




Qualidade dos registros de óbitos infantis para espaços regionalizados: um percurso metodológico

Quality of children's death records for regionalized spaces: a methodological route

Neir Antunes Paes¹ , Carlos Sérgio Araújo dos Santos¹ , Tiê Dias de Farias Coutinho¹ 

RESUMO: *Objetivos:* Propor um percurso metodológico para investigar a cobertura e o preenchimento das informações materno-infantis dos óbitos infantis registrados no Sistema de Informação de Óbitos do Ministério da Saúde para espaços regionais. *Métodos:* Quatro etapas foram propostas: avaliação da completude das variáveis materno-infantis, a qual foi dimensionada por meio da técnica *linkage* determinístico entre o Sistema de Informações sobre Mortalidade (SIM) e o Sistema de Informações sobre Nascidos Vivos (SINASC); aplicação da técnica de imputação múltipla para alcançar a totalidade do preenchimento das informações faltantes das variáveis; estimação da cobertura dos óbitos; e medição do Índice de Informação Desconhecida da Variável (IIDV), que representa o efeito conjunto da completude e da cobertura dos dados. A proposta do percurso metodológico foi exemplificada para os óbitos neonatais dos municípios da Paraíba que fazem parte da nova classificação proposta pelo Instituto Brasileiro de Geografia e Estatística (IBGE) — rurais adjacentes —, em três triênios de 2009 a 2017. *Resultados:* O percentual de pareamento dos registros foi de 45%. A maior parte das variáveis teve percentual de não preenchimento inferior a 10% e em torno de 17% para a escolaridade da mãe. As coberturas variaram de 75 a 83%. O IIDV para todas as variáveis ficou entre 21 e 36% após o *linkage*. *Conclusão:* O percurso da proposta metodológica mostrou-se efetivo, pode ser replicado para outras regiões e é extensivo a outras categorias de óbitos, como os pós-neonatais. A combinação dos procedimentos propostos demanda baixos custos operacionais, e seus usos são relativamente simples de serem aplicados pelos gestores e técnicos dos sistemas de informações das estatísticas vitais.

Palavras-chave: Morte do lactente. Confiabilidade dos dados. Estatísticas vitais. Sub-registro.

¹Programa de Pós-Graduação em Modelos de Decisão e Saúde, Universidade Federal da Paraíba – João Pessoa (PB), Brasil.

Autor correspondente: Neir Antunes Paes. Programa de Pós-Graduação em Modelos de Decisão e Saúde, Universidade Federal da Paraíba. Campus I, Cidade Universitária, CEP: 58033-455, João Pessoa, PB, Brasil. E-mail: antunes@de.ufpb.br

Conflito de interesses: nada a declarar – **Fonte de financiamento:** nenhuma.

ABSTRACT: Objectives: To propose a methodological path to investigate the coverage and information filling of maternal-infant deaths recorded in the Ministry of Health's Mortality Information System for regional spaces. **Methods:** Four steps were proposed: 1) Assessment of the completeness of the maternal and child variables, which was measured using the deterministic linkage technique between the Mortality Information System (*Sistema de Informações sobre Mortalidade – SIM*) and the Live Birth Information System (*Sistema de Informações sobre Nascidos Vivos – SINASC*); 2) Application of the multiple imputation technique to achieve the total filling of the missing information of the variables; 3) Estimation of death coverage; 4) The Unknown Variable Information Index (*Índice de Informação Desconhecida da Variável – IIDV*) was measured, which represents the combined effect of data completeness and coverage of deaths. The proposal of the methodological path was exemplified for neonatal deaths in the municipalities of Paraíba that are part of the new classification proposed by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística – IBGE), as adjacent rural areas, in three triennium periods from 2009 to 2017. **Results:** The percentage of matching records was 45%. Most of the variables had a percentage of non-completion below 10% and around 17% for the mother's education. Coverages ranged from 75 to 83%. The IIDV for all variables was between 21 and 36% after the linkage. **Conclusion:** The path of the methodological proposal proved to be effective, which can be replicated to other regions, and can be extended to other categories of deaths such as post-neonatal. The combination of the proposed procedures demands low operating costs and their uses are relatively simple to be applied by the managers and technicians of the vital statistics information systems.

