



Universidade Federal do Rio Grande do Sul
Faculdade de Medicina
Programa de Pós-Graduação em Epidemiologia

**Avaliação de métodos de calibração para a análise dos Exames
Laboratoriais da Pesquisa Nacional de Saúde de 2013**

Juliana Sena de Souza

Orientador: Prof. Dr. Rodrigo Citton Padilha dos Reis

Coorientadora: Prof^a. Dr^a. Márcia Helena Barbian

Porto Alegre, Outubro de 2022



Universidade Federal do Rio Grande do Sul
Faculdade de Medicina
Programa de Pós-Graduação em Epidemiologia

Avaliação de métodos de calibração para a análise dos Exames Laboratoriais da Pesquisa Nacional de Saúde de 2013

Juliana Sena de Souza

Orientador: Prof. Dr. Rodrigo Citton Padilha dos Reis

Coorientadora: Prof^a. Dr^a. Márcia Helena Barbian

Dissertação a ser apresentada como requisito parcial à obtenção do título de mestre em Epidemiologia pelo Programa de Pós-graduação em Epidemiologia da Faculdade de Medicina da Universidade Federal do Rio Grande do Sul.

Porto Alegre
2022

CIP - Catalogação na Publicação

Souza, Juliana Sena de
Avaliação de métodos de calibração para a análise
dos exames laboratoriais da Pesquisa Nacional de Saúde
de 2013 / Juliana Sena de Souza. -- 2022.
58 f.
Orientador: Rodrigo Citton Padilha dos Reis.

Coorientadora: Márcia Helena Barbian.

Dissertação (Mestrado) -- Universidade Federal do
Rio Grande do Sul, Faculdade de Medicina, Programa de
Pós-Graduação em Epidemiologia, Porto Alegre, BR-RS,
2022.

1. peso amostral. 2. calibração. 3.
pós-estratificação. 4. raking. I. dos Reis, Rodrigo
Citton Padilha, orient. II. Barbian, Márcia Helena,
coorient. III. Título.

Banca Examinadora

Prof^a. Dr^a. Luciana Neves Nunes, Programa de Pós-Graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul.

Prof. Dr. Gustavo Guimarães de Castro Amorim, Department of Biostatistics, Vanderbilt University Medical Center.

Dr. Luís Gustavo Silva e Silva, DATALAB/Food and Agriculture Organization of the United Nations.

“We do not remember days, we remember moments.”

- Jennifer Niven, *All the Bright Places*

Agradecimentos

O maior e maior agradecimento de todos vai para minha mãe, a pessoa que mais amo no mundo e que fez de tudo nesse mundo por mim e pelos meus irmãos, além de todo apoio incondicional ao longo de toda a minha vida.

Aos meus irmãos, por sempre estarem ao meu lado e vibrarem comigo nas minhas conquistas. Elton, que sempre fala com muito orgulho sobre mim e Valdir, que está presente na minha vida desde antes de eu nascer quando dividíamos o útero juntos.

Aos amores da minha vida, meus sobrinhos, Arthur e Gabriel, por me terem como a tia preferida, por me encherem de amor e carinho.

Ao restante da minha família, que apesar de morarem longe, torcem por mim e pelo meu sucesso e sempre estão presentes comigo, mesmo à distância.

Ao meu orientador, Rodrigo Citton, por ter aceitado me orientar e por ter me dado um ótimo suporte ao longo de todo o mestrado, ainda mais nos tempos de pandemia em que foi um pouco difícil manter o foco na dissertação.

À minha co-orientadora, Márcia Barbian, a quem eu tive o prazer de conhecer em 2017 como minha professora e desde então não me abandonou na minha jornada na graduação e agora aqui no mestrado, tanto o de epidemiologia, quanto o de estatística.

À Luciana Nunes, Gustavo Amorim e Luís Gustavo Silva, por terem aceitado serem parte da minha banca e por contribuírem da melhor forma possível com a minha dissertação. Sou muito grata a vocês por todas as considerações feitas.

Aos professores que tive ao longo da graduação e da pós-graduação, pois todos foram importantes no meu crescimento dentro da universidade.

À universidade, pelo ensino gratuito e de qualidade, mesmo depois de quatro anos de adversidades provocadas por um governo contra a educação.

Ao melhor namorado do mundo, Leandro, por me fazer tão feliz e sempre me apoiar, exaltar e que a partir de agora só vai me chamar de mestre.

E, por fim, mas não menos importante, à todos os amigos que fizeram parte dessa jornada comigo. Não vou mencionar por nome porque são muitos, mas vocês sabem quem são. Muito obrigada por todo o suporte que vocês me deram e ainda me dão.

Sumário

	Lista de abreviaturas e siglas	7
	Resumo	8
	Abstract	9
1	APRESENTAÇÃO	10
2	INTRODUÇÃO	11
2	REVISÃO DA LITERATURA	14
2.1	A Pesquisa Nacional de Saúde de 2013	14
2.1.1	A subamostra de Exames Laboratoriais da PNS	20
2.2	Amostragem	21
2.2.1	Ajustes para não-resposta	23
2.2.2	Calibração dos Pesos	23
2.2.2.1	Pós-Estratificação	25
2.2.2.2	<i>Raking</i>	26
2.2.3	Comparação dos Métodos de Calibração	28
2.2.4	Avaliação do desempenho da calibração	29
3	OBJETIVOS	31
3.1	Objetivo geral	31
3.2	Objetivos específicos	31
	Referências	32
4	ARTIGO	35
5	CONCLUSÕES E CONSIDERAÇÕES FINAIS	56

Lista de abreviaturas e siglas

AAS Amostra Aleatória Simples

AVC Acidente Vascular Cerebral

CNEFE Cadastro Nacional de Endereços para Fins Estatísticos

COPIS Coordenação de População e Indicadores Sociais

CV Coeficiente de Variação

DCNT Doenças Crônicas Não-Transmissíveis

DF Distrito Federal

EPA Efeito do Plano Amostral

ER Erro Relativo

Fiocruz Fundação Oswaldo Cruz

GREG *Generalised Regression Estimator*

HDL *High Density Lipoprotein*

IBGE Instituto Brasileiro de Geografia e Estatística

IMC Índice de Massa Corporal

LDL *Low Density Lipoprotein*

LFS *Labor Force Survey*

PDA *Personal Digital Assistance*

PNAD Pesquisa Nacional por Amostra de Domicílios

PNS Pesquisa Nacional de Saúde

PSA *Propensity Score Adjustment*

SIM Sistema de Informações sobre Mortalidade

SIPD Sistema Integrado de Pesquisas Domiciliares

UPA Unidades Primárias de Amostragem

Vigitel Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico

Resumo

Introdução: A Pesquisa Nacional de Saúde (PNS) é uma pesquisa domiciliar a nível nacional que possui duas edições: 2013 e 2019. Especificamente na PNS de 2013, foram coletadas amostras biológicas através de exames complementares. Devido a dificuldades no trabalho de campo, houve uma perda superior à perda inicialmente prevista e alguns estratos ficaram sem representação amostral. Para contornar esse problema fez-se necessário a utilização de métodos de pós-estratificação para o ajuste dos pesos amostrais.

Objetivo: Comparar os métodos de calibração para os pesos amostrais com os pesos disponibilizados juntamente aos dados da PNS 2013, utilizando como métrica de comparação um conjunto de sete indicadores que auxiliam na avaliação de aspectos da calibração. Além disso, identificar as principais diferenças nas estimativas de proporção e total para as variáveis auxiliares, assim como a prevalência de diabetes *mellitus* na população.

Métodos: O conjunto de microdados da subamostra de exames laboratoriais consiste em uma amostra de 8.952 indivíduos adultos, disponível no *website* da PNS. Para verificar se os objetivos da calibração foram alcançados, foram utilizados sete indicadores que auxiliam na avaliação da calibração. Em adição, os erros relativos (ER) das estimativas dos totais populacionais para as variáveis auxiliares, estimativas de média e coeficiente de variação para um conjunto específico de variáveis também foram estimados. Por fim, para a estimação de diabetes *mellitus*, foi calculado sua prevalência para algumas características de interesse, juntamente com seu intervalo de confiança de 95% (IC95%).

Resultados: Com o método de pós-estratificação foi identificado uma sub-representação de quase todas as categorias. As estimativas do total populacional estimadas através dos pesos laboratoriais diferem bastante da projeção populacional, apesar de não haver diferenças grandes entre as estimativas de proporção, média, prevalências e intervalos de confiança. O *raking* apresentou melhor desempenho em seis dos sete indicadores.

Conclusão: Foi possível observar que os totais populacionais estimados pelos pesos laboratoriais produzem estimativas para o total que não são generalizáveis para a população alvo da PNS. Em termos das comparações com os métodos de calibração, o *raking* apresentou, no geral, melhor desempenho em relação à pós-estratificação.

Palavras-Chave: peso amostral, calibração, pós-estratificação, *raking*.

Abstract

Introduction: The National Health Survey (PNS) is a nationwide household survey that has two editions: 2013 and 2019. Specifically, in the 2013 PNS, biological samples were collected through complementary exams. Due to difficulties in the fieldwork, there was a loss greater than the initially predicted loss and, some strata were left without sample representation. To overcome this problem, it was necessary to use post-stratification methods to adjust sample weights.

Objective: To compare the calibration methods for the sample weights with the weights available together with the 2013 PNS data, using a set of seven indicators as a comparison metric that helps in the evaluation of calibration aspects. Furthermore, identify the main differences in the proportion and total estimates for the auxiliary variables, as well as the prevalence of diabetes mellitus in the population.

Methods: The subsample microdata set of laboratory tests consists of a sample of 8,952 adult individuals, available on the PNS website. To check whether the calibration objectives were achieved, seven indicators were used to help evaluate the calibration. In addition, relative errors (RE) of estimates of population totals for auxiliary variables, mean estimates, and coefficient of variation for a specific set of variables were also estimated. Finally, for the estimation of diabetes mellitus, its prevalence was calculated for some characteristics of interest, along with its 95% confidence interval (95%CI).

