



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

**Associação entre Saúde Cardiovascular e
Determinantes Socioeconômicos: Uma Aplicação
de modelos de regressão logística multinomial e
ordinal**

por

Ana Carolina Souto Valente Motta

Brasília, 14 de novembro de 2022

Associação entre Saúde Cardiovascular e Determinantes Socioeconômicos: Uma Aplicação de modelos de regressão logística multinomial e ordinal

por

Ana Carolina Souto Valente Motta

Dissertação apresentada ao Departamento de Estatística da Universidade de Brasília, como requisito parcial para obtenção do título de Mestre em Estatística.

Orientadora: Profa. Dra. Joanlise Marco de Leon Andrade

Coorientadora: Profa. Dra. Kelb Bousquet-Santos

Brasília, 14 de novembro de 2022

Dissertação submetida ao Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade de Brasília como parte dos requisitos para a obtenção do título de Mestre em Estatística.

Texto aprovado por:

Profa. Joanlise Marco De Leon Andrade
Orientador, EST/UnB

Prof. Eduardo Yoshio Nakano
EST/UnB

Profa. Júlia Pavan Soler
IME/USP

*It is more important to know what sort of person has a disease
than to know what sort of disease a person has.*

(Hipócrates)

Com gratidão, dedico este trabalho a todos que
estiveram ao meu lado nessa caminhada.

Agradecimentos

A conclusão do mestrado em estatística é um grande sonho realizado e, por esse motivo, gostaria de agradecer a todos que me auxiliaram nessa jornada.

Agradeço, primeiramente, a UnB, ao departamento de Estatística e aos professores do PP-GEST/UnB, que realizam um trabalho excepcional e buscam engradecer a profissão. Em especial, meus sinceros agradecimentos às professoras Joanlise e Kelb, pela dedicação ao longo de todo o processo, por todo apoio, compreensão e conhecimentos compartilhados; e à minha colega de pesquisa Isabela pelo suporte.

Agradeço à minha família que sempre esteve ao meu lado e me apoiou nos momentos mais difíceis. Em especial, agradeço aos meus pais, Rogério e Renata, meus irmãos, Geovana e Rodrigo, e meu namorado, Pedro, que sempre estiveram torcendo por mim.

Agradeço também à Deus por todas as boas oportunidades que tive na vida e por ter colocado as pessoas certas para me acompanhar nessa caminhada. Por fim, gostaria de agradecer a todos os amigos que estiveram na minha vida durante esses anos e que me apoiaram de alguma forma.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Objetivo: Estimar a prevalência de Saúde Cardiovascular (SCV) Ideal na população adulta brasileira e avaliar a associação entre SCV e determinantes sociais com base na Pesquisa Nacional de Saúde 2019. **Métodos:** Inquérito de saúde de âmbito nacional (n=77.494). O escore de SCV proposto pela Associação Americana do Coração incluiu 4 metas comportamentais (tabagismo, índice de massa corpórea, exercício e dieta) e 3 metas biológicas (colesterol, pressão arterial e glicose). Prevalências (e intervalos de 95% de confiança) de SCV ideal e suas metas individuais foram estimadas utilizando expansão da amostra. Associações entre SCV e determinantes socioeconômicos (Escolaridade, Índice de riqueza e status de ocupação) foram avaliadas por modelos de regressão logística, ordinal e multinomial, ajustando-se por variáveis sociodemográficas. **Resultados:** Apenas 0,5% (IC95% 0,4;0,6) da população apresentou SCV Ideal (7 metas favoráveis) e 8,9% (IC95% 8,5;9,3), SCV superior (6-7 metas favoráveis), com pior desempenho em metas comportamentais. Escolaridade, riqueza e ocupação, além das covariáveis faixa etária, estado civil, presença de doenças crônicas, região e situação de domicílio apresentaram associação significativa com SCV. Modelos de regressão logística binária, multinomial e ordinal identificaram praticamente as mesmas variáveis independentes significantes, sendo o multinomial mais interessante clinicamente e o ordinal de difícil interpretabilidade e avaliação no contexto de amostragem complexa. **Conclusão:** A baixíssima prevalência de SCV Ideal e associações entre SCV e características sociodemográficas observadas na população adulta brasileira evidenciam a necessidade de políticas públicas para promoção, vigilância e atenção à SCV com intervenções mais direcionadas e eficazes visando um aumento na prevalência de

SCV.

Palavras-chave: Estudos Populacionais em Saúde Pública; Pesquisa Nacional de Saúde (PNS) 2019; Estudo de Prevalência; Escore de Saúde cardiovascular; Regressão logística multinomial.

Abstract

Objective: To estimate the prevalence of Ideal Cardiovascular Health (CVH) in the Brazilian adult population and to evaluate the association between CVH and social determinants based on the 2019 National Health Survey. **Methods:** Nationwide health survey (n=77,494). The CVH score proposed by the American Heart Association includes 4 behavioral metrics (smoking, body mass index, exercise, and diet) and 3 biological metrics (cholesterol, blood pressure, and glucose). Prevalences (and 95% confidence intervals) of ideal CVH and their individual metrics were estimated using sample expansion. Associations between CVH and socioeconomic determinants (Education, Wealth and Occupation Index) were evaluated by logistic, ordinal and multinomial regression models, adjusting for sociodemographic variables. **Results:** Only 0.5% (95%CI 0.4;0.6) of the population presented Ideal CVH (7 favorable metrics) and 8.9% (95%CI 8.5;9.3) presented superior CVH (6-7 favorable metrics), with worse performance in behavioral metrics. Education, wealth index and occupation status, in addition to the covariates age group, marital status, presence of chronic diseases, region, and urban-rural classification were significantly associated with CVH. Binary, multinomial and ordinal logistic regression models identified practically the same significant independent variables, with the multinomial being more interesting clinically and the ordinal being difficult to interpret and evaluate in the context of complex sampling. **Conclusion:** The very low prevalence of Ideal CVH and associations between CVH and sociodemographic characteristics observed in the Brazilian adult population highlight the need for public policies to promote, monitor and care for CVH with more targeted and effective interventions to increase the prevalence of CVH.

Keywords: Public Health Population-based studies, National Health Survey 2019, Prevalence study, ideal cardiovascular health, Life's Simple 7th, Multinomial logistic regression.

Sumário

1	Introdução Geral	1
2	Artigo 1: Prevalência de saúde cardiovascular ideal na população adulta brasileira - Pesquisa Nacional de Saúde 2019	3
3	Artigo 2: Associação entre saúde cardiovascular e determinantes socioeconômicos na população adulta brasileira – Pesquisa Nacional de Saúde 2019	31
4	Conclusão Geral	63
A	Metodologia	65
A.1	Regressão Logística para amostragem complexa	65
A.2	Regressão Logística Multinomial	66
A.3	Regressão Logística Ordinal	69
A.3.1	Regressão Logística Ordinal com função logito cumulativo	69
A.3.1.1	Logito Cumulativo com Chances Proporcionais	70
A.3.1.2	Teste de hipóteses para a suposição de chances proporcionais	72
A.3.1.3	Logito Cumulativo sem Chances Proporcionais	73
A.3.1.4	Qualidade de ajuste	73
A.3.2	Regressão Logística com Logito de Categorias Adjacentes	74
A.3.2.1	Conexão com o Logito de Categoria de Referência	76

A.3.3	Regressão Logística com logito de Razão Contínua	76
A.3.4	Modelo de Esteriótipo	77
A.3.5	Referências	79
B	Resultados de modelos de regressão logística ordinal	80
C	Construção do Índice de Riqueza	84
D	Códigos de <i>software</i> R e SAS utilizados	87
D.1	Códigos em linguagem R	87
D.2	Códigos em linguagem SAS	103

Lista de Tabelas

2.1	Características dos indivíduos brasileiros adultos participantes, PNS 2019 (n = 77.494).	26
2.2	Prevalência das metas de saúde cardiovascular, com respectivos intervalos de 95% de confiança, estimados para a população adulta brasileira em 2019 por expansão da amostra (Dados da PNS 2019).(N=77.494)	28
3.1	Distribuições de variáveis sociodemográficas por categorias de saúde cardiovascular – PNS 2019.	55
3.2	Razões de chances, IC95% e p-valores de modelos de regressão logística para a variável resposta saúde cardiovascular ideal (score de SCV ≥ 6 versus demais categorias) – PNS 2019.	57
3.3	Razões de chances, IC95% e p-valores de modelos de regressão multinomial para a variável resposta SCV ideal em 3 categorias (inferior 0-2, intermediária 3-5 e superior ou ideal 6-7) – PNS 2019.	58
3.4	Razões de chances, IC95% e p-valores de modelos de regressão multinomial, minimamente e moderadamente ajustados, para a variável resposta SCVa ideal em 3 categorias (inferior 0-2, intermediária 3-5 e superior ou ideal 6-7) – PNSb 2019.	60

B.1	Razões de chances, IC95% e p-valores de modelos de regressão logística ordinal com logito cumulativo e chances proporcionais para a variável resposta SCV ^a em 3 categorias (inferior 0-2, intermediária 3-5 e superior ou ideal 6-7) – PNS ^b 2019	82
B.2	Razões de chances, IC95% e p-valores de modelos de regressão logística ordinal com logito cumulativo e chances proporcionais para a variável resposta SCV ^a ideal em 8 categorias (0-7 metas atingidas) – PNS ^b 2019	83
C.1	Variáveis utilizadas na construção do índice de riqueza	85

Lista de Figuras

2.1	Metas de saúde cardiovascular.	25
2.2	Distribuição do Escore de SCVa com respectivos Intervalos de 95% de Confi- ança, estimados para a população adulta brasileira em 2019 (Dados da PNS 2019).	27
2.3	Construção e métodos de cálculo das variáveis, Pesquisa Nacional de Saúde 2019, Brasil	30
3.1	Metas de saúde cardiovascular.	53
3.2	Distribuição do Escore de SCV com respectivos Intervalos de 95% de Confi- ança, estimados para a população adulta brasileira em 2019 (Dados da PNS 2019).	54
3.3	Construção e métodos de cálculo das variáveis, Pesquisa Nacional de Saúde 2019, Brasil	59

Abreviações e Siglas

SCV	Saúde Cardiovascular
PNS	Pesquisa Nacional de Saúde
DCV	Doenças Cardiovasculares
IBGE	Instituto Brasileiro de Geografia e Estatística
PNAD	Pesquisa Nacional de Amostra por Domicílios
UPA	Unidades Primárias de Amostragem
n	tamanho de uma amostra
IMC	Índice de Massa Corporal
IC95%	Intervalo de Confiança de 95%
RC	Razão de Chances
p-valor	nível descritivo de um teste de hipóteses
log	Logaritmo
exp	Exponencial
MV	Máxima Verossimilhança
EMV	Estimador de Máxima Verossimilhança
DP	Desvio-padrão
χ^2	Qui-quadrado
gl	Grau de liberdade
AIC	Critério de informação Akaike

Capítulo 1

Introdução Geral

O presente trabalho envolve a avaliação da associação entre determinantes socioeconômicos e um escore de saúde cardiovascular (SCV) ideal, proposto pela *American Heart Association*, na população adulta brasileira, utilizando dados da Pesquisa Nacional de Saúde (PNS) de 2019. O escore consiste na soma de quantas de 7 metas de comportamentos e fatores de SCV são atingidas pelo indivíduo, variando assim de 0 a 7. Tal escore também será utilizado de forma binária, considerando como favorável 6 ou 7 metas atingidas, ou de forma categorizada em 3 níveis de SCV: inferior (0 a 2 metas atingidas), intermediário (3 a 5 metas atingidas) e superior (6 ou 7 metas atingidas).

O capítulo 2 consiste em um artigo descritivo no qual foram calculadas prevalências de saúde cardiovascular ideal e de seus componentes na população brasileira adulta e em subgrupos sociodemográficos com base em dados da PNS de 2019.

O capítulo 3 apresenta um artigo cujo objetivo envolve a avaliação de associações entre SCV (em duas e três categorias) e determinantes socioeconômicos, ajustando-se ainda por variáveis sociodemográficas em modelos de regressão logística binária e multinomial. Para tanto, utilizou-se também dados da PNS de 2019.

O Apêndice A descreve de forma mais detalhada diferentes modelos de regressão logística multinomial e ordinal. O Apêndice B apresenta os resultados da análise de regressão logística

ordinal com logito cumulativo e chances proporcionais para a variável resposta SCV (em três categorias) e as mesmas variáveis e covariáveis explicativas utilizadas no segundo artigo (capítulo 3). Finalmente, o Apêndice C explica a metodologia para a construção do Índice de riqueza, um dos determinantes socioeconômicos utilizados nos modelos apresentados no segundo artigo e no Apêndice B.

Capítulo 2

Artigo 1: Prevalência de saúde cardiovascular ideal na população adulta brasileira - Pesquisa Nacional de Saúde 2019

Artigo submetido e em processo de revisão no periódico **Epidemiologia e Serviços de Saúde** (<https://ress.iec.gov.br/p/home>).

ARTIGO ORIGINAL

**Prevalência de saúde cardiovascular ideal na população adulta brasileira –
Pesquisa Nacional de Saúde 2019**

Prevalence of ideal cardiovascular health in the Brazilian adult population – National
Health Survey 2019

Prevalencia de salud cardiovascular ideal en la población adulta brasileña – Encuesta
Nacional de Salud 2019

Título resumido em português: Saúde cardiovascular ideal em adultos no Brasil

Ana Carolina Souto Valente Motta¹ - orcid.org/0000-0001-9016-4794

Kelb Bousquet-Santos² - orcid.org/0000-0002-4687-4681

Isabela Harumi Lopes Motoki¹ – orcid.org/0000-0002-8316-921X

Joanlise Marco De Leon Andrade¹ - orcid.org/0000-0002-3984-3799

¹ Universidade de Brasília, Departamento de Estatística, Brasília, DF, Brasil.

² Universidade de Brasília, Colegiado de Bases Biológicas e da Saúde, Brasília, DF,
Brasil.

CORRESPONDÊNCIA

Ana Carolina Souto Valente Motta | e-mail: anacarolina_souto@hotmail.com |

Endereço: Universidade de Brasília – Prédio CIC/EST – Campus Darcy Ribeiro – Asa Norte, Brasília – DF. 70910-900.

FINANCIAMENTO

Não se aplica.

TRABALHO ACADÊMICO ASSOCIADO

Artigo derivado de dissertação de mestrado intitulada “Associação entre Saúde Cardiovascular e Determinantes Socioeconômicos: Uma Aplicação de modelos de regressão logística multinomial e ordinal”, a ser defendida por ‘Ana Carolina Souto Valente Motta’ no Programa de Pós-graduação em Estatística, da Universidade de Brasília, em novembro de 2022.

CONFLITOS DE INTERESSE

Os autores declararam não possuir conflitos de interesse.

CONTRIBUIÇÃO DOS AUTORES

Todos os autores contribuíram de forma significativa no desenho do estudo, na análise e interpretação dos dados e na elaboração da versão preliminar do manuscrito. Motta ACSV e Motoki IHL contribuíram na análise e interpretação dos dados, redação e revisão crítica do conteúdo do manuscrito. Bousquet-Santos KB e Andrade JML contribuíram na concepção e delineamento do estudo, análise e interpretação dos resultados, redação e revisão crítica do conteúdo do manuscrito. Todos os autores aprovaram a versão final do

manuscrito e são responsáveis por todos os seus aspectos, incluindo a garantia de sua precisão e integridade.

AGRADECIMENTO

Não se aplica.

RESUMO

Objetivo: Analisar a prevalência de Saúde Cardiovascular (SCV) ideal na população adulta brasileira com base na Pesquisa Nacional de Saúde 2019. **Métodos:** Estudo transversal de base populacional (n=77.494). Conforme proposto pela Associação Americana do Coração, foram estimadas as prevalências, e respectivos intervalos de confiança de 95% (IC95%) de SCV ideal de forma global (sete metas alcançadas simultaneamente) e por metas individuais (quatro metas comportamentais e três biológicas). As prevalências também foram estimadas por características sociodemográficas. **Resultados:** Apenas 0,5% (IC95% 0,4;0,6) da população apresentou SCV Ideal, observando-se maior prevalência entre aqueles com maior escolaridade (1,3%; IC95% 0,9;1,6) e entre residentes em áreas urbanas (0,6%; IC95% 0,5;0,7). A prevalência das metas comportamentais e biológicas foi 0,7% (IC95% 0,6;0,8) e 63,3% (IC95% 62,7;63,9), respectivamente. **Conclusão:** A prevalência de SCV ideal foi muito baixa, evidenciando a necessidade de políticas públicas para promoção, vigilância e atenção à SCV na população adulta brasileira.

Palavras-chave: Estudos Populacionais em Saúde Pública; Inquérito Nacional de Saúde; Fatores de Risco de Doenças Cardiovasculares; Desigualdade em Saúde; Taxa de Prevalência; Epidemiologia Descritiva; Estudos Transversais.

ABSTRACT

Objective: To analyze the prevalence of Ideal Cardiovascular Health (CVH) in the Brazilian adult population based on the 2019 National Health Survey. **Methods:** Population-based cross-sectional study (N=77,494). As proposed by the American Heart Association, the prevalence and 95% confidence intervals (95%CI) of ideal CVH were estimated globally (7 metrics achieved simultaneously) and by individual metrics (four behavioral and three biological). Prevalence was also estimated by sociodemographic characteristics. **Results:** Only 0.5% (0.4;0.6) of the population presented ideal CVH, with a higher prevalence those with higher education (1.3%; 95%CI 0.9;1.6) and those living in urban areas (0.6%; 0.5;0.7). The prevalence of behavioral and biological metrics was 0.7% (0.6;0.8) and 63.3% (62.7;63.9), respectively. **Conclusion:** The prevalence of ideal CVH was very low, which highlights the importance of public policies for the promotion, surveillance and attention to CVH in the Brazilian adult population.

Keywords: Population Studies in Public Health; Health Survey; Cardiovascular Risk Factors; Health Status Disparity; Prevalence; Epidemiology, Descriptive; Cross-Sectional Studies.

RESUMÉN

Objetivo: Analizar la prevalencia de Salud Cardiovascular (SCV) Ideal en la población adulta brasileña con base en la Encuesta Nacional de Salud de 2019. **Métodos:** Estudio transversal de base poblacional ($n = 77,495$). Según lo propuesto por la Asociación Americana del Corazón, la prevalencia y los intervalos de confianza del 95% (IC95%) del SCV ideal se estimaron globalmente (siete metas alcanzadas simultáneamente) y por metas individuales (cuatro metas de comportamiento y tres metas). Se calculó la prevalencia de las metas por variables sociodemográficas. **Resultados:** Sólo el 0,5% (0,4;0,6) de la población presentó SCV ideal, siendo más frecuente en aquellos con educación superior (1,3%; IC95% 0,9;1,6) y entre los residentes de áreas urbanas (0,6%; IC95% 0,5;0,7). La prevalencia de las metas de comportamiento y biológicas fue 0,7% (0,6;0,8) y 63,3% (62,7;63,9), respectivamente. **Conclusión:** La prevalencia de SCV Ideal en adultos brasileños es muy baja, destacando la necesidad de políticas públicas para la promoción, vigilancia y atención a la SCV en la población adulta brasileña.

Palabras clave: Estudios Poblacionales en Salud Pública; Encuesta de Salud; Factores de Riesgo de Enfermedad Cardiovascular; Desigualdad en Salud; Tasa de Prevalencia; Epidemiología Descriptiva; Estudios Transversales.

Contribuições do estudo

Principais resultados

A prevalência de saúde cardiovascular (SCV) ideal na população brasileira foi estimada em 0,5% em 2019. As metas comportamentais e biológicas de SCV apresentaram prevalência de 0,7% e 63,3%, respectivamente. De todas as metas de SCV, a meta de dieta foi a que apresentou menor prevalência na população estudada (4,3%).

Implicações para os serviços

Os resultados obtidos podem contribuir na elaboração de políticas públicas para promoção, vigilância e atenção à no Brasil. As estimativas de prevalência das metas de SCV apontam para quais metas e segmentos populacionais os esforços deverão ser direcionados.

Perspectivas

Temas para estudos futuros incluem a estimação de prevalências de metas de SCV para subgrupos da população, a identificação de variáveis demográficas e socioeconômicas que afetem tais prevalências e a avaliação de novos critérios de SCV.

INTRODUÇÃO

As doenças cardiovasculares (DCV) constituem a principal causa de morte no mundo e no Brasil, correspondendo a um terço do total de óbitos.¹ Em 2019, as DCV responderam por 27,0% do total de mortes no Brasil, com a maior proporção na região Sudeste e a menor na região Norte.¹ Embora a taxa de mortalidade por DCV ajustada por idade tenha diminuído no Brasil entre 1990 e 2019, o número total de mortes por DCV aumentou. Isto se deve ao aumento da expectativa de vida da população e ao processo de globalização e urbanização, os quais por sua vez promoveram mudanças no estilo de vida e resultaram em maior prevalência de maus hábitos alimentares, obesidade e sedentarismo.² Os fatores de risco modificáveis são responsáveis por 70% das mortes por DCV.³ A dieta inadequada, a obesidade abdominal, sedentarismo e tabagismo representam os fatores de risco comportamentais, enquanto os fatores metabólicos incluem níveis de colesterol elevados, hipertensão arterial e diabetes. Em vários países, a identificação dos fatores de risco para DCV que apresentam maior prevalência na população permitiu que programas de prevenção cardiovascular conseguissem reduzir de forma expressiva a mortalidade por DCV.⁴ No Brasil, diferentes fontes de dados possibilitam o monitoramento contínuo dos fatores de risco para DCV, que incluem o Sistema de Informações sobre Mortalidade (SIM), o Sistema de Informações Hospitalares (SIH), e as pesquisas de saúde periódicas, como a Pesquisa Nacional de Saúde (PNS).²

Em 2010, a Associação Americana do Coração (*American Heart Association*) propôs o conceito de Saúde Cardiovascular (SCV) ideal com o objetivo de melhorar a saúde cardiovascular da população americana e reduzir a mortalidade por DCV.⁵ Tendo como premissa a prevenção primordial, as sete metas que compõem o escore de SCV

estão divididas em três metas biológicas, que incluem níveis adequados de pressão arterial, colesterol e glicose sérica, e quatro metas comportamentais, que incluem dieta adequada, prática de atividade física, peso corporal adequado e não fumar. Desde a sua criação, em 2010, o escore de SCV vem sendo amplamente utilizado na literatura científica internacional em associação com desfechos cardiovasculares ou outros tipos de desfechos.⁶⁻⁸

A prevalência de SCV ideal na população brasileira em 2013 foi inferior a 1%, sendo menor em homens, nos indivíduos com 60 anos ou mais e entre residentes da região Sul do país.⁹ Resultados semelhantes foram obtidos em estudos internacionais,^{10,11} que mostraram uma prevalência de SCV menor que 1% e nula em países do Cone Sul (Argentina, Chile e Uruguai) e Estados Unidos, respectivamente.¹²⁻¹⁴ Considerando a prevalência muito baixa de SCV ideal no Brasil, em 2013, e que o escore de SCV ideal demonstrou ser um instrumento eficiente de vigilância cardiovascular, conclui-se que o acompanhamento de indicadores de SCV, representados pelas metas biológicas e comportamentais do escore de SCV ideal, representa uma estratégia necessária para prevenção primordial das DCV. No entanto, não se tem conhecimento até o momento de estudos recentes que avaliaram a prevalência de SCV ideal em uma amostra representativa da população brasileira. Neste sentido, o objetivo do presente estudo foi analisar a prevalência de SCV ideal na população adulta brasileira.

MÉTODOS

Desenho do estudo

Estudo com um delineamento transversal, utilizando dados da PNS de 2019, empregado para determinar prevalências de indivíduos com SCV ideal na população brasileira adulta. As prevalências e seus respectivos intervalos de confiança de 95% foram calculados para as metas de SCV e também foram apresentadas por variáveis de características sociodemográficas.

Contexto

A Pesquisa Nacional de Saúde (PNS) é um inquérito de saúde de base domiciliar, de âmbito nacional, realizada pelo Ministério da Saúde (MS) em parceria com o Instituto Brasileiro de Geografia e Estatística (IBGE), nos anos de 2013 e 2019.¹⁵ A população pesquisada corresponde aos moradores de domicílios particulares permanentes do Brasil, exceto os localizados nos setores censitários especiais.¹⁵ A amostra da PNS é uma subamostra da Amostra Mestra da Pesquisa Nacional de Amostra por Domicílios (PNAD), com avanços tanto no espalhamento geográfico como na precisão das estimativas.¹⁵

Para a seleção da amostra, primeiramente selecionou-se as unidades primárias de amostragem (UPAs), por amostragem aleatória simples, mantendo-se a estratificação da Amostra Mestra da Pesquisa Nacional por Amostra de Domicílios (PNAD). Dessa forma, uma quantidade fixa de domicílios particulares permanentes foi selecionada aleatoriamente em cada UPA, a partir do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE), totalizando 108.525 domicílios. Por fim, um morador (com idade de 15 anos ou mais) foi selecionado em cada domicílio da amostra, a partir de uma lista

de moradores elegíveis construída no momento da entrevista, para responder à entrevista individual. A amostra final teve um total de 94.114 entrevistas individuais realizadas. Após preencher uma lista de todos os indivíduos residentes no domicílio, o entrevistador identificou o morador que proveria as informações sobre os questionários domiciliar e de todos os moradores do domicílio, além do sorteio do morador de 15 anos e mais para responder à entrevista individual. O questionário do morador selecionado continha questões relacionadas ao estilo de vida, doenças crônicas não transmissíveis, dentre outros. Maior detalhamento sobre o delineamento da pesquisa pode ser encontrado na literatura.^{15,16}

O cálculo do tamanho da amostra de domicílios e de pessoas levou em consideração: os vários indicadores de interesse, a estimação de proporções com nível de precisão desejado em intervalos de 95% de confiança, o efeito do plano amostral, o número de domicílios selecionados por UPA, a proporção de domicílios com pessoas na faixa etária de interesse e a possibilidade de estimação dos principais indicadores em níveis estaduais e de capitais.¹⁵

Participantes

Os critérios de exclusão do presente estudo foram (1) idade inferior a 18 anos completos, sendo esse critério também considerado para os fatores de expansão; (2) entrevistas individuais não realizadas; (3) questionários individuais respondidos por terceiros; (4) mulheres grávidas ou que não souberam informar se estavam grávidas e (5) resposta ausente de uma ou mais variáveis de SCV. A amostra final deste estudo constituiu-se por 77.494 indivíduos.

