Imds
Instituto Mobilidade e Desenvolvimento Social

MAIO DE 2026

RELATÓRIO

Expectativa de Vida por Cor ou Raça no Brasil

Estimativas Regionais

Diretor-Presidente
Paulo Tafner

Diretor de Pesquisa
Fernando Veloso

Gerente Administrativo-Financeira
Carolina Roiter

Responsáveis Técnicos
Paulo Tafner
Sergio Guimarães Ferreira

Equipe de elaboração de conteúdo
Cássio Maldonado Turra
Laura Carvalho Andrade

Instituto Mobilidade e Desenvolvimento Social

Relatório; Expectativa de Vida por Cor ou Raça no Brasil: Estimativas Regionais

Rio de Janeiro, RJ, 2026. 35 p.

Imds
Instituto Mobilidade e Desenvolvimento Social

Sumário

1. INTRODUÇÃO **4**

2. BASE DE DADOS **5**

3. MÉTODOS E RESULTADOS PRINCIPAIS **7**

3.1. Método 1: Cedeplar/Imds fixo com ponderação pela composição racial 7

3.2. Método 2: Ajuste regional por raça a partir da mortalidade infantil do Atlas PNUD 10

3.3. Método 3: Ajuste regional por raça ancorado simultaneamente nas estimativas do Atlas/Pnud e no IBGE 16

3.4 Método 4: Ajuste vertical Brasil x Imds a partir das tabelas regionais por raça/cor 18

3.5. Método 5: Ajuste de segunda ordem por UF e sexo em torno do IBGE, com tolerância percentual 22

4. CONSISTÊNCIA GERAL E SÍNTESE DOS RESULTADOS **29**

5. CONCLUSÃO **33**

Apêndice 1 34

1. INTRODUÇÃO

Este terceiro relatório dá continuidade ao projeto “Expectativa de Vida por Cor ou Raça no Brasil: Diferenciais Regionais, Socioeconômicos e suas Interações com a Saúde”, desenvolvido em parceria entre o Cedeplar/UFMG e o Instituto de Mobilidade e Desenvolvimento Social (Imds). À semelhança dos dois primeiros relatórios, o foco permanece na quantificação e na interpretação das desigualdades de mortalidade entre brancos e negros no país, entendidas como expressão das assimetrias estruturais de acesso a oportunidades, recursos e bem-estar.

No primeiro relatório, concentramos a análise em dois eixos principais. Primeiro, construímos tabelas de vida anuais por sexo e raça/cor no Brasil entre 2005 e 2015, a partir da interpolação linear das funções lx e nLx das tabelas intercensitárias estimadas para 2000–2009 e 2010–2019. Em segundo lugar, discutimos preliminarmente os desafios de desagregar essas estimativas para os níveis subnacionais. Mostramos que a combinação de heterogeneidades na autodeclaração racial, diferenças regionais na composição dos grupos (em particular no papel da categoria “parda”) e a relevância dos fluxos migratórios inter-regionais impõem limitações substantivas ao uso direto, em nível regional ou estadual, do método intercensitário iterativo, que se mostrou adequado para o nível nacional. Destacamos, em especial, que a migração seletiva e a variação na cobertura censitária tendem a afetar de forma mais intensa as estimativas quando se introduzem simultaneamente as dimensões territoriais e raciais.

No segundo relatório, organizamos as estimativas regionais em torno de cinco métodos (M1 a M5), desde uma hipótese mais restritiva, que assumiu diferenciais raciais homogêneos em todas as regiões, até especificações que incorporaram informação adicional sobre a mortalidade infantil por raça/cor, índices de intensidade regional dos diferenciais raciais e calibrações específicas por região e sexo. Nosso objetivo foi explorar formas de decomposição da mortalidade total regional, impondo, em cada caso, um conjunto de restrições numéricas sobre as expectativas de vida ao nascer regionais, a consistência vertical com os padrões nacionais e a aderência aos níveis oficiais de mortalidade do IBGE.

O objetivo deste terceiro e último relatório é estender a aplicação dos Métodos 1 a 5, desenvolvidos na etapa anterior apenas para as grandes regiões, às unidades federativas, sempre reconhecendo as inúmeras dificuldades metodológicas já apontadas desde o primeiro relatório e reiteradas em nossas reuniões de pesquisa. Além disso,

apresentamos uma análise de sensibilidade dos resultados do Método 5, adotado desde o segundo relatório como o conjunto de estimativas mais consistente com as restrições e validações estabelecidas no projeto. Como enfatizado anteriormente, as estimativas aqui produzidas devem ser interpretadas como aproximações, ancoradas em hipóteses explícitas e em procedimentos de validação igualmente explícitos. Não são medidas observadas em sentido estrito, mas sim decomposições da mortalidade total, informadas pelos padrões nacionais de desigualdade racial e por indicadores regionais de mortalidade infantil. O propósito é oferecer, para cada UF e sexo, um conjunto de tabelas de vida por raça/cor que seja, ao mesmo tempo, comparável entre unidades, compatível com as estatísticas oficiais e transparente quanto às escolhas metodológicas envolvidas.

2. BASE DE DADOS

Tabelas Cedeplar/Imds por sexo e raça/cor

Assim, como no segundo relatório, o ponto de partida é o conjunto de tabelas de vida elaborado pela equipe do Cedeplar/UFMG para o Instituto de Mobilidade e Desenvolvimento Social (Imds). Utilizamos tabelas para o Brasil em 2015, desagregadas por sexo (homens, mulheres e ambos os sexos) e por raça/cor (brancos e negros). As estimativas contêm probabilidades de morte em grupos quinquenais de idade (q_x), número de sobreviventes (l_x), óbitos (d_x), pessoas-ano (L_x)e a expectativa de vida a partir de cada idade (e_x).

Atlas do Desenvolvimento Humano no Brasil (PNUD–IPEA–FJP)

A segunda fonte é o Atlas do Desenvolvimento Humano no Brasil, desenvolvido pela parceria entre o Programa das Nações Unidas para o Desenvolvimento (PNUD), o Instituto de Pesquisa Econômica Aplicada (IPEA) e a Fundação João Pinheiro (FJP). O objetivo central do Atlas é oferecer um conjunto abrangente de indicadores socioeconômicos e demográficos em diferentes níveis geográficos, com foco na mensuração do desenvolvimento humano em perspectiva subnacional. As estimativas resultam de um trabalho de harmonização e modelagem demográfica que combina, principalmente, informações das PNADs, dos Censos Demográficos do IBGE e de sistemas de estatísticas vitais, com correções para sub-registro e inconsistências, conforme documentado pelos autores do Atlas.

Para nossa aplicação, interessa especificamente o módulo que disponibiliza estimativas de mortalidade infantil por raça/cor, por Unidade da Federação e para o Brasil como um

todo. Para tal, os autores utilizam métodos indiretos, com base em perguntas sobre a sobrevivência dos filhos nascidos vivos em pesquisas domiciliares. Para o ano de 2015, as informações baseiam-se nos dados da PNAD. Ao utilizar o Atlas/PNUD como referência, ancoramos explicitamente nossas estimativas em uma síntese que já incorpora o esforço metodológico conjunto do PNUD, do IPEA e da FJP

População de 2015 por UF, sexo, raça/cor e idade

A terceira base é a distribuição da população residente por Unidade da Federação, sexo, raça/cor (branca e negra) e grupos etários quinquenais de 0–4 a 80 anos e mais. A partir dessa base, derivamos três estruturas principais: i) a população estadual por sexo, raça/cor e idade, utilizada para ponderar mortalidade específica por raça/cor e para obter tabelas agregadas "todas as raças" em cada UF; ii) a população infantil de 0–4 anos por UF, necessária para utilizar as estimativas de mortalidade infantil do Atlas/PNUD como referência de nível nas idades jovens; e iii) agregações regionais e nacionais (obtidas a partir das UFs), empregadas como checagens de consistência e como base para validações verticais com as referências nacionais. Em todas as etapas, essas populações funcionam como pesos que traduzem, em termos demográficos, a composição racial observada em cada unidade territorial.

Como na etapa regional, embora o ideal fosse utilizar a composição racial dos nascidos vivos, optamos por não incorporar dados do DATASUS, dada a possibilidade de introduzir novos vieses. Assim, empregamos a distribuição observada no grupo etário de 0 a 4 anos como proxy para nossas estimativas nas idades iniciais

Tabelas de Vida Regionais do IBGE (2015)

Por fim, no Método 5, utilizamos as tabelas de vida por Unidade da Federação e sexo produzidos pelo IBGE para 2015, que fornecem, para cada UF e sexo, a expectativa de vida ao nascer e os perfis de mortalidade por idade. Essas tabelas não entram diretamente na decomposição por raça/cor, mas servem como referência externa para controlar o nível final de mortalidade agregado em cada UF (todas as raças), garantindo compatibilidade com as estatísticas oficiais. Em termos práticos, comparamos a expectativa de vida resultante dos nossos modelos por UF e sexo com os valores do IBGE, definimos alvos dentro de tolerâncias previamente estabelecidas e, a partir daí, ajustamos parâmetros de nível para assegurar maior consistência com a mortalidade oficialmente estimada para cada UF.

Com essas bases de dados, passamos a dispor simultaneamente de um padrão nacional de mortalidade por idade, sexo e raça/cor (Cedeplar/Imds), de referências subnacionais

para o nível de mortalidade nas idades iniciais (Atlas/PNUD) e de informações de composição populacional por UF, sexo, raça/cor e idade para agregações e ponderações. A seção seguinte descreve como esses insumos são combinados nos cinco métodos (M1 a M5), agora aplicados sistematicamente ao nível das Unidades da Federação, preservando a mesma lógica geral da etapa regional e explicitando, quando necessário, os ajustes operacionais e as validações adicionais requeridas pelo maior grau de desagregação territorial.

3. MÉTODOS E RESULTADOS PRINCIPAIS

3.1. Método 1: Cedeplar/Imds fixo com ponderação pela composição racial

O Método 1 é o ponto de partida da estratégia de estimação. Ele assume deliberadamente que as diferenças espaciais de mortalidade decorrem apenas da composição por raça/cor e sexo, mantendo fixos, em todo o território, os perfis de mortalidade estimados pelo Cedeplar/Imds para 2015. Em outras palavras, as tabelas por raça/cor e sexo são nacionais; o que varia entre Unidades da Federação (e, por agregação, entre grandes regiões) é exclusivamente a composição desses grupos na população.

Nesse método, não se utiliza nenhuma informação externa sobre o nível de mortalidade (como Atlas/PNUD ou IBGE): todo o nível por raça/cor e sexo é o implícito nas tabelas nacionais Cedeplar/Imds. Assim, para cada unidade geográfica g (UF ou região), replicamos o vetor de probabilidades de morte do Imds por idade, sexo e raça/cor. Em termos práticos, para toda idade x, sexo s e raça r, tem-se $q_{x,g}^{(r,s)} = q_{x,IMDS}^{(r,s)}$. A partir desse vetor e utilizando os parâmetros demográficos derivados das próprias tabelas do Imds, em particular, o tamanho do do intervalo etário n_x, o tempo médio de contribuição dos mortos a_x e a condição do grupo aberto (via e_ω), reconstruímos, para cada combinação $g \times r \times s$, a tabela de vida completa $(q_x, l_x, d_x, L_x, T_x, e_x)$, garantindo que, por construção, as tabelas subnacionais específicas por raça/cor reproduzam exatamente o padrão nacional do Imds. O mesmo procedimento é aplicado para ambos os sexos. Isto é, replicamos diretamente as probabilidades q_x do Imds para ambos os sexos, por raça/cor, para cada UF e para cada região, reconstruindo as respectivas tabelas completas com os parâmetros correspondentes.