Results: With the post-stratification method, an underrepresentation of almost all categories was identified. Estimates of the total population estimated through laboratory weights differ significantly from the population projection, although there are no major differences between estimates of proportion, mean, prevalence, and confidence intervals. Raking performed better in six of the seven indicators.

Conclusion: It was possible to observe that the population totals estimated by laboratory weights produce estimates for the total that are not generalizable to the target population of the PNS. In terms of comparisons with the calibration methods, raking generally performed better than post-stratification.

Keywords: sampling weights, calibration, post-stratification, raking.

1 Apresentação

Este trabalho consiste na dissertação de mestrado intitulada “Avaliação de métodos de calibração para a análise dos Exames Laboratoriais da Pesquisa Nacional de Saúde de 2013”, apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, em 26 de outubro de 2022. O trabalho é apresentado em três partes, na ordem que segue:

1. Introdução, Revisão da Literatura e Objetivos
2. Artigo
3. Conclusões e Considerações Finais.

2 Introdução

As doenças crônicas não transmissíveis (DCNT) são morbidades não transmissíveis que afetam as funções do organismo a longo prazo e englobam a diabetes *mellitus*, as doenças respiratórias crônicas, cerebrovasculares, cardiovasculares e as neoplasias (Júnior et al., 2020). As DCNT são determinadas por diversos fatores sociais e individuais, afetam predominantemente países em desenvolvimento e tem a maior carga de morbimortalidade no mundo (Malta et al., 2019).

Em 2013, de acordo com os dados do Sistema de Informação de Mortalidade (SIM), foram registrados 829.916 óbitos por doenças crônicas não transmissíveis no Brasil, o que corresponde a 68,6% do total bruto no país e 72,6% quando feito correção para sub-registro e redistribuição de causas mal definidas. Do total dessas mortes, estima-se que 85% seja relacionado aos quatro principais grupos de DCNT: doenças cardiovasculares (29,7%), neoplasias (16,8%), doenças respiratórias crônicas (5,9%) e diabetes (5,1%) (Malta et al., 2019).

Portanto, fazem-se necessárias estratégias de prevenção e controle sobre os principais fatores de risco das doenças crônicas não transmissíveis (Malta et al., 2017). O “Plano de ações estratégicas para o enfrentamento das DCNT no Brasil, 2011-2022”, coordenado pelo Ministério da Saúde, foi desenvolvido com o objetivo de enfrentar e deter, em dez anos, essas doenças por meio de três diretrizes: vigilância, informação, avaliação e monitoramento; promoção da saúde; e cuidado integral (Malta et al., 2011).

Nesse contexto, a Pesquisa Nacional de Saúde (PNS) entrou como uma importante e essencial fonte de informação para auxiliar as diretrizes, políticas públicas e ações estratégicas específicas. A PNS é uma pesquisa domiciliar a nível nacional, desenvolvida pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e realizada em convênio com o Ministério da Saúde. É fundamentada em três eixos principais: desempenho do sistema nacional de saúde, condições de saúde da população brasileira e vigilância das doenças crônicas não transmissíveis e fatores de risco associados (IBGE, 2014).

De acordo com resultados obtidos através das análises dos dados da PNS de 2013, estima-se em 45,1% a prevalência de indivíduos adultos com pelo menos uma DCNT, com maior prevalência entre as mulheres (50,4%), em relação aos homens (39,2%) (Malta et al., 2015).

É a pesquisa mais ampla já realizada no país sobre saúde e seus determinantes (Malta et al., 2015). As informações de saúde coletadas dizem respeito a perguntas sobre a percepção do estado de saúde, estilo de vida e doenças crônicas, além de aferições antropométricas e fatores econômicos, sociodemográficos, culturais, comportamentais e

ambientais que influenciam na saúde (Szwarcwald et al., 2019b).

A PNS conseguiu dimensionar o acesso ao diagnóstico e à assistência prestada às doenças crônicas, pois foi possível aprofundar os módulos temáticos dos questionários de saúde. Entre as doenças, foi dado maior detalhamento para a hipertensão arterial, diabetes e depressão, através da investigação do acesso a exames complementares de diagnóstico, consultas com especialistas e medicamentos de uso contínuo. Também foram examinadas as limitações das atividades habituais, sequelas e internações devidas à doença (IBGE, 2014).

Através da PNS foi estimado, por meio do diagnóstico referido, que a prevalência de hipertensão arterial foi de 21,4%, sendo que 81,4% destes indivíduos fizeram uso de medicamentos para tratamento da doença nas duas semanas anteriores à entrevista (Tavares et al., 2015). Em relação à diabetes e a depressão, as prevalências foram de 6,2% e 7,6%, enquanto que o uso de medicamentos orais (ou insulina, no caso da diabetes) foram de 80,2% e 52%, respectivamente (Tavares et al., 2015).

Com o objetivo de traçar o perfil bioquímico de condições pré-clínicas ou clínicas de indivíduos que necessitam de intervenção, amostras biológicas foram coletadas através da PNS de 2013 para a realização de exames complementares (IBGE, 2015b). As coletas foram feitas posteriormente devido ao seu alto custo. Como subamostra laboratorial, foi selecionada aleatoriamente 25% da amostra da PNS. A amostra foi estimada em 12 mil indivíduos, entretanto, devido a taxa de não resposta maior que o previamente esperado de 20%, a amostra final foi de 8.952 indivíduos.

Devido às perdas mencionadas acima, alguns estratos ficaram sem representação amostral. Para contornar esse problema fez-se necessário a utilização de métodos de pós-estratificação para o ajuste dos pesos amostrais (Szwarcwald et al., 2019b). Essa técnica, já utilizada em outras pesquisas como a Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel) (Bernal et al., 2014, 2013), ajusta os pesos de modo que os totais populacionais para as variáveis utilizadas sejam iguais aos totais populacionais conhecidos, o que torna a amostra mais representativa da população (Lumley, 2011).

Com as medidas de colesterol total, HDL, LDL, hemoglobina glicosilada e creatinina obtidas através da subamostra de exames laboratoriais da PNS, foi possível encontrar, pela primeira vez no Brasil, valores de referência para marcadores biológicos, segundo raça/cor, idade e sexo (Szwarcwald et al., 2019a). A hemoglobina glicosilada, por exemplo, é um importante marcador de diabetes e pré-diabetes (Imbeault et al., 2003) enquanto que colesterol LDL elevado aumenta o risco de doenças cardiovasculares (Malta et al., 2019).

Ao ser observado a construção do delineamento da subamostra de exames laboratoriais da PNS 2013, essa pesquisa poderia ser vista como resultado de um delineamento

de amostragem em duas fases (Neyman, 1938), que consiste na obtenção do desfecho de interesse e das covariáveis de baixo custo numa primeira fase da pesquisa, disponíveis para todos os indivíduos na amostra e, em seguida, para a segunda fase, as informações são utilizadas para selecionar uma subamostra de indivíduos para medição de covariáveis de alto custo (Amorim et al., 2021). No entanto, dificuldades na coleta dos exames complementares laboratoriais resultaram na não divulgação dos pesos de delineamento em conjunto com os dados da subamostra. A única variável de peso fornecida é a variável de peso resultante do método de pós-estratificação (variável `peso_lab`).

Como alternativa, outras técnicas de calibração dos pesos amostrais poderiam ter sido empregadas nos dados da subamostra de exames laboratoriais da PNS, como por exemplo o *raking* (Deming e Stephan, 1940), cujo método já foi utilizado na comparação com a pós-estratificação nos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) (Ruiz e Silva, 2014). Em relação à PNS, nota-se na literatura uma ausência de trabalhos com teor crítico referente à metodologia de calibração empregado na pesquisa. Assim, tendo em vista a riqueza das informações coletadas através da PNS e a importância de se ter estimativas representativas da população para o fomento de políticas de saúde pública, os métodos de estimação devem ser empregados da melhor maneira possível.

Além disso, a Pesquisa Nacional de Saúde tem uma continuidade, cujo os microdados da pesquisa de 2019 já encontram-se disponíveis e, apesar de não ter sido coletado uma subamostra para realização de exames laboratoriais complementares na PNS de 2019, essa discussão ainda ganha relevância, não somente para os levantamentos amostrais subsequentes da PNS, mas também para outros levantamentos amostrais. Sendo assim, em meio a diferentes técnicas disponíveis na literatura para lidar com a estimação dos pesos amostrais, é de suma importância entender as vantagens e limitações que cada método carrega consigo.

Portanto, sob este contexto, o que será apresentado nesta dissertação é a comparação dos métodos de calibração para os pesos amostrais com o já apresentado no conjunto de dados da PNS 2013 utilizando como métrica de comparação um conjunto de sete indicadores que auxiliam na avaliação de vários aspectos da calibração. Além disso, buscamos identificar as principais diferenças nas estimativas de proporção e total para as variáveis auxiliares, bem como também a prevalência de diabetes *mellitus* na população de acordo com algumas características de interesse.

2 Revisão da Literatura

Este capítulo apresenta uma breve revisão sobre a Pesquisa Nacional de Saúde de 2013, focando nos aspectos gerais da pesquisa, bem como o questionário, aspectos da amostragem, fatores de expansão, coleta dos dados, divulgação dos resultados e também a subamostra de exames laboratoriais da PNS. Posteriormente, a revisão de literatura foca na amostragem, com destaque para o estimador de Horvitz-Thompson, ajustes para não-resposta, calibração dos pesos, mais especificamente a pós-estratificação e o *raking* e também comparação e avaliação de desempenho da calibração.

2.1 A Pesquisa Nacional de Saúde de 2013

A Pesquisa Nacional de Saúde é uma pesquisa a nível nacional desenvolvida pelo IBGE e realizada em convênio com o Ministério da Saúde. Foi planejada sob um delineamento próprio para coleta de informações de saúde referentes a perguntas sobre a percepção do estado de saúde, estilo de vida e doenças crônicas, além de aferições antropométricas e fatores econômicos, sociodemográficos, culturais, comportamentais e ambientais que influenciam na saúde do indivíduo (Szwarcwald et al., 2019b).

