TRUE)
test <- subset(ruth, sample == FALSE)

tree_train = tree(s_metab ~ imc + c_pesq + c_cint + c_abd + c_quadril + conicidade + rce + iac, data =
train)
summary(tree_train)
plot(tree_train)
text(tree_train, cex=.8, use.n=TRUE, all=TRUE)

plot(tree_train);text(tree_train,pretty=0)
tree.pred = predict(tree_train,test,type="class")
with(test,table(tree.pred,s_metab))
mean(tree.pred==test$s_metab)
(5 + 3 +3)/(5 + 3 +3 + 1 + 2 + 1)

tree_ruth = tree(s_metab ~ imc + c_pesq + c_cint + c_abd + c_quadril + conicidade + rce + iac, data =
ruth)
summary(tree_ruth)
plot(tree_ruth)
text(tree_ruth, cex=.8, use.n=TRUE, all=TRUE)
```

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa com número de CAEE: 61780216800008079 e parecer nº 1822242.

2.1 Análise de dados

Realizou-se análise descritiva por meio de frequências absolutas e relativas das variáveis categóricas que também foram avaliadas por meio do teste Qui Quadrado, considerando valores significativos quando $p < 0,05$. Para as variáveis numéricas foi realizada estatística com medidas de tendência central (média e mediana) e variabilidade (amplitude interquartilica e desvio padrão).

Optou-se por um método de aprendizagem de máquina (*Machine learning*), a árvore de decisão. Os dados de teste foram apresentados ao modelo após a sua criação, simulando previsões reais que o modelo realizará, permitindo assim que o desempenho real seja verificado por meio por meio do Programa R version 3.6.1.

A variável sexo seria uma candidata potencial ao modelo, mas não foi incluída na construção gráfica devido à desproporção da amostra entre homens/mulheres (57:19 -> 3:1) e ao pequeno volume de homens.

Para a predição de risco de desenvolvimento da SM, utilizou-se como variável dependente ter ou não SM. Foi realizada a análise bivariada comparando-se as variáveis independentes (sexo, idade, classificação de risco cardiovascular, ERF e dos indicadores antropométricos IMC, CC, CA, CP, CQ, RCQ, IC, RCEst) em relação a presença de SM, assim como para as variáveis categóricas.

A análise inferencial da SM em relação às variáveis independentes numéricas foi realizada com o teste de Análise de Variância (ANOVA) ou Kruskal-Wallis conforme a tendência à normalidade avaliada por meio do teste de Shapiro-Wilk.

3. RESULTADOS

O grupo populacional de quilombolas em estudo é composto por 19 (25%) homens e 57 (75%) mulheres, com idade média entre $47,2 \pm 18,3$ e $49,0 \pm 14,7$ respectivamente. Dentre as características socioeconômicas e de escolaridade, aproximadamente 60% dos quilombolas não têm nenhum tipo de formação escolar e 43 (56,6%) possuem renda familiar média mensal \leq R\$ 600,00 reais diante de 33 (43,4%) que possuem renda \geq R\$ 1.320 (salário mínimo no Brasil).

Valores de pressão arterial $\geq 140 \times 90$ mmHg foram constatados em 12 (21,0%) mulheres e 3 (15,8%) homens e a ocorrência de diabetes tipo II foi verificada, por meio de diagnóstico médico na Unidade Básica de Saúde (UBS), em 4 (7%) mulheres e 2 (10,5%) homens. O uso do tabaco foi identificado em apenas em 18 (23,6%) sujeitos. Dentre os fatores intervenientes da obesidade neste grupo de quilombolas, identificou-se o consumo de refrigerante ou suco artificial com periodicidade de, pelo menos, 2 vezes por semana e consumo de carne com gordura por 47,36% dos sujeitos avaliados.

Dos 76 (100%) quilombolas, 26 (34,2%) possuem SM, destes 16 (61,5%) são irregularmente ativos ou sedentários e apresentaram insegurança alimentar moderada ou grave, 12 (46,1%) foram classificados pelo IAC como obesos, 15 (57,6%) apresentaram peso em excesso por meio do IMC e 11 (42,3%) apresentaram risco alto ou intermediário por meio do ERF (Tabela 1).

Tabela 1: Distribuição de fatores relacionados à Síndrome Metabólica (SM) - Comunidades Quilombolas (n=76). Capela, SE, Brasil. 2022.

