



ARTIGO ORIGINAL



Comparação de métodos de calibração na análise de dados da Pesquisa Nacional de Saúde de 2013

Comparison of calibration methods in the analysis of 2013 Brazilian National Health Survey data

Juliana Sena de Souza^I , Márcia Helena Barbian^{II} , Rodrigo Citton Padilha dos Reis^{I,II}

^IUniversidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Epidemiologia – Porto Alegre (RS), Brasil.

^{II}Universidade Federal do Rio Grande do Sul, Departamento de Estatística – Porto Alegre (RS), Brasil.

RESUMO

Objetivo: Este estudo visa comparar métodos de calibração de pesos na subamostra de Exames Laboratoriais da Pesquisa Nacional de Saúde (PNS) 2013, buscando avaliar sua representatividade e precisão. **Métodos:** Duas propostas alternativas de construção de pesos calibrados foram realizadas com base nos métodos de pós-estratificação e *raking*. Uma comparação entre os pesos fornecidos para a subamostra de Exames Laboratoriais e os dois pesos sugeridos foi realizada por meio de estimativas de parâmetros de interesse utilizando os dados da subamostra da PNS 2013. Além disso, foram utilizadas sete medidas para avaliar o desempenho dos sistemas de pesos propostos. **Resultados:** Os pesos de pós-estratificação e *raking* alternativos produziram estimativas generalizáveis para a população-alvo da PNS 2013, enquanto os pesos originais não. Os métodos alternativos apresentaram desempenho similar ao método original, com pequena vantagem para o *raking* em algumas medidas de avaliação. **Conclusão:** Recomenda-se que os pesos básicos de delineamento sejam documentados e incluídos nos arquivos de dados de uso público da PNS. Além disso, sugere-se o chaveamento das informações entre a amostra e a subamostra da PNS 2013 para possibilitar a exploração de métodos como a imputação de dados, visando obter estimativas mais precisas e representativas. Essas melhorias são essenciais para garantir a qualidade e utilidade dos dados da PNS em estudos epidemiológicos e de saúde coletiva.

Palavras-chave: Estimativas populacionais. Estudos de amostragem. Estudos epidemiológicos. Estudos populacionais em saúde pública. Métodos e procedimentos estatísticos.

AUTOR CORRESPONDENTE: Rodrigo Citton Padilha dos Reis. Programa de Pós-graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul. Rua Ramiro Barcelos, 2.400, sala 215, CEP: 90035-003, Porto Alegre (RS), Brasil. E-mail: citton.padilha@ufrgs.br

CONFLITO DE INTERESSES: nada a declarar.

COMO CITAR ESTE ARTIGO: Souza JS, Barbian MH, Reis RCP. Comparação de métodos de calibração na análise de dados da Pesquisa Nacional de Saúde de 2013. Rev Bras Epidemiol. 2025; 28: e250005. <https://doi.org/10.1590/1980-549720250005.2>

EDITOR ASSOCIADO: José Cazuya de Farias Júnior

EDITOR CIENTIFICO: Cassia Maria Buchalla , Juraci Almeida Cesar

Esse é um artigo aberto distribuído sob licença CC-BY 4.0, que permite cópia e redistribuição do material em qualquer formato e para qualquer fim desde que mantidos os créditos de autoria e de publicação original.

Recebido em: 16/05/2024

Revisado em: 09/10/2024

Aceito em: 07/11/2024



INTRODUÇÃO

A Pesquisa Nacional de Saúde (PNS), conduzida pelo Ministério da Saúde em parceria com a Fundação Oswaldo Cruz (Fiocruz) e o Instituto Brasileiro de Geografia e Estatística (IBGE), é a pesquisa mais ampla já realizada no país sobre a situação de saúde e seus determinantes¹, permitindo o dimensionamento do acesso ao diagnóstico e à atenção à saúde² e a produção de dados dos estilos de vida da população brasileira.

Um grande diferencial da PNS de 2013 foi a inclusão da coleta de material biológico (amostras de sangue e urina dos participantes) para uma subamostra de 25% dos indivíduos que responderam à primeira fase da pesquisa, o que permitiu a realização de exames laboratoriais e a elaboração de estudos referentes à prevalência de anemia, colesterol total, insuficiência renal e diabetes, além de outros temas e fatores associados da população brasileira³⁻⁷. Além disso, esse estudo é de abrangência nacional, o que nunca havia acontecido até então⁸.

No entanto, devido a dificuldades no trabalho de campo, houve uma perda superior a 20% na subamostra de indivíduos indicados para a realização dos exames laboratoriais, resultando em 8.952 participantes que forneceram material biológico. Devido a essas perdas, o método de ponderação de pós-estratificação foi proposto para a análise dos dados e para garantir a validade dos resultados⁹.

Técnicas alternativas à pós-estratificação foram propostas na literatura, entre as quais se destacam o *raking*¹⁰ e a calibração de amostragem em duas fases^{11,12}. A avaliação do desempenho dos estimadores considerando pesos amostrais e demais características do delineamento é uma questão recorrente no campo de pesquisa de levantamentos por amostragem de planos complexos^{13,14}.

Neste cenário, o objetivo do presente estudo foi comparar diferentes métodos de calibração de pesos, buscando não apenas corrigir as distorções da subamostra de Exames Laboratoriais da PNS 2013, mas também melhorar a precisão e a confiabilidade das estimativas. Foram utilizadas medidas de avaliação de desempenho¹⁵ para identificar a estratégia mais eficaz de calibração dos pesos para garantir que os dados coletados pela PNS possam ser plenamente aproveitados, contribuindo para uma compreensão mais profunda da saúde pública no Brasil.