A população investigada consiste em moradores de domicílios particulares do Brasil, desconsiderando as localidades em setores censitários especiais (quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos e hospitais). A amostra da PNS é uma subamostra da Amostra Mestra do Sistema Integrado de Pesquisas Domiciliares (SIPD) do IBGE (IBGE, 2014).

Na primeira etapa da PNS, foram selecionados 81.254 domicílios, dos quais 69.994 eram domicílios ocupados. Foram 64.348 domicílios incluídos na pesquisa e, por fim, 60.202 indivíduos adultos realizaram a entrevista (Szwarcwald et al., 2019b). A Figura 1 abaixo apresenta o diagrama contendo o tamanho amostral de todas as fases das etapas da PNS 2013. A segunda etapa da pesquisa é abordada na [subseção 2.1.1](#).

Questionário

De acordo com o desenho da amostra (Souza-Júnior et al., 2015), o questionário da pesquisa é dividido em três partes: domiciliar, moradores do domicílio e morador selecionado. A primeira parte, que diz respeito a questões domiciliares, foi respondida pelo responsável do domicílio ou por qualquer morador no momento da entrevista. Foi

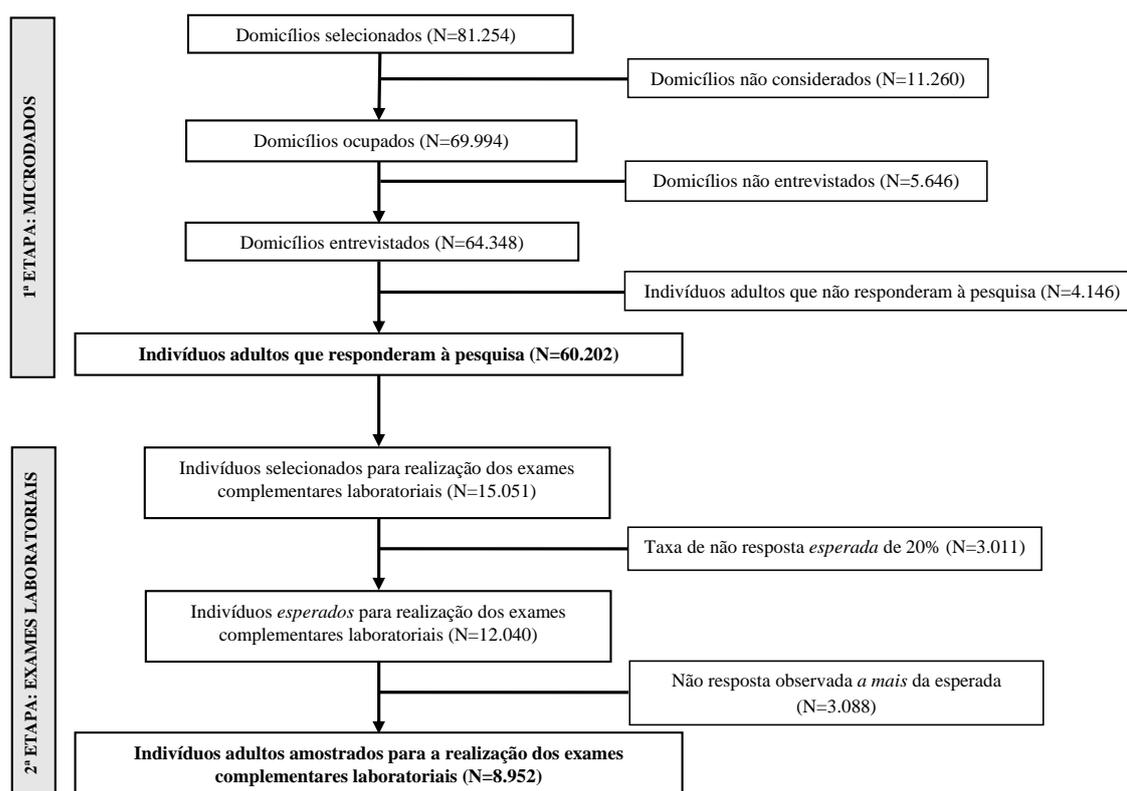


Figura 1 – Diagrama de fluxo das etapas da Pesquisa Nacional de Saúde, 2013.

perguntado questões domiciliares referente a informações do domicílio e visitas da Equipe de Saúde da Família e agentes de endemias.

A segunda parte aborda nível de educação, trabalho, rendimento, deficiências, cobertura de plano de saúde, utilização dos serviços de saúde, saúde do idoso, cobertura de mamografia e características de crianças menores de 2 anos de idade, que é respondida por todos os moradores do domicílio. A pessoa responsável poderia responder o questionário por outro morador, caso este estivesse ausente ou impossibilitado de responder.

Na terceira parte somente o morador adulto, selecionado aleatoriamente, deve responder ao questionário. As questões dizem respeito a características de trabalho e apoio social, percepção do estado de saúde, acidentes e violências, estilos de vida, doenças crônicas, saúde da mulher, atendimento pré-natal, saúde bucal e atendimento médico.

Aspectos da Amostragem

O plano amostral definido para a PNS de 2013 foi a amostragem por conglomerados em três estágios de seleção, com estratificação das unidades primárias de amostragem (UPA). Mantendo a estratificação da Amostra Mestra da PNAD, o primeiro estágio

consistiu na realização de uma amostragem aleatória simples (AAS) para seleção das UPA (setores censitários ou composição de setores). Posteriormente, para o segundo estágio, por meio da atualização mais recente do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE), um número fixo de domicílios particulares permanentes foi selecionado, de forma aleatória, em cada UPA selecionada no estágio anterior. Por fim, no terceiro estágio, a partir de uma lista de moradores elegíveis construída no momento da entrevista dentro de cada domicílio da amostra, da residência maior de 18 anos foi selecionado, com equiprobabilidade, para responder à entrevista individual (Szwarcwald et al., 2019b; FIOCRUZ, 2022; IBGE, 2014).

O tamanho da amostra da PNS considerou a estimação de proporções com nível de confiança de 95%, efeito do plano amostral, número de domicílios selecionados por UPA e proporção de pessoas adultas com 18 anos ou mais de idade (Szwarcwald et al., 2014). Os cálculos iniciais do tamanho da amostra foram baseados em AAS, fixando os níveis de precisão desejados para cada domínio de divulgação, que são os níveis geográficos (Brasil, Grande Região, Unidade Federativa (UF), Região Metropolitana, Capital e restante da UF). A fórmula de cada domínio é dada por (IBGE, 2014):

$$n_{AAS} = \frac{N}{N-1} \times P \times Q \times \frac{1}{CV^2 \times P^2 + \frac{P \times Q}{N-1}}, \quad (2.1)$$

em que N é o tamanho populacional, CV é o coeficiente de variação desejado da estimativa de proporção, P a proporção de interesse e Q seu complementar.

Como o plano amostral adotado na pesquisa foi o de amostragem por conglomerados, foi necessário fazer um ajuste no tamanho amostral, de modo a considerar o efeito de plano amostral (EPA), que foram estimados com base no suplemento saúde da PNAD de 2008 (IBGE, 2014), portanto:

$$n_{AC} = n_{AAS} \times \text{EPA}, \quad (2.2)$$

sendo que n_{AAS} é o tamanho de amostra de pessoas sob amostragem aleatória simples apresentado anteriormente.

Ao estimar o tamanho mínimo da amostra por UF em 1.800, possibilitou-se estimar 9 dos indicadores com a precisão desejada. O restante dos indicadores, apesar de possuir precisão inferior a desejada para as estimativas, encontram-se dentro das faixas consideradas aceitáveis (IBGE, 2014). Espera-se, no mínimo, dois domínios de divulgação por UF, já que nem todos possuem região metropolitana e, portanto, possuem apenas capital e restante da UF como domínios. Com isto, determinou-se que o tamanho mínimo da amostra fosse de 900 domicílios por domínio (Souza-Júnior et al., 2015).

Foi fixado em 10 o número de domicílios selecionados em cada UPA, com exceção dos domínios em que haviam limitações no tamanho da amostra de UPA pela Amostra Mestra. Nestes casos, o número de domicílios foi de 14 (IBGE, 2014). Para incorporar a

taxa de não resposta e a taxa de má classificação das unidades no cadastro de seleção, considerou-se aproximadamente 23% a mais de domicílios selecionados, em relação à amostra mínima (Souza-Júnior et al., 2015).

Fatores de Expansão

Foram definidos fatores de expansão para as UPAs, para os domicílios e todos seus moradores e para o morador selecionado (IBGE, 2014).

Peso das unidades primárias de amostragem

Para os pesos amostrais das UPAs foram considerados o inverso das probabilidades de seleção da UPA para a Amostra Mestra e as de seleção para a amostra da PNS (IBGE, 2014). O peso básico da UPA i do estrato h da PNS é dado por (Souza-Júnior et al., 2015):

$$W_{hi} = \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}}, \quad (2.3)$$

onde m_h é o número de UPAs selecionadas no estrato h para a Amostra Mestra, N_{hi} é o número de domicílios particulares permanentes ocupados, ocupados sem entrevistas realizadas (equivalentes aos domicílios fechados) e vagos na UPA i do estrato h^1 , N_h é o número de domicílios particulares permanentes ocupados, ocupados sem entrevistas realizadas (equivalentes aos domicílios fechados) e vagos no estrato h^1 e m_h^{PNS} é o número de UPAs selecionadas no estrato h para a PNS.

Peso dos domicílios e todos seus moradores

Para os pesos dos domicílios e todos seus moradores levou-se em conta o peso da UPA, ajustes para correção de não resposta e também calibração das estimativas com os totais populacionais estimados pela COPIS do IBGE (IBGE, 2014). Como os domicílios foram selecionados com equiprobabilidade dentro de cada UPA, o peso de seleção do domicílio j dentro da UPA i e estrato h é definido como (Souza-Júnior et al., 2015):

$$W_{j|hi} = \frac{N_{hi}^*}{n_{hi}}, \quad (2.4)$$

em que N_{hi}^* é o número de domicílios particulares permanentes ocupados e fechados na UPA i do estrato h^2 e n_{hi} é o número de domicílios selecionados na UPA i do estrato h .

