

Edson Melo Muniz

**Avaliação nutricional de adultos no Brasil e
sua associação com características
sociodemográficas, comportamentais e de
saúde**

Niterói – RJ, Brasil
05 de janeiro de 2017



Universidade Federal Fluminense

Edson Melo Muniz

**Avaliação nutricional de adultos no Brasil e
sua associação com características
sociodemográficas, comportamentais e de
saúde**

Trabalho de Conclusão de Curso

Projeto Final II apresentado para obtenção do grau de Bacharel em
Estatística pela Universidade Federal Fluminense.

Orientador: Prof. José Rodrigo de Moraes

Niterói - RJ, Brasil

05 de janeiro de 2017



Universidade Federal Fluminense

Edson Melo Muniz

Avaliação nutricional de adultos no Brasil e sua associação com características sociodemográficas, comportamentais e de saúde

Projeto Final II de graduação sob o título “*Avaliação nutricional de adultos no Brasil e sua associação com características sociodemográficas, comportamentais e de saúde*”, defendido por Edson Melo Muniz e aprovada em 05 de janeiro de 2017, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

Dr. José Rodrigo de Moraes

Orientador

Departamento de Estatística – UFF

Me. Patrícia Carvalho de Jesus

Núcleo de Pesquisa em Micronutrientes – UFRJ

Dra. Cecília de Souza Fernandez

Departamento de Análise – UFF

Niterói, 05 de janeiro de 2017

M965 Muniz, Edson Melo

Avaliação nutricional de adultos no Brasil e sua associação com características sociodemográficas, comportamentais e de saúde / Edson Melo Muniz. – Niterói, RJ : [s.n.], 2017. 43f.

Orientador: Prof. José Rodrigo de Moraes.

Trabalho de Conclusão de Curso (Estatística) - Universidade Federal Fluminense, 2017.

1. Saúde pública. 2. Regressão logística multinomial. 3. Regressão logística binária. 4. Índice de massa corporal. 5. Pesquisa nacional de saúde. I. Título.

CDD 614.07

Agradecimentos

Primeiramente, aos meus falecidos pais Carlos Edson dos Santos Muniz e Marinilsa Melo Muniz, por terem me educado de forma adequada dando toda liberdade de escolha para os caminhos a seguir na vida e por terem me dado o exemplo para alcançar cada uma das qualidades que adquiri na vida.

Ao meu irmão Márcio de Melo Muniz, que carrega o nome da nossa família ao meu lado. À minha namorada Viny Meiriño do Val, por toda a força que sempre me transmite e por aturar todos os meus momentos de estresse. E aos meus amigos por seguirem me apoiando apesar das distâncias impostas.

Ao meu orientador e professor durante todo curso de graduação, Dr. José Rodrigo de Moraes pelos ensinamentos. A quem atribuo o meu interesse por modelagem estatística.

A todos os professores do departamento de Estatística da UFF pelos conhecimentos aprendidos e pelo auxílio no processo de amadurecimento acadêmico.

A Universidade Federal Fluminense por me acolher e possibilitar meu conhecimento adquirido e o título de Bacharel em Estatística.

A Vale S/A por possibilitar meu ingresso no mercado de trabalho e meu desenvolvimento profissional.

Resumo

O índice de massa corporal, embora não indique a composição corporal, é um indicador de avaliação nutricional que apresenta vantagens em estudo epidemiológicos, entre as quais pode-se mencionar a facilidade de cálculo e a sua relação com a morbimortalidade. Além de ser um método barato e não invasivo para avaliar o estado nutricional em estudos populacionais. Neste trabalho foi utilizado primeiramente o modelo logístico multinomial, usando os dados da PNS 2013, para explicar o IMC com três níveis (magreza, eutrofia, excesso de peso), entretanto foi necessário dicotomizar este desfecho. Assim o presente trabalho estabeleceu associação entre as características sociodemográficas, comportamentais e de saúde e a chance de excesso de peso em adultos. No ajuste do modelo, foram incorporados os pesos amostrais, estratos e conglomerados. Usando o teste de Wald de significância geral, verificou-se que sete variáveis explicativas estavam associadas significativamente com a chance de excesso de peso, considerando o nível de significância de 5%. Observou-se maior chance de excesso de peso entre adultos do sexo masculino, mais velhos (de 50 a 59 anos em comparação a faixa de 20 a 29 anos), com ocupação profissional, com nível de instrução médio ou superior incompleto (comparativamente ao nível superior completo ou pós-graduação) e entre adultos residentes em domicílios com esgotamento sanitário adequado. Além disso, há maior de chance de excesso de peso entre adultos não fumantes e com diagnóstico de alguma doença. Assim, destaca-se a necessidade de realização de ações voltadas, sobretudo, para promover mudanças comportamentais adequadas, a fim de prevenir doenças e promover a saúde.

Palavras-chaves: Regressão logística multinomial, regressão logística binária, índice de massa corporal e Pesquisa Nacional de Saúde

Sumário

Índice de Tabelas	vii
Índice de Quadros	viii
Índice de Figuras	ix
1. Introdução	1
1.1. Epidemiologia	1
1.2. Avaliação nutricional com foco em prevenção da saúde	1
1.3. Índice de Massa Corporal	3
1.4. Modelo conceitual para o estado nutricional	5
2. Objetivos	9
2.1. Objetivo Geral	9
2.2. Objetivos Específicos	9
3. Material e Métodos	10
3.1. Pesquisa Nacional de Saúde 2013	10
3.2. Plano Amostral da PNS 2013	11
3.3. População de estudo	13
3.4. Variáveis de estudo	13
3.4.1. Índice de Massa Corporal	13
3.4.2. Variáveis sociodemográficas, comportamentais e de saúde	15
3.5. Modelo de Regressão Logística Multinomial	16
3.5.1. Especificação do modelo	16
3.5.2. Método de estimação dos parâmetros	19
3.5.2.1. Método de Máxima Pseudo-Verossimilhança (MPV)	20
3.5.3. Razão de chance e sua interpretação	23
3.5.4. Inferência sobre os parâmetros do modelo	24
3.5.4.1. Teste de Wald de significância individual	24
3.5.4.2. Teste de Wald de significância geral	25
3.5.4.3. Medidas de qualidade do ajuste	27
4. Resultados	30
4.1. Distribuições de frequências	30
4.2. Resultados do ajuste do modelo logístico multinomial	32
4.3. Resultados do ajuste do modelo logístico binário	35
5. Discussão e Considerações Finais	41
6. Referências	44

Índice de Tabelas

Tabela 1: Classificação do estado nutricional dos adultos segundo o IMC.....	14
Tabela 2: Distribuição dos adultos segundo os três níveis do estado nutricional.....	30
Tabela 3: Distribuição (%) dos adultos por níveis de estado nutricional, segundo as variáveis sociodemográficas, comportamentais e de saúde.....	31
Tabela 4: Teste F de Wald de significância geral dos parâmetros dos modelos logísticos multinomiais univariados e multivariados considerando o estado nutricional com três níveis (magreza, eutrofia e excesso de peso).....	33
Tabela 5: Resultados do ajuste do modelo logístico multinomial explicativo do estado nutricional de adultos de 20 a 59 anos de idade, considerando apenas as variáveis selecionadas.....	34
Tabela 6: Teste F de Wald de significância geral dos parâmetros dos modelos logísticos binários univariados e multivariados considerando o estado nutricional com dois níveis (magreza/eutrofia e excesso de peso).....	36
Tabela 7: Resultados do ajuste do modelo logístico binário explicativo da chance de excesso de peso, considerando apenas as variáveis explicativas selecionadas.....	37
Tabela 8: Avaliação da capacidade preditiva do modelo.....	40

Índice de Quadros

Quadro 1: Listagem das Variáveis sociodemográficas, comportamentais e de saúde de adultos de 20 a 59 anos de idade.....	15
Quadro 2: Número de indivíduos classificados segundo as categorias observadas e preditas de <i>Y</i>	28

Índice de Figuras

Figura 1: Determinantes do estado nutricional.....	6
Figura 2: Curva ROC para o modelo logístico binário selecionado (incluindo a área sob esta curva).....	39
Figura 3: Análise gráfica das medidas de sensibilidade e especificidade, considerando diferentes pontos de cortes.....	40

1. Introdução

1.1. Epidemiologia

Atualmente, entre as principais causas de mortalidade no mundo, destacam-se as doenças crônicas não transmissíveis, que são consideradas evitáveis por meio de hábitos e comportamentos que contribuem para caracterizar um estilo de vida saudável como alimentação adequada, prática regular de exercícios físicos, controle de estresse, administração de períodos de lazer e o cuidado pessoal e com o meio ambiente (BARBOSA et al., 2016).

No Brasil, as doenças crônicas não transmissíveis têm um impacto econômico devido aos gastos para o controle de enfermidades como diabetes, hipertensão e doenças cardiovasculares. Além disso, essas doenças contribuem para o surgimento de longas filas e períodos de espera para o atendimento médico básico, principalmente nas unidades de saúde pública, e também em muitas unidades particulares, assim como para o aumento da necessidade de procedimentos médicos mais aprofundados como exames especializados e cirurgias, entre outros problemas (BARBOSA et al., 2016).

Nesse contexto, a avaliação do estado nutricional tem papel fundamental na promoção da saúde e na prevenção destas doenças. Segundo Theme-Filha et al. (2005), um conjunto de fatores de risco de doenças crônicas não transmissíveis já foram bem estabelecidos na literatura, entre eles destacam o uso do tabaco, consumo de álcool, obesidade, hipertensão arterial, níveis elevados de colesterol, baixo consumo de frutas e vegetais, e sedentarismo. O desenvolvimento de ações que visem a prevenção das doenças crônicas não transmissíveis e de seus fatores de risco é fundamental para promoção de saúde da população.

1.2. Avaliação nutricional com foco em prevenção da saúde

O diagnóstico e o acompanhamento do estado nutricional dos indivíduos possibilitam identificar precocemente os fatores determinantes de agravos nutricionais e assim prevenir os riscos à saúde (LOURENÇO et al., 2011).

Segundo Acuña & Cruz (2004), as alterações do estado nutricional (desnutrição e obesidade, entre outros) apresentam relação com sérios problemas de saúde.

A avaliação nutricional tem como objetivo identificar distúrbios nutricionais e os pacientes com risco elevado de apresentar complicações associadas ao estado nutricional, possibilitando uma intervenção adequada para que os indivíduos possam receber terapia nutricional, de forma a auxiliar na recuperação ou manutenção do estado de saúde individual, assim como monitorizar a eficácia da intervenção dietoterápica (CUPPARI, 2002; BARBOSA, 2016).

A avaliação nutricional pode ser feita através de métodos convencionais ou de métodos não convencionais. Os não convencionais, apesar de serem métodos altamente sensíveis, considerados como de referência para medição dos compartimentos corporais, apresentam limitações no seu uso, pois costumam necessitar de equipamentos de custo elevado e de alto nível técnico para sua execução (ACUÑA & CRUZ, 2004; CUPPARI, 2002). Além disso, alguns desses métodos são invasivos e necessitam para a sua realização de local apropriado e avaliadores especializados. Entre os métodos não convencionais estão a hidrodensitometria, tomografia computadorizada e a ressonância magnética (CUPPARI, 2002).

Os métodos convencionais, por sua vez, são tradicionalmente usados e amplamente estimulados pelo fato de sua utilização já estar consolidada na prática clínica e em estudos epidemiológicos de base populacional, além de apresentarem custo aceitável e refletirem adequadamente o estado nutricional do indivíduo.

Na avaliação nutricional um dos conhecimentos básicos são os dados antropométricos. Tais dados são obtidos a partir da antropometria que fornece informações a respeito dos compartimentos corporais, mais especificamente em relação à gordura e ao músculo. A antropometria é a medida do tamanho corporal e de suas proporções. No que diz respeito à composição corporal, sabe-se que a massa corporal humana pode ser separada em dois grupos: massa gorda (ou gordura corporal) e massa magra (massa livre de gordura). A massa magra é basicamente constituída por proteínas, água intra e extracelular e conteúdo mineral ósseo (CUPPARI, 2002). A avaliação da composição corporal possibilita que possíveis anormalidades nutricionais sejam diagnosticadas,

contribuindo para melhorar a eficiência das intervenções nutricionais (CUPPARI, 2002).

A avaliação antropométrica é considerada um método simples e de baixo custo, seguro e prático de ser executado com o objetivo de identificar o estado nutricional dos indivíduos (LOURENÇO et al., 2011). Apresenta uma boa confiabilidade, refletindo adequadamente o estado nutricional dos indivíduos.

As medidas mais adotadas na avaliação antropométrica são o peso corporal, a altura, o índice de massa corporal, as pregas cutâneas e as circunferências do braço, cintura e quadril (CUPPARI, 2002).