A etapa seguinte introduz o elemento territorial do método, mas apenas por meio da composição racial. Para obter tabelas para todas as raças em cada UF (e, analogamente, em cada região), combinamos as taxas específicas de mortalidade m_x das tabelas Imds

por raça/cor, ponderando-as pela distribuição populacional observada em 2015 por idade, sexo e raça/cor. Para cada unidade g, sexo s e idade x, definimos a taxa agregada como uma média ponderada: $m_{x,g}^{(tot,s)} = \frac{\sum_r m_{x,IMDS}^{(r,s)} P_{x,g}^{(r,s)}}{\sum_r P_{x,g}^{(r,s)}}$.

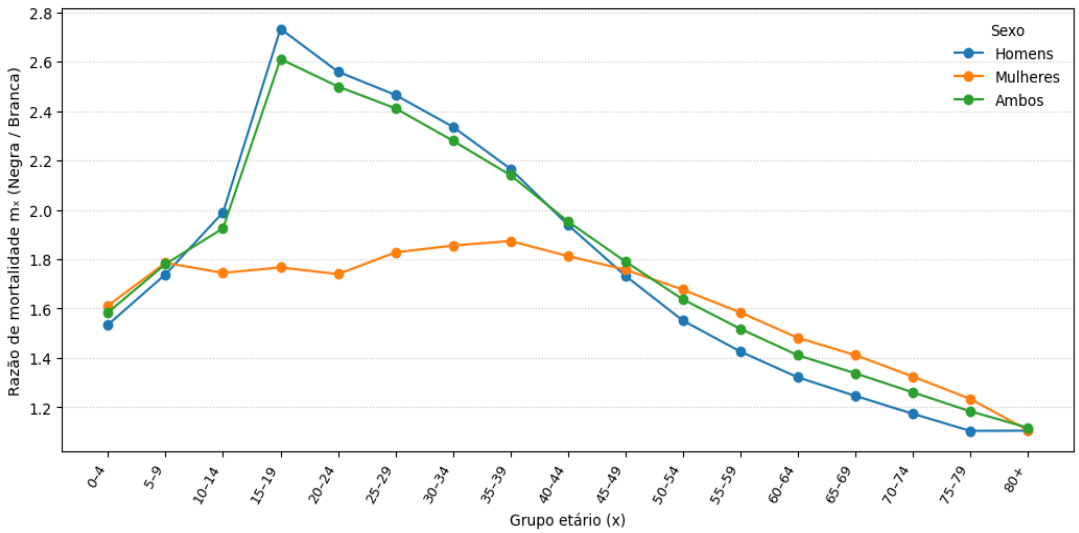
No procedimento operacional, além de m_x, calculamos por ponderação populacional os parâmetros necessários à reconstrução da tabela agregada (em particular, a_x e a condição do grupo aberto via e_ω), e então reconstruímos a tabela para todas as raças a partir do vetor $m_{x,g}^{(tot,s)}$. Esse mesmo raciocínio é aplicado às tabelas para todas as raças e ambos os sexos.

Como produto do Método 1, geramos um conjunto completo de tabelas por UF e por região, cobrindo: (i) sexo × raça/cor; (ii) ambos os sexos ×raça/cor; (iii) sexo × todas as raças; além de (iv) ambos os sexos × todas as raças.

Padrões principais observados (UFs e regiões) a partir do Método 1

A Figura 1 apresenta a razão das taxas de mortalidade entre a população negra e a branca (Negra/Branca), por grupo etário e sexo, derivada das tabelas nacionais Cedeplar/Imds. Essas funções foram amplamente discutidas na etapa anterior deste projeto. No escopo do Método 1, esse perfil é fixo e nacional: a curva não varia entre regiões nem entre UFs, pois o método não incorpora qualquer informação territorial sobre o nível ou o padrão etário de mortalidade. Ele apenas replica, para cada unidade geográfica, as tabelas nacionais por sexo e raça/cor.

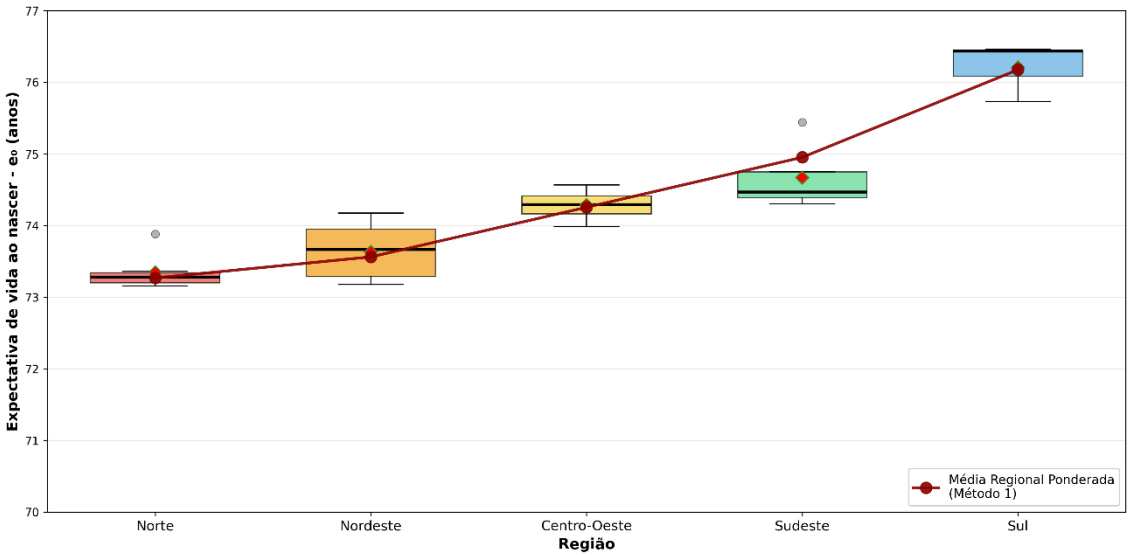
Figura 1. Razões das Taxas de Mortalidade por raça/cor e idade, Cedeplar/Imds



A figura explicita que a mortalidade negra é sistematicamente superior à branca (razão $m_x > 1$) ao longo de todo o ciclo de vida, com padrão etário bem definido. Entre homens, o excesso relativo de mortalidade negra é mais pronunciado nas idades jovens e adultas, e declina nas idades avançadas; entre as mulheres, o diferencial é menor e varia de forma mais suave, convergindo para valores próximos de 1 nas idades mais altas. Coerentemente com esse resultado, o diferencial racial em e_0 também é nacional e constante no Método 1: para homens, o gap (branca - negra) é de 5,87 anos; para mulheres, de 4,05 anos; e para ambos os sexos, de 5,36 anos.

A Figura 2 sintetiza a distribuição de e_0 entre as Unidades da Federação dentro de cada grande região, para ambos os sexos e todas as raças, destacando também a média regional ponderada. O gráfico evidencia um aspecto central do Método 1: ao manter fixos os perfis nacionais de mortalidade por raça/cor, a variação de e_0 entre UFs dentro de uma mesma região é relativamente pequena e decorre exclusivamente de diferenças na composição racial e sexual das populações estaduais. Em termos descritivos, a amplitude intrarregional de e_0 é de 0,72 anos no Norte, 1,00 no Nordeste, 0,58 no Centro-Oeste, 1,14 no Sudeste e 0,73 no Sul. Os pontos assinalados como valores atípicos (outliers) correspondem, neste método, a UFs cuja composição populacional implica um e_0 $para$ todas as raças acima (ou abaixo) do padrão típico da região. No caso ilustrado, Rondônia ($e_0 = 73{,}88$) se destaca como outlier no Norte, e São Paulo (e inferior à linha 0 é igual a 75,44) no Sudeste. Esses desvios devem ser lidos, estritamente, como efeitos composicionais no interior das regiões, e não como evidência de padrões próprios de mortalidade estadual.

Figura 2. Variação de e_0 entre as unidades da federação em cada região. Método 1.



Tomado em conjunto, o Método 1 estabelece um cenário contrafactual de referência, um efeito composicional puro, que responde à pergunta: quais seriam os níveis de mortalidade por região e por UF se as únicas diferenças territoriais decorressem da composição por raça/cor e sexo, mantendo-se fixos os perfis nacionais de mortalidade por raça/cor? Os resultados indicam que a composição racial explica parcela importante das diferenças entre grandes regiões: a média regional ponderada de e_0 no Sul é 76,18 anos, frente a 73,27 no Norte, uma diferença de 2,91 anos. Ao mesmo tempo, a dispersão intrarregional permanece limitada nesse cenário, sugerindo que diferenças empiricamente observadas entre UFs, quando se permite variação no nível de mortalidade, tendem a refletir também heterogeneidades reais nos padrões etários e nos níveis de mortalidade, dimensões que serão incorporadas progressivamente nos Métodos 2 a 5.

3.2. Método 2: Ajuste regional por raça a partir da mortalidade infantil do Atlas PNUD

O Método 2 combina três fontes de dados. Em primeiro lugar, as tabelas de vida nacionais Cedeplar/Imds para 2015, desagregadas por sexo e raça/cor, que fornecem o padrão etário da mortalidade e os parâmetros demográficos necessários à reconstrução das tabelas por idade quinquenal. Em segundo lugar, as estimativas de mortalidade infantil do Atlas/PNUD, por raça/cor e unidade geográfica, expressas como probabilidade de morrer antes de completar 1 ano de vida, q inferior à linha 1, calculadas para ambos os sexos. Em terceiro lugar, a população de 2015 por unidade geográfica, sexo, raça/cor e grupos etários quinquenais, utilizada como base de ponderação nas agregações e na construção de tabelas totais.

O objetivo é obter tabelas por raça/cor e sexo em cada unidade geográfica, preservando o padrão etário implícito nas tabelas Cedeplar/Imds, mas ajustando o nível de mortalidade para refletir os diferenciais geográficos observados no Atlas/PNUD. Em termos substantivos, o método introduz heterogeneidade territorial apenas no nível da mortalidade em idades jovens (ancorada em q_1), propagando esse ajuste para todas as idades por meio de um modelo relacional, sem alterar o formato etário original do Cedeplar/Imds.

As Unidades da Federação são agrupadas nas cinco grandes regiões: Norte, Nordeste, Sudeste, Sul e Centro-Oeste. Para cada região r e raça/cor c, calcula-se a mortalidade infantil regional como média ponderada das estimativas estaduais $q_1^{(u,c)}$, utilizando como peso a população de 0 a 4 anos da própria raça/cor na UF (somando-se homens e mulheres). Definindo $N_{u,0-4}^{(c)}$ como a população de 0 a 4 anos da raça c na UF u, tem-se:

$$q_1^{(r,c)} = \frac{\sum_{u \in r} q_1^{(u,c)} \, N_{u,0-4}^{(c)}}{\sum_{u \in r} N_{u,0-4}^{(c)}}$$

Essa ponderação produz um indicador regional coerente com a distribuição racial da população infantil na região. Embora o ideal fosse utilizar a composição racial dos nascidos vivos, optamos por não empregar dados do DATASUS, dado o risco de introduzir vieses adicionais. Assim, utilizamos a distribuição observada no grupo etário de 0 a 4 anos como proxy da exposição ao risco em idades muito jovens.