Variáveis

a) SCV

Neste estudo, as metas de SCV avaliadas foram adaptadas das metas originais propostas pela Sociedade Americana do Coração em 2010.⁵ As quatro metas comportamentais e as três metas biológicas foram avaliadas por meio de informações autorreferidas, como apresentado no Quadro 1. Cada meta (variável) foi categorizada como resultado favorável (=1) ou desfavorável (=0). O escore de SCV foi obtido pela soma dos resultados das sete metas, com variação de 0 a 7. A SCV ideal foi considerada quando o indivíduo apresentou valor de escore=7, ou seja, quando obteve resultado favorável (=1) nas sete metas (Quadro 1).⁵ Os códigos originais (indicados no dicionário da PNS 2019)¹⁷ das variáveis utilizadas e os resultados considerados favoráveis estão disponíveis no material suplementar.⁵ O consumo de grãos integrais não foi considerado para a meta de dieta adequada, pois essa informação não foi coletada pela PNS 2019.

b) Sociodemográficas

Sexo (masculino e feminino), região do Brasil (norte; nordeste; centro-oeste; sudeste e sul), situação de domicílio (urbano e rural), faixa etária (18-24 anos; 25-39 anos; 40-59 anos e 60 ou mais), raça (cor da pele) (parda; branca; preta e amarela/indígena/ignorada), estado civil (solteiro; casado e separado/ divorciado/viúvo), nível de escolaridade (até o ensino fundamental incompleto; do ensino fundamental completo ao ensino médio incompleto; do ensino médio completo ao ensino superior incompleto e ensino superior completo) e ocupação (com e sem emprego).

c) Doenças crônicas

Uma variável "doenças crônicas" foi definida quando o indivíduo relatasse pelo menos uma das seguintes doenças crônicas: asma (ou bronquite asmática); artrite ou

reumatismo; dor crônica nas costas ou no pescoço, lombalgia, dor ciática e problemas nas vértebras ou disco; distúrbios osteomusculares relacionados ao trabalho; doença no pulmão ou doença pulmonar obstrutiva crônica; câncer; insuficiência renal crônica; depressão; doença mental, como esquizofrenia, transtorno bipolar, psicose ou transtorno obsessivo compulsivo; no geral, doença crônica (física ou mental) ou doença de longa duração (de mais de 6 meses).

Fontes de dados/Mensuração

Os dados são públicos e estão disponíveis na página eletrônica do IBGE¹⁵ ou da PNS (<https://www.pns.icict.fiocruz.br/bases-de-dados/>).

Métodos estatísticos

Inicialmente foram realizadas análises descritivas das características da população de estudo, com cálculo das proporções (para variáveis categóricas) além das médias e desvio-padrão (variáveis contínuas). Em seguida, as prevalências e respectivos IC95% foram estimados para as seguintes condições: (1) SCV ideal; (2) escore de SCV (0-7 metas); (3) metas comportamentais e seu respectivo escore (0-4); (4) metas biológicas e seu respectivo escore (0-3). Por fim, as prevalências de SCV ideal e as metas comportamentais e biológicas foram descritas segundo as características sociodemográficas e presença de doença crônica.

As diferenças nas proporções foram consideradas estatisticamente significativas quando não houve sobreposição nos respectivos intervalos de confiança.¹⁸ As análises foram conduzidas utilizando-se o *software* R¹⁹ versão 3.6.2. O pacote *survey* foi utilizado para ajuste por amostragem complexa e o delineamento do estudo da PNS 2019.¹⁵

Aspectos éticos

A PNS 2019 foi aprovada pela Comissão Nacional de Ética em Pesquisa (CONEP) do Conselho Nacional de Saúde - CNS, em agosto de 2019, sob o número nº 3.529.376.¹⁵ A pesquisa obedeceu à Resolução 196/96 do Conselho Nacional de Saúde, assegurando aos sujeitos de pesquisa sua voluntariedade, anonimato e possibilidade de desistência a qualquer momento do estudo.

RESULTADOS

A amostra final para neste estudo foi constituída por 77.494 indivíduos, com idade média de 48 (± 17) anos. Foram excluídos 2.431 indivíduos com idade inferior a 18 anos completos, 837 indivíduos com entrevistas individuais não completadas, 1.711 questionários individuais respondidos por terceiros, 3.131 mulheres grávidas ou que não souberam informar se estavam grávidas e 8.508 indivíduos com dados faltantes para uma ou mais variáveis de SCV. As características sociodemográficas da amostra são apresentadas na Tabela 1. Observou-se maiores frequências para participantes do sexo feminino (54,7%), pardos (49,5%), sem doença crônica (57,7%), com nível de escolaridade até o ensino fundamental incompleto (38,5%), empregados (59,6%), residentes na região Nordeste (33,9%) e em área urbana (79,0%). A Figura 1 apresenta a distribuição do escore de SCV na população estudada. A prevalência de SCV ideal foi de 0,5% (IC95% 0,4;0,6) na população adulta brasileira. Observou-se prevalência de 8,9% da categoria superior de SCV, 81,5% na intermediária e 9,6% na inferior.

A prevalência estimada de SCV ideal, das metas comportamentais e biológicas é apresentada na Tabela 2. A tabela apresenta ainda os resultados por variáveis

sociodemográficas. Foram observadas prevalências de SVC ideal mais elevadas no sexo feminino (0,7%, IC95% 0,5;0,8), em relação ao masculino; entre os com nível de escolaridade superior completo (1,3%, IC95% 0,9;1,6), em relação aos demais níveis; sem doenças crônicas prévias e entre residentes de áreas urbanas (0,6%, IC95% 0,5;0,7) comparados aos de área rural.

Observou-se uma prevalência de 63,3% (IC95% 62,7;63,9) de indivíduos que reportaram resultados favoráveis em relação às três metas biológicas e de 0,7% (IC95% 0,6;0,8) em relação às quatro metas comportamentais, na população de estudo. Em relação às metas biológicas, a prevalência de indivíduos que reportaram glicemia normal (resultado favorável) foi a mais elevada (91,4%, IC95% 91,0;91,7), enquanto a prevalência de relato de pressão arterial normal (resultado favorável) foi a que apresentou menor nível (73,0%, IC95% 72,4;73,6).

Em relação às metas comportamentais, a que apresentou maior prevalência foi o hábito de não fumar (favorável) (86,0%, IC95% 85,6;86,5) e a dieta adequada (favorável) (4,3%, IC95% 4,0;4,5) a que obteve a menor prevalência na população estudada.

DISCUSSÃO

Os resultados do presente estudo demonstraram uma baixa prevalência de SCV ideal na população adulta brasileira. Dentre os componentes do escore de SCV ideal, observou-se maior prevalência de cumprimento das metas biológicas (63,3%) na população, em detrimento das metas comportamentais (0,7%). A meta relativa à dieta adequada foi a principal responsável pela baixa prevalência das metas comportamentais e da condição de SCV ideal propriamente dita. A análise da SCV ideal de acordo com o

perfil sociodemográfico mostrou maiores níveis de prevalência de SCV ideal na população com maior nível de escolaridade e entre residentes na área urbana.

A principal limitação do presente estudo é que a PNS usa medidas autorreferidas da população. A prevalência de indivíduos que alcançaram metas biológicas e comportamentais favoráveis podem estar subestimadas, uma vez que os dados de morbidade autorreferida, incluindo diagnóstico médico de diabetes mellitus, hipercolesterolemia e hipertensão arterial, dependem do acesso aos serviços de saúde para o diagnóstico. Portanto, aqueles usuários com acesso restrito a esses serviços têm menor oportunidade de diagnóstico médico. No entanto, o uso de dados autorreferidos para avaliação das metas de SCV já foi utilizado em outros estudos²⁰⁻²² e possui boa correlação com as medidas biológicas²³. Outra limitação inclui a comparação entre subgrupos de variáveis sociodemográficas que podem não ter quantidade suficiente de representantes, como por exemplo, raça/cor da pele. Como pontos positivos do estudo ressalta-se o fato de ter sido realizado em uma amostra representativa da população adulta brasileira e a apresentação de dados atualizados acerca da prevalência de SCV no Brasil.

O conceito de SCV ideal é baseado em uma combinação de metas comportamentais e biológicas, as quais estão associadas a evidências de sobrevivência livre de DCVs, longevidade saudável, redução de morbidade e dos custos de saúde.⁵ Embora atingir cinco a sete metas esteja associada ao menor risco de incidência de DCV, um efeito protetor intermediário para DCV pode ser alcançado ao atingir 3-4 metas do escore de SCV.²⁴ No presente estudo, 37,4% da população alcançou 5-7 metas de SCV. No entanto, a prevalência de SCV ideal – quando as sete metas que compõem o escore são alcançadas simultaneamente – foi inferior a 1%. Esses resultados estão em conformidade com estudo anterior que utilizou dados da PNS 2013.⁹ Apesar de algumas

diferenças metodológicas nas duas edições da pesquisa e na construção da amostra, os autores relataram uma prevalência de 0,3% (IC95% 0,2;0,5) de SCV ideal na PNS 2013. Estudos internacionais também relatam uma prevalência baixa de SCV ideal na população adulta, com valores que variam de 0% a 0,1%.⁶⁻⁸

Os resultados do presente estudo demonstraram que a prevalência de metas biológicas favoráveis na população estudada – considerando as três metas simultaneamente – foi maior em comparação às metas comportamentais – quatro metas simultaneamente (63,3% vs 0,7%, respectivamente). As metas de glicemia normal e ausência de hábito de fumar foram as que apresentaram as maiores proporções de indivíduos em suas respectivas categorias, confirmando os resultados da PNS 2013.⁹ No que se refere ao tabagismo, tem-se observado uma tendência de queda em sua prevalência entre adultos brasileiros nas últimas décadas, em resposta às ações desenvolvidas pela Política Nacional de Controle do Tabaco²⁵. Quanto à meta de glicose normal, é necessário destacar que as informações obtidas a partir de autorrelato tendem a ser menos precisas. De fato, um estudo²⁶ baseado nos dados da PNS reportou um aumento de 24% na prevalência de diabetes em 2019 em relação a 2013. Múltiplos fatores contribuíram para este achado, incluindo um aumento na incidência e diagnóstico de diabetes. No entanto, a obesidade e os maus hábitos alimentares se destacam como fatores de risco associados ao aumento nos números do diabetes assim como das DCVs no Brasil.

A meta de dieta adequada foi a principal responsável pela baixa prevalência de SCV ideal no presente estudo. Isto se deve ao fato de apenas 4,3% da população ter apresentado, simultaneamente, um consumo de sal, açúcar, peixe, frutas e hortaliças de acordo com as recomendações. Outros estudos nacionais e internacionais^{27,28} encontraram resultados semelhantes. Por exemplo, no estudo ELSA-Brasil²⁸ – que

avaliou servidores com 35 a 74 anos – a meta de dieta adequada foi a que apresentou pior resultado (1,3%). Sabe-se que indivíduos adultos que aderem às recomendações para uma dieta saudável apresentam menores taxas de morbidade e mortalidade cardiovascular.²⁹ Sendo assim, os resultados do presente estudo confirmam os achados anteriores acerca da importância de intervenções para promoção de hábitos alimentares saudáveis, como parte de uma estratégia para prevenção de DCVs.

Conclui-se que a prevalência de brasileiros com SCV ideal foi baixíssima. Esses resultados, baseados na PNS 2019, seguem a tendência dos achados da PNS 2013. As metas comportamentais de SCV, principalmente a dieta adequada, apresentaram os piores resultados. Estes achados podem ser úteis no embasamento de ações para promoção da SCV e prevenção das DCV. O sucesso na execução dessas ações implicará numa redução das mortes prematuras por DCV, sendo essa uma das pautas prioritárias da Agenda 2030 para o Desenvolvimento Sustentável, implementada pela Organização das Nações Unidas e da qual o Brasil é partícipe.³⁰ Além disso, espera-se como resultado dessas ações uma redução dos custos em saúde no Brasil, já que as DCV são responsáveis pelos maiores gastos com hospitalização no Sistema Único de Saúde.³¹

REFERÊNCIAS

1. Malta DC, Teixeira R, de Oliveira GMM, Ribeiro ALP. Cardiovascular disease mortality according to the Brazilian information system on mortality and the global burden of disease study estimates in Brazil, 2000-2017. *Arq Bras Cardiol.* 2020;115(2):152-160. doi:10.36660/abc.20190867
2. Oliveira GMM de, Brant LCC, Polanczyk CA, et al. Estatística Cardiovascular – Brasil 2021. *Arq Bras Cardiol.* 2022;118(1):115-373. doi:10.36660/abc.20211012
3. Yusuf S, Joseph P, Rangarajan S, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *The Lancet.* 2020;395(10226):795-808. doi:10.1016/S0140-6736(19)32008-2
4. Yusuf PS, Hawken S, Ôunpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet.* 2004;364(9438):937-952. doi:10.1016/S0140-6736(04)17018-9
5. Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: The American heart association's strategic impact goal through 2020 and beyond. *Circulation.* 2010;121(4):586-613. doi:10.1161/CIRCULATIONAHA.109.192703
6. Younus A, Aneni EC, Spatz ES, et al. A Systematic Review of the Prevalence and Outcomes of Ideal Cardiovascular Health in US and Non-US Populations. *Mayo Clin Proc.* 2016;91(5):649-670. doi:10.1016/j.mayocp.2016.01.019
7. Joseph JJ, Bennett A, Echouffo Tcheguigui JB, et al. Ideal cardiovascular health, glycaemic status and incident type 2 diabetes mellitus: the REasons for Geographic and Racial Differences in Stroke (REGARDS) study. *Diabetologia.* 2019;62(3):426-437. doi:10.1007/s00125-018-4792-y
8. Bousquet-Santos K, Chen R, Kubzansky LD. A sad heart: Depression and favorable cardiovascular health in Brazil. *Prev Med (Baltim).* 2021;142:106378. doi:https://doi.org/10.1016/j.ypmed.2020.106378
9. Velasquez-Melendez G, Felisbino-Mendes MS, Matozinhos FP, Claro R, Gomes CS, Malta DC. Prevalência de saúde cardiovascular ideal na população Brasileira – pesquisa nacional de saúde (2013). *Revista Brasileira de Epidemiologia.* 2015;18:97-108. doi:10.1590/1980-5497201500060009
10. Janković J, Davidović M, Bjegović-Mikanović V, Janković S. Status of cardiovascular health in the Republic of Serbia: Results from the National health Survey. *PLoS One.* 2019;14(3). doi:10.1371/journal.pone.0214505
11. Xanthakis V, Enserro DM, Murabito JM, et al. Ideal Cardiovascular Health. *Circulation.* 2014;130(19):1676-1683. doi:10.1161/CIRCULATIONAHA.114.009273
12. Nguyen XMT, Quaden RM, Wolfrum S, et al. Prevalence of Ideal Cardiovascular Health Metrics in the Million Veteran Program. *American Journal of Cardiology.* 2018;122(2):347-352. doi:10.1016/j.amjcard.2018.04.002
13. Djoussé L, Petrone AB, Blackshear C, et al. Prevalence and changes over time of ideal cardiovascular health metrics among African-Americans: The Jackson Heart Study. *Prev Med (Baltim).* 2015;74:111-116. doi:https://doi.org/10.1016/j.ypmed.2015.02.006

14. Seron P, Irazola V, Rubinstein A, et al. Ideal Cardiovascular Health in the southern cone of Latin America. *Public Health*. 2018;156:132-139. doi:10.1016/j.puhe.2017.12.017
15. IBGE - Coordenação de Trabalho e Rendimento. *Pesquisa Nacional de Saúde 2019: Informações Sobre Domicílios, Acesso e Utilização Dos Serviços de Saúde: Brasil, Grandes Regiões e Unidades Da Federação*. IBGE; 2020.
16. Stopa SR, Szwarcwald CL, Oliveira MM de, et al. Pesquisa Nacional de Saúde 2019: histórico, métodos e perspectivas. *Epidemiol Serv Saude*. 2020;29(5):e2020315. doi:10.1590/S1679-49742020000500004
17. Instituto Brasileiro de Geografia e Estatística (IBGE). PNS - Pesquisa Nacional de Saúde. Accessed May 3, 2022. <https://www.ibge.gov.br/estatisticas/sociais/saude/29540-2013-pesquisa-nacional-de-saude.html?edicao=9177&t=microdados>
18. Bonita R, Beaglehole R, Kjellström T, Organization WH. Basic epidemiology. Published online 2006:xi, 212 p.
19. R Development Core Team (2009). R: A language and environment for statistical computing. <http://www.R-project.org>
20. Shockey TM, Sussell AL, Odom EC. *Morbidity and Mortality Weekly Report*; 2016. http://www.cdc.gov/mmwr/cme/conted_info.html#weekly. http://www.cdc.gov/bfss/annual_data/2013/pdf/overview_2013.pdf.
21. Burroughs Peña MS, Mbassa RS, Slopen NB, Williams DR, Buring JE, Albert MA. Cumulative psychosocial stress and ideal cardiovascular health in older women: Data by race/ethnicity. *Circulation*. 2019;139(17):2012-2021. doi:10.1161/CIRCULATIONAHA.118.033915
22. Thomas SC, Umer A, Commodore-Mensah Y, Davidov D, Abildso CG. Length of Residence and Cardiovascular Health among Afro-Caribbean Immigrants in New York City. *J Racial Ethn Health Disparities*. 2019;6(3):487-496. doi:10.1007/s40615-018-00547-8
23. Moreira AD, Gomes CS, Machado ÍE, Carvalho D, Felisbino-Mendes MS. Saúde cardiovascular e validação do escore autorreferido no Brasil: uma análise da Pesquisa Nacional de Saúde. *Cien Saude Colet*. 2020;25(11):4259-4268. doi:10.1590/1413-812320202511.31442020
24. Ramírez-Vélez R, Saavedra JM, Lobelo F, Celis-Morales CA, Pozo-Cruz B del, García-Hermoso A. Ideal Cardiovascular Health and Incident Cardiovascular Disease Among Adults: A Systematic Review and Meta-analysis. *Mayo Clin Proc*. 2018;93(11):1589-1599. doi:10.1016/j.mayocp.2018.05.035
25. INCA - Observatório da Política Nacional de Controle do Tabaco. Dados e números da prevalência do tabagismo. Última modificação: 13/05/2022. Published 2022. Accessed July 31, 2022. <https://www.inca.gov.br/observatorio-da-politica-nacional-de-controle-do-tabaco/dados-e-numeros-prevalencia-tabagismo#:~:text=Segundo%20dados%20do%20Vigitel%202021,6%2C7%25%20entre%20mulheres>.
26. dos Reis RCP, Duncan BB, Malta DC, Iser BPM, Schmidt MI. Evolution of diabetes in Brazil: prevalence data from the 2013 and 2019 Brazilian National Health Survey. *Cad Saude Publica*. 2022;38. doi:10.1590/0102-311X00149321

27. Pilkerton CS, Singh SS, Bias TK, Frisbee SJ. Changes in cardiovascular health in the United States, 2003-2011. *J Am Heart Assoc.* 2015;4(9). doi:10.1161/JAHA.114.001650
28. Machado LBM, Silva BLS, Garcia AP, et al. Ideal cardiovascular health score at the ELSA-Brasil baseline and its association with sociodemographic characteristics. *Int J Cardiol.* 2018;254:333-337. doi:10.1016/j.ijcard.2017.12.037
29. Mangione CM, Barry MJ, Nicholson WK, et al. Behavioral Counseling Interventions to Promote a Healthy Diet and Physical Activity for Cardiovascular Disease Prevention in Adults Without Cardiovascular Disease Risk Factors. *JAMA.* 2022;328(4):367. doi:10.1001/jama.2022.10951
30. IPEA. Objetivos de Desenvolvimento Sustentável - 3. Saúde e Bem-estar. Accessed October 19, 2022. <https://www.ipea.gov.br/ods/ods3.html>
31. Moraes De Oliveira GM, Caldeira Brant LC, Anne Polanczyk C, et al. Cardiovascular statistics–brazil 2020. *Arq Bras Cardiol.* 2020;115(3):308-439. doi:10.36660/abc.20200812

TABELAS, QUADROS E FIGURAS

Quadro 1- Metas de saúde cardiovascular adaptadas da Sociedade Americana do Coração.

Metas	Métrica utilizada
Comportamentais	4 metas comportamentais atingidas simultaneamente.
Tabagismo	Nunca ter fumado ou não fumar há mais de 12 meses.
IMC ^a	< 25 kg/m ² (a partir de medidas de peso e altura autorreferidas e imputadas).
Atividade Física	Prática de exercício físico ou esporte \geq 150 min/semana.
Dieta adequada	4 metas de dieta atingidas simultaneamente:
Dieta	1) Frutas e vegetais: consumo \geq 1 vez/dia; 2) Peixe: consumo \geq 2 vezes/semana; 3) Sódio: consumo de sal na comida preparada na hora e nos alimentos industrializados em quantidade adequada, baixa ou muito baixa; 4) Açúcar: consumo de refrigerantes e/ou sucos industrializados <5 dias/semana.
Biológicas	3 metas biológicas atingidas simultaneamente.
Colesterol Total	Ausência de diagnóstico médico autorreferido de colesterol alto.
Pressão arterial	Ausência de diagnóstico médico autorreferido de hipertensão arterial.
Glicose	Ausência de diagnóstico médico autorreferido de diabetes.

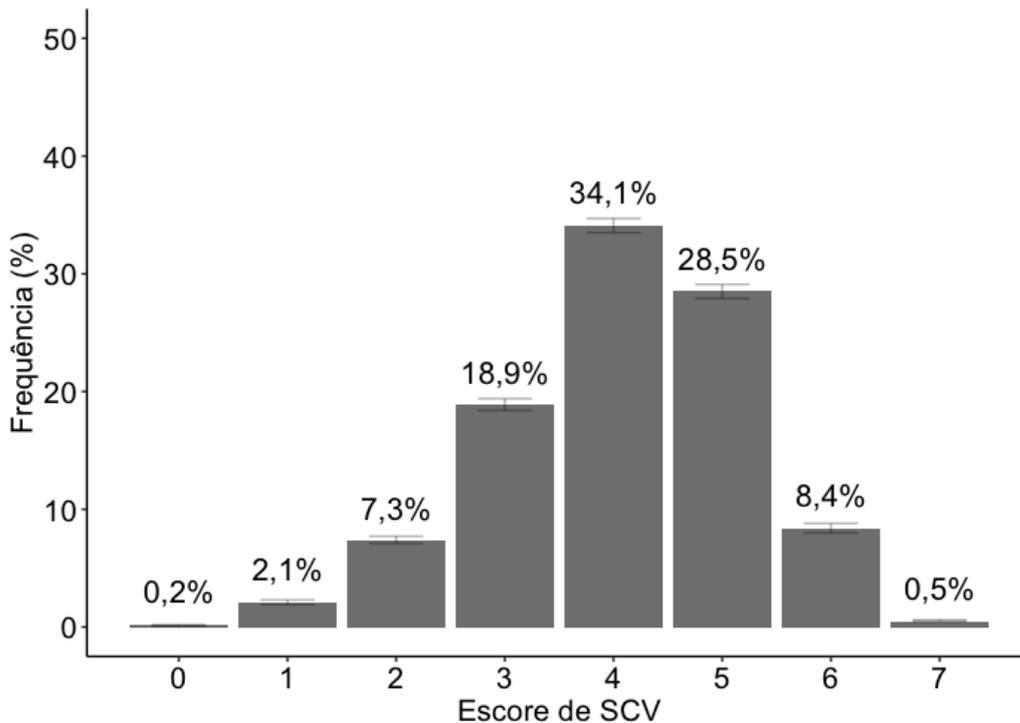
a) IMC – Índice de Massa Corporal.

Tabela 1: Características da população brasileira adulta estudada, PNS 2019 (n = 77.494).