1.3. Índice de Massa Corporal

A avaliação do estado nutricional de adultos era classicamente feita através do conceito de "peso ideal" obtido pela relação da massa corporal em função da estatura em comparação com um padrão antropométrico. Basicamente, se o indivíduo tinha massa corporal acima de 20% do padrão, era classificado como obeso. Há consenso de que a utilização de padrão único universal para adultos é inadequada, tendo em vista a grande variação das estaturas médias das populações adultas de diferentes regiões ou países.

Ficou evidente a necessidade do desenvolvimento de indicadores antropométricos do estado nutricional capazes de refletir adequadamente a composição corporal. Além disso, tais indicadores deveriam ser simples de serem obtidos e que não necessitassem de padrão para comparação. Assim, os indicadores antropométricos foram desenvolvidos baseados nos dados de massa corporal e estatura, com o objetivo de haver, teoricamente, uma fraca correlação dos mesmos com a estatura e uma forte correlação com a massa corporal e outras medidas de gordura corporal (ANJOS, 1992).

A partir da década de 70, a razão da massa corporal pela estatura ao quadrado passou a ser denominada Índice de Massa Corporal (IMC).

O Índice de Massa Corporal (IMC) foi calculado a partir da equação criada por Quételet, e por isso também conhecido como índice de Quételet, sendo considerado um indicador adequado para diagnósticos nutricionais e calculado de forma simples através da razão entre o peso corporal e a altura ao quadrado, podendo ser expresso por kg/m^2 , como mostrado abaixo:

$$IMC = \frac{Peso}{Altura^2}$$

O peso corporal (Kg) é a soma de todos os componentes corporais e reflete o equilíbrio proteico-energético do indivíduo. Dessa forma, alterações no peso do corpo refletem mudanças nos níveis de proteínas e energia que compreendem a reserva adiposa e muscular. O peso corporal reflete uma condição nutricional recente do indivíduo, enquanto a altura é um indicador que reflete uma condição nutricional de longa duração (WAITZBERG, 2000).

A medição do peso corporal e da altura é feita por métodos diretos utilizando-se equipamentos específicos, como a balança, o estadiômetro ou o antropômetro. A aferição dessas medidas pode ser feita por meio de alguns métodos indiretos (alternativos) para indivíduos impossibilitados de utilizarem os métodos diretos. Entre os métodos indiretos para a estimativa da estatura do indivíduo pode-se citar a altura do joelho, a extensão ou envergadura dos braços, a altura sentada e a estatura recumbente¹, esta uma alternativa para pacientes confinados no leito (WAITZBERG, 2000; RODRIGUES et al., 2010).

A altura é resultado de fatores genéticos ou de fatores nutricionais e ambientais, sendo os primeiros responsáveis por 80% da definição de uma pessoa ser alta ou baixa. O ser humano tem se tornado mais alto nas últimas gerações por causa de fatores externos como a melhoria da nutrição.

Apesar do IMC não indicar a composição corporal, evidenciando a relação massa magra versus massa gorda, o IMC apresenta uma série de vantagens em estudos epidemiológicos, entre elas pode-se citar a facilidade de cálculo através de dados de massa corporal e estatura facilmente mensuráveis, e a sua relação com morbimortalidade. Dessa forma, considera-se a utilização do IMC aceitável como indicador adequado do estado nutricional em associação ou não com outras medidas antropométricas (ANJOS, 1992). O uso do IMC em comparação a um exame de diagnóstico nutricional é um método prático, barato e bem menos invasivo para avaliar o estado nutricional (WAITZBERG, 2000).

Mesmo quando o IMC é calculado usando medidas autorreferidas de peso e altura, mostrou-se ser um indicador válido e útil na avaliação da situação

¹ Posição supina, deitado de face para cima.

nutricional dos indivíduos em estudos populacionais (OLIVEIRA et al., 2012). Thomaz et al. (2013) também concluíram que as medidas de peso, altura e IMC referidas pelos indivíduos são adequadas para o acompanhamento populacional do excesso de peso.

O IMC é considerado o melhor indicador de massa corporal para adultos, sobretudo em estudos populacionais, porém alguns autores destacam limitações importantes deste índice. As principais limitações podem ser evidenciadas através de três relações: A relação com a proporcionalidade do corpo, visto que pessoas com as pernas significativamente curtas para a sua altura terão como consequência valores maiores de IMC. A relação com a estatura, que pode ser ainda mais significativa em algumas situações, principalmente em pré-adolescentes. Por fim, a relação com a massa livre de gordura ou massa magra, sobretudo em homens, pois atletas e indivíduos musculosos podem através do IMC serem classificados numa faixa de obesidade (ACUÑA & CRUZ, 2004). Como o IMC não distingue o peso associado ao músculo ou à gordura corporal, considera-se importante investigar a composição corporal, sobretudo quando os valores do IMC estiverem nos limites ou fora da normalidade (eutrofia). Ao interpretar os pontos de corte de IMC, é recomendável associar a outros fatores de risco (CUPPARI, 2002).

1.4. Modelo conceitual para o estado nutricional

A Figura 1 ilustra os principais determinantes do estado nutricional dos indivíduos, isto é, mostra que o estado nutricional do indivíduo é resultado de uma função de fatores de diferentes tipos (BRASIL, 2008). Assim, existe uma multiplicidade de fatores que podem propiciar a ocorrência de um estado nutricional inadequado.

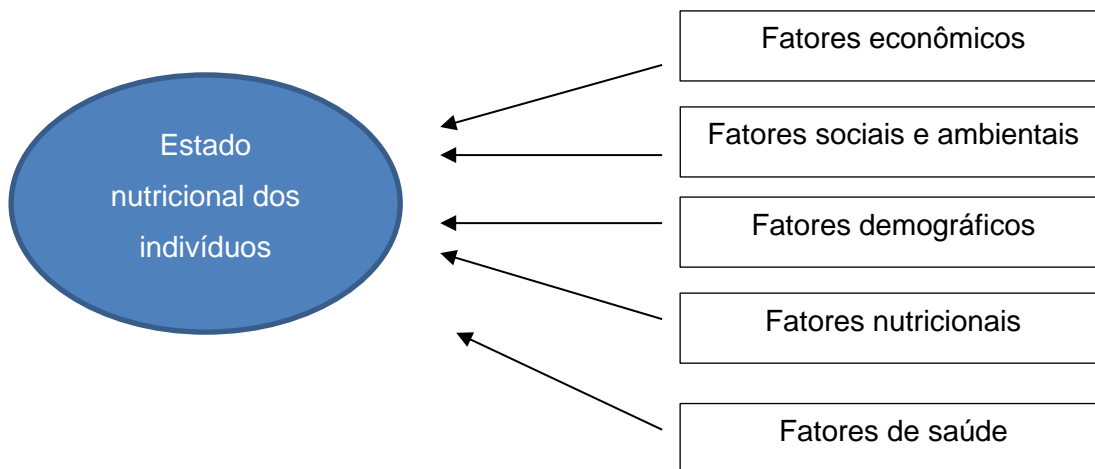


Figura 1: Determinantes do estado nutricional
(Elaborado pelo autor)

Os fatores econômicos reúnem fatores tais como a renda, preço de alimentos e preço de outros itens do orçamento familiar considerados essenciais, tais como habitação, escola, saúde e vestuário (CAMPINO, 1986). Indivíduos que residem em domicílios com acesso precário a renda estão mais vulneráveis a situação de insegurança alimentar e nutricional e às consequências dessa situação, tais como inadequações nutricionais (CABRAL et al., 2014).

Os fatores sociais mostram a importância dos aspectos da relação social para a explicação do estado nutricional, pois os indivíduos encontram-se inseridos em seu meio social e influenciam e são influenciados por estilos de vida, hábitos e comportamentos adquiridos de sua família e de seus amigos (BRASIL, 2008). Entre os fatores sociais e ambientais, estão ainda às condições sociais das famílias, o nível educacional dos pais e aqueles referentes as condições habitacionais (número de cômodos, material usado na parede e teto) e acesso a serviços sociais básicos, como saneamento básico e água potável (CAMPINO, 1986). O estado nutricional é afetado pelas condições de saneamento básico precárias, devido a ocorrências de parasitas intestinais (GRILLO et al, 2000; SEIXAS et al., 2011).

Os fatores demográficos também são considerados como determinantes do estado nutricional de adultos, como sexo, faixa-etária, situação conjugal

(presença ou ausência de companheiro), área de localização do domicílio (urbano, rural), região de residência (Norte, Nordeste, Sul, Sudeste, Centro-Oeste), entre outros. Moraes et al. (2011) mencionam que o estabelecimento de relações entre indicadores de excesso de peso e fatores demográficos e sociais possibilita um efetivo planejamento das ações de saúde e uma adequada intervenção sobre fatores de risco potencialmente modificáveis.

Os fatores nutricionais, por sua vez, estão relacionados a qualidade e a quantidade do alimento consumido e o modo de se alimentar (BRASIL, 2008). A alimentação inadequada influencia tanto na saúde da população quanto no bom desempenho das atividades diárias (TONINI et al., 2013). A adoção de hábitos alimentares adequados diminui os riscos de doenças crônicas e melhoram a saúde e a qualidade de vida. A situação de insegurança alimentar e nutricional pode afetar a saúde, em função dos seus efeitos negativos no estado nutricional dos indivíduos e dos problemas de exclusão social, perda de autoestima, estresse e o sofrimento emocional causados por esta situação (KEPPLE & SEGALL-CORRÊA, 2011).

Há também fatores diretamente relacionados à saúde que são determinantes do estado nutricional dos indivíduos. Pode-se citar a redução da atividade física, motivada por um estilo de vida sedentário que é muito comum nos tempos atuais devido a algumas comodidades que fazem parte da rotina de grande parte da população. A redução do tempo de sono pode acarretar mudanças no organismo que podem ter influência no risco de doenças, tais como diabetes e problemas cardíacos. Além disso, o uso de medicamentos para ansiedade, depressão, hipertensão e diabetes podem contribuir para o aumento de peso (MALFITANO, 2011).

Muitos desses fatores vêm sendo incluídos em estudos sobre estado nutricional de crianças, adolescentes, adultos e idosos. Silveira et al (2009) considerando um estudo de natureza transversal e modelo de regressão de Poisson, considerou fatores sociodemográficos (sexo, idade, escolaridade, renda per capita, estado civil, cor da pele) e fatores relativos à estilo de vida (tabagismo, uso de bebida alcoólica, prática de exercício físico) para explicar a obesidade em idosos, avaliada com base no IMC. Estes autores observaram maiores prevalências de obesidade entre mulheres, nas categorias de mais baixo quartil de renda per capita, entres os menos escolarizados (menos de 1

ano de escolaridade), brancos, solteiros ou separados e não fumantes. Também observaram que idosos com idade mais avançada apresentam menor prevalência de obesidade.

Moraes et al. (2011), também utilizando um estudo transversal estabeleceram, por meio de modelo de regressão de Poisson, a associação entre fatores sociodemográficos e dois desfechos de estado nutricional (sobrepeso e obesidade), em adultos residentes em Ribeirão Preto – SP, Brasil, em 2006. Considerando o desfecho de obesidade, todas as faixas-etárias e a situação conjugal (com companheiro) se mostraram positivamente associados, enquanto que para o desfecho de sobrepeso verificou-se que os seguintes fatores estão positivamente associados com este estado nutricional: ser do sexo masculino, ter 50 ou não mais de idade; estar no “1º terço” de renda e viver “com companheiro”.

Pinto et al. (2010), usando métodos não paramétricos, em estudo transversal envolvendo adolescentes de 10-14 anos matriculados em escolas do município de Recife – PE, no desfecho relacionado ao excesso de peso e obesidade abdominal não houve associação com o sexo do indivíduo. Contudo, o aumento da idade se mostrou negativamente associado com o excesso de peso. Já considerando o grau de maturação sexual (precoce e normal/tardio), as meninas com maturação precoce apresentaram maior prevalência de excesso de peso e obesidade abdominal do que as meninas que apresentaram grau de maturação sexual normal/tardio. Nos meninos, o grau de maturação sexual se mostrou associado positivamente apenas ao excesso de peso.

2. Objetivos

2.1. Objetivo Geral

Estudar a relação entre o índice de massa corporal (utilizado como uma proxy simplificada do estado nutricional) de adultos (com idade de 20 a 59 anos) no Brasil e suas características sociodemográficas, comportamentais e de saúde, usando os dados provenientes da Pesquisa Nacional de Saúde 2013.

2.2. Objetivos Específicos

- Classificar os adultos segundo o seu estado nutricional com base no Índice de Massa Corporal;
- Estudar o perfil nutricional dos adultos levando em consideração as suas características sociodemográficas, comportamentais e de saúde;
- Identificar as características sociodemográficas, comportamentais e de saúde dos adultos relacionadas com seu estado nutricional;
- Estimar a chance dos níveis do estado nutricional dos adultos, a partir das características selecionadas.