	SM*		p-value
	Sim n(%)	Não n(%)	
Sexo			
Feminino	20(77)	37(74)	< 0,001
Masculino	6(23)	13(26)	
Classificação IAC[†]			
Peso saudável	7(26,9)	13(26)	0,229
Peso em excesso	7(26,9)	17(34)	
Obesidade	12(46,2)	20(40)	
Classificação IMC[‡]			
Obesidade Grau I	6(23)	8(16)	< 0,001
Obesidade Grau II	3(11,6)	4(8,0)	
Peso em excesso	15(57,7)	20(40)	
Peso saudável	2(7,7)	18(36)	
Escore de Framingham			
Risco Alto	5(19,3)	2(4,0)	< 0,001
Risco Baixo	15(57,7)	45(90)	
Risco Intermediário	6(23)	3(6,0)	
Escala Brasileira de Insegurança Alimentar			
Insegurança Alimentar Grave	3(11,6)	5(10)	< 0,001
Insegurança Alimentar moderada	13(50)	22(44)	
Insegurança Alimentar Leve	7(26,9)	16(32)	
Segurança Alimentar	3(11,6)	7(14)	
Classificação do nível de atividade física (IPAQ[§])			
Sedentário	4(15,4)	8(16)	< 0,001
Irregularmente ativo B	7(26,9)	9(18)	
Irregularmente ativo A	5(19,3)	8(16)	
Ativo	10(38,4)	21(42)	
Muito Ativo	0	4(8,0)	

*SM - Síndrome Metabólica; [†]IAC - Índice de Adiposidade Corporal; [‡]IMC - Índice de Massa Corporal; [§]IPAQ - International Physical Activity Questionnaire

A análise da amplitude interquartílica demonstrou que as variáveis peso, CP, CC, CA, CQ, IMC, IAC, RCEst e IC foram significativas ($p < 0,001$) em relação a presença ou não de SM nos quilombolas. Ao serem analisadas para compor o modelo preditivo, estas mesmas variáveis apresentaram-se estatisticamente significativas.

Após a identificação das variáveis significativas para a elaboração do modelo preditivo em relação à presença de SM, as variáveis IAC, IMC, CC, CP, CA foram escolhidas pelo algoritmo para compor a árvore decisória.

A árvore de decisão (Figura 1) apresentou uma acurácia de $(11 + 1) / (11 + 1 + 3 + 1) = 0.75$ (75%), sensibilidade de $1 / (1 + 1) = 0.50$ (50%) e especificidade de $11 / (11 + 3) = 0.79$ (79%), tendo como nó raiz a CA e como nós terminais CC, CP, IAC e IMC sendo escolhida para gerar o gráfico em relação ao processo de decisão, ilustrada por meio de um fluxograma (Figura 2).

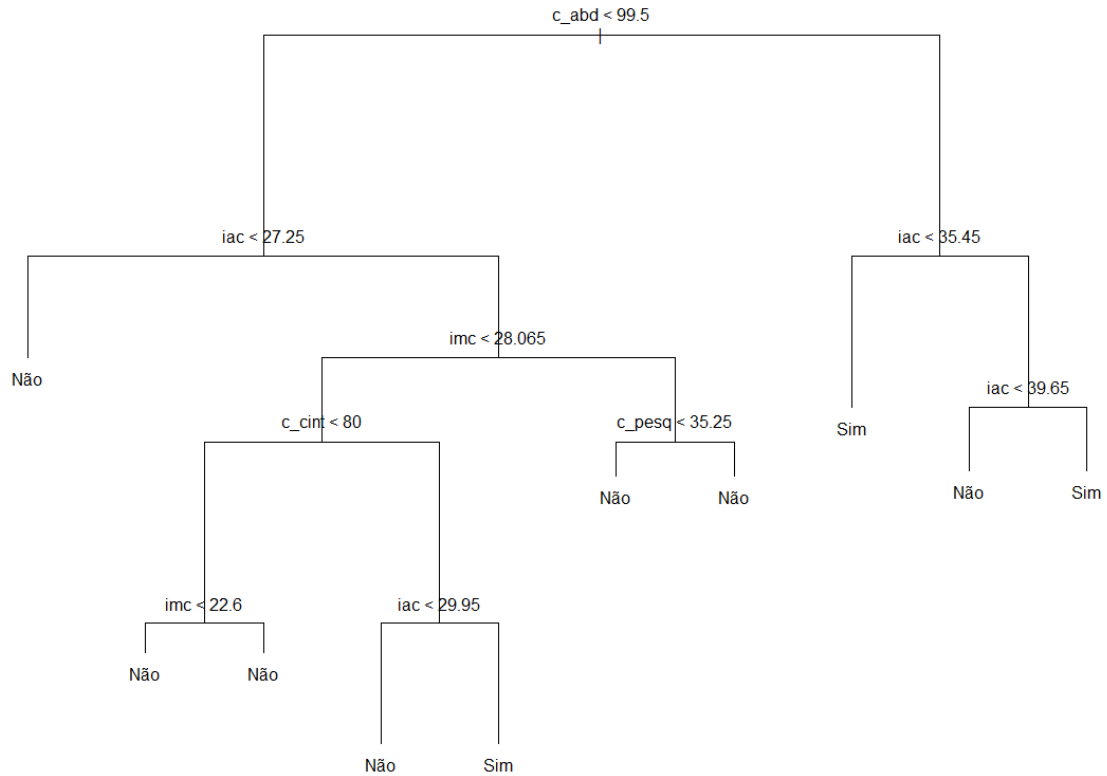


Figura 1. Árvore de decisão gerada pelo algoritmo

A árvore foi composta por variáveis decisórias em um modelo hierárquico, a partir do nó raiz, que conduz das regras decisórias para os nós internos e aos seus ramos, seguindo em direção aos nós terminais com as decisões [28]. As variáveis utilizadas na construção da árvore em relação a ter ou não SM (Figura 2) pelo algoritmo foram IAC, IMC, CC, CP, CA e CQ, apresentando 10 nós terminais, Desvio residual médio (*Residual mean deviance*) de 0.6024 (30.12/50), taxa de erro na classificação (*Misclassification error rate*) de 0.1167 (7/ 60).