Keywords: Infant death. Data accuracy. Vital statistics. Underregistration.

INTRODUÇÃO

As informações disponibilizadas pelo Ministério da Saúde (MS) servem para subsidiar análises da situação sanitária, tomadas de decisão e elaboração de programas de ações de saúde, com a finalidade de obter conhecimento local, regional e nacional das necessidades mais urgentes e futuras intervenções em saúde como forma de elucidação dos problemas existentes¹.

No Brasil, o Sistema de Informações sobre Mortalidade (SIM) e o Sistema de Informações sobre Nascidos Vivos (SINASC), criados pelo MS, têm sido a base da produção das estatísticas vitais e são amplamente utilizados, particularmente para o estudo da saúde materna e infantil^{2,3}.

A mortalidade infantil é considerada um indicador sensível capaz de capturar as condições gerais de vida de uma população. Contudo, sua precisão não é tida como inteiramente confiável quando calculada a partir dessas fontes de dados para várias regiões do Brasil. Ainda persistem sub-registros nas regiões Norte e Nordeste, apesar de eles estarem decrescendo ao longo das décadas⁴.

Por sua vez, a qualidade das informações não é homogênea quando se analisa as divisões administrativas, como rural e urbana — haja vista que a qualidade dos registros depende das condições de desenvolvimento humano e tecnológico de cada região.

A mortalidade neonatal é sensível a fatores endógenos ou biológicos relacionados à gestação e ao parto. Assim, mortalidade neonatal se constitui em um dos principais componentes da mortalidade infantil, relevante para estudar fatores materno-infantis, como peso ao nascer, idade da mãe, tipo de parto, entre outros^{5,6}.

Ao lidar com a qualidade dos registros, vários aspectos merecem atenção, dentre os quais se destacam a completude do preenchimento das variáveis, que sinaliza para o dimensionamento das respostas, e a cobertura que aponta o grau de subnotificação^{7,8}.

Uma das grandes vantagens das bases de dados dos Sistemas de Informação em Saúde (SIS) é a possibilidade de relacionamento de dados entre pelo menos dois sistemas quando existe um código em comum. Alguns estudos empregaram o relacionamento de bases de dados (*linkage*) como estratégia para o aperfeiçoamento da qualidade das informações, dado que esse procedimento permite a recuperação de registros incompletos ou inconsistentes^{1,9,10}.

No entanto, a técnica de relacionar banco de dados possui limitações, já que se refere apenas aos casos que foram pareados. Ainda assim, para aqueles que o foram, é possível não obter sucesso na captação da informação ausente. Uma técnica alternativa que permite atingir a completude das informações por meio de métodos de inferência estatística é a imputação de dados¹¹.

O dimensionamento da cobertura de óbitos, ademais de informar sobre o sub-registro, é muito útil para assinalar a falta de informação das variáveis da Declaração de Óbito (DO) que não foram registradas. Nesse sentido, é muito útil aplicar uma medida complementar que agregue os indicadores cobertura dos óbitos e completude das variáveis, como a proposta por Paes⁸.

Os diferentes níveis de desenvolvimento entre as áreas urbanas e rurais impactam nas condições de vulnerabilidade social da população. Nesse sentido, o menor acesso aos serviços de saúde atinge principalmente a população que vive em áreas rurais¹². Visando ampliar o entendimento das áreas urbanas e rurais do Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) divulgou, em 2017, uma nova tipologia de classificação dos espaços rurais e urbanos do país, que divide os municípios brasileiros em urbano, intermediário adjacente, intermediário remoto, rural adjacente e rural remoto. A definição da tipologia efetuou-se segundo um processo de classificações com base nos seguintes critérios: população em áreas de ocupação densa, proporção da população em áreas de ocupação densa em relação à população total e localização. Esse aspecto contribui para a construção de uma tipologia que rompa com a abordagem dicotômica que separa os espaços rurais dos espaços urbanos¹³.

