Instituto de MATEMÁTICA E ESTATÍSTICA UFRGS

DEPARTAMENTO DE ESTATÍSTICA
E-mail: dest@mat.ufrgs.br

Decomposição de taxas com o método de Das Gupta: conceitos e aplicações na avaliação dos determinantes sociais do diabetes mellitus na população adulta brasileira

Tânia Cristina Bispo Gonçalves

09 de dezembro de 2025

Tânia Cristina Bispo Gonçalves

Decomposição de taxas com o método de Das Gupta: conceitos e aplicações na avaliação dos determinantes sociais do diabetes mellitus na população adulta brasileira

Trabalho de Conclusão de Curso apresentado à Comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador: Prof. Dr. Rodrigo Citton Padilha dos Reis

Porto Alegre
09 de dezembro de 2025

Tânia Cristina Bispo Gonçalves

Decomposição de taxas com o método de Das Gupta: conceitos e aplicações na avaliação dos determinantes sociais do diabetes mellitus na população adulta brasileira

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador:_____
Prof. Dr. Rodrigo Citton Padilha dos Reis,
Doutor pela Universidade Federal de Minas Gerais,
Belo Horizonte, MG

Banca Examinadora:

Prof. Rogério Boff Borges, UFRGS
Mestre pela Universidade Federal do Rio Grande do Sul – Porto Alegre, RS

Porto Alegre
09 de dezembro de 2025

"O território é um espaço
vivido e desigual"

Milton Santos

Dedico este trabalho às mulheres da minha vida: minha mãe, Vanete (In Memoriam), que continua sendo minha referência de princípios e valores, e às minhas filhas, Shayene e Jhennyfer, das quais tenho muito orgulho e que constantemente me inspiram a ser uma pessoa melhor.

Agradecimentos

Agradeço ao professor Rodrigo Citton, que aceitou o desafio de me orientar e que, mesmo com as rasteiras que levamos da vida no último ano, acredito que formamos uma parceria de sucesso. Por todos os ensinamentos, não só na orientação, mas desde o primeiro semestre da graduação, e por todas as palavras de incentivo e as tentativas de acalmar minhas inseguranças.

Ao professor Rogério Borges por ter aceitado dedicar seu valioso tempo para avaliar e contribuir para esse trabalho.

Ao grupo Racionais MC's, que em *Capítulo 4, versículo 3* me apresentou os primeiros dados estatísticos que, infelizmente, fizeram sentido.

Ao meu grupo oficial de seminários, Eduardo Kempfer e Valentina Bezpalko, parceiros que dividiram as ansiedades, as preocupações e as risadas. E, claro, ao Rodrigo Villa, que foi além: generoso, paciente e sempre disposto a me ajudar a entender o que parecia impossível.

Aos meus amigos da Biblioteca de Ciências Sociais e Humanidades da UFRGS (colegas, bolsistas e usuárias), que colaboraram com muitas pitadas de incentivo, conhecimento e correções ortográficas.

Agradeço aos familiares e amigos que entenderam minha ausência e distanciamento nesse período e também aos que não entenderam, mas, ainda assim, garantiram a minha saúde mental, como, por exemplo, meus "pimos" Vitor de Jesus e Megui Rodrigues, e, em especial, aos meus "xuxus", Éder Bispo e Lilian Valente, que asseguraram que nossos encontros regulares acontecessem, sempre com boa comida e boas risadas; e ao Damião dos Santos e à Raimunda Gonçalves, que, nos fins de semana, faziam questão de tentar me tirar de casa, preparando meus pratos preferidos com aquele cuidado de quem personaliza os afetos.

Ao meu irmão Alexandre Bispo, pelos incentivos, reconhecimentos e pela sua persistência em evoluir, que tanto me motiva.

Ao Leonardo Echevarria, Preto, meu companheiro, por todo o incentivo desde antes do ingresso até agora, sempre me encorajando. Obrigada, Paixão!

E agradeço às políticas públicas e afirmativas brasileiras, que me garantiram o acesso e a permanência em uma universidade pública, gratuita e de excelência: a UFRGS.

Resumo

Este estudo aplica o método de decomposição de Das Gupta para investigar a contribuição de determinantes sociais na variação das taxas de prevalência de diabetes mellitus na população adulta brasileira entre 2013 e 2019. Utilizando dados da Pesquisa Nacional de Saúde (PNS), foram estimadas taxas de prevalência específicas e padronizadas e, em seguida, aplicada a decomposição considerando os fatores idade, sexo, raça/cor, escolaridade e renda, organizados em estrutura de classificação cruzada, permitindo separar a diferença total entre efeitos estruturais e efeitos das taxas específicas. Os resultados evidenciaram que o envelhecimento populacional foi o principal responsável pelo aumento da prevalência de diabetes, contribuindo com a maior parte da variação observada em todos os cenários analisados. Fatores como escolaridade e renda atuaram no sentido oposto, reduzindo parte do crescimento esperado. A implementação computacional foi realizada em R, com códigos reproduzíveis incluídos nos apêndices. O estudo demonstra a utilidade de considerar múltiplos fatores simultaneamente para compreender suas influências e mostra que o método de Das Gupta é uma ferramenta confiável e aplicável à avaliação de desigualdades em saúde.

Palavras-Chave: Decomposição, Demografia, Padronização, Epidemiologia, Determinantes sociais, Diabetes mellitus, Das Gupta, PNS, Pesquisa Nacional de Saúde.

Abstract

This study applies the Das Gupta decomposition method to investigate the contribution of social determinants to the variation in prevalence rates of diabetes mellitus in the Brazilian adult population between 2013 and 2019. Using data from the National Health Survey (PNS), specific and standardized prevalence rates were estimated, and then decomposition was applied considering the factors age, sex, race/color, education, and income, organized in a cross-classification structure, allowing the separation of the total difference between structural effects and the effects of specific rates. The results showed that population aging was the main driver of the increase in diabetes prevalence, contributing most of the variation observed in all scenarios analyzed. Factors such as education and income acted in the opposite direction, reducing part of the expected growth. The computational implementation was performed in R, with reproducible code included in the appendices. The study demonstrates the usefulness of considering multiple factors simultaneously to understand their influences and shows that the Das Gupta method is a reliable and applicable tool for evaluating health inequalities.

Keywords: Decomposition, Demographics, Standardization, Epidemiology, Social determinants, Diabetes mellitus, Das Gupta, PNS, National Health Survey.

Sumário

1 INTRODUÇÃO **13**

1.1 Revisão da Literatura 14

1.2 Justificativa 16

1.3 Objetivos 17

1.4 Organização do TCC 17

2 MATERIAL E MÉTODOS **19**

2.1 Pesquisa Nacional de Saúde 19

2.2 Taxas Gerais e Específicas 20

2.3 Métodos de Padronização 21

2.4 Decomposição e Método Das Gupta 21

2.5 Implementação Computacional 24

3 RESULTADOS **25**

3.1 Caracterização da amostra 25

3.2 Taxas Específicas 27

3.3 Taxas Padronizadas 29

3.4 Decomposição das diferenças 35

4 CONCLUSÃO E CONSIDERAÇÕES FINAIS **47**

REFERÊNCIAS **48**

APÊNDICE A – CÓDIGO R: PADRONIZAÇÃO POR IDADE **53**

APÊNDICE B – FUNÇÃO PARA TABELAS CRUZADAS **55**

APÊNDICE C – CLASSIFICAÇÃO CRUZADA: IDADE E SEXO **59**

APÊNDICE D – CÓDIGO R: DECOMPOSIÇÃO POR IDADE E SEXO **60**

APÊNDICE E – CÓDIGO R: TABELA CRUZADA PARA DADOS COM COMBINAÇÕES AUSENTES 63

APÊNDICE F – DISTRIBUIÇÃO POPULACIONAL ESTIMADA 66

APÊNDICE G – TAXAS PADRONIZADAS 67

Lista de Figuras

Figura 3.1: Pirâmide etária da população brasileira em 2013 e 2019 27
Figura 3.2: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e sexo 37
Figura 3.3: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e raça/cor 39
Figura 3.4: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e escolaridade 41
Figura 3.5: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e renda 43
Figura 3.6: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade, renda e escolaridade 44
Figura 3.7: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade, sexo, renda e escolaridade 46

Lista de Tabelas

Tabela 3.1: Características da amostra, Brasil, PNS 2013 e 2019 26
Tabela 3.2: Prevalência de diabetes - taxas específicas, Brasil, PNS 2013 e 2019 28
Tabela 3.3: Prevalência de diabetes padronizada por idade, Brasil, PNS 2013 e 2019 30
Tabela 3.4: Prevalência de diabetes padronizada por escolaridade[a], Brasil, PNS 2013 e 2019 32
Tabela 3.5: Prevalência de diabetes padronizada por idade e sexo, Brasil, PNS 2013 e 2019 34
Tabela 3.6: Padronização e decomposição das taxas de prevalência segundo idade e sexo (n_2013 = 60.202; n_2019 = 88.531) 36
Tabela 3.7: Padronização e decomposição das taxas de prevalência segundo idade e raça/cor (n_2013 = 60.199; n_2019 = 88.522) 38
Tabela 3.8: Padronização e decomposição das taxas de prevalência segundo idade e escolaridade (n_2013 = 60.202; n_2019 = 88.531) 40
Tabela 3.9: Padronização e decomposição das taxas de prevalência segundo idade e faixas de renda (n_2013 = 60.191; n_2019 = 88.509) . . 42
Tabela 3.10: Padronização e decomposição das taxas de prevalência segundo idade, faixas de renda e escolaridade (n_2013 = 60.191; n_2019 = 88.509) 44
Tabela 3.11: Padronização e decomposição das taxas de prevalência segundo idade, sexo, faixas de renda e escolaridade (n_2013 = 60.191; n_2019 = 88.509) 45

1 INTRODUÇÃO

Demógrafos e outros cientistas sociais tradicionalmente utilizam a técnica de padronização direta para eliminar os efeitos composicionais em duas ou mais populações (Das Gupta, 1993). Efeitos como idade e sexo podem ser eliminados de indicadores como as taxas de mortalidade, fertilidade ou morbidade observadas em duas ou mais populações (ANDERSON; ROSENBERG, 1998). Outra técnica amplamente empregada é a análise de decomposição de taxas, que lida com a descoberta das contribuições aditivas das diferenças nos fatores composicionais, como idade, sexo e taxas específicas, para a diferença nas taxas gerais entre essas populações (SMALEN; BAVEL; LANCKER, 2024; TAYE *et al.*, 2024). Os temas da padronização e da decomposição estão interligados, pois ambos lidam com a separação dos efeitos da estrutura populacional daqueles associados a fatores específicos. Segundo Das Gupta (1991), tais técnicas compartilham uma base conceitual comum e não devem ser tratadas de forma independente.

A análise de decomposição de taxas é uma técnica estatística bastante utilizada em demografia e epidemiologia para entender como diferentes fatores contribuem para variações em indicadores populacionais. Essa abordagem tem sido fundamental, por exemplo, para interpretar mudanças ao longo do tempo (PRESTON; HEUVELINE; GUILLOT, 2001; ARRIAGA, 1984).

Em termos estatísticos, segundo Fonseca e Martins (2012), populações representam conjuntos de indivíduos com características em comum. Essas populações podem se referir a regiões distintas em um mesmo momento ou à mesma região em momentos diferentes no tempo. Como a população é dinâmica no tempo e no espaço, a sua composição muda, e isso pode influenciar diretamente as taxas observadas (Grupo de Foz, 2021).

Quando é identificada uma diferença não nula entre taxas calculadas para populações distintas, surge a necessidade de compreender quais fatores explicam essa mudança. Essas variações podem estar associadas, por exemplo, ao aumento do tamanho da população, envelhecimento populacional ou à maior exposição a fatores de risco (REIS *et al.*, 2022). O objetivo da análise de decomposição é justamente

quantificar a contribuição de cada um desses fatores para a variação observada nas taxas.

Nesse contexto, o desenvolvimento de métodos de decomposição remonta a trabalhos clássicos como o de Kitagawa (1955), que propôs uma abordagem para decompor diferenças entre duas taxas em componentes atribuíveis à estrutura populacional e às taxas específicas. Posteriormente, diversos autores expandiram essas métodos, incluindo Das Gupta, que introduziu técnicas mais gerais para decomposição envolvendo múltiplos fatores e múltiplas populações (Das Gupta, 1978).

Uma técnica mais abrangente e sistematicamente unificada de análise de decomposição foi proposta por Das Gupta (1993). O autor argumenta que, antes desse sistema unificado, abrangendo as duas áreas (padronização e decomposição) como partes do mesmo sistema consistente, a escolha arbitrária de populações-padrão levava a resultados indefensáveis. No mesmo estudo, são apresentadas metodologias para tratar problemas de padronização e decomposição envolvendo múltiplos fatores e populações, inclusive para taxas de dados classificados de forma cruzada.

No caso do método proposto por Das Gupta (1993), o grande diferencial está na formulação algébrica que permite decompor a diferença entre taxas agregadas, como a taxa bruta de mortalidade, morbidade ou fecundidade, em componentes atribuíveis a múltiplos fatores.

1.1 Revisão da Literatura

Desde sua publicação, a técnica de Das Gupta tem sido amplamente utilizada, especialmente, em estudos que avaliam a carga de doenças. Exemplo disso pode ser observados pelos estudos realizados pelo *Global Burden of Disease* (GBD), um estudo colaborativo que monitora as principais causas de mortalidade e morbidade global (Global Burden of Disease Collaborative Network, 2024). Burkart *et al.* (2022) demonstram que, além do aumento da exposição à poluição do ar ambiente, o crescimento populacional e o envelhecimento também contribuíram para o aumento expressivo da carga de diabetes tipo 2 atribuível à poluição do ar; ou e em Naghavi *et al.* (2024), que, ao decompor as contribuições específicas por causa e idade observa-se que, apesar dos efeitos da pandemia, houve progresso contínuo no combate a diversas causas de morte, resultando em melhorias na expectativa de vida global durante o período analisado.

Diversos trabalhos aplicam a decomposição para analisar dados em saúde, especialmente quando estão lidando com grandes bases de dados como o GBD. Por exemplo, Zhou *et al.* (2025) utilizaram o método de Das Gupta para investigar as mudanças na carga de cânceres do sistema nervoso central (SNC) entre 1990 e 2021, eles identificaram que o crescimento populacional e o envelhecimento foram

os principais fatores responsáveis pelo aumento da carga global desses cânceres. De modo semelhante, Zhang *et al.* (2025) aplicaram a decomposição para analisar a dor lombar, também destacando o envelhecimento populacional como principal determinante do aumento na prevalência da condição.

Outros estudos recentes utilizaram dados do Global Burden of Disease Collaborative Network (2020, 2024) (GBD) combinados com abordagens de decomposição de Das Gupta para identificar a contribuição de fatores como envelhecimento e crescimento populacional para a carga de doenças como, por exemplo, Dong *et al.* (2024) que avaliaram a evolução da doença renal crônica entre 1990 e 2019 e observaram que alterações na estrutura populacional explicam grande parte do aumento observado. Os mesmos fatores foram observados em Ghalichi *et al.* (2024) ao investigarem os determinantes do crescimento dos transtornos mentais no Irã, e por Khanmohammadi *et al.* (2023), ao examinarem doenças respiratórias, como câncer de traqueia, brônquio e pulmão, no norte da África, ambos destacando o papel das transformações inéditas no perfil etário e no tamanho da população.