O passo seguinte estabelece uma conexão entre o indicador do Atlas/PNUD (morte até 1 ano) e o indicador disponível diretamente nas tabelas do Cedeplar/Imds (morte até 5 anos). Do lado do Atlas/PNUD, considera-se $q_1^{(\mathrm{BR},c)}$para o Brasil por raça/cor. Do lado do Cedeplar/Imds, utiliza-se $q_5^{(\mathrm{Imds},c)}$, entendido aqui como a probabilidade de morte no intervalo 0 a 4 (isto é, $5q_0$) para ambos os sexos e raça c. Define-se, então, um fator de proporcionalidade por raça/cor: $R_c = \frac{q_5^{(\mathrm{Imds},c)}}{q_1^{(\mathrm{BR},c)}}$.

Assume-se, como aproximação de primeira ordem, que essa razão também se aplica fora do agregado nacional. Assim, para cada unidade geográfica g (região r ou UF u) e raça c, define-se o valor alvo para a mortalidade até 5 anos como:

$$q_5^{(g,c,\mathrm{target})} = R_c \cdot q_1^{(g,c)},$$

onde $q_1^{(g,c)}$é a mortalidade infantil observada no Atlas/PNUD para a unidade geográfica g (ou, no caso regional, a média ponderada definida acima).

Para garantir que esse ajuste seja aplicado a todas as idades, utilizamos um modelo relacional na escala de logito. Portanto, seja $q_5^{(IMDS,c)}$a probabilidade de morrer até os 5 anos do padrão Imds para a raça c, definimos, para cada região re raça c, um parâmetro de nível dado por

$$\alpha_{r,c} = \mathrm{logito}(q_5^{(r,c,\mathrm{target})}) - \mathrm{logito}(q_5^{(Imds,c)}),$$

onde $\text{logito}(p) = \ln(\frac{p}{1-p})$. Em termos intuitivos, $\alpha_{r,c}$ indica quanto é preciso ajustar, na escala de logito, a curva de mortalidade Imds para que o q_5 resultante coincida com o valor geográfico desejado, preservando o padrão etário da função de mortalidade.

Dessa maneira, preserva-se o diferencial por sexo implícito nas tabelas Imds, enquanto o nível geográfico de mortalidade, por raça/cor, é calibrado de forma consistente com a informação do Atlas/PNUD para ambos os sexos. No grupo aberto, mantém-se a convenção do sistema de tabelas adotado, preservando-se o parâmetro correspondente do padrão Imds.

A partir do novo vetor de probabilidades $q_x^{(g,s,c)}$, reconstrói-se, para cada combinação de unidade geográfica, sexo e raça/cor, a tabela de vida completa, $(q_x, l_x, d_x, L_x, T_x, e_x)$, mantendo-se a mesma estrutura de intervalo, por unidade geográfica e sexo, por meio de médias ponderadas para todas as raças por unidade geográfica e sexo por meio de média ponderada em m_x, utilizando as populações correspondentes de 2015 como pesos. Para cada unidade geográfica g, sexo s e idade x:

$$m_x^{(g,s)} = \frac{\sum_c m_x^{(g,s,c)} \; N_{g,s,c,x}}{\sum_c N_{g,s,c,x}}.$$

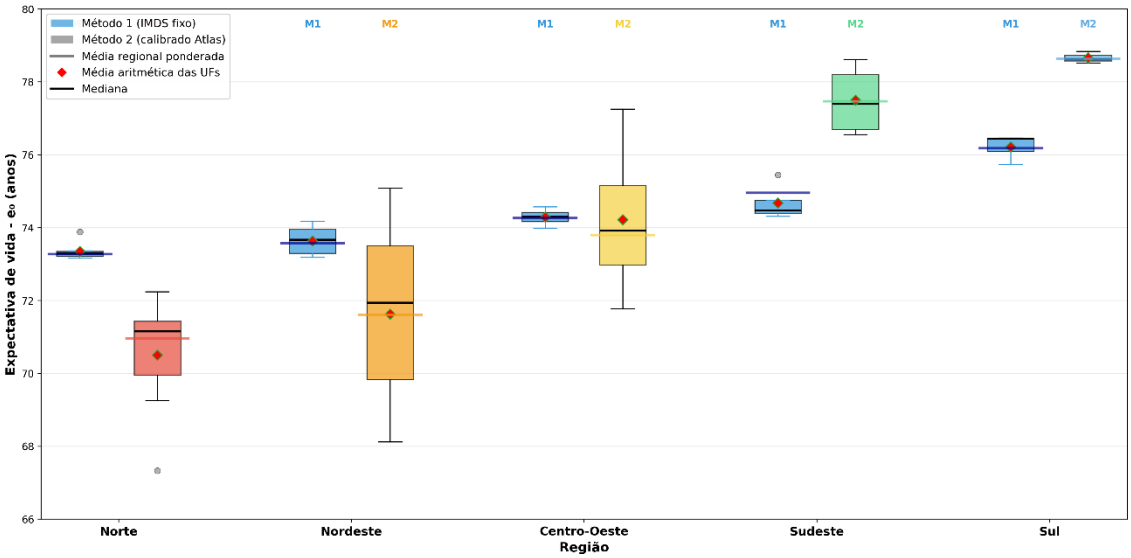
O procedimento análogo é aplicado para obter resultados para ambos os sexos, combinando homens e mulheres com base nas respectivas populações. As tabelas totais resultantes servem como base de comparação com as referências oficiais e como insumo para os métodos subsequentes.

Padrões principais observados (UFs e regiões) a partir do Método 2

As Figuras 3 e 4 comparam diretamente os resultados dos Métodos 1 e 2 para a expectativa de vida ao nascer (e_0) estimada para ambos os sexos e todas as raças, evidenciando a natureza progressiva do desenvolvimento metodológico. Como observado, no Método 1 o padrão etário e o nível de mortalidade por raça/cor e sexo estão inteiramente ancorados nas tabelas nacionais Cedeplar/Imds. Assim, a variação territorial de e_0 emerge apenas da composição por raça/cor (e, secundariamente, por sexo), uma vez que os perfis de mortalidade por idade são, por construção, comuns a todo o território. O diferencial racial implícito nas tabelas nacionais é, nesse cenário, constante no espaço: para ambos os sexos, o gap e_0(Branca − Negra) é de 5,36 anos; por sexo, ele é de 4,05 anos entre mulheres e 5,87 anos entre homens.

No Método 2, mantém-se o padrão etário do Imds, mas o nível de mortalidade passa a ser calibrado territorialmente com base nas estimativas de mortalidade infantil do Atlas/PNUD. A Figura 3 mostra que a calibração por mortalidade infantil altera de forma importante a escala da heterogeneidade entre UFs. Enquanto no Método 1 a amplitude nacional de e_0 entre UFs é de 3,30 anos (de 73,16 no Pará a 76,46 em Santa Catarina), no Método 2 ela aumenta para 11,50 anos (de 67,33 no Amapá a 78,83 em Santa Catarina), isto é, aproximadamente 3,5 vezes maior. A mudança não é homogênea entre as regiões. No Método 2, a dispersão intrarregional se expande no Norte (4,90 anos), Nordeste (6,96) e Centro-Oeste (5,49), permanece relativamente contida no Sudeste (2,06) e se contrai no Sul (0,31), sugerindo que parte importante da desigualdade territorial oculta no Método 1 aparece justamente nas regiões em que as condições de saúde e a capacidade de resposta institucional são mais heterogêneas entre UFs. Em paralelo, as médias regionais ponderadas passam a refletir um gradiente muito mais intenso do que no Método 1: no Método 2, o $e_0$regional (ambos os sexos, todas as raças) é de 78,64 no Sul, 77,47 no Sudeste, 73,80 no Centro-Oeste, 71,61 no Nordeste e 70,96 no Norte, com amplitude regional de 7,68 anos.

Figura 3. Variação de e_0 entre as unidades da federação em cada região. Métodos 1 e 2.

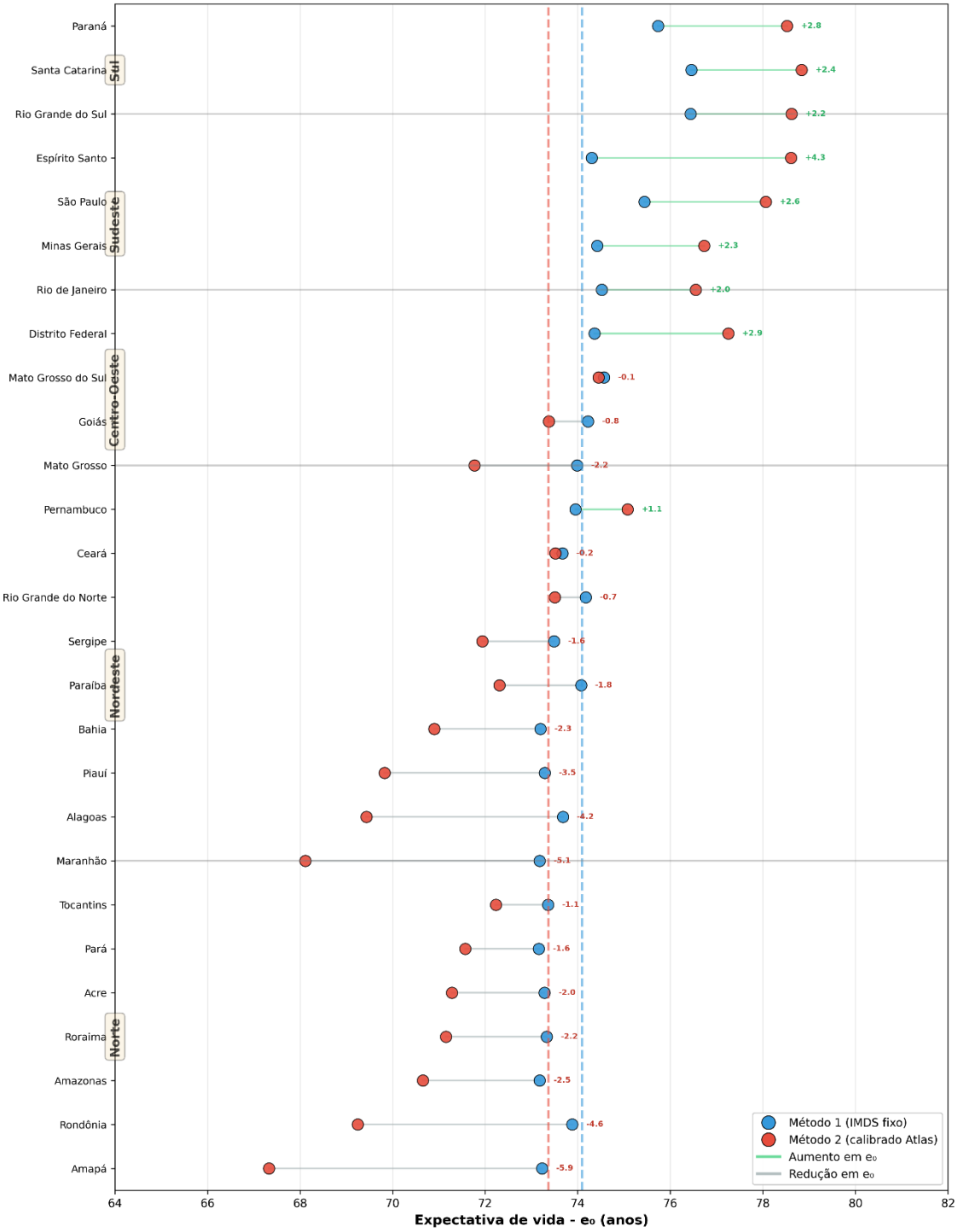


A Figura 4 complementa nossa análise ao explicitar o sentido e a magnitude dos deslocamentos nas expectativas de vida. Ela evidencia que a calibração do Método 2 impõe, para a maior parte das UFs, revisões no nível de mortalidade implícito no Método 1. O padrão regional é claro: Norte, Nordeste e Centro-Oeste deslocam-se, em média, para níveis de e_0 inferiores aos do padrão Imds fixo (interpretação compatível com $\alpha > 0$, isto é, mortalidade infantil pior do que a sugerida pelo Imds nacional), ao passo que

Sudeste e Sul se deslocam para níveis superiores ao do Imds (interpretação compatível com $\alpha < 0$, mortalidade infantil melhor que a média nacional). Esse contraste é o sinal mais direto de que, ao contrário do Método 1, o Método 2 não se limita a descrever um efeito composicional. Ele também incorpora variação espacial no nível de mortalidade, ancorada em um indicador observacional externo (q_1 do Atlas/PNUD), e a propaga ao longo das idades preservando o padrão etário de referência.