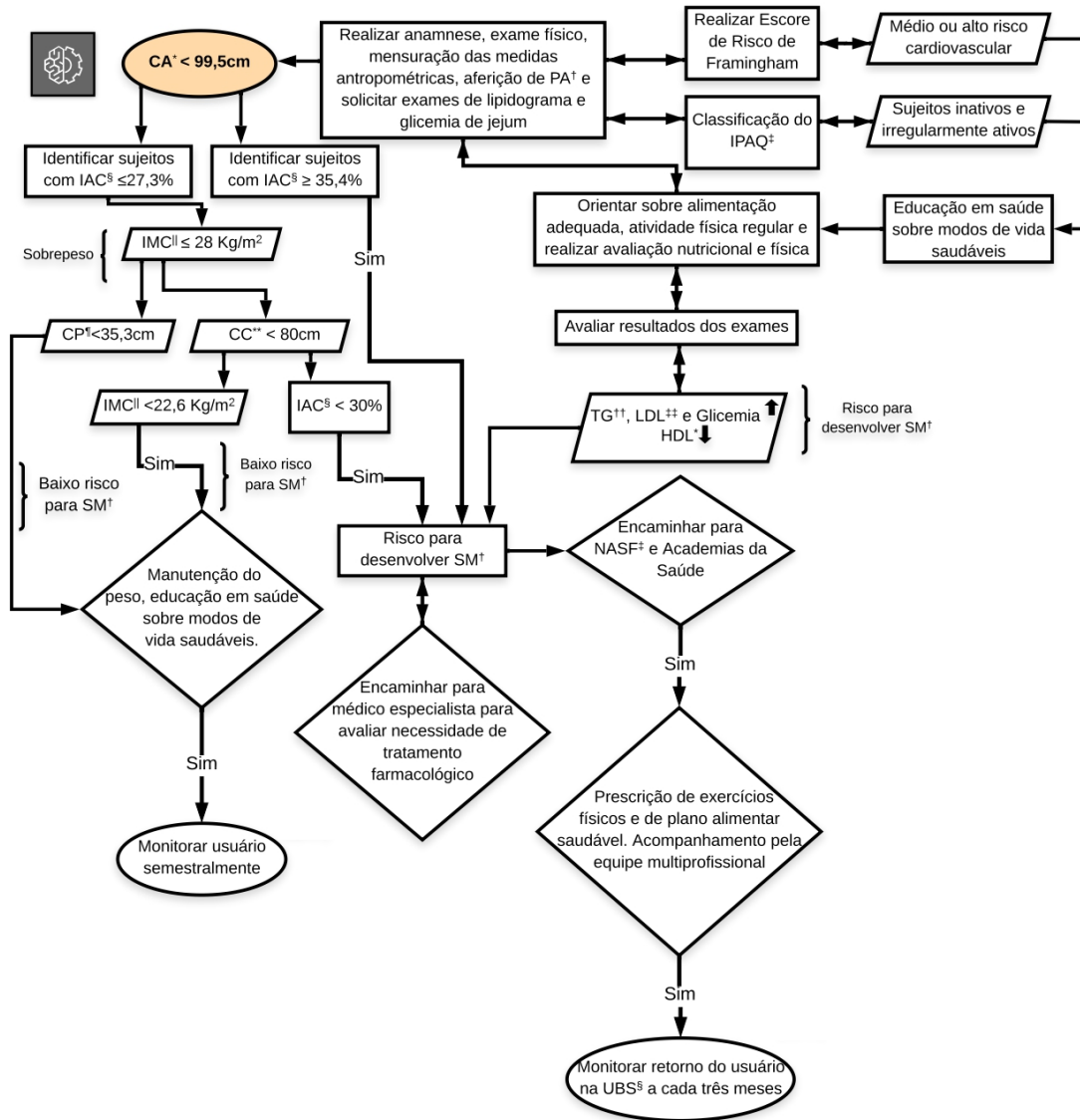


Figura 2. Fluxograma de árvore de decisão para predição de Síndrome Metabólica em quilombolas - Capela, SE, Brasil. 2022

Legenda:

- = Indica início e final do processo
- = Indica ação/processo
- = Dados
- = Indica tomada de decisão
- = Indica as direções a serem seguidas

}= Nota

= Machine learning

*CA - Circunferência abdominal; †PA - Pressão Arterial; ‡IPAQ - International Physical Activity Questionnaire; §IAC - Índice de Adiposidade Corporal; ||IMC - Índice de Massa Corporal; ¶CP - Circunferência de pescoço; **CC - Circunferência de cintura; ††TG - Triglicerídeos; ‡‡LDL - Low Density Lipoproteins; *HDL - High Density Lipoproteins; †SM - Síndrome Metabólica; ‡NASF - Núcleo de Apoio à Saúde da Família; §UBS - Unidade Básica de Saúde

Para facilitar a compreensão e aplicação do fluxograma, foram elencadas 14 regras de decisão (Quadro 1) que evidenciam a relevância das variáveis em cada etapa do modelo preditivo para o atendimento em saúde de quilombolas com chances de desenvolvimento da SM. As variáveis mais próximas ao nó raiz da árvore têm maior relevância na distinção e classificação dos grupos dos quilombolas usuários da rede de serviços de atenção à saúde. A variável CA é a principal para a tomada de decisão e responsável pelo desfecho (ter ou não SM).