3. Material e Métodos

3.1. Pesquisa Nacional de Saúde 2013

A Pesquisa Nacional de Saúde (PNS) foi realizada pelo IBGE em convênio com o Ministério da Saúde. Anteriormente a temática de saúde era abordada apenas nos suplementos da Pesquisa Nacional por Amostra de Domicílios (PNADs 1998, 2003 e 2008) do IBGE, sendo essa pesquisa integrante do Sistema Integrado de Pesquisas Domiciliares (SIPD) do IBGE (DAMACENA et al, 2015; IBGE,2015). A implantação do SIPD deveu-se a necessidade de utilização de uma mesma infraestrutura amostral, de um mesmo cadastro de seleção e de uma amostra comum, chamada “Amostra Mestra”, sendo esta, um dos elementos fundamentais para a implementação deste sistema (FREITAS & ANTONACI, 2014). Entre as vantagens de implantação do SIPD, está o maior espalhamento geográfico e uma maior eficiência amostral

A PNS foi realizada a fim de avaliar o desempenho do sistema nacional de saúde, as condições de saúde da população brasileira (autoavaliação geral de saúde, acesso e utilização de serviços de saúde; diagnóstico de alguma doença crônica, física ou mental), estilos de vida e a vigilância das doenças crônicas não transmissíveis e os fatores de risco associados (IBGE, 2015).

Com a finalidade de descrever adequadamente as características relacionadas às condições de saúde da população adulta brasileira, a PNS realizou uma abordagem da percepção individual da saúde em várias dimensões. Teve como foco investigativo a autoavaliação de saúde, que é considerado um importante indicador que vem sendo frequentemente utilizado, nacional e internacionalmente, para diferenciar a população com relação a morbidade em subgrupos e comparar as necessidades de serviços e recursos de saúde por área geográfica (IBGE, 2015).

Além do levantamento de informações em inquéritos populacionais por meio de questionário, a PNS aferiu medidas físicas, como pressão arterial, peso, altura e circunferência da cintura. Além dessas medidas, a PNS 2013 coletou nos domicílios, pela primeira vez em caráter investigativo de saúde de âmbito nacional, amostras de material biológico (sangue e urina) dos entrevistados para realização de exames complementares, a fim de avaliar as condições clínicas ou

pré-clínicas que necessitam de intervenção (DAMACENA et al, 2015; IBGE). Espera-se que as experiências de campo alcançadas com a realização da Pesquisa Nacional de Saúde 2013 sirvam de subsídio ao desenvolvimento e aprimoramento de novos inquéritos populacionais de saúde no Brasil (DAMACENA et al, 2015).

Inquéritos populacionais são cada vez mais utilizados como meio de se obter informações que permitam esclarecer questões epidemiológicas relevantes. Através desses inquéritos, é possível analisar diferentes aspectos de saúde, tais como a percepção do indivíduo em relação à saúde, o estilo de vida, o acesso e utilização dos serviços de saúde, além de características demográficas e socioeconômicas, o que possibilita a realização de estudos voltados para estabelecer a associação de um conjunto de variáveis com diferentes desfechos de saúde (THEME-FILHA et al, 2005).

No caso da PNS, os dados levantados permitirão a realização do monitoramento das metas de redução das doenças crônicas não transmissíveis, em acordo com a OMS, que incluem o tratamento de hipertensos e diabéticos, a redução do consumo do sal, do tabaco e do álcool, o combate à inatividade física, bem como a prevenção e o controle da obesidade (IBGE, 2015; DAMACENA, 2015).

No que se refere ao instrumento de coleta dos dados da PNS, é importante mencionar que o questionário pode ser dividido em três partes: domiciliar, geral e individual. O domiciliar refere-se à localização, características estruturais do domicílio e outras informações gerais. O geral refere-se a todos os moradores do domicílio quanto à situação socioeconômica e de saúde. O questionário individual, respondido por um morador maior de idade selecionado aleatoriamente entre todos os moradores adultos do domicílio, refere-se informações mais específicas sobre as características de trabalho, percepção de saúde, diagnósticos de doenças crônicas, atendimento médico e estilos de vida (DAMACENA, 2015).

3.2. Plano Amostral da PNS 2013

O plano amostral da PNS foi desenvolvido a partir da Amostra Mestra do Sistema Integrado de Pesquisas Domiciliares (SIPD), do IBGE. A amostra da

PNS é, portanto, uma subamostra selecionada a partir da Amostra Mestre. Amostra Mestre é um conjunto de unidades de áreas que são selecionadas para atender a diversas pesquisas que compõem o SIPD, sendo esta seleção feita do cadastro composto pelos setores censitários da Base Operacional Geográfica do Censo Demográfico de 2010, com exceção dos setores classificados considerados especiais, que compreendem aldeias indígenas, quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos, hospitais e agrovilas de projetos de assentamentos rurais (FREITAS & ANTONACI, 2014).

Com relação ao plano amostral da Amostra Mestre, considerou-se uma estratificação das unidades de áreas (setores censitários) e a seleção dessas unidades de áreas com probabilidade proporcional ao tamanho (PPT), utilizando como medida de tamanho o número de domicílios particulares permanentes (DPPs). De acordo com Freitas & Antonaci (2014), os setores foram estratificados segundo quatro critérios: o administrativo, o geográfico, o de situação censitária e o estatístico. O *critério administrativo* contempla a divisão da UF em capital, municípios da Região Metropolitana (RM) e demais municípios da UF. O *critério geográfico* subdivide as capitais e outros municípios de grande porte em mais estratos, buscando garantir um maior espalhamento territorial, além de facilitar a operação de coleta, controle e redução dos custos de deslocamento entre as unidades da amostra. O *critério de situação censitária* consiste na divisão em área rural e urbana. O *critério estatístico*, corresponde a estratificação estatística dos setores censitários considerando as informações sociodemográficas do Censo Demográfico 2010, como rendimento total e total de domicílios particulares permanentes, com objetivo de agrupar os setores de forma melhorar a precisão das estimativas.

A PNS é uma a pesquisa domiciliar de abrangência nacional que utiliza um plano amostral complexo, isto é, a amostra de indivíduos foi selecionada por meio de uma amostragem de conglomerados em três estágios de seleção, com estratificação dos setores censitários, sendo o procedimento de estratificação o mesmo adotado para a seleção da Amostra Mestre. No primeiro estágio, são selecionados os setores censitários (ou conjunto de setores) que representam as unidades primárias de amostragem (UPA). No segundo estágio, para cada setor da amostra, foram selecionados os domicílios, que correspondem as

unidades secundárias de amostragem (USA). E no terceiro estágio de seleção, para cada domicílio, foi selecionado um morador adulto (18 anos ou mais) para responder ao questionário individual, com probabilidade igual entre todos os moradores adultos residentes no domicílio (DAMACENA et al, 2015; IBGE, 2015). Sendo assim os moradores correspondem as unidades terciárias de amostragem (UTA).

Cabe mencionar que foi sempre adotado o plano aleatório simples sem reposição para a seleção das unidades primárias (setor censitário), secundárias (domicílio) e terciárias (morador) de amostragem (IBGE, 2015).

3.3. População de estudo

A população de estudo refere-se a adultos de ambos os sexos com idade de 20 a 59 anos, no Brasil, residentes em domicílios particulares permanentes. Mulheres grávidas, por terem suas medidas antropométricas alteradas, não foram consideradas no estudo. Além disso, indivíduos que não apresentaram informação sobre o índice de massa corporal ou de alguma das variáveis de interesse não foram incluídos na análise.

3.4. Variáveis de estudo

3.4.1. Índice de Massa Corporal

No módulo P sobre estilos de vida do questionário da PNS, foram levantadas informações sobre o peso corporal (em Kg) e altura (em cm) autorreferidas pelo indivíduo.

Segundo a Organização Mundial de Saúde (OMS), o estado nutricional de adultos pode ser interpretado a partir da classificação do IMC conforme a tabela 1.

Tabela 1: Classificação do estado nutricional dos adultos segundo o IMC.

IMC (kg/m²)	Classificação
IMC < 16,0	Magreza grau III
16,0 ≤ IMC < 17,0	Magreza grau II
17,0 ≤ IMC < 18,5	Magreza grau I
18,5 ≤ IMC < 25,0	Eutrofia
25,0 ≤ IMC < 30,0	Pré-obesidade
30,00 ≤ IMC < 35,0	Obesidade grau I
35,0 ≤ IMC < 40,0	Obesidade grau II
IMC ≥ 40,0	Obesidade grau III

Adaptado: Organização Mundial de Saúde (OMS, 1998)

No presente trabalho, o estado nutricional foi avaliado de maneira simplificada pelo IMC, calculado a partir dos dados autorreferidos pelos entrevistados.

Para fins de modelagem estatística, serão considerados a princípio três categorias de estado nutricional obtidas pela agregação de algumas categorias definidas na tabela 1: Magreza (IMC<18,5); Eutrofia (18,5≤IMC<25); Excesso de peso (Pré-obesidade/Obesidade) (IMC≥25). Caso a capacidade preditiva do modelo logístico multinomial considerando este desfecho politômico (com 3 categorias) não seja adequada, o procedimento a ser adotado consistirá na dicotomização do desfecho, focando na identificação dos fatores associados ao excesso de peso.

A pré-obesidade e a obesidade são definidas como acúmulo anormal de gordura corporal, podendo trazer sérias implicações para a saúde das pessoas. Souza et al. (2014) e Pinto et al. (2010), classificaram obesidade e sobrepeso (pré-obesidade) como excesso de peso, considerando este um sério agravo que atinge diferentes populações e pode acarretar diversas doenças crônicas não transmissíveis.

O módulo P do questionário contém algumas perguntas que permitem calcular o IMC e conhecer um pouco sobre o histórico do indivíduo, com relação aos hábitos alimentares, prática de atividade física, uso de bebidas alcoólicas e uso de cigarros. Ainda neste módulo, para mulheres com idade compreendida

entre 18 e 49 anos de idade, é perguntado se estão ou não grávidas (IBGE, 2015).

3.4.2. Variáveis sociodemográficas, comportamentais e de saúde

O Quadro 1 fornece a listagem das variáveis relativas aos adultos a serem consideradas na análise, bem como as categorias de cada variável.

Quadro 1: Listagem das Variáveis sociodemográficas, comportamentais e de saúde de adultos de 20 a 59 anos de idade.

Variáveis	Categorias
Região de residência	Norte, Nordeste, Sudeste, Sul, Centro-Oeste.
Área de localização do domicílio	Urbano, Rural
Abastecimento de água	Adequado (ligado à rede geral de distribuição), Inadequado (poço, carro-pipa, rios, lagos e igarapés ou água da chuva armazenada)
Esgotamento sanitário	Adequado (Rede geral de esgoto), Inadequado (vala, direto para rio, lago ou mar)
Sexo	Masculino, Feminino
Faixa Etária	20 a 29 anos, 30 a 39 anos, 40 a 49 anos, 50 a 59 anos.
Nível de instrução	Sem instrução, fundamental, médio, superior ou pós-graduação
Condição de ocupação	Ocupado, não ocupado
Autoavaliação geral de saúde	Bom, regular, ruim

Variáveis	Categorias (Continuação)
Plano de saúde	Sim, não
Diagnóstico de alguma doença crônica, física ou mental, ou doença de longa duração	Sim, não
Uso de bebida alcoólica	Sim, não
Frequência do exercício físico ou esporte	Até 1 vez por semana, Mais de 1 vez por semana
Tabagismo	Sim, não

3.5. Modelo de Regressão Logística Multinomial

3.5.1. Especificação do modelo

O modelo de regressão logística multinomial é um modelo linear generalizado (MLG), onde a função de ligação é a função logística e a distribuição de probabilidade da variável resposta Y é a distribuição multinomial. Este modelo é utilizado quando a variável resposta Y (desfecho) é uma variável categórica politômica (com três ou mais categorias), isto é, a variável resposta Y apresenta J categorias mutuamente exclusivas.