Essa transição tem implicações imediatas para a desigualdade racial. No Método 1, o gap racial em e_0é constante no território (5,36 anos para ambos os sexos), por construção. No Método 2, como o ajuste de nível é feito separadamente por raça/cor com base em q_1, o diferencial racial deixa de ser invariável e passa a variar entre regiões. Em termos regionais, o gap (Branca - Negra) passa a ser de 4,88 no Norte, 5,21 no Nordeste, 4,40 no Centro-Oeste, 3,91 no Sudeste e 3,56 no Sul. Isso implica uma redução de aproximadamente 33% no Sul (3,56 versus 5,36) e reduções também relevantes no Sudeste, em contraste com o Nordeste, onde o diferencial permanece próximo ao padrão nacional.

Figura 4. Mudança de e_0 por unidade da federação. Método 1 -> Método 2.



3.3. Método 3: Ajuste regional por raça ancorado simultaneamente nas estimativas do Atlas/Pnud e no IBGE

O Método 3 mantém a lógica geral do Método 2, mas introduz um passo adicional de calibração para garantir consistência com as estimativas oficiais do IBGE. Em termos simples, continuamos a usar o Atlas/PNUD para capturar os diferenciais de mortalidade infantil por raça/cor entre territórios, mas passamos a ancorar o nível agregado de mortalidade infantil de cada grande região nos valores de q_1 obtidos a partir das tabelas regionais do IBGE. Assim, o Atlas/PNUD fornece a estrutura dos diferenciais por raça/cor dentro de cada região, enquanto o IBGE define o nível de q_1 que o sistema deve reproduzir para o conjunto de todas as raças em cada região.

Do ponto de vista das bases de dados, o Método 3 utiliza as mesmas informações centrais já descritas para os Métodos 1 e 2: (i) as tabelas nacionais do Cedeplar/Imds por sexo e raça/cor, que fornecem o padrão etário da mortalidade e os parâmetros necessários para reconstruir tabelas completas; (ii) as estimativas de mortalidade infantil do Atlas/PNUD por raça/cor e Unidade da Federação (UF), expressas como q_1; e (iii) a população de 2015 por UF, sexo, raça/cor e grupos etários quinquenais, utilizada como peso nas agregações. A esses insumos soma-se agora q_1 por grande região produzido pelo IBGE para 2015 (todas as raças, por região), derivado das tabelas regionais por sexo. Essas estimativas do IBGE são utilizadas estritamente como referência para o nível agregado de mortalidade infantil regional e não interferem no padrão etário das funções de mortalidade.

A primeira etapa do Método 3 reproduz o procedimento do Método 2 para organizar as informações do Atlas/PNUD. Para cada UF u e raça/cor c, dispomos de $q_1^{(u,c)}$. Em seguida, para cada grande região r, calculam-se médias ponderadas dessas probabilidades, usando como peso a população infantil (0 a 4 anos) da UF, $N_{u,0-4}$, obtendo-se:

$$q_1^{(r,c),\text{Atlas}} = \frac{\sum_{u\in r} q_1^{(u,c)}\ N_{u,0-4}}{\sum_{u\in r} N_{u,0-4}}.$$

A partir dessas quantidades por raça/cor, construímos também, para cada região, um q_1 total implícito no Atlas, $q_1^{(r),\text{Atlas}}$, como média ponderada por raça/cor, usando como pesos as populações infantis regionais por raça/cor.

O passo distintivo do Método 3 é a calibração do nível regional do Atlas para coincidir com o nível regional do IBGE. Para cada região r, define-se um fator de calibração:

$$s_r = \frac{q_1^{(r),\mathrm{IBGE}}}{q_1^{(r),\mathrm{Atlas}}}.$$

Esse fator é então aplicado às probabilidades do Atlas por raça/cor dentro da região, de modo que, após a correção, a média ponderada por raça/cor reproduza exatamente o q_1 regional do IBGE:

$$q_1^{(r,c),\mathrm{cal}} = s_r \cdot q_1^{(r,c),\mathrm{Atlas}}$$

Na extensão às UFs, o mesmo princípio é aplicado no interior de cada região: para cada UF $u \in r$, obtém-se $q_1^{(u,c),\mathrm{cal}} = s_r \cdot q_1^{(u,c)}$. Em termos intuitivos, o Atlas/PNUD continua determinando o diferencial entre brancos e negros (e sua variação territorial), mas o nível médio de mortalidade infantil de cada grande região passa a ser definido pelo IBGE.

Uma vez obtidos os $q_1$calibrados por raça/cor, o Método 3 retoma o mesmo mecanismo de conexão com as tabelas Imds já introduzido no Método 2. Como as tabelas Imds não dispõem diretamente de q_1(mas de q_5, ou $5q_0$), utiliza-se novamente, para cada raça/cor c, um fator nacional que relaciona os dois indicadores, e converte-se o q_1 calibrado em um alvo para q_5. Em seguida, estima-se um parâmetro de nível na escala de logito que desloca, para cada unidade territorial g(onde g pode ser uma grande região r ou uma UF u), a curva de probabilidades de morte por idade do padrão Imds. O ajuste é aplicado a todas as idades por inversão do logito, preservando o padrão etário do Imds e alterando apenas o nível, de modo que o q_5 resultante coincida com o alvo territorial definido a partir de q_1. No grupo aberto, procede-se de forma análoga ao modelo relacional já descrito nos métodos anteriores.

O mesmo raciocínio é estendido às tabelas específicas por sexo. Partindo de $q_x^{(\mathrm{Imds};s,c)}$, aplica-se o mesmo $\alpha_{g,c}$ calculado com base em $q_1$de ambos os sexos, preservando o diferencial por sexo implícito no Imds e ajustando o nível territorial de mortalidade para cada raça/cor. Ao final, agregam-se as tabelas por raça/cor para obter tabelas totais (todas as raças), por sexo e para ambos os sexos, por meio de médias ponderadas das taxas específicas por idade, usando as populações territoriais de 2015 como pesos, como nos Métodos 1 e 2.

Do ponto de vista substantivo, o Método 3 pode ser interpretado como um aperfeiçoamento do Método 2 quanto à consistência vertical regional. Continuamos a explorar o Atlas/PNUD para capturar desigualdades raciais em mortalidade infantil, preservando o padrão etário do Cedeplar/Imds. No entanto, ao impor que o q_1 total

resultante em cada grande região coincida com o q_1 das tabelas regionais do IBGE, garantimos que nossas tabelas territoriais por raça/cor e sexo sejam compatíveis com o nível de mortalidade infantil oficialmente reportado para a população total em cada região. Em suma, se o Método 2 ancora o nível em q_1 do Atlas/PNUD, o Método 3 adiciona uma segunda âncora no IBGE, resultando em um sistema que combina, de forma coerente, diferenciais raciais (Atlas/PNUD) e níveis regionais agregados (IBGE).

3.4 Método 4: Ajuste vertical Brasil x Imds a partir das tabelas regionais por raça/cor

O Método 4 parte diretamente das tabelas territoriais por raça/cor produzidas no Método 3 e introduz um segundo nível de ancoragem, agora no nível nacional. A ideia central é garantir que, ao agregar as tabelas territoriais por raça/cor e idade, ponderadas pela distribuição populacional de 2015, a expectativa de vida ao nascer do Brasil para brancos e negros reproduza exatamente os valores das tabelas nacionais do Cedeplar/Imds. Em outras palavras, enquanto o Método 3 assegura consistência regional com o IBGE no nível de q_1(todas as raças), o Método 4 acrescenta uma restrição vertical adicional: as tabelas por raça/cor, quando reagrupadas do nível territorial para o Brasil, devem ser compatíveis com o nível nacional por raça/cor do próprio Imds.

Do ponto de vista das bases de dados, o Método 4 não introduz novas fontes externas em relação aos métodos anteriores. Continuamos a utilizar: (i) as tabelas nacionais do Cedeplar/Imds por sexo e raça/cor, que fornecem o padrão etário e os parâmetros demográficos por idade; (ii) a distribuição da população de 2015 por UF, sexo, raça/cor e idade, previamente organizada para permitir agregações; e (iii) as tabelas territoriais por raça/cor resultantes do Método 3, que já incorporam simultaneamente os diferenciais raciais do Atlas/PNUD e a calibração regional do nível agregado via IBGE. A novidade está na forma como utilizamos essas tabelas para reconstruir, de baixo para cima, uma tabela nacional por raça/cor, impondo que sua expectativa de vida ao nascer coincida com a referência do Imds.

Operacionalmente, para cada raça/cor c (branca e negra), define-se um parâmetro de nível ∂_c, aplicado na escala de logito às probabilidades de morte por idade das tabelas territoriais por raça/cor produzidas no Método 3. Mantém-se o padrão etário e os diferenciais territoriais já estimados; varia apenas o nível geral da mortalidade por raça/cor, deslocado por um valor constante ∂_c em todas as idades (com tratamento do grupo aberto consistente com o modelo relacional). Para um valor dado de ∂_c, recalculam-se as tabelas territoriais por raça/cor e agrega-se, por idade, a mortalidade territorial usando como pesos as populações de 2015, obtendo-se assim uma tabela

nacional implícita para a raça/cor c. Em termos substantivos, esse procedimento garante que o conjunto de tabelas territoriais por raça/cor permaneça coerente com os níveis nacionais do Imds quando agregado ao Brasil.

Uma vez estimados ∂_{branca}e ∂_{negra}, aplica-se o mesmo ajuste vertical às tabelas específicas por sexo. Mantém-se intacto o diferencial por sexo implícito no Imds dentro de cada raça/cor, deslocando apenas o nível de mortalidade de forma uniforme por raça/cor. Em seguida, agregam-se as tabelas por raça/cor para obter tabelas territoriais totais (todas as raças) por sexo e, por fim, para ambos os sexos, por meio de médias ponderadas das taxas específicas, utilizando as populações territoriais por idade, sexo e raça/cor de 2015, como descrito nos Métodos 1 a 3.

Do ponto de vista substantivo, o Método 4 pode ser interpretado como um refinamento adicional da consistência vertical do sistema de tabelas. O Atlas/PNUD continua definindo os diferenciais de mortalidade infantil entre brancos e negros nos territórios. O IBGE permanece como referência para o nível de mortalidade infantil agregado por grande região, e o Imds passa a atuar também como âncora explícita para o nível nacional por raça/cor. Ao permitir um pequeno ajuste uniforme do nível de mortalidade por raça/cor em todos os territórios, preservando os diferenciais territoriais e de sexo construídos no Método 3, o Método 4 produz um conjunto de tabelas territoriais por raça/cor e sexo que é simultaneamente coerente com a estrutura etária e os níveis nacionais do Imds e compatível com os diferenciais territoriais de mortalidade infantil observados nas fontes externas.