¹ Dados atualizados do CNEFE no momento da seleção das UPA para Amostra Mestra.

² Dados da última atualização do CNEFE no momento da seleção dos domicílios

Com isto, o peso básico do domicílio é obtido pela multiplicação dos pesos da Equação 2.3 e Equação 2.4, ou seja (Souza-Júnior et al., 2015):

$$\begin{aligned} W_{hij} &= W_{hi} \times W_{j|hi} \\ &= \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}} \times \frac{N_{hi}^*}{n_{hi}} \end{aligned} \quad (2.5)$$

O peso básico da Equação 2.5 é multiplicado pela razão entre n_{hi}^* , que é o número de domicílios selecionados ocupados (com morador) na UPA i do estrato h e n_{hi}^{**} , que é o número de domicílios selecionados e ocupados com entrevista realizada na UPA i do estrato h . Essa razão é um ajuste feito nos pesos para que seja compensado as perdas por não resposta, seja por recusa, impossibilidade de contato com o morador ou outro motivo para perda (Souza-Júnior et al., 2015).

$$W_{hij}^* = W_{hij} \times \frac{n_{hi}^*}{n_{hi}^{**}}. \quad (2.6)$$

Um novo ajuste é feito no peso, de modo que a Equação 2.6 também recebe um fator multiplicador, denominado de calibração ou pós-estratificação, para que as estimativas dos totais populacionais dos níveis geográficos sejam iguais aos totais populacionais produzidos pela COPIS. Com isso, o peso amostral final do domicílio é dado por:

$$\begin{aligned} W_{hij}^{**} &= W_{hij}^* \times \frac{P_d^{tri}}{\hat{P}_d^{tri}} \\ &= \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}} \times \frac{N_{hi}^*}{n_{hi}} \times \frac{n_{hi}^*}{n_{hi}^{**}} \times \frac{P_d^{tri}}{\hat{P}_d^{tri}}, \end{aligned} \quad (2.7)$$

em que P_d^{tri} é o total populacional produzido pela COPIS para o nível geográfico d no dia 27 de julho de 2013 e \hat{P}_d^{tri} é a estimativa populacional obtida com os dados da pesquisa para o nível geográfico d (Souza-Júnior et al., 2015).

Peso do morador selecionado

O peso do morador selecionado considerou o peso do domicílio, probabilidade de seleção do morador, ajuste de não resposta por sexo e calibração pelos totais populacionais por sexo e faixa etária (18 a 24 anos, 25 a 39 anos, 40 a 59 anos e 60 anos e mais) estimados com peso de todos os moradores (IBGE, 2014).

Como a seleção morador se dá por AAS, então o peso é o produto entre o peso do domicílio da Equação 2.5 e o número de moradores elegíveis no domicílio (indivíduos adultos com idade igual ou superior à 18 anos). Desse modo, o peso básico sem ajuste é dado por (Souza-Júnior et al., 2015):

$$\begin{aligned} W_{hijk} &= W_{hij} \times O_{hij} \\ &= \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}} \times \frac{N_{hi}^*}{n_{hi}} \times O_{hij}, \end{aligned} \quad (2.8)$$

onde k é o índice do morador selecionado e O_{hij} é o número de moradores elegíveis no domicílio j na UPA i e estrato h .

Considerou-se também o ajuste de não resposta nos pesos dos moradores selecionados, portanto o peso do domicílio utilizado é o da [Equação 2.6](#), ou seja:

$$\begin{aligned} W'_{hijk} &= W_{hij}^* \times O_{hij} \\ &= \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}} \times \frac{N_{hi}^*}{n_{hi}} \times \frac{n_{hi}^*}{n_{hi}^{**}} \times O_{hij}. \end{aligned} \quad (2.9)$$

Como houve uma perda maior entre os homens em relação às mulheres, optou-se por um novo ajuste por sexo. Com isto, os pesos são ([Souza-Júnior et al., 2015](#)):

$$W_{hijk}^M = \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}} \times \frac{N_{hi}^*}{n_{hi}} \times \frac{n_{hi}^*}{n_{hi}^{**}} \times O_{hij} \times \frac{\sum_j^{n_{hijk}^{**}} W'_{hijk} \times \alpha_{hijk}^M}{\sum_i^{n_{hijk}^{***}} W'_{hijk} \times \alpha_{hijk}^M} \quad (2.10)$$

$$W_{hijk}^F = \frac{1}{m_h} \times \frac{N_h}{N_{hi}} \times \frac{m_h}{m_h^{PNS}} \times \frac{N_{hi}^*}{n_{hi}} \times \frac{n_{hi}^*}{n_{hi}^{**}} \times O_{hij} \times \frac{\sum_j^{n_{hijk}^{**}} W'_{hijk} \times \alpha_{hijk}^F}{\sum_i^{n_{hijk}^{***}} W'_{hijk} \times \alpha_{hijk}^F}, \quad (2.11)$$

em que n_{hijk}^{**} é o número de domicílios selecionados com entrevista realizada na UPA i do estrato h , n_{hijk}^{***} é o número de moradores selecionados com entrevista realizada na UPA i do estrato h , α_{hijk}^M indica se o morador selecionado no domicílio j da UPA i do estrato h é do sexo masculino e α_{hijk}^F indica se o morador selecionado no domicílio j da UPA i do estrato h é do sexo feminino ([Souza-Júnior et al., 2015](#)).

Os totais populacionais dos moradores selecionados, considerando sexo e faixa etária (18 a 24 anos, 25 a 39 anos, 40 a 59 anos e 60 anos ou mais) foram calibrados de maneira que correspondessem com os totais populacionais obtidos com os pesos dos domicílios, pois como a amostra dos moradores dos domicílios é maior, então possui uma estimativa mais precisa dos totais populacionais. O peso final do morador selecionado do sexo masculino é dado pela seguinte fórmula:

$$W_{hijk}^{M*} = W_{hijk}^M \times \frac{\hat{P}_{d,M,c}^M}{\hat{P}_{d,M,c}^S}, \quad (2.12)$$

onde $\hat{P}_{d,M,c}^M$ é a estimativa populacional obtida com os dados dos moradores dos domicílios da pesquisa para o nível geográfico d , do sexo masculino e faixa etária c , enquanto que $\hat{P}_{d,M,c}^S$ é a estimativa populacional obtida com os dados dos moradores selecionados da pesquisa para o nível geográfico d , do sexo masculino e faixa etária c . A fórmula dos pesos para o sexo feminino é análoga à fórmula apresentada na [Equação 2.12](#) ([Souza-Júnior et al., 2015](#)).

Coleta dos Dados

O IBGE foi responsável pela organização e coordenação do trabalho de campo. Os entrevistadores utilizaram PDAs para a realização das entrevistas, além destes terem sido treinados para aferir as medidas de peso, altura, circunferência da cintura e pressão arterial dos indivíduos (IBGE, 2016).

O agente de coleta entrou em contato com uma pessoa do domicílio selecionado, que descreveu ao morador o estudo, seus objetivos, procedimentos e a importância de participar da pesquisa, além de ter sido elaborada uma lista contendo todos os moradores adultos do respectivo domicílio. O morador adulto que responderia a entrevista individual foi selecionado através do programa de seleção aleatória no PDA e assim as entrevistas eram agendadas de modo que as datas e horários fossem convenientes para os entrevistados (IBGE, 2014).

Divulgação dos Resultados

Quatro volumes com os resultados obtidos através da Pesquisa Nacional de Saúde de 2013 foram divulgados:

1. percepção do estado de saúde, estilo de vida e doenças crônicas (IBGE, 2014),
2. acesso e utilização dos serviços de saúde, acidentes e violências (IBGE, 2015a),
3. ciclos de vida (IBGE, 2015b) e
4. indicadores de saúde e mercado de trabalho (IBGE, 2016).

Para facilitar a interpretação dos resultados, foram apresentados indicadores e intervalos de confiança, cujo os resultados são cruzamentos de recortes geográficos de grandes regiões, UF, municípios das capitais e DF, com características sociodemográficas de sexo, faixa etária, grau de escolaridade, raça/cor e condição em relação a força de trabalho. Devido a necessidade de análises mais criteriosas do MS ou por ter tamanho amostral insuficiente nas desagregações, alguns indicadores não foram divulgados (IBGE, 2014).

2.1.1 A subamostra de Exames Laboratoriais da PNS

Por ter alto custo, a coleta de material biológico foi feita em uma subamostra dos setores censitários pesquisados. Selecionou-se 25% do total de entrevistados para a realização dos exames laboratoriais, visando a uma taxa de não resposta de 20%, o que

consiste em, aproximadamente, 12 mil sujeitos como amostra final (Szwarcwald et al., 2019b). Para facilitar a logística, a seleção dos setores censitários foi feita estabelecendo probabilidade inversamente proporcional à dificuldade de coleta (Szwarcwald et al., 2014).

Para o morador selecionado, foi medido peso, altura, circunferência da cintura e pressão arterial. Os exames laboratoriais da PNS foram realizados por um consórcio de laboratórios privados que teve como responsabilidades: o agendamento dos exames com o morador, a orientação e preparação prévia adequada de cada exame, a entrega do frasco de urina, a coleta de sangue e urina, o transporte do material coletado, o fornecimento dos resultados para cada morador selecionado (seja por internet ou carta pessoal, no caso de indivíduos sem acesso a internet), transmissão dos resultados ao IBGE e sigilo das informações (IBGE, 2014).

Foi feita a coleta de sangue para realização de exames de hemoglobina glicada, colesterol total, LDL (método direto), HDL (método direto), hemograma, hemoglobina S e outras hemoglobinopatias, creatinina e sorologia de dengue. Os exames realizados com a coleta de urina foram os de: dosagem de sódio, dosagem de potássio e creatinina (IBGE, 2014).