Variáveis	N	%	IC95%
Sexo			
Masculino	63.359.986	45,2	(44,6; 45,9)
Feminino	76.676.523	54,8	(54,1; 55,4)
Faixa etária (anos)			
18 a 24	16.797.883	12,0	(11,5; 12,5)
25 a 39	40.026.391	28,6	(28,0; 29,2)
40 a 59	51.822.239	37,0	(36,4; 37,6)
60 ou mais	31.389.995	22,4	(21,8; 23,0)
Raça/Cor da pele			
Parda	59.763.814	42,7	(42,0; 43,4)
Branca	62.440.066	44,6	(43,8; 45,4)
Preta	15.752.117	11,2	(10,8; 11,7)
Amarela/Indígena/Ignorada	2.080.511	1,5	(1,3; 1,7)
Estado Civil			
Solteiro	55.627.974	39,7	(39,0; 40,4)
Casado	64.089.556	45,8	(45,1; 46,5)
Separado/Divorciado/Viúvo	20.318.979	14,5	(14,1; 14,9)
Escolaridade (Nível de ensino)			
Até fundamental incompleto	46.739.391	33,4	(32,7; 34,1)
Fundamental completo à médio incompleto	19.367.870	13,8	(13,4; 14,3)
Médio completo à superior incompleto	49.745.301	35,5	(34,9; 36,2)
Superior Completo	24.183.947	17,3	(16,6; 18,0)
Situação de Ocupação			
Empregado	86.692.821	61,9	(61,3; 62,5)
Sem ocupação	53.343.688	38,1	(37,5; 38,7)
Doença Crônica			
Pelo menos 1	59.969.579	42,8	(42,1; 43,5)
Sem doença	80.066.930	57,2	(56,5; 57,9)
Região			
Norte	10.178.075	7,3	(7,0; 7,6)
Nordeste	35.417.215	25,3	(24,7; 25,9)
Sudeste	63.001.662	45,0	(44,1; 45,9)
Sul	20.852.598	14,9	(14,4; 15,3)
Centro-Oeste	10.586.959	7,6	(7,3; 7,9)
Situação de Domicílio			
Urbano	122.412.326	87,4	(87,0; 87,8)
Rural	17.624.183	12,6	(12,2; 13,0)

Fonte: Pesquisa Nacional de Saúde de 2019.

Figura 1: Distribuição da frequência da população de acordo com a pontuação obtida no escore de SCV^a com respectivos intervalos de 95% de confiança^b, estimados para a população adulta brasileira em 2019^c (Dados da PNS 2019).



a) SCV – Saúde Cardiovascular

b) Os intervalos de confiança de 95% para os valores do Escore de SCV foram, respectivamente: 0,1 a 0,2; 1,9 a 2,3; 7,1 a 7,7; 18,4 a 19,4; 33,5 a 34,7; 27,9 a 29,1; 8,0 a 8,8; e 0,5 a 0,6.

Fonte: Pesquisa Nacional de Saúde 2019.

Tabela 2: Prevalência de brasileiros adultos com metas comportamentais e biológicas adequadas e SCV ideal. PNS 2019 (n=77.494).

(Continua)

Característica sociodemográfica	Metas comportamentais					Metas biológicas				Saúde Cardiovascular Ideal (soma das 7 metas)
	Metas comportamentais (soma das 4 metas)	Não fumar	Peso corporal adequado	Prática de atividade física	Dieta adequada	Metas biológicas (soma das 3 metas)	Níveis normais de colesterol	Pressão Arterial normal	Glicemia normal	
Total	0,7 (0,6; 0,8)	86,0 (85,6; 86,5)	34,1 (33,4; 34,7)	27,4 (26,8; 28,1)	4,3 (4,0; 4,5)	63,3 (62,7; 63,9)	84,1 (83,6; 84,6)	73,0 (72,4; 73,6)	91,4 (91,0; 91,7)	0,5 (0,4; 0,6)
Sexo										
Masculino	0,5 (0,4; 0,6)	82,7 (82,0; 83,4)	38,3 (37,4; 39,2)	29,3 (28,4; 30,2)	3,1 (2,8; 3,3)	68,1 (67,3; 69,0)	87,5 (86,8; 88,1)	77,2 (76,4; 77,9)	92,3 (91,8; 92,8)	0,4 (0,3; 0,5)
Feminino	0,9 (0,8; 1,1)	88,8 (88,3; 89,3)	41,7 (40,9; 42,6)	25,9 (25,1; 26,6)	5,2 (4,9; 5,6)	59,2 (58,4; 60,1)	81,3 (80,6; 81,9)	69,6 (68,8; 70,3)	90,6 (90,1; 91,1)	0,7 (0,5; 0,8)
Faixa etária (anos)										
18 a 24	0,5 (0,2; 0,9)	86,6 (85,0; 88,2)	61,6 (60,0; 63,7)	35,7 (33,4; 37,9)	1,6 (1,0; 2,1)	92,0 (90,8; 93,1)	95,8 (94,9; 96,7)	96,4 (95,6; 97,2)	99,0 (98,6; 99,4)	0,5 (0,2; 0,9)
25 a 39	0,6 (0,5; 0,8)	86,7 (85,9; 87,5)	39,9 (38,6; 41,2)	31,3 (30,2; 32,4)	2,6 (2,3; 2,9)	83,3 (82,4; 84,2)	92,6 (91,9; 93,3)	90,1 (89,4; 90,8)	97,8 (97,5; 98,2)	0,6 (0,4; 0,7)
40 a 59	0,8 (0,7; 1,1)	84,6 (83,9; 85,4)	34,1 (33,2; 35,1)	26,4 (25,3; 27,4)	5,0 (4,6; 5,3)	58,3 (57,2; 59,3)	81,2 (80,4; 82,0)	70,4 (69,4; 71,4)	91,2 (90,7; 91,8)	0,6 (0,4; 0,7)
60 ou mais	0,8 (0,7; 1,0)	87,3 (86,5; 88,0)	39,1 (38,1; 40,2)	19,9 (18,9; 20,9)	6,7 (6,1; 7,2)	30,6 (29,6; 31,6)	71,8 (70,7; 72,9)	43,0 (41,9; 44,2)	79,3 (78,4; 80,2)	0,3 (0,2; 0,4)
Raça										
Parda	0,7 (0,5; 0,8)	85,4 (84,7; 86,1)	40,5 (39,7; 41,3)	26,1 (25,3; 26,9)	3,9 (3,6; 4,2)	64,6 (63,7; 65,4)	85,1 (84,5; 85,6)	73,4 (72,6; 74,2)	91,5 (91,1; 92,0)	0,5 (0,3; 0,6)
Branca	0,9 (0,7; 1,0)	87,0 (86,3; 87,6)	40,2 (39,2; 41,3)	29,0 (27,9; 30,0)	4,6 (4,2; 5,0)	62,4 (61,4; 63,3)	82,8 (82,0; 83,6)	73,3 (72,4; 74,2)	91,4 (90,9; 91,9)	0,7 (0,5; 0,8)
Preta	0,4 (0,2; 0,5)	85,0 (83,6; 86,3)	37,8 (36,1; 39,6)	26,5 (24,9; 28,1)	3,9 (3,3; 4,5)	62,4 (60,6; 64,2)	85,5 (84,2; 86,9)	70,5 (68,9; 72,1)	91,0 (90,0; 91,9)	0,3 (0,2; 0,4)
Amarela/ Indígena/ Ignorada	1,3 (0,6; 2,0)	84,3 (80,6; 87,9)	47,7 (41,8; 53,7)	27,0 (22,1; 31,8)	5,9 (3,6; 8,1)	59,7 (53,9; 65,4)	83,1 (78,5; 87,7)	70,1 (65,0; 75,2)	87,9 (84,2; 91,7)	0,8 (0,2; 1,3)
Estado civil										
Solteiro	0,7 (0,5; 0,8)	82,4 (81,5; 83,2)	47,2 (46,2; 48,2)	29,8 (28,8; 30,8)	3,1 (2,8; 3,4)	76,7 (75,9; 77,5)	89,7 (89,1; 90,3)	84,0 (83,3; 84,7)	95,4 (95,1; 95,8)	0,5 (0,4; 0,7)
Casado	0,8 (0,7; 1,0)	89,8 (89,3; 90,4)	35,0 (34,1; 35,9)	26,9 (26,0; 27,8)	4,7 (4,3; 5,1)	57,6 (56,7; 58,5)	81,7 (81,0; 82,5)	68,9 (68,0; 69,8)	89,9 (89,4; 90,5)	0,5 (0,4; 0,7)
Separado/ Divorciado/ Viúvo	0,9 (0,7; 1,1)	84,1 (83,1; 85,2)	37,4 (36,0; 38,8)	22,5 (21,3; 23,7)	6,1 (5,5; 6,7)	44,5 (43,1; 45,9)	76,0 (74,8; 77,3)	55,8 (54,4 - 57,3)	84,7 (83,8; 85,7)	0,4 (0,3; 0,6)

Tabela 2: Prevalência de brasileiros adultos com metas comportamentais e biológicas adequadas e SCV ideal. PNS 2019 (n=77.494).

(Conclusão)

Característica sociodemográfica	Metas comportamentais					Metas biológicas				Saúde Cardiovascular Ideal (soma das 7 metas)
	Metas comportamentais (soma das 4 metas)	Não fumar	Peso corporal adequado	Prática de atividade física	Dieta adequada	Metas biológicas (soma das 3 metas)	Níveis normais de colesterol	Pressão Arterial normal	Glicemia normal	
Escolaridade (Nível de ensino)										
Até fundamental incompleto	0,4 (0,3; 0,5)	81,7 (80,9; 82,5)	39,1 (38,1; 40,0)	15,6 (14,8; 16,3)	3,7 (3,3; 4,0)	48,4 (47,4; 49,4)	78,7 (77,8; 79,5)	58,9 (57,9; 59,9)	85,6 (84,9; 86,2)	0,2 (0,1; 0,3)
Fundamental completo à médio incompleto	0,5 (0,3; 0,7)	82,8 (81,5; 84,1)	40,4 (38,7; 42,0)	22,9 (21,5; 24,4)	3,2 (2,7; 3,7)	67,3 (65,7; 69,0)	87,4 (86,3; 88,5)	75,3 (73,9; 76,8)	92,6 (91,7; 93,4)	0,3 (0,1; 0,5)
Médio completo à superior incompleto	0,8 (0,6; 0,9)	88,7 (88,0; 89,4)	40,1 (39,8; 41,2)	32,0 (31,0; 33,1)	3,8 (3,5; 4,2)	72,9 (71,9; 73,9)	87,7 (86,9; 88,5)	81,8 (81,0; 82,7)	94,7 (94,2; 95,1)	0,6 (0,4; 0,7)
Superior Completo	1,7 (1,4; 2,1)	91,6 (90,8; 92,4)	40,6 (39,0; 41,1)	44,5 (43,0; 46,0)	7,2 (6,4; 7,9)	68,9 (67,5; 70,3)	84,4 (83,4; 85,5)	80,2 (79,0 - 81,3)	94,9 (94,3 - 95,5)	1,3 (0,9 - 1,6)
Ocupação										
Empregado	0,7 (0,6; 0,8)	86,1 (85,5; 86,6)	39,0 (38,2; 39,8)	29,6 (28,8; 30,4)	3,8 (3,6; 4,1)	71,3 (70,5; 72,0)	87,1 (86,5; 87,7)	80,5 (79,8; 81,1)	94,6 (94,2; 94,9)	0,6 (0,5; 0,7)
Sem Emprego	0,8 (0,7; 1,0)	86,0 (85,3; 86,7)	42,0 (41,1; 43,1)	23,9 (23,0; 24,7)	4,9 (4,6; 5,3)	50,3 (49,3; 51,3)	79,2 (78,4; 79,9)	60,8 (59,9; 61,8)	86,2 (85,5; 86,8)	0,5 (0,3; 0,6)
Doença crônica										
Sem Doença	0,8 (0,7; 1,0)	86,8 (86,2; 87,4)	42,4 (41,5; 43,3)	29,0 (28,2; 29,8)	4,1 (3,8; 4,4)	71,7 (70,9; 72,4)	89,5 (89,0; 90,1)	79,5 (78,8; 80,1)	93,3 (92,9; 93,7)	0,7 (0,5; 0,8)
Com Doença	0,6 (0,5; 0,7)	85,1 (84,3; 85,8)	37,2 (36,3; 38,2)	25,3 (24,5; 26,2)	4,5 (4,2; 4,8)	52,0 (51,1; 53,0)	76,8 (76,0; 77,6)	64,4 (63,4; 65,3)	88,8 (88,3; 89,4)	0,4 (0,3; 0,5)
Região										
Norte	0,7 (0,5; 0,9)	89,5 (88,7; 90,3)	42,4 (41,6; 43,7)	25,9 (24,7; 27,1)	4,9 (4,3; 5,5)	69,9 (68,6; 71,2)	86,3 (85,4; 87,2)	79,4 (78,4; 80,5)	93,4 (92,8; 94,1)	0,4 (0,3; 0,6)
Nordeste	0,9 (0,7; 1,0)	88,4 (87,8; 89,1)	42,7 (41,7; 43,7)	27,2 (26,4; 28,1)	5,4 (5,0; 5,8)	63,9 (63,0; 64,9)	83,9 (83,2; 84,7)	73,4 (72,6; 74,2)	91,9 (91,4; 92,4)	0,6 (0,4; 0,7)
Sudeste	0,8 (0,6; 0,9)	84,9 (84,1; 85,7)	39,8 (38,6; 41,0)	28,1 (26,9; 29,3)	4,0 (3,6; 4,4)	61,5 (60,4; 62,6)	83,3 (82,4; 84,2)	71,7 (70,7; 72,8)	90,7 (90,1; 91,4)	0,6 (0,4; 0,7)
Sul	0,6 (0,5; 0,8)	83,9 (82,9; 85,0)	36,1 (34,7; 37,5)	25,1 (23,8; 26,4)	3,2 (2,8; 3,7)	63,0 (61,6; 64,3)	84,7 (83,8; 85,7)	72,2 (70,9; 73,4)	91,2 (90,5; 92,0)	0,5 (0,3; 0,6)
Centro-Oeste	0,8 (0,5; 1,1)	85,9 (84,9; 86,9)	40,0 (38,6; 41,4)	29,8 (28,4; 31,3)	3,6 (3,1; 4,2)	65,8 (64,3; 67,4)	85,6 (84,4; 86,7)	74,6 (73,3; 76,0)	91,7 (90,9; 92,6)	0,5 (0,3; 0,8)
Situação de domicílio										
Urbano	0,8 (0,7; 0,9)	86,1 (85,6; 86,6)	39,5 (38,7; 40,2)	29,1 (28,4; 29,8)	4,4 (4,2; 4,7)	63,2 (62,6; 63,9)	83,9 (83,3; 84,4)	73,1 (72,5; 73,8)	91,2 (90,9; 91,6)	0,6 (0,5; 0,7)
Rural	0,4 (0,2; 0,5)	85,5 (84,5; 86,6)	45,3 (44,1; 46,6)	15,8 (14,8; 16,9)	3,1 (2,7; 3,5)	63,4 (62,2; 64,6)	85,7 (84,8; 86,5)	71,9 (70,8; 73,1)	92,5 (91,9; 93,1)	0,2 (0,1; 0,3)

Fonte: Pesquisa Nacional de Saúde de 2019.

MATERIAL SUPLEMENTAR

Variáveis	Descrição das variáveis do questionário da PNS ^a 2019	Método de cálculo
Tabagismo	P050. Atualmente, o(a) Sr.(a) fuma algum produto do tabaco? Opções: 1- Sim, diariamente; 2- Sim, menos que diariamente; 3- Não fumo atualmente. P052. E no passado, o(a) Sr.(a) fumou algum produto do tabaco? Opções: 1- Sim, diariamente; 2- Sim, menos que diariamente; 3- Não nunca fumei. P05901. Número de anos que parou de fumar. Opções: Número de anos.	Resultado Favorável (1): Se indivíduo é não fumante [P050 = 3 e P052 = 3] ou é ex-fumante há mais de um ano [P050 = 3 e P052 = (1 ou 2) e P05901 > 1]
IMC	P00104. Peso; Final (em kg) Opções: 3 inteiros e 1 casa decimal P00404. Altura; Final (em cm) Opções: 3 inteiros	Resultado Favorável (1): Peso [P00104] dividido pelo quadrado da Altura [P00404] < 0,0025
Atividade Física	P035. Quantos dias por semana o(a) Sr(a) costuma (costumava) praticar exercício físico ou esporte? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P03701. Em geral, no dia que o(a) Sr(a) pratica exercício ou esporte, quanto tempo em horas dura essa atividade? Opções: Quantidade de horas. P03702. Em geral, no dia que o(a) Sr(a) pratica (praticava) exercício ou esporte, quanto tempo em minutos dura essa atividade? Opções: Quantidade de minutos.	Resultado Favorável (1): Quantidade de exercícios físicos praticados em minutos por dia [P03701*60+P03702] vezes a quantidade de dias por semana [P035] ≥ 150
Dieta	P015. Em quantos dias da semana o(a) Sr(a) costuma comer peixe? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P02001. Em quantos dias da semana o(a) Sr(a) costuma tomar suco de caixinha/lata ou refresco em pó? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P02002. Em quantos dias da semana o(a) Sr(a) costuma tomar refrigerante? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P018. Em quantos dias da semana o(a) Sr(a) costuma comer frutas? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P00901. Em quantos dias da semana, o(a) Sr(a) costuma comer pelo menos um tipo de verdura ou legume (sem contar batata, mandioca, cará ou inhame) como alface, tomate, couve, cenoura, chuchu, berinjela, abobrinha? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P02601. Considerando a comida preparada na hora e os alimentos industrializados, o(a) Sr(a) acha que o seu consumo de sal é: Opções: 1- Muito Alto; 2- Alto; 3- Adequado; 4- Baixo; 5- Muito Baixo.	Resultado Favorável (1): Indivíduos que atingiram as 4 metas de dieta: (1) Consumo de frutas e vegetais todos os dias [P018 = 7 e P00901 = 7]; (2) Consumo de peixe 2 vezes por semana [P015 ≥ 2]; (3) Consumo de sal adequado, baixo ou muito baixo [P02601 ≥ 3]; e (4) Consumo de refrigerantes e/ou sucos industrializados em menos de 5 dias da semana [P02001+P02002 < 5].
Colesterol Total	Q060. Algum médico já lhe deu o diagnóstico de colesterol alto? Opções: 1- Sim; 2- Não.	Resultado Favorável (1): Ausência de diagnóstico médico de colesterol alto [Q060 = 2].
Pressão arterial	Q00201. Algum médico já lhe deu o diagnóstico de hipertensão arterial (pressão alta)? Opções: 1- Sim; 2- Não.	Resultado Favorável (1): Ausência de diagnóstico médico de hipertensão arterial [Q00201 = 2].
Glicose	Q03001. Algum médico já lhe deu o diagnóstico de diabetes? Opções: 1- Sim; 2- Não.	Resultado Favorável (1): Ausência de diagnóstico médico de diabetes [Q03001 = 2].

Quadro Suplementar 1. Construção e métodos de cálculo das variáveis, Pesquisa Nacional de Saúde 2019, Brasil

Legenda:

^a PNS – Pesquisa Nacional de Saúde

Capítulo 3

Artigo 2: Associação entre saúde cardiovascular e determinantes socioeconômicos na população adulta brasileira – Pesquisa Nacional de Saúde 2019

Artigo em preparo para submissão em periódico científico.

ARTIGO ORIGINAL

Associação entre saúde cardiovascular e determinantes socioeconômicos na população adulta brasileira – Pesquisa Nacional de Saúde 2019

Association of cardiovascular health and socioeconomic determinants in the Brazilian adult population – National Health Survey 2019

Asociación de la salud cardiovascular y los determinantes socioeconómicos en la población adulta brasileña – Encuesta Nacional de Salud 2019

Título resumido em português: Associação de saúde cardiovascular e determinantes socioeconômicos em adultos no Brasil.

Ana Carolina Souto Valente Motta¹; orcid.org/0000-0001-9016-4794

Kelb Bousquet-Santos²; orcid.org/0000-0002-4687-4681

Isabela Harumi Lopes Motoki¹ – orcid.org/0000-0002-8316-921X

Joanlise Marco De Leon Andrade¹; orcid.org/0000-0002-3984-3799

¹ Universidade de Brasília, Departamento de Estatística, Brasília, DF, Brasil.

² Universidade de Brasília, Colegiado de Bases Biológicas e da Saúde, Brasília, DF, Brasil.

CORRESPONDÊNCIA

Ana Carolina Souto Valente Motta | e-mail: anacarolina_souto@hotmail.com

FINANCIAMENTO

Não se aplica.

TRABALHO ACADÊMICO ASSOCIADO

Artigo derivado de dissertação de mestrado intitulada “Associação entre Saúde Cardiovascular e Determinantes Socioeconômicos: Uma Aplicação de modelos de regressão logística multinomial e ordinal”, defendida por ‘Ana Carolina Souto Valente Motta’ no Programa de Pós-graduação em Estatística, da Universidade de Brasília, em novembro de 2022.

CONFLITOS DE INTERESSE

Os autores declararam não possuir conflitos de interesse.

CONTRIBUIÇÃO DOS AUTORES

Todos os autores contribuíram de forma significativa no desenho do estudo, na análise e interpretação dos dados e na elaboração da versão preliminar do manuscrito. Todos os autores aprovaram a versão final do manuscrito e são responsáveis por todos os seus aspectos, incluindo a garantia de sua precisão e integridade.

AGRADECIMENTO

Não se aplica.

RESUMO

Objetivo: Avaliar a associação entre Saúde Cardiovascular (SCV) e determinantes socioeconômicos com base em dados da Pesquisa Nacional de Saúde (PNS) 2019.

Métodos: Inquérito de saúde de âmbito nacional. Escore de SCV obtido pela soma de sete metas favoráveis (tabagismo, índice de massa corpórea, exercício, dieta, colesterol, pressão arterial e glicose). Tal escore foi categorizado em 2 (6 ou 7 metas atingidas vs. demais valores) ou 3 níveis de SCV (inferior: 0 a 2 metas, intermediário: 3 a 5 metas e superior: 6 ou 7 metas atingidas). Análises realizadas utilizando modelos de regressão logística e multinomial. **Resultados:** Escolaridade, índice de riqueza e situação de ocupação estão associados a SCV (em 3 categorias) após ajuste por covariáveis sociodemográficas. Para o modelo binário ocupação não foi significativa. **Conclusão:** Há associação entre variáveis socioeconômicas e SCV na população adulta brasileira, o que pode auxiliar no desenvolvimento de intervenções mais direcionadas e eficazes.

Palavras-chave: Estudos Populacionais em Saúde Pública; Escore de Saúde cardiovascular; Desigualdade em Saúde; Determinantes socioeconômicos; Regressão logística; Regressão multinomial.

ABSTRACT

Objective: To assess the association between Cardiovascular Health (CVH) and socioeconomic determinants based on data from the National Health Survey (PNS) 2019.

Methods: Nationwide health survey. The CVH score was obtained by the sum of seven favorable goals (smoking, body mass index, exercise, diet, cholesterol, blood pressure and glucose). This score was categorized into 2 (6 or 7 goals achieved vs. other values) and 3 CVH levels (lower: 0 to 2 goals, intermediate: 3 to 5 goals and superior: 6 or 7 goals achieved). Analyzes were performed using logistic and multinomial regression models. **Results:** Education, wealth index and employment status were associated with SCV (in 3 categories) after adjustment for sociodemographic covariates. Employment status was not significant in the binary model. **Conclusion:** SCV is associated with socioeconomic variables in the Brazilian adult population, which can help the development of more targeted and effective interventions.

Keywords: Population Studies in Public Health; Cardiovascular Health Score; Health Inequality; Socioeconomic determinants; Logistic regression; Multinomial regression.

INTRODUÇÃO

As doenças cardiovasculares (DCV) constituem a principal causa de morte no mundo e no Brasil, correspondendo a um terço do total de óbitos¹. Em países em desenvolvimento, as taxas de mortalidade por DCVs costumam ser elevadas devido ao acesso mais limitado de grupos desfavorecidos das populações a serviços de saúde eficazes e à programas para detecção e tratamento precoce dos fatores que favorecem tais doenças².

Em 2010, a Associação Americana do Coração propôs o conceito de Saúde Cardiovascular (SCV) Ideal com o objetivo de melhorar a saúde cardiovascular da população americana e reduzir a mortalidade por DCV.³ Tendo como premissa a prevenção primordial, as sete metas que compõem o escore de SCV incluem três fatores biológicos (níveis adequados de pressão arterial, colesterol e glicose sérica) e quatro fatores comportamentais (dieta adequada, prática de atividade física, peso corporal adequado e não fumar). Desde a sua criação, o escore de SCV vem sendo amplamente utilizado na literatura científica internacional em associação com desfechos cardiovasculares ou outros tipos de desfechos.⁴⁻⁶

O acesso às informações sobre prevenção de DCVs e hábitos de vida saudável é diferente entre os grupos sociais⁷. Alguns fatores socioeconômicos favoráveis, como maiores níveis de riqueza, escolaridade e empregabilidade, podem contribuir para a ampliação do percentual de indivíduos com SCV ideal, pois essas condições sociais podem estar fortemente associadas a escolhas de alimentação mais saudável e de prática regular de atividade física, por exemplo⁸.

Assim, a coleta de informações em saúde é fundamental para a compreensão do perfil de saúde e socioeconômico dos indivíduos, os riscos envolvidos, as fragilidades do

sistema e as oportunidades de melhorias, para assim otimizar o planejamento e a gestão da saúde individual e coletiva. No Brasil, diferentes fontes de dados possibilitam o monitoramento contínuo dos fatores de risco para DCV, incluindo o Sistema de Informações sobre Mortalidade (SIM), o Sistema de Informações Hospitalares (SIH), e as pesquisas de saúde periódicas, como a Pesquisa Nacional de Saúde (PNS).⁹

Prevalências estimadas de SCV Ideal na população brasileira em 2013 e em 2019 foram inferiores a 1%, sendo menores em homens, indivíduos com 60 anos ou mais e residentes na região Sul do país.¹⁰ Considerando ainda que o escore de SCV demonstrou ser um instrumento eficiente de vigilância cardiovascular, há indicação de que o acompanhamento de indicadores de SCV, representados pelas metas biológicas e comportamentais do escore de SCV Ideal, representa uma estratégia necessária para prevenção primordial das DCV.