Já Zhou, Zhang e Li (2025) avaliaram as variações de tendências da malária no arquipélago de Comores que, apesar de ser uma área endêmica de malária e de sofrido picos de aumento de casos, teve uma redução substancial no período observado (1990-2021). Da mesma forma, Yu *et al.* (2025) analisaram a evolução da doença de Alzheimer e outras demências, identificando o envelhecimento populacional como componente decisivo na variação temporal dessas condições.

Nota-se, então, que essa abordagem é empregada na investigação da influência de múltiplos fatores na saúde populacional, permitindo isolar a contribuição desses fatores ao longo do tempo.

Além dos estudos baseados em dados do GBD, há aplicações importantes do método de Das Gupta em contextos distintos, abordando fatores diversos. Um exemplo é o trabalho de Fanelli (2023), que analisou o aumento da taxa de nascimento fora do casamento na Itália entre 1999 e 2019. O estudo utilizou a decomposição para identificar a contribuição relativa de mudanças no comportamento reprodutivo (como o aumento da fecundidade entre mulheres solteiras) e transformações na composição populacional, como idade, estado civil e nacionalidade das mulheres em idade fértil. A análise permitiu evidenciar que o crescimento da fecundidade não marital, e não apenas alterações demográficas, constituiu um dos principais responsáveis pela mudança no perfil dos nascimentos italianos.

Outro exemplo é na análise da evolução do diabetes no Brasil entre 2013 e 2019, Reis *et al.* (2022) aplicaram o método de decomposição de Das Gupta para identificar a contribuição relativa de diferentes fatores no aumento do número de casos diagnosticados. Utilizando dados da Pesquisa Nacional de Saúde (PNS), os autores estimaram quanto do crescimento observado se devia ao aumento e envelhecimento

populacional ou a outros determinantes como a mudança na detecção de casos, aumento da incidência e na redução da mortalidade por diabetes. O estudo destaca ainda que a maior prevalência está associada à obesidade, hipertensão e menor nível educacional, reforçando a importância de políticas públicas voltadas para a prevenção e controle desses fatores de risco.

Outros autores têm usado a decomposição para investigar padrões menos usuais. Pandian e Allendorf (2022), por exemplo, estudaram o aumento de famílias com filhos de um único sexo em países da Ásia e do Norte da África, em resposta à queda da fertilidade e a normas culturais patrilineares. A partir de uma abordagem baseada em Kitagawa (1955), associada a princípios semelhantes aos de Das Gupta, os autores quantificaram a contribuição da queda da fecundidade para essas mudanças e demonstraram, com o auxílio de cenários hipotéticos, a influência de preferências por sexo dos filhos. Já no campo da saúde, Stucki (2021) aplicou uma versão expandida do método para decompor os custos hospitalares no cantão de Zurique entre 2013 e 2017. A análise investigou seis fatores, entre eles tempo médio de internação, custos por dia e prevalência de pacientes e demostrou que o principal responsável pelo aumento dos gastos foi o encarecimento do tratamento diário, com variações importantes por tipo de doença.

Outro estudo relevante é o de Li (2017), que analisou os gastos com saúde na China entre 1993 e 2012. A equipe utilizou a função `rdecompose` do software `Stata` para dividir o crescimento das despesas em cinco fatores distintos. No entanto, o trabalho não apresenta os detalhes algébricos da decomposição, o que limita a transparência do procedimento adotado. Situação semelhante é observada em Chevan e Sutherland (2009), que propõem uma extensão do método de Das Gupta para variáveis politômicas, mas também sem entrar na explicação operacional do método original.

Em resumo, a literatura demostra que o método de Das Gupta tem sido amplamente aplicado em contextos variados, tanto em estudos demográficos quanto em investigações epidemiológicas e sociais. Sua flexibilidade para lidar com múltiplos fatores e populações o destaca como ferramenta relevante às demandas atuais, especialmente com a crescente disponibilidade de bases de dados extensas e detalhadas.

1.2 Justificativa

A padronização e a decomposição de taxas são ferramentas fundamentais em demografia e epidemiologia, pois permitem distinguir variações reais nas condições de saúde das diferenças decorrentes da estrutura populacional. Apesar do uso consolidado do método de Das Gupta em estudos internacionais, sua aplicação prática com dados brasileiros ainda é limitada, sobretudo em análises de determinantes sociais

da saúde. O aprofundamento metodológico e a demonstração de sua implementação com dados nacionais podem contribuir para ampliar o uso dessa abordagem em pesquisas aplicadas e em avaliações de políticas públicas. Além disso, a compreensão dos fatores que explicam o aumento da prevalência de doenças crônicas, como o diabetes, é essencial para o planejamento em saúde. Assim, a aplicação do método de Das Gupta a dados da Pesquisa Nacional de Saúde (PNS) oferece uma oportunidade de integrar a análise estatística a um problema de relevância epidemiológica e social.

1.3 Objetivos

Objetivo geral

Aplicar o método de decomposição proposto por Das Gupta para quantificar a contribuição de fatores demográficos e socioeconômicos na variação das taxas de prevalência de diabetes no Brasil entre 2013 e 2019, utilizando dados da Pesquisa Nacional de Saúde (PNS).

Objetivos específicos

- Calcular taxas específicas e padronizadas de prevalência de diabetes segundo idade, sexo, raça/cor, escolaridade e renda.
- Implementar a decomposição das diferenças nas taxas de prevalência entre 2013 e 2019, estimando a contribuição relativa de cada fator.
- Avaliar a influência do envelhecimento populacional e das mudanças na estrutura social sobre as diferenças observadas.
- Elaborar código reprodutível em linguagem R para as etapas de padronização e decomposição, disponibilizando o procedimento para uso em futuras análises populacionais.

1.4 Organização do TCC

O restante do trabalho está organizado da seguinte forma: O Capítulo 2 descreve o material e os métodos utilizados, apresentando os conceitos fundamentais de taxas gerais e específicas, os dados da Pesquisa Nacional de Saúde (PNS), a metodologia empregada para cálculo das taxas, os procedimentos de padronização e decomposição, o método de Das Gupta e a implementação computacional utilizada. O Capítulo 3 apresenta e discute os resultados obtidos a partir das aplicações empíricas, incluindo as taxas padronizadas, as diferenças entre as populações comparadas e as decomposições

realizadas segundo diferentes fatores. O Capítulo 4 reúne as considerações finais, sintetizando os principais achados, destacando as contribuições do estudo e apontando limitações e possibilidades de desdobramentos futuros. Ao final, são apresentados os Apêndices, que reúnem o código em R utilizando, tabelas complementares, funções auxiliares e demais materiais de apoio.

2 MATERIAL E MÉTODOS

2.1 Pesquisa Nacional de Saúde

Foram utilizados dados da Pesquisa Nacional de Saúde (PNS), conduzida pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em parceria com o Ministério da Saúde, nos anos de 2013 e 2019. A PNS é uma pesquisa domiciliar de abrangência nacional que visa produzir informações sobre as condições de saúde, estilos de vida e acesso a serviços da população brasileira (Instituto Brasileiro de Geografia e Estatística (IBGE) (2015, 2020)).

A amostra da PNS é uma subamostra da Amostra Mestra do IBGE, que de acordo com Souza Júnior *et al.* (2022), utiliza um plano amostral probabilístico por conglomerados em três estágios de seleção, com estratificação das unidades primárias de amostragem (UPAs). No primeiro estágio, as UPAs são formadas por setores censitários ou por conjuntos de setores. No segundo estágio, são selecionados domicílios particulares permanentes dentro de cada UPA. No terceiro estágio, é selecionado um morador adulto (≥ 18 anos) de cada domicílio, a partir de uma lista de residentes elaborada no momento da entrevista.

Além disso, como explica Souza-Júnior *et al.* (2015), o IBGE calcula pesos amostrais considerando as probabilidades de seleção e ajustes por não resposta, o que garante a representatividade populacional das estimativas.

Apesar da representatividade garantida pelo delineamento amostral, algumas estratificações, como raça/cor, apresentam tamanhos reduzidos, o que resulta em intervalos de confiança mais amplos.

Em 2013, dos 81.167 domicílios selecionados, 11.173 estavam vazios, 5.646 não responderam à pesquisa e 4.146 indivíduos não concordaram em responder ao questionário individual, restando 60.202 residentes com 18 anos ou mais que responderam ao questionário individual. Isso corresponde a uma taxa de resposta de 86% do total de domicílios selecionados não vazios. Em 2019, dos 108.525 domicílios selecionados, 7.984 estavam vazios, 6.427 não responderam à pesquisa e 3.268 não concordaram em responder ao questionário individual, restando 90.846 residentes com 15 anos ou

mais como respondentes (aproximadamente 90% do total de domicílios selecionados não vazios). Para este estudo, incluímos respondentes com 18 anos ou mais, o que corresponde a 88.531 residentes. O desfecho de interesse nesta análise é a prevalência auto-referida de diabetes, obtida a partir da questão Q30: *"Algum médico já lhe deu o diagnóstico de diabetes?"*, cujas respostas possíveis são "Sim" ou "Não".

As análises desenvolvidas neste trabalho têm caráter exclusivamente ilustrativo e não têm como objetivo produzir interpretações causais ou epidemiológicas dos achados. Os dados da PNS foram utilizados apenas como exemplo para demonstrar a aplicação dos métodos de padronização e do procedimento de decomposição proposto por Das Gupta (1993). Assim, os valores apresentados servem para mostrar a operacionalização das técnicas, sem discussão quanto aos determinantes reais das diferenças observadas na prevalência de diabetes.

2.2 Taxas Gerais e Específicas

A demografia permite analisar eventos populacionais - como nascimentos, óbitos e migrações - considerando não apenas sua frequência, mas também sua distribuição entre diferentes grupos populacionais.

Na demografia, as taxas expressam a frequência de ocorrência de determinado evento em relação à população exposta ao risco, em um intervalo de tempo específico. Foram utilizadas, neste estudo, taxas de prevalência específicas e taxas gerais padronizadas, como base para as análises comparativas Grupo de Foz (2021), as taxas são expressas por mil habitantes. Por isso, a constante multiplicativa não é apresentada nas equações, embora seja aplicada no cálculo.

A Taxa Brutas (TB) é dada por:

$$\text{TB} = \frac{\text{nº total de pessoas com a doença}}{\text{nº total da população em risco}}.$$

Taxas específicas por idade e sexo:

$$_n\text{TP}_{x,s} = \frac{\text{nº total de pessoas com a doença com idade entre } x \text{ e } x+n \text{ do sexo } s}{\text{população em risco com idades entre } x \text{ e } x+n \text{ do sexo } s}.$$

Onde o denominador corresponde ao número de indivíduos expostos ao risco de apresentar o evento.

Este estudo aplicou métodos demográficos clássicos de padronização e decomposição de taxas para investigar diferenças em indicadores de morbidade entre populações. A abordagem considera tanto a estrutura etária das populações quanto variações nas taxas específicas por grupo.

2.3 Métodos de Padronização

A padronização é uma técnica demográfica empregada para eliminar o efeito de diferenças estruturais entre populações, como idade e sexo, de modo que as comparações reflitam apenas variações reais nas taxas específicas.

A padronização direta foi utilizada para eliminar o efeito da estrutura etária nas comparações de morbidade entre populações. Considera-se uma população-padrão comum a todos os grupos analisados. A técnica segue os princípios descritos em Grupo de Foz (2021), onde a taxa padronizada (TB_{pd}) é obtida ponderando-se as taxas específicas ($_nTP_x$) pela estrutura da população-padrão ($_nP_x$), conforme a expressão:

$$TB_{pd} = \frac{\Sigma_x \ {}_n\mathrm{TP}_x \ {}_n\mathrm{P}_x}{\Sigma_x \ {}_n\mathrm{P}_x},$$

A morbidade esperada foi estimada multiplicando-se as taxas específicas de 2019 pela população de 2010 correspondente a cada grupo etário.

$$P_{2019} = {}_n\mathrm{TP}_x \text{ de } 2019 \ \times \ \text{população de } 2010 \text{ com idade entre } x \text{ e } x+n$$

Neste trabalho, foram realizadas padronizações diretas considerando diferentes classificadores: idade, escolaridade, renda, raça/cor e a combinação idade–sexo. Todas as padronizações utilizaram como população padrão a distribuição do Censo Demográfico de 2010.

2.4 Decomposição e Método Das Gupta

A técnica de decomposição é usada para isolar os efeitos atribuíveis à estrutura populacional e às taxas específicas. A decomposição permite entender o quanto da diferença entre duas taxas padronizadas se deve à composição etária ou às taxas específicas.

Após a padronização, aplicou-se a técnica de decomposição, conforme o método geral proposto por Das Gupta (1991, 1993). Método que representa uma evolução em relação às técnicas anteriores de decomposição entre pares, como a de Kitagawa (1955). Sua principal contribuição está na formulação de um procedimento sistemático e simétrico que permite decompor a diferença entre taxas agregadas - como taxas brutas de mortalidade, morbidade ou fecundidade - considerando dois ou mais fatores simultaneamente e, quando necessário, mais de duas populações.

O objetivo central da técnica é quantificar a contribuição individual de cada fator (como idade, sexo ou qualquer outro classificador) para a diferença observada

em uma taxa agregada entre populações. Para isso, o método parte da seguinte estrutura: seja uma taxa agregada r, que depende de n fatores $x_1, x_2, ..., x_n$, podendo ser expressa como:

$$r = f(x_1, x_2, ..., x_n).$$

Dada a diferença entre duas taxas agregadas, $\Delta r = r^{(1)} - r^{(0)}$, onde $r^{(1)}$ e $r^{(0)}$ são, respectivamente, as taxas das populações 1 e 0, Das Gupta (1993) propõe que essa diferença pode ser decomposta na soma de contribuições marginais de cada fator x_i:

$$\Delta r = \sum_{i=1}^{n} C_i.$$

No entanto, quando há mais de dois fatores ou mais de duas populações, o método utiliza médias aritméticas das variações marginais entre fatores para garantir uma decomposição exata e simétrica, independente da escolha da população-padrão.

Se houver dois fatores α e β, a taxa F(α, β) neste caso é uma função desses fatores. Se os fatores assumirem valores A, B e a,b nas populações 1 e 2, respectivamente, então a diferença entre as taxas pode ser expressa como:

$$\begin{aligned} F(a,b) - F(A,B) &= \frac{1}{2}[F(a,b) - F(A,b) + F(a,B) - F(A,B)] \\ &\quad + \frac{1}{2}[F(a,b) - F(a,B) + F(A,b) - F(A,B)] \\ &= \alpha\text{-efeito} + \beta\text{-efeito}. \end{aligned}$$

Quando há três fatores α, β e γ , podemos usar uma notação semelhante para decompor a diferença como:

$$F(a,b,c) - F(A,B,C) = \alpha\text{-efeito} + \beta\text{-efeito} + \gamma\text{-efeito}$$

$$\begin{aligned} \alpha\text{-efeito} &= \frac{1}{3}[F(a,b,c) - F(A,b,c) + F(a,B,C) - F(A,B,C)] \\ &\quad + \frac{1}{6}[F(a,b,c) - F(a,B,c) + F(A,b,C) - F(A,B,C)], \end{aligned}$$

$$\begin{aligned} \beta\text{-efeito} &= \frac{1}{3}[F(a,b,c) - F(a,B,c) + F(A,b,C) - F(A,B,C)] \\ &\quad + \frac{1}{6}[F(a,b,C) - F(a,B,C) + F(A,b,c) - F(A,B,c)], \end{aligned}$$

$$\begin{aligned} \gamma\text{-efeito} &= \frac{1}{3}[F(a,b,c) - F(a,b,C) + F(A,B,c) - F(A,B,C)] \\ &\quad + \frac{1}{6}[F(a,B,c) - F(a,B,C) + F(A,b,c) - F(A,b,C)]. \end{aligned}$$

Quando há quatro fatores α, β, γ e δ, a diferença F(a,b,c,d)-F(A, B, C, D) pode

ser escrita como a soma dos quatro efeitos dos fatores, onde o efeito de α é dado por:

$$\begin{aligned}\alpha\text{-efeito} = \frac{1}{4}\Big[&F(a,b,c,d) - F(A,b,c,d) + F(a,B,C,D) - F(A,B,C,D)\Big] \\ &+ \frac{1}{12}\Big[F(a,b,c,D) - F(A,b,c,D) + F(a,b,C,D) - F(A,b,C,D) \\ &\quad + F(a,B,c,D) - F(A,B,c,D) + F(a,B,C,d) - F(A,B,C,d) \\ &\quad + F(A,b,C,d) - F(A,b,C,d) + F(a,B,c,D) - F(A,B,c,D)\Big],\end{aligned}$$

e os demais fatores (β, γ e δ) têm expressão semelhante.