Desse modo, objetivou-se, neste estudo, propor um percurso metodológico para investigar a cobertura e o preenchimento das informações materno-infantis dos óbitos infantis registrados no SIM do MS para espaços regionais.

A literatura tem revelado que a temática da qualidade dos registros de óbitos é pouco abordada em regiões menos desenvolvidas ou omissa, a exemplo da região rural adjacente do estado da Paraíba. Considerando esse espaço regional, bem como os óbitos neonatais

dessa região, o percurso metodológico proposto foi aplicado como ilustração, tendo em vista esse contexto.

MÉTODOS

O percurso metodológico proposto para avaliação da qualidade dos registros de óbitos infantis em espaços regionalizados foi realizado em três etapas, ilustrado com estimativas da cobertura dos óbitos neonatais registrados e medição em termos percentuais da completude do preenchimento de variáveis neonatais presentes nas DOs. Essas, disponíveis no SIM no período de 2009 a 2017, e os registros de nascimentos de 2008 até 2017 disponíveis no SINASC para os municípios do estado da Paraíba classificados pelo IBGE¹³ como “rural adjacente”. A Paraíba é formada por 223 municípios, dos quais 166 foram classificados nessa categoria. A inclusão do ano 2008 para o SINASC permitiu captar nascimentos ocorridos no final desse ano, cujos óbitos ocorreram no início de 2009.

As variáveis selecionadas para a investigação da qualidade do preenchimento das informações foram: sexo e raça/cor da criança, idade da mãe, escolaridade materna, número de filhos nascidos vivos, número de filhos nascidos mortos, tipo de gravidez, duração da gestação, peso ao nascer e tipo de parto.

As etapas do percurso foram realizadas considerando três triênios: 2009–2011, 2012–2014 e 2015–2017. Optou-se pela utilização de três triênios com o objetivo de evitar flutuações anuais nos dados analisados nos dois sistemas de informações.

ETAPA 1: COMPLETUDE DO PREENCHIMENTO DAS VARIÁVEIS DOS ÓBITOS NEONATAIS

A completude do preenchimento das variáveis para os óbitos neonatais foi abordada utilizando duas técnicas estatísticas: *linkage* determinístico e imputação múltipla.

Linkage determinístico

O uso do *linkage* determinístico exige que se tenha pelo menos uma variável unificadora comum em bancos de dados distintos. Para essa situação, fez-se uso da variável unificadora comum aos sistemas SIM e SINASC, o “número da Declaração de Nascido Vivo” (DNV). Foi empregada a função de pesquisa e referência (PROCV) do Microsoft Office Excel 2010. A partir do pareamento de registros do SIM e do SINASC, foi possível resgatar nos registros do SINASC informações de variáveis dos registros que não constavam no SIM. Na ausência de um identificador, o *linkage* probabilístico é uma das alternativas mais utilizadas. No entanto, seu uso envolve cálculos que incorporam incertezas no dimensionamento do pareamento entre as bases de dados e que raramente garantem que todos

os dados sejam pareados. Desse modo, para complementar o *linkage* determinístico, foi utilizada a imputação múltipla.

Imputação dos dados incompletos

Para aplicação dessa técnica, é necessário avaliar o mecanismo e o padrão de dados ausentes dos dados a serem imputados. Os três mecanismos de não resposta são: completamente aleatório, aleatório e não aleatório. Os padrões de não resposta se referem à forma com que os valores ausentes estão distribuídos em uma base de dados e podem ser classificados em: padrão monotônico ou padrão não monotônico. Após essas verificações, devem ser imputadas as informações ausentes das variáveis materno-infantis supracitadas. Caso a proporção de dados ausentes para as variáveis seja maior do que 0,05, recomenda-se a imputação múltipla. Sendo igual ou abaixo de 0,05, a imputação única pode ser aplicada^{14,15}.