O modelo de regressão logística multinomial (modelo teórico) é representado por:

$$\ln\left(\frac{p_{ij}}{p_{iJ}}\right) = \mathbf{x}'_i \boldsymbol{\beta}_j ; \quad (3.0)$$

$$i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, J - 1$$

onde:

- $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{i,p-1}) \rightarrow$ vetor coluna, de dimensão $p \times 1$, composto pelas variáveis explicativas (covariáveis ou variáveis *dummy* para os níveis dos fatores) referentes ao i -ésimo elemento;

- $\beta_j = (\beta_{j0}, \beta_{j1}, \beta_{j2}, \dots, \beta_{j,p-1}) \rightarrow$ vetor coluna, de dimensão $p \times 1$, composto pelos parâmetros desconhecidos referentes a j -ésima categoria da variável resposta do modelo;
- $p_{ij} = P(Y_i = j) \rightarrow$ probabilidade do i -ésimo elemento pertencer a j -ésima categoria da variável resposta do modelo;
- $p_{ij} = P(Y_i = J) \rightarrow$ probabilidade do i -ésimo elemento pertencer a J -ésima categoria (base ou referência) da variável resposta do modelo;

O modelo da equação (3.0) pode ser representado, alternativamente, como em (3.1):

$$\ln\left(\frac{p_{ij}}{p_{iJ}}\right) = \beta_{j0} + \beta_{j1}x_{i1} + \beta_{j2}x_{i2} + \dots + \beta_{j,p-1}x_{i,p-1} ; \quad (3.1)$$

$$i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, J - 1$$

Sejam $p_{i1}, p_{i2}, \dots, p_{iJ}$ as probabilidades referentes ao i -ésimo elemento para cada uma das J categorias da variável resposta, tal que a soma dessa probabilidade é igual a unidade, como indicado em (3.2):

$$\sum_{j=1}^J p_{ij} = p_{i1} + p_{i2} + \dots + p_{iJ} = 1 \quad (3.2)$$

No modelo de regressão logística multinomial, os resultados de cada uma das categorias são relativos à categoria de referência, também chamada de categoria base. Assim, a chance do i -ésimo elemento pertencer à categoria j da variável resposta Y , em vez de pertencer a sua categoria de referência J , é dada em (3.3):

$$\frac{p_{ij}}{p_{iJ}} = e^{x_i' \beta_j} ; \quad (3.3)$$

$$i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, J - 1$$

Da equação (3.3), pode-se explicitar a probabilidade p_{ij} como mostrado em (3.4):

$$p_{ij} = p_{iJ} \cdot e^{x_i' \beta_j} \quad (3.4)$$

Substituindo a equação (3.4) na equação (3.2), é possível obter a probabilidade p_{ij} :

$$\sum_{j=1}^J p_{ij} \cdot e^{x_i' \beta_j} = 1$$

$$p_{iJ} = \frac{1}{\sum_{j=1}^J e^{x_i' \beta_j}} \quad (3.5)$$

Substituindo agora a equação (3.5) na equação (3.4), é possível obter a probabilidade p_{ij} :

$$p_{ij} = \frac{e^{x_i' \beta_j}}{\sum_{j=1}^J e^{x_i' \beta_j}} \quad (3.6)$$

Uma vez que J é a categoria de referência ($\beta_J = 0$) a probabilidade p_{ij} definida em (3.6) pode ser expressa alternativamente como em (3.7):

$$p_{ij} = \frac{e^{x_i' \beta_j}}{e^{x_i' \beta_J} + \sum_{j=1}^{J-1} e^{x_i' \beta_j}}$$

$$p_{ij} = \frac{e^{x_i' \beta_j}}{1 + \sum_{j=1}^{J-1} e^{x_i' \beta_j}}; \quad (3.7)$$

$$j = 1, 2, \dots, J$$

A especificação do modelo apresentada neste trabalho de conclusão de curso também é válida para o caso em que se considera uma variável resposta com apenas duas categorias ($J = 2$). Entretanto, o modelo estatístico passa a ser chamado de modelo de regressão logística binária.

3.5.2. Método de estimação dos parâmetros

O método de máxima verossimilhança (MV) consiste em escolher um valor para o parâmetro desconhecido que maximize a probabilidade de obter a amostra particularmente observada, isto é, um valor que torne esta amostra a mais verossímil. Este método de estimação, empregado no caso das observações amostrais serem independentes e identicamente distribuídas (Amostragem Aleatória Simples com reposição), permite obter os denominados estimadores de máxima verossimilhança, que segundo a literatura possui boas propriedades (BUSSAB & MORETTIN, 2013).

Segundo o método de MV, a função de verossimilhança da amostra é definida como o produto das funções de probabilidade de cada variável aleatória componente da amostra, sendo dada por:

$$L(\theta; y_1, y_1 \dots, y_n) = P(Y_1 = y_1; \theta) \cdot (Y_2 = y_2; \theta) \dots \cdot (Y_n = y_n; \theta)$$

que deve ser entendida como uma função de θ . O estimador de máxima verossimilhança de θ é o valor $\hat{\theta}_{MV}$ que maximiza $L(\theta; y_1, y_1 \dots, y_n)$ (BUSSAB & MORETTIN, 2013).

No caso de uma amostragem complexa como a adotada na PNS, é necessário incorporar na modelagem estatística pelo menos os pesos amostrais, os estratos e os conglomerados (unidades primárias de amostragem) (MORAES et al., 2012). No Método de Máxima Pseudo-Verossimilhança (MPV), os pesos amostrais são incorporados na análise através das equações de estimação dos parâmetros e da estimação da matriz de variância-covariância dos estimadores dos parâmetros do modelo (PESSOA & SILVA, 1998).

Sendo assim, para considerar o plano amostral da pesquisa é necessário que esteja disponível no banco de dados informações detalhadas sobre os elementos da amostra, especificamente os estratos e as unidades primárias de amostragem, além dos pesos amostrais. Cabe ressaltar que mesmo quando o estimador pontual de MPV coincide com o estimador usual de MV, o estimador da variância obtida pelo procedimento de MPV deve ser o adotado, ao invés do estimador usual da variância baseado no método de MV, que não considera todas as informações do plano amostral da pesquisa (PESSOA & SILVA, 1998).

A seguir será descrito o procedimento de estimação dos parâmetros do modelo pelo método de MPV, que é o método adotado neste trabalho para

ajustar o modelo logístico multinomial, considerando o plano amostral da PNS 2013. Cabe ressaltar, que o mesmo procedimento de estimação descrito a seguir pode ser aplicado no caso do modelo logístico binário, fazendo algumas adaptações.

3.5.2.1. Método de Máxima Pseudo-Verossimilhança (MPV)

Suponha que y_1, y_2, \dots, y_N são os valores observados das respectivas variáveis aleatórias de Y_1, Y_2, \dots, Y_N independentes e todas com distribuição multinomial, cuja função de probabilidade da variável aleatória Y_i é dada por:

$$P(Y_i = y_i) = p_{i1}^{d_{i1}} \cdot p_{i2}^{d_{i2}} \cdot \dots \cdot p_{ij}^{d_{ij}}; \quad (3.8)$$

$$i = 1, 2, \dots, N \quad e \quad j = 1, 2, \dots, J$$

$$P(Y_i = y_i) = \prod_{j=1}^J p_{ij}^{d_{ij}}$$

onde:

$$p_{ij} = p_{ij}(\boldsymbol{\beta}) = \frac{e^{x_i' \boldsymbol{\beta}_j}}{1 + \sum_{j=1}^{J-1} e^{x_i' \boldsymbol{\beta}_j}}$$

Como a variável resposta Y do modelo de regressão logística multinomial possui J categorias, foi conveniente definir J variáveis binárias, denotadas da seguinte forma:

$$d_{ij} = \begin{cases} 1, & \text{se } Y_i = j \\ 0, & \text{se } Y_i \neq j \end{cases} \quad (3.9); \text{ o que resulta em um e apenas um } d_{ij} = 1 \text{ para}$$

cada observação (POWERS & XIE, 2000).

Se todos os elementos da população finita fossem conhecidos, o que seria possível através de um censo, a função de verossimilhança populacional seria dada por:

$$L(\boldsymbol{\beta} | \mathbf{y}) = \prod_{n=1}^N P(Y_i = y_i) = \prod_{i=1}^N \prod_{j=1}^J p_{ij}^{d_{ij}}, \quad (3.10)$$

e o logaritmo da função de verossimilhança populacional, seria dado por:

$$\ln L(\boldsymbol{\beta}|\mathbf{y}) = \ln \prod_{i=1}^N \prod_{j=1}^J p_{ij}^{d_{ij}}$$

$$\ln L(\boldsymbol{\beta}|\mathbf{y}) = \sum_{i=1}^N \sum_{j=1}^J \ln(p_{ij})^{d_{ij}}$$

$$\ln L(\boldsymbol{\beta}|\mathbf{y}) = \sum_{i=1}^N \sum_{j=1}^J d_{ij} \cdot \ln(p_{ij}) \quad (3.11)$$

Para maximizar o logaritmo da função de verossimilhança populacional, basta derivá-lo em relação ao vetor $\boldsymbol{\beta}$ e igualar a derivada a $\mathbf{0}$, obtendo assim as equações de verossimilhança populacionais, dadas por:

$$\frac{\partial}{\partial \boldsymbol{\beta}} \ln L(\boldsymbol{\beta}|\mathbf{y}) = \sum_{i=1}^N \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{j=1}^J d_{ij} \cdot \ln(p_{ij}) = \mathbf{0}$$

Então:

$$\sum_{i=1}^N \mathbf{u}_i(\boldsymbol{\beta}) = \mathbf{0} \quad (3.12)$$

onde

$$\mathbf{u}_i(\boldsymbol{\beta}) = \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{j=1}^J d_{ij} \cdot \ln(p_{ij})$$

é um vetor de dimensão $P \times 1$ dos escores da i -ésima unidade populacional.

O lado esquerdo da equação (3.12) é um vetor de totais populacionais, podendo ser denotado por $\mathbf{T}(\boldsymbol{\beta}) = \sum_{i=1}^N \mathbf{u}_i(\boldsymbol{\beta})$ e a solução desse sistema de equações de verossimilhanças populacionais é o estimador de MV do vetor $\boldsymbol{\beta}$, que na abordagem de MPV representa um vetor $\boldsymbol{\beta}_{pp}$ de pseudo-parâmetros que só poderiam ser calculados no caso de realização de um censo. No entanto, apenas tem-se informação sobre as unidades da amostra selecionadas da

população de interesse usando o plano amostral complexo adotado na PNS. Sendo assim, torna-se necessário estimar o vetor de pseudo-parâmetros. Para tanto, estima-se $T(\boldsymbol{\beta})$ usando um estimador linear ponderado da forma $\hat{T}(\hat{\boldsymbol{\beta}}) = \sum_{i=1}^N w_i \mathbf{u}_i(\hat{\boldsymbol{\beta}})$, onde w_i é o peso amostral referente a i -ésima unidade da amostra.

Ao fazer $\hat{T}(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ obtém-se um sistema de equações de pseudo-verossimilhança, cuja solução é o estimador de MPV de $\boldsymbol{\beta}_{pp}$, e conseqüentemente do vetor $\boldsymbol{\beta}$ de parâmetros de interesse. Como em pesquisas por amostragem o tamanho da população é geralmente grande, um estimador adequado para $\boldsymbol{\beta}_{pp}$ também é considerado adequado para $\boldsymbol{\beta}$ (PESSOA & SILVA, 1998).

Uma vez obtido o estimador de MPV, denotado por $\hat{\boldsymbol{\beta}}_{MPV}$, pode-se estimar a probabilidade p_{ij} da i -ésima unidade da amostra pertencer à categoria j da variável resposta Y . Dado o vetor x_i composto pelos valores observados das variáveis explicativas, a probabilidade estimada é dada por:

$$\hat{p}_{ij} = \frac{e^{x_i' \hat{\boldsymbol{\beta}}_j}}{1 + \sum_{j=1}^{J-1} e^{x_i' \hat{\boldsymbol{\beta}}_j}}; \quad (3.13)$$

$$j = 1, 2, \dots, J$$

Através da linearização de Taylor pode-se obter o estimador da matriz de variância-covariância assintótica do estimador $\hat{\boldsymbol{\beta}}_{MPV}$ através da expressão (3.14):

$$\widehat{VAR}(\hat{\boldsymbol{\beta}}_{MPV}) = [J(\hat{\boldsymbol{\beta}}_{MPV})]^{-1} \widehat{VAR} \left[\sum_{i=1}^n w_i \mathbf{u}_i(\hat{\boldsymbol{\beta}}_{MPV}), \right] [J(\hat{\boldsymbol{\beta}}_{MPV})]^{-1} \quad (3.14)$$

onde

$$J(\hat{\boldsymbol{\beta}}_{MPV}) = \left. \frac{\partial \hat{T}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{MPV}} = \sum_{i=1}^n w_i \left. \frac{\partial \mathbf{u}_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}_{MPV}}$$

$\widehat{VAR}[\sum_{i=1}^n w_i \mathbf{u}_i(\hat{\boldsymbol{\beta}}_{MPV})]$ é o estimador da matriz de variância-covariância do estimador do total populacional dos escores.