MÉTODOS

Calibração na subamostra de Exames Laboratoriais da PNS 2013

Tendo em vista que a subamostra de Exames Laboratoriais da PNS 2013 foi delineada com base na informação de distância entre o setor selecionado para a PNS 2013 (primeira fase) e os municípios de grande porte populacional (aqueles municípios com 80 mil habitantes ou mais) dentro da UF do setor⁹, esperava-se que um sistema de pesos de

um delineamento em duas fases ($\{d_k^*; k \in S_2\}$, em que S_2 representa a subamostra de Exames Laboratoriais) fosse fornecido aos estatísticos, epidemiologistas e outros pesquisadores que fazem uso dos dados da PNS 2013. Tal informação poderia ser utilizada para a obtenção de estimativas a partir do estimador de expansão dupla, disponível na Equação 3 do Material Suplementar 1, ou ainda, com base no estimador de calibração, permitindo aos usuários dos dados a construção de sistemas de pesos calibrados a partir dos pesos básicos do delineamento em duas fases d_k^* .

Por outro lado, pesos “pós-estratificação” foram fornecidos em conjunto com os dados da subamostra de Exames Laboratoriais (*peso_lab*). Denotaremos os pesos fornecidos para a subamostra de Exames Laboratoriais por $W_k^{(lab)}$, cuja definição pode ser encontrada no artigo de metodologia da subamostra⁹. No cálculo dos pesos “pós-estratificação”, foram considerados os dados dos 60.202 participantes selecionados para as entrevistas individuais na primeira fase da PNS 2013. As seguintes variáveis auxiliares foram consideradas para a composição dos estratos: sexo (dois níveis: masculino e feminino); idade (4 faixas etárias: 18 a 29 anos, 30 a 44 anos, 45 a 59 anos e 60 anos ou mais); raça/cor (quatro níveis: branca, preta, parda e outros); grau de instrução (três níveis: fundamental incompleto, fundamental completo e/ou ensino médio incompleto e ensino médio completo ou mais) e macrorregião geográfica (cinco níveis: Sul, Sudeste, Centro-Oeste, Norte e Nordeste), resultando em um total de 480 pós-estratos. Os pesos “pós-estratificação” foram então definidos por⁹ (Equação 1):

$$w_k^{(lab)} = \frac{N_h}{n_h} \times \frac{\sum n_h}{\sum N_h}, \text{ para } k \text{ pertencente ao estrato } h, \quad (1)$$

Em que:

N_h : o número de moradores selecionados da PNS 2013 em cada estrato,

h e n_h : o número de observações correspondentes na subamostra de Exames Laboratoriais⁹.

Note que, utilizando a notação de delineamentos em duas fases, N_h e n_h seriam denotados por n_{1h} e n_{2h} , respectivamente.

Sobre a construção e utilização dos pesos “pós-estratificação” propostos, destacamos três observações:

- Não é possível saber ao certo, mas a quantidade $\sum n_h / \sum N_h$ parece ser aproximadamente $n_1/n_2 = 8952/60.202$, a “fração de amostragem” em relação à amostra da primeira fase da PNS 2013. Ainda, notamos que os pesos (1) não parecem ser pesos pós-estratificação genuínos, pois estes não levam em consideração os pesos básicos de delineamento;
- As estimativas produzidas com os dados da subamostra de Exames Laboratoriais e os pesos $W_k^{(lab)}$ são generalizáveis para a amostra da PNS 2013 de 60.202 participantes, e não para a população-alvo da PNS, composta por adultos residentes em domicílios particulares permanentes;

- Na análise de dados da subamostra de Exames Laboratoriais, o uso dos pesos definidos em (1) é feito em conjunto com a definição do delineamento da subamostra. Em geral, assume-se uma amostra com probabilidades desiguais de inclusão, em que os pesos básicos são dados pelos pesos $W_k^{(lab)}$. Isso resulta em um cálculo da variância dos estimadores a partir desses pesos $W_k^{(lab)}$.

Em relação a essa última observação, nós argumentamos que a variância dos estimadores deveria utilizar os pesos d_k , ou um sistema de pesos que se aproxime de tal expressão. Maiores informações referente às definições básicas dos delineamentos por amostragem, definição dos pesos básicos e sistema de pesos calibrados estão disponíveis no Material Suplementar 1.

Métodos de calibração alternativos

Como as variáveis utilizadas para a construção dos pesos $W_k^{(lab)}$ são disponibilizadas nos arquivos de microdados da PNS 2013, para a amostra de 60.202 participantes e para a subamostra de Exames Laboratoriais de 8.952 participantes, propomos como forma alternativa de análise dos dados da subamostra da PNS 2013 os seguintes delineamento e construção dos pesos: assumimos que a subamostra de Exames Laboratoriais foi selecionada por amostragem aleatória simples sem reposição da amostra da primeira fase da PNS 2013. Os pesos básicos assumem a forma $d_k = n_1/n_2 = 60.202/8.952 \approx 6,72$ para todos os 8.952 participantes da segunda fase da PNS 2013. Pesos pós-estratificação foram construídos utilizando as mesmas variáveis auxiliares dos pesos $W_k^{(lab)}$, porém adotando como referência a variável “Projeção da população”, disponível nos microdados da PNS 2013 (variável V00282). Com essa projeção, o total de adultos brasileiros é estimado em 145.572.210. Os pesos pós-estratificação são calculados por Equação 2:

$$w_k^{PEAAS} = d_k \frac{N_h}{\hat{N}_{h,\pi}} = \frac{n_1}{n_2} \times \frac{N_h}{\hat{N}_{h,\pi}} = \frac{n_1}{n_2} \times \frac{N_h}{n_{2h}} \times \frac{n_2}{n_1} = \frac{N_h}{n_{2h}}, \quad (2)$$

para k pertencente ao estrato h em que n_{2h} corresponde ao número de participantes no estrato h na subamostra S_2 . A penúltima igualdade em (2) se justifica, pois $\hat{N}_{h,\pi} = \sum_{K \in S_2 \cap U_h} d_k = n_{2h} \times (n_1/n_2)$, uma vez que assumimos $d_k = n_1/n_2$.

Note que a expressão (2) é idêntica à primeira equação do artigo de metodologia da subamostra de Exames Laboratoriais⁹, com a diferença que em (2) N_h representa o número de moradores em cada estrato h na população brasileira, e não mais o número de moradores selecionados da PNS 2013 em cada estrato h . Dessa forma, propomos um sistema de pesos que produz estimativas generalizáveis para a população-alvo da PNS 2013.

Uma segunda alternativa na criação do sistema de pesos calibrados, ainda assumindo que os pesos básicos correspondem à amostragem aleatória simples, foi obtida por meio do método conhecido como *raking*. O processo envol-

ve a pós-estratificação de cada conjunto de variáveis por vez e a repetição desse processo até que os pesos se estabilizem¹⁶. O *raking* permite que várias variáveis de agrupamento sejam usadas sem a necessidade de construir uma classificação cruzada completa. As variáveis auxiliares utilizadas para construção desses pesos foram as mesmas utilizadas para construir os pesos $W_k^{(PEAAS)}$, e a variável “Projeção da população” também foi utilizada. Denotaremos por $W_k^{(rakeAAS)}$ os pesos obtidos a partir do *raking*.

Aspectos éticos

Este estudo utilizou dados de acesso públicos e anonimizados da PNS do IBGE, portanto, não foi necessária a aprovação do Comitê de Ética em Pesquisa.

Análise estatística

Para avaliar a sub e sobre representação dos grupos das variáveis auxiliares (utilizadas para construção dos pesos calibrados), supondo que as projeções populacionais (a partir dos dados da PNS 2013) são os valores verdadeiros da população, foram comparadas as frequências relativas estimadas pelos três sistemas de pesos $W_k^{(lab)}$, $W_k^{(PEAAS)}$ e $W_k^{(rakeAAS)}$ com os delineamentos já apresentados.

Como forma de comparar as propostas de pesos calibrados com os pesos fornecidos em conjunto com a subamostra de Exames Laboratoriais, foram estimados parâmetros populacionais para certas variáveis de interesse, conforme a descrição da Tabela 1. As variáveis com código iniciando com a letra “Z” foram obtidas na fase de coleta de Exames Laboratoriais da PNS 2013. As demais variáveis (código iniciando com a letra “Q” ou “J”) foram obtidas na primeira fase de coleta da PNS 2013, e estão disponíveis para os 60.202 participantes.

Para avaliar as estimativas produzidas com os sistemas de pesos propostos na $W_k^{(PEAAS)}$ e $W_k^{(rakeAAS)}$ foram calculadas as 7 medidas apresentadas no Material Suplementar 2. As análises foram realizadas no *software* R¹⁷ em conjunto com o pacote *survey*¹⁸ para incorporar o delineamento de amostragem. O código em R com a especificação do objeto de delineamento (*svydesign*) utilizado nas análises é apresentado no Material Suplementar 3.

RESULTADOS

Distribuição dos pesos

As distribuições dos pesos de $W_k^{(lab)}$ e $W_k^{(PEAAS)}$ possuem forma similar, no entanto, os valores que cada sistema de pesos assume é muito distinta, refletindo o que cada um dos sistemas de pesos representa: se, por um lado, $W_k^{(lab)}$ tem como alvo representar a amostra da primeira fase da PNS 2013 (Figura 1A); por outro lado, $W_k^{(PEAAS)}$ tem como objetivo representar a população-alvo da pesquisa, ou seja, adultos brasileiros residentes em domicílios particulares permanentes (Figura 1B). A distribuição de pesos do *raking* apresenta uma forma distinta das demais distribuições de

Tabela 1. Características de interesse da subamostra de Exames Laboratoriais da PNS 2013 selecionadas para avaliação das estimativas de calibração.

Código da variável	Descrição da variável	Tipo de parâmetro
Q030	Diabetes (<i>sim, não, apenas na gravidez</i>)	Total de prevalência
Q002	Hipertensão (<i>sim, não, apenas na gravidez</i>)	Total de prevalência
Q060	Colesterol alto (<i>sim, não</i>)	Total de prevalência
Q063	Doença do coração [infarto, angina, insuficiência cardíaca ou outra] (<i>sim, não</i>)	Total de prevalência
Q068	Acidente vascular cerebral – AVC ou derrame (<i>sim, não</i>)	Total de prevalência
Q120	Câncer (<i>sim, não</i>)	Total de prevalência
Q124	Insuficiência renal crônica – IRC (<i>sim, não</i>)	Total de prevalência
J058	Dengue (<i>sim, não</i>)	Total de prevalência
Z025	Creatinina (mg/dL)	Total e média
Z031	Colesterol total (mg/dL)	Total e média
Z034	Hemoglobina glicosilada – HbA1c (%)	Total e média