A taxa amostral de não resposta observada na coleta de material laboratorial foi de 40%, o dobro da esperada, resultando em uma amostra de 8.952 adultos. Além disso, a distribuição das perdas foi não uniforme, culminando em estratos sem representação amostral, sendo necessária a utilização de um ajuste compensador de ponderação para corrigir a falta de representatividade dos grupos. Assim, através de um conjunto de variáveis auxiliares utiliza-se técnicas usando os totais populacionais conhecidos para ajustar os pesos amostrais, reduzindo o viés das estimativas que são causados pela não resposta (Szwarcwald et al., 2019b).

2.2 Amostragem

Lumley (2011) cita que a grande maioria dos desenvolvimentos matemáticos estatísticos são baseados-no-modelo, o que significa que para o processo aleatório que gera os dados, é especificado um modelo de probabilidade, como por exemplo, especificar que os dados seguem uma distribuição normal. Dessa forma, como o modelo representa o processo que gera os dados, então as conclusões tiradas podem ser generalizadas para outras situações em que se tem o mesmo processo.

Em contraste, as análises de pesquisas amostrais são usualmente baseadas-no-delineamento, ou seja, o pesquisador especifica uma população, cujo os dados possuem valores desconhecidos, mas considerados fixos. Como a amostra depende da seleção aleatória dos indivíduos da população, ela é aleatória. O delineamento amostral, processo de seleção

aleatória dos indivíduos, está sob o controle do pesquisador, portanto as probabilidades envolvidas podem, a princípio, serem conhecidas precisamente.

Essa análise tem como objetivo estimar características dessa população, então as inferências baseadas-no-delineamento não podem ser generalizadas para outras populações. Essas inferências têm como conceito estatístico fundamental a amostra probabilística ou aleatória. As propriedades necessárias para a amostragem baseada-no-delineamento são:

1. Cada indivíduo i na população deve ter probabilidade não nula de pertencer a amostra, ou seja, $\pi_i > 0$.
2. Essa probabilidade π_i precisa ser conhecida para cada indivíduo que pertencer a amostra.
3. Cada par de indivíduos (i, j) na amostra deve ter probabilidade não nula de pertencerem a amostra, ou seja, $\pi_{ij} > 0$.
4. Essa probabilidade do par de indivíduos, π_{ij} precisa ser conhecida para cada par que pertencer a amostra.

As primeiras duas condições são necessárias para se ter estimativas válidas para a população, enquanto que as duas últimas são necessárias para a acurácia das estimativas. Os pesos amostrais podem ser interpretados como o número de indivíduos na população representados pelo i -ésimo indivíduo da amostra (Levy et al., 2011).

O Estimador de Horvitz-Thompson

Seja X_i uma observação qualquer de uma variável X de um indivíduo i , define-se:

$$\check{X}_i = \frac{1}{\pi_i} X_i,$$

para denotar a *observação ponderada*, seguindo a notação de Lumley (2011).

Horvitz e Thompson (1952) propuseram uma técnica para tratar amostras selecionadas sem reposição quando são usadas probabilidades de seleção diferentes em uma população finita. Ao considerar um tamanho de amostra n , o estimador de Horvitz-Thompson para o total populacional e sua respectiva variância é

$$\hat{T}_X = \sum_{i=1}^n \frac{1}{\pi_i} X_i = \sum_{i=1}^n \check{X}_i \quad \widehat{\text{Var}}(\hat{T}_X) = \sum_{ij} \left(\frac{X_i X_j}{\pi_{ij}} - \frac{X_i}{\pi_i} \times \frac{X_j}{\pi_j} \right), \quad (2.13)$$

em que $1/\pi_i$ é o peso amostral, que diz respeito ao inverso da probabilidade do indivíduo i pertencer a amostra. Esse estimador é uma generalização para o total populacional,

portanto pode ser utilizado sob qualquer delineamento, com ou sem reposição. Além disso, ele é não viesado.

Mesmo que os indivíduos da amostra sejam selecionados com igual probabilidade, seus pesos amostrais podem variar. Isso ocorre devido as reponderações feitas para ajustar os dados com a finalidade de reduzir erros padrões e vieses das estimativas compensando pelos efeitos de não-resposta, não cobertura e instabilidade nos estimadores. Estes ajustes podem aumentar os pesos amostrais para alguns respondentes que possuem características similares com indivíduos que não responderam ou não tiveram a chance de serem selecionados para a pesquisa (Levy et al., 2011).

2.2.1 Ajustes para não-resposta

O propósito de grande parte das pesquisas é o de estimar os parâmetros populacionais com grande acurácia e confiabilidade (Levy et al., 2011), porém praticamente todas as pesquisas possuem algum nível de não resposta e seu efeito pode variar de acordo com cada pesquisa (Särndal et al., 1992). Se as estimativas forem baseadas somente nos respondentes, as inferências irão representar somente essa subpopulação e, mesmo que tecnicamente as estimativas possam ser não viciadas para essa subpopulação em específico, ainda assim o vício existirá em relação à população alvo. Portanto, terão níveis significativos de viés se não forem ajustadas para a não resposta, sendo que, quanto maior essa taxa, maior é o viés nas estimativas (Levy et al., 2011).

Devido ao aumento no uso de pesquisas por amostragem e a grande dificuldade de se obter altas taxas de resposta, esse problema vem recebendo uma atenção considerável, o que levou ao desenvolvimento de diversas técnicas para lidar com a não resposta (Levy et al., 2011). Dependendo do que se sabe sobre os respondentes, os procedimentos a serem aplicados podem ser simples ou mais elaborados (Valliant et al., 2013).

Essas técnicas têm como objetivo fazer ajustes de modo que os totais populacionais para as variáveis utilizadas sejam os mesmos dos totais populacionais conhecidos, fazendo a amostra ser mais representativa da população (Lumley, 2011). Existem diversos métodos de ponderação possíveis de aplicação para lidar com esses problemas, como por exemplo a pós-estratificação, *raking*, ponderação linear, estimadores de regressão generalizada (do inglês *Generalised Regression Estimator*, GREG), entre outros. Há também técnicas que englobam misturas destes métodos (Kalton e Flores-Cervantes, 2003).

2.2.2 Calibração dos Pesos

Variáveis auxiliares são variáveis que estão disponíveis a nível populacional, seja de forma individual ou agregada, e são empregadas com a finalidade de melhorar o plano

amostral ou aumentar a precisão da estimativa das variáveis de interesse (Lavrakas, 2008). De acordo com Valliant et al. (2013), a ideia geral da calibração é o de utilizar as variáveis auxiliares para melhorar a eficiência dos estimadores, cujos benefícios são a diminuição das variâncias, correção do viés de cobertura e outros erros de pesquisa e ajuste para não resposta.

Seguindo a explicação de Valliant et al. (2013), suponha que $\mathbf{d} = \{d_i\}_{i \in n}$ é o conjunto de pesos amostrais, chamado de pesos básicos. Para o calculo destes, é necessário ter-se as probabilidades de seleção π_i de cada indivíduo i , visto que o peso é dado pelo inverso dessa probabilidade. O objetivo da calibração é o de encontrar um conjunto de novos pesos $\mathbf{w} = \{w_i\}_{i \in n}$, próximo de \mathbf{d} que, ao ser utilizado para estimar os totais das variáveis auxiliares, reproduz exatamente os totais populacionais.

A ideia de estabelecer que \mathbf{w} seja próximo de \mathbf{d} tem o propósito de que assim o conjunto de \mathbf{w} terá aproximadamente as mesmas propriedades inerentes de \mathbf{d} , sejam estas propriedades boas ou ruins. Por exemplo, se os pesos \mathbf{d} estiverem associados a uma estimativa não viciada, então essa mesma estimativa de \mathbf{w} será aproximadamente não viciada. Da mesma forma que, se estes pesos \mathbf{d} produzirem estimativas com alta variância, então \mathbf{w} também irá possuir uma alta variância.

Portanto, a calibração consiste em achar um conjunto de pesos \mathbf{w} que minimizam uma medida de distância $G(\mathbf{w}, \mathbf{d})$ entre os dois pesos, sujeito a restrição

$$\sum_{i \in n} w_i \mathbf{x}_i = \sum_{k \in U} \mathbf{x}_k, \quad (2.14)$$

em que U é o conjunto de todos elementos da população, $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$ é o conjunto de p variáveis auxiliares para o indivíduo i e $w_i = d_i g_i$ é uma função em relação ao peso amostral e um ajuste (peso \mathbf{g}) que satisfaz a Equação 2.14. Uma escolha para a distância é a função de mínimos quadrados dada por

$$G(\mathbf{w}, \mathbf{d}) = \sum_n \frac{(w_i - d_i)^2}{d_i},$$

em que minimizar essa distância $G(\mathbf{w}, \mathbf{d})$, sujeita a restrição (2.14) leva ao GREG.

Em suma, tem-se que o total calibrado (referente a cada variável auxiliar) é igual ao total populacional. Para determinar os pesos finais \mathbf{w} é necessário ter os valores de \mathbf{x} para os indivíduos amostrados e os totais populacionais das variáveis auxiliares, que podem ser tanto quantitativas quanto qualitativas. A popularidade da calibração em aplicações reais deu-se pelo fato dela depender apenas de restrições naturais, além de fornecer estimativas fáceis de interpretar.

Além disso, estas são aproximadamente não viesadas. Contudo, alguns estimadores podem levar a pesos negativos, sendo necessário a utilização de pós-estratificação ou *raking*

para evitar a ocorrência desses pesos negativos (Amorim, 2014). Essas duas técnicas são bastante utilizadas em pesquisas domiciliares, em que as variáveis auxiliares são usualmente indicadores de grupos demográficos (Valliant et al., 2013).

2.2.2.1 Pós-Estratificação

Como citado por Lumley (2011), a pós-estratificação é a técnica de ajuste de pesos amostrais mais simples. Suponha que a população possa ser dividida em grupos chamados estratos, amostrando n_k indivíduos da população do estrato k que contém N_k indivíduos. Os pesos amostrais \mathbf{d} são definidos como $d_i = N_k/n_k$ e o estimador de Horvitz-Thompson \hat{N}_k do tamanho da população desse estrato é correta e seu erro padrão é zero. Como os estratos k são aplicados depois que a amostra é selecionada e os dados são coletados, são chamados também de pós-estratos (Valliant et al., 2013).