Quadro 1: Regras de decisão construídas a partir da Arvore Decisória para Predição de SM em quilombolas.

Regras de decisão		
Atributo - condição necessária (Se)		Decisão - (Então) é o resultado obtido
1	Se realizar anamnese, exame físico, mensuração das medidas antropométricas e PA*, solicitar exames de lipidograma e glicemia de jejum	Então deve-se realizar escore de risco de Framingham e classificação do IPAQ†
2	Se sujeitos com médio ou alto risco cardiovascular e inativos ou irregularmente ativos	Então deve-se realizar educação em saúde sobre modos de vida saudáveis, orientar sobre alimentação adequada, atividade física regular e realizar avaliação nutricional e física
3	Se orientar sobre alimentação adequada, atividade física regular e realizar avaliação nutricional e física	Então deve-se avaliar paralelamente os resultados dos exames laboratoriais
4	Se TG‡, LDL§ e Glicemia de jejum aumentados e HDL baixo	Então sujeito possui risco de desenvolver SM¶. Encaminhar para médico especialista, NASF** e academias da Saúde.
5	Se Encaminhar para NASF** e academias da Saúde.	Então deve-se prescrever exercícios físicos e plano alimentar saudável e sujeito deve ser acompanhado por equipe multiprofissional com retorno na UBS†† a cada três meses
6	Se realizar anamnese, exame físico, mensuração das medidas antropométricas e PA*, solicitar exames de lipidograma e glicemia de jejum	Então deve-se identificar sujeitos com CA‡‡ <99,5 cm
7	Se CA‡‡ <99,5 cm	Então deve-se verificar a classificação do IAC*
8	Se IAC* ≤35,4%	Então sujeito possui risco de desenvolver SM¶. Encaminhar para médico especialista, NASF** e academias da Saúde.
9	Se Encaminhar para NASF** e academias da Saúde.	Então deve-se prescrever exercícios físicos e plano alimentar saudável e sujeito deve ser acompanhado por equipe multiprofissional com retorno na UBS†† a cada três meses
10	Se CA‡‡ <99,5 cm e IAC* ≤27,3%	Então deve-se identificar sujeitos com IMC† ≤28 kg/m ²
11	Se CP‡ ≤35,3 cm	Então sujeito possui baixo risco para desenvolver SM¶. Deve-se realizar a manutenção do peso, educação em saúde sobre modos de vida saudáveis e monitorar sujeito semestralmente
12	Se IMC† ≤28 kg/m ² e CC§ < 80cm	Então deve-se verificar o IAC*
13	Se IAC* <30%	Então sujeito possui risco de desenvolver SM¶. Encaminhar para médico especialista, NASF** e academias da Saúde.
14	Se Encaminhar para NASF** e academias da Saúde.	Então deve-se prescrever exercícios físicos e plano alimentar saudável e sujeito deve ser acompanhado por equipe multiprofissional com retorno na UBS†† a cada três meses

*PA - Pressão Arterial; †IPAQ- International Physical Activity Questionnaire; ‡TG - Triglicerídeos; §LDL - Low Density Lipoproteins; ||HDL - High Density Lipoproteins; ¶SM - Síndrome Metabólica; **NASF - Núcleo de Apoio à Saúde da Família; ††UBS - Unidade Básica de Saúde; ‡‡CA - Circunferência abdominal; *IAC - Índice de Adiposidade Corporal; †IMC - Índice de Massa Corporal; ‡CP - Circunferência de pescoço; §CC - Circunferência de cintura

4. DISCUSSÃO

O modelo para Predição de SM em Quilombolas construído demonstrou as variáveis independentes que conduzem as decisões inerentes à cada condição presente na árvore decisória. Foi assim destacada a importância da mensuração dos índices antropométricos e bioquímicos nos quilombolas para predição de SM, visto que se trata de um recurso de fácil execução e de baixo custo.

A utilização de métodos simples e baratos como a mensuração de indicadores de obesidade permitem a avaliação de um maior número de indivíduos contribuindo para o rastreamento de SM em diferentes grupos étnico-raciais, visto que a literatura científica ainda é escassa em relação à predição de SM para esta população [17].

A CA como nó raiz do modelo preditivo da Árvore de decisão para Predição de SM em Quilombolas, configurou-se como a variável mais importante para as tomadas de decisão em relação às probabilidades de desenvolvimento de SM, evidenciando a relevância da mensuração deste indicador para a predição de risco cardiovascular e metabólico, corroborando com estudos que afirmaram que o excesso de gordura abdominal significa um importante fator de risco cardiovascular [12].