Segundo Binder (1983), $\hat{\boldsymbol{\beta}}_{MPV}$ tem distribuição assintoticamente normal multivariada, isto é:

$$[\widehat{VAR}(\widehat{\beta}_{MPV})]^{-1/2}(\widehat{\beta}_{MPV} - \beta) \sim N(\mathbf{0}; \mathbf{I}) \quad (3.15)$$

Esta estatística (3.15) é utilizada para realização de inferências sobre β usando amostras grandes.

3.4.3. Razão de chance e sua interpretação

A partir do modelo logístico multinomial representado em (3.0), pode-se obter a chance de um determinado elemento i pertencer a categoria j da variável resposta Y , relativamente a categoria de referência J , é dada por:

$$\frac{p_{ij}}{p_{iJ}} = \frac{P(Y_i = j)}{P(Y_i = J)} = e^{x_i' \beta_j}; \quad (3.16)$$

$$j = 1, 2, \dots, J - 1$$

Além disso, é possível obter também a medida de razão de chances (*odds ratio* – *OR*), que é uma medida de associação que tem como objetivo avaliar o sentido e a força da associação entre as variáveis. Na regressão logística multinomial, calcula-se a razão de chance de se pertencer a cada uma das $J - 1$ categorias da variável resposta Y em relação à categoria de referência J .

A medida de razão de chance pode ser calculada tanto para uma variável explicativa numérica quanto categórica, entretanto como serão consideradas somente variáveis explicativas categóricas na modelagem estatística, nesta seção será descrita a interpretação dessa medida no caso de uma variável categórica (particularmente de uma variável binária).

Desse modo, no caso em que x_k seja uma variável categórica binária ($x_{ik} = 0, 1$), a medida de razão de chances de se pertencer à categoria j em vez de pertencer a categoria de referência J , pode ser interpretada pela comparação entre $x_{ik} = 1$ e $x_{ik} = 0$:

$$OR(j, J | x_{ik}) = \frac{(p_{ij} | x_{ik} = 1) / (p_{iJ} | x_{ik} = 1)}{(p_{ij} | x_{ik} = 0) / (p_{iJ} | x_{ik} = 0)}$$

$$OR(j, J | x_{ik}) = \frac{e^{\beta_{j0} + \beta_{j1}x_{i1} + \dots + \beta_{jk}(1) + \dots + \beta_{j,p-1}x_{i,p-1}}}{e^{\beta_{j0} + \beta_{j1}x_{i1} + \dots + \beta_{jk}(0) + \dots + \beta_{j,p-1}x_{i,p-1}}} = e^{\beta_{jk}} \quad (3.17)$$

Se o coeficiente da variável explicativa for positivo ($\beta_{jk} > 0$), então $OR(j, J | x_{ik}) > 1$, indicando que no grupo 1 ($x_{ik} = 1$) a chance do elemento pertencer à categoria j em vez da categoria J é $[OR(j, J | x_{ik}) - 1] \cdot 100\%$ maior do que no grupo de referência ($x_{ik} = 0$).

Se o coeficiente da variável explicativa for negativo ($\beta_{jk} < 0$), então $OR(j, J | x_{ik}) < 1$, indicando que no grupo 1 ($x_{ik} = 1$) a chance do elemento pertencer à categoria j em vez da categoria J é $[OR(j, J | x_{ik}) - 1] \cdot 100\%$ menor do que no grupo de referência ($x_{ik} = 0$).

Se o coeficiente da variável explicativa for zero ($\beta_{jk} = 0$), a medida de razão e chance é igual a unidade ($OR(j, J | x_{ik}) = 1$), significando que não existe diferença na chance do elemento pertencer à categoria j em vez da categoria J entre os dois grupos.

Assim como o modelo logístico multinomial, o modelo logístico binário também fornece medidas de razão de chance como medidas de associação entre cada variável explicativa e o desfecho, sendo a representação e a interpretação dessas medidas similares as descritas anteriormente.

3.4.4. Inferência sobre os parâmetros do modelo

3.4.4.1. Teste de Wald de significância individual

A seguir são apresentadas as principais etapas de realização do teste de Wald para avaliar a significância individual dos parâmetros do modelo logístico multinomial, no que se refere as hipóteses a serem testadas (hipótese nula H_0

versus a hipótese alternativa H_1), estatística de teste, região crítica (região de rejeição de H_0) e critério de decisão.

I) Hipóteses a serem testadas:

$$\begin{cases} H_0: \beta_{jk} = 0 \\ H_1: \beta_{jk} \neq 0 \end{cases}$$

onde:

β_{jk} é o efeito principal do k -ésimo nível da variável explicativa x referente a j -ésima categoria da variável resposta (equação j); $j = 1, 2, \dots, J - 1$ e $k = 1, \dots, K$.

II) Estatística de teste:

$$\text{Sob } H_0: \beta_{jk} = 0, \text{ temos que: } Z = \frac{\hat{\beta}_{jk}}{\sqrt{\text{VAR}(\hat{\beta}_{jk})}} \sim N(0,1),$$

III) Região crítica:

$$RC = \left\{ z \in \mathbb{R} / z \leq -z_{\alpha/2} \text{ ou } z \geq z_{\alpha/2} \right\}$$

IV) Critério de decisão:

Se $z_{obs} \in RC$ rejeita-se $H_0: \beta_{jk} = 0$ ao nível de significância α , ou seja, o efeito β_{jk} é significativamente diferente de zero. Equivalentemente, rejeita-se $H_0: \beta_{jk} = 0$ se $p - \text{valor} = 2 \cdot P(|Z| > z_{obs})$ é menor que o nível de significância α . Neste trabalho, o nível de significância do teste vai ser fixado em 5%. Por outro lado, se $z_{obs} \notin RC$ (ou $p - \text{valor} > \alpha = 5\%$) não rejeita-se $H_0: \beta_{jk} = 0$ ao nível de significância α , e neste caso o efeito β_{jk} não é significativamente diferente de zero.

3.4.5. Teste de Wald de significância geral

A significância dos parâmetros do modelo também pode ser avaliada através do teste de Wald de significância geral. Seja $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ um vetor de

parâmetros desconhecidos, de dimensão $P \times 1$, de um modelo de regressão. As etapas desse teste são descritas a seguir:

I) Hipóteses a serem testadas:

$$\begin{cases} H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0} \\ H_1: \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0} \end{cases}$$

onde:

\mathbf{C} é uma matriz de constantes de dimensão $R \times P$ de posto completo R ; e $\mathbf{0}$ é um vetor $R \times 1$.

II) Estatística de teste:

Sob $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$, a estatística de teste é dada por:

$$X_W^2 = (\mathbf{C}\hat{\boldsymbol{\beta}})[\mathbf{C}\widehat{\text{VAR}}(\hat{\boldsymbol{\beta}})\mathbf{C}']^{-1}(\mathbf{C}\hat{\boldsymbol{\beta}}),$$

onde:

$\hat{\boldsymbol{\beta}}$ é o estimador de MPV do vetor $\boldsymbol{\beta}$ de parâmetros do modelo e $\widehat{\text{VAR}}(\hat{\boldsymbol{\beta}})$ é o estimador da matriz de variância-covariância do estimador de MPV do vetor $\boldsymbol{\beta}$ de parâmetro do modelo.

De acordo com Binder (1983), a normalidade assintótica do estimador $\hat{\boldsymbol{\beta}}$ e a consistência do estimador da matriz de variância-covariância de $\hat{\boldsymbol{\beta}}$, permite afirmar que:

$$X_W^2 \sim X_R^2$$

Com o objetivo de melhorar a aproximação de estatística de teste, pode-se utilizar a *estatística F-corrigida de Wald*, que está implementada em pacotes estatísticos para a análise de dados provenientes de pesquisas amostrais complexas, dada por:

$$F = \frac{X_W^2}{R} \sim F(R, v),$$

onde:

$v = m - H$, sendo m o número de unidades primárias de amostragem (UPA) e H o número de estratos da amostra, indicando o número de graus de liberdade.

III) Região crítica:

$$RC = \{f \in \mathbb{R} / f \geq f_{\alpha, R, v}\}$$

IV) Critério de decisão:

Se $f_{obs} \in RC$ rejeita-se $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ ao nível de significância $\alpha = 5\%$. Equivalentemente, rejeita-se $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ se $p - valor = P(F > f_{obs})$ é menor que o nível de significância $\alpha = 5\%$, e conclui-se que pelo menos um dos níveis da variável explicativa tem efeito significativo.

Com relação a estratégia de modelagem, primeiramente foram ajustados modelos univariados, isto é, modelos ajustados considerando uma única variável explicativa. Apenas as variáveis que apresentaram associação com o desfecho no máximo ao nível de significância de 20% ($p - valor \leq 0,20$) foram consideradas na análise multivariada. Foram mantidas no modelo final apenas as variáveis com associação significativa ao nível de 5%.

3.4.6. Medidas de qualidade do ajuste

As medidas de qualidade do ajuste servem para avaliar a adequação do modelo para representar os dados observados. Entre estas medidas, destacam-se as taxas de classificações corretas global e condicionais.

Para obter as taxas de classificações corretas pode-se construir uma tabela de dupla-entrada (ou de contingência) para classificar os elementos segundo as categorias observadas e preditas da variável resposta Y . As categorias preditas são obtidas com base nas probabilidades estimadas de cada elemento pertencer a cada uma das J categorias da variável resposta Y . Assim, para cada elemento da amostra é possível obter sua categoria predita como sendo aquela que apresentar a maior probabilidade estimada a partir do modelo selecionado.

O quadro 2 apresenta a classificação dos elementos segundo as categorias observadas e preditas da variável resposta (desfecho) do modelo.

Quadro 2: Número de indivíduos classificados segundo as categorias observadas e previstas de Y .

Categorias observadas	Categorias previstas				Total
	$\hat{Y} = 1$	$\hat{Y} = 2$...	$\hat{Y} = J$	
$Y = 1$	O_{11}	O_{12}	...	O_{J1}	$O_{1.}$
$Y = 2$	O_{21}	O_{22}	...	O_{J2}	$O_{2.}$
\vdots	\vdots	\vdots	...	\vdots	\vdots
$Y = J$	O_{J1}	O_{J2}	...	O_{JJ}	$O_{J.}$
Total	$O_{.1}$	$O_{.2}$...	$O_{.J}$	$O_{..}$

sendo:

O_{ij} → número de elementos referentes a i -ésima categoria observada e a j -ésima categoria prevista da variável resposta do modelo;

O_i → número total de elementos referentes a i -ésima categoria observada da variável resposta do modelo;

O_j → número total de elementos referentes a j -ésima categoria prevista da variável resposta do modelo;

$O_{..}$ → número total de elementos, sendo dado por: $O_{..} = \hat{N} = \sum_{i=1}^J \sum_{j=1}^J O_{ij}$

A taxa global de classificações corretas do modelo, denotada por G , é dada por:

$$G = \left(\frac{O_{11} + O_{22} + \dots + O_{JJ}}{O_{..}} \right) \cdot 100$$

As taxas de classificações corretas condicionais (sensibilidades), isto é, calculadas dentro de cada categoria de Y são dadas por:

Sensibilidade da categoria 1: $S_1 = P(\hat{Y} = 1|Y = 1) = \left(\frac{O_{11}}{O_{1.}} \right) \cdot 100$

Sensibilidade da categoria 2: $S_2 = P(\hat{Y} = 2|Y = 2) = \left(\frac{O_{22}}{O_{2.}} \right) \cdot 100$

...

Sensibilidade da categoria J : $S_J = P(\hat{Y} = J|Y = J) = \left(\frac{o_{JJ}}{o_{J.}}\right) \cdot 100$

As especificidades, por sua vez, são dadas por:

Especificidade da categoria 1: $E_1 = P(\hat{Y} \neq 1|Y \neq 1)$

Especificidade da categoria 2: $E_2 = P(\hat{Y} \neq 2|Y \neq 2)$

...

Especificidade da categoria J : $E_J = P(\hat{Y} \neq J|Y \neq J)$

Desse modo, no caso de $J=2$ (modelo logístico binário), a sensibilidade e a especificidade poderiam ser representadas por: $S = P(\hat{Y} = 1|Y = 1)$ e $E = P(\hat{Y} = 2|Y = 2)$

Na seção 4, são apresentados os resultados obtidos a partir das distribuições de frequências das variáveis explicativas e da variável resposta, bem como os principais resultados do ajuste dos modelos logísticos multinomial e binário.

4. Resultados

4.1. Distribuições de frequências

Observa-se que 54,6% dos adultos apresentaram excesso de peso, enquanto 44,3% são eutróficos. Apenas 1,1% dos adultos pertencem a categoria “magreza”.

Tabela 2: Distribuição dos adultos segundo os três níveis do estado nutricional.