Os determinantes sociais desempenham um papel significativo no desenvolvimento de fatores de risco de DCVs bem como na morbidade e mortalidade por DCV. Ainda assim, o impacto dessas variáveis na SCV e nos desfechos de DCV permanece pouco estudado e compreendido²¹. Resultados de revisão sistemática recente mostraram que há uma associação entre SCV e os determinantes sociais de saúde tanto em países de baixa e média renda, quanto em países de alta renda, ou seja, a tendência da associação é a mesma independente da condição do país, sendo a situação do indivíduo o mais importante.²² Por exemplo, um estudo com os participantes do ELSA-Brasil²³ – que analisou a associação entre o escore de SCV e características sociodemográficas usando modelos de regressão quase-Poisson para servidores com 35 a 74 anos mostrou associação significativa do escore com faixa etária, sexo, raça, educação e riqueza.

Não se tem conhecimento até o momento de estudos que avaliaram associações entre SCV e determinantes socioeconômicos em uma amostra representativa da população brasileira. Neste sentido, o objetivo do presente estudo foi avaliar tais associações na população adulta brasileira com base em dados da Pesquisa Nacional de Saúde de 2019 utilizando modelos de regressão logística binária e multinomial.

MÉTODOS

Desenho do estudo

Estudo com um delineamento transversal, utilizando dados da PNS de 2019. Busca-se determinar associações entre o escore de SCV e determinantes socioeconômicos na população brasileira adulta.

Contexto

A Pesquisa Nacional de Saúde é um inquérito de saúde de base domiciliar, de âmbito nacional, realizada pelo Ministério da Saúde (MS) em parceria com o Instituto Brasileiro de Geografia e Estatística (IBGE) nos anos de 2013 e 2019.¹¹ A população pesquisada corresponde aos moradores de domicílios particulares permanentes do Brasil, exceto os localizados nos setores censitários especiais. A amostra da PNS de 2019 é uma subamostra da Amostra Mestra da Pesquisa Nacional de Amostra por Domicílios (PNAD), com avanços tanto no espalhamento geográfico como na precisão das estimativas.¹¹

Para a construção da amostra, primeiramente selecionou-se as unidades primárias de amostragem (UPAs), por amostragem aleatória simples, mantendo-se a estratificação da Amostra Mestra da Pesquisa Nacional por Amostra de Domicílios (PNAD). Então, uma quantidade fixa de domicílios particulares permanentes foi selecionada aleatoriamente em cada UPA, a partir do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE). Por fim, um morador (com idade de 15 anos ou mais) foi selecionado em cada domicílio da amostra, a partir de uma lista de moradores elegíveis construída no momento da entrevista, para responder à entrevista individual. A amostra

final para a PNS 2019 teve um total de 94.114 entrevistas individuais realizadas. Maior detalhamento sobre o delineamento da pesquisa pode ser encontrado na literatura.^{11,12}

Participantes

Os critérios de exclusão (e as respectivas quantidades de indivíduos excluídos em 2019) do presente estudo foram (1) idade inferior a 18 anos completos (n=2.431 indivíduos); (2) entrevistas individuais não realizadas (n=837); (3) questionários individuais respondidos por terceiros (n=1.711); (4) mulheres grávidas ou que não souberam informar se estavam grávidas (n= 3.131) e (5) resposta ausente de uma ou mais variáveis de SCV (n=8.510). Assim, a amostra final deste estudo constitui-se por 77.494 adultos com idade maior ou igual da 18 anos.

Variáveis

As metas de SCV utilizadas neste estudo foram adaptadas das metas originais propostas pela Sociedade Americana do Coração em 2010.³ As quatro metas comportamentais assim como as três metas biológicas foram avaliadas por meio de informações autorreferidas, como apresentado no Quadro 1. Cada uma das metas foi tratada como variável indicadora para resultado favorável (igual a 1) ou desfavorável (igual a 0). O escore de SCV foi obtido pela soma dos resultados das sete metas, variando de 0 a 7. Os códigos originais (indicados no dicionário da PNS 2019) das respectivas variáveis utilizadas e os resultados considerados favoráveis estão disponíveis no material suplementar.

Considera-se que o indivíduo apresenta SCV Ideal quando as sete metas são atingidas simultaneamente.³ O escore é dividido em metas comportamentais e biológicas.

As metas comportamentais incluem: (1) nunca ter fumado ou não fumar há mais de 12 meses (Tabagismo); (2) Índice de Massa Corpórea (IMC) menor que 25kg/m²; (3) prática de exercício físico ou esporte pelo menos 150 minutos por semana; (4) atingir as quatro metas de dieta relativas ao consumo recomendado de peixe, sal, bebidas açucaradas, frutas e hortaliças (Quadro 1). O consumo de grãos integrais não foi considerado para a meta de dieta, pois essa informação não foi coletada pela PNS 2019. As metas biológicas consideram: (1) ausência de diagnóstico médico de colesterol alto, (2) de hipertensão arterial e (3) de diabetes.

Utilizou-se 2 categorizações distintas para a variável resposta score de SCV. A primeira, denominada **SCV2cat** (binária), envolve a criação de uma variável indicadora de “saúde superior”, quando 6 a 7 metas são atingidas. A segunda forma de categorização, denominada **SCV3cat** (ternária), foi obtida dividindo-se os indivíduos em três níveis: superior (6 a 7 metas atingidas); intermediário (3 a 5 metas); e inferior (0 a 2 metas).

Os determinantes socioeconômicos avaliados (variáveis de exposição, com a categoria de referência em negrito) foram o nível de escolaridade (**até o ensino fundamental incompleto**; do ensino fundamental completo ao ensino médio incompleto; do ensino médio completo ao ensino superior incompleto e ensino superior completo), a situação de ocupação (com e **sem emprego**) e o índice de riqueza (5 classes separadas pelos quintis, sendo a 1^a referente aos domicílios com menor riqueza e a 5^a, aos domicílios com maior riqueza).¹³

As covariáveis sociodemográficas analisadas (com a categoria de referência em negrito) incluem: sexo (**masculino** e feminino), região do Brasil (**nordeste**; norte; centro-oeste; sudeste e sul), situação de domicílio (**urbano** e rural), faixa etária (**18-24 anos**; 25-39 anos; 40-59 anos e 60 ou mais), raça (cor da pele) (**parda**; branca; preta e

amarela/indígena/ignorada), estado civil (**solteiro**; casado e separado/ divorciado/viúvo). Ainda como covariável foi utilizada a variável status de doença crônica, ou simplesmente “doenças crônicas”, indicando se os indivíduos possuíam pelo menos uma das seguintes doenças crônicas (autorreferidas): asma (ou bronquite asmática); artrite ou reumatismo; dor crônica nas costas ou no pescoço, lombalgia, dor ciática e problemas nas vertebrae ou disco; distúrbios osteomusculares relacionados ao trabalho; doença no pulmão ou doença pulmonar obstrutiva crônica; câncer; insuficiência renal crônica; depressão; doença mental, como esquizofrenia, transtorno bipolar, psicose ou transtorno obsessivo compulsivo; no geral, doença crônica (física ou mental) ou doença de longa duração (de mais de 6 meses). A categoria de referência é não ter doenças crônicas.

Fonte de dados/Mensuração

Os dados são públicos e estão disponíveis na página eletrônica do IBGE¹¹ ou da PNS (<https://www.pns.icict.fiocruz.br/bases-de-dados/>).

Métodos Estatísticos

Para todas as análises deste estudo, utilizou-se informações do morador selecionado e a expansão da amostra foi realizada a partir do peso do morador selecionado com correção de não entrevista e calibração pela projeção de população para morador selecionado (variável V00291 da PNS 2019).¹¹

Prevalências e respectivos intervalos de 95% de confiança do escore de SCV (em 8 níveis, Figura 1) e de SCV3cat total e por características sociodemográficas ((Tabela 1)) foram estimados com base em expansão da amostra. As diferenças nas proporções

foram consideradas estatisticamente significantes quando não houve sobreposição de intervalos de confiança.¹⁴

Para avaliar a associação entre SCV2cat e determinantes socioeconômicos utilizou-se Regressão Logística Binária. e para SCV3cat, Regressão Logística Multinomial. Em ambas as abordagens, ajustou-se por covariáveis sociodemográficas e doenças crônicas, considerando ainda a amostragem complexa da PNS de 2019 na estimação dos parâmetros. Em tais análises estima-se razões de chances (RC) com seus respectivos os intervalos de 95% de confiança e os p-valores.

O modelo de regressão logística binária (variável SCV2cat) estimou as razões de chances de indivíduos apresentarem SCV superior entre diferentes categorias das variáveis explicativas (Tabela 2). Já para as análises da variável SCV3cat (Tabela 3 e Tabela Suplementar 1) optou-se por utilizar a categoria inferior de SCV (escore de 0 a 2, representando pior SCV) como categoria de referência. Dessa forma, os modelos de regressão logística multinomial, estimaram as razões de chances de indivíduos apresentarem SCV superior (em relação a inferior) e de apresentarem SCV intermediária (em relação a inferior) entre categorias das variáveis explicativas.

As variáveis sociodemográficas e status de doenças crônicas foram utilizadas como covariáveis nos modelos. O modelo cujos resultados estão apresentados na Tabela 3 foi o mais completo e incluiu todas as covariáveis selecionadas. O modelo denominado moderadamente ajustado não incluiu a variável doenças crônicas e o modelo denominado minimamente ajustado não incluiu as variáveis estado civil, região, situação de domicílio e doenças crônicas (Tabela Suplementar 1).

As análises foram conduzidas utilizando-se o *software* SAS versão 9.4 (SAS Institute Inc., NC, USA), com os comandos (*procedures*) *surveyfreq* e

surveylogistic, ou o *software* R¹⁵ versão 3.6.2, com auxílio do pacote survey para ajuste por amostragem complexa do delineamento de estudo da PNS 2019.

Aspectos éticos

A PNS 2019 foi aprovada pela Comissão Nacional de Ética em Pesquisa (CONEP) do Conselho Nacional de Saúde; CNS, em agosto de 2019, sob o número nº 3.529.376.¹¹ A pesquisa obedeceu à Resolução 196/96 do Conselho Nacional de Saúde, assegurando aos sujeitos de pesquisa sua voluntariedade, anonimato e possibilidade de desistência a qualquer momento do estudo.

RESULTADOS

A amostra final para este estudo foi constituída por 77.494 indivíduos. A Figura 1 apresenta a distribuição do escore de SCV (em 8 categorias) na população estudada. Apenas 0,5% (IC_{95%} 0,4; 0,6) da população atingiu a condição de SCV Ideal, ou seja, as sete metas simultaneamente. Considerando a categorização em 3 níveis (SCV3cat), tem-se que 8,9% (IC_{95%} 8,5; 9,3) dos participantes atingiu 6 a 7 metas (classe superior), 81,5% (IC_{95%} 81,0; 81,9) atingiu 3 a 5 metas (classe intermediária) e 9,6% (IC_{95%} 9,3; 10,0) atingiu 0 a 2 metas (classe inferior).

As características sociodemográficas da amostra são apresentadas na Tabela 1. Observou-se, a partir da expansão da amostra para a população brasileira adulta, maiores proporções de indivíduos com nível de escolaridade entre ensino médio completo e ensino superior incompleto (35,5%, IC_{95%} 34,9; 36,2), empregados (61,9%, IC_{95%} 61,3; 62,5), na maior classe do índice de riqueza (29,2%, IC_{95%} 28,2; 30,1), do sexo feminino (54,8%, IC_{95%} 54,1; 55,4), com idades entre 40 e 59 anos (37%, IC_{95%} 36,4; 37,6), brancos (44,6%, IC_{95%} 43,8; 45,4), casados (45,8%, IC_{95%} 45,1; 46,5), sem doenças crônicas (57,2%, IC_{95%} 56,5; 57,9), residentes na região Sudeste (45%, IC_{95%} 44,1; 45,9) e em área urbana (87,4%, IC_{95%} 87,0; 87,8).

Ainda na Tabela 1, considerando as distribuições conjuntas de SCV3cat com as demais variáveis, é possível observar que a maior prevalência de SCV superior está entre indivíduos com maior escolaridade (17,2%, IC_{95%} 16,1; 17,9), empregados (10,2%, IC_{95%} 9,6; 10,7), na maior classe do índice de riqueza (13,2%, IC_{95%} 12,3; 14,2), com idade entre 18 a 24 anos (18,3%, IC_{95%} 16,6; 20,1), solteiros (12%, IC_{95%} 11,3; 12,7), sem doença crônica (10,6%, IC_{95%} 10,0; 11,1) e da região urbana (9,5%, IC_{95%} 9,0; 9,9).

A Tabela 2 apresenta as estimativas das razões de chances (com respectivos intervalos de 95% de confiança e p-valores) de SCV2cat por determinantes socioeconômicos e covariáveis em um modelo de regressão logística binária. Observou-se que escolaridade e índice de riqueza estão associados à SCV2cat, após ajuste por covariáveis. Dessa maneira, interpreta-se que o nível de escolaridade, por exemplo, apresenta relação positiva entre os níveis (ajustando-se pelas outras variáveis), ou seja, à medida que o nível de escolaridade aumenta, a chance de ter escore de SCV favorável aumenta. Faixa etária, raça, estado civil, região e doenças crônicas foram covariáveis significantes.

A Tabela 3 apresenta as estimativas das razões de chances (com respectivos intervalos de 95% de confiança e p-valores) de SCV3cat por determinantes socioeconômicos e covariáveis em um modelo de regressão logística multinomial. Razões de chances da categoria superior de SCV vs. a inferior foram semelhantes aos observados no modelo com SCV2cat, com a exceção de ocupação e situação de domicílio que passaram a ser significantes. Indivíduos empregados vs. sem ocupação e residentes em área rural vs. urbana passaram a ter maior chance de apresentar SCV na categoria superior (em relação à inferior). Os resultados da comparação entre as categorias de SCV intermediária vs. inferior seguiram a mesma tendência (em relação à significância) observada na comparação Superior vs. Inferior exceto em algumas categorias das seguintes variáveis: escolaridade, riqueza, sexo, estado civil e região.

Para a avaliação de qualidade de ajuste (referente aos modelos cujos resultados são apresentados nas Tabela 3 e 2, respectivamente), utilizou-se o coeficiente de determinação R^2 de McFadden¹⁶ que, para o modelo multinomial ajustado por todas as covariáveis, foi de 0,11 e, para o modelo binário, foi de 0,09. Associações entre SCV3cat

e determinantes sociais apresentados na Tabela 3 foram similares aos de modelos minimamente e moderadamente ajustados, apresentados na Tabela Suplementar 1.

DISCUSSÃO

Os resultados do presente estudo demonstraram haver associação entre SCV e determinantes socioeconômicos na população adulta brasileira. Observou-se que estar empregado (ocupado) e possuir maiores níveis de escolaridade e riqueza aumentaram as chances de indivíduos apresentarem melhor saúde cardiovascular (com base em SCV3cat), ajustando-se por faixa etária, raça, estado civil, região, situação de domicílio e doenças crônicas (covariáveis significantes).

A prevalência do nível superior de SCV (6-7 metas) na população brasileira adulta foi de apenas 8,9%. Para o nível intermediário foi de 81,5% e o nível inferior foi de 9,6%. Contrastou-se um modelo com variável resposta binária (SCV2cat), mais utilizado e mais simples, com um modelo de resposta ternária (SCV3cat), de maior complexidade e detalhamento, mas ainda bem mais parcimonioso que um modelo que utilizasse o escore de SCV em 8 categorias (resultados não apresentados). O modelo binário combina em uma mesma categoria os mais de 90% que possuem SCV intermediária e inferior, o que não nos pareceu apropriado. A comparação entre os modelos de regressão binária e multinomial (Tabelas 2 e 3, respectivamente) mostra que, apesar das direções das associações para as variáveis de exposição (escolaridade, índice de riqueza e situação de ocupação) se manterem, o modelo multinomial apresenta maior poder de discriminação entre os subgrupos populacionais. Ainda, quando SCV é separado em 3 categorias, a variável situação de ocupação e a covariável situação de domicílio passam a ser significantes para ambas as comparações (entre categorias Superior vs Inferior e Intermediária vs. Inferior de SCV).

A semelhança entre os resultados do modelo ajustado por todas as covariáveis selecionadas (Tabela 3) com resultados de modelos moderadamente e minimamente

ajustados (Tabela suplementar 1) evidencia a robustez dos modelos. A comparação de diferentes tipos de ajustes é utilizada com certa frequência em estudos epidemiológicos.

A principal limitação do estudo é a utilização de medidas autorreferidas da população. Por tal razão, as prevalências das metas biológicas e comportamentais de SCV podem estar subestimadas. Como os dados de morbidade autorreferida, incluindo diagnóstico médico de diabetes mellitus, hipercolesterolemia e hipertensão arterial, dependem do acesso aos serviços de saúde para o diagnóstico, aqueles usuários com acesso restrito a tais serviços têm menor oportunidade de diagnóstico médico. No entanto, o uso de dados autorreferidos para avaliação das metas de SCV já foi utilizado em outros estudos¹⁷⁻¹⁹ e possui boa correlação com as medidas biológicas.²⁰ Tem-se que o desenho transversal do estudo que não permite se estabelecer uma relação causal entre SCV e determinantes sociais bem o número reduzido de representantes de subgrupos de variáveis sociodemográficas como, por exemplo, a categoria de raça Amarela/Indígena/Ignorada. Ainda, cita-se como limitação que os métodos usuais de diagnóstico de ajuste de modelos não estão disponíveis para modelagem considerando planos amostrais complexos. O coeficiente de determinação R^2 de McFadden para os modelos logísticos multinomial e binário indicam que a redução proporcional na estatística $-2\log(\text{verossimilhança})$ ²¹ foram de 11% e 9% respectivamente. Tais coeficientes também podem ser interpretados como a razão entre a) e b), sendo a) a informação adquirida estimada quando o modelo ajustado é utilizado (em comparação ao modelo sem covariáveis) e b) a informação estimada adquirida potencialmente se o modelo fosse ajustado por todas as possíveis covariáveis (comparado ao modelo sem covariáveis).^{22,23} Geralmente modelos logísticos apresentam valores de coeficientes R^2

bem inferiores aos observados em modelos de regressão lineares.²⁴ Como pode-se observar, a interpretação desses coeficientes é também bastante distinta nessas situações.

Mesmo com tais limitações, o presente estudo é de suma importância por utilizar dados de uma pesquisa de âmbito nacional desenvolvida pelo Instituto Brasileiro de Geografia e Estatística em parceria com o Ministério da Saúde, com delineamento cuidadosamente planejado e executado e de metodologia validada. A amostra selecionada é representativa da população adulta brasileira e apresenta dados atualizados acerca da saúde bem como de determinantes sociais, econômicos e demográficos no Brasil.

Diferentes fatores podem estar atrelados ao alcance das metas de SCV. Nos Estados Unidos, em 2015, as oportunidades mais significativas para reduzir morbidade e mortalidade por DCV envolvem em determinantes sociais da saúde.⁷ Tais condições sociais podem estar fortemente associadas a escolhas de alimentação mais saudáveis e de atividades físicas mais frequentes.⁷ A escolaridade, por exemplo, é um dos indicadores mais utilizados nos Estados Unidos e fornece os resultados mais consistentes em relação aos desfechos de DCV.^{7,25,26} Inclusive, maiores prevalência de fatores de riscos cardiovasculares foram encontradas em pessoas com baixos níveis de escolaridade, em um estudo anterior.^{7,26} A combinação desfavorável de se ter um menor nível de escolaridade e estar sem ocupação durante a vida adulta pode criar uma forma de estresse crônico, que se mostrou associado a um aumento da carga alostática (quantidade de energia metabólica necessária para manter o equilíbrio fisiológico) e suas consequências para a saúde.⁷

Os determinantes sociais parecem desempenhar um papel significativo no desenvolvimento de fatores de risco de DCVs bem como na morbidade e mortalidade por DCV.^{26,27} Esse é um tema que precisa ser mais estudado no Brasil tendo em vista a grande

disparidade socioeconômica observada na população. Recentemente, o estudo ELSA-Brasil²⁸ reportou associações entre o escore de SCV e características sociodemográficas (como faixa etária, sexo, raça, renda mensal familiar e escolaridade) para servidores com 35 a 74 anos. Tais resultados apontam para a mesma direção dos observados no presente estudo, porém foram obtidos utilizando o modelo de regressão quase-Poisson para a variável resposta escore de SCV em 8 níveis. Além disso, a amostra é grande (n = 13.356) porém se limita a servidores de 6 universidades de grandes cidades do país (São Paulo, Belo Horizonte, Porto Alegre, Salvador, Rio de Janeiro, e Vitória), sendo menor e menos representativa da população brasileira que a do presente. Em termos de metodologia adotada, avaliou-se mais variáveis de exposição e covariáveis e utilizou-se outros modelos estatísticos para as diferentes formas da variável resposta.

Portanto, o presente estudo representa um importante conjunto de resultados que servem para auxiliar no desenvolvimento de intervenções mais direcionadas e eficazes, para o alcance de maiores prevalências de SCV ideal na população brasileira adulta.

REFERÊNCIAS

1. Malta DC, Teixeira R, de Oliveira GMM, Ribeiro ALP. Cardiovascular disease mortality according to the Brazilian information system on mortality and the global burden of disease study estimates in Brazil, 2000-2017. *Arq Bras Cardiol.* 2020;115(2):152-160. doi:10.36660/abc.20190867
2. WHO - World Health Organization. Cardiovascular diseases (CVDs). Accessed October 30, 2022. Cardiovascular diseases (CVDs)
3. Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: The American Heart Association's strategic impact goal through 2020 and beyond. *Circulation.* 2010;121(4):586-613. doi:10.1161/CIRCULATIONAHA.109.192703
4. Younus A, Aneni EC, Spatz ES, et al. A Systematic Review of the Prevalence and Outcomes of Ideal Cardiovascular Health in US and Non-US Populations. *Mayo Clin Proc.* 2016;91(5):649-670. doi:10.1016/j.mayocp.2016.01.019
5. Joseph JJ, Bennett A, Echouffo Tcheguie JB, et al. Ideal cardiovascular health, glycaemic status and incident type 2 diabetes mellitus: the REasons for Geographic and Racial Differences in Stroke (REGARDS) study. *Diabetologia.* 2019;62(3):426-437. doi:10.1007/s00125-018-4792-y
6. Bousquet-Santos K, Chen R, Kubzansky LD. A sad heart: Depression and favorable cardiovascular health in Brazil. *Prev Med (Baltim).* 2021;142:106378. doi:https://doi.org/10.1016/j.ypmed.2020.106378
7. Havranek EP, Mujahid MS, Barr DA, et al. Social determinants of risk and outcomes for cardiovascular disease: A scientific statement from the American Heart Association. *Circulation.* 2015;132(9):873-898. doi:10.1161/CIR.0000000000000228
8. Harper S, Lynch J, Smith GD. Social determinants and the decline of cardiovascular diseases: Understanding the links. *Annu Rev Public Health.* 2011;32:39-69. doi:10.1146/annurev-publhealth-031210-101234
9. Oliveira GMM de, Brant LCC, Polanczyk CA, et al. Estatística Cardiovascular – Brasil 2021. *Arq Bras Cardiol.* 2022;118(1):115-373. doi:10.36660/abc.20211012
10. Velasquez-Melendez G, Felisbino-Mendes MS, Matozinhos FP, Claro R, Gomes CS, Malta DC. Prevalência de saúde cardiovascular ideal na população Brasileira – pesquisa nacional de saúde (2013). *Revista Brasileira de Epidemiologia.* 2015;18:97-108. doi:10.1590/1980-5497201500060009
11. IBGE - Coordenação de Trabalho e Rendimento. *Pesquisa Nacional de Saúde 2019: Informações Sobre Domicílios, Acesso e Utilização Dos Serviços de Saúde: Brasil, Grandes Regiões e Unidades Da Federação.* IBGE; 2020.
12. Stopa SR, Szwarcwald CL, Oliveira MM de, et al. Pesquisa Nacional de Saúde 2019: histórico, métodos e perspectivas. *Epidemiol Serv Saude.* 2020;29(5):e2020315. doi:10.1590/S1679-49742020000500004
13. Powell-Wiley TM, Baumer Y, Baah FO, et al. Social Determinants of Cardiovascular Disease. *Circ Res.* 2022;130(5):782-799. doi:10.1161/CIRCRESAHA.121.319811
14. Bonita R, Beaglehole R, Kjellström T, Organization WH. Basic epidemiology. Published online 2006:xi, 212 p.