Diferentemente da abordagem de modelagem estatística, o presente método decompõe a diferença entre duas taxas e não considera interações entre fatores, tratando-os como efeitos principais aditivos. Os efeitos dos fatores não implicam necessariamente quaisquer relações causais, eles simplesmente indicam a natureza da associação dos fatores com o fenômeno medido.

Para aplicação do método aos dados da PNS é preciso que as taxas sejam organizadas em formato de classificação cruzada, ou seja, estratificada por todos os fatores que forem utilizados na decomposição. Esse arranjo permite que o método produza, de forma simétrica, a decomposição da diferença entre as taxas globais nos componentes atribuíveis a cada fator.

A decomposição no caso de dados classificados de forma cruzada envolve um efeito adicional: o efeito das diferenças nas taxas específicas, chamado de efeito taxa. Em outras palavras, a decomposição vai fornecer um número de efeito a mais do que o número de fatores analisados.

Dessa forma, o método de Das Gupta permite isolar a contribuição de cada fator demográfico ou socioeconômico sobre a diferença total entre duas taxas agregadas, constituindo uma ferramenta confiável para análises comparativas, que não exige o uso de uma população padrão e sendo aplicável em diversos contextos.

Os fatores incluídos nas decomposições (idade, sexo, escolaridade, renda e raça/-cor) foram selecionados devido à sua disponibilidade na PNS e por sua relevância em estudos demográficos e epidemiológicos.

Nas aplicações empíricas, foi utilizado a notação semelhante ao manual de Das Gupta (1993), em que as letras I, J, K e L representam os efeitos dos fatores incluídos na decomposição e R representa o efeito das taxas específicas. Como a ordem dos fatores é arbitrária no método, cada letra corresponde a um fator diferente a depender do cenário analisado. Por esse motivo, em cada resultado apresentado no Capítulo 3 há uma indicação explícita (nas notas de rodapé) identificando qual fator está associado a cada símbolo.

2.5 Implementação Computacional

A implementação foi realizada no software e R Core Team (2025), utilizando scripts desenvolvidos pela autora para este trabalho. Para garantir a consistência das estimativas, considerando o desenho complexo da PNS, foram utilizados os pacotes `survey` (Lumley, 2024) e `srvyr` (FREEDMAN ELLIS; SCHNEIDER, 2024), responsáveis pelo delineamento amostral e pela estimação de prevalências ponderadas. O delineamento foi obtido por meio da função `svydesign`, incorporando pesos, estratos e unidades primárias, conforme recomendação do IBGE.

Todas as estimativas — taxas específicas, padronizações e tabelas cruzadas — foram calculadas respeitando o delineamento amostral complexo da PNS, com uso de pesos, estratos e unidades primárias.

Os intervalos de confiança de 95% foram estimados pelo método linearizado ('linearized variance'), conforme implementado no pacote `survey`, incorporando pesos, estratos e unidades primárias.

O pacote `PNSIBGE` (ASSUNCAO; HIDALGO, 2024) foi utilizado para o download e organização dos microdados da PNS 2013 e 2019, garantindo a compatibilidade entre as edições da pesquisa.

As padronizações direta por idade, escolaridade, raça/cor, faixas de renda e sexo foram implementados com a função `svystandardize`, conforme o método apresentado na Seção 2.3. O código completo encontra-se no Apêndice A.

Para o processo de decomposição, foi utilizado o pacote `DasGuptaR` (KING; MATTHEWS, 2025) que implementa o método geral de Das Gupta 1993 para dados com múltiplos fatores, populações e dados com classificações cruzadas. As funções `dgnpop()` e `dgcc()` foram utilizadas para calcular efeitos dos fatores e gerar tabelas detalhadas. O procedimento completo de decomposição, incluindo a geração das tabelas cruzadas, estão descritos nos Apêndices B, C, D e E.

Em resumo, o presente capítulo descreveu os dados, conceitos e métodos empregados na análise. Toda a implementação foi estruturada de acordo com a abordagem original de Das Gupta, a combinação das técnicas de padronização direta e de decomposição de Das Gupta possibilita quantificar, de forma consistente e reprodutível, as contribuições de diferentes fatores demográficos e socioeconômicos para as variações nas taxas de prevalência de diabetes entre 2013 e 2019.

3 RESULTADOS

Neste capítulo são apresentados os resultados descritivos, as taxas padronizadas e as decomposições segundo os fatores analisados, além das respectivas representações gráficas.

Os resultados deste capítulo têm o objetivo de demonstrar o funcionamento dos métodos apresentados no Capítulo 2, sem interpretação epidemiológica aprofundada.

3.1 Caracterização da amostra

A Tabela 3.1 apresenta os resultados brutos por sexo, idade, raça/cor, renda, índice de massa corporal (IMC) e hipertensão dos inquéritos da PNS nos anos de 2013 e 2019 com amostras de tamanho 60.202 e 88.531 indivíduos, respectivamente e as estimativas ponderadas conforme delineamento amostral complexo da PNS.

Tabela 3.1: Características da amostra, Brasil, PNS 2013 e 2019

Categoria	2013 n(%)	2019 n(%)
Sexo	**n = 60.202**	**n = 88.531**
Masculino	25.920 (47,1%)	41.662 (46,84%)
Feminino	34.282 (52,90%)	46.869 (53,16%)
Idade (anos)	**n = 60.202**	**n=88.531**
18-24	7.823 (15,93%)	8.145 (13,87%)
25-34	13.923 (21,65%)	15.970 (18,10%)
35-44	12.817 (19,16%)	18.033 (20,24%)
45-54	10.246 (17,52%)	15.885 (17,85%)
55-64	7.681 (13,45%)	14.572 (15,05%)
65-74	4.825 (7,73%)	9.965 (9,37%)
≥75	2.887 (4,55%)	5.961 (5,53%)
Raça/Cor da pele	**n = 60.199**	**n=88.522**
Branca	24.106 (47,56%)	32.409 (43,26%)
Preta	5.631 (9,13%)	10.132 (11,47%)
Amarela	533 (0,94%)	665 (0,92%)
Parda	29.512 (41,96%)	44.646 (43,80%)
Indígena	417 (0,42%)	670 (0,54%)
Escolaridade	**n = 60.202**	**n=88.531**
Analfabeto/Fundamental incompleto	24.083 (38,98%)	35.572 (34,76%)
Fundamental completo/Médio incompleto	9.215 (15,52%)	12.005 (14,48%)
Médio completo/Superior incompleto	19.149 (32,79%)	27.337 (34,94%)
Superior completo	7.755 (12,72%)	13.617 (15,83%)
Renda domiciliar per capita	**n = 60.191**	**n=88.509**
Até ¼ salário mínimo	4.873 (6,38%)	9.056 (7,71%)
Mais de ¼ até ½ salário mínimo	8.482 (12,70%)	13.482 (14,31%)
Mais de ½ até 1 salário mínimo	15.737 (27,57%)	25.758 (29,18%)
Mais de 1 até 2 salários mínimos	17.905 (31,63%)	22.160 (28,18%)
Mais de 2 até 3 salários mínimos	5.580 (9,57%)	7.515 (9,07%)
Mais de 3 até 5 salários mínimos	3.857 (6,43%)	5.490 (6,36%)
Mais de 5 salários mínimos	3.757 (5,70%)	5.048 (5,16%)
Índice de massa corpórea	**n = 59.402**	**n=87.678**
Baixo peso/Normal (< 25 kg/m^2)	25.446 (43,07%)	36.908 (42,03%)
Sobrepeso (entre 25 e 29,9 kg/m^2)	21.598 (36,13%)	32.744 (36,72%)
Obesidade (≥ 30 kg/m^2)	12.358 (20,80%)	18.026 (21,25%)
Hipertensão	**n = 60.202**	**n = 88.531**
Sim	12.500 (21,38%)	23.819 (25,40%)
Não	47.702 (78,62%)	64.712 (74,60%)

As amostras apresentaram composição semelhante entre 2013 e 2019, com discreta predominância feminina em ambos os inquéritos. Observou-se um envelhecimento populacional, com o aumento das faixas etárias acima de 55 anos e redução das mais novas.

A distribuição por raça/cor se manteve estável, com predominância de pessoas brancas e pardas. Não houve um avanço relevante nos níveis de escolaridade e renda per capita. A prevalência de obesidade e hipertensão aumentou, passando de 20,80% para 21,25% e de 21,38% para 25,40%, respectivamente. A tabela com a distribuição estimada com base no delineamento da PNS encontra-se no Apêndice F.

A Figura 3.1 ilustra visualmente o envelhecimento populacional observado entre os inquéritos.

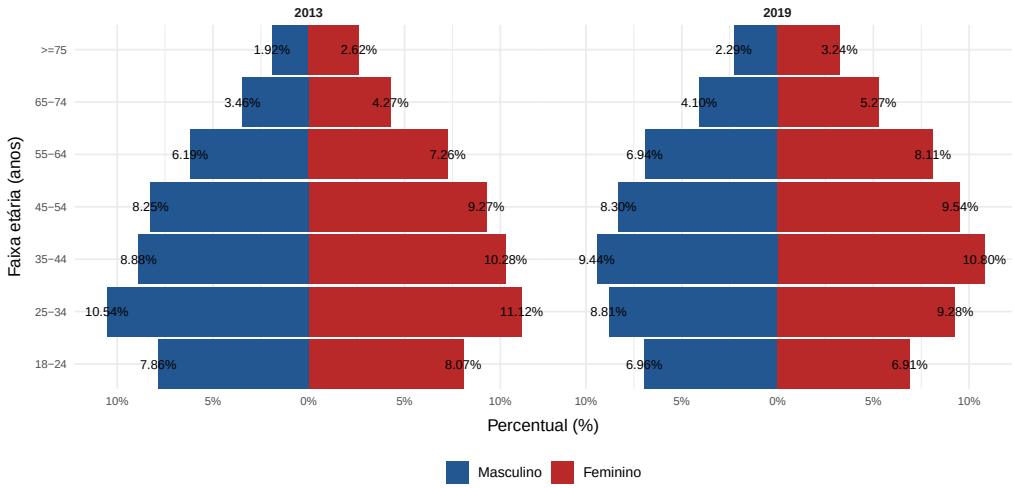


Figura 3.1: Pirâmide etária da população brasileira em 2013 e 2019

A pirâmide etária revela um claro processo de envelhecimento populacional entre 2013 e 2019. As faixas etárias mais altas aumentam visualmente, enquanto as faixas mais jovens diminuem, indicando uma redução da base da pirâmide e uma maior proporção de pessoas em idades avançadas.

O uso da população do Censo 2010 como padrão permite comparar as taxas entre 2013 e 2019 eliminando o efeito do envelhecimento.

3.2 Taxas Específicas

As Tabelas 3.2 a 3.5 apresentam as taxas específicas e as taxas padronizadas segundo os principais fatores demográficos e socioeconômicos.

A prevalência total de diabetes apresentada na Tabela 3.2 aumentou de 6,22% [5,9–6,6] em 2013 para 7,72% [7,4–8,0] em 2019. A doença foi mais frequente entre as mulheres (7,03% para 8,44%) e aumentou progressivamente com a idade, atingindo 21,12% entre indivíduos com 75 anos ou mais em 2019.

Observou-se maior prevalência entre indivíduos com menor escolaridade e renda e entre aqueles com obesidade ou hipertensão.

Tabela 3.2: Prevalência de diabetes - taxas específicas, Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	6,22 (5,88 - 6,56)	7,72 (7,43 - 8,02)
Sexo		
Masculino	5,31 (4,80 - 5,82)	6,91 (6,47 - 7,35)
Feminino	7,03 (6,55 - 7,52)	8,44 (8,03 - 8,85)
Idade (anos)		
18-24	0,53 (0,28 - 0,79)	0,73 (0,39 - 1,06)
25-34	0,82 (0,59 - 1,05)	0,91 (0,66 - 1,16)
35-44	2,93 (2,41 - 3,45)	3,15 (2,70 - 3,60)
45-54	6,60 (5,77 - 7,42)	7,72 (6,94 - 8,50)
55-64	13,50 (12,07 - 14,92)	14,76 (13,78 - 15,74)
65-74	19,94 (17,92 - 21,95)	21,93 (20,49 - 23,37)
≥75	19,49 (17,02 - 21,96)	21,12 (19,35 - 22,88)
Raça/Cor da pele		
Branca	6,69 (6,14 - 7,24)	8,05 (7,56 - 8,53)
Preta	7,27 (5,96 - 8,58)	7,84 (7,01 - 8,67)
Amarela	6,31 (3,00 - 9,63)	12,79 (7,93 - 17,66)
Parda	5,45 (5,03 - 5,88)	7,27 (6,86 - 7,68)
Indígena	6,93 (2,75 - 11,11)	7,49 (4,39 - 10,59)
Escolaridade		
Analfabeto/Fundamental incompleto	9,62 (8,96 - 10,28)	12,89 (12,30 - 13,48)
Fundamental completo/Médio incompleto	5,42 (4,52 - 6,32)	6,26 (5,48 - 7,04)
Médio completo/Superior incompleto	3,37 (2,93 - 3,81)	4,58 (4,17 - 5,00)
Superior completo	4,15 (3,28 - 5,02)	4,65 (4,10 - 5,20)
Renda		
Até ¼ salário mínimo	4,06 (3,01 - 5,12)	5,07 (4,23 - 5,91)
Mais de ¼ até ½ salário mínimo	4,63 (3,79 - 5,47)	6,03 (5,42 - 6,64)
Mais de ½ até 1 salário mínimo	5,89 (5,19 - 6,58)	8,38 (7,82 - 8,94)
Mais de 1 até 2 salários mínimos	7,48 (6,80 - 8,16)	8,21 (7,63 - 8,79)
Mais de 2 até 3 salários mínimos	7,23 (6,05 - 8,41)	8,60 (7,32 - 9,88)
Mais de 3 até 5 salários mínimos	5,56 (4,35 - 6,77)	9,06 (7,79 - 10,33)
Mais de 5 salários mínimos	5,93 (4,50 - 7,35)	6,87 (5,85 - 7,89)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m^2)	3,60 (3,23 - 3,97)	4,92 (4,57 - 5,28)
Sobrepeso (entre 25 e 29,9 kg/m^2)	6,57 (5,97 - 7,17)	8,31 (7,81 - 8,81)
Obesidade (≥ 30 kg/m^2)	11,31 (10,30 - 12,32)	12,56 (11,70 - 13,42)
Hipertensão		
Sim	18,23 (17,01 - 19,45)	19,95 (19,11 - 20,78)
Não	2,96 (2,68 - 3,23)	3,56 (3,32 - 3,80)

Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%

3.3 Taxas Padronizadas

Conforme apresentado no Capítulo 2, a padronização por idade permite eliminar o efeito das diferenças estruturais entre os anos.