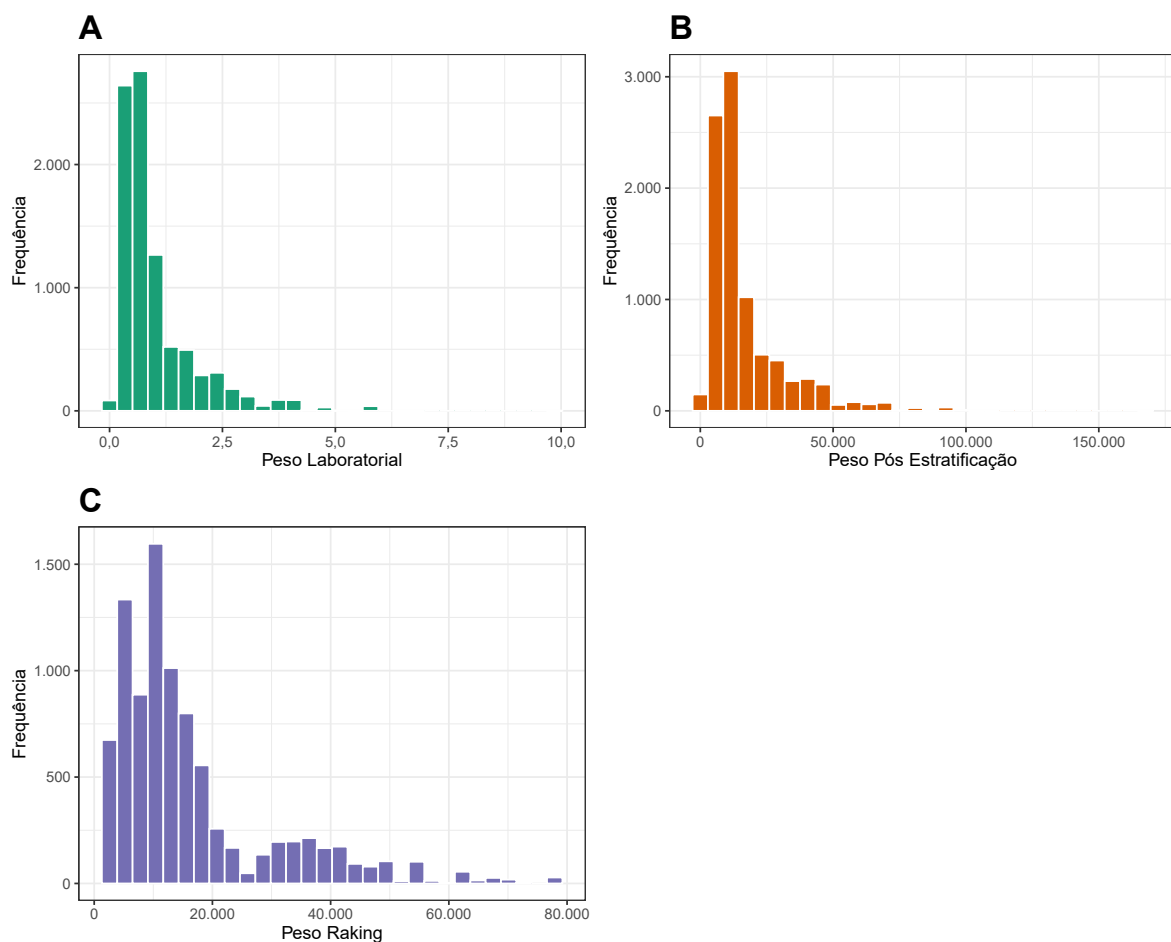


Figura 1. Distribuição dos pesos de (A) pós-estratificação disponibilizados junto aos dados da subamostra de exames laboratoriais da PNS 2013; e de (B) pós-estratificação e (C) raking construídos a partir da projeção populacional – subamostra de Exames Laboratoriais, PNS 2013.

pesos $W_k^{(lab)}$ e $W_k^{(PEAS)}$, mas com valores parecidos com os pesos $W_k^{(PEAS)}$. Assim, os pesos $W_k^{(rakeAS)}$ também produzem estimativas generalizáveis para a população-alvo da PNS 2013 (Figura 1C).

Representatividade dos pós-estratos

Os pesos $W_k^{(lab)}$ apresentam estimativas das proporções das variáveis auxiliares próximas da projeção po-

pulacional (Figura S1, Material Suplementar). Em geral, as estimativas com base em $W_k^{(lab)}$ apresentam diferença (erro absoluto) de, no máximo, 0,15%, com exceção das categorias raça/cor parda (0,22%), raça/cor outra (0,42%) e região Nordeste (-0,26%). Já as estimativas obtidas pelos pesos $W_k^{(PEAS)}$ apresentaram erro absoluto, em geral, também não maior que 0,15%, com exceção das categorias raça/cor branca (0,21%), parda (-0,20%) e outra (0,43%).

Com os pesos calibrados pelo método do *raking*, $W_k^{(rakeAAS)}$, as estimativas das proporções das variáveis auxiliares coincidem com as projeções populacionais. Resultados semelhantes foram obtidos ao se comparar as estimativas dos totais populacionais das variáveis auxiliares (Tabela S1, Material Suplementar 4). Consideramos calcular o erro relativo (ER) de $W_k^{(lab)}$, no entanto, é possível notar que as estimativas populacionais oriundas desses pesos são muito diferentes das projeções populacionais, o que resulta em um ER extremamente alto — por isto optamos por não apresentar esses resultados.