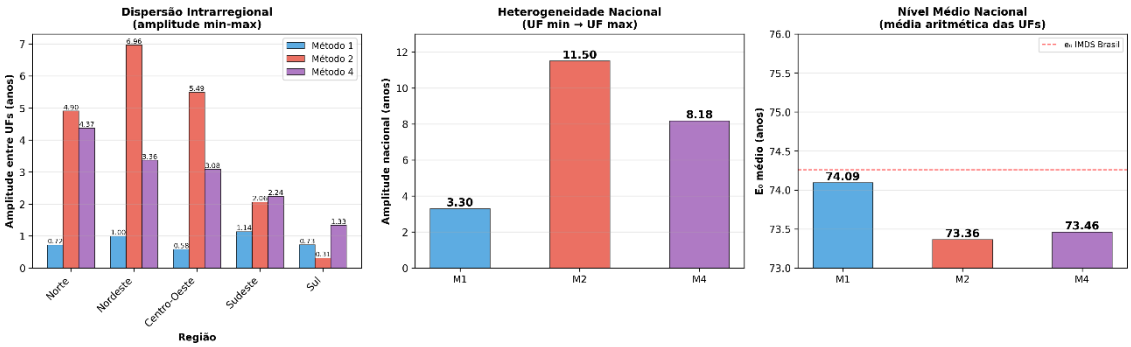
Padrões principais observados (UFs e regiões) a partir do Método 4

As Figuras 5 a 7 sintetizam a evolução dos resultados quando comparamos diretamente os Métodos 1, 2 e 4 para e_0(ambos os sexos, todas as raças) e para o diferencial racial em e_0(Branca - Negra, ambos os sexos). Como discutido nas seções metodológicas, os três métodos compartilham o mesmo padrão etário de mortalidade de referência, mas diferem quanto às âncoras utilizadas para definir o nível de mortalidade no território. O Método 1 constitui um cenário-base em que o nível e o padrão etário por raça/cor e sexo são inteiramente ancorados nas tabelas nacionais Cedeplar/Imds, de modo que a variação territorial decorre apenas do efeito composicional. O Método 2 mantém o padrão etário Imds, mas calibra o nível de mortalidade territorialmente a partir do $q_1$do Atlas/PNUD, separadamente por raça/cor, ampliando a heterogeneidade espacial. O Método 4, aplicado de forma acumulativa ao Método 3, acrescenta uma restrição vertical adicional: parte das tabelas regionais por raça/cor já calibradas para reproduzir, em cada região, o $q_1$total do IBGE e ajusta, por raça/cor, o nível de mortalidade de modo que a

reagregação nacional (ponderada pela população) reproduza exatamente as expectativas de vida nacionais do Cedeplar/Imds para brancos e negros. Com isso, preservam-se os diferenciais regionais e o padrão etário construídos nas etapas anteriores, ao mesmo tempo em que se garante coerência do sistema no nível do Brasil por raça/cor.

A primeira evidência dessa progressão aparece na dispersão intrarregional de e_0 entre UFs (Figura 5, painel esquerdo). No Método 1, a amplitude min–max de $e_0$dentro de cada grande região é baixa, variando de 0,58 anos (Centro-Oeste) a 1,14 anos (Sudeste), o que reflete a natureza estritamente composicional do método. No Método 2, essa dispersão se expande fortemente, sobretudo nas regiões onde as condições de saúde e a capacidade de resposta institucional são mais heterogêneas entre as UFs: Norte (4,90 anos), Nordeste (6,96) e Centro-Oeste (5,49). Ao mesmo tempo, a dispersão permanece relativamente contida no Sudeste (2,06) e se contrai no Sul (0,31), sugerindo que, ao incorporar um indicador externo de nível (mortalidade infantil), o Método 2 revela desigualdades territoriais que estavam ocultas sob o cenário homogeneizante do Método 1. No Método 4, a dispersão intrarregional se reduz em relação ao Método 2 no Norte (4,37), Nordeste (3,36) e Centro-Oeste (3,08), sinalizando o efeito disciplinador do ajuste vertical nacional. No entanto, permanece bem acima do padrão do Método 1 nessas regiões. Em contraste, no Sudeste e no Sul, o Método 4 resulta em dispersões ligeiramente superiores às do Método 2 (2,24 no Sudeste e 1,33 no Sul), um padrão compatível com o fato de que, no Método 2, o Sul já se apresenta extremamente homogêneo e, no Método 4, o ajuste vertical pode redistribuir os níveis de forma não uniforme quando observado no agregado por UF.

Figura 5. Comparação de e_0 entre as unidades da federação e regiões segundo os Métodos 1, 2 e 4.



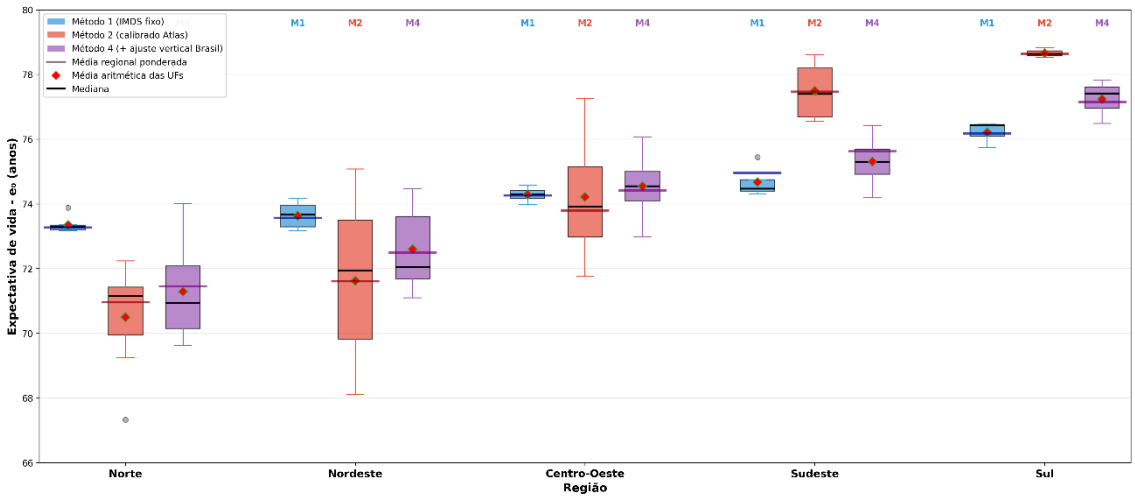
Essa mesma leitura fica ainda mais clara quando observamos a heterogeneidade nacional de $e_0$entre UFs (Figura 5, painel central). No Método 1, a amplitude nacional é de 3,30 anos, enquanto, no Método 2, ela salta para 11,50 anos, isto é, mais de três vezes maior. No Método 4, a amplitude nacional cai para 8,18 anos, o que confirma que o ajuste vertical recupera parte da coerência do sistema em relação à âncora nacional por

raça/cor, mas sem eliminar a heterogeneidade territorial introduzida pelo uso do Atlas/PNUD como referência de nível. Em termos substantivos, isso indica que uma parcela importante das diferenças espaciais sugeridas pelo Método 2 não é meramente composicional: decorre de variações territoriais no nível de mortalidade consistentes com um indicador externo de mortalidade infantil, mas que precisam ser compatibilizadas com as restrições nacionais do Imds no Método 4.

O painel direito complementa essa síntese ao comparar o nível médio nacional de e_0 quando tomamos a média aritmética das UFs. Nesse critério, o Método 1 produz e_0 médio de 74,09 anos, ao passo que o Método 2 reduz esse valor para 73,36 e o Método 4 o eleva ligeiramente para 73,46. Essa diferença entre métodos é coerente com o fato de que, no Método 2, a calibração pelo Atlas/PNUD rebaixa substancialmente o nível de mortalidade implícito para um amplo conjunto de UFs, especialmente no Norte e no Nordeste, enquanto o Método 4 recupera parte desse rebaixamento ao impor consistência vertical nacional por raça/cor com o Imds. É importante notar que esse painel resume médias não ponderadas por população. Por isso, ele deve ser interpretado como uma estatística descritiva da distribuição entre UFs, e não como uma estimativa direta da e_0 nacional ponderada.

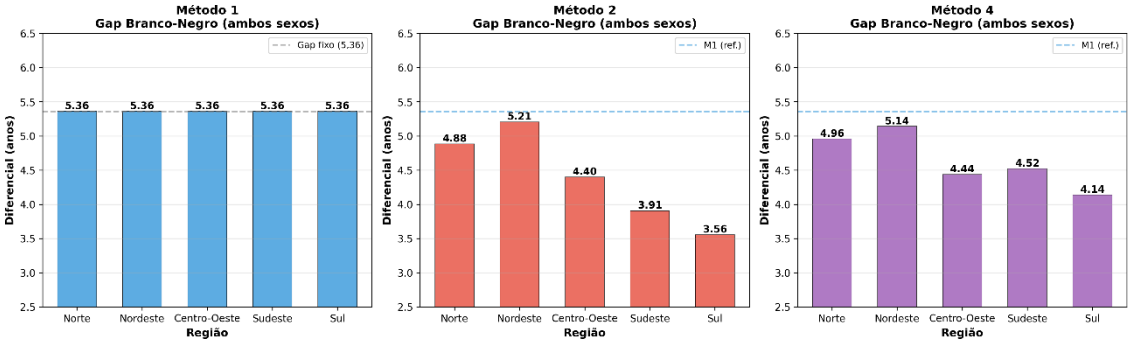
A Figura 6 reforça visualmente os mesmos achados por meio de boxplots regionais, ao mostrar simultaneamente os deslocamentos de nível e as mudanças na dispersão entre os três métodos. Em particular, ela evidencia que a transição do Método 2 para o Método 4 reduz a dispersão em regiões com grande heterogeneidade (especialmente no Nordeste e no Centro-Oeste), mas preserva um padrão de desigualdade espacial consideravelmente mais intenso do que o observado no cenário-base do Método 1.

Figura 6. Variação de e_0 entre as unidades da federação em cada região. Métodos 1, 2 e 4.



Por fim, a Figura 7 descreve a evolução do diferencial racial em e_0(Branca – Negra, ambos os sexos) ao longo dos métodos. No Método 1, o diferencial é constante por construção, fixado em 5,36 anos em todas as regiões. No Método 2, como o ajuste de nível é feito separadamente por raça/cor com base no $q_1$do Atlas/PNUD, o gap deixa de ser invariável e passa a variar regionalmente, assumindo 4,88 anos no Norte, 5,21 no Nordeste, 4,40 no Centro-Oeste, 3,91 no Sudeste e 3,56 no Sul. No Método 4, o diferencial racial regional se desloca novamente, refletindo a restrição vertical que impõe consistência nacional por raça/cor com o Imds. Os gaps tornam-se 4,96 anos no Norte, 5,14 no Nordeste, 4,44 no Centro-Oeste, 4,52 no Sudeste e 4,14 no Sul. Em comparação ao Método 2, observa-se uma recomposição do diferencial racial, sobretudo no Sudeste e no Sul, onde o gap aumenta e aproxima-se, parcialmente, do padrão nacional, embora permaneça abaixo do valor fixo do Método 1. Em termos interpretativos, isso é compatível com o objetivo do Método 4: preservar a informação espacial do Atlas/PNUD e a coerência regional com a âncora do IBGE (na etapa anterior), mas garantindo, simultaneamente, que a decomposição por raça/cor, quando reagrupada nacionalmente, permaneça ancorada nos níveis nacionais por raça/cor do Imds.