Após a padronização por idade, apresentada na Tabela 3.3, a prevalência de diabetes em 2019 reduziu de 7,72% para 6,12%, e, em 2013, de 6,22% para 5,51%, sugerindo que parte do aumento observado se deve ao envelhecimento populacional. Mesmo após o ajuste, permanece um aumento residual, sugerindo influência de outros fatores além da idade. O código computacional foi gerado conforme os métodos apresentados nas Seções 2.3 e 2.5.

Tabela 3.3: Prevalência de diabetes padronizada por idade, Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	5,51 (5,22 - 5,80)	6,12 (5,88 - 6,36)
Sexo		
Masculino	4,68 (4,24 - 5,12)	5,47 (5,10 - 5,83)
Feminino	6,25 (5,82 - 6,68)	6,71 (6,37 - 7,05)
Raça/Cor da pele		
Branca	5,94 (5,46 - 6,42)	6,45 (6,04 - 6,87)
Preta	6,40 (5,28 - 7,53)	6,22 (5,55 - 6,90)
Amarela	5,49 (2,56 - 8,41)	10,20 (6,08 - 14,31)
Parda	4,82 (4,45 - 5,20)	5,71 (5,37 - 6,05)
Indígena	5,88 (2,38 - 9,38)	5,80 (3,25 - 8,34)
Escolaridade		
Analfabeto/Fundamental incompleto	8,83 (8,24 - 9,42)	11,20 (10,66 - 11,74)
Fundamental completo/Médio incompleto	4,71 (3,94 - 5,48)	4,89 (4,24 - 5,53)
Médio completo/Superior incompleto	2,98 (2,58 - 3,38)	3,57 (3,24 - 3,91)
Superior completo	3,69 (2,91 - 4,47)	3,69 (3,25 - 4,14)
Renda		
Até ¼ salário mínimo	3,59 (2,65 - 4,53)	4,06 (3,37 - 4,75)
Mais de ¼ até ½ salário mínimo	4,09 (3,37 - 4,82)	4,68 (4,18 - 5,17)
Mais de ½ até 1 salário mínimo	5,19 (4,59 - 5,79)	6,69 (6,22 - 7,16)
Mais de 1 até 2 salários mínimos	6,66 (6,05 - 7,26)	6,55 (6,08 - 7,03)
Mais de 2 até 3 salários mínimos	6,46 (5,39 - 7,52)	6,98 (5,86 - 8,09)
Mais de 3 até 5 salários mínimos	4,92 (3,84 - 6,01)	7,07 (6,08 - 8,07)
Mais de 5 salários mínimos	5,30 (4,02 - 6,59)	5,66 (4,76 - 6,56)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m^2)	3,12 (2,81 - 3,43)	3,77 (3,48 - 4,06)
Sobrepeso (entre 25 e 29,9 kg/m^2)	5,84 (5,32 - 6,37)	6,64 (6,23 - 7,05)
Obesidade (≥ 30 kg/m^2)	10,35 (9,41 - 11,30)	10,47 (9,73 - 11,21)
Hipertensão		
Sim	17,53 (16,35 - 18,70)	18,45 (17,65 - 19,25)
Não	2,62 (2,38 - 2,86)	2,82 (2,62 - 3,03)

a) Padronizada por idade pelo método direto, tendo como população de referência censo demográfico de 2010.
Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%

Como discutido no Capítulo 2, a padronização direta pode ser aplicada a diferentes classificadores, permitindo avaliar o impacto de fatores como sexo, renda, escolaridade e raça/cor.

Após o ajuste por escolaridade, apresentado na Tabela 3.4, observou-se um aumento da prevalência padronizada de 6,66% para 8,66%, indicando que a melhora no nível educacional contribuiu para atenuar o crescimento da prevalência.

Tabela 3.4: Prevalência de diabetes padronizada por escolaridade[a], Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	6,66 (6,29 - 7,03)	8,66 (8,33 - 8,99)
Sexo		
Masculino	5,51 (4,97 - 6,05)	7,43 (6,95 - 7,90)
Feminino	7,69 (7,16 - 8,23)	9,78 (9,29 - 10,26)
Idade (anos)		
18-24	0,56 (0,29 - 0,82)	0,80 (0,41 - 1,19)
25-34	0,83 (0,60 - 1,05)	0,94 (0,69 - 1,20)
35-44	3,02 (2,49 - 3,55)	3,35 (2,86 - 3,85)
45-54	6,68 (5,85 - 7,50)	8,19 (7,36 - 9,02)
55-64	13,93 (12,43 - 15,42)	15,28 (14,24 - 16,32)
65-74	20,17 (18,10 - 22,25)	22,71 (21,20 - 24,22)
≥75	19,54 (17,03 - 22,05)	21,13 (19,31 - 22,96)
Raça/Cor da pele		
Branca	7,24 (6,64 - 7,83)	9,21 (8,65 - 9,78)
Preta	7,81 (6,34 - 9,27)	8,76 (7,81 - 9,72)
Amarela	6,25 (3,25 - 9,25)	14,00 (8,54 - 19,46)
Parda	5,78 (5,33 - 6,24)	8,05 (7,59 - 8,51)
Indígena	7,43 (2,90 - 11,96)	8,39 (4,86 - 11,92)
Renda		
Até ¼ salário mínimo	4,15 (3,07 - 5,22)	5,24 (4,40 - 6,08)
Mais de ¼ até ½ salário mínimo	4,90 (3,99 - 5,81)	6,58 (5,91 - 7,25)
Mais de ½ até 1 salário mínimo	6,27 (5,50 - 7,03)	9,43 (8,81 - 10,04)
Mais de 1 até 2 salários mínimos	8,08 (7,36 - 8,81)	9,54 (8,84 - 10,23)
Mais de 2 até 3 salários mínimos	8,08 (6,75 - 9,42)	9,91 (8,31 - 11,51)
Mais de 3 até 5 salários mínimos	6,09 (4,75 - 7,42)	10,58 (9,04 - 12,11)
Mais de 5 salários mínimos	6,29 (4,82 - 7,75)	7,84 (6,66 - 9,02)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m^2)	3,85 (3,46 - 4,24)	5,63 (5,23 - 6,04)
Sobrepeso (entre 25 e 29,9 kg/m^2)	7,09 (6,43 - 7,75)	9,35 (8,77 - 9,93)
Obesidade (≥ 30 kg/m^2)	11,96 (10,89 - 13,03)	13,76 (12,82 - 14,70)
Hipertensão		
Sim	18,79 (17,51 - 20,07)	21,06 (20,18 - 21,94)
Não	3,15 (2,86 - 3,44)	3,92 (3,64 - 4,19)

a) Padronizada por escolaridade pelo método direto, tendo como população de referência censo demográfico de 2010.
Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%

Na Tabela 3.5, observa-se que a padronização conjunta por idade e sexo resulta em taxas ligeiramente menores, indicando que essa composição contribuiu para o aumento da prevalência.

Tabela 3.5: Prevalência de diabetes padronizada por idade e sexo, Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	5,50 (5,21 - 5,79)	6,11 (5,88 - 6,35)
Raça/Cor da pele		
Branca	5,94 (5,46 - 6,42)	6,45 (6,04 - 6,87)
Preta	6,40 (5,27 - 7,53)	6,21 (5,54 - 6,88)
Amarela	5,50 (2,56 - 8,44)	10,15 (6,03 - 14,26)
Parda	4,80 (4,43 - 5,18)	5,69 (5,35 - 6,03)
Indígena	5,89 (2,36 - 9,41)	5,79 (3,25 - 8,33)
Escolaridade		
Analfabeto/Fundamental incompleto	8,80 (8,22 - 9,38)	11,15 (10,62 - 11,69)
Fundamental completo/Médio incompleto	4,71 (3,94 - 5,48)	4,88 (4,23 - 5,52)
Médio completo/Superior incompleto	2,98 (2,58 - 3,38)	3,57 (3,24 - 3,91)
Superior completo	3,70 (2,91 - 4,48)	3,71 (3,26 - 4,16)
Renda		
Até ¼ salário mínimo	3,59 (2,64 - 4,54)	4,03 (3,36 - 4,71)
Mais de ¼ até ½ salário mínimo	4,08 (3,35 - 4,80)	4,66 (4,17 - 5,16)
Mais de ½ até 1 salário mínimo	5,18 (4,58 - 5,79)	6,68 (6,20 - 7,15)
Mais de 1 até 2 salários mínimos	6,64 (6,04 - 7,25)	6,53 (6,06 - 7,01)
Mais de 2 até 3 salários mínimos	6,44 (5,38 - 7,50)	6,98 (5,87 - 8,08)
Mais de 3 até 5 salários mínimos	4,93 (3,84 - 6,02)	7,10 (6,10 - 8,11)
Mais de 5 salários mínimos	5,32 (4,03 - 6,61)	5,68 (4,77 - 6,58)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m^2)	3,12 (2,81 - 3,43)	3,77 (3,48 - 4,06)
Sobrepeso (entre 25 e 29,9 kg/m^2)	5,83 (5,30 - 6,35)	6,63 (6,23 - 7,04)
Obesidade (≥ 30 kg/m^2)	10,36 (9,42 - 11,31)	10,43 (9,70 - 11,17)
Hipertensão		
Sim	17,54 (16,36 - 18,71)	18,43 (17,63 - 19,23)
Não	2,61 (2,37 - 2,86)	2,83 (2,62 - 3,03)

a) Padronizada conjunta por idade e sexo pelo método direto, tendo como população de referência censo demográfico de 2010.
Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%

A padronização apenas por sexo apresentou diferenças discretas, assim como as padronizações por raça/cor e por renda, o que sugere que as mudanças na estrutura populacional, segundo essas variáveis, sofreu efeito mínimo sobre a taxa global da prevalência de diabetes. Os resultados detalhados dessas padronizações encontram-se no Apêndice G.

3.4 Decomposição das diferenças

Antes de examinar os resultados, vale retomar que o método de Das Gupta permite separar a diferença total entre as prevalências de 2013 e 2019 em dois componentes: (i) parcela atribuível às mudanças na estrutura da população (por exemplo, variações na distribuição etária ou de sexo) (ii) parcela atribuível às diferenças nas taxas específicas dentro de cada grupo.

A seguir, apresentam-se os resultados dessa decomposição para os diferentes fatores considerados, permitindo quantificar a contribuição relativa de cada componente para a variação total da prevalência de diabetes no período.

A Tabela 3.6 apresenta os resultados da decomposição dos fatores idade (I) e sexo (J), esses resultados podem ser lidos da seguinte forma:

1. As taxas de prevalência de diabetes para 2013 e 2019 são de 6,22% e 7,72%, respectivamente. A diferença (Efeito total) é de 1,50 ponto percentual.

2. Se a estrutura por sexo (J) e as taxas de prevalência específicas por idade e sexo (R) fossem as mesmas para os anos de 2013 e 2019, as taxas de prevalência de diabetes seriam de 6,56% e 7,34%. A diferença, 0,78 p.p., apresenta a diferença das taxas de prevalência para o efeito idade (I).

3. Se a estrutura etária (I) e as taxas de prevalência específicas por idade e sexo (R) fossem as mesmas para os anos de 2013 e 2019, as taxas de prevalência de diabetes seriam de 6,9532% e 6,95804%. A diferença, 0,0048 p.p., apresenta a diferença das taxas de prevalência para o efeito sexo (J).

4. Se a estrutura etária (I) e a estrutura por sexo (J) fossem as mesmas para os anos de 2013 e 2019, as taxas de prevalência de diabetes seriam de 6,59% e 7,31%. A diferença, 0,72 p.p., apresenta a diferença das taxas de prevalência para o efeito taxa (R). O efeito taxa pode ser considerado a diferença real ou, ainda, pode ser composto por outros fatores não considerados nessa decomposição.

5. O efeito total (1) é a soma dos efeitos em (2), (3) e (4). Os efeitos podem ser expressados como porcentagens do efeito total. Assim, 51,93% do aumento das taxas de prevalência de diabetes para os anos de 2013 e 2019 pode ser explicado pela mudança na estrutura etária (I) das duas populações.

Ou, em outras palavras, a taxa bruta de prevalência de diabetes foi 1,50% maior em 2019 do que em 2013. Se apenas a estrutura etária mudar nos dois anos e as estruturas por sexo e taxas de prevalência especificas por idade e sexo fossem as mesmas para 2013 e 2019, então a taxa geral de prevalência de 2019 seria de apenas 0,78 p.p., maior do que em 2013.

Os dados de classificação cruzada usados para essa decomposição estão disponíveis no Apêndice C.

Tabela 3.6: Padronização e decomposição das taxas de prevalência segundo idade e sexo (n_2013 = 60.202; n_2019 = 88.531)

	Padronização		Decomposição	
Taxas de prevalência	2013 (%)	2019 (%)	Efeitos (p.p.)	Distribuição dos efeitos (%)
Taxas (J,R) padronizadas	6,565	7,346	0,78 (I)	51,93
Taxas (I,R) padronizadas	6,953	6,958	0,00 (J)	0,32
Taxas (I,J) padronizadas	6,597	7,314	0,72 (R)	47,75
Taxas brutas de prevalência de diabetes	6,221	7,724	1,50 (Efeito total)	100,00

(I) = efeito idade; (J) = efeito sexo; (R) = efeito taxa
A soma dos efeitos corresponde à diferença total entre os anos comparados.
Os resultados foram arredondados, podendo ocorrer pequenas diferenças decorrentes do arredondamento.

A Figura 3.2 representa graficamente a decomposição da variação absoluta na prevalência de diabetes entre 2013 e 2019, conforme os resultados da Tabela 3.6:

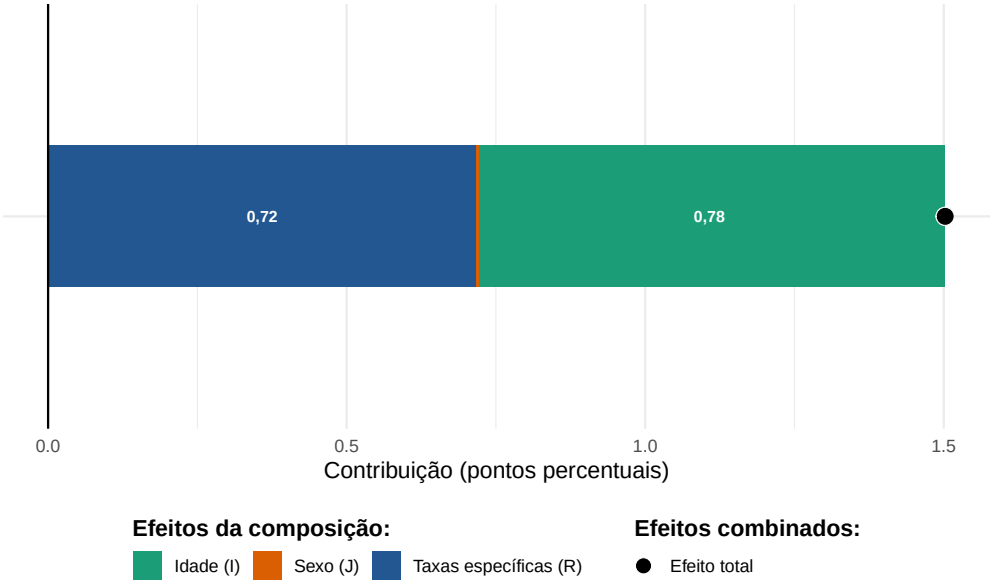


Figura 3.2: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e sexo

A Figura 3.2 apresenta como a diferença total é separada entre variações na estrutura populacional (idade e sexo) e mudanças nas taxas específicas. O objetivo é apenas ilustrar o mecanismo da decomposição.