Medidas de avaliação da calibração

As 7 medidas de avaliação dos métodos de calibração são apresentadas na Tabela 2. A média do ER absoluto de calibração foi $M_1 = 2,16$ para o método de pós-estratificação e $M_1 = 0$ para o *raking*. Esses resultados já eram esperados, tendo em vista que os totais estimados para as variáveis auxiliares foram, em geral, inferiores aos totais populacionais. A média dos coeficientes de variação dos totais das variáveis auxiliares $M_2 = 0\%$ para ambos os métodos indica que estimativas dos totais das variáveis auxiliares produzidos pelos dois métodos são não tendenciosas.

As medidas M_3 e M_4 (proporção de pesos extremos) indicam a presença de pesos g extremos tanto para o método de pós-estratificação quanto para o *raking*. Isso era esperado, tendo em vista a distância entre os pesos assumidos como pesos de delineamento d_k e os pesos construídos $W_k^{(PEAAS)}$ e $W_k^{(rakeAAS)}$ ($M_6 = 13.440,07$ para a pós-estratificação e $M_6 = 11.447,09$ para o *raking*). O coeficiente de variação dos pesos g também indica essa característica da construção dos pesos calibrados e foi alto para os dois métodos ($M_5 = 91,37\%$ para a pós-estratificação e $M_5 = 83,94\%$ para o *raking*), apresentando uma leve vantagem para o método *raking*.

A eficiência média das estimativas (M_7) pelos métodos alternativos de calibração para o conjunto de variáveis apresentadas na Tabela 1 indicam que há uma pequena vantagem do método de pós-estratificação em comparação ao *raking*.

Precisão das estimativas dos parâmetros de interesse

A Tabela 3 apresenta as estimativas para os parâmetros relacionados às características de interesse (listados

na Tabela 1), juntamente com os coeficientes de variação obtidos a partir dos três sistemas de pesos calibrados. Observa-se que as estimativas pontuais (totais, prevalências e médias) produzidas pelos três métodos, em geral, são muito próximas. A exceção diz respeito à estimativa dos totais populacionais a partir dos pesos pós-estratificação da própria subamostra de Exames Laboratoriais, os pesos $W_k^{(lab)}$. Como já observado, tais resultados não são generalizáveis para a população-alvo da PNS 2013. Por fim, as estimativas de coeficiente de variação das estimativas apresentam maior precisão dos métodos de calibração alternativos propostos neste trabalho. Este resultado era esperado, uma vez que a estimação do erro-padrão das estimativas leva em consideração os aspectos do delineamento amostral assumido.

Com o propósito de avaliar o desempenho dos métodos de calibração na estimativa em subgrupos da população, foram obtidas estimativas de prevalência de diabetes de acordo com algumas características de interesse (Tabela 4). Mais uma vez, percebe-se que as estimativas pontuais são muito semelhantes para os três sistemas de peso, e que o ganho aparente dos métodos alternativos de calibração se dá nas estimativas de precisão. Os intervalos de confiança de 95% para a prevalência de diabetes nos subgrupos da população produzidos pelos métodos de pós-estratificação ($W_k^{(PEAAS)}$) e *raking* ($W_k^{(rakeAAS)}$) são ligeiramente mais estreitos que aqueles produzidos pelo peso $W_k^{(lab)}$.

DISCUSSÃO

A Pesquisa Nacional de Saúde em 2013 envolveu a coleta de uma subamostra de exames laboratoriais, importante contribuição para estudos da saúde da população brasileira. As técnicas de amostragem sugerem que um delineamento em duas fases poderia ter sido empregado na subamostra da PNS 2013 para a construção de sistemas de pesos básicos e calibrados. No entanto, dificuldades na coleta da amostra da segunda fase resultaram na não divulgação dos pesos de delineamento em conjunto com os microdados da subamostra de Exames Laboratoriais. Na ausência de pesos básicos de amostragem, os gestores dos dados da subamostra forneceram pesos calibrados pelo método de pós-estratificação.

Tabela 2. Medidas de desempenho dos métodos de calibração por pós-estratificação e *raking* – subamostra de Exames Laboratoriais, PNS 2013.

Medida de avaliação	$W_k^{(PEAAS)}$	$W_k^{(rakeAAS)}$
M1: média do erro relativo absoluto de calibração	2,16	0,00
M2: média dos coeficientes de variação (CV) dos totais das variáveis auxiliares	0,00%	0,00%
M3: proporção dos pesos extremos em relação a 1/3	19,02%	18,81%
M4: proporção dos pesos extremos em relação a 3	4,39%	3,24%
M5: coeficiente de variação dos pesos g	91,37%	83,94%
M6: distância χ^2 entre w e d	13.440,07	11.447,09
M7: eficiência média das estimativas de calibração	1,005	1,018

Tabela 3. Totais e prevalências (%)[‡] estimados e coeficiente de variação (CV%) das características de interesse obtidos a partir dos três sistemas de pesos calibrados — subamostra de Exames Laboratoriais, PNS 2013.