Uma questão relevante na imputação múltipla é a escolha da quantidade de imputações m a serem realizadas. Alguns autores argumentam que um m entre 3 e 5 já é suficiente para gerar resultados satisfatórios^{14,15}. A decisão sobre a quantidade de imputações se baseia em um indicador denominado por Rubin¹⁴ de eficiência relativa, expresso como função da quantidade de imputações e do percentual de dados ausentes da variável. O resultado do indicador aponta o percentual de eficiência dos valores imputados de cada variável.

A técnica de imputação múltipla cria m cópias da base de dados em que os valores ausentes são substituídos por valores plausíveis imputados por meio de técnicas adequadas de estimação. Os valores imputados para o padrão monotônico são obtidos por meio de métodos estatísticos inferenciais, como Método da Regressão Linear Bayesiana ou Método da Média Preditiva, e para o padrão não monotônico tem-se o método de Monte Carlo, baseado em Cadeias de Markov. Um número m de bancos distintos e completos é gerado, e cada um deles deve ser analisado^{14,15}. Para a combinação entre todas as m estimativas individuais de todas as imputações realizadas, recorreu-se às Regras de Rubin¹⁶, que se utiliza de estimativas da média e da variância entre as imputações. O procedimento de imputação múltipla foi realizado pelo *software* estatístico R versão 3.6.2 de acesso livre.

ETAPA 2: COBERTURA DOS ÓBITOS NEONATAIS

Para estimar a cobertura dos óbitos infantis neonatais ($C_{\text{óbitos(neo)}}$), foram realizados dois passos, detalhados a seguir.

Cálculo do número de nascidos vivos estimados

Neste estudo, foi calculada, para cada triênio, a média das taxas de mortalidade infantil (TMI), disponibilizadas pelo MS¹⁷, dos 166 municípios classificados como rurais adjacentes.

Com o produto das TMI pela proporção de óbitos neonatais obtidos com os dados registrados no SIM, foram obtidas as taxas de mortalidade infantil neonatais (TMI_{neo}).

Com o produto das TMI_{neo} pelo número de nascidos vivos estimados, obteve-se o número de óbitos neonatais estimados — onde o número de nascidos vivos estimados foi calculado por meio do quociente entre o número de nascidos vivos observados pelo SINASC do MS (agrupados para os municípios rurais adjacentes) e a cobertura dos nascidos vivos. Essas coberturas variaram no período de 2009 a 2017 de 93 a 99% para o estado da Paraíba, foram consideradas de boa qualidade para os municípios rurais adjacentes, conforme classificação proposta por Paes⁸, e foram obtidas por meio da Busca Ativa do MS.

Cálculo da cobertura dos óbitos neonatais

Obteve-se a $C_{\text{óbitos(neo)}}$ por meio da seguinte fórmula (Equação 1):

$$C_{\text{óbitos(neo)}} = \frac{\text{óbitos(neo)}_{\text{observados}}}{\text{óbitos(neo)}_{\text{esperados}}} \times 100 \quad (1)$$

ETAPA 3: ÍNDICE DE INFORMAÇÃO DESCONHECIDA DA VARIÁVEL

O Índice de Informação Desconhecida da Variável (IIDV) é um indicador complementar proposto por Paes⁸ que auxilia na obtenção de uma estimativa mais aproximada do verdadeiro percentual de informações ausentes nas variáveis de estudo das DOs, o qual é dado pela seguinte expressão (Equação 2):

$$IIDV = (100 - C) + \frac{C \cdot V_d}{100} \quad (2)$$

Em que:

C = cobertura dos óbitos;

V_d = percentual de não resposta das variáveis.