Figura 7. Gap racial em e_0 segundo métodos e regiões.



3.5. Método 5: Ajuste de segunda ordem por UF e sexo em torno do IBGE, com tolerância percentual

O Método 5 parte diretamente das tabelas estimadas no Método 4, que, por sua vez, já incorporam, de forma cumulativa, os dois níveis anteriores de ancoragem: (i) os diferenciais raciais de mortalidade infantil do Atlas/Pnud, condicionados ao nível regional do IBGE no Método 3; e (ii) a restrição vertical nacional por raça/cor do Imds no Método 4. O objetivo do Método 5 é introduzir um ajuste adicional, de segunda ordem, agora diferenciado por unidade federativa e sexo, para aproximar os níveis territoriais de expectativa de vida ao nascer das referências oficiais do IBGE, sem alterar o padrão etário das funções de mortalidade nem desancorar o sistema das restrições já impostas por raça/cor.

Do ponto de vista dos dados, o Método 5 não acrescenta novas fontes às já descritas. Mantêm-se as tabelas nacionais Cedeplar/Imds por sexo e raça/cor, as estimativas do Atlas/Pnud por raça/cor e UF (já utilizadas para construir os diferenciais raciais e regionais nos Métodos 2 e 3), a população de 2015 por UF, sexo, raça/cor e idade (usada como peso nas agregações), e as referências de $e_0$do IBGE por UF e sexo para 2015, que passam a operar como alvo de aproximação para o nível territorial por sexo.

Na versão anterior do Método 5, apresentada no Relatório 2, essa aproximação ao IBGE era operacionalizada por meio de bandas de tolerância fixas (em anos) definidas por grande região, aplicadas à diferença entre $e_{0,r,s}^{(4)}$e $e_{0,r,s}^{IBGE}$. A intuição era reconhecer, de forma explícita, que a incerteza e a heterogeneidade dos dados de mortalidade variam regionalmente, adotando-se tolerâncias mais estritas nas regiões com melhor qualidade de informação (Sul e Sudeste) e mais amplas nas regiões onde erros de idade e problemas de cobertura tendem a ser mais relevantes (Centro-Oeste, Nordeste e Norte). Contudo, ao estendermos o procedimento ao nível das Unidades da Federação, essa parametrização regional fixa mostrou-se menos adequada para controlar, de forma comparável entre unidades, o grau de aderência horizontal ao IBGE e o trade-off com as restrições verticais já incorporadas ao sistema.

Portanto, em vez de bandas fixas por região, as tolerâncias passam a ser definidas, neste relatório, como uma fração do próprio e_0 oficial do IBGE, para cada UF u e sexo s. Seja $e_{0,u,s}^{IBGE}$ a expectativa de vida ao nascer do IBGE e seja $e_{0,u,s}^{(4)}$ a expectativa de vida total (todas as raças) produzida no Método 4 para a mesma UF e sexo, define-se uma tolerância percentual comum p, e a banda admissível em anos é dada por

$$\tau_{u,s} = p \cdot e_{0,u,s}^{IBGE}.$$

A partir daí, define-se um alvo intermediário $e_{0,u,s}^{*}$ por uma regra de truncamento: se o nível do Método 4 já está dentro da banda em torno do valor oficial, nenhuma correção adicional é aplicada; caso contrário, o alvo é deslocado até a borda da banda mais próxima do valor do Método 4. Formalmente,

$$e_{0,u,s}^{*} = \begin{cases} e_{0,u,s}^{(4)}, & \text{se } \mid e_{0,u,s}^{(4)} - e_{0,u,s}^{IBGE} \mid \leq \tau_{u,s}, \\ e_{0,u,s}^{IBGE} + \tau_{u,s} \cdot \operatorname{sign}\left(e_{0,u,s}^{(4)} - e_{0,u,s}^{IBGE}\right), & \text{se } \mid e_{0,u,s}^{(4)} - e_{0,u,s}^{IBGE} \mid > \tau_{u,s}. \end{cases}$$

Essa construção tem duas implicações importantes. Primeiro, quando $e_{0,u,s}^{(4)}$já é compatível com o IBGE dentro do grau de tolerância escolhido, o Método 5 é inerte: preserva integralmente o nível do Método 4. Segundo, quando há discrepância maior, o ajuste aproxima o resultado do IBGE, mas impõe explicitamente um limite à correção,

evitando reescalonamentos excessivos em contextos de maior incerteza e heterogeneidade territorial.

Uma vez fixado $e^*_{0,u,s}$, o ajuste é implementado por um deslocamento uniforme, na escala de logito, aplicado às probabilidades de morte por idade das tabelas do Método 4, separadamente para cada combinação UF–sexo, e de forma comum às raças dentro dessa combinação. Denotando por $q^{(4)}_{x,u,s,c}$as probabilidades de morte por idade x, UF u, sexo se raça/cor $c$no Método 4, define-se

$$\text{logito}\left(q^{(5)}_{x,u,s,c}\right) = \text{logito}\left(q^{(4)}_{x,u,s,c}\right) + \gamma_{u,s},$$

para todas as idades abaixo do grupo aberto, com tratamento do último grupo etário consistente com os métodos anteriores. Para cada UF e sexo, o parâmetro $\gamma_{u,s}$é determinado numericamente de modo que, após reconstruir as tabelas por raça/cor e agregá-las para todas as raças por mistura ponderada das taxas por idade, a expectativa de vida ao nascer resultante coincida com o alvo definido:

$$e^{(5)}_{0,u,s}(\gamma_{u,s}) = e^*_{0,u,s}.$$

O Método 5 é, portanto, um ajuste de segunda ordem em dois sentidos. Em primeiro lugar, ele opera sobre um sistema que já foi calibrado para reproduzir, simultaneamente, os diferenciais raciais de mortalidade infantil do Atlas/Pnud condicionados ao nível regional do IBGE (Método 3) e a consistência vertical nacional por raça/cor com o Imds (Método 4). Em segundo lugar, o ajuste é restrito a um deslocamento homogêneo por UF e sexo, preservando o padrão etário e os diferenciais raciais previamente construídos, enquanto corrige apenas o nível territorial total por sexo dentro de uma banda controlada.

Por fim, a escolha do parâmetro p é tratada como uma análise de sensibilidade: o procedimento pode ser repetido para diferentes valores de p, do caso-limite $p = 0$(maior aproximação pontual ao IBGE, na medida em que o alvo coincide com o valor oficial) até valores mais permissivos (bandas mais largas), permitindo mapear o trade-off entre aderência às referências oficiais e preservação das restrições acumuladas ao longo dos Métodos 3 e 4.

Para facilitar o entendimento dos métodos desenvolvidos nesta pesquisa, o Quadro A1, em anexo, sintetiza as hipóteses, os tipos de ajuste e as restrições impostas a cada um deles.

Método 5: Resultados

O Método 5 introduz um ajuste de segunda ordem por unidade da federação (UF) e por sexo, com o objetivo de aproximar os níveis territoriais estimados às referências oficiais do IBGE, sem perder a consistência vertical assegurada pelo Método 4. Como discutido, o critério de ajuste é definido por uma banda de tolerância proporcional ao nível oficial de expectativa de vida ao nascer do IBGE.

A Tabela 1 sintetiza os principais indicadores do trade-off entre a aderência horizontal (IBGE territorial) e a consistência vertical (Imds nacional) nos cenários analisados. Os resultados estão vinculados aos valores de tolerância em relação às expectativas de vida ao nascer estimadas pelo IBGE, variando entre 0,0% (tolerância nula) e 3,0% (tolerância mais permissiva).

Tabela 1. Indicadores de trade-off por cenário de tolerância (Método 5).

p (%)	N ajustes	|γ| médio	Δe0 negra vs Imds (anos)	RMSE log(mx) (negra)
0,0	81 (100%)	0,193	+1,45	0,126
0,5	70 (86%)	0,157	+1,26	0,110
1,0	60 (74%)	0,126	+1,06	0,091
1,5	50 (62%)	0,100	+0,87	0,076
2,0	42 (52%)	0,078	+0,72	0,062
2,5	34 (42%)	0,062	+0,60	0,052
3,0	31 (38%)	0,048	+0,48	0,041

O cenário p = 0,0% representa a máxima aderência ao IBGE, ao custo de um desvio vertical na expectativa de vida ao nascer de +1,45 anos para a população negra (agregada) em relação ao Imds nacional, e de ajustes em todas as 81 combinações de UFs e sexo (homens, mulheres e ambos os sexos). Em contraste, valores de p entre 2,5% e 3,0% produzem um equilíbrio mais favorável: reduzem o desvio vertical em 59% a 67%, mantêm o RMSE de log(mx) baixo (0,041 a 0,052), aplicam ajustes apenas onde estritamente necessário (31 a 34 combinações UF×sexo) e preservam 58% a 62% das combinações sem alterações (M5 = M4).

O padrão geográfico dos ajustes reforça a coerência do critério de tolerância percentual. Na parametrização p = 2,5%, todas as 7 UFs do Norte e 6 das 9 UFs do Nordeste recebem

algum ajuste, assim como uma UF do Sul (Rio Grande do Sul. Homens); as regiões Sudeste e Centro-Oeste permanecem inalteradas. Esse comportamento indica que o Método 4, ancorado no Imds nacional, já produzia níveis territoriais próximos ao IBGE nas regiões de melhor qualidade de dados, ao passo que Norte e Nordeste concentravam discrepâncias residuais maiores. Portanto, para os resultados apresentados a seguir, adota-se p = 2,5%. Em termos práticos, a escolha é muito próxima de p = 3,0%, mas preserva um grau adicional de aderência às tabelas oficiais do IBGE para as UFs.

A Tabela 2 compara o Método 5 (p = 2,5%) aos métodos anteriores quanto à heterogeneidade territorial da expectativa de vida ao nascer (e_0), considerando ambos os sexos e todas as raças. A amplitude de 11,50 anos do Método 2 reflete a incorporação, sem mediação adicional, dos diferenciais territoriais de mortalidade infantil provenientes do Atlas/PNUD. O Método 4 (8,18 anos) modera esses extremos ao impor consistência vertical com o Imds nacional por raça/cor. O Método 5 reduz a amplitude em 47% em relação ao M2 e em 25% em relação ao M4, situando-se em um patamar intermediário: 1,9 vezes a variação puramente composicional do M1, mas substancialmente inferior à heterogeneidade observada quando se aplica o Atlas/PNUD sem restrições verticais adicionais.

Tabela 2. Amplitude nacional da expectativa de vida (e_0) por método (ambos os sexos, todas as raças)

Método	Amplitude nacional (anos)	UF mín (e_0)	UF máx (e_0)
M1	3,30	Pará (73,16)	Santa Catarina (76,46)
M2	11,50	Amapá (67,33)	Santa Catarina (78,83)
M4	8,18	Roraima (69,63)	Santa Catarina (77,82)
M5 (p=2,5%)	6,14	Alagoas (71,67)	Santa Catarina (77,81)

A dispersão intrarregional, medida pela amplitude entre UFs de uma mesma grande região, segue um padrão semelhante ao descrito anteriormente (Tabela 3). No Norte e no Nordeste, regiões em que o Método 5 aplicou ajustes em maior proporção de UFs, a dispersão recua para 1,67 e 2,78 anos, valores 66% e 60% menores do que no Método 2. Ainda assim, permanece acima do baseline composicional do M1. Em contraste, Sudeste e Centro-Oeste mantêm dispersões idênticas ao Método 4, confirmando que a banda de tolerância percentual preserva integralmente os resultados nessas regiões. No Sul, a dispersão é baixa em todos os métodos e a razão M5/M2 é elevada, pois o denominador (M2) é particularmente pequeno.