Características de interesse	$W_k^{(lab)}$ (N = 8.952)		$W_k^{(PEAAS)}$ (N = 144.908.922)		$W_k^{(rakeAAS)}$ (N = 145.572.210)	
	Estimativa (%)	CV%	Estimativa (%)	CV%	Estimativa (%)	CV%
Diabetes	573 (7,69)	5,04	9.275.930 (7,7)	4,76	9.266.527 (7,65)	4,85
Hipertensão	1.930 (25,91)	2,53	31.236.815 (25,92)	2,18	31.171.025 (25,75)	2,25
Colesterol alto	1.160 (15,57)	3,42	18.773.491 (15,58)	3,2	18.842.610 (15,56)	3,25
Doença do coração	370 (4,96)	6,15	5.971.338 (4,95)	5,83	6.092.130 (5,03)	5,98
AVC	146 (1,96)	11,57	2.367.716 (1,96)	10,47	2.266.318 (1,87)	10,3
Câncer	158 (2,11)	9,87	2.540.009 (2,11)	9,45	2.549.079 (2,11)	9,57
IRC	111 (1,48)	12,04	1.785.325 (1,48)	11,78	1.815.305 (1,5)	11,73
Dengue	1.406 (18,88)	3,06	22.735.194 (18,86)	2,9	22.732.548 (18,78)	2,91
Creatinina (Média, mg/dL)	0,881	0,44	0,881	0,38	0,882	0,44
Colesterol total (Média, mg/dL)	185,031	0,31	185,01	0,29	184,815	0,29
HbA1c (Média, %)	5,513	0,23	5,513	0,22	5,51	0,23

[‡]Para as características Creatinina, Colesterol total e HbA1c foram estimados média e CV.

Tabela 4. Prevalência de diabetes (em %) e intervalo de confiança de 95% (IC95%) de acordo com subgrupos da população obtidos a partir dos três sistemas de pesos calibrados – subamostra de Exames Laboratoriais, PNS 2013.

Características	$W_k^{(lab)}$		$W_k^{(PEAAS)}$		$W_k^{(rakeAAS)}$	
	%	IC95%	%	IC95%	%	IC95%
Idade						
18 a 29 anos	1,4	0,5–2,4	1,4	0,6–2,3	1,3	0,5–2,0
30 a 44 anos	2,4	1,6–3,3	2,4	1,6–3,2	2,4	1,6–3,1
45 a 59 anos	10,8	9,1–12,5	10,8	9,2–12,5	11,1	9,4–12,8
≥60 anos	18,2	15,9–20,5	18,2	16,0–20,5	17,9	15,6–20,2
Raça/cor						
Branca	7,3	6,2–8,4	7,3	6,3–8,3	7,3	6,3–8,4
Preta	9,6	6,8–12,5	9,7	7,1–12,2	9,7	6,9–12,4
Parda	7,2	6,2–8,3	7,2	6,2–8,2	7,1	6,1–8,2
Outra	6,0	1,1–10,8	45,5	2,0–8,9	4,6	1,1–8,0
Índice de Massa Corporal – IMC (kg/m ²)						
Baixo peso/normal (IMC<25)	4,2	3,4–4,1	4,2	3,4–5,0	4,3	3,4–5,2
Sobrepeso (25≤IMC<30)	7,5	6,2–8,7	7,5	6,2–8,7	7,2	6,0–8,4
Obesidade (IMC≥30)	13,5	11,4–15,5	13,5	11,5–15,4	13,5	11,6–15,5
Plano de Saúde						
Sim	7,1	5,8–8,4	7,1	5,8–8,3	7,0	5,8–8,3
Não	7,6	6,8–8,5	7,7	6,8–8,5	7,6	6,7–8,4

Neste trabalho, propusemos dois métodos de calibração alternativos com base na pós-estratificação e no *raking*. Os sistemas de pesos obtidos a partir destes métodos apresentaram desempenho concorrente ao sistema de pesos disponibilizado junto aos dados dos Exames Laboratoriais. Em particular, destaca-se que as estimativas obtidas a partir dos sistemas de pesos sugeridos são generalizáveis

para a população-alvo da PNS 2013: a população brasileira de adultos residentes em domicílios particulares.

Outro aspecto a ser enfatizado é que os dois métodos de calibração propostos apresentaram maior precisão para as estimativas consideradas neste trabalho. Uma possível explicação para esse comportamento das estimativas é o fato de termos utilizado alguns aspectos

importantes, embora de forma suposta, do delineamento do levantamento por amostragem, que resultam no cálculo adequado da estimativa da variância das estimativas de parâmetros de interesse.

Ao utilizar medidas de avaliação do desempenho dos métodos de calibração, os dois sistemas de pesos sugeridos apresentaram um bom desempenho, como uma vantagem para os pesos baseados no *raking*. Alguns trabalhos anteriores apontam para as mesmas conclusões¹⁹⁻²². No entanto, é preciso salientar que medidas referentes à comparação entre os pesos calibrados e “pré-calibrados” identificam uma distância muito grande entre essas duas ponderações. É provável que esse comportamento seja observado pelo fato de termos assumido um delineamento de amostragem aleatória simples a partir da amostra da primeira fase da PNS 2013 para os dados da subamostra de Exames Laboratoriais, enquanto os pós-estratos foram construídos a partir das projeções da população brasileira.