A decomposição mostra que a idade é o fator que mais contribuiu para a variação total da prevalência de diabetes entre 2013 e 2019 (0,78 pontos percentuais, de um total de 1,50). O efeito do sexo apresenta magnitude muito pequena (0,0048 pontos percentuais) e, por isso, praticamente não altera o resultado final. Isso indica que a variação observada no período é explicada principalmente pelo envelhecimento da população, e não por diferenças entre homens e mulheres. Como todas as contribuições são positivas, o limite da representação gráfica dos fatores coincide com a variação total indicada pelo ponto preto.

A Tabela 3.7 apresenta os resultados da decomposição dos fatores idade (I) e raça/cor (J). A taxa bruta de prevalência de diabetes foi 1,50 p.p. maior em 2019 que em 2013. Se apenas a estrutura etária mudar nos dois anos enquanto a estrutura racial e taxas de prevalência especificas por idade e raça fossem as mesmas para 2013 e 2019, então a taxa geral de prevalência de 2019 seria de apenas 0,78 p.p. maior do que em 2013. A estrutura etária explica 53,16% do aumento na taxa de prevalência para o ano de 2019.

Tabela 3.7: Padronização e decomposição das taxas de prevalência segundo idade e raça/cor (n_2013 = 60.199; n_2019 = 88.522)

	Padronização		Decomposição	
Taxas de prevalência	2013 (%)	2019 (%)	Efeitos (p.p.)	Distribuição dos efeitos (%)
Taxas (J,R) padronizadas	6,555	7,354	0,80 (I)	53,11
Taxas (I,R) padronizadas	6,952	6,956	0,00 (J)	0,30
Taxas (I,J) padronizadas	6,603	7,304	0,70 (R)	46,59
Taxas brutas de prevalência de diabetes	6,220	7,724	1,50 (Efeito total)	100,00

(I) = efeito idade; (J) = efeito raça/cor; (R) = efeito taxa.
A soma dos efeitos corresponde à diferença total entre os anos comparados.
Os resultados foram arredondados, podendo ocorrer pequenas diferenças decorrentes do arredondamento.

A Figura 3.3 representa graficamente a decomposição da variação absoluta na prevalência de diabetes entre 2013 e 2019, conforme os resultados da Tabela 3.7:

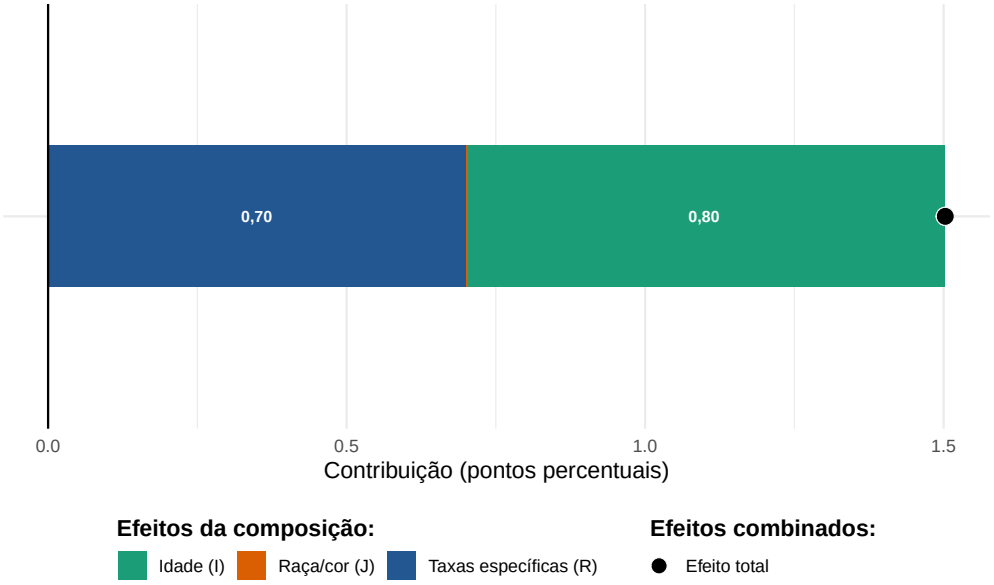


Figura 3.3: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e raça/cor

Os resultados da decomposição por idade (I) e escolaridade (J) estão apresentados na Tabela 3.8. A diferença negativa nos efeitos da escolaridade (J) de -0,22 p.p. indica que as mudanças na estrutura educacional da população contribuíram para reduzir o aumento da taxa de prevalência de diabetes na população em 2019. Em outra palavras, se distribuição por escolaridade tivesse permanecido igual à de 2013, o aumento na prevalência de diabetes teria sido ainda maior. No entanto, apesar de contribuir negativamente, essa contribuição não foi suficiente para reduzir a prevalência global, contribuindo com uma redução de apenas 14,53% da variação total.

Tabela 3.8: Padronização e decomposição das taxas de prevalência segundo idade e escolaridade (n_2013 = 60.202; n_2019 = 88.531)

	Padronização		Decomposição	
Taxas de prevalência	2013 (%)	2019 (%)	Efeitos (p.p.)	Distribuição dos efeitos (%)
Taxas (J,R) padronizadas	6,540	7,415	0,88 (I)	58,29
Taxas (I,R) padronizadas	7,087	6,869	-0,22 (J)	-14,53
Taxas (I,J) padronizadas	6,554	7,398	0,84 (R)	56,24
Taxas brutas de prevalência de diabetes	6,223	7,724	1,50 (Efeito total)	100,00

(I) = efeito idade; (J) = efeito escolaridade; (R) = efeito taxa.
A soma dos efeitos corresponde à diferença total entre os anos comparados.
Os resultados foram arredondados, podendo ocorrer pequenas diferenças decorrentes do arredondamento.

A Figura 3.4 representa graficamente a decomposição da variação absoluta na prevalência de diabetes entre 2013 e 2019, conforme os resultados da Tabela 3.8:

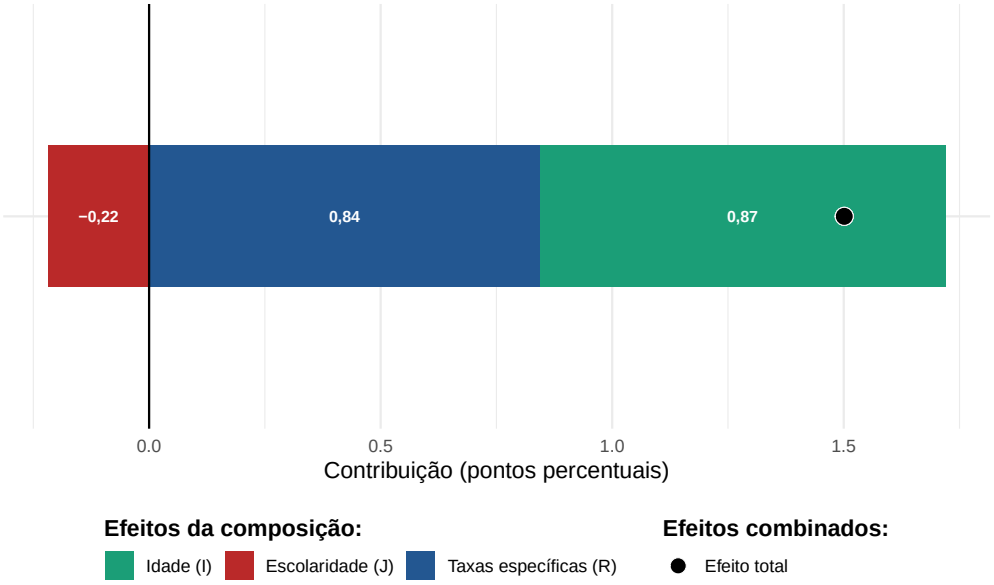


Figura 3.4: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e escolaridade

Diferentemente das decomposições anteriores, neste caso observa-se uma contribuição negativa associada à escolaridade, indicando que esse fator atuou no sentido de reduzir parte da variação total entre os anos analisados.

A decomposição absoluta mostra que a idade continua sendo o componente que mais contribuiu para o aumento da prevalência de diabetes entre 2013 e 2019 (0,88 pontos percentuais, de um total de 1,50). Já o efeito da escolaridade apresenta contribuição negativa (–0,22 pontos percentuais), indicando que a melhora no nível educacional da população atuou como um fator protetor, reduzindo parte do aumento que seria esperado apenas pelo envelhecimento populacional. Assim, a soma das contribuições positivas e negativas explica o efeito total representado pela linha de referência no gráfico.

Como o efeito da escolaridade contribui negativamente para a variação da prevalência, ele compensa parcialmente os efeitos positivos, resultando em uma soma das contribuições individuais superior à variação total observada.Isso significa que, se a estrutura educacional da população tivesse permanecido a mesma ao longo do período, a prevalência de diabetes teria aumentado ainda mais, ou seja, o cenário seria pior na ausência desse efeito.

A padronização por idade (I) e renda (J) apresentados na Tabela 3.9 indicam que a mudança na estrutura etária foi a principal responsável pelo aumento na taxa global, contribuindo com 57% enquanto a contribuição da mudança na estrutura por renda (J), apesar de agir no sentido contrário, foi muito pequena, contribuindo com a redução de apenas 0,6% da variação total.

Tabela 3.9: Padronização e decomposição das taxas de prevalência segundo idade e faixas de renda (n_2013 = 60.191; n_2019 = 88.509)

	Padronização		Decomposição	
Taxas de prevalência	2013 (%)	2019 (%)	Efeitos (p.p.)	Distribuição dos efeitos (%)
Taxas (J,R) padronizadas	6,511	7,367	0,86 (I)	57,09
Taxas (I,R) padronizadas	6,943	6,934	-0,01 (J)	-0,59
Taxas (I,J) padronizadas	6,628	7,280	0,65 (R)	43,50
Taxas brutas de prevalência de diabetes	6,223	7,723	1,50 (Efeito total)	100,00

(I) = efeito idade; (J) = efeito faixas de renda; (R) = efeito taxa.
A soma dos efeitos corresponde à diferença total entre os anos comparados.
Os resultados foram arredondados, podendo ocorrer pequenas diferenças decorrentes do arredondamento.

A Figura 3.5 representa graficamente a decomposição da variação absoluta na prevalência de diabetes entre 2013 e 2019, conforme os resultados da Tabela 3.9:

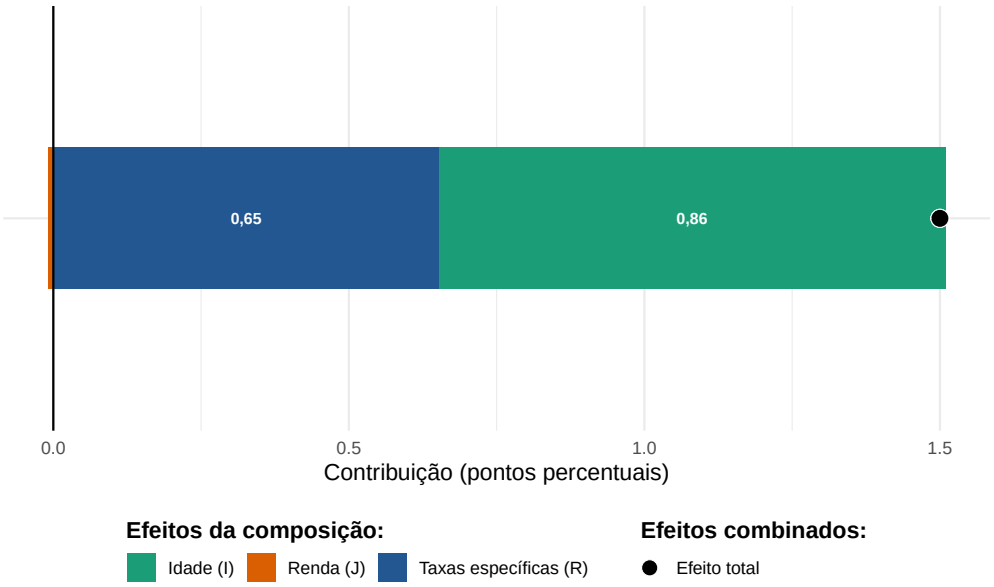


Figura 3.5: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade e renda

Como destacado no Capítulo 2, a atribuição de letras aos fatores (I, J, K, L) segue convenções arbitrárias e é definida de acordo com o conjunto de fatores analisado em cada decomposição.

A Tabela 3.10 traz resultados da decomposição considerando um terceiro fator: idade (I), renda (J) e escolaridade (K). A decomposição conjunta indica que a mudança na estrutura etaria é o principal responsável pelo aumento da prevalência de diabetes entre os anos de 2013 e 2019, correspondendo à 67% da variação global enquanto os efeitos de renda (J) e escolaridade (K) contribuem negativamente sendo responsáveis por aproximadamente 20% da redução total. Logo, o aumento global das taxas de prevalência de diabetes para o ano de 2019 em relação ao ano de 2013 deve-se, principalmente, pelo envelhecimento populacional e o aumento real das taxas especificas.

A Figura 3.6 representa graficamente a decomposição da variação absoluta na prevalência de diabetes entre 2013 e 2019, conforme os resultados da Tabela 3.10:

Tabela 3.10: Padronização e decomposição das taxas de prevalência segundo idade, faixas de renda e escolaridade (n_2013 = 60.191; n_2019 = 88.509)

	Padronização		Decomposição	
Taxas de prevalência	2013 (%)	2019 (%)	Efeitos (p.p.)	Distribuição dos efeitos (%)
Taxas (J,K,R) padronizadas	6,460	7,468	1,01 (I)	67,08
Taxas (I,K,R) padronizadas	6,984	6,937	-0,05 (J)	-3,11
Taxas (I,J,R) padronizadas	7,093	6,839	-0,25 (K)	-16,87
Taxas (I,J,K) padronizadas	6,590	7,384	0,79 (R)	52,90
Taxas brutas de prevalência de diabetes	6,222	7,724	1,50 (Efeito total)	100,00

(I) = efeito idade; (J) = efeito faixas de renda; (K) = efeito escolaridade (R) = efeito taxa.
A soma dos efeitos corresponde à diferença total entre os anos comparados.
Os resultados foram arredondados, podendo ocorrer pequenas diferenças decorrentes do arredondamento.