O IAC em suas três classes foi importante na construção do modelo preditivo, visto que se encontra próximo ao nó raiz da árvore de decisão. O IAC tem sido utilizado na predição de alterações metabólicas e cardiovasculares, inclusive na população negra [16, 29, 30]. A presente pesquisa apresentou resultados similares a outro estudo em relação à capacidade de predição de SM por meio do IAC que constatou que indivíduos colombianos com níveis mais altos de IAC apresentaram obesidade positivamente associada a componentes da SM, dentre eles HDL-c ≤ 40 mg/dL, CC ≥ 90 cm, triglicérides ≥ 150 mg/dL, Colesterol Total (CT) ≥ 200 mg/dL e elevação dos níveis de LDL [31].

Em relação a RCEst, no presente estudo, apesar desta ser uma variável estatisticamente significativa para compor o modelo preditivo, na árvore decisória não se mostrou importante para prever as chances de desenvolvimento de SM.

A capacidade preditiva de SM por meio de indicadores de obesidade e o percentual de gordura (%G) por bioimpedância em adultos quilombolas foi estudada no Estado da Bahia, nordeste brasileiro, evidenciando que os indicadores %G, IAC e IMC apresentaram capacidade satisfatória para a predição da SM em homens e mulheres neste grupo populacional, com maior acurácia do percentual de gordura [17].

Variáveis antropométricas e fatores de risco cardiovascular foram analisados em quilombolas de ambos os sexos no norte do Brasil. Foi constatado que IMC e CC em conjunto são ótimos indicadores de obesidade, e quando aumentados configuram-se como fatores de risco cardiovascular [31]. O perfil destes quilombolas foi semelhante ao dos quilombolas sergipanos, pois as medidas de CC e IMC demonstraram risco cardiovascular e foram importantes componentes do modelo preditivo, apresentando-se entre as variáveis a serem consideradas para as tomadas de decisão, em relação às probabilidades de desenvolvimento de SM.

A avaliação da CP também se mostrou importante para a tomada de decisão em relação ao risco de desenvolver SM, uma vez que a referida medida obteve proximidade aos nós mais próximos da raiz da árvore de decisão do modelo proposto. Os resultados deste estudo corroboram com pesquisadores que afirmaram que a CP é um potencial preditor da SM [32].

Evidenciou-se que apenas as mulheres quilombolas apresentaram índices risco alto para o desenvolvimento de DCV em relação a CA. Pesquisadores demonstraram por meio de revisão sistemática, que as alterações mais expressivas em mulheres foram nas medidas de CC, assim como relataram uma maior influência da menopausa na ocorrência de SM independentemente da população [33].

Ao se analisar a idade média das mulheres quilombolas sergipanas do município de Capela, Brasil, sugere-se que elas se encontram no período da menopausa, o que pode favorecer a ocorrência de SM. Vale ressaltar que redução das concentrações de estrogênios no período pós-menopausa são condições que propiciam um aumento na gordura abdominal visceral e na dosagem de lipídeos séricos, o que está correlacionado com resistência à insulina, hipertensão e potencialização do risco cardiovascular em mulheres [34].

O predomínio de obesidade e sobrepeso já foi constatado na população quilombola no Brasil até mesmo associada a baixa escolaridade [31, 35]. O nível de escolaridade \leq ao ensino

fundamental incompleto e a raça/cor preta são fatores de risco para obesidade [36]. Este perfil assemelha-se com a situação dos quilombolas avaliados, pois os resultados do IMC mostraram, em sua maioria, excesso de peso e obesidade. Constatou-se que a maioria dos sujeitos que não possui SM foram classificados pelo IPAQ como ativos, condição que contribui para a prevenção de SM.

No Estado do Piauí, região nordeste do Brasil, quilombolas de ambos os sexos foram estudados para verificar a presença de SM por meio da avaliação das medidas antropométricas, diário alimentar, modos de vida (uso de fumo, ingestão de bebida alcoólica e prática regular de atividade física), avaliação de lipidograma e glicemia de jejum, e constatou-se que a SM foi mais frequente entre os quilombolas obesos ou com sobrepeso, sendo predominante em mulheres, e menos frequente em quilombolas que praticavam atividade física regular. Diante desse achado quanto ao maior risco de SM em quilombolas com sobrepeso ou obesos e não praticantes de atividade física regular, os autores abordaram a necessidade de investimento em políticas públicas de saúde voltadas para a prevenção da SM em quilombolas [36].

Ressalta-se que modificações comportamentais visando modos de vida saudáveis são medidas eficientes para a prevenção de alguns fatores de risco para SM, o que podem contribuir para a redução dos custos de assistência médica [37]. Os resultados do presente estudo apontaram através dos relatos das entrevistas que a maioria dos quilombolas estudados não praticava atividade física regular e possuem insegurança alimentar moderada. Evidencia-se assim que o sedentarismo e a alimentação inadequada são condições que devem ser abordadas nas estratégias de prevenção de DCNT e SM no cenário Atenção Primária para os quilombolas.