Como pode ser observado, a subamostra de Exames Laboratoriais da PNS 2013 poderia ser vista como resultado de um delineamento de amostragem em duas fases¹⁶. No entanto, as probabilidades de seleção (ou, de forma análoga, os pesos básicos de amostragem) não foram disponibilizadas no arquivo de microdados dos Exames Laboratoriais, e assim, as vantagens de um delineamento em duas fases não puderam ser empregadas. A disponibilização dos pesos básicos de delineamento da subamostra de Exames Laboratoriais pode auxiliar estatísticos e epidemiologistas a construir sistemas de pesos calibrados de tal forma a melhorar o desempenho dos estimadores de calibração.

Com respeito à estimativa de totais populacionais (e.g., o total de indivíduos com diabetes), os dois métodos propostos por este artigo fornecem estimativas para a população-alvo da PNS 2013, generalizando os resultados para a população de adultos brasileiros. Já o sistema de pesos fornecidos em conjunto com os dados da subamostra de Exames Laboratoriais resultou em estimativas para a amostra da PNS 2013. Embora seja possível obter estimativas para a população-alvo da PNS de forma indireta, multiplicando-se a proporção amostral pelo tamanho da população, é preferível que os métodos, em conjunto com os *softwares* adequados, forneçam diretamente essas estimativas para evitar erros de interpretação dos resultados.

A abordagem adotada neste artigo foi a da ponderação por meio de métodos de calibração. Algumas limitações do nosso trabalho são inerentes à referida abordagem, por exemplo, a aplicação de pesos para estimativas mais complexas, como coeficientes de regressão; a dificuldade em avaliar os erros-padrões das estimativas ponderadas; e as decisões envolvidas na criação de sistemas de pesos²³. Ainda assim, recomendamos que estimativas ponderadas sejam utilizadas no lugar de estimativas não ponderadas a partir de dados de inquéritos populacionais de delineamento de amostragem complexo. Uma outra limitação deste trabalho foi a suposição de certos aspectos do de-

lineamento empregado para a coleta dos dados da subamostra de Exames Laboratoriais da PNS 2013. Como já enfatizado na literatura, as análises devem considerar os aspectos relevantes do plano amostral¹⁴, e a disponibilização de tais aspectos referentes ao plano da subamostra de Exames Laboratoriais da PNS proporcionaria aos pesquisadores maior precisão às estimativas das quantidades populacionais de interesse.

Finalizamos com duas sugestões para os amostristas e gestores das Pesquisas Nacionais de Saúde que serão utilizadas por outros pesquisadores. Primeiro, que todos os pesos referentes ao delineamento da PNS sejam documentados e incluídos nos arquivos de dados de uso público. Isso possibilitaria que os usuários construíssem estimativas a partir dos pesos básicos ou calibrados, utilizando as suas próprias variáveis auxiliares. A correta precisão das estimativas poderia ser obtida com a utilização adequada de *softwares* de análise de dados oriundos de planos amostrais complexos.

A nossa segunda recomendação diz respeito ao chaveamento das informações obtidas na amostra e subamostra (de Exames Laboratoriais) da PNS 2013. Tal informação permitiria a exploração de métodos como a imputação de dados para obter estimativas mais precisas. 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.

REFERÊNCIAS

1. Malta DC, Stopa SR, Szwarcwald CL, Gomes NL, Silva Júnior JB, Reis AAC. Surveillance and monitoring of major chronic diseases in Brazil - National Health Survey, 2013. *Rev Bras Epidemiol.* 2015; 18(Supl. 2): 3-16. <https://doi.org/10.1590/1980-5497201500060002>
2. IBGE. Pesquisa nacional de saúde: 2013: percepção do estado de saúde, estilos de vida e doenças crônicas: Brasil, grandes regiões e unidades da federação [Internet]. Rio de Janeiro: IBGE; 2014 [acessado em 21 set. 2022]. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?id=291110&view=detalhes>
3. Machado IE, Malta DC, Bacal NS, Rosenfeld LGM. Prevalence of anemia in Brazilian adults and elderly. *Rev Bras Epidemiol.* 2019; 22(Supl. 2): E190008.Supl.2. <https://doi.org/10.1590/1980-549720190008.supl.2>
4. Malta DC, Szwarcwald CL, Machado IE, Pereira CA, Figueiredo AW, Sá ACMGN, et al. Prevalence of altered total cholesterol and fractions in the Brazilian adult population: National Health Survey. *Rev Bras Epidemiol.* 2019; 22(Supl. 2): E190005. SUPL.2. <https://doi.org/10.1590/1980-549720190005.supl.2>
5. Malta DC, Machado IE, Pereira CA, Figueiredo AW, Aguiar LK de, Almeida WS, et al. Evaluation of renal function in the Brazilian adult population, according to laboratory criteria from the National Health Survey. *Rev Bras Epidemiol.* 2019; 22(Supl. 2): E190010.SUPL.2. <https://doi.org/10.1590/1980-549720190010.supl.2>