Esse indicador representa o efeito conjunto de dois indicadores: o primeiro trata-se dos óbitos neonatais que não foram registrados (o complemento da cobertura) e que, portanto, não se tem nenhuma informação sobre o preenchimento das variáveis; o segundo incorpora o problema da incompletude (complemento da completude). Assim, calculou-se o IIDV antes e após a aplicação do *linkage*.

RESULTADOS

De 2009 a 2017, foram registrados 5.149 óbitos neonatais no estado da Paraíba. Desses, 1.507 (29,3%) ocorreram nos municípios rurais adjacentes. Na Tabela 1, são apresentados

Tabela 1. Percentual de registros de óbitos infantis pareados entre todos os registros dos municípios rurais adjacentes da Paraíba, de 2009 a 2017.

Ano	Rural adjacente (%)	Total (%)
2009	13,9	16,0
2010	25,6	29,2
2011	40,9	44,4
2012	44,4	50,8
2013	47,1	52,3
2014	51,5	54,0
2015	52,6	58,8
2016	59,5	56,8
2017	46,1	53,4
Total	41,6	45,1

Fonte: Sistema de Informações sobre Mortalidade (SIM).

os percentuais de pareamento entre o SIM e o SINASC de 2009 a 2017. Observou-se um aumento nos percentuais de pareamento entre 2009 e 2015. No total, o percentual de pareamento das informações foi de 41,6%. Houve tendência de crescimento dos percentuais de pareamento de 2009 com 13,9% de declarações pareadas, chegando a 59,5% no ano de 2016, apresentando queda apenas em 2017, para 46,1%. Esse comportamento é semelhante ao percentual médio do estado da Paraíba, porém os níveis nos municípios rurais adjacentes são menores que o do estado em todos os anos, com exceção de 2016, em que esses municípios apresentaram 59,5% de declarações pareadas e o estado, 56,8%.

A completude das variáveis estudadas foi verificada a partir dos percentuais de não preenchimento ou informações ignoradas antes e depois da aplicação do *linkage* determinístico (Tabela 2). Observou-se que o percentual de não preenchimento ou ignorados na maioria das variáveis foi superior a 10%, com destaque para a variável “escolaridade da mãe”, com percentual superior a 20% antes da realização do *linkage*. A variável “duração da gestação” destacou-se como a segunda variável que apresentou maiores percentuais de não preenchimento ou ignorados (17,3%).

A aplicação do *linkage* determinístico permitiu a captação de informações não preenchidas ou ignoradas que estavam presentes nas DNVs. Os percentuais de não preenchimento ou ignorados após o pareamento estão apresentados na Tabela 2. Pode-se observar que, após o *linkage*, os percentuais de informações vazias ou ignoradas reduziram nas variáveis estudadas. As variáveis “quantidade de filhos mortos” e “duração da gestação” apresentaram as maiores reduções dos percentuais após o *linkage*: de 16,1 para 13,4% e de 17,3 para 13,3%, respectivamente. Destaca-se que, das dez variáveis analisadas, antes do *linkage*, quatro

Tabela 2. Número e percentual de informações ignoradas ou não preenchidas antes e após o uso do *linkage* dos municípios rurais adjacentes da Paraíba, de 2009 a 2017.

Variáveis	Rural adjacente			
	Antes		Depois	
	n	%	n	%
Sexo	12	0,8	11	0,7
Raça/Cor	126	8,4	117	7,8
Idade da mãe	250	16,6	217	14,4
Escolaridade da mãe	345	22,9	294	19,5
Quantidade de filhos vivos	212	14,1	183	12,1
Quantidade de filhos mortos	242	16,1	202	13,4
Tipo de gravidez	117	7,8	116	7,7
Duração da gestação	260	17,3	200	13,3
Tipo de parto	127	8,4	124	8,2
Peso ao nascer	169	11,2	163	10,8
Total de registros de óbitos	1.507	-	1.507	-

Fonte: Sistema de Informações sobre Mortalidade (SIM).

apresentaram percentuais de não preenchimento ou ignoradas abaixo dos 10% — essas também obtiveram melhoras com o pareamento.