Dentre as ações relevantes determinadas pelo modelo preditivo, encontra-se a educação sobre modos de vida saudáveis que reforçam a importância da prática regular de exercícios físicos e alimentação adequada para a prevenção de SM. Neste sentido a classificação do nível de atividade física dos sujeitos pode contribuir para as tomadas de decisão em relação ao encaminhamento deles para os Núcleos de Apoio à Saúde da Família (NASF) e polos do programa Academia da Saúde do Ministério da Saúde nas Unidades de Saúde da Família dos territórios referentes à população quilombola adscrita, bem como para a prescrição de exercícios físicos considerando as condições clínicas individuais.

O ERF é um instrumento relevante na predição das DCV, que através de variáveis clínicas e laboratoriais, estima o risco cardiovascular em busca do rastreamento de condição coronariana no período de 10 anos [22]. Destaca-se que os resultados do ERF foram significativos em relação às ter ou não SM (risco alto e intermediário), reforça a importância de intervenções educativas direcionadas para a prevenção das DCV nas comunidades quilombolas do município de Capela/SE, Brasil.

Cumprido salientar que a acurácia da árvore decisória reflete as melhores escolhas diante das chances de os sujeitos desenvolverem SM, entretanto, como todo modelo possui alguma fragilidade, podem ocorrer pequenas dispersões em relação aos dados das variáveis numéricas próximas aos nós terminais. Apesar de algumas variáveis previamente significativas não estarem presentes no fluxograma, não significa que não tenham relevância clínica, mas que no modelo de classificação da árvore pelo algoritmo, não tiveram consistência para provocar as decisões que refletissem uma boa acurácia.

A estratificação dos riscos cardiovasculares segundo o sexo não foi possível em função das características demográficas do universo populacional, entretanto, o modelo preditivo apresentou-se relevante e com acurácia que permite a predição do desfecho em 75% dos casos.

O predomínio de mulheres na amostra foi semelhante a outros estudos com populações quilombolas em diferentes regiões do Brasil [31, 36, 38, 39]. Provavelmente a maior participação de mulheres nestes estudos pode estar relacionada a disponibilidade de tempo e interesse para participarem das pesquisas ou que a população quilombola no Brasil seja predominantemente composta por mulheres. Contudo, dados oficiais que caracterizem as populações quilombolas no Brasil ainda estão em período de análise estatística de acordo com o Censo do IBGE [5]. Este fato poderá favorecer o melhor entendimento sobre distribuição por sexo e demais características sociodemográficas dos quilombolas brasileiros e fornecer melhores subsídios para a construção de políticas públicas mais assertivas para os quilombolas.

5. CONSIDERAÇÕES FINAIS

Constatou-se o predomínio de obesidade, sobrepeso e risco cardiovascular entre os quilombolas que podem desenvolver SM. O modelo Árvore de Decisão para Predição de Síndrome Metabólica se mostrou um artifício estatístico de rápida execução que pode contribuir antecipadamente para as tomadas de decisão específicas no atendimento de saúde na Atenção Primária aos quilombolas com chances de desenvolverem SM.

Neste modelo, a CA é um importante preditor da SM em quilombolas, desta forma, sugere-se que esta mensuração seja incluída no exame físico dos sujeitos durante o atendimento de saúde, bem como as demais medidas antropométricas que apresentaram capacidade preditiva na árvore de decisão. Sugere-se que a acurácia do modelo preditivo permitirá a aplicação do fluxograma em outras comunidades quilombolas, visto que as características socioculturais, econômicas e clínicas são similares nestas populações.

O fato de a população quilombola estudada ter apresentado fatores de risco para o desenvolvimento de SM e eventos cardiovasculares, bem como associação com a obesidade, reforça a importância de estratégias de prevenção de agravos, tais como educação em saúde sobre modos de vida saudáveis, no contexto da orientação nutricional, prática regular de exercícios físicos, não ingestão de bebidas alcoólicas em excesso e uso de fumo.

A Árvore de decisão para Predição de SM em Quilombolas permitiu identificar os pontos de corte específicos dos indicadores antropométricos mais importantes a serem mensurados no primeiro atendimento de saúde nesta população, evidenciando que a CA, CP, CC, o IAC e IMC avaliados em conjunto, possibilitam uma boa acurácia na predição de SM neste grupo populacional. A utilização destas medidas associadas, principalmente o uso do IAC e da CA é uma ferramenta tecnológica oportuna e facilitadora para tomada de decisão em saúde, visto que os dados sobre estas medidas na população negra ainda são escassos.

6. REFERÊNCIAS

1. Dorigny M. Atlas das escravidões: da Antiguidade até os nossos dias. Tradução de João Guilherme de Freitas; Cartografia de Fabrice le Golf. Petrópolis (RJ): Vozes; 2017.
2. Tanure RGA. Da necessidade da implementação de políticas públicas no combate ao racismo estrutural. *Boletim Científico ESMPU, Brasília*. 2021;20(57):265-84.
3. Bailey SR, Loveman M, Muniz JO. Measures of "Race" and the analysis of racial inequality in Brazil. *Soc Sci Res*. 2013;42(1):106-19. doi: 10.1016/j.ssresearch.2012.06.006
4. Brasil. Ministério da Saúde, Secretaria de Gestão Estratégica e Participativa, Departamento de Apoio à Gestão Participativa e ao Controle Social. Política Nacional de Saúde Integral da População Negra: Uma política do SUS. 3. ed. Brasília (DF): Ministério da Saúde; 2017. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/politica_nacional_saude_populacao_negra_3d.pdf.