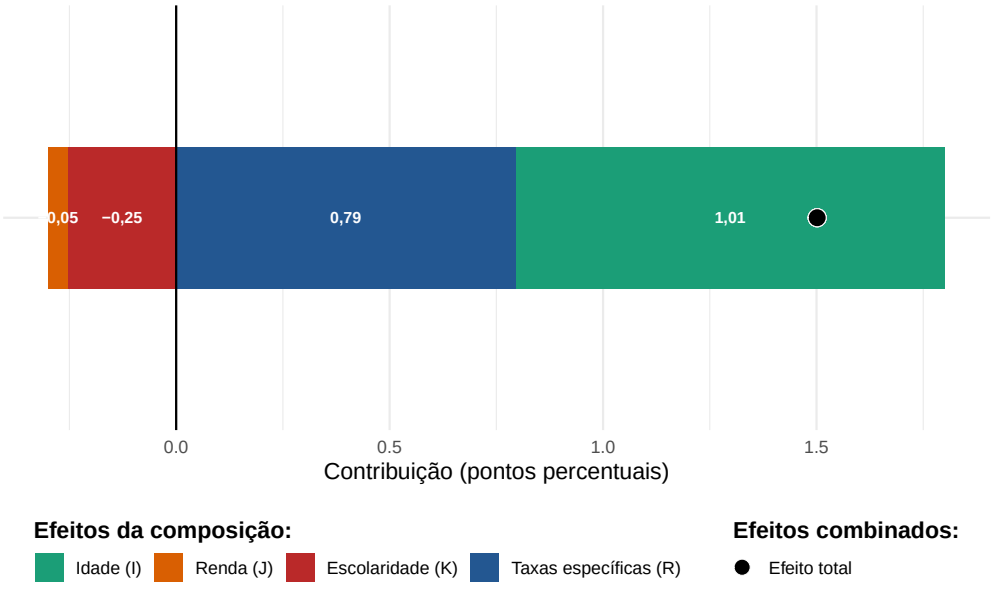


Figura 3.6: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade, renda e escolaridade

Tabela 3.11: Padronização e decomposição das taxas de prevalência segundo idade, sexo, faixas de renda e escolaridade (n_2013 = 60.191; n_2019 = 88.509)

	Padronização		Decomposição	
Taxas de prevalência	2013 (%)	2019 (%)	Efeitos (p.p.)	Distribuição dos efeitos (%)
Taxas (J,K,L,R) padronizadas	6,433	7,430	1,00 (I)	66,52
Taxas (I,K,L,R) padronizadas	6,934	6,940	0,01 (J)	0,36
Taxas (I,K,L,R) padronizadas	6,951	6,906	-0,05 (K)	-2,99
Taxas (I,J,K,R) padronizadas	7,062	6,807	-0,26 (L)	-16,98
Taxas (I,J,K,L) padronizadas	6,569	7,365	0,80 (R)	53,09
Taxas brutas de prevalência de diabetes	6,223	7,723	1,50 (Efeito total)	100,00

(I) = efeito idade; (J) = efeito sexo; (K) = efeito faixas de renda;
(L) = efeito escolaridade; (R) = efeito taxa.
A soma dos efeitos corresponde à diferença total entre os anos comparados.
Os resultados foram arredondados, podendo ocorrer pequenas diferenças decorrentes do arredondamento.

A Figura 3.7 representa graficamente a decomposição da variação absoluta na prevalência de diabetes entre 2013 e 2019, segundo os fatores idade, sexo, renda e escolaridade, conforme os os resultados da Tabela 3.11:

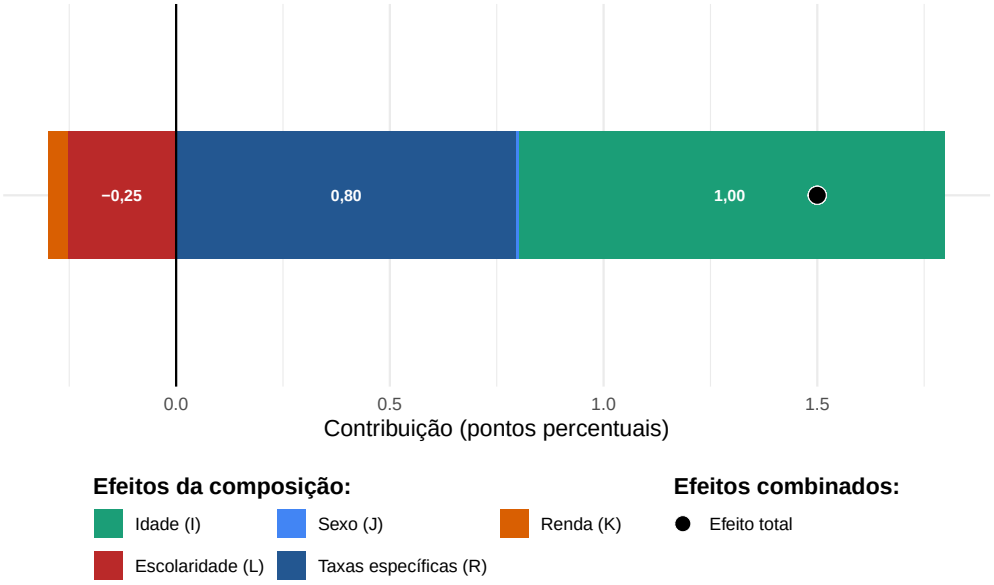


Figura 3.7: Decomposição da diferença total na prevalência de diabetes entre 2013 e 2019 segundo idade, sexo, renda e escolaridade

Ao comparar as diversas decomposições aplicadas é percebida uma pequena variação no efeito total, acredita-se que essas diferenças se deve ao fato de que alguns indicadores apresentam um tamanho de amostra diferente, devido à dados ausentes, gerando combinações de fatores sem população para algumas estratificações. Para esse estudo as simulações foram propostas em alguns cenários variando os fatores e verificando os seus comportamentos. Em uma análise em que a decomposição fosse apenas a ferramenta e não o objeto de estudo, é provável que uma única decomposição com fatores adequadamente escolhidos seja suficiente e então a comparação entre efeitos totais não será necessária.

De modo geral, os resultados evidenciaram o envelhecimento populacional como principal determinante do aumento da prevalência de diabetes no período analisado.

4 CONCLUSÃO E CONSIDERAÇÕES FINAIS

Com este trabalho foi possível abordar o tema da decomposição de taxas focando no método de Das Gupta voltado para a aplicação na avaliação dos determinantes sociodemográficos na prevalência do diabetes na população adulta brasileira, por meio dos dados da PNS das edições de 2013 e 2019.

O método de Das Gupta foi aplicado aos dados da Pesquisa Nacional de Saúde (PNS) quantificando a contribuição dos fatores demográficos e socioeconômicos nas variações de taxas de prevalência de diabetes no Brasil entre os anos de 2013 e 2019.

Taxas especificas e padronizadas foram calculadas possibilitando identificar que algumas mudanças de estruturas populacionais podem influenciar na variação de prevalência de diabetes.

A aplicação do método nos proporcionou conhecer a contribuição de cada fator na variação das taxas de prevalência alguns cenários, decompondo os fatores idade, sexo, faixas de renda, raça/cor e escolaridade, combinados dois à dois ou até os quatro fatores de uma só vez.

Os resultados mostraram que as mudanças na estrutura populacional impactam nas variações das taxas analisadas, especialmente a alteração na estrutura etária, que apresentou uma forte influência na prevalência de diabetes em todos os cenários avaliados. E, por outro lado, se observou que alguns fatores que agiram no sentido oposto, amenizando a variação total entre os anos estudados. A estrutura educacional é o exemplo mais evidente indicando que se a distribuição por escolaridade tivesse permanecido inalterada o aumento da taxa de prevalência teria sido ainda maior. Dessa forma, a decomposição evidencia tantos os fatores que colaboram para o aumento quanto para a redução da taxa analisada.

Neste trabalho, os códigos computacionais utilizados para gerar os resultados estão disponibilizados de forma organizada, sequencial e comentada, o que possibilita que as análises sejam reproduzidas integralmente, seja para os dados da PNS ou para outras fontes de dados com objetivos semelhantes.

Neste trabalho não foi possível garantir a veracidade das informações no sentido de ter valores fidedignos em relação ao diagnóstico de diabetes, por se tratar de dados autorreferidos estão sujeitos a erro de classificação ou viés de memória, gerando uma subnotificação que pode afetar as estimativas das prevalências. Por outro lado, a edição de 2013 contou com uma subamostra na segunda fase da pesquisa que coletava material biológico para exames laboratoriais. Se essa abordagem for repetida nas edições futuras, será possível uma análise mais confiável.

Outra limitação percebida é que a decomposição apresenta os efeitos apenas para os fatores incluídos explicitamente na análise, enquanto mudanças em outros determinantes sociais não considerados podem influenciar na diferença entre as taxas, mas não são isoladas pelo método. Esse é um prejuízo que pode ser minimizado pelo conhecimento prévio de quais fatores podem ter influência na variação pesquisada.

Este estudo apresenta fortalezas metodológicas importantes como a utilização dos dados da Pesquisa Nacional de Saúde (PNS), um inquérito populacional com amostragem probabilística e representatividade nacional, validos à toda a população brasileira. A comparação entre as edições de 2013 e 2019, com metodologias semelhantes, possibilitou examinar tendências relevantes.

Outro ponto forte deste trabalho está na aplicação de uma técnica confiável que permite separar as contribuições de diferentes fatores demográficos e socioeconômicos para a variação das taxas observadas, sem a dependência da escolha de uma população padrão. O uso da estrutura de classificação cruzada proporciona estimativas dos efeitos dos fatores controlado pelos demais fatores.

Em conjunto, essas características tornam este estudo metodologicamente sólido e capaz de contribuir com evidências relevantes para compreender as variações nas condições de saúde entre 2013 e 2019.

Por fim, a utilização da técnica de decomposição de Das Gupta aplicada aos dados da PNS de 2013 e 2019 permitiu identificar a contribuição relativa de diferentes fatores sociodemográficos para a variação observada nos dois períodos. Quando separados, o efeito das mudanças estruturais da população e o efeito das mudanças nas taxas específicas, possibilitaram a compreensão mais precisa dos determinantes da diferença total. E, apesar das limitações, os resultados oferecem uma visão metodológica consistente das transformações entre os períodos, contribuindo para o entendimento das desigualdades e para o planejamento de políticas públicas adequadas de acordo com as evidências informadas.

Os resultados reforçam a utilidade da decomposição como ferramenta didática e analítica para dados populacionais.

Referências Bibliográficas

ANDERSON, R. N.; ROSENBERG, H. M. Age standardization of death rates: implementation of the year 2000 standard. *National Vital Statistics Reports*, v. 47, n. 3, p. 1–16, 1998. Disponível em: https://pubmed.ncbi.nlm.nih.gov/9796247/. Acesso em: 27 abr. 2025.

ARRIAGA, E. E. Measuring and explaining the change in life expectancies. *Demography*, v. 21, n. 1, p. 83–96, fev. 1984. Disponível em: https://doi.org/10.2307/2061029. Acesso em: 26 jun. 2025.

ASSUNCAO, G.; HIDALGO, L. *PNSIBGE: Downloading, Reading and Analyzing PNS Microdata*. [S.l.], 2024. R package version 0.2.1. Disponível em: https://CRAN.R-project.org/package=PNSIBGE. Acesso em: 26 ago. 2025.

BURKART, K. *et al.* Estimates, trends, and drivers of the global burden of type 2 diabetes attributable to pm2 · 5 air pollution, 1990–2019: an analysis of data from the global burden of disease study 2019. *The Lancet Planetary Health*, v. 6, n. 7, p. e586–e600, 2022. Disponível em: https://doi.org/10.1016/S2542-5196(22)00122-X. Acesso em: 27 abr. 2025.

CHEVAN, A.; SUTHERLAND, M. Revisiting das gupta: Refinement and extension of standardization and decomposition. *Demography*, v. 46, n. 3, p. 429–449, 2009. Disponível em: https://doi.org/10.1353/dem.0.0060. Acesso em: 5 jun. 2025.

DAS GUPTA, P. A general method of decomposing a difference between two rates into several components. *Demography*, v. 15, p. 99–112, 1978.

DAS GUPTA, P. Decomposition of the difference between two rates and its consistency when more than two populations are involved. *Mathematical Population Studies*, v. 3, n. 2, p. 105–125, 1991. Disponível em: https://doi.org/10.1080/08898489109525329. Acesso em: 27 abr. 2025.

DAS GUPTA, P. *Standardization and Decomposition of Rates: A User's Manual*. Washington, D.C.: U.S. Bureau of the Census, 1993.

DONG, B. *et al.* Epidemiological analysis of chronic kidney disease from 1990 to 2019 and predictions to 2030 by bayesian age-period-cohort analysis. *Renal Failure*, v. 46, n. 2, 2024. Disponível em: https://doi.org/10.1080/0886022X.2024.2403645. Acesso em: 26 jun. 2025.

FANELLI, E. Italian nonmarital fertility ratio: components of an unexpected rise. *Population Research and Policy Review*, v. 42, n. 5, 2023. Disponível em: https://doi.org/10.1007/s11113-023-09825-8. Acesso em: 28 jun. 2025.

FONSECA, J. S. d.; MARTINS, G. d. A. *Curso de estatística*. 6. ed. Rio de Janeiro: Atlas, 2012. E-book. p. 111. Disponível em: https://app.minhabiblioteca.com.br/reader/books/9788522477937. Acesso em: 6 ago. 2025. ISBN 9788522477937.

FREEDMAN ELLIS, G.; SCHNEIDER, B. *srvyr: 'dplyr'-like syntax for summary statistics of survey data*. [S.l.], 2024. R package version 1.3.0. Disponível em: https://CRAN.R-project.org/package=srvyr. Acesso em: 26 ago. 2025.

GHALICHI, L. *et al.* National and subnational burden of mental disorders in iran (1990–2019): findings of the global burden of disease 2019 study. *The Lancet Global Health*, v. 12, n. 12, p. e1984–e1992, 2024. Disponível em: https://doi.org/10.1016/S2214-109X(24)00342-5. Acesso em: 26 jun. 2025.

GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK. *Global Burden of Disease Study 2019 (GBD 2019)*. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2020. Disponível em: https://ghdx.healthdata.org/series/global-burden-disease-gbd. Acesso em: 23 abr. 2025.

GLOBAL BURDEN OF DISEASE COLLABORATIVE NETWORK. *Global Burden of Disease Study 2021 (GBD 2021)*. Seattle, Estados Unidos: Institute for Health Metrics and Evaluation (IHME), 2024. Disponível em: https://www.healthdata.org/research-analysis/gbd. Acesso em: 23 abr. 2025.

GRUPO DE FOZ. *Métodos demográficos: uma visão desde os países de língua portuguesa*. São Paulo, Brasil: Blucher, 2021. Versão impressa. Também disponível em formato eletrônico: ISBN 978-65-5550-083-7. Open Access. ISBN 978-65-5550-082-0.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Pesquisa Nacional de Saúde: 2013 – Ciclos de vida: Brasil e grandes regiões*. Rio de Janeiro: IBGE, 2015. Disponível em: https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html?edicao=9166&t=publicacoes. Acesso em: 22 maio 2025.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Pesquisa Nacional de Saúde: 2019*. Rio de Janeiro: IBGE, 2020. Disponível em: https://www.ibge.gov.br/estatisticas/sociais/saude/9160-pesquisa-nacional-de-saude.html. Acesso em: 22 maio 2025.

KHANMOHAMMADI, S. *et al.* Burden of tracheal, bronchus, and lung cancer in north africa and middle east countries, 1990 to 2019: results from the gbd study 2019. *Frontiers in Oncology*, v. 12, 2023. Disponível em: https://doi.org/10.3389/fonc.2022.1098218. Acesso em: 26 ago. 2025.

KING, J.; MATTHEWS, B. *DasGuptR: Das Gupta standardisation and decomposition*. [S.l.], 2025. R package version 2.1.0, commit 7d4ac84ff9db7fed039d0bbb697ff737de32978a. Disponível em: https://github.com/josiahpjking/DasGuptR. Acesso em: 26 ago. 2025.