Após o *linkage*, o passo seguinte foi a imputação múltipla, em que foi verificado o padrão de dados faltantes presentes no conjunto de dados monotônico, uma vez que os dados faltantes foram observados em todas as variáveis estudadas. Quanto ao mecanismo gerador dos dados ausentes no banco de dados estudado, os dados omissos ocorreram pelo mecanismo completamente aleatório.

De acordo com os dados observados e o percentual de dados faltantes nos municípios rurais adjacentes, verificou-se que a eficiência relativa para cada variável foi superior a 96%, optando-se por cinco imputações para os dados faltantes das variáveis estudadas. Como os percentuais de informações omissas das variáveis estudadas não foram tão elevados, o uso de um número maior que cinco imputações não implicaria em benefício prático nas estimativas dos valores a serem imputados.

Após a verificação do padrão e do mecanismo de dados ausentes, deu-se a cabo a imputação múltipla, gerando cinco bancos de dados completos utilizando o Método da Regressão Linear Bayesiana para gerar os valores imputados. Ao final da imputação, realizou-se a combinação dos resultados por meio das Regras de Rubin¹⁶. Assim, pela combinação das técnicas de *linkage* determinístico e da imputação múltipla foi possível atingir a completude das

informações do banco de dados sobre a mortalidade neonatal nos municípios rurais adjacentes da Paraíba.

Dando prosseguimento às etapas traçadas, na Tabela 3 encontram-se os percentuais das estimativas pontuais de cobertura dos registros de óbitos neonatais e seus respectivos intervalos de confiança de 95% (IC95%) para os municípios rurais adjacentes nos três triênios do período de 2009 a 2017. No geral, observa-se que as coberturas pontuais variaram em todos os períodos de 75,9 a 83,5%, com elevação do primeiro para o segundo triênio e redução no último.

Na eventualidade de que a imputação dos dados das variáveis de estudo não houvesse sido realizada, calculou-se o IIVD das variáveis estudadas antes e depois do *linkage* (Tabela 4).

Tabela 3. Cobertura e intervalo de confiança de 95% dos óbitos dos municípios rurais adjacentes da Paraíba nos triênios de 2009 a 2017.

Triênio	Rural adjacente	
	%	IC95%
2009–2011	75,9	64,7 – 87,1
2012–2014	83,5	73,8 – 93,2
2015–2017	77,7	70,8 – 84,6

IC95%: intervalo de confiança de 95%.

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Tabela 4. Índice de Informação Desconhecida da Variável, ignorada ou não preenchida, antes e depois do *linkage*, dos municípios rurais adjacentes da Paraíba, de 2009 a 2017.

Variáveis	Rural adjacente	
	Antes (%)	Depois (%)
Sexo	21,6	21,5
Raça/Cor	27,6	27,1
Idade da mãe	34,1	32,3
Escolaridade da mãe	39,1	36,4
Quantidade de filhos vivos	32,1	30,6
Quantidade de filhos mortos	33,7	31,6
Tipo de gravidez	27,1	27,1
Duração da gestação	34,6	31,5
Tipo de parto	27,6	27,5
Peso ao nascer	29,8	29,5

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Esse indicador evidencia uma apreciação mais realista do dimensionamento da incompletude das variáveis. Por exemplo, para a variável “escolaridade da mãe” (Tabela 1), o percentual de informações ausentes antes do *linkage* seria de 22,9%, baixando para 19,5% após o *linkage*. Ao se levar em conta o percentual de sub-registro (do qual todas as variáveis são consideradas ignoradas porque não houve declaração do registro), esses percentuais aumentariam para 39,1 e 36,4%, respectivamente, os quais são considerados, em qualquer situação, elevados.