5. Instituto Brasileiro de Geografia e Estatística (IBGE). Censo demográfico 2022 [Internet]; 2023 [acesso em 04 ago 2023]. Disponível em: <https://censo2022.ibge.gov.br/panorama/indicadores.html?loalidade=BR>
6. Brasil. Ministério da Cultura. Fundação Cultural dos Palmares (FCP). Certidões expedidas às comunidades remanescentes de quilombos (CRQs). Brasília (DF): Diário Oficial da União; 22 ago 2022. Disponível em: <https://www.palmares.gov.br/wp-content/uploads/2015/07/quadro-geral-por-estados-e-regioes-22-08-2022.pdf>
7. Brasil. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos, Departamento de Ciência e Tecnologia. Agenda de Prioridades de Pesquisa do Ministério da Saúde – APPMS [recurso eletrônico]. Brasília (DF): Ministério da Saúde; 2018. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/agenda_prioridades_pesquisa_ms.pdf
8. Rodrigues MC, Silva Maciel ES, Quaresma FRP, Sesti LFC, Paiva LS, Junior HM, et al. Prevalence and factors associated with metabolic syndrome in a vulnerable population in northern Brazil: a cross-sectional study. *J Hum Growth Dev*. 2021;31(2):291-301. doi: 10.36311/jhgd.v31.11410
9. Vitor AD, Alvarez TS. Health eating: preventing the metabolic syndrome. 2016;23(2):944-54. doi: 10.20396/san.v23i2.8647782
10. López-Jaramillo P, Sánchez RA, Diaz M, Cobos L, Bryce A, Parra-Carrillo JZ, et al. Latin American consensus on hypertension in patients with diabetes type 2 and metabolic syndrome. *Arq Bras Endocrinol Metab*. 2014;58(3):205-25. doi: 10.1590/0004-2730000003019

11. Maurer P, Gullich AAC, Retamoso VR, Freitas VER, Bruno JB, Berro LF, et al. Components for diagnosing Metabolic Syndrome using NCEP-ATP III in an Afro-Brazilian population. *Rev Bras Pesq Saúde*. 2016;18(4):55-60.
12. Associação Brasileira para o Estudo da Obesidade e da Síndrome Metabólica. *Diretrizes Brasileiras de Obesidade* 2016. 4. ed. São Paulo (SP): Abeso; 2016.
13. Kroll C, Mastroeni SSBS, Czarnobay SA, Ekwaru JP, Veugelers PJ, Mastroeni MF. The accuracy of neck circumference for assessing overweight and obesity: a systematic review and meta-analysis. *Ann Hum Biol*. 2017;44(8):667-77. doi: 10.1080/03014460.2017.1390153
14. Kahn HS, Bullard KM, Barker LE, Imperatore G. Differences between adiposity indicators for predicting all-cause mortality in a representative sample of United States non-elderly adults. *PLoS One*. 2012;7(11):1-11. doi: 10.1371/journal.pone.0050428
15. Lopez-Lopez JP, Cohen DD, Ney-Salazar D, Martinez D, Otero J, Gomez-Arbelaez D, et al. The prediction of Metabolic Syndrome alterations is improved by combining waist circumference and handgrip strength measurements compared to either alone. *Cardiovasc Diabetol*. 2021;20:68. doi: 10.1186/s12933-021-01256-z
16. Bergman RN, Stefanovski D, Buchanan TA, Sumner AE, Reynolds JC, Sebring NG, et al. A better index of body adiposity. *The Obesity Society*. 2011;19(5):1083-9. doi:10.1038/oby.2011.38
17. Mussi RFF, Petroski EL. Predictive capacity of obesity indicators for metabolic syndrome in adult quilombolas (inhabitants of black communities). *Ciência & Saúde Coletiva*. 2019;24(7):2471-80. doi: 10.1590/1413-81232018247.19032017
18. Mussi RFF, Petroski EL. Metabolic syndrome and associated factors in quilombolas (inhabitants of black communities) from Bahia, Brazil. *Ciênc Saúde Coletiva*. 2019;24(7):2481-90. doi: 10.1590/1413-81232018247.13982017
19. Dudi P, Goyal B, Saxena V, Rabari K, Mirza AA, Naithani M, et al. Single point insulin sensitivity estimator as an index for insulin sensitivity for metabolic syndrome: A study in North Indian population. *J Lab Physicians*. 2019;11(3):244-8. doi: 10.4103/JLP.JLP_163_18
20. Lobo LC. Artificial Intelligence and Medicine. *Rev Bras de Educação Médica*. 2017;41(2):185-93. doi: 10.1590/1981-52712015v41n2esp
21. Faludi AA, Izar MCO, Saraiva JFK, Chacra APM, Bianco HT, Afiune NA, et al. Atualização da diretriz brasileira de dislipidemias e prevenção da aterosclerose – 2017. *Arq Bras Cardiol*. 2017;109:1-76. doi: 10.5935/abc.20170121
22. Fernandes PV, Castro MM, Fuchs A, Machado MCR, Oliveira FD, Silva LB et al. Predictive value of the framingham risk score in identifying high cardiovascular risk. *Int J Cardiovasc Sci*. 2015;28(1):4-8. doi: 10.5935/2359-4802.20150002
23. Matsudo S, Araujo T, Matsudo V, Andrade D, Andrade E, Oliveira LC. et al. Questionário Internacional de Atividade Física (IPAQ): estudo de validade e reprodutibilidade no Brasil. *Revista brasileira Atividade Física e Saúde*. 2001;6(2): 5-18.