15. R Development Core Team (2009). R: A language and environment for statistical computing. <http://www.R-project.org>
16. Shtatland E, Kleinman K, Cain E. One more time about R2 measures of fit in logistic regression. *NESUG 15 Proceedings*. 2002;15.
17. Shockey TM, Sussell AL, Odom EC. *Morbidity and Mortality Weekly Report.*; 2016. http://www.cdc.gov/mmwr/cme/conted_info.html#weekly.http://www.cdc.gov/brfss/annual_data/2013/pdf/overview_2013.pdf.
18. Burroughs Peña MS, Mbassa RS, Slopen NB, Williams DR, Buring JE, Albert MA. Cumulative psychosocial stress and ideal cardiovascular health in older women: Data by race/ethnicity. *Circulation*. 2019;139(17):2012-2021. doi:10.1161/CIRCULATIONAHA.118.033915
19. Thomas SC, Umer A, Commodore-Mensah Y, Davidov D, Abildso CG. Length of Residence and Cardiovascular Health among Afro-Caribbean Immigrants in New York City. *J Racial Ethn Health Disparities*. 2019;6(3):487-496. doi:10.1007/s40615-018-00547-8
20. Moreira AD, Gomes CS, Machado ÍE, Carvalho D, Felisbino-Mendes MS. Saúde cardiovascular e validação do escore autorreferido no Brasil: uma análise da Pesquisa Nacional de Saúde. *Cien Saude Colet*. 2020;25(11):4259-4268. doi:10.1590/1413-812320202511.31442020
21. Menard S. Coefficients of Determination for Multiple Logistic Regression Analysis. *Am Stat*. 2000;54(1):17-24. doi:10.2307/2685605
22. Kent JT. Information Gain and a General Measure of Correlation. *Biometrika*. 1983;70(1):163-173. doi:10.2307/2335954
23. Hastie T. *A Closer Look at the Deviance*. Vol 41.; 1987.
24. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. John Wiley & Sons, Inc.; 2000. doi:10.1002/0471722146
25. Winkleby MA, Jatulis DE, Frank E, Fortmann SP. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am J Public Health*. 1992;82(6):816-820. doi:10.2105/AJPH.82.6.816
26. Kaplan GA, Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation*. 1993;88(4):1973-1998. doi:10.1161/01.CIR.88.4.1973
27. Qureshi F, Bousquet-Santos K, Okuzono SS, et al. The social determinants of ideal cardiovascular health: A global systematic review. *Ann Epidemiol*. 2022;76:20-38. doi:<https://doi.org/10.1016/j.annepidem.2022.09.006>
28. Machado LBM, Silva BLS, Garcia AP, et al. Ideal cardiovascular health score at the ELSA-Brasil baseline and its association with sociodemographic characteristics. *Int J Cardiol*. 2018;254:333-337. doi:10.1016/j.ijcard.2017.12.037

TABELAS, QUADROS E FIGURAS

Metas ^a	Métrica utilizada
Comportamentais	
Tabagismo	Nunca ter fumado ou não fumar há mais de 12 meses.
IMC ^b	< 25 kg/m ² (a partir de medidas de peso e altura autorreferidas e imputadas).
Atividade Física	Prática de exercício físico ou esporte ≥ 150 min/semana.
Dieta	4 metas de dieta atingidas simultaneamente: 1) Frutas e vegetais: consumo ≥ 1 vez/dia; 2) Peixe: consumo ≥ 2 vezes/semana; 3) Sódio: consumo de sal na comida preparada na hora e nos alimentos industrializados em quantidade adequada, baixa ou muito baixa; 4) Açúcar: consumo de refrigerantes e/ou sucos industrializados <5 dias/semana.
Biológicas	
Colesterol Total	Ausência de diagnóstico médico autorreferido de colesterol alto.
Pressão arterial	Ausência de diagnóstico médico autorreferido de hipertensão arterial.
Glicose	Ausência de diagnóstico médico autorreferido de diabetes.

Quadro 1- Metas de saúde cardiovascular.

Notas:

^a As metas propostas neste estudo foram adaptadas de AHA³;

^b IMC: Índice de Massa Corporal.

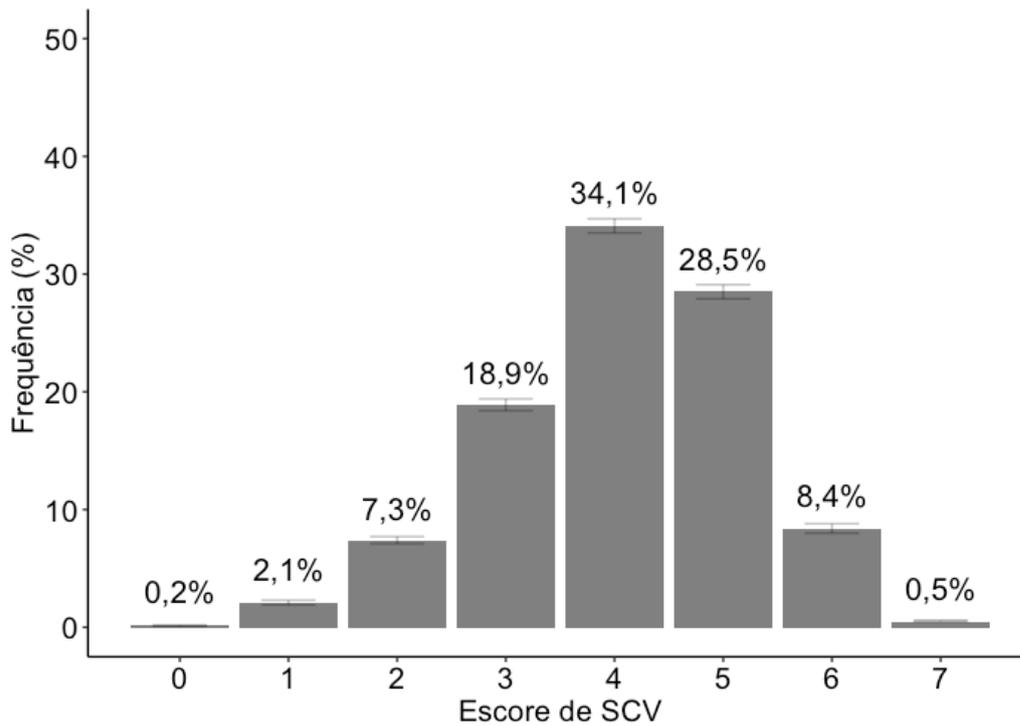


Figura 1: Distribuição do Escore de SCV^a com respectivos Intervalos de 95% de Confiança^b, estimados para a população adulta brasileira em 2019^c (Dados da PNS 2019).

Notas:

^a SCV: Saúde Cardiovascular;

^b Os intervalos de 95% de confiança para as prevalências de valores do Escore de SCV (de 0 a 7) foram, respectivamente: (0,1 a 0,2); (1,9 a 2,3); (7,1 a 7,7); (18,4 a 19,4); (33,5 a 34,7); (27,9 a 29,1); (8,0 a 8,8); e (0,5 a 0,6);

^c Estimativas populacionais obtidas a partir da expansão da amostra.

Tabela 1: Distribuições de variáveis sociodemográficas por categorias de saúde cardiovascular – PNS^a 2019.

Variáveis	Total ^b N (%; IC)	Categorias de Saúde Cardiovascular ^c		
		Inferior (Escore ≤ 2) % (IC)	Intermediária (Escore 3; 5) % (IC)	Superior (Escore ≥ 6) % (IC)
Total	140.036.509 (100,0)	9,6 (9,3; 10,0)	81,5 (81,0; 81,9)	8,9 (8,5; 9,3)
Escolaridade (Nível de ensino)				
Até fundamental incompleto	46.739.391 (33,4; 32,7; 34,1)	16,3 (15,6; 17,1)	80,5 (79,7; 81,3)	3,1 (2,8; 3,5)
Fundamental completo à médio incompleto	19.367.870 (13,8; 13,4; 14,3)	9,3 (8,2; 10,3)	83,6 (82,2; 84,9)	7,2 (6,3; 8,1)
Médio completo à superior incompleto	49.745.301 (35,5; 34,9; 36,2)	5,5 (5,0; 6,0)	83,5 (82,6; 84,3)	11,0 (10,3; 11,7)
Superior completo	24.183.947 (17,3; 16,6; 18,0)	5,4 (4,8; 6,0)	77,4 (76,2; 78,6)	17,2 (16,1; 18,4)
Situação de Ocupação				
Empregado	86.692.821 (61,9; 61,3; 62,5)	6,5 (6,1; 6,9)	83,3 (82,7; 83,9)	10,2 (9,6; 10,7)
Sem ocupação	53.343.688 (38,1; 37,5; 38,7)	14,7 (14,1; 15,3)	78,4 (77,7; 79,2)	6,9 (6,4; 7,4)
Índice de Riqueza				
1ª classe	16.671.019 (11,9; 11,5; 12,3)	10,7 (9,9; 11,5)	83,9 (83,0; 84,9)	5,3 (4,7; 6,0)
2ª classe	22.204.156 (15,9; 15,3; 16,4)	11,6 (10,8; 12,4)	82,3 (81,3; 83,2)	6,1 (5,6; 6,7)
3ª classe	27.159.566 (19,4; 18,8; 20,0)	10,3 (9,4; 11,1)	82,8 (81,8; 83,8)	6,9 (6,3; 7,5)
4ª classe	33.166.362 (23,7; 23,0; 24,4)	9,3 (8,5; 10,0)	81,8 (80,7; 82,9)	8,9 (8,2; 9,7)
5ª classe	40.835.406 (29,2; 28,2; 30,1)	7,9 (7,2; 8,6)	78,8 (77,8; 79,9)	13,2 (12,3; 14,2)
Sexo				
Masculino	63.359.986 (45,2; 44,6; 45,9)	7,9 (7,4; 8,4)	83,4 (82,7; 84,1)	8,7 (8,2; 9,3)
Feminino	76.676.523 (54,8; 54,1; 55,4)	11,1 (10,5; 11,6)	79,9 (79,2; 80,5)	9,1 (8,6; 9,6)
Faixa Etária (anos)				
18 a 24	16.797.883 (12,0; 11,5; 12,5)	0,6 (0,3; 0,8)	81,1 (79,4; 82,9)	18,3 (16,6; 20,1)
25 a 39	40.026.391 (28,6; 28,0; 29,2)	2,9 (2,5; 3,3)	85,7 (84,9; 86,5)	11,4 (10,6; 12,1)
40 a 59	51.822.239 (37,0; 36,4; 37,6)	10,8 (10,2; 11,5)	82,3 (81,5; 83,0)	6,9 (6,4; 7,5)
60 ou mais	31.389.995 (22,4; 21,8; 23,0)	21,0 (20,1; 22,0)	74,9 (73,9; 75,9)	4,1 (3,6; 4,5)
Raça/Etnia				
Parda	59.763.814 (42,7; 42,0; 43,4)	9,7 (9,1; 10,2)	82,1 (81,4; 82,8)	8,2 (7,7; 8,8)
Branca	62.440.066 (44,6; 43,8; 45,4)	9,3 (8,7; 9,8)	80,8 (80,1; 81,6)	9,9 (9,3; 10,6)
Preta	15.752.117 (11,2; 10,8; 11,7)	10,6 (9,6; 11,6)	81,7 (80,4; 83,0)	7,7 (6,7; 8,6)
Amarela/Indígena/Ignorada	2.080.511 (1,5; 1,3; 1,7)	11,4 (7,8; 14,9)	79,6 (75,2; 83,9)	9,0 (6,4; 11,7)
Estado Civil				
Solteiro	55.627.974 (39,7; 39,0; 40,4)	5,7 (5,3; 6,2)	82,3 (81,4; 83,1)	12,0 (11,3; 12,7)
Casado	64.089.556 (45,8; 45,1; 46,5)	10,5 (9,9; 11,0)	82,1 (81,5; 82,8)	7,4 (6,9; 7,9)

Separado/ Divorciado/ Viúvo	20.318.979 (14,5; 14,1; 14,9)	17,5 (16,4; 18,6)	77,1 (75,9; 78,3)	5,4 (4,8; 6,0)
Doença Crônica				
Pelo menos 1	59.969.579 (42,8; 42,1; 43,5)	14,2 (13,6; 14,8)	79,1 (78,3; 79,8)	6,7 (6,2; 7,2)
Sem doença	80.066.930 (57,2; 56,5; 57,9)	6,2 (5,8; 6,5)	83,2 (82,6; 83,8)	10,6 (10,0; 11,1)
Região				
Norte	10.178.075 (7,3; 7,0; 7,6)	6,8 (6,2; 7,4)	83,7 (82,7; 84,7)	9,5 (8,7; 10,3)
Nordeste	35.417.215 (25,3; 24,7; 25,9)	8,7 (8,1; 9,2)	81,9 (81,2; 82,6)	9,5 (8,9; 10,1)
Sudeste	63.001.662 (45,0; 44,1; 45,9)	10,3 (9,6; 10,9)	81,1 (80,2; 82,0)	8,7 (8,0; 9,4)
Sul	20.852.598 (14,9; 14,4; 15,3)	10,7 (10,0; 11,5)	81,3 (80,2; 82,3)	8,0 (7,1; 8,9)
Centro-Oeste	10.586.959 (7,6; 7,3; 7,9)	9,6 (8,7; 10,5)	80,5 (79,3; 81,7)	9,9 (8,9; 11,0)
Situação de Domicílio				
Urbano	122.412.326 (87,4; 87,0; 87,8)	9,7 (9,3; 10,1)	80,8 (80,3; 81,4)	9,5 (9,0; 9,9)
Rural	17.624.183 (12,6; 12,2; 13,0)	9,0 (8,3; 9,7)	85,7 (84,9; 86,5)	5,3 (4,7; 5,8)

Notas:

^a PNS: Pesquisa Nacional de Saúde;

^b Os percentuais calculados da primeira coluna são referentes ao total de participantes;

^c Os demais são referentes ao total da linha correspondente.

Tabela 2. Razões de chances, IC95% e p-valores de modelos de regressão logística para a variável resposta saúde cardiovascular ideal (escore de SCV^a ≥ 6 versus demais categorias) – PNS^b 2019.

Variáveis	RC ^c	IC95%	p-valor
Escolaridade (Nível de ensino)			<0,001
Fundamental completo a médio incompleto vs. Até FI ^d	1,59	(1,31; 1,91)	<0,001
Médio completo a superior incompleto vs. Até FI	2,16	(1,85; 2,51)	<0,001
Superior completo vs. Até FI	3,97	(3,34; 4,73)	<0,001
Situação de Ocupação			0,422
Empregado	1,05	(0,94; 1,17)	0,422
Índice de Riqueza			<0,001
2 ^a vs. 1 ^a classe	1,00	(0,85; 1,19)	0,972
3 ^a vs. 1 ^a classe	1,07	(0,89; 1,29)	0,478
4 ^a vs. 1 ^a classe	1,35	(1,11; 1,64)	0,003
5 ^a vs. 1 ^a classe	1,85	(1,51; 2,28)	<0,001
Sexo			0,074
Feminino vs. Masculino	1,09	(0,99; 1,21)	0,074
Faixa Etária (anos)			<0,001
25 a 39 vs. 18 a 24	0,53	(0,46; 0,62)	<0,001
40 a 59 vs. 18 a 24	0,39	(0,34; 0,46)	<0,001
60 ou mais vs. 18 a 24	0,30	(0,25; 0,36)	<0,001
Raça/Etnia			0,095
Branca vs. Parda	1,11	(1,00; 1,24)	0,049
Preta vs. Parda	0,93	(0,79; 1,10)	0,397
Amarela/Indígena/Ignorada vs. Parda	0,99	(0,70; 1,42)	0,973
Estado Civil			<0,001
Casado vs. Solteiro	0,78	(0,70; 0,86)	<0,001
Separado/Divorciado/Viúvo vs. Solteiro	0,78	(0,67; 0,91)	0,001
Doença Crônica			<0,001
Pelo menos uma vs. Nenhuma	0,70	(0,64; 0,78)	<0,001
Região			<0,001
Norte vs. Nordeste	0,97	(0,86; 1,09)	0,558
Sudeste vs. Nordeste	0,63	(0,56; 0,70)	<0,001
Sul vs. Nordeste	0,62	(0,53; 0,72)	<0,001
Centro-Oeste vs. Nordeste	0,79	(0,69; 0,90)	<0,001
Situação de Domicílio			0,137
Rural vs. Urbano	0,89	(0,77; 1,04)	0,137

Notas:

^a SCV: Saúde Cardiovascular;

^b PNS: Pesquisa Nacional de Saúde;

^c RC: Razão de chance;

^d FI: Fundamental Incompleto.

Tabela 3: Razões de chances, IC95% e p-valores de modelos de regressão multinomial para a variável resposta SCV^a ideal em 3 categorias (inferior 0-2, intermediária 3-5 e superior ou ideal 6-7) – PNS^b 2019.

Variáveis	p-valor geral ^c	Comparações de Categorias de Saúde Cardiovascular					
		Superior vs. Inferior			Intermediária vs. Inferior		
		RC ^d	IC95%	p-valor	RC	IC95%	P-valor
Escolaridade (Nível de ensino)	<0,001						
Fundamental completo a médio incompleto vs. Até FI ^e		1,65	(1,31; 2,08)	<0,001	1,07	(0,92; 1,24)	0,407
Médio completo a superior incompleto vs. Até FI		3,08	(2,55; 3,72)	<0,001	1,50	(1,32; 1,70)	<0,001
Superior completo vs. Até FI		6,57	(5,24; 8,23)	<0,001	1,75	(1,50; 2,05)	<0,001
Situação de Ocupação	<0,001						
Empregado		1,37	(1,18; 1,59)	<0,001	1,33	(1,21; 1,47)	<0,001
Índice de Riqueza	<0,001						
2 ^a vs. 1 ^a classe		0,99	(0,81; 1,22)	0,941	0,98	(0,86; 1,12)	0,800
3 ^a vs. 1 ^a classe		1,20	(0,95; 1,52)	0,136	1,12	(0,96; 1,31)	0,165
4 ^a vs. 1 ^a classe		1,57	(1,23; 2,01)	<0,001	1,17	(1,00; 1,38)	0,055
5 ^a vs. 1 ^a classe		2,33	(1,80; 3,03)	<0,001	1,27	(1,06; 1,52)	0,010
Sexo	<0,001						
Feminino vs. Masculino		0,91	(0,80; 1,04)	0,184	0,83	(0,75; 0,91)	<0,001
Faixa Etária (anos)	<0,001						
25 a 39 vs. 18 a 24		0,11	(0,07; 0,19)	<0,001	0,20	(0,12; 0,33)	<0,001
40 a 59 vs. 18 a 24		0,03	(0,02; 0,05)	<0,001	0,06	(0,04; 0,11)	<0,001
60 ou mais vs. 18 a 24		0,01	(0,01; 0,02)	<0,001	0,04	(0,02; 0,07)	<0,001
Raça/Etnia	<0,001						
Branca vs. Parda		1,30	(1,13; 1,51)	<0,001	1,18	(1,06; 1,31)	0,002
Preta vs. Parda		0,87	(0,71; 1,07)	0,180	0,93	(0,81; 1,06)	0,278
Amarela/Índigena/Ignorada vs. Parda		0,85	(0,53; 1,38)	0,520	0,85	(0,59; 1,23)	0,390
Estado Civil	<0,001						
Casado vs. Solteiro		0,75	(0,65; 0,86)	<0,001	0,96	(0,86; 1,07)	0,427
Separado/Divorciado/Viúvo vs. Solteiro		0,70	(0,58; 0,85)	<0,001	0,89	(0,79; 1,00)	0,058
Doença Crônica	<0,001						
Pelo menos 1		0,41	(0,36; 0,47)	<0,001	0,56	(0,51; 0,61)	<0,001
Região	<0,001						
Norte vs. Nordeste		1,13	(0,96; 1,33)	0,136	1,18	(1,05; 1,33)	0,006
Sudeste vs. Nordeste		0,48	(0,41; 0,56)	<0,001	0,75	(0,67; 0,84)	<0,001
Sul vs. Nordeste		0,46	(0,38; 0,56)	<0,001	0,73	(0,65; 0,82)	<0,001
Centro-Oeste vs. Nordeste		0,62	(0,51; 0,74)	<0,001	0,77	(0,68; 0,88)	<0,001
Situação de Domicílio	<0,001						
Rural vs. Urbano		1,25	(1,03; 1,50)	0,023	1,42	(1,27; 1,60)	<0,001

Notas:

^a SCV: Saúde Cardiovascular;

^b PNS: Pesquisa Nacional de Saúde;

^c Análise do efeito global da variável;

^d RC: Razão de chance;

^e FI: Fundamental Incompleto.

MATERIAL SUPLEMENTAR

Variáveis	Descrição das variáveis do questionário da PNS ^a 2019	Método de cálculo
Tabagismo	P050. Atualmente, o(a) Sr.(a) fuma algum produto do tabaco? Opções: 1- Sim, diariamente; 2- Sim, menos que diariamente; 3- Não fumo atualmente. P052. E no passado, o(a) Sr.(a) fumou algum produto do tabaco? Opções: 1- Sim, diariamente; 2- Sim, menos que diariamente; 3- Não nunca fumei. P05901. Número de anos que parou de fumar. Opções: Número de anos.	Resultado Favorável (1): Se indivíduo é não fumante [P050 = 3 e P052 = 3] ou é ex-fumante há mais de um ano [P050=3 e P052 = (1 ou 2) e P05901 > 1]
IMC	P00104. Peso; Final (em kg) Opções: 3 inteiros e 1 casa decimal P00404. Altura; Final (em cm) Opções: 3 inteiros	Resultado Favorável (1): Peso [P00104] dividido pelo quadrado da Altura [P00404] < 0,0025
Atividade Física	P035. Quantos dias por semana o(a) Sr(a) costuma (costumava) praticar exercício físico ou esporte? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P03701. Em geral, no dia que o(a) Sr(a) pratica exercício ou esporte, quanto tempo em horas dura essa atividade? Opções: Quantidade de horas. P03702. Em geral, no dia que o(a) Sr(a) pratica (praticava) exercício ou esporte, quanto tempo em minutos dura essa atividade? Opções: Quantidade de minutos.	Resultado Favorável (1): Quantidade de exercícios físicos praticados em minutos por dia [P03701*60+P03702] vezes a quantidade de dias por semana [P035] ≥ 150
Dieta	P015. Em quantos dias da semana o(a) Sr(a) costuma comer peixe? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P02001. Em quantos dias da semana o(a) Sr(a) costuma tomar suco de caixinha/lata ou refresco em pó? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P02002. Em quantos dias da semana o(a) Sr(a) costuma tomar refrigerante? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P018. Em quantos dias da semana o(a) Sr(a) costuma comer frutas? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P00901. Em quantos dias da semana, o(a) Sr(a) costuma comer pelo menos um tipo de verdura ou legume (sem contar batata, mandioca, cará ou inhame) como alface, tomate, couve, cenoura, chuchu, berinjela, abobrinha? Opções: 1 a 7 dias ou 0; Nunca ou menos de uma vez por semana. P02601. Considerando a comida preparada na hora e os alimentos industrializados, o(a) Sr(a) acha que o seu consumo de sal é: Opções: 1- Muito Alto; 2- Alto; 3- Adequado; 4- Baixo; 5- Muito Baixo.	Resultado Favorável (1): Indivíduos que atingiram as 4 metas de dieta: (1) Consumo de frutas e vegetais todos os dias [P018 = 7 e P00901 = 7]; (2) Consumo de peixe 2 vezes por semana [P015 ≥ 2]; (3) Consumo de sal adequado, baixo ou muito baixo [P02601 ≥ 3]; e (4) Consumo de refrigerantes e/ou sucos industrializados em menos de 5 dias da semana [P02001+P02002 < 5].
Colesterol Total	Q060. Algum médico já lhe deu o diagnóstico de colesterol alto? Opções: 1- Sim; 2- Não.	Resultado Favorável (1): Ausência de diagnóstico médico de colesterol alto [Q060 = 2].
Pressão arterial	Q00201. Algum médico já lhe deu o diagnóstico de hipertensão arterial (pressão alta)? Opções: 1- Sim; 2- Não.	Resultado Favorável (1): Ausência de diagnóstico médico de hipertensão arterial [Q00201 = 2].
Glicose	Q03001. Algum médico já lhe deu o diagnóstico de diabetes? Opções: 1- Sim; 2- Não.	Resultado Favorável (1): Ausência de diagnóstico médico de diabetes [Q03001 = 2].

Quadro Suplementar 1. Construção e métodos de cálculo das variáveis, Pesquisa Nacional de Saúde 2019, Brasil

Legenda:

^a PNS – Pesquisa Nacional de Saúde

Tabela Suplementar 1. Razões de chances, IC95% e p-valores de modelos de regressão multinomial, minimamente e moderadamente ajustados, para a variável resposta SCV^a ideal em 3 categorias (inferior 0-2, intermediária 3-5 e superior ou ideal 6-7) – PNS^b 2019.