Utilizando a notação de Valliant et al. (2013), formalmente, o estimador de pós-estratificação para o total populacional é definido como:

$$\hat{t}_{yPS} = \sum_{k=1}^K N_k \left(\frac{\hat{t}_{yk}}{\hat{N}_k} \right), \quad (2.15)$$

em que $\hat{t}_{yk} = \sum_{i \in k} d_i y_i$ é o total estimado de y e \hat{N}_k é o tamanho populacional estimado, ambos referentes ao pós-estrato k e baseados nos pesos fornecidos. Já N_k representa o total populacional.

Com a pós-estratificação, os pesos \mathbf{d} são substituídos por \mathbf{w} , em que $g_i = N_k/\hat{N}_k$ é o ajuste de pós-estratificação para o pós-estrato k do indivíduo i , então:

$$w_i = d_i \frac{N_k}{\hat{N}_k}. \quad (2.16)$$

Portanto, a Equação 2.15 pode ser reescrita de forma que

$$\hat{t}_{yPS} = \sum_{i \in n} w_i y_i, \quad (2.17)$$

isto é, uma soma ponderada dos valores.

Se for utilizada mais de uma variável auxiliar na pós-estratificação, os grupos de pós-estratos devem ser construídos como uma classificação cruzada completa entre as variáveis, o que pode se tornar indesejável conforme o número de variáveis auxiliares ou categorias aumentam, porque essa classificação cruzada pode resultar em grupos muito distintos, aumentando o risco de alguns destes não serem amostrados, ou então de não haver os totais populacionais destes pós-estratos (Lumley, 2011).

Como exemplo, suponha uma pesquisa domiciliar em que possa ser interessante utilizar as variáveis de raça/cor (categorias: preto, pardo, branco, indígena e outros), nível

de escolaridade (categorias: ensino fundamental incompleto ou menos, ensino fundamental completo ou ensino médio incompleto, ensino médio completo ou superior incompleto e ensino superior completo) e gênero (categorias: masculino ou feminino), como pós-estratos. Então, ao todo, serão 40 grupos, pois $k = 5 \times 4 \times 2 = 40$, o que significa ter os totais populacionais para essas 40 classificações cruzadas das variáveis auxiliares, além de ter indivíduos amostrados dentro de cada pós-estrato.

Se não houver nenhum indivíduo amostrado do grupo k , então não será possível realizar a reponderação, sendo necessária a mescla de dois estratos ou então uma correção *ad hoc* (Lumley, 2011).

2.2.2.2 Raking

Tanto o *raking* quanto a pós-estratificação são estimadores de calibração bastante utilizados, sendo frequentemente prevalentes em pesquisas domiciliares em que as variáveis auxiliares são variáveis demográficas (Valliant et al., 2013), porém o *raking* não necessita que se faça uma classificação cruzada completa entre todas as variáveis auxiliares.

O processo consiste na pós-estratificação de cada conjunto de variáveis por vez, repetindo esse processo iterativamente até a convergência dos pesos. O *raking* também pode ser aplicado quando se tem disponível somente algumas classificações cruzadas das variáveis auxiliares (Lumley, 2011).

Abaixo é apresentado os sete passos para o algoritmo do *raking*. O algoritmo consiste em dois ciclos: externo e interno. O ciclo externo verifica o critério de convergência, enquanto que o ciclo interno realiza as iterações sob as variáveis auxiliares. A notação $w_j^{k,\nu}$ indica o peso para a j -ésima unidade amostral, na iteração k (de um total de K iterações) após pós-estratificar a ν -ésima variável auxiliar (de um total de p variáveis auxiliares) (Kolenikov, 2014).

1. Considere $k \leftarrow 0$ e os pesos iniciais ($w_j^{0,p}$) como sendo os pesos básicos (w_{1j}) para inicialização do algoritmo.
2. Atualize o valor da iteração para $k \leftarrow k + 1$ e use o resultado da iteração anterior ($w_j^{k-1,p}$) para inicializar os pesos para a iteração atual ($w_j^{k,0}$).
3. Repasse as variáveis auxiliares $\nu = 1, \dots, p$ e atualize os pesos de acordo com

$$w_j^{k,\nu} = \begin{cases} \frac{w_j^{k,\nu-1} T_{X_\nu}}{\sum_{l \in S} w_l^{k,\nu-1} x_{\nu,l}}, & x_{\nu,l} \neq 0 \\ w_j^{k,\nu-1}, & x_{\nu,l} = 0 \end{cases}$$

4. Se as diferenças entre os totais ponderados ($\sum_{j \in S} w_j^{k,p} x_\nu$) e os totais populacionais (T_{X_ν}) forem menores do que um valor pré-especificado de tolerância (tol) para

todas as variáveis auxiliares $\nu = 1, \dots, p$, então declare que o algoritmo atingiu a convergência e vá para o passo 7.

5. Se o número de iterações k atingir o valor pré-definido de K iterações especificado inicialmente, então declare que o algoritmo não atingiu a convergência e vá para o passo 7.
6. Caso contrário, retorne ao passo 2.
7. Retorne os últimos pesos calculados ($w_j^{k,p}$) como os pesos calibrados.

Posteriormente, o [Algoritmo 1](#) fornece o pseudo-código para o algoritmo. Como informação de entrada é preciso determinar o número máximo de iterações K e a tolerância tol desejada. O sobrescrito 0 é usado apenas para consistência com a notação.

Algoritmo 1 Pseudo-código do procedimento *Raking*

Input: K, tol

$k \leftarrow 0$

while ($k \leq K$ & $|\sum_{j \in S} w_j^{k,p} x_\nu - T_{X_\nu}| > tol$) **do**

$k \leftarrow k + 1$

for ν in $1 : p$ **do**

$w_j^{k,0} \leftarrow w_j^{k-1,p}$

$$w_j^{k,\nu} = \begin{cases} \frac{w_j^{k,\nu-1} T_{X_\nu}}{\sum_{l \in S} w_l^{k,\nu-1} x_{\nu,l}}, & x_{\nu,l} \neq 0 \\ w_j^{k,\nu-1}, & x_{\nu,l} = 0 \end{cases}$$

end for

if ($k > K$) **then**

return “Algoritmo não atingiu a convergência”

else if ($|\sum_{j \in S} w_j^{k,p} x_\nu - T_{X_\nu}| < tol$) **then**

return “Algoritmo atingiu a convergência”

end if

end while

return $w_j^{k,p}$

Em [Deming e Stephan \(1940\)](#) são citados diversos motivos do porque o raking é de interesse prático. É relativamente mais fácil encontrar os totais populacionais marginais, do que é encontrar os totais populacionais das classificações cruzadas entre as variáveis auxiliares. Além do mais, estes totais marginais podem ser provenientes de diferentes pesquisas. Outro motivo, que já foi mencionado anteriormente, é que alguns grupos criados a partir do cruzamento dos pós-estratos podem ter poucas unidades amostradas, ou até mesmo nenhuma, o que faz com que a [Equação 2.17](#) seja indefinida ou até mesmo instável. Portanto, mesmo que viável, a pós-estratificação é preterida em favor de uma calibração mais confiável baseada nas marginais.

2.2.3 Comparação dos Métodos de Calibração

Em seu livro, Lumley (2011) faz uma breve comparação entre os métodos de pós-estratificação e *raking* para verificar possíveis diferenças entre eles. No exemplo apresentado, o erro padrão pelo *raking* foi menor, em comparação com a pós-estratificação. Entretanto, o benefício desse procedimento diminui para pequenas subpopulações porque a informação completa é menos relevante para esses subgrupos menores. As informações das variáveis auxiliares só fornecem extra precisão se essas informações auxiliares dizem respeito à essas subpopulações.

Na literatura existem alguns estudos de comparação de métodos de calibração em cenários distintos. Em Tu (2015), as técnicas de pós-estratificação e *raking* são comparadas ao estimar o uso de internet em uma amostra telefônica de dados coletados em 2008. Adicionalmente à esses dois métodos, o autor considerou na estimativa o uso de *Propensity Score Adjustment* (PSA) e concluiu que, além da utilização de PSA obter maior acurácia do que outras estimativas, o método com o *raking* obteve maior redução no vício, em comparação com a pós-estratificação.

Em Djerf (1997), é examinado o método de pós-estratificação utilizado para calcular as estimativas do *Labor Force Survey* (LFS), um inquérito sobre mercado de trabalho e o compara com um método alternativo de *raking*. A conclusão apresentada pelo autor é de que o procedimento de estimação utilizado na LFS seja substituído pelo *raking*, que se mostrou mais eficiente.

No contexto de análises com dados do Brasil, pode-se citar o artigo de Bernal et al. (2014), que apresentou a metodologia utilizada na construção dos pesos do Vigitel de 2012 e uma comparação da tendência de indicadores estimados pelos métodos de pós-estratificação e *raking*. Em relação aos resultados, para ambos os métodos não houve diferença nas tendências temporais dos indicadores estimados, porém houve algumas diferenças nas estimativas de prevalências. Por exemplo, as prevalências de fumantes foram inferiores e as de consumo de fruta, verduras e legumes foram superiores pelo método de *raking*, em comparação com a pós-estratificação. Para a precisão das estimativas, o *raking* foi mais preciso, apesar de ter apresentado vício de cobertura de cadastro.

Outra pesquisa brasileira utilizada para a comparação dos dois métodos foi a PNAD, em Ruiz e Silva (2014). Com o método de pós-estratificação observou-se viés de disponibilidade, com alguns grupos de pessoas amostradas sendo sub e sobre-representados nos dados. Além disso, os coeficientes de variação dos indicadores foram, em geral, maiores. Sub e sobre-representações nas estimativas dos totais populacionais não foram observadas no *raking*, visto a forma que o método opera. Quando calculado os indicadores de desempenho da calibração pelos dois métodos, tanto a média do coeficiente de variação do estimador de Horvitz–Thompson do total da variável auxiliar e a eficiência média das estimativas

de calibração foram menores para o *raking*, enquanto que o coeficiente de variação dos pesos \mathbf{g} e a distância qui-quadrado entre os pesos calibrados e os pesos do desenho foram menores para a pós-estratificação.