KITAGAWA, E. M. Components of a difference between two rates. *Journal of the American Statistical Association*, v. 50, p. 1168–1194, 1955.

LI, J. Rate decomposition for aggregate data using das gupta's method. *The Stata Journal*, v. 17, n. 2, p. 490–502, 2017. Disponível em: https://doi.org/10.1177/1536867X1701700213. Acesso em: 26 ago. 2025.

LUMLEY, T. *survey: analysis of complex survey samples*. [S.l.], 2024. R package version 4.4. Disponível em: https://CRAN.R-project.org/package=survey. Acesso em: 26 ago. 2025.

NAGHAVI, M. *et al.* Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the global burden of disease study 2021. *The Lancet*, v. 403, n. 10440, p. 2100–2132, 2024. Disponível em: https://doi.org/10.1016/S0140-6736(24)00367-2. Acesso em: 4 jun. 2025.

PANDIAN, R.; ALLENDORF, K. The rise of sonless families in asia and north africa. *Demography*, v. 59, n. 2, p. 761–786, 2022. Disponível em: https://doi.org/10.1215/00703370-9815547. Acesso em: 28 jun. 2025.

PRESTON, S. H.; HEUVELINE, P.; GUILLOT, M. *Demography: Measuring and Modeling Population Processes*. Malden, MA, USA; Oxford, UK; Carlton, Australia: Blackwell Publishing, 2001. Reprinted 2002, 2003, 2004. ISBN 1557864519.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2025. Disponível em: https://www.R-project.org/. Acesso em: 26 ago. 2025.

REIS, R. C. P. D. *et al.* Evolution of diabetes in brazil: prevalence data from the 2013 and 2019 brazilian national health survey. *Cadernos de Saúde Pública*, v. 38, n. 3, p. e00218021, 2022. Disponível em: https://doi.org/10.1590/0102-311X00218021. Acesso em: 12 jun. 2025.

SMALEN, D. de; BAVEL, J. V.; LANCKER, W. V. How much does immigration contribute to national poverty rates? a decomposition analysis for 17 european welfare states. *Socio-Economic Review*, v. 22, n. 2, p. 603–624, 2024. ISSN 1475-1461. Disponível em: https://doi.org/10.1093/ser/mwae012. Acesso em: 4 jun. 2025.

SOUZA-JÚNIOR, P. R. B. d. *et al.* Desenho da amostra da pesquisa nacional de saúde 2013. *Epidemiologia e Serviços de Saúde*, v. 24, n. 2, p. 207–216, 2015. Disponível em: https://doi.org/10.5123/S1679-49742015000200003. Acesso em: 26 ago. 2025.

SOUZA JÚNIOR, P. R. B. d. *et al.* Comparison of sampling designs from the two editions of the brazilian national health survey, 2013 and 2019. *Cadernos de Saúde Pública*, v. 38, p. e00164321, 2022. Disponível em: https://doi.org/10.1590/0102-311X00164321. Acesso em: 26 ago. 2025.

STUCKI, M. Factors related to the change in swiss inpatient costs by disease: a 6-factor decomposition. *European Journal of Health Economics*, v. 22, n. 2, p. 195–221, 2021. Disponível em: https://doi.org/10.1007/s10198-020-01243-3. Acesso em: 28 jun. 2025.

TAYE, B. A. *et al.* Exploring the determinants of fertility rates in ethiopia: a decomposition analysis using count regression models with a focus on urban and rural residence, based on the 2019 ethiopian demographic health survey. *Journal of Health, Population and Nutrition*, v. 43, p. 162, 2024. Disponível em: https://doi.org/10.1186/s41043-024-00659-4. Acesso em: 27 abr. 2025.

YU, D.-T. *et al.* Global mortality, prevalence and disability-adjusted life years of alzheimer's disease and other dementias in adults aged 60 years or older, and the impact of the covid-19 pandemic: A comprehensive analysis for the global burden of disease 2021. *BMC Psychiatry*, v. 25, n. 1, p. 503, 2025. Disponível em: https://doi.org/10.1186/s12888-025-06661-2. Acesso em: 28 jun. 2025.

ZHANG, Y. K. *et al.* Low back pain among the working-age population: from the global burden of disease study 2021. *BMC Musculoskeletal Disorders*, v. 26, p. 441, 2025. Disponível em: https://doi.org/10.1186/s12891-025-08704-x. Acesso em: 04 jun. 2025.

ZHOU, J. *et al.* The global, regional, and national brain and cns cancers burden and trends from 1990 to 2021. *Scientific Reports*, v. 15, p. 19228, 2025. Disponível em: https://doi.org/10.1038/s41598-025-04636-7. Acesso em: 28 jun. 2025.

ZHOU, S.; ZHANG, J.; LI, C. Burden of malaria in the comoros, 1990–2021: findings from the global burden of disease study 2021. *Frontiers in Public Health*, v. 13, 2025. Disponível em: https://doi.org/10.3389/fpubh.2025.1470021. Acesso em: 25 jun. 2025.

APÊNDICE A – CÓDIGO R: PADRONIZAÇÃO POR IDADE

Taxas específicas padronizadas por idade

```
library(survey)
library(readr)
library(dplyr)
library(srvyr)

# ---- Padronização por idade ----
# dist. etaria de 2010 (projeção de 2018 - tabela 7358)
# https://sidra.ibge.gov.br/tabela/7358#resultado
# "18 a 24 anos","25 a 34 anos","35 a 44 anos",
# "45 a 54 anos","55 a 64 anos","65 a 74 anos",
# "75 anos ou mais"

pns.svy <- read_rds(file = here::here("dados",
      "pns_svy.rds"))

pop_2010 <- c(24365799, 33290059,
      27259278,22269212,
      14984305,8698030,
      5573154)

padronizadas <- svystandardize(design = pns.svy,
           by = ~ idade_cat,
           over = ~ ano,
           population = pop_2010)

svyby(~ diabetes_binn, ~ ano,
```

```
        FUN = svymean, design = padronizadas, na.rm = TRUE,
        vartype = "ci")

    padronizadas.srvyr <- as_survey(padronizadas)

***
```

APÊNDICE B – FUNÇÃO PARA TABELAS CRUZADAS

Função para as tabelas cruzadas

```
library(dplyr)
library(srvyr)
library(readr)
library(here)

cruzar_variaveis <- function(dados, vars,
                         nome_arquivo = NULL) {

# --- checagem
if (length(vars) < 1)
        stop("Informe pelo menos uma variável em ‘vars‘.")

# --- converte para símbolos para uso no dplyr
vars_syms <- syms(vars)

# ---- Prevalência ----
prev <- dados %>%
   filter(!is.na(ano)) %>%
   filter(across(all_of(vars), ~ !is.na(.))) %>%
   group_by(ano, !!!vars_syms) %>%
   summarise(
      prev = 100 * survey_mean(diabetes_binn,
                           vartype = "ci", na.rm = TRUE),
      .groups = "drop"
)
```

```
# ---- População ponderada ----
pop <- dados %>%
   filter(!is.na(ano)) %>%
   filter(across(all_of(vars), ~ !is.na(.))) %>%
   group_by(ano, !!!vars_syms) %>%
   summarise(
   populacao = survey_total(na.rm = TRUE),
   .groups = "drop"
)

# ---- Junta as tabelas ----

tabela <- left_join(prev, pop, by = c("ano", vars)) %>%
mutate(
  # Formatar população com separador de milhar
  populacao = format(round(populacao), big.mark = ".",
                decimal.mark = ","),
  # Formata porcentagem com 2 casa decimal
  prev = formatC(prev, format = "f", digits = 2,
                decimal.mark = ","),
  prev_low  = formatC(prev_low, format = "f", digits = 2,
                decimal.mark = ","),
  prev_upp  = formatC(prev_upp, format = "f", digits = 2,
                decimal.mark = ","),
  IC = paste0("(", prev_low, "-",  prev_upp, ")"),
  tamanho = populacao
  ) %>%
  select(ano, all_of(vars), prev, IC, tamanho)

  # ---- Salvamento opcional ----
  if (!is.null(nome_arquivo)) {
  caminho <- here::here("taxas_cruzadas",
                 paste0("tabela_prev_",
                 nome_arquivo, ".csv"))
  write_csv(tabela, caminho)
  message("Tabela salva em: ", caminho)
  }
```

```
return(tabela)
}

tabela_idade_sexo<- cruzar_variaveis(pns.srvyr,
                     vars=c("idade_cat", "sexo"),
                            "idadeSexo")

# ----  Variações de taxas cruzadas ----
# idade e Raça/cor
tabela_idade_raca <- cruzar_variaveis(pns.srvyr,
                    vars=c("idade_cat",
                    "cor_raca_5"), "idade_raca")
# idade e Escolaridade
tabela_idade_esc <- cruzar_variaveis(pns.srvyr,
                    vars=c("idade_cat",
                    "escola_cat"), "idade_esc")
# idade e Faixas de renda
tabela_idade_renda <- cruzar_variaveis(pns.srvyr,
                    vars=c("idade_cat",
                    "renda_pc_cat"), "idade_renda")
# idade e IMC
tabela_idade_imc <- cruzar_variaveis(pns.srvyr,
                    vars=c("idade_cat",
                    "imc_cat"), "idade_imc")
# idade e Hipertensão
tabela_idade_hip <- cruzar_variaveis(pns.srvyr,
                     vars=c("idade_cat",
                     "hipertensao_bin" ), "idade_hip")
# idade, Faixas de renda e Escolaridade
tabela_idade_r_esc <- cruzar_variaveis(pns.srvyr,
                    vars=c("idade_cat","renda_pc_cat",
                    "escola_cat" ), "idade_r_esc")
# idade, Faixas de renda e IMC
tabela_idade_r_imc <- cruzar_variaveis(pns.srvyr,
                    vars=c("idade_cat",
                    "renda_pc_cat","imc_cat" ),
                    "idade_r_imc")
# idade, Raça/cor, Faixas de renda e IMC
```

```
      tabela_idade_rc_r_imc <- cruzar_variaveis(pns.srvyr,
                           vars=c("idade_cat",
                          "cor_raca_5",renda_pc_cat",
                          "imc_cat"), "idade_rc_r_imc")
***
```

APÊNDICE C – CLASSIFICAÇÃO CRUZADA: IDADE E SEXO

Tabela C.1: Classificação cruzada - população e taxas de prevalência de diabetes por idade e sexo: PNS, 2013 e 2019

Sexo	Idade	2013 (população 1)		2019 (população 2)	
j	i	(Tamanho) N_{ij}	(Taxa) T_{ij}	(Tamanho) n_{ij}	(Taxa) t_{ij}
1	1	11.438.717	0,41	11.078.323	1,04
1	2	15.341.755	0,73	14.028.276	0,67
1	3	12.933.046	2,43	15.025.496	3,12
1	4	12.010.612	5,77	13.216.281	6,95
1	5	9.008.738	12,08	11.038.815	13,58
1	6	5.036.208	18,22	6.518.963	20,00
1	7	2.799.451	16,8	3.646.545	20,63
2	1	11.748.060	0,65	10.993.765	0,41
2	2	16.180.822	0,90	14.776.025	1,14
2	3	14.963.815	3,36	17.196.219	3,17
2	4	13.500.785	7,33	15.191.321	8,39
2	5	10.567.213	14,70	12.914.559	15,77
2	6	6.222.523	21,32	8.390.231	23,43
2	7	3.820.466	21,46	5.156.494	21,46
$j = .$	$i = .$	145.572.211	6,221	159.171.313	7,724

Idade $i = 1, 2, ..., 7$ correspondem a 18-24, 25-34, 35-44, 45-54, 55-64, 65-74 e ≥ 75.
Sexo $j = 1, 2$ correspondem a Masculino e Feminino.

APÊNDICE D – CÓDIGO R: DECOMPOSIÇÃO POR IDADE E SEXO

Decomposição com pacote "DasGuptR"com a função "dgnpop"

```
library(readr)
library(here)
# ----  Decomposição idade e Sexo ----
# idade e sexo
idadesexo <-  read_csv(here("taxas_cruzadas",
                            "tabela_prev_idadeSexo.csv"),
                       locale = locale(decimal_mark = ",",
                                       grouping_mark = "."))

library(DasGuptR)
DG_idadesexo <- dgnpop(idadesexo,
                       pop = "ano", factors = c("prev"),
                       id_vars = c("idade_cat", "sexo"),
                       crossclassified = "tamanho"
) |>
  dg_table()

DG_idadesexo

#                       2013     2019        diff decomp
# idade_cat_struct 6.565479 7.345719 0.780240038  51.93
# prev             6.597181 7.314469 0.717288029  47.75
# sexo_struct      6.953210 6.957989 0.004779389   0.32
# crude            6.221390 7.723697 1.502307457 100.00
***
```

Decomposição com pacote "DasGuptR"detalhado com a função "dgcc"

```
DG_idadesexoH <- dgcc(idadesexo,
                      pop = "ano",
                      id_vars = c("idade_cat", "sexo"),
                      crossclassified = "tamanho")

DG_idadesexoH
#
# ano idade_cat      sexo  prev            IC  tamanho
# 1  2013     18-24 Masculino  0.41   (0,13-0,69) 11438717
# 2  2013     18-24  Feminino  0.65   (0,22-1,07) 11748060
# 3  2013     25-34 Masculino  0.73   (0,41-1,05) 15341755
# 4  2013     25-34  Feminino  0.90   (0,60-1,21) 16180822
# 5  2013     35-44 Masculino  2.43   (1,71-3,15) 12933046
# 6  2013     35-44  Feminino  3.36   (2,61-4,11) 14963815
# 7  2013     45-54 Masculino  5.77   (4,60-6,94) 12010612
# 8  2013     45-54  Feminino  7.33   (6,24-8,43) 13500785
# 9  2013     55-64 Masculino 12.08  (9,82-14,34)  9008738
# 10 2013     55-64  Feminino 14.70 (12,84-16,57) 10567213
# 11 2013     65-74 Masculino 18.22 (14,76-21,68)  5036208
# 12 2013     65-74  Feminino 21.32 (18,82-23,82)  6222523
# 13 2013      >=75 Masculino 16.80 (13,10-20,50)  2799451
# 14 2013      >=75  Feminino 21.46 (18,09-24,83)  3820466
# 15 2019     18-24 Masculino  1.04   (0,41-1,66) 11078323
# 16 2019     18-24  Feminino  0.41   (0,21-0,61) 10993765
# 17 2019     25-34 Masculino  0.67   (0,32-1,02) 14028276
# 18 2019     25-34  Feminino  1.14   (0,77-1,50) 14776025
# 19 2019     35-44 Masculino  3.12   (2,48-3,77) 15025496
# 20 2019     35-44  Feminino  3.17   (2,53-3,81) 17196219
# 21 2019     45-54 Masculino  6.95   (5,85-8,05) 13216281
# 22 2019     45-54  Feminino  8.39   (7,20-9,59) 15191321
# 23 2019     55-64 Masculino 13.58 (12,19-14,98) 11038815
# 24 2019     55-64  Feminino 15.77 (14,48-17,07) 12914559
# 25 2019     65-74 Masculino 20.00 (17,96-22,03)  6518963
# 26 2019     65-74  Feminino 23.43 (21,43-25,43)  8390231
# 27 2019      >=75 Masculino 20.63 (17,87-23,39)  3646545
# 28 2019      >=75  Feminino 21.46 (19,13-23,78)  5156494
```

```
#
# idade_cat_struct sexo_struct
# 1        0.16300736   0.4820495
# 2        0.15588642   0.5177015
# 3        0.22011353   0.4787953
# 4        0.21331243   0.5210818
# 5        0.19011940   0.4673001
# 6        0.19297624   0.5326722
# 7        0.17520562   0.4709108
# 8        0.17528778   0.5290892
# 9        0.13292044   0.4655794
# 10       0.13584594   0.5343617
# 11       0.07536937   0.4590186
# 12       0.07905566   0.5406984
# 13       0.04308847   0.4463066
# 14       0.04749955   0.5525205
# 15       0.14354717   0.4848580
# 16       0.13422382   0.5145790
# 17       0.18452990   0.4776093
# 18       0.17776331   0.5222166
# 19       0.20198757   0.4673469
# 20       0.20282685   0.5326510
# 21       0.17787210   0.4668062
# 22       0.17899863   0.5331888
# 23       0.14927274   0.4645978
# 24       0.15155064   0.5353737
# 25       0.09050077   0.4525447
# 26       0.09637152   0.5469661
# 27       0.05201074   0.4404775
# 28       0.05805347   0.5580351
***
```