Variáveis	Modelo minimamente ajustado						Modelo moderadamente ajustado							
	P geral ^c	Superior vs. Inferior			Intermediária vs. Inferior			P geral	Superior vs. Inferior			Intermediária vs. Inferior		
		RC ^d	IC95%	p	RC	IC95%	p		RC	IC95%	p	RC	IC95%	p
Escolaridade	<0,001							<0,001						
EF ^e completo a EM ^f incompleto vs. Até EF ^g		1,72	(1,37; 2,16)	<0,001	1,09	(0,95; 1,26)	0,235		1,72	(1,37; 2,16)	<0,001	1,10	(0,95; 1,27)	0,199
EM completo a ES ^h incompleto vs. Até EF ⁱ		3,42	(2,83; 4,13)	<0,001	1,57	(1,39; 1,79)	<0,001		3,21	(2,66; 3,88)	<0,001	1,54	(1,36; 1,75)	<0,001
ES completo vs. Até EF ⁱ		7,51	(6,00; 9,40)	<0,001	1,86	(1,59; 2,18)	<0,001		6,66	(5,31; 8,34)	<0,001	1,78	(1,52; 2,08)	<0,001
Situação de Ocupação	<0,001							<0,001						
Empregado		1,40	(1,21; 1,62)	<0,001	1,38	(1,24; 1,52)	<0,001		1,43	(1,23; 1,65)	<0,001	1,37	(1,24; 1,52)	<0,001
Índice de Riqueza	<0,001							<0,001						
2 ^a vs. 1 ^a classe		0,77	(0,64; 0,94)	0,010	0,80	(0,70; 0,90)	<0,001		0,97	(0,79; 1,19)	0,787	0,96	(0,84; 1,10)	0,589
3 ^a vs. 1 ^a classe		0,79	(0,64; 0,98)	0,030	0,82	(0,71; 0,94)	0,005		1,19	(0,94; 1,50)	0,148	1,11	(0,94; 1,29)	0,214
4 ^a vs. 1 ^a classe		0,88	(0,72; 1,09)	0,244	0,80	(0,70; 0,91)	<0,001		1,53	(1,20; 1,95)	<0,001	1,15	(0,98; 1,35)	0,092
5 ^a vs. 1 ^a classe		1,17	(0,93; 1,47)	0,173	0,83	(0,71; 0,97)	0,016		2,23	(1,72; 2,89)	<0,001	1,23	(1,03; 1,47)	0,025
Sexo	<0,001							<0,001						
Feminino vs. Masculino		0,81	(0,71; 0,92)	0,001	0,75	(0,69; 0,82)	<0,001		0,82	(0,72; 0,94)	0,004	0,77	(0,70; 0,84)	<0,001
Faixa Etária (anos)	<0,001							<0,001						
25 a 39 vs. 18 a 24		0,10	(0,06; 0,16)	<0,001	0,19	(0,12; 0,32)	<0,001		0,11	(0,06; 0,18)	<0,001	0,20	(0,12; 0,33)	<0,001
40 a 59 vs. 18 a 24		0,02	(0,01; 0,03)	<0,001	0,06	(0,03; 0,09)	<0,001		0,02	(0,01; 0,04)	<0,001	0,06	(0,04; 0,10)	<0,001
60 ou mais vs. 18 a 24		0,01	(0,01; 0,02)	<0,001	0,03	(0,02; 0,05)	<0,001		0,01	(0,01; 0,02)	<0,001	0,04	(0,02; 0,06)	<0,001
Raça/Etnia	0,059							0,002						
Branca vs. Parda		1,07	(0,94; 1,23)	0,322	1,09	(0,99; 1,20)	0,082		1,28	(1,11; 1,48)	<0,001	1,17	(1,06; 1,30)	0,003
Preta vs. Parda		0,83	(0,68; 1,02)	0,071	0,90	(0,79; 1,03)	0,128		0,88	(0,72; 1,08)	0,210	0,93	(0,82; 1,07)	0,303
Amarela/Indígena/Ignorada vs. Parda		0,75	(0,46; 1,21)	0,237	0,82	(0,57; 1,17)	0,272		0,86	(0,54; 1,39)	0,545	0,87	(0,60; 1,25)	0,436
Estado Civil	<0,001							<0,001						
Casado vs. Solteiro			0,74	(0,64; 0,85)	<0,001	0,95	(0,86; 1,06)	0,347
Separado/Divorciado/Viúvo vs. Solteiro			0,68	(0,57; 0,82)	<0,001	0,87	(0,77; 0,99)	0,030
Região	<0,001							<0,001						
Norte vs. Nordeste			1,16	(0,99; 1,37)	0,068	1,20	(1,07; 1,35)	0,003
Sudeste vs. Nordeste			0,47	(0,40; 0,55)	<0,001	0,74	(0,66; 0,83)	<0,001
Sul vs. Nordeste			0,44	(0,37; 0,54)	<0,001	0,71	(0,63; 0,80)	<0,001
Centro-Oeste vs. Nordeste			0,63	(0,53; 0,76)	<0,001	0,78	(0,69; 0,88)	<0,001
Situação de Domicílio	<0,001							<0,001						
Rural vs. Urbano			1,23	(1,02; 1,48)	0,033	1,40	(1,25; 1,58)	<0,001
Doença Crônica														
Pelo menos 1	

Notas:

^a SCV – Saúde Cardiovascular;

^b PNS – Pesquisa Nacional de Saúde;

^c Análise do efeito global da variável;

^d RC – Razão de chance;

^e EF – Ensino Fundamental;

^f EM – Ensino Médio;

^g EFI – Ensino Fundamental Incompleto;

^h ES – Ensino Superior.

Capítulo 4

Conclusão Geral

O presente trabalho envolveu o preparo de 2 artigos e teve por objetivos avaliar o escore de saúde cardiovascular, proposto pela *American Heart Association* e suas metas individuais na população brasileira adulta e investigar, em um segundo momento, a associação entre categorias de saúde cardiovascular e determinantes sociais. Os dados utilizados foram da Pesquisa Nacional de Saúde de 2019.

O escore de SCV foi obtido pela soma de 7 metas comportamentais e biológicas atingidas pelo indivíduo, variando assim de 0 a 7. Tal escore foi categorizado de forma binária (SCV2cat), considerando como favorável 6 ou 7 metas atingidas, e em 3 níveis de SCV (SCV3cat): inferior (0 a 2 metas atingidas), intermediário (3 a 5 metas atingidas) e superior (6 ou 7 metas atingidas).

O primeiro artigo (capítulo 2) mostrou que a prevalência de indivíduos com SCV Ideal na população adulta brasileira em 2019 foi de apenas 0,5%. Dentre os componentes do escore, observou-se que a prevalência de indivíduos que atingiram as três metas biológicas foi bem maior que a de indivíduos que atingiram as quatro metas comportamentais. A meta de dieta favorável, por exemplo, foi a principal responsável pela baixa prevalência de SCV Ideal. Uma investigação mais ampliada do perfil sociodemográfico permitiu identificar maior prevalência de SCV Ideal em indivíduos com maior nível de escolaridade e residentes em áreas urbanas.

O segundo artigo apresentou os resultados de associações entre determinantes socioeconômicos

micos (fatores de exposição) e diferentes categorizações de SCV (SCV2cat e SCV3cat) em modelos de regressão logística binária e multinomial, ajustados por variáveis sociodemográficas. O modelo multinomial apresentou resultados mais interessantes que o modelo binário, possibilitando maior discriminação de associações por subgrupos populacionais. Os 3 determinantes socioeconômicos (escolaridade, índice de Riqueza e situação de ocupação) se mostraram associados à SCV em 3 categorias, após ajuste pelas covariáveis sexo, faixa etária, raça, estado civil, regiões, situação de domicílio e doenças crônicas (todas significantes).

Ressalta-se que o presente estudo é de suma importância por utilizar dados de uma pesquisa de âmbito nacional com delineamento e metodologia validados. A amostra selecionada é representativa da população adulta brasileira e apresenta dados atualizados acerca da saúde bem como de determinantes sociais, econômicos e demográficos no Brasil. Representa, portanto, um importante conjunto de resultados que servem para auxiliar no desenvolvimento de intervenções mais direcionadas e eficazes, para o alcance de maiores prevalências de SCV ideal na população brasileira adulta.

Sugere-se, para estudos futuros, maior exploração das informações contidas da PNS e análises mais aprofundadas das associações de cada determinante social de saúde com o escore de SCV. Ferramentas de avaliação de qualidade de ajuste também podem ser exploradas de forma mais aprofundada no contexto de análises de modelos de regressão logística (binária, ordinal e multinomial) em estudos com delineamentos envolvendo amostragem complexa. Ainda, é de suma importância o acompanhamento dessas prevalências ao longo do tempo e a verificação dos resultados quanto a prevenção das doenças cardiovasculares.

Apêndice A

Metodologia

Nesta seção, apresenta-se modelos de regressão logística multinomiais e ordinais. As definições e metodologias apresentadas são baseadas em Pessoa, D. e Nascimento Silva, P. (2018), Agresti, A. (2010) e Agresti, A. (2013).

A.1 Regressão Logística para amostragem complexa

O modelo de regressão logística é o mais importante para modelagem de variável resposta binária. O método de máxima pseudo-verossimilhança (MPV) pode produzir estimativas para os parâmetros do modelo levando em consideração planos amostrais complexos, como os utilizados nas Pesquisas Nacionais de Saúde. O estimador de MPV fornece estimativas mais robustas para a realização de inferência para a população, mesmo podendo ser muito próximo ao estimador de máxima verossimilhança (EMV) tradicional.

O vetor de parâmetros β do modelo é estimado pela MPV solucionando-se a relação a seguir:

$$\sum_{i \in s} w_i \frac{\partial L(\beta)}{\partial \beta_k} = \sum_{i \in s} w_i [y_i - \pi(\mathbf{x}_i)] \mathbf{x}_i = 0,$$

em que Y é a variável resposta binária (por simplicidade), \mathbf{x} é o conjunto de variáveis explicativas, $\pi(\mathbf{x})$ é a probabilidade de sucesso dado o conjunto de variáveis explicativas, β é o

vetor de parâmetros do modelo, $L(\beta)$ é a função de log verossimilhança e w é o peso amostral.

Os parâmetros são estimados por métodos numéricos de um sistema de equações. Para o cálculo do erro padrão dos estimadores podem ser utilizados métodos de linearização de Taylor e de replicação (BRR, Jackknife, bootstrap, multistage bootstrap, ou fornecido pelo usuário). Para os demais modelos lineares generalizados, como os apresentados a seguir, o cálculo de tais estimativas se dá de forma análoga.

Para situações de dados provenientes de amostras complexas, uma das principais dificuldades na aplicação da MPV é a necessidade de conhecimento dessas características do plano amostral, pois o estimador de MPV é obtido levando em consideração os pesos amostrais. Além disso, estudos com amostragem complexa limitam ainda mais a utilização de métodos usuais de diagnóstico de ajuste de modelos (tais como gráficos de resíduos) e outros procedimentos da inferência clássica (tais como testes estatísticos de Razões de Verossimilhança), os quais nem sempre são apropriados quando os modelos avaliados são logísticos. Alternativas como tabelas de classificação, curvas ROC (*receiver operator characteristic*), coeficientes de determinação (R^2) propostos especificamente para regressões logísticas (como o R^2 de McFadden) e validação cruzada são citadas na literatura.

A.2 Regressão Logística Multinomial

Modelos de Regressão Logística Multinomial são extensões de modelos de regressão logística binária. São utilizados para modelagem de variáveis resposta ordinais ou nominais com mais de duas categorias. Tais modelos permitem a estimação de chances do ocorrência de uma categoria da variável resposta em relação à categoria de referência (desfecho de interesse). Razões de chances do desfecho de interesse são então estimadas em função das variáveis explicativas, sejam elas contínuas ou categorizadas. Considerando uma variável resposta Y , com J categorias possíveis ($J > 2$), e suas respectivas probabilidades dadas por π_1, \dots, π_J , o modelo multinomial é equivalente a realizar, simultaneamente, $J-1$ modelos logísticos binários para to-

dos os pares de categorias, sempre com uma mesma categoria de referência (J), que representa o desfecho de interesse.

Considera-se $\mathbf{y}_i = (y_{i1}, \dots, y_{iJ})$ uma observação do processo multinomial pertencente à categoria j, com apenas $y_{ij} = 1$ e os demais $y_{ik} = 0, \forall(k \neq j)$. O vetor de variáveis explicativas para a i-ésima observação é $\mathbf{x}_i = (x_{i1}, \dots, x_{iJ})$. Ainda, $\boldsymbol{\beta}_j = (\beta_{j1}, \dots, \beta_{jp})'$ denota o vetor de parâmetros para o j-ésimo logito de categoria de referência. Assim, com $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$, dado um conjunto fixo de variáveis explicativas \mathbf{x} e com $\sum_j \pi_j(\mathbf{x}) = 1$, o modelo pode ser expresso como:

$$\log \frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})} = \alpha_j + \boldsymbol{\beta}'_j \mathbf{x}, \quad j = 1, \dots, J - 1. \quad (\text{A.1})$$

A função de ligação apresentada também é conhecida como logito de categoria de referência ou de referência (do inglês, *baseline-category logit*). Dessa maneira, o modelo descreve o efeito de \mathbf{x} nesses $J - 1$ logitos, mas ainda é possível utiliza-los para estimar o efeito em todos os $J * (J - 1)$ logitos de pares de categorias, da seguinte maneira:

$$\log \frac{\pi_a(\mathbf{x})}{\pi_b(\mathbf{x})} = \log \frac{\pi_a(\mathbf{x})}{\pi_J(\mathbf{x})} - \log \frac{\pi_b(\mathbf{x})}{\pi_J(\mathbf{x})}. \quad (\text{A.2})$$

Ainda, é possível descrever o modelo com uma expressão diretamente em termos da probabilidade de se observar a j-ésima categoria da variável resposta $\pi_j(\mathbf{x})$ como a seguir:

$$\pi_j(\mathbf{x}) = \frac{\exp(\alpha_j + \boldsymbol{\beta}'_j \mathbf{x})}{1 + \sum_{h=1}^{J-1} \exp(\alpha_h + \boldsymbol{\beta}'_h \mathbf{x})}. \quad (\text{A.3})$$

Como $\pi_J = 1 - (\pi_1 + \dots + \pi_{J-1})$ e $y_{iJ} = 1 - (y_{i1} + \dots + y_{i,J-1})$, então a função de

verossimilhança para a i -ésima observação é dada por:

$$\begin{aligned} \log \left[\prod_{j=1}^J \pi_j(\mathbf{x}_i)^{y_{ij}} \right] &= \sum_{j=1}^{J-1} y_{ij} \log \pi_j(\mathbf{x}_i) + \left(1 - \sum_{j=1}^{J-1} y_{ij} \right) \log \left[1 - \sum_{j=1}^{J-1} \pi_j(\mathbf{x}_i) \right] \\ &= \sum_{j=1}^{J-1} y_{ij} \log \frac{\pi_j(\mathbf{x}_i)}{1 - \sum_{j=1}^{J-1} \pi_j(\mathbf{x}_i)} + \log \left[1 - \sum_{j=1}^{J-1} \pi_j(\mathbf{x}_i) \right]. \end{aligned} \quad (\text{A.4})$$

Dessa maneira, a função de log-verossimilhança para n observações independentes é dada por:

$$\begin{aligned} \log \prod_{i=1}^n \left[\prod_{j=1}^J \pi_j(\mathbf{x}_i)^{y_{ij}} \right] & \\ &= \sum_{i=1}^n \left[\sum_{j=1}^{J-1} y_{ij} (\alpha_j + \beta_j' \mathbf{x}_i) - \log \left(1 + \sum_{j=1}^{J-1} \exp(\alpha_j + \beta_j' \mathbf{x}_i) \right) \right] \\ &= \sum_{j=1}^{J-1} \left[\alpha_j \left(\sum_{i=1}^n y_{ij} \right) + \sum_{k=1}^p \beta_{jk} \left(\sum_{i=1}^n x_{ik} y_{ij} \right) \right] \\ &\quad - \sum_{i=1}^n \log \left[1 + \sum_{j=1}^{J-1} \exp(\alpha_j + \beta_j' \mathbf{x}_i) \right]. \end{aligned} \quad (\text{A.5})$$

A estatística suficiente para β_{jk} é $\sum_i x_{ik} y_{ij}$, $j = 1, \dots, J-1$, $k = 1, \dots, p$ e para α_j é o número total de observações na categoria j , ou seja, $\sum_i y_{ij} = \sum_i x_{i0} y_{ij}$, com $x_{i0} = 1$.

Em geral, o método de Newton-Raphson é capaz de produzir as estimativas dos parâmetros de MV, já que a função de log-verossimilhança é côncava. Para amostras grandes, os estimadores seguem distribuição Normal e os desvios-padrões são a raiz quadrada dos elementos da diagonal da matriz de informação inversa.

Para verificar a adequação do modelo, pode-se utilizar os testes χ^2 ou G^2 de qualidade de ajuste, ou outras alternativas citadas anteriormente, caso os preditores sejam categóricos e os dados não sejam esparsos. Nos demais casos, essas estatísticas são válidas apenas para comparar modelos aninhados com pouca diferença.

A.3 Regressão Logística Ordinal

Os modelos de regressão logística ordinal também são extensões dos modelos de regressão logística específicos para variáveis resposta ordinais com 3 ou mais categorias. Tais modelos variam conforme a maneira que a função de ligação é estabelecida, como são medidos os efeitos por razão de chances e com qual abordagem utiliza-se os efeitos das variáveis explicativas.

Os modelos logísticos ordinais podem utilizar diferentes formas de definir a razão de chances na função de ligação logito: cumulativa; de categorias adjacentes e de razão contínua. Tais modelos podem associar a função de ligação com as variáveis explicativas com ou sem chances proporcionais (do inglês, *proportional odds*), com chances proporcionais parciais ou da forma do modelo de esteriótipo.

Existe ainda uma grande variedade de abordagens, como, por exemplo, a utilização de outros tipos de funções de ligação (*probit*, *log-log*), a aplicação de análises multivariada para as respostas ordinais, a realização de estudos longitudinais ou aplicação de métodos bayesianos.

A.3.1 Regressão Logística Ordinal com função logito cumulativo

Considera-se uma variável resposta Y com J categorias possíveis com probabilidades π_1, \dots, π_J . A probabilidade acumulada de Y para a categoria específica $j = 1, \dots, J - 1$ é dada por $P(Y \leq j)$. Ou seja, a chance de ocorrer uma categoria menor ou igual à j definida é $P(Y \leq j)/P(Y > j)$, para $P(Y > j) > 0$. Portanto, o chamado logito cumulativo (do inglês, *cumulative logit*) é dado por:

$$\begin{aligned} \text{logit}[P(Y \leq j)] &= \log \left(\frac{P(Y \leq j)}{1 - P(Y \leq j)} \right) \\ &= \log \left(\frac{\pi_1 + \dots + \pi_j}{\pi_{j+1} + \dots + \pi_J} \right), \quad j = 1, \dots, J - 1. \end{aligned} \quad (\text{A.6})$$

Tal procedimento é análogo a agrupar as categorias em diferentes respostas binárias. Cada logito utiliza as respostas de todas as categorias.

A.3.1.1 Logito Cumulativo com Chances Proporcionais

Um modelo com logito cumulativo e chances proporcionais pode ser ajustado considerando que variáveis explicativas possuem o mesmo efeito para as diferentes categorias da variável resposta. Para cada observação i , considera-se y_i a categoria resposta e \mathbf{x}_i o vetor de valores das variáveis explicativas. O modelo utiliza, simultaneamente, todos os $J - 1$ logits cumulativos, da seguinte maneira:

$$\text{logit}[P(Y_i \leq j | \mathbf{x}_i)] = \alpha_j + \boldsymbol{\beta}' \mathbf{x}_i = \alpha_j + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots \quad (\text{A.7})$$

para $j = 1, \dots, J - 1$ e para $\boldsymbol{\beta}$ sendo um vetor de parâmetros que descreve o efeitos das respectivas variáveis explicativas.

Uma expressão equivalente para o modelo, em função da probabilidade acumulada, é dada por:

$$P(Y_i \leq j | \mathbf{x}_i) = \frac{\exp(\alpha_j + \boldsymbol{\beta}' \mathbf{x}_i)}{1 + \exp(\alpha_j + \boldsymbol{\beta}' \mathbf{x}_i)}, \quad j = 1, \dots, J - 1. \quad (\text{A.8})$$

Já para expressar o modelo em função da probabilidade de cada uma das categorias, tem-se:

$$P(Y_i = j | \mathbf{x}_i) = \frac{\exp(\alpha_j + \boldsymbol{\beta}' \mathbf{x}_i)}{1 + \exp(\alpha_j + \boldsymbol{\beta}' \mathbf{x}_i)} - \frac{\exp(\alpha_{j-1} + \boldsymbol{\beta}' \mathbf{x}_i)}{1 + \exp(\alpha_{j-1} + \boldsymbol{\beta}' \mathbf{x}_i)}, \quad (\text{A.9})$$

com $\alpha_0 = -\infty$, $\alpha_J = \infty$ e $j = 1, \dots, J - 1$.

O modelo admite chances proporcionais quando a magnitude do efeito de uma variável explicativa é a mesma entre todas as categorias da variável resposta, ou seja, quando os coeficientes $\boldsymbol{\beta}$ são iguais para cada logito cumulativo conforme expressão à seguir:

$$\begin{aligned} & \text{logit}[P(Y \leq j | \mathbf{x}_1)] - \text{logit}[P(Y \leq j | \mathbf{x}_2)] \\ &= \log \left(\frac{P(Y \leq j | \mathbf{x}_1) / P(Y > j | \mathbf{x}_1)}{P(Y \leq j | \mathbf{x}_2) / P(Y > j | \mathbf{x}_2)} \right) = \boldsymbol{\beta}' (\mathbf{x}_1 - \mathbf{x}_2), \end{aligned} \quad (\text{A.10})$$

ou seja, a razão de chances de se obter uma resposta $Y \leq j$ em $\mathbf{x} = \mathbf{x}_1$ é $\exp[\boldsymbol{\beta}'(\mathbf{x}_1 - \mathbf{x}_2)]$ vezes a chance em $\mathbf{x} = \mathbf{x}_2$.

Considera-se y_{i1}, \dots, y_{iJ} variáveis indicadoras para cada categoria da variável resposta, ou seja, se $Y_i = j$ então $y_{ij} = 1$ e $y_{ik} = 0, \forall k \neq j$. Para representar a probabilidade $P(Y_i = j | \mathbf{X} = \mathbf{x}_i)$ será utilizada a notação $\pi_j(\mathbf{x}_i)$. Assumindo que as observações são independentes e que $P(Y_i \leq 0 | \mathbf{x}_i) = 0$, a função de verossimilhança, que envolve o produto de distribuições multinomiais das n observações, apresenta-se da seguinte maneira:

$$\begin{aligned} l(\{\alpha_j\}, \boldsymbol{\beta}) &= \prod_{i=1}^n \left[\prod_{j=1}^J \pi_j(\mathbf{x}_i)^{y_{ij}} \right] = \prod_{i=1}^n \left\{ \prod_{j=1}^J [P(Y_i \leq j | \mathbf{x}_i) - P(Y_i \leq j - 1 | \mathbf{x}_i)]^{y_{ij}} \right\} \\ &= \prod_{i=1}^n \left\{ \prod_{j=1}^J \left[\frac{\exp(\alpha_j + \boldsymbol{\beta}'\mathbf{x}_i)}{1 + \exp(\alpha_j + \boldsymbol{\beta}'\mathbf{x}_i)} - \frac{\exp(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}_i)}{1 + \exp(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}_i)} \right]^{y_{ij}} \right\}. \end{aligned} \quad (\text{A.11})$$

Os parâmetros do modelo são estimados pelo método de máxima verossimilhança (MV) no qual a função de log-verossimilhança $L(\{\alpha_j\}, \boldsymbol{\beta}) = \log(l(\{\alpha_j\}, \boldsymbol{\beta}))$ é derivada e igualada a zero. Desse modo, a primeira derivada de um parâmetro β_k é dada por:

$$\frac{\partial L(\{\alpha_j\}, \boldsymbol{\beta})}{\partial \beta_k} = \sum_{i=1}^n \sum_{j=1}^J y_{ij} x_{ik} \frac{g(\alpha_j + \boldsymbol{\beta}'\mathbf{x}_i) - g(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}_i)}{G(\alpha_j + \boldsymbol{\beta}'\mathbf{x}_i) - G(\alpha_{j-1} + \boldsymbol{\beta}'\mathbf{x}_i)} = 0, \quad (\text{A.12})$$

em que $G(z) = \exp(z)/1 + \exp(z)$ e $g(z) = \exp(z)/[1 + \exp(z)]^2$. Assim como em modelos de regressão logística binária, essas equações não são lineares nos parâmetros e não possuem uma solução fechada. Então, métodos iterativos são utilizados na solução de tais equações para a estimação dos parâmetros do modelo. Um tamanho de amostra grande o suficiente garante um único valor máximo para a função.

Com isso, a matriz de covariância dos parâmetros estimados do modelo de MV, para uma amostra grande é o inverso da matriz de informação, calculada em relação às estimativas de MV. A matriz de informação é obtida por $-\partial^2 L(\{\alpha_j\}, \boldsymbol{\beta})/\partial \beta_k^2$, descrevendo a curvatura da função. Essa matriz de informação possui duas versões possíveis, sendo uma a observada que realmente

utiliza as segundas derivadas parciais e a outra, com seus valores esperados. O valor estimado para os desvios-padrões é a raiz quadrada da diagonal principal dessa matriz de covariâncias, independente da versão escolhida.

Ainda, tem-se o intervalo de 95% de confiança de Wald para $\hat{\beta}_k$ dado por $\hat{\beta}_k \pm 1,96 * (DP_{\hat{\beta}_k})$. Já para testar a hipótese nula $H_0 : \beta_k = 0$, utiliza-se a estatística de teste $z = \hat{\beta}_k / (DP_{\hat{\beta}_k})$ que, sob H_0 , possui distribuição normal padrão, ou z^2 que, sob H_0 , possui distribuição assintótica Qui-quadrado (χ^2) com 1 grau de liberdade (*g.l.*), porém quando a amostra for pequena ou a maior parte das observações estiverem concentradas na maior (ou na menor) categoria da variável resposta é interessante utilizar o teste de razão de verossimilhança, no qual a estatística é dada por $-2(L_0 - L_1)$, sendo L_0 a função $L(\{\alpha_j\}, \beta)$ sob a hipótese nula e L_1 , sob a hipótese alternativa.

A.3.1.2 Teste de hipóteses para a suposição de chances proporcionais

O teste para a suposição de chances proporcionais verifica se os efeitos das variáveis explicativas são os mesmos para todos os logitos cumulativos contra a alternativa de efeitos diferentes, ou seja, compara os modelos com e sem chances proporcionais. A estatística para o teste escore sob H_0 possui distribuição Qui-quadrado com $gl = p(J - 2)$, em que p é o número de variáveis preditoras e J de categorias da variável resposta.