Estado nutricional	% de adultos (N=24.988.437)	IC 95%
Magreza	1,1	[0,8-1,5]
Eutrofia	44,3	[42,5-46,2]
Excesso de peso	54,6	[52,7-56,4]
Total	100,0	-

*Excesso de peso = pré-obesidade/obesidade

De acordo com a tabela 3, destaca-se maior percentual de adultos residentes na região sudeste (44,2%) e em área urbana (93,7%). A maioria dos adultos tem acesso, em suas residências, a abastecimento de água adequado (90,5%), assim como a esgotamento sanitário adequado (84,3%). Quanto às características dos adultos, observa-se um maior percentual de homens (61,4%), de adultos na faixa etária de 20 a 29 anos (33,8%), de adultos com nível médio completo ou superior incompleto (46,3%) e de adultos com ocupação (95,4%). Em relação às variáveis de saúde, observa-se um maior percentual de adultos que autoavaliaram sua saúde como boa (85,2%), que não possuem plano de saúde (52,0%) e não tiveram nenhum diagnóstico de doença crônica, física ou mental (86,7%). Já em relação as variáveis comportamentais, observa-se um maior percentual de uso de bebida alcoólica (41,9%), de frequência de atividade física maior que uma vez por semana (75,0%) e de não fumantes (89,9%).

Na tabela 3, pode-se destacar também um maior percentual de adultos com excesso de peso na área urbana (54,8%) e residindo em domicílios com esgotamento sanitário adequado (55,2%). Quanto às características sócio-demográficas entre adultos, nas faixas etárias mais avançadas, com ocupação profissional (55,4%) e com ensino médio ou superior incompleto, bem

como no grupo dos sem instrução ou fundamental incompleto (em torno de 57%). No que se refere as características comportamentais e de saúde, observou-se maior percentual de excesso de peso entre os adultos com autoavaliação de saúde ruim (65,9%), com diagnóstico de alguma doença crônica, física ou mental (67,9%) e entre os não fumantes (55,5%).

Tabela 3: Distribuição (%) dos adultos por níveis de estado nutricional, segundo as variáveis sociodemográficas, comportamentais e de saúde.

Variáveis	% Adultos	Estado nutricional			Total
		Excesso de peso	Eutrofia	Magreza	
Região de residência					
Norte	6,3	54,8	44,7	0,6	100
Nordeste	22,3	53,2	45,5	1,3	100
Sudeste	44,2	53,5	44,8	1,7	100
Sul	18,5	55,7	43,4	0,9	100
Centro-Oeste	8,8	54,3	45,0	0,7	100
Área de localização do domicílio					
Urbana	93,7	54,8	44,1	1,1	100
Rural	6,3	50,7	47,6	1,7	100
Abastecimento de água					
Adequado	90,5	54,6	44,3	1,1	100
Inadequado	9,5	54,1	44,9	1,0	100
Esgotamento sanitário					
Adequado	84,3	55,2	43,7	0,9	100
Inadequado	15,7	51,2	47,6	0,9	100
Sexo					
Masculino	61,4	59,4	39,7	0,9	100
Feminino	38,6	46,9	51,7	1,4	100
Faixa etária					
De 20 a 29 anos	33,8	44,1	54,3	1,6	100
De 30 a 39 anos	31,3	55,7	43,3	1,1	100
De 40 a 49 anos	20,9	64,0	35,2	0,8	100
De 50 a 59 anos	13,9	63,4	36,2	0,4	100
Nível de instrução					
Sem instrução/Fund. Incomp.	12,9	57,3	41,3	1,4	100
Fundamental/Médio Incomp.	12,6	52,3	47,2	0,5	100
Médio/Sup. Incomp.	46,3	56,7	42,2	1,1	100
Superior completo ou pós-graduação	28,2	50,9	47,9	1,2	100

Variáveis	% Adultos	Estado nutricional (Continuação)			
		Excesso de peso	Eutrofia	Magreza	Total
Condição de ocupação					
Sim	95,4	55,4	43,6	1,1	100
Não	4,6	38,1	59,9	2,0	100
Autoavaliação geral de saúde					
Bom	85,2	53,4	45,5	1,0	100
Regular	13,8	60,8	38,0	1,2	100
Ruim	1,1	65,9	30,6	3,5	100
Plano de saúde					
Sim	48,0	55,4	43,8	0,8	100
Não	52,0	53,8	44,8	1,4	100
Diagnóstico de alguma doença					
Sim	13,3	67,9	30,4	1,6	100
Não	86,7	52,5	46,5	1,0	100
Uso de bebida alcoólica					
Sim	58,1	55,0	44,0	1,0	100
Não	41,9	54,0	44,8	1,3	100
Frequência de exercício físico					
Até 1 vez por semana	25,0	53,5	45,1	1,4	100
Mais de 1 vez por semana	75,0	54,9	44,1	1,0	100
Tabagismo					
Sim	10,1	46,4	52,7	0,9	100
Não	89,9	55,5	43,4	1,1	100

4.2. Resultados do ajuste do modelo logístico multinomial

A tabela 4 apresenta os resultados do teste F de Wald de significância geral para os parâmetros dos modelos logísticos multinomiais univariados e multivariados.

No que tange aos modelos logísticos multinomiais univariados, verificou-se com base no teste *F de Wald de significância geral*, que cinco variáveis explicativas não tiveram associação ao nível de significância de 20% ($p - valor > 0,20$) com o estado nutricional com três níveis (magreza, eutrofia e pré-obesidade/obesidade): região de residência, área de localização do domicílio, abastecimento de água, uso de bebida alcóolica e frequência do exercício físico (Tabela 4).

Um modelo multivariado foi ajustado considerando conjuntamente aquelas variáveis que tiveram associação significativa com o estado nutricional com três níveis na análise bruta (Tabela 4), e verificou-se que esgotamento sanitário, autoavaliação geral de saúde e plano de saúde não apresentaram significância estatística ao nível de 5%. Cabe ressaltar, que a condição de ocupação passou a apresentar associação significativa com a retirada das três variáveis mencionadas. Então, o modelo multivariado selecionado foi composto pelas seguintes variáveis explicativas: sexo, faixa etária, nível de instrução, condição de ocupação, diagnóstico de doença e tabagismo (Tabela 4).

Tabela 4: Teste F de Wald de significância geral dos parâmetros dos modelos logísticos multinomiais univariados e multivariados considerando o estado nutricional com três níveis (magreza, eutrofia e excesso de peso).

Variável	Modelos univariados* (p-valor)	Modelo multivariado inicial** (p-valor)	Modelo multivariado final** (p-valor)
Região de residência	0,494	-	-
Área de localização do domicílio	0,223	-	-
Abastecimento de água	0,927	-	-
Esgotamento sanitário	0,148	0,114	-
Sexo	<0,001	<0,001	<0,001
Faixa etária	<0,001	<0,001	<0,001
Nível de instrução	0,047	0,006	0,014
Condição de ocupação	<0,001	0,054	0,048
Autoavaliação geral de saúde	0,007	0,412	-
Plano de saúde	0,167	0,169	-
Diagnóstico de alguma doença	<0,001	<0,001	<0,001
Uso de bebida alcoólica	0,628	-	-
Frequência de exercício físico	0,635	-	-
Tabagismo	0,006	<0,001	<0,001

*Análise bruta **Análise multivariada

A Tabela 5 fornece as razões de chance e seus intervalos de 95% de confiança, além dos p-valores do teste de Wald de significância individual para o modelo logístico multinomial selecionado.

Tabela 5: Resultados do ajuste do modelo logístico multinomial explicativo do estado nutricional de adultos de 20 a 59 anos de idade, considerando apenas as variáveis selecionadas.

Variáveis	Excesso de peso <i>vs</i> Eutrofia (Equação 1)			Magreza <i>vs</i> Eutrofia (Equação 2)		
	Razão de chance (OR)	IC 95% para OR	p-valor	Razão de chance (OR)	IC 95% para OR	p-valor
Sexo						
Masculino	1,863	(1,614;2,151)	<0,001	0,939	(0,500;1,763)	0,844
Feminino	1	-	-	1	-	-
Faixa etária						
De 20 a 29 anos	0,479	(0,374;0,613)	<0,001	3,639	(1,148;11,540)	0,028
De 30 a 39 anos	0,797	(0,628;1,012)	0,063	2,705	(0,905;8,088)	0,075
De 40 a 49 anos	1,118	(0,868;1,441)	0,389	2,202	(0,709;6,837)	0,172
De 50 a 59 anos	1	-	-	1	-	-
Nível de instrução						
Sem instrução/Fundamental Incompleto	1,130	(0,886;1,441)	0,325	1,523	(0,433;5,353)	0,512
Fundamental/Médio Incompleto	1,088	(0,844;1,012)	0,516	0,445	(0,174;1,139)	0,091
Médio/Superior Incompleto	1,358	(1,138;1,621)	0,001	1,027	(0,487;2,166)	0,944
Superior completo ou pós graduação	1	-	-	1	-	-
Condição de ocupação						
Sim	1,612	(1,090;2,384)	0,017	0,833	(0,222;3,118)	0,786
Não	1	-	-	1	-	-
Diagnóstico de alguma doença						
Sim	1,927	(1,545;2,403)	<0,001	2,969	(1,036;8,506)	0,043
Não	1	-	-	1	-	-
Tabagismo						
Sim	0,632	(0,499;0,799)	<0,001	0,697	(0,289;1,682)	0,422
Não	1	-	-	1	-	-

Com base nos resultados do ajuste do modelo logístico multinomial selecionado (equação 1), pode-se observar uma maior chance de excesso de peso (pré-obesidade/obesidade) *versus* eutrofia entre os homens, indivíduos com nível médio ou superior incompleto, ocupados, com diagnóstico de doença crônica, física ou mental. Observou-se também maior chance de excesso de peso *versus* eutrofia entre os indivíduos de 50 a 59 anos (em comparação a faixa

de 20 a 29 anos) e entre os não fumantes. Quanto a equação 2, observa-se uma maior chance de magreza *versus* eutrofia entre indivíduos de 50 a 59 anos (em comparação a faixa de 20 a 29 anos e entre adultos com diagnóstico de doença crônica, física ou mental (Tabela 5). Com relação à equação 2, observou-se significância estatística somente para a faixa de 20 a 29 anos e diagnóstico de alguma doença. Assim adultos de 20 a 29 anos e adultos com diagnóstico de doença tem maior chance de magreza *versus* eutrofia.

Com relação à capacidade preditiva do modelo multinomial selecionado (Tabela 5) verificou-se que o modelo não conseguiu prever todos três os níveis do estado nutricional. Apesar da taxa global de classificações corretas ter sido 60,3%, isto é, mais da metade dos adultos terem sido classificados corretamente pelo modelo multinomial selecionado, a sensibilidade referente a categoria magreza foi de 0%. Já as sensibilidades referentes as categorias eutrofia e excesso de peso (pré-obesidade/obesidade) foram de 41,6% e 76,7%, respectivamente. Isso pode ser explicado pela baixa frequência da categoria magreza, como mostrado na tabela 2. Desta forma, optou-se por agregar os níveis eutrofia e magreza e ajustar o modelo logístico binário para estimar a chance de o adulto ter excesso de peso *versus* eutrofia/magreza.

4.3. Resultados do ajuste do modelo logístico binário

A mesma estratégia adotada para seleção do modelo logístico multinomial foi utilizada na seleção do modelo logístico binário. No que tange aos modelos logísticos binários univariados, verificou-se com base no teste F de Wald de significância geral, que cinco variáveis explicativas não tiveram associação significativa ao nível de significância de 20% ($p - valor > 0,20$) com a chance de o adulto ter excesso de peso: região de residência, abastecimento de água, plano de saúde, uso de bebida alcoólica e frequência do exercício físico (Tabela 6).

Ao ajustar o modelo considerando conjuntamente aquelas variáveis que tiveram associação significativa com chance de excesso de peso na análise bruta (Tabela 6), verificou-se que área de localização do domicílio, esgotamento sanitário e autoavaliação geral de saúde não apresentaram significância estatística ao nível de 5%. Entretanto, após a exclusão das variáveis autoavaliação de saúde e área, a variável esgotamento sanitário passou a ter

efeito significativo ao nível de 5%. Desse modo, o modelo multivariado selecionado foi composto pelas sete variáveis restantes: esgotamento sanitário, sexo, faixa etária, nível de instrução, condição de ocupação, diagnóstico de doença e tabagismo.

Tabela 6: Teste F de Wald de significância geral dos parâmetros dos modelos logísticos binários univariados e multivariados considerando o estado nutricional com dois níveis (magreza/eutrofia e excesso de peso).