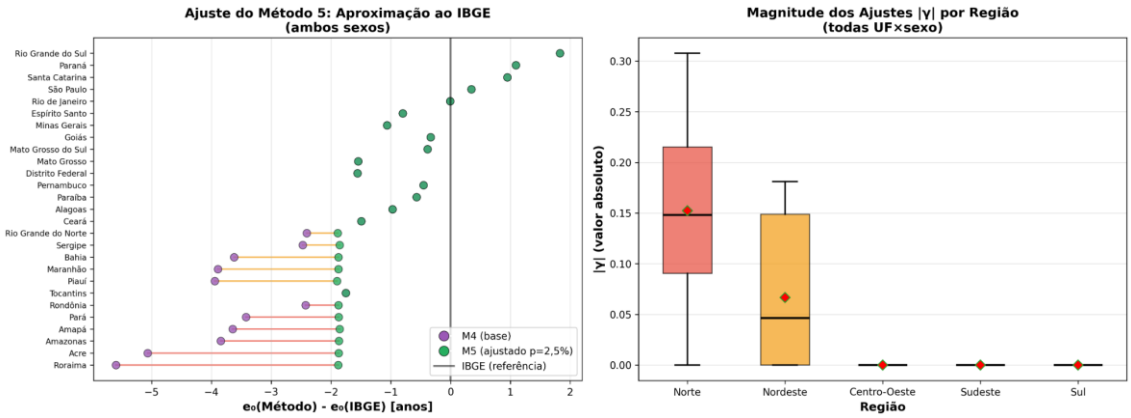
Tabela 3. Dispersão intrarregional (amplitude entre UFs, em anos) por método.

Região	M1	M2	M4	M5	M5/M2
Norte	0,72	4,90	4,37	1,67	0,34×
Nordeste	1,00	6,96	3,36	2,78	0,40×
Centro-Oeste	0,58	5,49	3,08	3,08	0,56×
Sudeste	1,14	2,06	2,24	2,24	1,09×
Sul	0,73	0,31	1,33	1,34	4,29×

A Figura 8 detalha o comportamento do ajuste de M4 para M5 em relação às referências do IBGE por UF, evidenciando a lógica de truncamento na borda da banda de tolerância. No painel de ambos os sexos, as estimativas por UF situam-se dentro, ou exatamente na borda de 2,5%. As sete UFs que atingem numericamente a borda (Amapá, Pará, Roraima, Bahia, Maranhão, Piauí e Sergipe) apresentam um excesso desprezível (de 0,000 a 0,005 anos), compatível com o arredondamento numérico.

As maiores aproximações absolutas ao IBGE ocorrem justamente onde o Método 4 apresentava os maiores desvios negativos: Roraima (de -5,60 para -1,88 anos), Acre (de -5,07 para -1,87), Piauí (de -3,94 para -1,90) e Maranhão (de -3,89 para -1,87). Em contraste, as UFs do Sul exibem desvios positivos moderados após o ajuste (+0,95 a +1,83 anos), indicando que o sistema de tabelas no Método 5 mantém níveis ligeiramente superiores aos do IBGE nessas localidades. Esse resultado é consistente com o trade-off embutido na escolha de p: o algoritmo prioriza corrigir discrepâncias residuais maiores (concentradas no Norte e Nordeste) sem impor compressão adicional onde o Método 4 já era aderente.

Figura 8. Indicadores de ajuste do Método 5 (p = 2,5%)



Ao permitir ajustes independentes por UF×sexo, o Método 5 introduz uma flexibilização ausente no Método 4, que operava deslocamentos uniformes por raça/cor. A média nacional da diferença entre M5 e IBGE é de -0,87 anos para homens e -1,38 anos para mulheres (de ambas as raças), com -1,16 anos no agregado de ambos os sexos. Em média, portanto, as estimativas femininas permanecem mais abaixo do IBGE do que as

masculinas. Esse resultado é compatível com a interação entre (i) diferenças de nível entre as tabelas do Imds e as implícitas no IBGE territorial por sexo, (ii) a forma como o Atlas/PNUD captura diferenciais territoriais de mortalidade por sexo e (iii) a possibilidade de cada combinação UF×sexo atingir sua própria borda de tolerância sem restrições cruzadas. Algumas UFs exibem discrepâncias mais pronunciadas entre o comportamento masculino e feminino na aderência ao IBGE. No Ceará (Δ = +1,61), no Espírito Santo (+1,32) e na Paraíba (+1,23), a diferença entre o M5 e o IBGE é sistematicamente maior entre homens do que entre mulheres. Esse padrão não indica inconsistência interna, mas a capacidade do Método 5 de acomodar variação territorial específica por sexo, que os métodos anteriores, por construção, não incorporavam.

Os gaps raciais (população branca menos população negra) no Método 5 variam entre 3,38 anos (Pernambuco) e 8,47 anos (Roraima), para ambos os sexos. A Tabela 4 resume a distribuição regional desses diferenciais. O Norte apresenta gaps médios mais elevados e maior variabilidade intrarregional, com Roraima em posição extrema. Esse resultado pode refletir tanto as limitações dos dados de entrada para uma UF de população reduzida quanto as desigualdades socioeconômicas mais pronunciadas entre grupos raciais. Em contraste, Sul, Sudeste e Centro-Oeste exibem gaps mais homogêneos, convergindo para a faixa de 3,7 a 4,9 anos. Esse padrão pode ser interpretado como menor heterogeneidade nas condições socioeconômicas e de acesso à saúde entre grupos raciais nessas regiões ou, alternativamente, como menor sensibilidade dos diferenciais territoriais do Atlas/PNUD em contextos de maior cobertura e melhor qualidade dos registros.

Tabela 4. Gaps raciais (anos) por região - Método 5, ambos os sexos.

Região	Gap médio (anos)	Gap mín	Gap máx
Norte	5,25	3,89 (Tocantins)	8,32 (Roraima)
Nordeste	4,53	3,38 (Pernambuco)	6,10 (Maranhão)
Centro-Oeste	4,09	3,73 (Mato Grosso do Sul)	4,92 (Goiás)
Sudeste	4,61	3,80 (São Paulo)	5,68 (Rio de Janeiro)
Sul	4,16	3,87 (Santa Catarina)	4,62 (Rio Grande do Sul)

4. CONSISTÊNCIA GERAL E SÍNTESE DOS RESULTADOS

A progressão metodológica de M1 a M5 pode ser entendida como um processo de incorporação gradual de restrições e referências externas, visando compatibilizar a heterogeneidade territorial observada com a consistência vertical em relação aos padrões nacionais de mortalidade. A Tabela 5 sintetiza os níveis regionais finais no Método 5 (p = 2,5%) e suas diferenças em relação aos métodos anteriores. O Método 5 posiciona as regiões Norte e Nordeste 2,14 e 1,83 anos acima do Método 2, aproximando-as do baseline composicional do Método 1, ao mesmo tempo em que mantém Sul e Sudeste apenas ligeiramente abaixo dos níveis elevados do Método 2. Com isso, a amplitude entre grandes regiões no Método 5 (4,05 anos) situa-se entre a compressão do Método 1 (2,91) e a expansão observada no Método 2 (7,68).

Tabela 5. Expectativa de vida regional (anos) no Método 5 e diferenças em relação aos métodos anteriores (ambos os sexos, todas as raças).

Região	M5 (p=2,5%)	M5 - M1	M5 - M2	M5 - M4
Norte	73,10	-0,17	+2,14	+1,65
Nordeste	73,44	-0,12	+1,83	+0,95
Centro-Oeste	74,43	+0,17	+0,64	+0,01
Sudeste	75,60	+0,65	-1,87	-0,03
Sul	77,15	+0,97	-1,49	+0,00

A análise de sensibilidade evidenciou que p = 2,5% constitui um ponto de equilíbrio entre objetivos potencialmente conflitantes: reduz em 67% o desvio vertical da população negra agregada em relação ao Imds nacional (de +1,45 para +0,60 anos), aproxima controladamente as estimativas ao IBGE territorial dentro de uma regra explícita de truncamento, aplica ajustes parcimoniosos em 42% das combinações UF×sexo e preserva heterogeneidade territorial em patamar intermediário (amplitude nacional de 6,14 anos).

Como referência externa para a escala de variação territorial por sexo, nos dados do IBGE 2015 a maior diferença mulheres-homens entre UFs é observada em Sergipe, com 78,70 anos para mulheres e 69,77 para homens (**+**8,93 anos), seguida por Ceará (+8,38), Bahia (+8,28), Rio Grande do Norte (+8,27) e Alagoas (+8,26). Considerando as 54

combinações UF×sexo, o maior nível de e_0 no IBGE ocorre no Distrito Federal (mulheres: 80,77**)** e o menor em Alagoas (homens: 68,52), resultando em uma amplitude de 12,26 anos. Esse intervalo é útil como ordem de grandeza do contraste territorial observado nas estatísticas oficiais quando se permite variação plena por UF e sexo.

No Método 5 (p = 2,5%), considerando todas as raças agregadas e, novamente, as 54 combinações UF×sexo, os extremos permanecem praticamente no mesmo patamar do IBGE: o maior e_0 ocorre em Santa Catarina (mulheres: 80,51) e o menor em Alagoas (homens: 68,09), com amplitude de 12,42 anos. A proximidade entre as amplitudes (IBGE: 12,26; M5: 12,42) é consistente com o objetivo do Método 5: aproximar controladamente as estimativas ao IBGE por UF, sem desorganizar o sistema verticalmente ancorado, preservando uma variação territorial por sexo que segue a escala observada nas referências oficiais, mas com uma regra explícita de truncamento pela banda τ.

Quando se reincorpora explicitamente raça/cor (brancos e negros separados), a faixa observada no Método 5 se amplia, como esperado. Considerando todas as combinações UF×sexo×raça, o maior e_0 ocorre em Roraima, mulheres brancas (81,57) e o menor em Alagoas, homens negros (66,69), resultando em uma amplitude de 14,88 anos. Como Roraima aparece como extremo superior e pode influenciar a leitura de amplitude, é útil também reportar o cenário sem Roraima. Excluindo essa UF, a maior expectativa de vida passa a ser Santa Catarina, mulheres brancas (80,94), enquanto o mínimo permanece em Alagoas, homens negros (66,69), com amplitude de 14,25 anos. A redução é modesta (−0,63 ano), indicando que a ampliação da faixa não depende exclusivamente do caso extremo de Roraima. Ela reflete, sobretudo, que além da heterogeneidade territorial por sexo, o sistema incorpora desigualdades raciais que variam substancialmente entre UFs, sem violar o critério horizontal, já que o ajuste foi calibrado para manter as combinações UF×sexo dentro da banda em torno do IBGE.