APÊNDICE E– CÓDIGO R: TABELA CRUZADA PARA DADOS COM COMBINAÇÕES AUSENTES

Tabela cruzada para dados com combinações ausentes

```
# ----  Decomposição idade, Faixas de renda e Escolaridade ----
# idade, Faixas de renda e Escolaridade
library(readr)
library(here)
idaderesc <- read_csv(here("taxas_cruzadas",
                           "tabela_prev_idade_r_esc.csv"),
                      locale = locale(decimal_mark = ",",
                                      grouping_mark = "."))

library(DasGuptR)
DG_idaderesc <- dgnpop(idaderesc,
                       pop = "ano",
                       factors = c("prev"),
                       id_vars = c("idade_cat", "renda_pc_cat",
                       "escola_cat"),
                       crossclassified = "tamanho"
) |>
  dg_table()

# Error in data.frame(rate = c(popAi, popBi), pop = rep(pops,
# e = length(popAi)),:
# argumentos implicam em número de linhas distintos: 416, 408,
# 424, 1
continua ...
```

Tabela cruzada para dados com combinações ausentes (continuação)

```
# Além disso: Mensagen de aviso:
# In popAi - popBi :
# comprimento do objeto maior não é múltiplo do comprimento do
# objeto menor

# algumas combinações não tinham população e não foi gerada.
# Como por exemplo:
# pessoas de 18-24 anos, sem renda informado, escolaridade =
# Fundamental completo/Médio incompleto

library(dplyr)
idaderesc <- idaderesc %>%
  mutate(
    ano = as.factor(ano),
    idade_cat = as.factor(idade_cat),
    renda_pc_cat = as.factor(renda_pc_cat),
    escola_cat = as.factor(escola_cat)
  )

# criando todas as combinações, independente da  população
library(tidyr)
combinacoes <- expand_grid(
  ano = levels(idaderesc$ano),
  idade_cat = levels(idaderesc$idade_cat),
  renda_pc_cat = levels(idaderesc$renda_pc_cat),
  escola_cat = levels(idaderesc$escola_cat)
)

# atribuindo valor pequeno para quando não houver população
idaderesc_completo <- combinacoes %>%
  left_join(idaderesc, by = c("ano", "idade_cat",
  "renda_pc_cat", "escola_cat")) %>%
  mutate(
    prev = ifelse(is.na(prev), 1e-6, prev),
    tamanho = ifelse(is.na(tamanho), 1e-6, tamanho)
  )
continua ...
```

Tabela cruzada para dados com combinações ausentes (final)

```
DG_idaderesc <- dgnpop(idaderesc_completo,
                       pop = "ano",
                       factors = c("prev"),
                       id_vars = c("idade_cat",
                       "renda_pc_cat", "escola_cat"),
                       crossclassified = "tamanho"
) |>
  dg_table()

DG_idaderesc


#                          2013     2019       diff decomp
# escola_cat_struct   7.092629 6.839343 -0.2532863 -16.87
# idade_cat_struct    6.460311 7.467563  1.0072522  67.08
# prev                6.590048 7.384352  0.7943036  52.90
# renda_pc_cat_struct 6.984059 6.937299 -0.0467603  -3.11
# crude               6.222481 7.723990  1.5015092 100.00

***
```

APÊNDICE F – DISTRIBUIÇÃO POPULACIONAL ESTIMADA

Tabela F.1: Distribuição populacional estimada por faixa etária — PNS 2013 e 2019

Faixa etária	Ano	
	2013 N(%)	2019 N(%)
18-24	23.186.777 (15,93%)	22.072.088 (13,87%)
25-34	31.522.577 (21,65%)	28.804.301 (18,10%)
35-44	27.896.861 (19,16%)	32.221.715 (20,24%)
45-54	25.511.397 (17,52%)	28.407.601 (17,85%)
55-64	19.575.951 (13,45%)	23.953.373 (15,05%)
65-74	11.258.731 (7,73%)	14.909.195 (9,37%)
≥75	6.619.917 (4,55%)	8.803.038 (5,53%)
Total	145.572.210	159.171.311

APÊNDICE G – TAXAS PADRONIZADAS

Tabela G.1: Prevalência de diabetes padronizada por sexo, Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	6,21 (5,87 - 6,54)	7,70 (7,41 - 8,00)
Sexo		
Masculino	5,31 (4,80 - 5,82)	6,91 (6,47 - 7,35)
Feminino	7,03 (6,55 - 7,52)	8,44 (8,03 - 8,85)
Idade (anos)		
18-24	0,53 (0,28 - 0,78)	0,73 (0,40 - 1,07)
25-34	0,82 (0,59 - 1,04)	0,90 (0,65 - 1,15)
35-44	2,92 (2,41 - 3,44)	3,15 (2,70 - 3,60)
45-54	6,58 (5,76 - 7,40)	7,70 (6,93 - 8,47)
55-64	13,47 (12,04 - 14,90)	14,74 (13,75 - 15,72)
65-74	19,90 (17,88 - 21,93)	21,88 (20,45 - 23,32)
≥75	19,44 (16,98 - 21,90)	21,11 (19,34 - 22,87)
Raça/Cor da pele		
Branca	6,68 (6,13 - 7,22)	8,04 (7,55 - 8,53)
Preta	7,24 (5,93 - 8,55)	7,82 (6,99 - 8,64)
Amarela	6,33 (2,98 - 9,68)	12,81 (7,94 - 17,69)
Parda	5,43 (5,01 - 5,86)	7,24 (6,83 - 7,65)
Indígena	6,91 (2,71 - 11,10)	7,43 (4,34 - 10,52)
Escolaridade		
Analfabeto/Fundamental incompleto	9,56 (8,91 - 10,21)	12,83 (12,24 - 13,41)
Fundamental completo/Médio incompleto	5,42 (4,51 - 6,33)	6,23 (5,45 - 7,01)
Médio completo/Superior incompleto	3,37 (2,93 - 3,81)	4,58 (4,17 - 5,00)
Superior completo	4,17 (3,29 - 5,05)	4,68 (4,12 - 5,24)
Renda		
Até ¼ salário mínimo	4,07 (3,00 - 5,13)	5,04 (4,21 - 5,87)
Mais de ¼ até ½ salário mínimo	4,61 (3,77 - 5,45)	6,01 (5,40 - 6,62)
Mais de ½ até 1 salário mínimo	5,86 (5,16 - 6,56)	8,35 (7,80 - 8,91)
Mais de 1 até 2 salários mínimos	7,45 (6,78 - 8,13)	8,17 (7,60 - 8,75)
Mais de 2 até 3 salários mínimos	7,22 (6,04 - 8,40)	8,61 (7,33 - 9,88)
Mais de 3 até 5 salários mínimos	5,55 (4,34 - 6,75)	9,08 (7,80 - 10,36)
Mais de 5 salários mínimos	5,96 (4,52 - 7,40)	6,90 (5,87 - 7,92)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m2)	3,60 (3,23 - 3,96)	4,91 (4,56 - 5,27)
Sobrepeso (entre 25 e 29,9 kg/m2)	6,54 (5,94 - 7,14)	8,29 (7,79 - 8,80)
Obesidade (>= 30 kg/m2)	11,30 (10,28 - 12,31)	12,51 (11,66 - 13,37)
Hipertensão		
Sim	18,21 (16,98 - 19,43)	19,92 (19,09 - 20,76)
Não	2,95 (2,68 - 3,23)	3,57 (3,32 - 3,81)

Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%

Tabela G.2: Prevalência de diabetes padronizada por raça/cor, Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	6,22 (5,88 - 6,56)	7,76 (7,46 - 8,07)
Sexo		
Masculino	5,33 (4,82 - 5,84)	7,01 (6,55 - 7,47)
Feminino	7,02 (6,53 - 7,50)	8,42 (8,00 - 8,84)
Idade (anos)		
18-24	0,53 (0,27 - 0,78)	0,73 (0,38 - 1,07)
25-34	0,82 (0,59 - 1,05)	0,87 (0,62 - 1,12)
35-44	2,92 (2,40 - 3,44)	3,15 (2,69 - 3,62)
45-54	6,61 (5,77 - 7,44)	7,74 (6,93 - 8,55)
55-64	13,43 (12,02 - 14,84)	14,48 (13,51 - 15,45)
65-74	19,93 (17,90 - 21,96)	21,97 (20,49 - 23,45)
≥75	19,45 (16,98 - 21,93)	21,09 (19,29 - 22,90)
Raça/Cor da pele		
Branca	6,69 (6,14 - 7,24)	8,05 (7,56 - 8,53)
Preta	7,27 (5,96 - 8,58)	7,84 (7,01 - 8,67)
Amarela	6,31 (3,00 - 9,63)	12,79 (7,93 - 17,66)
Parda	5,45 (5,03 - 5,88)	7,27 (6,86 - 7,68)
Indígena	6,93 (2,75 - 11,11)	7,49 (4,39 - 10,59)
Escolaridade		
Analfabeto/Fundamental incompleto	9,62 (8,96 - 10,29)	13,03 (12,43 - 13,64)
Fundamental completo/Médio incompleto	5,43 (4,54 - 6,32)	6,34 (5,53 - 7,16)
Médio completo/Superior incompleto	3,39 (2,95 - 3,84)	4,65 (4,22 - 5,08)
Superior completo	4,17 (3,29 - 5,06)	4,72 (4,14 - 5,29)
Renda		
Até ¼ salário mínimo	4,12 (3,04 - 5,21)	5,24 (4,31 - 6,17)
Mais de ¼ até ½ salário mínimo	4,64 (3,79 - 5,48)	6,08 (5,45 - 6,72)
Mais de ½ até 1 salário mínimo	5,85 (5,16 - 6,54)	8,44 (7,86 - 9,02)
Mais de 1 até 2 salários mínimos	7,47 (6,79 - 8,16)	8,16 (7,57 - 8,74)
Mais de 2 até 3 salários mínimos	7,26 (6,07 - 8,46)	8,59 (7,32 - 9,86)
Mais de 3 até 5 salários mínimos	5,54 (4,33 - 6,75)	9,00 (7,71 - 10,30)
Mais de 5 salários mínimos	5,90 (4,48 - 7,33)	6,88 (5,84 - 7,92)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m2)	3,59 (3,22 - 3,96)	4,95 (4,58 - 5,31)
Sobrepeso (entre 25 e 29,9 kg/m2)	6,60 (5,99 - 7,20)	8,32 (7,80 - 8,84)
Obesidade (>= 30 kg/m2)	11,28 (10,27 - 12,28)	12,70 (11,81 - 13,59)
Hipertensão		
Sim	18,22 (17,00 - 19,45)	19,96 (19,10 - 20,81)
Não	2,96 (2,69 - 3,24)	3,62 (3,36 - 3,87)

Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%

Tabela G.3: Prevalência de diabetes padronizada por faixas de renda per capita, Brasil, PNS 2013 e 2019

Categoria	2013 %(IC95%)	2019 %(IC95%)
Total	6,14 (5,81 - 6,47)	7,70 (7,40 - 8,00)
Sexo		
Masculino	5,28 (4,77 - 5,79)	6,91 (6,47 - 7,35)
Feminino	6,90 (6,43 - 7,38)	8,39 (7,99 - 8,80)
Idade (anos)		
18-24	0,52 (0,27 - 0,77)	0,70 (0,39 - 1,01)
25-34	0,83 (0,61 - 1,06)	0,90 (0,65 - 1,15)
35-44	2,91 (2,40 - 3,41)	3,15 (2,70 - 3,61)
45-54	6,58 (5,76 - 7,41)	7,64 (6,87 - 8,41)
55-64	13,45 (12,03 - 14,87)	14,67 (13,69 - 15,66)
65-74	19,91 (17,90 - 21,91)	21,80 (20,36 - 23,23)
≥75	19,48 (17,01 - 21,94)	21,06 (19,30 - 22,82)
Raça/Cor da pele		
Branca	6,68 (6,13 - 7,22)	8,01 (7,52 - 8,50)
Preta	7,10 (5,83 - 8,36)	7,87 (7,02 - 8,71)
Amarela	6,09 (2,91 - 9,27)	12,76 (8,08 - 17,43)
Parda	5,33 (4,92 - 5,74)	7,23 (6,83 - 7,64)
Indígena	6,61 (2,66 - 10,56)	7,52 (4,40 - 10,63)
Escolaridade		
Analfabeto/Fundamental incompleto	9,38 (8,74 - 10,02)	12,87 (12,28 - 13,46)
Fundamental completo/Médio incompleto	5,37 (4,46 - 6,27)	6,31 (5,53 - 7,08)
Médio completo/Superior incompleto	3,35 (2,91 - 3,78)	4,65 (4,23 - 5,07)
Superior completo	4,18 (3,30 - 5,05)	4,67 (4,12 - 5,22)
Renda		
Até ¼ salário mínimo	4,06 (3,01 - 5,12)	5,07 (4,23 - 5,91)
Mais de ¼ até ½ salário mínimo	4,63 (3,79 - 5,47)	6,03 (5,42 - 6,64)
Mais de ½ até 1 salário mínimo	5,89 (5,19 - 6,58)	8,38 (7,82 - 8,94)
Mais de 1 até 2 salários mínimos	7,48 (6,80 - 8,16)	8,21 (7,63 - 8,79)
Mais de 2 até 3 salários mínimos	7,23 (6,05 - 8,41)	8,60 (7,32 - 9,88)
Mais de 3 até 5 salários mínimos	5,56 (4,35 - 6,77)	9,06 (7,79 - 10,33)
Mais de 5 salários mínimos	5,93 (4,50 - 7,35)	6,87 (5,85 - 7,89)
Índice de massa corpórea		
Baixo peso/Normal (< 25 kg/m2)	3,55 (3,19 - 3,92)	4,90 (4,55 - 5,25)
Sobrepeso (entre 25 e 29,9 kg/m2)	6,51 (5,90 - 7,11)	8,28 (7,78 - 8,78)
Obesidade (>= 30 kg/m2)	11,18 (10,19 - 12,17)	12,54 (11,68 - 13,41)
Hipertensão		
Sim	18,12 (16,90 - 19,34)	19,86 (19,03 - 20,70)
Não	2,92 (2,65 - 3,19)	3,57 (3,32 - 3,81)

Valores expressos em porcentagem.
IC95%: Intervalo de confiança de 95%