24. Santos LP, Lindemann IL, Motta JVS, Mintem G, Bender E, Gigante DP. Proposta de versão curta da Escala Brasileira de Insegurança Alimentar. *Revista de Saúde Pública*. 2014;48(5):783-789. doi:10.1590/S0034-8910.2014048005195
25. Ibrahim Q, Ahsan M. Measurement of visceral fat, abdominal circumference and waist-hip ratio to predict health risk in males and females. *Pak J Biol Sci*. 2019;22(4):168-73. doi: 10.3923/pjbs.2019.168.173. PMID: 31930817
26. Valdez RA. Simple model-based index of abdominal adiposity. *J. Clin. Epidemiol*. 1991;44(9):955-6. doi: 10.1016/0895-4356(91)90059-i
27. Revathy S, Bharathi B, Jeyanthi P, Ramesh M. Chronic kidney disease prediction using machine learning models. *Int J Eng Adv Technol (IJEAT)*. 2019;9(1):6364-7. doi: 10.35940/ijeat.A2213.109119
28. Medeiros LB, Trigueiro DRSG, Silva DM, Nascimento JÁ, Monroe AA, Nogueira JÁ, et al. Integration of health services in the care of people living with aids: an approach using a decision tree. *Ciência & Saúde Coletiva*. 2016;21(2):543-52. doi: 10.1590/1413-81232015212.06102015
29. Alvim RDE O, Mourão-Junior CA, de Oliveira CM, Krieger JE, Moinho JG, Pereira AC. Body mass index, waist circumference, body adiposity index, and risk for type 2 diabetes in two populations in Brazil: General and amerindian. *PLoS One*. 2014;9(6):1-6. doi: 10.1371/journal.pone.0100223
30. González-Ruiz K, Correa-Bautista JE, Ramírez-Vélez R. Body adiposity and its relationship of metabolic syndrome components in Colombian adults. *Nutr Hosp*. 2015;32(4):1468-75. doi: 10.3305/nh.2015.32.4.9164
31. Cordovil YF, Almeida SS. Anthropometric variables and associated cardiovascular risk factors in Quilombolas Marajoaras. *Rev Bra Obesidade, Nutr Emagr*. 2018;12(71):406-15.
32. Ataie-Jafari A, Namazi N, Djalalinia S, Chaghamirzayi P, Abdar ME, Zadehe SS, et al. Neck circumference and its association with cardiometabolic risk factors: a systematic review and meta-analysis. *Diabetol Metab Syndr*. 2018;29;10:72. doi: 10.1186/s13098-018-0373-y

33. Pucci G, Alcidi R, Tap L, Battista F, Mattace-Raso F. Sex- and gender-related prevalence, cardiovascular risk and therapeutic approach in metabolic syndrome: a review of the literature. *Pharmacol Res.* 2017;120:34-42. doi: 10.1016/j.phrs.2017.03.008
34. Soares DA, Kochergin CN. Factors associated with obesity in elderly maroons in Bahia, Brazil. *Rev APS.* 2017;20(2):174-84. doi: 10.34019/1809-8363.2017.v20.15768
35. Ferreira APS, Szwarcwald CL, Damacena GN. Prevalence of obesity and associated factors in the Brazilian population: a study of data from the 2013 National Health Survey. *Rev Bras Epidemiol.* 2019;22:1-14. doi: 10.1590/1980-549720190024
36. Aragão JA, Bós AJG, Souza GC. Síndrome metabólica em adultos e idosos de comunidades quilombolas do centro-sul do Piauí, Brasil. *Estud Interdiscipl Envelhec.* 2014;19(2):501-12.
37. Šoštarič A, Jenko B, Kozjek NR, Ovijač D, Šuput D, Milisav I, et al. Detection of metabolic syndrome burden in healthy young adults may enable timely introduction of disease prevention. *Arch Med Sci.* 2019;15(5):1184-94. doi: 10.5114/aoms.2019.87462
38. Oliveira SKM, Pereira MM, Guimarães ALS, Caldeira AP. Self-perceived health among 'quilombolas' in northern Minas Gerais, Brazil. *Cien Saude Colet.* 2015;29(9):2879-90. doi: 10.1590/1413-81232015209.20342014
39. Pauli S, Bairros FS, Nunes LN, Neutzling MB. Self-reported prevalence of hypertension and associated factors in quilombola communities in the State of Rio Grande do Sul, Brazil. *Ciênc Saúde Coletiva.* 2019;24(9):3293-303. doi: 1590/1413-81232018249.28002017