Variável	Modelos univariados* (p-valor)	Modelo multivariado inicial** (p-valor)	Modelo multivariado final** (p-valor)
Região de residência	0,859	-	-
Área de localização do domicílio	0,155	0,197	-
Abastecimento de água	0,859	-	-
Esgotamento sanitário	0,057	0,096	0,046
Sexo	<0,001	<0,001	<0,001
Faixa etária	<0,001	<0,001	<0,001
Nível de instrução	0,032	0,007	0,005
Condição de ocupação	<0,001	0,016	0,015
Autoavaliação geral de saúde	0,005	0,264	-
Plano de saúde	0,372	-	-
Diagnóstico de alguma doença	<0,001	<0,001	<0,001
Uso de bebida alcoólica	0,562	-	-
Frequência de exercício físico	0,531	-	-
Tabagismo	0,002	<0,001	<0,001

*Análise bruta **Análise multivariada

A Tabela 7 apresenta os resultados do ajuste do modelo logístico binário multivariado somente com as variáveis explicativas selecionadas.

Tabela 7: Resultados do ajuste do modelo logístico binário explicativo da chance de excesso de peso, considerando apenas as variáveis explicativas selecionadas.

Variáveis	Excesso de peso <i>versus</i> Eutrofia/Magreza		
	Razão de chance (OR)	IC 95% para o OR	p-valor
Esgoto			
Adequado	1,194	(1,003; 1,420)	0,046
Inadequado	1	-	-
Sexo			
Masculino	1,872	(1,622; 2,160)	<0,001
Feminino	1	-	-
Faixa etária			
De 20 a 29 anos	0,471	(0,368; 0,603)	<0,001
De 30 a 39 anos	0,789	(0,622; 1,001)	0,051
De 40 a 49 anos	1,108	(0,859; 1,430)	0,430
De 50 a 59 anos	1	-	-
Nível de instrução			
Sem instrução/Fundamental Incompleto	1,165	(0,911; 1,488)	0,223
Fundamental/Médio Incompleto	1,132	(0,876; 1,463)	0,345
Médio/Superior Incompleto	1,378	(1,153; 1,647)	<0,001
Superior completo ou pós-graduação	1	-	-
Condição de ocupação			
Sim	1,616	(1,097; 2,382)	0,015
Não	1	-	-
Diagnóstico de alguma doença			
Sim	1,864	(1,502; 2,314)	<0,001
Não	1	-	-
Tabagismo			
Sim	0,636	(0,503; 0,803)	<0,001
Não	1	-	-

Analisando as medidas de associação do modelo logístico binário selecionado (Tabela 7), observa-se que os adultos residentes em domicílios com esgotamento sanitário adequado têm uma chance de ter excesso de peso 19,4% maior que os adultos que moram em domicílios considerados inadequados quanto a forma de esgotamento sanitário (OR=1,194; p – valor = 0,046).

Quanto ao sexo do adulto, os homens têm uma chance de ter excesso peso 87,2% maior do que as mulheres (OR=1,872; $p - valor < 0,001$). A chance dos adultos na faixa etária de 20 a 29 anos é 52,9% menor do que a dos adultos com idade mais avançada (de 50 a 59 anos) (OR=0,471; $p - valor = 0,001$).

Com relação ao nível de instrução, observou-se que os adultos com ensino médio completo ou superior incompleto têm chance de excesso de peso 37,8% maior em comparação aos adultos com nível de instrução mais elevado (superior completo ou pós-graduação) (OR=1,378; $p - valor < 0,001$). Quanto a condição de ocupação, verificou-se que os adultos ocupados têm chance de excesso de peso 61,6% maior que os adultos desocupados (OR=1,616; $p - valor = 0,015$). No que se refere ao diagnóstico de saúde, os adultos portadores de doença crônica, física ou mental têm chance de excesso de peso 86,4% maior comparativamente aos adultos sem alguma dessas doenças (OR=1,864; $p - valor < 0,001$).

Quanto ao tabagismo, verificou-se que adultos fumantes têm chance de excesso de peso 36,4% menor do que os adultos não fumantes (OR=0,636; $p - valor < 0,001$).

Quanto à capacidade preditiva do modelo logístico binário selecionado (Tabela 8) verificou-se que o modelo conseguiu predizer os dois níveis do estado nutricional considerados (excesso de peso e eutrofia/magreza). Para um ponto de corte de 0,5; a taxa global de classificações corretas foi de 61,2% e as medidas de sensibilidade e especificidade foram de 74,0% e 45,8%, respectivamente.

Usando a curva ROC para avaliar o poder discriminatório do modelo, obteve-se uma área sob a curva de $A=0,647$ (Figura 2). O intervalo de 95% de confiança para esta área foi de [0,6469;0,6473] mostrando a não inclusão do valor 0,5, o que significa que o modelo selecionado tem capacidade para discriminar os adultos com excesso de peso (pré-obesidade/obesidade) e sem excesso de peso (eutrofia/magreza).

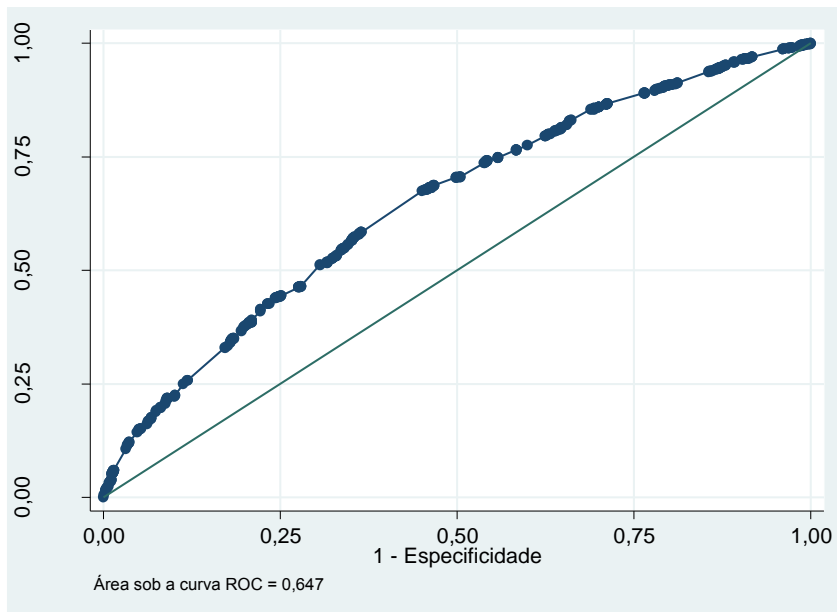


Figura 2: Curva ROC para o modelo logístico binário selecionado (incluindo a área sob esta curva).

Para um ponto de corte de 0,5 observou-se que a especificidade foi baixa ($E < 0,5$) (Tabela 8) o que mostrou a necessidade de se considerar outros valores de ponto de corte. O cruzamento entre as curvas de sensibilidade e especificidade indicou que o ponto de corte ótimo foi de 0,535 (Figura 3).

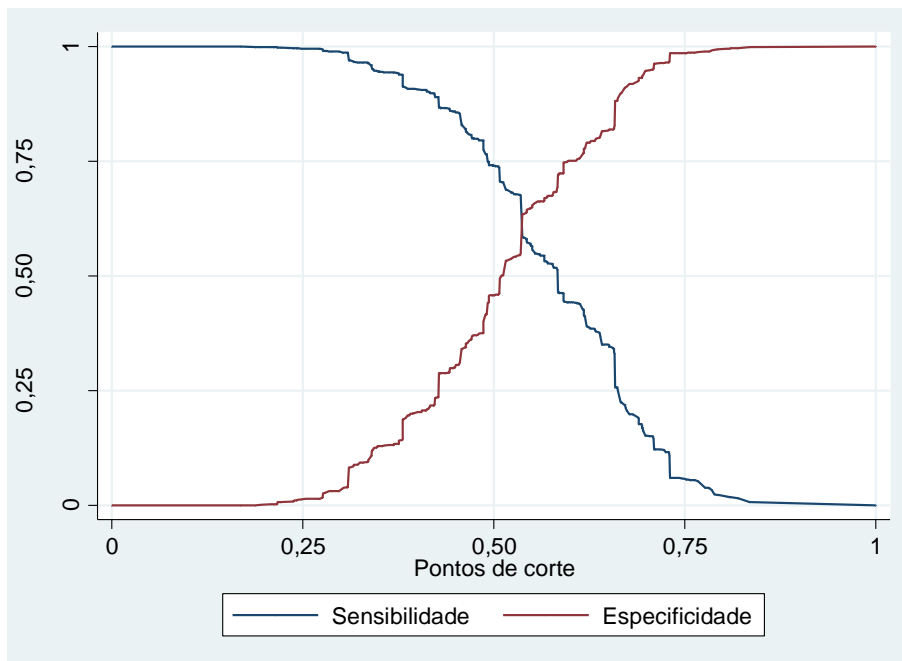


Figura 3: Análise gráfica das medidas de sensibilidade e especificidade, considerando diferentes pontos de cortes.

Para este ponto de corte ótimo ($\gamma=0,535$), a taxa global de classificações corretas foi de 61,8% e as medidas de sensibilidade e especificidade foram de 67,5% e 55,0%, respectivamente (Tabela 8). Segundo Marôco (2010), modelos com sensibilidade e especificidade entre 50% e 80% já apresentam boa capacidade preditiva.

Tabela 8: Avaliação da capacidade preditiva do modelo.

Medidas (%)	Ponto de corte ($\gamma=0,5$)	Ponto de corte ($\gamma=0,535$)
Taxa global de classificações corretas	61,2	61,8
Sensibilidade	74,0	67,5
Especificidade	45,8	55,0

5. Discussão e Considerações Finais

O excesso de peso é considerado um sério problema de crescente evolução em todo o mundo, atingindo populações e pode carregar doenças crônicas não transmissíveis (SOUZA, 2014). No presente estudo, observou-se que a maioria dos adultos (54,6%) apresentaram excesso de peso.

No presente trabalho, indivíduos que residem em domicílios com carência de infraestrutura de serviço básico (esgotamento sanitário inadequado) tendem a ter menos excesso de peso. Uma explicação para isto é que estes indivíduos são em geral mais pobres e, por este motivo, podem apresentar um padrão de gasto energético elevado devido a ocupação e locomoção, que os protegeriam contra o excesso de peso (VIEIRA & SICHIERI, 2008).

O resultado deste trabalho no que se refere a associação entre a faixa etária e o excesso de peso, parece corroborar com o estudo de Teichmann et al. (2006) no sentido que a probabilidade de obesidade aumentou com a idade, mas os autores consideraram na análise mulheres adultas residentes em São Leopoldo – RS. Ainda no estudo de Teichmann et al. (2006), mulheres com idade entre 50 e 60 anos tiveram probabilidade de obesidade três vezes maior que as mulheres na faixa etária de 20 a 29 anos. Na pesquisa de Vedana et al. (2008) a obesidade esteve positivamente associada com a faixa etária para ambos os sexos e esteve negativamente associado com a escolaridade (apenas para mulheres). No presente estudo, observou-se apenas diferença de excesso de peso entre indivíduos nas faixas etárias “de 20 a 29 anos” e de “50 a 59 anos”, sendo a chance de excesso de peso maior para indivíduos mais velhos (50 a 59 anos).

A maior chance de excesso de peso em indivíduos ocupados encontrada neste estudo pode ser explicada em parte pelo ambiente de trabalho impróprio, que pode implicar na adoção de práticas como sedentarismo e hábitos alimentares inadequados neste grupo de indivíduos (SALVE, 2009).

Com relação ao diagnóstico de alguma doença crônica, física ou mental, a associação com o excesso de peso encontrada pode ser explicada, por um lado, pela maior presença de doenças crônicas em adultos obesos, tais como doenças cardiovasculares, que se manifestam na forma de doença cardíaca isquêmica, hipertensão e insuficiência cardíaca (MANCINI, 2001). Por outro lado, a

utilização de medicamentos de uso contínuo por indivíduos com diagnóstico de doenças psiquiátricas ou doenças crônicas como hipertensão e diabetes, por exemplo, podem levar a um ganho excessivo de peso.

Em relação ao tabagismo, foi observado que adultos fumantes têm menor chance de excesso de peso. Isto pode ser comprovado a partir de estudos clínicos e epidemiológicos que demonstram que fumantes têm menor peso corporal que não fumantes e ganham peso quando param de fumar (FARIA et al., 2012). Faria et al. (2012) mostrou em seu estudo que o IMC foi significativamente menor nos fumantes do que nos não fumantes e nos ex-fumantes. Já no primeiro ano após parar de fumar, os indivíduos apresentaram um ganho de peso. Contudo, a interrupção do tabagismo pode reduzir os riscos de morbidade e mortalidade.