6. Malta DC, Duncan BB, Schmidt MI, Machado IE, Silva AG da, Bernal RTI, et al. Prevalence of diabetes mellitus as determined by glycated hemoglobin in the Brazilian adult population, National Health Survey. *Rev Bras Epidemiol.* 2019; 22(Supl. 2): E190006.SUPL.2. <https://doi.org/10.1590/1980-549720190006.supl.2>
7. dos Reis RCP, Duncan BB, Szwarcwald CL, Malta DC, Schmidt MI. Control of glucose, blood pressure, and cholesterol among adults with diabetes: the Brazilian National Health Survey. *J Clin Med.* 2021; 10(15): 3428. <https://doi.org/10.3390/jcm10153428>
8. Malta DC, Szwarcwald CL, Silva JBD. First results of laboratory analysis in the National Health Survey. *Rev Bras Epidemiol.* 2019; 22(Supl. 02): E190001.SUPL.2. <https://doi.org/10.1590/1980-549720190001.supl.2>
9. Szwarcwald CL, Malta DC, Souza PRB de, Almeida WS de, Damascena GN, Pereira CA, et al. Laboratory exams of the National Health Survey: methodology of sampling, data collection and analysis. *Rev Bras Epidemiol.* 2019; 22(Supl. 2): E190004.SUPL.2. <https://doi.org/10.1590/1980-549720190004.supl.2>
10. Deville JC, Sarndal CE, Sautory O. Generalized raking procedures in survey sampling. *J Am Stat Assoc.* 1993; 88(423): 1013-20. <https://doi.org/10.2307/2290793>
11. Amorim G, Tao R, Lotspeich S, Shaw PA, Lumley T, Shepherd BE. Two-phase sampling designs for data validation in settings with covariate measurement error and continuous outcome. *J R Stat Soc Ser A Stat Soc.* 2021; 184(4): 1368-89. <https://doi.org/10.1111/rssa.12689>
12. Neyman J. Contribution to the theory of sampling human populations. *J Am Stat Assoc.* 1938; 33(201): 101-16. <https://doi.org/10.2307/2279117>
13. Korn EL, Graubard BI. Epidemiologic studies utilizing surveys: accounting for the sampling design. *Am J Public Health.* 1991; 81(9): 1166-73. <https://doi.org/10.2105/AJPH.81.9.1166>
14. Silva, PLN, Pessoa DGC, Lila MF. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Ciênc Saúde Coletiva.* 2002; 7(4): 659-70. <https://doi.org/10.1590/S1413-81232002000400005>
15. Silva PLN. Calibration estimation: when and why, how much and how [Internet]. Rio de Janeiro: IBGE; 2004 [acessado em 21 set. 2022]. Disponível em: <https://biblioteca.ibge.gov.br/biblioteca-catalogo?id=281040&view=detalhes>
16. Haziza D, Beaumont JF. Construction of weights in surveys: a review. *Statist Sci.* 2017; 32(2): 206-26. <https://doi.org/10.1214/16-STS608>
17. R Core Team. R: A language and environment for statistical computing [Internet]. Viena: R Foundation for Statistical Computing; 2021 [acessado em 16 maio 2021]. Disponível em: <https://www.r-project.org/>
18. Lumley T. Analysis of complex survey samples. *J Stat Softw.* 2004; 9(8): 1-19. <https://doi.org/10.18637/jss.v009.i08>
19. Djerf K. Effects of post-stratification on the estimates of the finnish labour force survey. *J Off Stat.* 1997; 13(1): 29-39.
20. Ruiz CMM, Silva PLN. Explorando alternativas para a calibração dos pesos amostrais da Pesquisa Nacional por Amostra de Domicílios. In: Proceedings of the Conference Name. Lima, Peru; 2014.
21. Tu SH. A comparison of propensity score sub-classification and other calibration methods based on a telephone sample to estimate internet usage. *Taiwanese J Sociol.* 2015; 56: 115-50. [https://doi.org/10.6786/TJS.201506_\(56\).0003](https://doi.org/10.6786/TJS.201506_(56).0003)
22. Bernal RTI, Iser BPM, Malta DC, Claro RM. Sistema de vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico (Vigitel): mudança na metodologia de ponderação. *Epidemiol Serv Saúde.* 2017; 26(4): 701-12. <https://doi.org/10.5123/S1679-49742017000400003>
23. Gelman A. Struggles with survey weighting and regression modeling. *Statist Sci.* 2007; 22(2): 153-64. <https://doi.org/10.1214/088342306000000691>

ABSTRACT

Objective: This study aims to compare calibration methods for weights in the subsample of Laboratory Exams from the 2013 Brazilian National Health Survey (PNS), seeking to assess their representativeness and precision. **Methods:** Two alternative proposals for constructing calibrated weights were performed based on post-stratification and raking methods. A comparison between the weights provided for the Laboratory Exams subsample and the two suggested weights was conducted through parameter estimates using the 2013 PNS subsample data. Additionally, seven measures were used to assess the performance of the proposed weighting systems. **Results:** The alternative post-stratification and raking weights produced generalizable estimates for the target population of the 2013 PNS, while the original weights did not. The alternative methods showed similar performance to the original method, with a slight advantage for raking in some evaluation measures. **Conclusion:** It is recommended that basic design weights be documented and included in the public-use data files of the PNS. Furthermore, it is suggested to cross-reference information between the sample and subsample of the 2013 PNS to enable the exploration of methods such as data imputation, aiming to obtain more accurate and representative estimates. These improvements are essential to ensure the quality and usefulness of PNS data in epidemiological and public health studies.

Keywords: Population forecast. Sampling studies. Epidemiologic studies. Population studies in public health. Statistics as topic.

CONTRIBUIÇÕES DOS AUTORES: Souza, J. S.: Conceituação, Curadoria de Dados, Análise Formal, Metodologia, Software, Visualização, Escrita – Primeira Redação, Escrita – Revisão e Edição. Barbian, M. H.: Conceituação, Supervisão, Escrita – Primeira Redação, Escrita – Revisão e Edição. Reis, R. C. P.: Conceituação, Metodologia, Supervisão, Escrita – Primeira Redação, Escrita – Revisão e Edição.

FONTE DE FINANCIAMENTO: nenhuma.