2.2.4 Avaliação do desempenho da calibração

Em [Silva \(2004\)](#) é sugerido o uso de sete medidas para verificar se os objetivos da calibração foram alcançados. Através do uso dessas medidas é possível avaliar diferentes aspectos associados aos estimadores de calibração, por exemplo, a quantidade de erro de calibração restante para o conjunto completo de variáveis auxiliares e o viés do estimador. Aspectos da distribuição dos pesos calibrados, como a proporção de pesos extremos, o coeficiente de variação dos pesos calibrados e a distância dos pesos calibrados para os pesos básicos também são características avaliadas pelas sete medidas.

A média do erro relativo absoluto de calibração para os totais populacionais estimados das variáveis auxiliares \mathbf{x} é a primeira medida, denotada por M_1 :

$$M_1 = \frac{1}{J} \sum_{j=1}^J \frac{|\hat{t}_{x_j, C} - t_{x_j}|}{t_{x_j}},$$

em que t_{x_j} é o total populacional da variável auxiliar x_j e $\hat{t}_{x_j, C}$ a sua estimativa proveniente do método de calibração. O valor ideal para M_1 é zero, que indica que o erro de calibração foi totalmente eliminado. Valores diferentes de zero podem ocorrer dependendo da abordagem de calibração adotada, ou se alguma das variáveis auxiliares forem descartadas durante o processo de calibração, conforme apresentado em [Silva \(2004\)](#).

A segunda medida, a média dos coeficientes de variação dos totais das variáveis auxiliares estimados pelo estimador de expansão (2.13), é importante para verificar se há necessidade de preocupação em relação ao vício com o estimador de calibração.

$$M_2 = \frac{1}{J} \sum_{j=1}^J \frac{[\widehat{\text{Var}}(\hat{t}_{x_j, \pi})]^{1/2}}{t_{x_j}},$$

em que $\widehat{\text{Var}}(\hat{t}_{x_j, \pi})$ é a estimativa da variância do estimador de Horvitz-Thompson do total da variável auxiliar x_j , apresentada também em (2.13).

As próximas quatro medidas de interesse dizem respeito à distribuição dos pesos g_k^3 , em que g_k é o peso \mathbf{g} para a k -ésima unidade. A terceira e quarta medida representam a proporção de pesos extremos em relação a um valor L e U , ambos pré-especificados:

$$M_3 = \frac{1}{n} \sum_{k \in S} I\{g_k < L\} \quad \text{e} \quad M_4 = \frac{1}{n} \sum_{k \in S} I\{g_k > U\},$$

³ $g_k = \tilde{w}_k/d_k$, ou seja, a razão entre o peso calibrado \tilde{w}_k e o peso básico d_k

em que $I\{\cdot\}$ é a função indicadora. Para M_3 e M_4 , pesos \mathbf{g} menores do que $1/3$ e maiores do que 3 são considerados extremos, respectivamente (Deville et al., 1993).

A quinta medida para avaliação de desempenho é o coeficiente de variação dos pesos \mathbf{g} :

$$M_5 = \left[\sqrt{\frac{\sum_{k \in S} (g_k - \bar{g})^2}{(n-1)}} \right] / \bar{g},$$

em que $\bar{g} = \frac{1}{n} \sum_{k \in S} g_k$ é a média dos pesos \mathbf{g} .

A distância qui-quadrado entre os pesos calibrados \tilde{w}_k e pesos básicos do delineamento d_k é definida como a sexta medida:

$$M_6 = \frac{1}{n} \sum_{k \in S} \frac{(\tilde{w}_k - d_k)^2}{d_k}.$$

O cenário ideal é aquele em que os pesos \tilde{w}_k não sejam tão distantes dos pesos d_k , pois o estimador de expansão é não enviesado. Portanto, quanto menor o valor da medida M_6 , melhor deve ser o desempenho do método de calibração.

A sétima e última medida é a eficiência média das estimativas de calibração para um conjunto especificado de variáveis de interesse $\mathbf{y} = (y_1, \dots, y_R)$. Este indicador é utilizado para acessar os ganhos da calibração em comparação com as estimativas obtidas pelo estimador de expansão:

$$M_7 = \frac{1}{R} \sum_{r=1}^R \frac{\widehat{\text{Var}}(\hat{t}_{y_r, C})}{\widehat{\text{Var}}(\hat{t}_{y_r, \pi})}.$$

A variância do numerador dos termos da soma é dada por

$$\widehat{\text{Var}}(\hat{t}_{y_r, C}) = \sum_{k \in S} \sum_{i \in S} \left(\frac{1 - \pi_k \pi_i}{\pi_{ki}} \right) (g_k d_k e_{ki})(g_i d_i e_{ij}), \quad (2.18)$$

em que $e_{ki} = y_{ki} - \mathbf{x}'_k \hat{\mathbf{B}}_j$. A variância do denominador é estimada usando a estimativa da variância do estimador de expansão (2.13), a mesma utilizada em M_2 .

3 Objetivos

3.1 Objetivo geral

O objetivo geral é o de investigar se os dados da subamostra da PNS de 2013 são utilizados da forma mais adequada, de acordo com o que se tem disponível na literatura sobre calibração dos pesos amostrais como ajuste de não-resposta.

3.2 Objetivos específicos

- Propor métodos alternativos de calibração dos pesos com base nos métodos de pós-estratificação e *raking* para os dados da subamostra de Exames Laboratoriais da PNS 2013;
- Comparar os métodos de calibração para os pesos amostrais com o já apresentado no conjunto de dados da PNS 2013 utilizando como métrica de comparação um conjunto de sete indicadores que auxiliam na avaliação de vários aspectos da calibração.
- Avaliar as estimativas de certas características populacionais de interesse a partir dos três métodos de calibração de pesos.

Referências

- Amorim, G., Tao, R., Lotspeich, S., Shaw, P. A., Lumley, T., e Shepherd, B. E. (2021). Two-phase sampling designs for data validation in settings with covariate measurement error and continuous outcome. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 184(4):1368–1389. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/rssa.12689>.
- Amorim, G. G. C. (2014). *Semiparametric methods for multiphase response-selective samples*. PhD thesis, The University of Auckland.
- Bernal, R. T. I., Malta, D. C., Araujo, T. S., e Silva, N. N. (2013). Inquérito por telefone: pesos de pos-estratificação para corrigir vícios de baixa cobertura em rio branco, AC. *Revista de Saúde Pública*, pages 316–325.
- Bernal, R. T. I., Malta, D. C., Morais Neto, O. L., Claro, R. M., Mendça, B. C. A., Oliveira, A. C. C., e Silva, N. N. (2014). Vigitel - aracaju, sergipe, 2008: the effects of post-stratification adjustments in correcting biases due to the small amount of households with a landline telephone. *Revista Brasileira de Epidemiologia*, 17(1):163–174.
- Deming, W. E. e Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4):427–444. Publisher: Institute of Mathematical Statistics.
- Deville, J.-C., Sarndal, C.-E., e Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, 88(423):1013–1020. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].
- Djerf, K. (1997). Effects of post-stratification on the estimates of the finnish labour force survey. *Journal of Official Statistics*, 13(1):29–39.
- FIOCRUZ (2022). Delineamento da PNS. <https://www.pns.icict.fiocruz.br/delineamento-da-pns/>. Acesso em: 26 de nov. 2022.
- Horvitz, D. G. e Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].
- IBGE (2014). Pesquisa nacional de saúde : 2013 : percepção do estado de saúde, estilos de vida e doenças crônicas : Brasil, grandes regiões e unidades da federação.
- IBGE (2015a). Pesquisa nacional de saúde : 2013 : acesso e utilização dos serviços de saúde, acidentes e violências : Brasil, grandes regiões e unidades da federação.

- IBGE (2015b). Pesquisa nacional de saúde : 2013 : ciclos de vida : Brasil e grandes regiões.
- IBGE (2016). Pesquisa nacional de saúde : 2013 : indicadores de saúde e mercado de trabalho : Brasil e grandes regiões.
- Imbeault, P., Prins, J. B., Stolic, M., Russell, A. W., O'Moore-Sullivan, T., Després, J.-P., Bouchard, C., e Tremblay, A. (2003). Aging per se does not influence glucose homeostasis: in vivo and in vitro evidence. *Diabetes Care*, 26(2):480–484.
- Júnior, A. F. J., Colares, G. C., Filho, I. B. M. R., e Silva, L. S. (2020). Doenças crônicas não transmissíveis na infância: Revisão integrativa de hipertensão arterial sistêmica, diabetes mellitus e obesidade. *Revista Saúde Dinâmica*, 2(2):38–56. Number: 2.
- Kalton, G. e Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19(2):81–97.
- Kolenikov, S. (2014). Calibrating survey data using iterative proportional fitting (raking). *The Stata Journal*, 14(1):22–59. Publisher: SAGE Publications.
- Lavrakas, P. J. (2008). *Encyclopedia of Survey Research Methods (2 Volume Set)*. Sage Publications, Inc, 1 edition.
- Levy, P., Lemeshow, S., e Ferketich, A. (2011). *Sampling of Populations: Methods and Applications*, volume 42. Wiley, 4 edition.
- Lumley, T. (2011). *Complex Surveys: A Guide to Analysis Using R*. Wiley Series in Survey Methodology. Wiley.
- Malta, D. C., Andrade, S. S. C. A., Oliveira, T. P., Moura, L., Prado, R. R., e Souza, M. F. M. (2019). Probabilidade de morte prematura por doenças crônicas não transmissíveis, brasil e regiões, projeções para 2025. *Revista Brasileira de Epidemiologia*, 22:e190030.
- Malta, D. C., Morais Neto, O. L., e Silva Junior, J. B. (2011). Apresentação do plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis no brasil, 2011 a 2022. *Epidemiologia e Serviços de Saúde*, 20(4):425–438. Publisher: Coordenação-Geral de Desenvolvimento da Epidemiologia em Serviços / Secretaria de Vigilância em Saúde / Ministério da Saúde.
- Malta, D. C., Silva, M. M. A., Moura, L., e Morais Neto, O. L. (2017). A implantação do sistema de vigilância de doenças crônicas não transmissíveis no brasil, 2003 a 2015: alcances e desafios. *Revista Brasileira de Epidemiologia*, 20(4):661–675.
- Malta, D. C., Stopa, S. R., Szwarcwald, C. L., Gomes, N. L., Silva Júnior, J. B., e Reis, A. A. C. d. (2015). A vigilância e o monitoramento das principais doenças crônicas não transmissíveis no brasil - pesquisa nacional de saúde, 2013. *Revista Brasileira de Epidemiologia*, 18:3–16.

- Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, 33(201):101–116. Publisher: [American Statistical Association, Taylor & Francis, Ltd.].
- Ruiz, C. M. M. e Silva, P. L. N. (2014). Explorando alternativas para a calibração dos pesos amostrais da pesquisa nacional por amostra de domicílios. *VI Congreso de la Asociación Latinoamericana de Población*.
- Silva, P. L. N. (2004). *Calibration estimation : when and why, how much and how*. IBGE.
- Souza-Júnior, P. R. B. d., Freitas, M. P. S., Antonaci, G. A., e Szwarcwald, C. L. (2015). Desenho da amostra da pesquisa nacional de saúde 2013. *Epidemiologia e Serviços de Saúde*, 24(2):207–216.
- Szwarcwald, C. L., Malta, D. C., Pereira, C. A., Figueiredo, A. W., Almeida, W. S., Machado, I. E., Bacal, N. S., Silva, A. G., Silva Júnior, J. B., e Rosenfeld, L. G. (2019a). Valores de referência para exames laboratoriais de colesterol, hemoglobina glicosilada e creatinina da população adulta brasileira. *Revista Brasileira de Epidemiologia*, 22:e190002.supl.2.
- Szwarcwald, C. L., Malta, D. C., Pereira, C. A., Vieira, M. L. F. P., Conde, W. L., Souza Júnior, P. R. B., Damacena, G. N., Azevedo, L. O., Silva, G. A., Theme Filha, M. M., Lopes, C. S., Romero, D. E., Almeida, W. S., e Monteiro, C. A. (2014). Pesquisa nacional de saúde no brasil: concepção e metodologia de aplicação. *Ciência & Saúde Coletiva*, 19(2):333–342.
- Szwarcwald, C. L., Malta, D. C., Souza Júnior, P. R. B., Almeida, W. S., Damacena, G. N., Pereira, C. A., e Rosenfeld, L. G. (2019b). Exames laboratoriais da pesquisa nacional de saúde: metodologia de amostragem, coleta e análise dos dados. *Revista Brasileira de Epidemiologia*, 22:E190004.SUPL.2.
- Särndal, C., Swensson, B., e Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer series in statistics. Springer-Verlag.
- Tavares, N. U. L., Costa, K. S., Mengue, S. S., Vieira, M. L. F. P., Malta, D. C., e Silva Júnior, J. B. (2015). Uso de medicamentos para tratamento de doenças crônicas não transmissíveis no brasil: resultados da pesquisa nacional de saúde, 2013. *Epidemiologia e Serviços de Saúde*, 24(2):315–323.
- Tu, S.-H. (2015). A comparison of propensity score sub-classification and other calibration methods based on a telephone sample to estimate internet usage. *Taiwan She Hui Xue Kan*, (56):115–150.
- Valliant, R., Dever, J., e Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples*, volume 51. Springer.

4 Artigo

COMPARAÇÃO DE MÉTODOS DE CALIBRAÇÃO NA ANÁLISE DE DADOS DA PESQUISA NACIONAL DE SAÚDE DE 2013

Juliana Sena de Souza, Márcia Helena Barbian, Rodrigo Citton P. dos Reis

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

A ser enviado a

Brazilian Journal of Probability and Statistics

5 Conclusões e Considerações Finais

O estudo identificou que na subamostra de exames laboratoriais da PNS de 2013, quando observados todos os pós-estratos definidos pela classificação cruzada completa entre as cinco variáveis auxiliares definidas na pós-estratificação, nem todos os pós-estratos possuíam representação amostral na subamostra. Ou seja, quando olhamos todas as possíveis combinações entre as categorias de todas as variáveis auxiliares, observamos que nem todas as combinações possíveis estão sendo representadas na subamostra de exames.

Uma consequência direta dessa falta de representação é que, quando estimados os pesos amostrais pelo método de pós-estratificação e estes são utilizados na estimação de totais e proporções populacionais das variáveis auxiliares, é encontrada uma sub-representação em praticamente todas as categorias. A maior sub-representação encontrada foi para indivíduos de raça/cor outra e, como consequência, a proporção de indivíduos estimada neste grupo, com o método de pós-estratificação, subestimou a verdadeira proporção populacional. Esse resultado já era esperado, visto que a maioria dos pós-estratos sem representação amostral eram justamente aqueles em que a raça/cor do indivíduo era outra.

Portanto, mesmo que para grande parte das variáveis auxiliares haja uma boa representação populacional, não é para todas. A própria distribuição relativa da população pelas variáveis auxiliares traz evidências dessa sub-representação. Ao calcular o ER da pós-estratificação, observou-se que para a raça/cor outra, o ER foi de -32,04%. A estimativa do total de indivíduos maiores de 18 anos nessa respectiva categoria foi aproximadamente 600 mil indivíduos a menos do que o encontrado na projeção populacional.

Posteriormente, ao comparar os indicadores de desempenho da calibração por pós-estratificação e por *raking*, os resultados sugerem que, no geral, o *raking* performa melhor. De um total de 7 indicadores, apenas para a eficiência média das estimativas de calibração que a pós-estratificação obteve melhor resultado, o que já era esperado, visto que as variâncias das estimativas utilizando o *raking* eram um pouco superiores, em comparação com a do outro método.

A grande discrepância em termos de estimativas é encontrada quando observa-se os totais estimados utilizando os três métodos de calibração. O peso laboratorial estimou em 8.952 a quantidade de adultos no país quando, na verdade, de acordo com a projeção populacional, o total é de 145.572.210. Essa discrepância no total populacional calculado através dos pesos laboratoriais nos leva a crer que a “fração de amostragem” que aparece na fórmula do peso laboratorial (Szwarcwald et al., 2019b) diminui drasticamente os valores dos pesos fornecidos pelo próprio IBGE, de modo que muitos deles acabam sendo

valores menores do que um, o que significa que cada indivíduo da subamostra de exames laboratoriais representa menos de um indivíduo da população, o que só deveria ocorrer se tivéssemos mais indivíduos na amostra do que na população.

As estimativas dos totais pelos métodos de pós-estratificação e *raking* definidos neste trabalho são mais condizentes com a projeção populacional. O *raking*, devido às próprias características do método, estima corretamente o total populacional, enquanto que a pós-estratificação estima aproximadamente 600 mil indivíduos a menos.

Nas estimativas de prevalência, média e variabilidade (CV) para o conjunto pré estabelecido de características, as estimativas foram similares nos três métodos, apesar dos pesos laboratoriais possuírem maior variabilidade e a pós-estratificação apresentar menor CV em dois terços das estimativas. Além disso, a prevalência estimada para diabetes *mellitus* forneceu estimativas similares em ambos os casos, nos três cenários distintos.

Uma limitação importante deste estudo é que o delineamento apresentado nos métodos alternativos de calibração não é o delineamento que de fato foi utilizado na amostra de exames laboratoriais da PNS 2013. Como não existe uma forma de ligar o conjunto de dados dos microdados da PNS com a subamostra de exames, não se tem disponível os pesos básicos de seleção da primeira fase da pesquisa. Além disso, no próprio banco de dados da subamostra, também não é fornecida a distância entre o setor selecionado para PNS 2013 (primeira fase) e os municípios de grande porte populacional (80 mil habitantes ou mais) dentro da UF do setor, sendo assim, não sendo possível replicar minimamente nenhum peso utilizado nos cálculos da pesquisa. Tendo em vista a falta de dados, optou-se por apresentar métodos alternativos seguindo um delineamento menos complexo, de amostragem aleatória simples, impossibilitando assim a utilização do delineamento em duas fases, o que seria o mais adequado, tendo em vista como os dados da PNS 2013 foram coletados.

Por outro lado, um ponto positivo deste estudo é que foi possível observar que, de acordo com os resultados empíricos deste estudo, apesar de não haver diferenças grandes entre as estimativas de prevalências, médias, CV e intervalos de confiança, quando apresentados os totais populacionais estimados pelos três métodos, observa-se que o peso laboratorial produz estimativas para o total que não são generalizáveis para a população alvo da PNS. Pode-se argumentar que, como as outras estimativas são relativamente semelhantes nos métodos, há a opção de, por exemplo, multiplicar o total da projeção populacional com a prevalência estimada e assim ter um total populacional mais representativo da população.

Por fim, a grande sugestão desta dissertação é de, se possível, os gestores responsáveis pelos dados das Pesquisas Nacionais de Saúde forneçam, de forma documentada e que sejam incluídos nos arquivos de dados públicos, todos os pesos referentes ao delineamento da PNS, além do chaveamento das informações obtidas na amostra e na subamostra de exames laboratoriais. Tais informações permitiriam, primeiramente, a possibilidade de

explorar outros métodos de calibração, como o delineamento por amostragem em duas fases, comumente utilizado em pesquisas com características similares às da PNS, além de dar maior autonomia aos pesquisadores para que utilizem também, outras variáveis auxiliares; além da exploração de métodos de imputação de dados para a obtenção de estimativas mais precisas. Visto que, dada a grande relevância destes dados para a pesquisa em epidemiologia e saúde coletiva, os métodos mais adequados devem ser empregados na análise dos mesmos.