O modelo mais complexo (sem chances proporcionais) tem o problema estrutural de que as probabilidades cumulativas podem estar fora de ordem em algumas configurações dos preditores. Por esse motivo, muitas vezes não é viável maximizar a função de verossimilhança. Assim, o teste de escore é mais amplamente aplicável que o teste de razão de verossimilhança ou o teste de Wald, mesmo o teste de escore podendo ser muito liberal (p-valores tendem a ser muito pequenos e as taxas reais do erro do tipo I tendem a ser maiores que o valor nominal) e não sendo tão robusto em algumas situações (com dados esparsos, com muitas observações nas categorias extremas ou quando envolve variáveis explicativas contínuas).

Mesmo que um modelo com efeitos diferentes para os logitos cumulativos pareça mais com-

pleto, é importante levar em consideração a parcimônia do modelo mais simples com estrutura de chances proporcionais. Em alguns casos, um p-valor pequeno, para uma amostra grande, pode refletir meramente significância estatística ao invés de significância prática.

A.3.1.3 Logito Cumulativo sem Chances Proporcionais

A suposição de chances proporcionais, descrita na seção anterior, geralmente é testada contra um modelo mais complexo, sem chances proporcionais, que inclui um parâmetro para cada logito. Tal modelo abrange $J - 1$ vetores de parâmetros β_j , permitindo incorporar efeitos diferentes das variáveis independentes para os diferentes logits cumulativos, é dado por

$$\text{logit}[P(Y_i \leq j | \mathbf{x}_i)] = \alpha_j + \beta_j' \mathbf{x}_i, \quad j = 1, \dots, J - 1. \quad (\text{A.13})$$

Contudo, o fato das curvas dessas funções não serem paralelas requer atenção, pois para um determinado valor de \mathbf{x} , as curvas podem se cruzar, violando o princípio da ordem entre as probabilidades acumuladas. De certo modo, a quantidade de preditores passa a ser limitada.

Existem ainda modelos intermediários, como o de chances proporcionais parciais, em que algumas variáveis preditoras possuem chances proporcionais e outras não.

A.3.1.4 Qualidade de ajuste

Testes globais de qualidade de ajuste e métodos mais direcionados, como comparações de modelos e análises de resíduos podem ser utilizados para a avaliação de adequação do ajuste do modelo.

O teste de qualidade de ajuste, compara os valores observados na tabela de contingência dos dados com as frequências esperadas baseadas no modelo ajustado. Dessa maneira, tem-se como hipótese nula que o modelo está adequado e como alternativa que não está, baseada no modelo saturado. Quando os dados são esparsos ou o modelo contém pelo menos uma variável explicativa contínua esses testes globais de qualidade de ajuste não são válidos e deve-se utilizar outras

abordagens, como o teste de Hosmer-Lemeshow para regressão logística binária generalizado.

Pode-se ainda tentar melhorar o ajuste do modelo com a adição de termos simples ou modificados (como transformações quadráticas ou logarítmicas) bem como interações. A comparação do modelo utilizado com outros modelos mais complexos pode ser realizada por meio de um teste de razão de verossimilhança ou pelo cálculo do AIC (critério de informação Akaike).

A análise de resíduos também pode auxiliar na avaliação da qualidade de ajuste. Para uma tabela de contingência, com contagem de células n_{ij} e valor ajustado $\hat{\mu}_{ij}$ para determinado conjunto i de variáveis explicativas, com categoria de resposta j , o resíduo padronizado é dado por:

$$r_{ij} = \frac{n_{ij} - \hat{\mu}_{ij}}{DP},$$

o qual segue distribuição Normal e tem DP representando o desvio-padrão estimado de $n_{ij} - \hat{\mu}_{ij}$, assumindo que o modelo é adequado. Pode-se também utilizar resíduos para totais acumulados, em vez de para contagem das células, substituindo o numerador por $(\sum_{k=1}^j n_{ik} - \sum_{k=1}^j \hat{\mu}_{ik})$.

A.3.2 Regressão Logística com Logito de Categorias Adjacentes

A Regressão Logística com logito de Categorias Adjacentes possibilita a estimação das razões de chances entre quaisquer 2 categorias da variável resposta e, assim, interpretações sobre as categorias separadamente.

Para a variável resposta Y com J categorias possíveis e com probabilidades π_1, \dots, π_J , os logitos de categorias adjacentes são dados por

$$\text{logit}[P(Y = j|Y = j \text{ ou } Y = j + 1)] = \log\left(\frac{\pi_j}{\pi_{j+1}}\right), \quad j = 1, \dots, J - 1. \quad (\text{A.14})$$

Utilizando um conjunto de variáveis explicativas \mathbf{x} , o modelo de regressão logística com

logito de categorias adjacentes geral é dado por

$$\log \left(\frac{\pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x})} \right) = \alpha_j + \boldsymbol{\beta}'_j \mathbf{x}, \quad j = 1, \dots, J-1, \quad (\text{A.15})$$

podendo assumir que as chances são proporcionais ou não. No modelo que envolve chances proporcionais o vetor de parâmetros é único, $\boldsymbol{\beta}_j = \boldsymbol{\beta}, \forall j = 1, \dots, J-1$. Já, quando as chances não são consideradas proporcionais o modelo é mais geral e envolve $J-1$ vetores de parâmetros $\boldsymbol{\beta}_j$.

A versão não proporcional, nesse caso, fornece probabilidades válidas independentemente dos valores das variáveis explicativas, diferentemente do modelo geral da seção anterior, Equação (A.7), pois os efeitos são descritos com razões de chances locais.

O modelo geral, Equação (A.15), referente às probabilidades de cada categoria, é

$$\pi_j(\mathbf{x}) = \frac{\exp(\alpha_j + \boldsymbol{\beta}'_j \mathbf{x})}{1 + \sum_{k=1}^{J-1} \exp(\alpha_k + \boldsymbol{\beta}'_k \mathbf{x})}, \quad j = 1, \dots, J-1. \quad (\text{A.16})$$

O modelo de chances proporcionais tem como vantagem uma maior parcimônia e mais fácil interpretação dos parâmetros. O aumento de uma unidade na variável preditora k implica em um aumento de $\exp(\beta_k)$ na chance de estar em uma categoria j , comparada à estar em uma categoria superior $j+1$ da variável resposta.

A função de log-verossimilhança é dada por

$$\begin{aligned} \log \left[\prod_{j=1}^J \pi_j(\mathbf{x}_i)^{y_{ij}} \right] &= \sum_{j=1}^{J-1} y_{ij} \log(\pi_j(\mathbf{x}_i)) + \left(1 - \sum_{j=1}^{J-1} y_{ij} \right) \log(\pi_J(\mathbf{x}_i)) \\ &= \sum_{j=1}^{J-1} y_{ij} \log \left(\frac{\pi_j(\mathbf{x}_i)}{\pi_J(\mathbf{x}_i)} \right) + \log(\pi_J(\mathbf{x}_i)), \end{aligned} \quad (\text{A.17})$$

em que y_{ij} é a variável indicadora se a resposta i está ou não na categoria j , x_{ik} é o valor da k -ésima variável explicativa, assumindo que as n observações são independentes e seguem uma distribuição multinomial.

A.3.2.1 Conexão com o Logito de Categoria de Referência

A Regressão Multinomial é construída utilizando o logito de categoria de referência, como foi discutido na Seção 4.1. Os logitos de categoria de referência são equivalentes a um conjunto de logitos de categorias adjacentes. Considerando a categoria de referência J , tem-se:

$$\log \frac{\pi_j}{\pi_J} = \log \frac{\pi_j}{\pi_{j+1}} + \log \frac{\pi_{j+1}}{\pi_{j+2}} + \cdots + \log \frac{\pi_{j-1}}{\pi_J}, \quad j = 1, \dots, J - 1.$$

Essa conexão é de suma importância, já que é muito mais comum o interesse no ajuste por categoria de referência. Então, como

$$\log \frac{\pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x})} = \log \frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})} - \log \frac{\pi_{j+1}(\mathbf{x})}{\pi_J(\mathbf{x})}, \quad j = 1, \dots, J - 1,$$

é possível se obter o EMV $\hat{\beta}_j$ para β_j , do modelo geral de categorias adjacentes (A.15), obtendo estimativas $\{\hat{\beta}_j^*\}$ do modelo de categoria de referência e utilizando $\hat{\beta}_j = \hat{\beta}_j^* - \hat{\beta}_{j+1}^*$, com $\hat{\beta}_J^* = 0$.

A.3.3 Regressão Logística com logito de Razão Contínua

Os logitos de razão contínua podem ser definidos de duas formas: considerando a chance de cada categoria em relação às mais altas ou, em relação às mais baixas. Estes logitos podem ser representados, respectivamente, por:

$$\log \left(\frac{\pi_j}{\pi_{j+1} + \cdots + \pi_J} \right) \quad e \quad \log \left(\frac{\pi_{j+1}}{\pi_1 + \cdots + \pi_j} \right), \quad j = 1, \dots, J - 1.$$

A ênfase nesta seção é no primeiro tipo e este é bastante utilizado em casos de processos sequenciais determinando o resultado da resposta, no sentido que uma observação deve potencialmente ocorrer na categoria j antes que possa ocorrer em uma categoria superior. Nesse caso,

considera-se

$$\omega_j = P(Y = j | Y \geq j) = \left(\frac{\pi_j}{\pi_j + \dots + \pi_J} \right), \quad j = 1, \dots, J - 1,$$

e então o modelo é representado por

$$\text{logit}[\omega_j(\mathbf{x})] = \alpha_j + \beta_j' \mathbf{x}, \quad j = 1, \dots, J - 1. \quad (\text{A.18})$$

No caso de chances proporcionais o modelo é simplificado, de maneira que $\beta_j = \beta, \forall j = 1, \dots, J - 1$.

Ainda nesse caso de chances proporcionais, é possível realizar uma fatoração binomial da função de distribuição multinomial, o que contribui para a construção da função de log-verossimilhança. Portanto, considerando S_j o conjunto de todas as categorias iguais ou acima de j e $n_j^* = n_j + \dots + n_J$, tem-se

$$L_j = \left(\sum_{S_j} y_{ij} \right) \alpha_j + \sum_k \left(\sum_{S_j} y_{ij} x_{ik} \right) \beta_k - \sum_{S_j} \log \left[1 + \exp \left(\alpha_j + \sum_k \beta_k x_{ik} \right) \right] \quad (\text{A.19})$$

e, assim, a função de log-verossimilhança é dada por

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = L_1 + L_2 + \dots + L_{L-1}.$$

A.3.4 Modelo de Esteriótipo

O modelo de estereótipo também é conhecido por usar logitos multiplicativos de categorias pareadas. Pode ser utilizado com o logito de categorias adjacentes ou o de categoria de referência. Tal modelo intermediário é mais flexível que o que admite chances proporcionais e mais parcimonioso em relação ao modelo geral, sem chances proporcionais.

Primeiramente, utilizando o logito de categoria de referência $\pi_j(\mathbf{x})/\pi_J(\mathbf{x})$ tem-se o modelo,

conhecido como modelo de estereótipo (do inglês, *stereotype model*), da seguinte maneira:

$$\log \left(\frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})} \right) = \alpha_j + \phi_j \boldsymbol{\beta}' \mathbf{x}, \quad j = 1, \dots, J - 1. \quad (\text{A.20})$$

Para o j -ésimo logito, o coeficiente da variável explicativa x_k é $\phi_j \beta_k$. Para a identificação desses novos parâmetros de escores $\{\phi_j\}$ é necessário se impor uma restrição. Determinados os valores para $\{\phi_j\}$, o modelo de esteriótipo passa a ter a mesma quantidade de parâmetros que um modelo com chances proporcionais, ou seja, com um único parâmetro para o efeito de cada variável explicativa.

Para o caso com logito de categorias adjacentes, por exemplo, o modelo é expresso por

$$\log \left(\frac{\pi_j(\mathbf{x})}{\pi_{j+1}(\mathbf{x})} \right) = \alpha_j + \nu_j \boldsymbol{\beta}' \mathbf{x}, \quad j = 1, \dots, J - 1, \quad (\text{A.21})$$

sendo que $\nu_j = \phi_j - \phi_{j+1}$ e, assim, $\phi_j = \nu_j + \nu_{j+1} + \dots + \nu_{J-1}$ com $j = 1, \dots, J - 1$.

Em termos das probabilidades da variável resposta, o modelo pode ser escrito como:

$$\pi_j(\mathbf{x}) = \frac{\exp(\alpha_j + \phi_j \boldsymbol{\beta}' \mathbf{x})}{\sum_{k=1}^J \exp(\alpha_k + \phi_k \boldsymbol{\beta}' \mathbf{x})}, \quad j = 1, \dots, J.$$

Portanto, modelos utilizando escores fixos e igualmente espaçados, como na situação em que $\{\phi_j = J - j\}$ e $\nu_j = 1$, são casos especiais do modelo de esteriótipo que equivale a utilizar um modelo de chances proporcionais com logito de categorias adjacentes. Para avaliar se os parâmetros de escore $\{\phi_j\}$ divergem significativamente de serem “igualmente espaçados” é interessante realizar um teste de razão de verossimilhança, comparando esse caso especial do modelo de estereótipo com a versão de chances proporcionais do modelo com logito de categorias adjacentes.

O modelo de esteriótipo pode apresentar complicações por possuir uma relação multiplicativa, em vez de linear, com os parâmetros. Assim, não é possível realizar o ajuste reaproveitando os padrões dos modelos lineares generalizados encontrados em programas para análises estatís-

ticas.

A realização da inferência dos parâmetros do modelo também envolve complicações adicionais. Pode-se adotar, como alternativa, uma versão iterativa de duas etapas: na primeira utiliza-se valores fixos para os parâmetros ϕ_j e estima-se os demais parâmetros (α_j e β) e, na segunda etapa, ao contrário, utiliza-se os parâmetros α_j e β estimados como fixos para então se estimar os parâmetros ϕ_j . O problema dessa abordagem é que os desvios-padrões das estimativas calculados no *software* não serão válidos, pois consideram ϕ_j fixo e não levam em consideração que este também é estimado a cada ciclo.

A.3.5 Referências

1. Agresti, A. (2010). Analysis of ordinal categorical data (Vol. 656). John Wiley Sons.
2. Agresti, A. (2013). Categorical data analysis. 3rd Edition, John Wiley Sons Inc., Hoboken.
3. Pessoa, D. e Nascimento Silva, P. (2018). Análise de Dados Amostrais Complexos.
<https://djalmapessoa.github.io/adac/>
4. Shtatland E, Kleinman K, Cain E. (2002). One more time about R2 measures of fit in logistic regression. NESUG 15 Proceedings.

Apêndice B

Resultados de modelos de regressão logística ordinal

Apresenta-se na Tabela 1 as estimativas das razões de chances (RC), seus respectivos intervalos de 95% de confiança (IC95%) e p-valores, considerando como variável resposta o escore de SCV em três categorias (inferior: 0 a 2 metas atingidas, intermediário: 3 a 5 metas atingidas e superior: 6 ou 7 metas atingidas) por diferentes variáveis explicativas em um modelo de regressão logística ordinal. Os resultados mostraram associações significantes entre SCV e as variáveis de exposição de interesse (Escolaridade, Status de ocupação e Índice de Riqueza), após ajuste por covariáveis. Observou-se que a chance de se estar em categorias superiores de SCV é maior entre indivíduos com maior escolaridade, ocupados e com maiores índices de riqueza, após ajustes por covariáveis. Observou-se ainda que todas as covariáveis foram significantes neste modelo. Maiores chances de estar em categorias superiores de SCV foram observadas em indivíduos do sexo feminino, mais jovens, de raça branca (comparada à parda), solteiros, sem doenças crônicas, residentes nas regiões norte ou nordeste, domiciliados em áreas urbanas.

A Tabela 2 apresenta as estimativas da regressão logística ordinal, considerando o escore de SCV em oito categorias (0 a 7 metas atingidas). Os resultados foram similares aos do modelo

logístico ordinal com a variável de SCV em três categorias (Tabela B.1) e aos obtidos com a regressão logística multinomial (Tabela 3.3 - Capítulo 3). Porém, com mais variáveis explicativas significantes que os outros modelos do segundo artigo (Capítulo 3).

O R^2 de McFadden para os modelos logísticos ordinais em 3 e 8 categorias foram, respectivamente, 0,09 e 0,06. Não foi possível construir o modelo de regressão logística ordinal com outras funções de ligação e nem testar a suposição de chances proporcionais, pois tais funções não estão implementadas nos *softwares* utilizados para o caso de planos amostrais complexos. De qualquer modo, a suposição de chances proporcionais não pareceu razoável para os modelos com as variáveis resposta de escore de SCV em 3 e 8 categorias. Por isso, o modelo multinomial foi utilizado no segundo artigo.

Tabela B.1: Razões de chances, IC95% e p-valores de modelos de regressão logística ordinal com logito cumulativo e chances proporcionais para a variável resposta SCV^a em 3 categorias (inferior 0-2, intermediária 3-5 e superior ou ideal 6-7) – PNS^b 2019

Variáveis	RC	IC95%	p-valor
Escolaridade (Nível de ensino)			<0,001
Fundamental completo a médio incompleto vs. Até FI ^c	1,21	(1,08 - 1,36)	<0,001
Médio completo a superior incompleto vs. Até FI	1,70	(1,55 - 1,87)	<0,001
Superior completo vs. Até FI	3,00	(2,65 - 3,41)	<0,001
Situação de Ocupação			<0,001
Empregado	1,20	(1,11 - 1,29)	<0,001
Índice de Riqueza			<0,001
2º vs. 1º quintil	0,98	(0,89 - 1,08)	0,705
3º vs. 1º quintil	1,08	(0,96 - 1,20)	0,208
4º vs. 1º quintil	1,22	(1,08 - 1,37)	0,001
5º vs. 1º quintil	1,54	(1,35 - 1,76)	<0,001
Sexo			0,109
Feminino vs. Masculino	0,95	(0,88 - 1,01)	0,109
Faixa Etária (anos)			<0,001
25 a 39 vs. 18 a 24	0,53	(0,47 - 0,60)	<0,001
40 a 59 vs. 18 a 24	0,29	(0,26 - 0,33)	<0,001
60 ou mais vs. 18 a 24	0,17	(0,15 - 0,20)	<0,001
Raça/Etnia			<0,001
Branca vs. Parda	1,14	(1,06 - 1,23)	<0,001
Preta vs. Parda	0,93	(0,84 - 1,03)	0,157
Amarela/Indígena/Ignorada vs. Parda	0,90	(0,68 - 1,19)	0,441
Estado Civil			<0,001
Casado vs. Solteiro	0,86	(0,80 - 0,93)	<0,001
Separado/Divorciado/Viúvo vs. Solteiro	0,79	(0,72 - 0,88)	<0,001
Doença Crônica			<0,001
Pelo menos 1	0,62	(0,57 - 0,66)	<0,001
Região			<0,001
Norte vs. Nordeste	1,04	(0,96 - 1,13)	0,315
Sudeste vs. Nordeste	0,69	(0,63 - 0,75)	<0,001
Sul vs. Nordeste	0,68	(0,61 - 0,75)	<0,001
Centro-Oeste vs. Nordeste	0,79	(0,72 - 0,88)	<0,001
Situação de Domicílio			<0,001
Rural vs. Urbano	1,17	(1,07 - 1,27)	<0,001

Fonte: Pesquisa Nacional de Saúde de 2019

Notas: ^a SCV - Saúde Cardiovascular; ^b PNS - Pesquisa Nacional de Saúde; ^c FI - Fundamental Incompleto.

Tabela B.2: Razões de chances, IC95% e p-valores de modelos de regressão logística ordinal com logito cumulativo e chances proporcionais para a variável resposta SCV^a ideal em 8 categorias (0-7 metas atingidas) – PNS^b 2019

Variáveis	RC	IC95%	p-valor
Escolaridade (Nível de ensino)			<0,001
Fundamental completo a médio incompleto vs. Até FI ^c	1,19	(1,10 - 1,29)	<0,001
Médio completo a superior incompleto vs. Até FI	1,58	(1,48 - 1,69)	<0,001
Superior completo vs. Até FI	2,47	(2,27 - 2,68)	<0,001
Situação de Ocupação			<0,001
Empregado	1,11	(1,05 - 1,17)	<0,001
Índice de Riqueza			<0,001
2° vs. 1° quintil	0,97	(0,90 - 1,04)	0,332
3° vs. 1° quintil	1,03	(0,95 - 1,12)	0,465
4° vs. 1° quintil	1,13	(1,04 - 1,24)	0,006
5° vs. 1° quintil	1,30	(1,19 - 1,43)	<0,001
Sexo			0,681
Feminino vs. Masculino	0,99	(0,94 - 1,04)	0,681
Faixa Etária (anos)			<0,001
25 a 39 vs. 18 a 24	0,50	(0,46 - 0,55)	<0,001
40 a 59 vs. 18 a 24	0,28	(0,26 - 0,31)	<0,001
60 ou mais vs. 18 a 24	0,17	(0,15 - 0,18)	<0,001
Raça/Etnia			<0,001
Branca vs. Parda	1,09	(1,04 - 1,15)	<0,001
Preta vs. Parda	0,93	(0,87 - 1,00)	0,064
Amarela/Indígena/Ignorada vs. Parda	1,01	(0,83 - 1,23)	0,905
Estado Civil			<0,001
Casado vs. Solteiro	0,90	(0,85 - 0,94)	<0,001
Separado/Divorciado/Viúvo vs. Solteiro	0,85	(0,79 - 0,91)	<0,001
Doença Crônica			<0,001
Pelo menos 1	0,65	(0,62 - 0,68)	<0,001
Região			<0,001
Norte vs. Nordeste	1,06	(1,00 - 1,12)	0,064
Sudeste vs. Nordeste	0,73	(0,68 - 0,77)	<0,001
Sul vs. Nordeste	0,69	(0,64 - 0,74)	<0,001
Centro-Oeste vs. Nordeste	0,85	(0,80 - 0,91)	<0,001
Situação de Domicílio			<0,001
Rural vs. Urbano	1,17	(1,10 - 1,24)	<0,001

Fonte: Pesquisa Nacional de Saúde de 2019

Notas: ^a SCV - Saúde Cardiovascular; ^b PNS - Pesquisa Nacional de Saúde; ^c FI - Fundamental Incompleto.

Apêndice C

Construção do Índice de Riqueza

A variável índice de riqueza (do inglês, *wealth index*) é uma variável latente obtida por análise fatorial com o propósito de representar os níveis de riqueza acumulada baseados em bens de consumo, conforme estabelecido pelo programa Demographic and Health Surveys (DHS). Essa variável é uma combinação linear de informações de bens e de outras características da habitação, que possam refletir a situação econômica do indivíduo. A construção dessa variável foi realizada com base nas instruções detalhadas por Rutstein, S. O. (2015) e pelo *World Food Programme* (VAM Guide - 2017).

Inicialmente, cada variável da PNS 2019 utilizada para a construção desse índice foi transformada em uma variável indicadora da(s) categoria(s) considerada(s) favorável(is). A Tabela C.1 apresenta as variáveis originais da PNS com seus respectivos códigos, a indicação das categorias favoráveis e suas frequências. Buscando discriminar melhor os domicílios entre “ricos” e “pobres”, foram excluídas da composição do índice as variáveis com mais de 95% ou menos de 5% da amostra em situação favorável.

Em seguida, calculou-se correlações tetracóricas (utilizadas para variáveis binárias) entre todos os pares de variáveis selecionadas para compor o Índice de Riqueza. Recomenda-se retirar uma das variáveis em pares de variáveis com coeficientes $> 0,95$ ou $< 0,5$. Nenhuma das variáveis selecionadas (Tabela C.1) foi retirada nesta etapa.

Tabela C.1: Variáveis utilizadas na construção do índice de riqueza

Ativo	Código da variável	Resultado Favorável	Percentual Favorável
Abastecimento de água	A005010	1	78,1%
Material do piso	A004010	1, 2	81,6%
Material das paredes externas	A002010	1, 4	91,4%
Material do telhado	A003010*	2, 4	30,4%
Empregado(a) Doméstico(a)	A02101	1	5,6%
Automóvel	A018027	1	45,1%
Telefone Celular	A018019	1	92,5%
Computador	A018023	1	38,0%
Micro-ondas	A018021	1	48,6%
Máquina de lavar roupa	A018015	1	63,1%
Água para beber	A009010	1, 5	66,2%
Esgoto	A01501	1,2	53,5%
Lixo	A016010	1	80,6%
Telefone Fixo	A018017	1	21,8%
Motocicleta	A018025	1	26,0%
Internet	A01901	1	78,0%
Pessoas/Dormitório	C001/A011	<3	91,8%
Pessoas/Banheiro	C001/A01401**	≤ 4	89,1%

* Se A001 = 2, o resultado favorável para A003010 pode ser 2, 3 ou 4;

** Se A01401 ≠ 0.

Para a obtenção da variável latente de riqueza acumulada utiliza-se um processo iterativo, no qual realiza-se análises fatoriais por componentes principais, considerando apenas 1 fator, retirando-se a variável que apresentar a menor comunalidade \hat{h}^2 e repetindo o processo até que todas as comunalidades sejam maiores que o limiar de 0,3. Assim, obtém-se o modelo final, baseado em apenas 1 componente da análise fatorial, já que esta explica a maior proporção da variância total. Nessa etapa as variáveis motocicleta, pessoas/dormitório, empregado(a) doméstico(a) e material das paredes foram excluídas.