Quanto à limitação deste estudo, aponta-se a não inclusão da variável renda, devido a não divulgação dessa informação no banco de dados da PNS 2013, e de outras variáveis nutricionais como, por exemplo, o perímetro braquial considerada como uma medida do estoque de energia do indivíduo. Outra limitação se refere à dificuldade em comparar os resultados de estudos que utilizaram o IMC, devido à diversificação da população de estudo e do uso de diferentes modelos estatísticos (VASCONCELLOS & PORTELA, 2001; MORAES, 2005; TEICHMANN et al., 2006; VEDANA et al., 2008). Neste trabalho, adotou-se inicialmente o modelo logístico multinomial, mas devido à baixa frequência da categoria “magreza” este modelo não apresentou uma boa capacidade preditiva, sendo necessário dicotomizar o desfecho e empregar o modelo de regressão logística binária.

Quanto às potencialidades deste trabalho, destaca-se o ajuste de modelos estatísticos incorporando os principais aspectos do plano amostral da PNS 2013. A inclusão da conglomeração, estratificação e pesos amostrais deve-se a obtenção de estimativas não viciadas para as medidas de associação (razões de chance) e para as medidas de precisão (MORAES et al., 2012; PESSOA & SILVA, 1998).

No que se refere a sugestão de estudos futuros, considera-se relevante fazer um estudo com objetivo de avaliar o efeito de variáveis relativas ao comportamento alimentar (consumo de refrigerantes, doces, salada, carnes, frutas, etc.) no estado nutricional de adultos, utilizando o IMC como uma *proxy* simplificada desse estado nutricional. Dessa forma, com esta sugestão, espera-

se estimular trabalhos que visem a utilização de técnicas de modelagem estatística na área de nutrição. Uma hipótese para a explicação do aumento da obesidade no mundo é a mudança nos padrões alimentares, observando-se diminuição no consumo de fibras, frutas, hortaliças, de arroz e de feijão. Por outro lado, observa-se também o aumento no consumo de produtos industrializados, ricos em gorduras, e açúcares, que caracterizam uma dieta mais calórica, o que predispõe o ganho de peso excessivo (VEDANA, 2008).

Finalmente, com a realização deste trabalho de conclusão de curso, conclui-se que há maior chance de excesso de peso entre adultos do sexo masculino, mais velhos (de 50 a 59 anos em comparação a faixa de 20 a 29 anos), com ocupação profissional, com nível de instrução médio ou superior incompleto (comparativamente ao nível superior completo ou pós-graduação) e entre adultos residentes em domicílios com esgotamento sanitário adequado. Além disso, há maior chance de excesso de peso entre adultos não fumantes e com diagnóstico de alguma doença.

Deste modo, destaca-se a necessidade de realização de ações voltadas, sobretudo, para promover mudanças comportamentais adequadas, a fim de prevenir doenças e promover a saúde. Além disso, destaca-se a importância de realização de estudos que contribuam para melhorar o entendimento sobre os fatores que se associam à obesidade em adultos, considerando-se uma variedade maior de fatores demográficos, socioeconômicos, comportamentais e de estilo de vida, e de saúde.

6. Referências

- ACUÑA, K.; CRUZ, T. Avaliação do Estado Nutricional de Adultos e Idosos e Situação Nutricional da População Brasileira. *Arq. Bras. Endocrinol. Metab.* 2004; 48(3):346-361.
- ANJOS, L.A. Índice de massa corporal ($\text{massa corporal} \cdot \text{estatura}^{-2}$) como indicador do estado nutricional de adultos: revisão da literatura. *Rev. Saúde públ.*, 26: 431-6, 1992.
- BARBOSA, L.B.; VASCONCELOS, S.M.L.; CORREIA, L.O.S.; FERREIRA, R.C. Estudos de avaliação do conhecimento nutricional de adultos: uma revisão sistemática. *Ciência e Saúde Coletiva*, 449-461, 2016.
- BINDER, D.A. On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review*, 51: 279-292, 1983.
- BUSSAB, W.O.; MORETTIN, P.A. Estatística Básica – 8ª edição. 2013.
- BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Ações Programáticas Estratégicas. Saúde do adolescente: competências e habilidades. Editora do Ministério da Saúde, 2008.
- CABRAL, C.S.; LOPES, A.G.; LOPES, J.M.; VIANNA, R.P.T. Segurança alimentar, renda e Programa Bolsa Família: estudo de coorte em municípios do interior da Paraíba, Brasil, 2005-2011. *Cad. Saúde Pública*, 30(2): 393-402, 2014.
- CAMPINO, A.C.C. Aspectos sócio-econômicos da desnutrição no Brasil. *Rev. Saúde públ.*, 20: 83-101, 1986.
- CUPPARI, L. Guia de Nutrição: Nutrição Clínica no Adulto. 2002.
- DAMACENA, G.N.; SZWARCOWALD C.L.; MALTA, D.C.; SOUZA JR, P.R.B.; VIEIRA, M.L.F.P.; PEREIRA, C.A.; NETO, O.L.M.; SILVA JR, J.B. O processo de desenvolvimento da Pesquisa Nacional de Saúde no Brasil, 2013. *Epidemiol. Serv. Saúde*, 24(2): 197-206, 2015.
- FARIA, C.S.; BOTELHO, C.; SILVA, R.M.V.G.; FERREIRA, M.G. Tabagismo e obesidade abdominal em doadores de sangue. *J. Bras. Pneumol.*, 38(3):356-363, 2012.
- FREITAS, M.P.S.; ANTONACI, G.A. Sistema Integrado de Pesquisas Domiciliares Amostra Mestra 2010 e Amostra da PNAD Contínua. IBGE. 2014.

GRILLO, L.P.; CARVALHO, L.R.; SILVA, A.C.; VERRESCHI, I.T.N.; SAWAYA, A.L. Influência das condições socioeconômicas nas alterações nutricionais e na taxa de metabolismo de repouso em crianças escolares moradoras em favelas no município de São Paulo. *Revista da Associação Médica Brasileira*, 46(1): 7-14, 2000.

Instituto Brasileiro de Geografia e Estatística (IBGE). Pesquisa Nacional de Saúde 2013 - Acesso e utilização dos serviços de saúde, acidentes e violências: Brasil, Grandes Regiões e Unidades da Federação. 2015.

KEPPLE, A.W.; SEGALL-CORRÊA, A.M. Conceituando e medindo segurança alimentar e nutricional. *Ciênc. saúde coletiva*, 16(1): 187-199, 2011.

LOURENÇO, A.M.; TAQUETTE, S.R.; HASSELMANN, M.H. Avaliação nutricional: antropometria e conduta nutricional na adolescência. *Adolescência & Saúde*, 8(1): 51-58, 2011.

MALFITANO, A.P.S.; ADORNO, R.C.F.; LOPES, R.E. Um relato de vida, um caminho institucional: juventude, medicalização e sofrimentos sociais. *Interface - Comunic. Saude, Educ.* 2011.

MANCINI, M.C. Obstáculos diagnósticos e desafios terapêuticos no paciente obeso. *Arq. Bras. Endocrinol. Metab.* 45(6): 584-608, 2001.

MAROCO, J. Análise estatística com o *PASW Statistics (ex-SPSS)*. 2010.

MORAES, J.R. Informações Estruturais do Plano Amostral do Estudo Nacional de Despesa Familiar: Critérios para sua Recuperação e Aplicação na Modelagem do Índice de Massa Corporal. 2005.

MORAES, J.R.; MOREIRA, J.P.L.; LUIZ, R.R. Efeito do plano amostral em modelo logístico ordinal: uma análise do estado de saúde autorreferido de adultos no Brasil usando a Pesquisa Nacional por Amostra de Domicílios de 2008. *Cad. Saúde Pública*, 28(5): 913-924, 2012.

MORAES, S.A.; HUMBERTO, J.S.M.; FREITAS, I.C.M. Estado nutricional e fatores sociodemográficos em adultos residentes em Ribeirão Preto, SP, 2006. Projeto OBEDIARP. *Rev. Bras. Epidemiol.*, 14(4): 662-76, 2011.

Organização Mundial de Saúde (OMS), 1998.

OLIVEIRA, L.P.M.; QUEIROZ, V.A.O.; SILVA, M.C.M.; PITANGUEIRA, J.C.D.; COSTA, P.R.F.; DEMÉTRIO, F.; ANJOS, M.C.G.; ASSIS, A.M.O. Índice de massa corporal obtido por medidas autorreferidas para a

classificação do estado antropométrico de adultos: estudo de validação com residentes no município de Salvador, estado da Bahia, Brasil. *Epidemiol. Serv. Saúde*, 21(2): 325-332, 2012.

PESSOA, D.G.C.; SILVA, P.L.N. Análise de Dados Amostrais Complexos. IBGE.1998.

PINTO, I.C.S.; ARRUDA, I.K.G.; DINIZ, A.S.; CAVALCANTI, A.M.T.S. Prevalência de excesso de peso e obesidade abdominal, segundo parâmetros antropométricos, e associação com maturação sexual em adolescentes escolares. *Cad. Saúde Pública*, 26(9): 1727-1737, 2010.

POWERS, D.A.; XIE, Y. Statistical Methods for Categorical Data Analysis. 2000.

RODRIGUES, L.P.F.; RONCADA, M.J. Educação nutricional no Brasil: evolução e descrição de proposta metodológica para escolas. 2008.

RODRIGUES, P.A.; RUFINO, M.C.B.; CORREIA, E.A.; LIMA, J.M.R.; LISBOA, A.Q. Correlação das medidas antropométricas reais do peso e da altura com os métodos de estimativa em pacientes adultos do Hospital Regional de Ceilândia. *Com. Ciências Saúde*, 21(3): 237-244, 2010.

SALVE, M.G.C. Ambiente de trabalho sedentarismo e doenças não transmissíveis de trabalhadores de bancas de jornal. *Salusvita*. 28(3):257-266, 2009.

SILVEIRA, E.A.; KAC, G.; BARBOSA, L.S. Prevalência e fatores associados à obesidade em idosos residentes em Pelotas, Rio Grande do Sul, Brasil: classificação da obesidade segundo dois pontos de corte do índice de massa corporal. *Cad. Saúde Pública*, 25(7): 1569-1577, 2009.

SEIXAS, M.T.L.; SOUZA, J.N.; SOUZA, R.P.; TEIXEIRA, M.C.A.; SOARES, N.M. Avaliação da Frequência de Parasitos Intestinais e do Estado Nutricional em Escolares de uma Área Periurbana de Salvador, Bahia, Brasil. *Rev. de Patologia Tropical*, 40 (4): 304-314, 2011.

SOUZA, M.C.C.; TIBÚRCIO, J.D.; BICALHO; RENNÓ, H.M.S.; DUTRA, J.S.; CAMPOS, L.G.; SILVA, E.S. Fatores Associados à Obesidade e Sobrepeso em Escolares. *Texto Contexto Enferm*. 23(3): 712-9, 2014.

TEICHMANN, L.; OLINTO, M.T.A.; COSTA, J.S.D.; ZIEGLER, D. Fatores de risco associados ao sobrepeso e a obesidade em mulheres de São Leopoldo, RS. *Rev. Bras. Epidemiol.*, 9(3): 360-73, 2006.

THEME-FILHA, M.M.; SZWARCOWALD, C.L.; SOUZA JR, P.R.B. Socio-demographic characteristics, treatment coverage, and self-rated health of individuals who reported six chronic diseases in Brazil, 2003. *Cad. Saúde Pública*, 21: S43-S53, 2005.

THOMAZ, P.M.D.; SILVA, E.F.; COSTA, T.H.M. Validade de peso, altura e índice de massa corporal autorreferidos na população adulta de Brasília. *Ver. Bras. Epidemiol.*, 16(1): 157-69, 2013.

TONINI, E.; BROLL, A.M.; CORRÊA, E.N. Avaliação do estado nutricional e hábito alimentar de funcionários de uma instituição de ensino superior do oeste de Santa Catarina. *O Mundo da Saúde*, 37(3): 268-279, 2013.

VASCONCELLOS, M.T.L.; PORTELA, M.C. Índice de Massa Corporal e sua relação com variáveis nutricionais e sócio-econômicas: um exemplo de uso de regressão linear para um grupo de adultos brasileiros. *Cad. Saúde Pública*, 17(6): 1425-1436, 2001.

VEDANA, E.H.B.; PERES, M.A.; NEVES, J.; ROCHA, G.C.; LONGA, G.Z. Prevalência de Obesidade e Fatores Potencialmente Causais em Adultos em Região do Sul do Brasil. *Arq. Bras. Endocrinol. Metab.*, 52/7:1156-1162, 2008.

VIEIRA, A.C.R.; SICHIERI, R. Associação do Status Socioeconômico com Obesidade. *Physis Revista de Saúde Coletiva*. 18(3): 415-426, 2008.

WAITZBERG, D.L.; FERRINI, M.T. Exame Físico e Antropometria. In: WAITZBERG, D.L. *Nutrição Oral, Enteral e Parenteral na Prática Clínica*. 1(16): 255-278, 2000.