Estimativas de Sub-Registro de Nascimentos e Óbitos

Nota técnica 01/2025

Retirada do selo de Estatísticas Experimentais e avaliação da atualização dos dados do modelo

Retirada do selo de Estatísticas Experimentais

O Instituto Brasileiro de Geografia e Estatística - IBGE, por meio da Coordenação de População e Indicadores Sociais - COPIS, iniciou, em 2016, os primeiros estudos para substituir a metodologia de cálculo das estimativas de sub-registro de nascimentos e óbitos. O novo método adotado aplicou técnicas de pareamento entre os dados da pesquisa Estatísticas do Registro Civil, do IBGE, e os do Sistema de Informações sobre Nascidos Vivos - SINASC e do Sistema de Informações sobre Mortalidade - SIM, do Ministério da Saúde, com base na Técnica de Captura-Recaptura, a partir de um Modelo Linear Generalizado (Generalized Linear Model - GLM) (Oliveira, 2018; IBGE, 2019).

Os primeiros resultados desse estudo foram disponibilizados, em 2017, para Unidades da Federação e Municípios, com a série histórica iniciando em 2015, e a metodologia adotada para tal foi validada por meio de comparações com outros estudos nacionais e internacionais, inclusive com publicação no periódico *Population Health Metrics* (Costa *et al*., 2020). Posteriormente, identificou-se a necessidade de obtenção de resultados desagregados por outras variáveis além das geográficas, cuja abordagem foi considerada uma inovação metodológica por produzir estatísticas novas, ainda em fase de teste e sob avaliação. Isso conduziu, então, à classificação desses novos resultados com o selo "Estatísticas Experimentais" sob o título Estudo Complementar à Aplicação da Técnica de Captura-Recaptura, com primeira divulgação em 2022 (Estudo..., 2022a, 2022b).

Para promover o debate científico, o estudo foi apresentado em diversos seminários e congressos e contou com a avaliação de um consultor especializado. Em todas essas oportunidades, o retorno dos especialistas foi positivo, com a avaliação de que os resultados obtidos eram confiáveis e atendiam ao padrão esperado pela comunidade demográfica.

Com base nessa ampla validação, a equipe técnica responsável pelo estudo solicitou à Coordenação de Métodos e Qualidade - COMEQ, também do IBGE, a retirada do selo "Estatísticas Experimentais" que lhe fora aplicado até então, pedido esse reforçado pelo Ministério da Saúde, em ofício enviado ao IBGE. Com isso, todos os resultados, tanto os geograficamente desagregados quanto os desagregados pelas demais variáveis, passariam a ser divulgados

como uma produção estatística regular do IBGE. Em novembro de 2024, tal pedido foi deferido pela COMEQ, incorporando-se essa operação estatística à produção institucional corrente a partir da divulgação referente aos resultados de 2023.

Avaliação da atualização dos dados do modelo

As estimativas dos totais de nascidos vivos e óbitos foram obtidas a partir de um Modelo Linear Generalizado (Generalized Linear Model - GLM). A modelagem parte de um pareamento bem-sucedido entre as Estatísticas do Registro Civil e os dados do SINASC (para nascidos vivos) e do SIM (para óbitos). O GLM é então utilizado para estimar a probabilidade de um indivíduo ser capturado, considerando-se as variáveis comuns às bases de dados do IBGE e do Ministério da Saúde, o que permite a obtenção dos totais populacionais com maior precisão.

Dentre as variáveis utilizadas nos modelos GLM, constam informações provenientes do Censo Demográfico. Inicialmente, o dado utilizado era do Censo Demográfico 2010. Com a disponibilização dos resultados do Censo Demográfico de 2022, foram realizados estudos para avaliar o impacto nas estimativas ao se atualizar as variáveis calculadas com os dados mais recentes.

As variáveis do modelo GLM para estimativa do total de nascidos vivos foram:

- Unidade da Federação de residência da mãe;
- Grupo de idade da mãe;
- Local de nascimento;
- Percentagem da população municipal de mulheres de 25 a 39 anos de idade que completou o ensino médio, com base no Censo Demográfico 2010; e
- Densidade populacional Município, com base no Censo Demográfico 2010.

Para estimativa do total de óbitos, as variáveis foram:

- Unidade da Federação de residência do falecido;
- Sexo do falecido;
- Grupo de idade do falecido;
- Local da morte;
- Causa da morte (natural/não natural);
- Percentagem da população municipal de mulheres de 25 a 39 anos de idade que completou o ensino médio, com base no Censo Demográfico 2010; e
- Densidade populacional Município, com base no Censo Demográfico 2010.