Finalmente, o índice de riqueza é obtido atribuindo-se grupos de 1 a 5 separados pelos quin-

tis da variável obtida no passo anterior. A primeira classe se refere aos domicílios com menor riqueza e a quinta, aos domicílios com maior riqueza. Assim, busca-se diferenciar famílias mais pobres de famílias mais ricas.

Apêndice D

Códigos de *software* R e SAS utilizados

D.1 Códigos em linguagem R

Códigos utilizados para a leitura dos dados, criação de variáveis (incluindo o escore de SCV e o índice de riqueza) e estimação de prevalências (Artigo 1).

```
##### PASSO 1 #####
### Importar Pesquisa Nacional de Saude
# library(SAScii) # - pacote para ler arquivos SAS
# pns <- read.SAScii(fn = "PNS_2019.txt",
                    sas_ri = "input_PNS_2019.sas",
                    beginline = 1) # 293726 1087

##### PASSO 2 #####
##### Criterios de Inclusao #####
table(pns$V0025A)
pns_ci <- pns[pns$V0025A==1,] #Entrevista Individual

### 1- M001 = `1` questionario realizado
```

```

##"Entrevista do adulto selecionado" = '1 - REALIZADA'
pns_ci <- pns_ci[pns_ci$M001 == "1",]

### 2- M00203= '1',
## "Morador apto é aquele capaz de responder (falar) por si
# próprio); 1=sim
pns_ci <- pns_ci[pns_ci$M00203 == "1",]

### 3- C008>=18 (idade)
pns_ci <- pns_ci[pns_ci$C008 >= 18,]

### 4- P005 = 2 ou missing - grávida
pns_ci <- pns_ci[pns_ci$P005 %in% c(" ", "2"),]

### 5- Missings
pns_ci <- pns_ci[pns_ci$Q03001 != " " & pns_ci$Q060 != " "
                & pns_ci$Q00201 != " " & !is.na(pns_ci$P00104),]

##### PASSO 3 #####
##### Recodificacao de Variaveis #####
pns_rv <- pns_ci

### 1- 'idade'
colnames(pns_rv)[colnames(pns_rv)=="C008"] <- "idade"
pns_rv$idadefac <- pns_rv$idade
pns_rv <- pns_rv %>% mutate(idadefac = case_when(
  (idade>= 18 & idade <= 24) ~ '0 18 a 24',

```

```
(idade >= 25 & idade <= 39) ~ '1 25 a 39',
(idade >= 40 & idade <= 59) ~ '2 40 a 59',
(idade >= 60) ~ '3 >= 60'
))
pns_rv$idedefac <- as.factor(pns_rv$idedefac)

### 2- 'sexo'
## Antes: 1, M; 2, F
colnames(pns_rv)[colnames(pns_rv)=="C006"] <- "sexo"
pns_rv <- pns_rv %>% mutate(sexo = case_when(
  (sexo == 1) ~ '0 Mas',
  (sexo == 2) ~ '1 Fem'))

### 3- 'raca'
## Antes: 1,Branca; 2,Preta; 3,Amarela; 4,Parda;
## 5,Indigena; 9,Ignorado;
## Dps: 0,Parda; 1,Branca; 2,Preta; 3,Amarela/Indigena/Ignorado
pns_rv$raca <- pns_rv$C009
pns_rv$raca[pns_rv$raca=="4"] <- "0"
pns_rv$raca[pns_rv$raca=="5"] <- "3"
pns_rv$raca[pns_rv$raca=="9"] <- "3"
pns_rv <- pns_rv %>% mutate(raca2 = case_when(
  raca == 1 ~ '1 Branca',
  raca == 2 ~ '2 Preta',
  raca == 3 ~ '3 Ama/Ind/Ign',
  raca == 0 ~ '0 Parda'
))
```

```
pns_rv$raca2 <- as.factor(pns_rv$raca2)

### 4- estado_civil
## Antes: 1,Casado(a); 2,Separado(a) ou desquitado(a)
## judicialmente ou divorciado(a); 3,Viuvo(a); 4,Solteiro(a)
## Dps: 0,Solteiro; 1,Casado; 2,Separado/Divorciado/Viuvo
colnames(pns_rv)[colnames(pns_rv)=="C011"] <- "estado_civil"
pns_rv$estado_civil[pns_rv$estado_civil=="4"] <- "0"
pns_rv$estado_civil[pns_rv$estado_civil=="3"] <- "2"
pns_rv <- pns_rv %>% mutate(estado_civil = case_when(
  (estado_civil == 0) ~ '0 solteiro',
  (estado_civil == 1) ~ '1 Casado',
  (estado_civil == 2) ~ '2 Sepa/Divo/Viu'
))
pns_rv$estado_civil <- as.factor(pns_rv$estado_civil)

### 5- `regioes`
## Antes: 11,Rondonia; 12,Acre; 13,Amazonas; 14,Roraima;
## 15,Para; 16,Amapa; 17,Tocantins
## 21,Maranhao; 22,Piaui; 23,Ceara; 24,Rio Grande do Norte;
## 25,Paraiba; 26,Pernambuco; 27,Alagoas; 28,Sergipe; 29,Bahia
## 31,Minas Gerais; 32,Espirito Santo; 33,Rio de Janeiro;
## 35,Sao Paulo
## 41,Parana; 42, Santa Catarina; 43,Rio Grande do Sul
## 50,Mato Grosso do Sul; 51,Mato Grosso; 52,Goias;
## 53,Distrito Federal
## Depois: 0 = `Norte`; 1 = `Nordeste`; 2 = `Sudeste`;
```

```
## 3 = 'Sul'; 4 = 'Centro-Oeste'
colnames(pns_rv)[colnames(pns_rv)=="V0001"] <- "regioes"
pns_rv$regioes <- as.numeric(pns_rv$regioes)
pns_rv$regioes[pns_rv$regioes %in% 11:17] <- 0
pns_rv$regioes[pns_rv$regioes %in% 21:29] <- 1
pns_rv$regioes[pns_rv$regioes %in% 31:35] <- 2
pns_rv$regioes[pns_rv$regioes %in% 41:43] <- 3
pns_rv$regioes[pns_rv$regioes %in% 50:53] <- 4
pns_rv$regioes <- as.character(pns_rv$regioes)
pns_rv <- pns_rv %>% mutate(regioes = case_when(
  (regioes == 0) ~ '1 Norte',
  (regioes == 1) ~ '0 Nordeste',
  (regioes == 2) ~ '2 Sudeste',
  (regioes == 3) ~ '3 Sul',
  (regioes == 4) ~ '4 Centro-Oeste'
))

### 6- Sit_domicilio
pns_rv$sit_domicilio <- pns_rv$V0026
pns_rv <- pns_rv %>% mutate(sit_domicilio = case_when(
  (sit_domicilio == 1) ~ '0 Urbano',
  (sit_domicilio == 2) ~ '1 Rural'
))
pns_rv$sit_domicilio <- as.factor(pns_rv$sit_domicilio)

##### PASSO 4 #####
##### Criando novas Variaveis #####
```

```

pns_cv <- pns_rv

### doenca_cronica -TEVE MUDANCA EM DUAS VARIAVEIS
## Q074 = '1' ou Q079 = '1' ou Q084 = '1' ou Q088 = '1'
## ou Q116 = '1' ou Q120 = '1' ou Q124 = '1' ou Q128 = '1'
## ou Q092='1' ou Q110='1'
pns_cv$doenca_cronica <- 0
pns_cv$doenca_cronica[pns_cv$Q074 == "1" |
  pns_cv$Q079 == "1" | pns_cv$Q084 == "1" |
  pns_cv$Q088 == "1" | pns_cv$Q11604 == "1" |
  pns_cv$Q120 == "1" | pns_cv$Q124 == "1" |
  pns_cv$Q128 == "1" | pns_cv$Q092 == "1" |
  pns_cv$Q11006 == "1"] <- 1
pns_cv <- pns_cv %>% mutate(doenca_cronica = case_when(
  (doenca_cronica == 0) ~ '0 Nao',
  (doenca_cronica == 1) ~ '1 Sim'
))
pns_cv$doenca_cronica <- as.factor(pns_cv$doenca_cronica)

### dcv (Doenca Cardiovascular)
## Q063 = '1' ou Q068 = '1'
## Perguntas: 1, Sim; 2, Nao
pns_cv$dcv <- 0
pns_cv$dcv[pns_cv$Q06306 == "1" | pns_cv$Q068 == "1"] <- 1

##### PASSO 5 #####
##### ESCORE DE SCV - FATORES DE SAUDE #####

```

```
### fav_pa (favoravel pressao arterial)
## Q00201 = '2' THEN 1 ELSE 0
pns_cv$fav_pa <- 0
pns_cv$fav_pa[pns_cv$Q00201 == "2"] <- 1

### fav_glu (favoravel glicose)
## Q03001 = '2' THEN 1 ELSE 0
pns_cv$fav_glu <- 0
pns_cv$fav_glu[pns_cv$Q03001 == "2"] <- 1

### fav_col (favoravel colesterol)
## Q060 = '2' THEN 1 ELSE 0
pns_cv$fav_col <- 0
pns_cv$fav_col[pns_cv$Q060 == "2"] <- 1

##### Escore - Fatores de Saude
### factors_score (soma dos fatores favoraveis)
pns_cv$factors_score <- pns_cv$fav_pa + pns_cv$fav_glu +
                        pns_cv$fav_col
### factors_bin (Fatores favoraveis binarios)
pns_cv$factors_bin <- 0
pns_cv$factors_bin[pns_cv$factors_score == 3] <- 1

##### ESCORE DE SCV - COMPORTAMENTOS SAUDAVEIS #####
### fav_imc (favorável imc) <25kg/m2
### NAO USAMOS VAR MEDIDA - W00103, W00203 - 5781 OBS
### NAO USAMOS VAR AUTORREF - P00103, P00403 - 12843 MISSINGS
```

```
### USAMOS AUTORREF IMPUT - P00104, P00404 - 1 MISSING
pns_cv$fav_imc <- 0
pns_cv$fav_imc[(pns_cv$P00104/((pns_cv$P00404)^2))<0.0025]<-1

### 'fav_exerc' (favorável exercício físico) >=150 min/week
pns_cv$fav_exerc <- 0
pns_cv$fav_exerc[(pns_cv$P03701*60+
                  pns_cv$P03702)*pns_cv$P035>= 150] <- 1

### DIETA
## 'fav_peixe' (favorável consumo de peixe) (>=2 times/week)
pns_cv$fav_peixe<- 0
pns_cv$fav_peixe[pns_cv$P015>= 2] <- 1
## 'fav_acucar' (favorável consumo de açúcar) (<5 cups/week)
pns_cv$fav_acucar<- 0
pns_cv$fav_acucar[(pns_cv$P02001+pns_cv$P02002)<5] <- 1
## 'fav_veg' (consumo de frutas e vegetais) (7 days/week)
pns_cv$fav_veg<- 0
pns_cv$fav_veg[pns_cv$P018==7 & pns_cv$P00901== 7] <- 1
## 'fav_sal' (favorável consumo de sal)
# (salt consumption levels rated low or appropriate)
pns_cv$fav_sal<- 0
pns_cv$fav_sal[pns_cv$P02601>= 3] <- 1
## 'dieta_score' (soma dos fatores favoráveis dieta)
pns_cv$dieta_score <- (pns_cv$fav_peixe+ pns_cv$fav_acucar+
                      pns_cv$fav_veg+ pns_cv$fav_sal)
### 'fav_dieta' (favorável dieta)
```

```
pns_cv$fav_dieta <- 0
pns_cv$fav_dieta[pns_cv$dieta_score==4] <- 1

### NAO FUMANTE
## 'fumante' (/* actual smoking */)
pns_cv$fumante<- 0
pns_cv$fumante[pns_cv$P050 == 3] <- 1
## ex_fumante'
pns_cv$ex_fumante<- 0
pns_cv$ex_fumante[pns_cv$P052 == 3] <- 1
pns_cv$ex_fumante[pns_cv$P052 %in% c(1,2) &
                    pns_cv$P05901 > '01'] <- 1
### 'fav_fumo' (favorável consumo de fumo) never smoked
# or quit smoking > 12 months
pns_cv$fav_fumo <- 0
pns_cv$fav_fumo[pns_cv$fumante==1 & pns_cv$ex_fumante==1]<-1

##### Escore - Comportamento Saudaveis
### `Saude_score`
pns_cv$saude_score <- (pns_cv$fav_imc+ pns_cv$fav_exerc+
                       pns_cv$fav_fumo+ pns_cv$fav_dieta)
### 'saude_bin' (saude em binário)
pns_cv$saude_bin <- 0
pns_cv$saude_bin[pns_cv$saude_score==4] <- 1

##### ESCORE DE SAUDE CARDIOVASCULAR - FINAL #####
### 'cvh_score' - 8 categorias
```

```

# factors_score + saude_score (favor verificar distribuição)
pns_cv$cvh_score <- (pns_cv$factors_score+pns_cv$saude_score)

##### PASSO 6 #####

##### DETERMINATES SOCIOECONOMICOS #####

#### ESCOLARIDADE ####
### 'esc_score'
# VDD004 e VDD004A são iguais em 2019
pns_cv$esc_score <- 1
pns_cv$esc_score[pns_cv$VDD004A %in% c(3,4)] <- 2
pns_cv$esc_score[pns_cv$VDD004A %in% c(5,6)] <- 3
pns_cv$esc_score[pns_cv$VDD004A >= 7] <- 4
pns_cv <- pns_cv %>% mutate(escolaridade = case_when(
  (esc_score == 1) ~ '0 < middleschool',
  (esc_score == 2) ~ '1 < highschool',
  (esc_score == 3)~ '2 < university',
  (esc_score == 4) ~ '3 university')
)
pns_cv$escolaridade <- as.factor(pns_cv$escolaridade)

#### EMPREGABILIDADE ####
### EMPREGO - VDE002
# 1 - PESSOAS OCUPADAS; 2 - PESSOAS NAO OCUPADAS;
# MISSING - PESSOAS FORA DA FORÇA DE TRABALHO
pns_cv$ocupacao <- 0
pns_cv$ocupacao[pns_cv$VDE002==1] <- 1

```

```
pns_cv <- pns_cv %>% mutate(ocupacao = case_when(
  (ocupacao == 0) ~ '0 Nao',
  (ocupacao == 1) ~ '1 Sim'
))
pns_cv$ocupacao <- as.factor(pns_cv$ocupacao)

##### PASSO 7 #####
##### WEALTH INDEX #####
wealth <- pns_cv
### Regiao Urbana e Rural - wealth$V0026

### water_supply - Fonte de Agua - A005010
wealth$water_supply <- 0
wealth$water_supply[wealth$A005010==1] <- 1

### floor - Material do piso principal- A004010
wealth$floor <- 0
wealth$floor[wealth$A004010 %in% c(1,2)] <- 1

### walls - Principal parede - A002010
wealth$walls <- 0
wealth$walls[wealth$A002010 %in% c(1,4)] <- 1

### roof - Principal telhado - A003010
wealth$roof <- 0
wealth$roof[wealth$A003010 %in% c(2,4)] <- 1
wealth$roof[wealth$A003010 == 3 & wealth$A001 == 2] <- 1
```

```
### Combustível de cozinha - não tem em 2019

### housekeeper - Serviços Domésticos - A02101
wealth$housekeeper <- 0
wealth$housekeeper[wealth$A02101 ==1 ] <- 1

### Eletricidade - Não está contemplada em 2019
### TV - A018011 # Variável retirada - mais de 95%
### Rádio - Não está contemplada em 2019
### Relógio - Não está contemplada em 2019

### car - Automóveis - A018027
wealth$car <- 0
wealth$car[wealth$A018027 ==1 ] <- 1

### Terra agrícola - Não está contemplada em 2019
### Animais domésticos - A02201 - variável retirada
### Conta bancária - Não está contemplada em 2019
### Tipos de janela - Não está contemplada em 2019
### DVD - Não está contemplada em 2019

### cell_phone - Telefone Celular - A018019
wealth$cell_phone <- 0
wealth$cell_phone[wealth$A018019 ==1 ] <- 1

### computer - Computador - A018023
```

```
wealth$computer <- 0
wealth$computer[wealth$A018023 ==1 ] <- 1

### microwave - Microondas - A018021
wealth$microwave <- 0
wealth$microwave[wealth$A018021 ==1 ] <- 1

### washing - Maquina de lavar roupa - A018015
wealth$washing <- 0
wealth$washing[wealth$A018015 ==1 ] <- 1

### water_drink - Agua para beber - A009010
wealth$water_drink <- 0
wealth$water_drink[wealth$A009010 %in% c(1,5) ] <- 1

### Cozinha - Não está contemplada em 2019

### sanitary_drain - Esgoto - A01501
wealth$sanitary_drain <- 0
wealth$sanitary_drain[wealth$A01501 %in% c(1,2) ] <- 1

### garbage - LIXO - A016010
wealth$garbage <- 0
wealth$garbage[wealth$A016010 %in% c(1) ] <- 1

### Geladeira - A018013 # Variável retirada - mais de 95%
```

```
### home_phone - Telefone Fixo - A018017
wealth$home_phone <- 0
wealth$home_phone[wealth$A018017 ==1 ] <- 1

### motorcycle - Motocicleta - A018025
wealth$motorcycle <- 0
wealth$motorcycle[wealth$A018025 ==1 ] <- 1

### internet - Internet - A01901
wealth$internet <- 0
wealth$internet[wealth$A01901 ==1 ] <- 1

### crowding - Pessoas/Comodo - C001/A011<3
wealth$crowding0 <- wealth$C001/wealth$A011
wealth$crowding <- 0
wealth$crowding[wealth$crowding0<3] <- 1

### bathrooms - Pessoas/Banheiro - C001/A011 <=4
wealth$bathrooms0 <- 15
wealth$bathrooms0[wealth$A01401>0] <- wealth$C001[
    wealth$A01401>0]/wealth$A01401[
    wealth$A01401>0]
wealth$bathrooms <- 0
wealth$bathrooms[wealth$bathrooms0 <= 4] <- 1

### Dados wealth index - 18 variáveis
wealth$id <- NA
```

```
wealth$id <- 1:nrow(wealth)
wealth_index <- subset(wealth,
                       select = c(id, V0026, water_supply,
                                   floor, walls, roof,
                                   housekeeper, car,
                                   cell_phone, computer,
                                   microwave, washing,
                                   water_drink, sanitary_drain,
                                   garbage, home_phone,
                                   motorcycle, internet,
                                   crowding, bathrooms))

library(psych)
### Critérios para adequação da base de dados: Correlações
# tetracóricas, teste de Barlett e Teste KMO
(corrTetra<- tetrachoric(wealth_index[c(-1,-2)]))
corPlot(corrTetra$rho, numbers=TRUE, upper=FALSE, diag=FALSE)
# Teste de Bartlett
cortest.bartlett(corrTetra$rho, n=nrow(wealth_index[c(-1,-2)]))
# significativa: critério atendido
#Teste de Kaiser-Meyer-Olkin (KMO)
KMO(corrTetra$rho)
#KMO overall > 0,8 e KMO dos itens > 0,8
#Critério atendido (exceto cell_phone, internet, crowding)

#2. Numero de fatores
# Calculo dos autovalores
round(eigen(corrTetra$rho)$values, 2)
```

```
# Calculo da proporcao explicada por cada fator
proporcao_explicacao<-eigen(corrTetra$rho)$values/
  sum(eigen(corrTetra$rho)$values)
# Calculo da proporcao de explicacao acumulada
proporcao_acumulada <- cumsum(proporcao_explicacao)
# Imprime a proporcao acumulada com 2 casas decimais
round(proporcao_acumulada,2)

#PCA do psych
pca0<-psych::principal(wealth_index[c(-1,-2)],rotate="varimax",
  nfactors=1, scores=T, cor="tet")
sub0 = subset(wealth_index, select = -c(id, V0026))
#Tirar sit_domicilio

pca1<-psych::principal(sub0, rotate="varimax", nfactors=1,
  scores=TRUE, cor="tet")
sub1 = subset(sub0, select = -c(motorcycle)) #0,01

pca2<-psych::principal(sub1, rotate="varimax", nfactors=1,
  scores=TRUE, cor="tet")
sub2 = subset(sub1, select = -c(crowding))

pca3<-psych::principal(sub2, rotate="varimax", nfactors=1,
  scores=TRUE, cor="tet")
sub3 = subset(sub2, select = -c(housekeeper))

pca4<-psych::principal(sub3, rotate="varimax", nfactors=1,
```

```
                scores=TRUE, cor="tet")
sub4 = subset(sub3, select = -c(walls))

pca5<-psych::principal(sub4, rotate="varimax", nfactors=1,
                scores=TRUE, cor="tet")

# pca5 todos h2 > 0,3
names(pca5$communality)
# "water_supply", "floor", "roof", "car", "cell_phone",
# "computer", "microwave", "washing", "water_drink", "garbage",
# "sanitary_drain", "home_phone", "internet", "bathrooms" .
#14 variáveis
# tirou: motorcycle, crowding, housekeeper e walls

#Construção do índice
pns_cv$pca_score_wealth_index=pca5$scores[,1]
library(fabricatr) # tamanho classes - 77494/5
pns_cv <- pns_cv %>%
    mutate(wealth_index = ntile(pca_score_wealth_index, 5))
```

D.2 Códigos em linguagem SAS

Códigos para a categorização da variável SCV em 3 categorias (SCV3cat), estimação de prevalências e realização de análises logísticas binárias, ordinais e multinomiais (Artigo 2).

```
/* CRIANDO SCV CATEGORIZADO */
PROC SQL;
CREATE TABLE DADOS AS
SELECT *,
```

```
(CASE WHEN CVH_SCORE IN (0,1,2) THEN '0 BAIXO'
      WHEN CVH_SCORE IN (3,4,5) THEN '1 INTERMED'
      WHEN CVH_SCORE IN (6,7) THEN '2 ALTO'
      ELSE '' END) AS CVH_CAT,
(CASE WHEN CVH_SCORE IN (6,7) THEN '1 FAV'
      ELSE '0 NFAV' END) AS CVH_BIN
FROM PNS_CV
ORDER BY cvh_cat;
QUIT;

proc surveyfreq data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  TABLES cvh_SCORE / CL;
run;

proc surveyfreq data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  TABLES cvh_cat / cl;
run;

proc surveyfreq data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  TABLES cvh_BIN / CL;
```

```
run;

/*CONTAGEM*/
proc surveyfreq data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  tables escolaridade ocupacao wealth_index
  sexo idadefac raca2 estado_civil
  doenca_cronica regioes sit_domicilio / cl;
run;

proc surveyfreq data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  tables escolaridade ocupacao wealth_index
  sexo idadefac raca2 estado_civil
  doenca_cronica regioes sit_domicilio;
  by cvh_cat;
run;

/* CVH LOGISTICA - 2 CAT - COMPLETO */
PROC SURVEYLOGISTIC data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
```

```
class escolaridade ocupacao wealth_index
sexo idadefac raca2 estado_civil
doenca_cronica regioes
    sit_domicilio / ref=first param=reference;
model cvh_BIN(descending) = escolaridade ocupacao wealth_index
    sexo idadefac raca2 estado_civil doenca_cronica
    regioes sit_domicilio ;
run;

/* CVH MULTINOMIAL - 3 CAT - COMPLETO */
PROC SURVEYLOGISTIC data=dados;
    weight V00291;
    strata v0024;
    cluster upa_pns;
    class escolaridade ocupacao wealth_index
sexo idadefac raca2 estado_civil
doenca_cronica regioes
sit_domicilio / ref=first param=reference;
    model cvh_cat(descending) = escolaridade ocupacao
        wealth_index sexo idadefac raca2 estado_civil
        doenca_cronica regioes sit_domicilio / LINK=glogit;
run;

/* CVH ORDINAL - 3 CAT - COMPLETO */
PROC SURVEYLOGISTIC data=dados;
    weight V00291;
    strata v0024;
```

```
cluster upa_pns;
class escolaridade ocupacao wealth_index
sexo idadefac raca2 estado_civil
doenca_cronica regioes
sit_domicilio / ref=first param=reference;
model cvh_cat(descending) = escolaridade ocupacao
wealth_index sexo idadefac raca2 estado_civil
doenca_cronica regioes sit_domicilio;
run;

/* CVH ORDINAL - 7 CAT - COMPLETO */
PROC SURVEYLOGISTIC data=dados;
weight V00291;
strata v0024;
cluster upa_pns;
class escolaridade ocupacao wealth_index
sexo idadefac raca2 estado_civil
doenca_cronica regioes
sit_domicilio / ref=first param=reference;
model CVH_SCORE(descending) = escolaridade ocupacao
wealth_index sexo idadefac raca2 estado_civil
doenca_cronica regioes sit_domicilio;
run;

/* VARIACOES DO MODELO MULTINOMIAL - 3 CAT */
/* minimamente ajustado */
```

```
PROC SURVEYLOGISTIC data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  class escolaridade ocupacao
wealth_index sexo
      idadefac raca2 / ref=first param=reference;
  model cvh_cat(descending) = escolaridade ocupacao wealth_index
  sexo idadefac raca2 / LINK=glogit;
run;

/* moderadamente ajustado */
PROC SURVEYLOGISTIC data=dados;
  weight V00291;
  strata v0024;
  cluster upa_pns;
  class escolaridade ocupacao wealth_index
sexo idadefac raca2 estado_civil
regioes sit_domicilio / ref=first param=reference;
  model cvh_cat(descending) = escolaridade ocupacao wealth_index
      sexo idadefac raca2 estado_civil regioes
      sit_domicilio / LINK=glogit;
run;
```