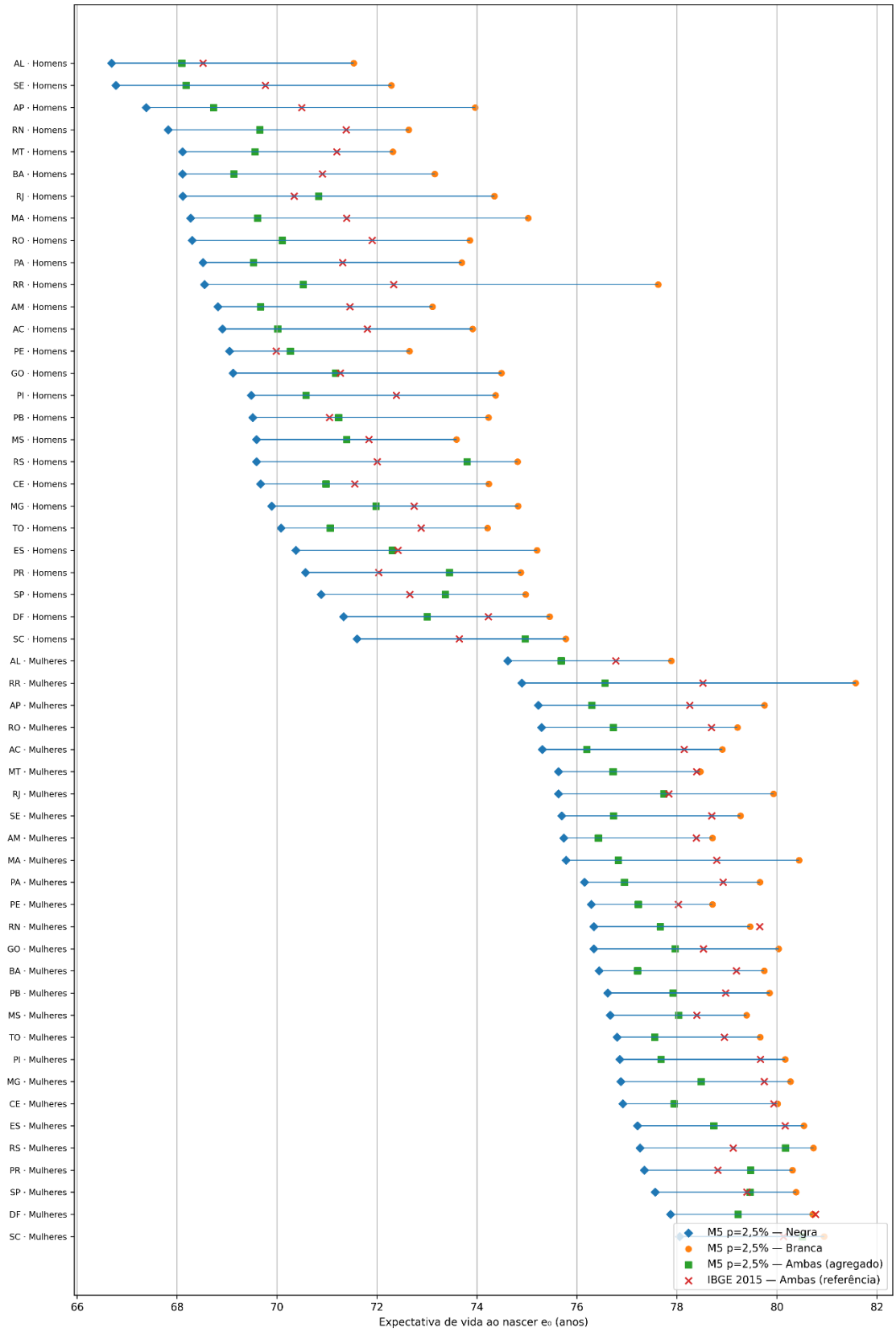
A Figura 9 apresenta uma comparação ordenada por e_0 da população negra (crescente). Ela é particularmente útil para visualizar, em um único painel, três dimensões centrais do resultado final: (i) a posição do agregado M5 (ambas as raças) em relação ao ponto de referência do IBGE por UF×sexo; (ii) a decomposição desse agregado nos níveis de e_0 de brancos e negros; e (iii) o tamanho do gap racial como segmento horizontal. Ao ordenar as UFs do menor para o maior e_0 negro, o gráfico evidencia que as combinações no pior extremo (linhas inferiores) concentram simultaneamente baixa longevidade na população negra e gaps raciais mais extensos em alguns estados, enquanto o melhor extremo (linhas superiores) combina maior e_0 negro e, em geral, gaps mais contidos. A figura também torna transparente onde o trade-off IBGE/consistência vertical é mais intenso, isto é, nos casos em que o ponto do IBGE se afasta do agregado M5 e o sistema

responde por meio de truncamento dentro de τ, mantendo o ajuste parcimonioso e territorialmente concentrado.

Para avaliar melhor a robustez do componente racial na etapa UF×sexo×raça, realizamos uma checagem de sensibilidade com base em dois critérios complementares. Primeiro, construímos um indicador de tamanho efetivo do Atlas/Pnud no recorte por raça, aproximando nascimentos por $Pop\ (0-4)/5$ e calculando o número esperado de óbitos infantis por raça $E[D_0] \approx (\mathrm{Pop}(0{-}4)/5) \times q_1^{Atlas}$. Esse indicador é conservador e serve como alerta para situações em que a estimativa do Atlas, por se basear na PNAD (amostra), pode apresentar maior variância quando estratificada por raça/cor em UFs específicas. Em seguida, combinamos esse indicador com um critério substantivo: verificamos se o gap racial do Método 5 se encontra acima da média da própria região, o que sinalizaria um resultado que se afasta do padrão regional de forma potencialmente anômala.

A aplicação conjunta desses critérios mostra que Roraima se destaca como o caso extremo: o componente racial do Atlas é o mais frágil em termos de informação efetiva, e o diferencial racial resultante é o mais elevado e atípico no Norte. Por essa razão, recomendamos que, para Roraima, as tabelas por raça/cor e sexo sejam substituídas pelas tabelas regionais do Norte (raça/cor e sexo), preservando a coerência do sistema e evitando que um ponto particularmente sensível ao ruído amostral (PNAD) influencie a interpretação substantiva. Para as demais UFs, embora existam variações e alguns sinais de menor sustentação estatística em subgrupos, os resultados não configuram discrepâncias que sejam notáveis quando comparados aos padrões regionais, de modo que não se justifica a substituição sistemática por médias regionais.

Figura 9. Comparação e_0 por UF/sexo (IBGE vs Método 5; p = 2,5%), ordenado por e_0 negra (crescente)



5. CONCLUSÃO

Ao longo dos três relatórios, o projeto avançou de um objetivo inicialmente circunscrito, produzir estimativas consistentes de mortalidade por sexo e raça/cor, para uma agenda metodológica bem mais ambiciosa: construir, para 2015, um sistema integrado de tábuas por UF capaz de reconciliar fontes e restrições externas distintas, sem perder consistência interna. Esse desafio é particularmente importante porque, no Brasil, não existe uma estatística oficial consolidada e estável de mortalidade por raça/cor no nível subnacional que permita utilizar, de forma simples, os diferenciais raciais observados nos registros vitais. Na prática, mesmo quando raça/cor consta em bases administrativas, sua qualidade é desigual e a comparabilidade territorial é limitada. Assim, qualquer tentativa de estimar e_0 por raça/cor nas UFs exige uma metodologia que combine informações indiretas e coerência demográfica, em vez de depender de contagens diretas. Nesse contexto, o que estava em jogo não era apenas a produção de estimativas, mas a criação de um mecanismo capaz de transformar evidências fragmentadas em resultados plausíveis, auditáveis e coerentes entre si.

Além disso, o próprio modelo oficial impõe dificuldades bem conhecidas na literatura demográfica. As tábuas do IBGE por UF e sexo são a melhor âncora disponível, mas são afetadas por limitações típicas de sistemas vitais e de censos: erros de declaração de idade, sub-registro e subenumeração, sobretudo em idades avançadas. Em populações menores, esses problemas se amplificam e podem distorcer o padrão da curva de mortalidade e, por consequência, o nível de e_0. É justamente por reconhecer essas limitações que o trabalho procurou não se limitar à escolha de uma fonte, mas explicitar e operacionalizar um sistema de restrições verticais e horizontais, documentar o trade-off entre flexibilidade e parcimônia e transformar um problema estrutural de dados (fragmentados, incompletos e, por vezes, inconsistentes) em um produto consistente, reprodutível e interpretável.

Por fim, o projeto enfrentou um problema pouco explicitado em estudos oficiais: ao descer ao nível de UF e, simultaneamente, desagregar por raça/cor, o risco de introdução de vieses por fontes auxiliares cresce significativamente, seja por variações amostrais (quando a informação deriva de pesquisas domiciliares), seja por instabilidade estatística em domínios pequenos, seja por diferenças territoriais na própria classificação racial. Esse último ponto é substantivamente importante: a composição branca/negra observada, sobretudo pela inclusão de pardos no grupo de negros, pode variar com o contexto regional, com padrões locais de identificação e com mudanças de autoclassificação, de modo que parte das diferenças regionais nos gaps raciais pode refletir, além do risco de mortalidade, variações na fronteira social e estatística entre

categorias raciais. A contribuição central foi transformar essas fragilidades potenciais em um procedimento controlado. Em vez de aceitar mecanicamente a heterogeneidade territorial disponível, o desenho metodológico passou a incorporar verificações de consistência e regras transparentes de ajuste, capazes de preservar a informatividade das fontes sem comprometer a coerência com padrões nacionais e referências oficiais. O resultado foi uma metodologia que não teve a pretensão de resolver a ausência de dados perfeitos, mas explicitou as escolhas e ofereceu um caminho replicável para estimar a mortalidade por raça/cor em unidades subnacionais, com documentação clara dos limites metodológicos.

Apêndice 1

Quadro A1. Síntese dos Métodos Desenvolvidos

Método	Hipótese central	Tipo de ajuste principal	Restrições sobre e_0 regionais
M2	Ajuste regional por raça com base no Atlas/PNUD: desigualdades raciais variam territorialmente conforme mortalidade infantil	Ajuste de nível por região e raça/cor (via logito), preservando padrão etário IMDS	e0 regional por raça/cor passa a refletir Atlas/PNUD; ainda sem calibração explícita com IBGE
M3	Atlas/PNUD define diferenciais raciais; IBGE define o nível agregado regional (todas as raças) de mortalidade infantil	Calibração do q1 regional total ao IBGE e preservação dos diferenciais raciais do Atlas/PNUD; em seguida, ajuste em logito das tabelas IMDS	Em cada região, o q1 total reproduz o IBGE; e0 regional por raça/cor torna-se compatível com esse nível infantil
M4	Restrição vertical nacional por raça/cor: ao agregar as regiões, deve-se recuperar o e0 nacional por raça/cor do IMDS	Ajuste de nível (em escala logito) com parâmetros por raça/cor, aplicado às tabelas regionais por raça/cor do Método 3, de modo que a agregação Brasil	Garante coerência Brasil–regiões por raça/cor; e0 regional é reescalado apenas o necessário para fechar a restrição nacional

		reproduza o IMDS por raça/cor	
M5	Os níveis de e_0 por UF e sexo devem ser compatíveis com as referências oficiais do IBGE, mas sem impor coincidência perfeita; o ajuste deve preservar o padrão etário e os diferenciais raciais construídos nos Métodos 3 e 4	Ajuste adicional (em escala logito) com parâmetros $\gamma_{u,s}$ específicos por UF e sexo, aplicado às tabelas do Método 4 (comum às raças dentro de cada UF/sexo), preservando o padrão etário e os diferenciais raciais previamente construídos	Para cada UF u e sexo s, $e_{0,u,s}$ (todas as raças) é mantido dentro de uma banda em torno do IBGE dada por $\tau_{u,s} = p \cdot e_{0,u,s}^{IBGE}$; se o Método 4 já estiver dentro da banda, não há ajuste; caso contrário, o resultado é truncado à borda mais próxima da banda