Os estudos realizados mostraram uma alta correlação da variável densidade populacional do Município, entre os Censos 2010 e 2022, indicando que não haveria grande impacto no modelo, ao se atualizar a variável. Padrão similar foi encontrado ao se comparar a variável percentagem da população municipal de mulheres de 25 a 39 anos de idade que completou o ensino médio, entre os Censos 2010 e 2022, com alta correlação entre ambos.

Os modelos foram ajustados, para cada ano, considerando-se os dois conjuntos das variáveis citadas. As medidas de ajuste dos modelos foram calculadas retroativamente até 2015, ano do primeiro resultado das estimativas de sub-registro, visando comparar qual valor – de 2010 ou de 2022 – apresentava melhor desempenho. As métricas utilizadas[1] servem para avaliar modelos estatísticos, sendo considerado o melhor ajuste aquele que apresentar o menor valor dessas medidas.

A comparação dos resultados das medidas de ajuste dos modelos, calculadas com base nas variáveis obtidas nos Censos 2010 e 2022, foi realizada para cada ano (entre 2015 e 2022) e para cada fonte de dados (nascidos vivos e óbitos), considerando-se tanto as bases do IBGE quanto as bases do Ministério da Saúde. A análise revelou que não há um padrão consistente de melhora no ajuste dos resultados, tanto para os anos mais próximos a 2010, quando se utilizam as variáveis provenientes do Censo 2010, quanto para os anos mais próximos a 2022, quando se utilizam as variáveis oriundas do Censo 2022.

Os resultados dos estudos indicam, portanto, que não há necessidade de se proceder à atualização das estimativas de sub-registro de nascimentos e óbitos já divulgadas, com os dados do Censo Demográfico 2022. Nas próximas divulgações (incluindo a de 2023), sim, serão utilizados os dados censitários mais recentes disponíveis, mantendo-se esse procedimento até a realização do próximo Censo, previsto para 2030, quando uma nova avaliação metodológica será realizada. Por conseguinte, a série histórica anteriormente divulgada, contemplando os resultados do estudo até 2022, não será alterada.

A consolidação desta metodologia e sua incorporação à produção regular do IBGE representam um avanço significativo para as estatísticas brasileiras de nascimentos e óbitos, pois oferecem estimativas mais precisas dos eventos vitais, a partir de fontes de dados já existentes, contribuindo, assim, para o planejamento de políticas públicas baseadas em evidências.

16 de maio de 2025

Diretoria de Pesquisas

[1] As métricas produzidas foram: -2 log-verossimilhança, AIC e BIC. Na prática, as três métricas partem dos mesmos parâmetros, portanto não há impacto na conclusão atingida com a escolha de uma das três.

Referências

COSTA, L. F. L. et al. Estimating completeness of national and subnational death reporting in Brazil: application of record linkage methods. *Population Health Metrics*, London: BioMed Central - BMC, v. 18, n. 22, p. 1-12, 2020. Disponível em: https://doi.org/10.1186/s12963-020-00223-2. Acesso em: abr. 2025.

ESTUDO complementar à aplicação da técnica de captura-recaptura: estimativas desagregadas dos totais de nascidos vivos e óbitos 2016-2019. Rio de Janeiro: IBGE, 2022a. 32 p. Investigações experimentais. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101927.pdf. Acesso em: abr. 2025.

ESTUDO complementar à aplicação da técnica de captura-recaptura: estimativas desagregadas dos totais de nascidos vivos e óbitos 2020. Rio de Janeiro: IBGE, 2022b. 26 p. Investigações experimentais. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101978.pdf. Acesso em: abr. 2025.

IBGE. [Sistema de Estatísticas Vitais]. Pareamento das Estatísticas do Registro Civil e dos *Sistemas de Informações sobre Nascidos Vivos e Mortalidade (Sinasc e SIM)*: aplicação da técnica de captura-recaptura para estimativa dos totais de nascidos vivos e óbitos. Rio de Janeiro, 4 dez. 2019. [12] p. Nota metodológica n. 01. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/periodicos/3098/rc_sev_pe_2015_2016_2017.pdf. Acesso em: abr. 2025.

OLIVEIRA, A. T. R. de (org.). *Sistemas de estatísticas vitais no Brasil*: avanços, perspectivas e desafios. Rio de Janeiro: IBGE, 2018. 70 p. (Estudos e análises. Informação demográfica e socioeconômica, n. 7). Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101575.pdf. Acesso em: abr. 2025.