DISCUSSÃO

Do ponto de vista da aplicação da proposta do percurso metodológico, a análise da qualidade dos registros de óbitos neonatais revelou deficiências nos municípios rurais adjacentes. No que se refere à análise da completude do preenchimento das variáveis, destaca-se a recuperação dos campos ignorados ou não preenchidos após a utilização do *linkage*. Entre as variáveis estudadas, destacaram-se: escolaridade da mãe, idade da mãe, quantidade de filhos mortos e duração da gestação, com percentuais de incompletude mais elevados, corroborando com os resultados da literatura^{1,18,19}.

A variável “número da DNV” é essencial para o sucesso do pareamento. O não preenchimento dela nas DOs infantis compromete a recuperação das informações presentes nas DOs. Verificou-se baixa proporção de pareamento entre SIM e SINASC em razão da deficiência no preenchimento do número da DNV nas DOs.

Como o percentual de não preenchimento da variável do número da DNV na DO foi elevado, o processo de relacionamento dos dois bancos não contemplou a totalidade dos registros. Barreto et al.¹⁹ também reportaram o mesmo problema em seu estudo sobre avaliação da completude das DOs neonatais em uma cidade do Rio de Janeiro. Maia et al.¹ encontraram problemas no relacionamento determinístico na cidade de Rio Branco, no Acre, em razão do elevado déficit de preenchimento da DNV no SIM. Em contrapartida, esses mesmos autores destacaram o sucesso do relacionamento determinístico nas cidades de Porto Alegre, Curitiba e Campo Grande.

Apesar do baixo número de informações recuperadas após o *linkage* determinístico, aquelas que foram recuperadas permitiram reduzir a proporção de dados faltantes das variáveis, melhorando os resultados da imputação múltipla. É sabido que ainda que o par seja formado, não há garantia de resgate das informações incompletas das declarações em razão da possível ausência em ambos os bancos de dados.

No que diz respeito à cobertura dos óbitos infantis, os resultados corroboraram com outros achados na literatura, que sinalizaram níveis de cobertura insuficiente em regiões como a Norte e a Nordeste^{20,21}. A queda nas coberturas no último triênio pode estar relacionada com o aumento das TMIs nos anos de 2016 e 2017 no estado da Paraíba, pois quanto menor a TMI melhor a qualidade dos dados. Segundo opiniões de especialistas nessa temática em entrevistas à Fundação Oswaldo Cruz (Fiocruz)²², vários fatores podem ter levado a essa redução, como a alta mortalidade em razão da emergência do vírus da Zika, o aumento

das causas diarreicas em 2016, entre outros, associados a falhas na linha de cuidado da gestante e desarticulação entre os níveis de atenção ambulatorial e hospitalar.

Conforme mostrado na Tabela 3, os intervalos de confiança das estimativas pontuais do percentual de cobertura tenderam a diminuir em amplitude entre os triênios de 2009 a 2017. Ressalta-se que a mortalidade neonatal para os municípios rurais adjacentes foi estimada a partir da TMI, disponibilizada pelo MS. De acordo com o MS¹⁷, no cálculo da TMI, tanto os óbitos quanto os nascimentos foram corrigidos pelo sub-registro.

Os percentuais de informações ausentes ou ignoradas das variáveis analisadas recalculados considerando as coberturas estimadas evidenciaram o grave problema da subnotificação dos óbitos — uma vez que as informações dos óbitos que não foram registrados não serão computadas, o que dificulta a formulação de medidas adequadas à saúde materno-infantil desses óbitos²⁰.

O arcabouço metodológico utilizado neste estudo auxiliou no melhoramento das informações dos registros de óbitos neonatais e avaliou o grau de cobertura desses registros. O *linkage*, que é uma técnica de fácil utilização somada ao uso da imputação múltipla, permitiu solucionar o problema das informações ausentes ou ignoradas presentes nos registros de óbitos infantis. Porém, é preciso cautela na avaliação do perfil materno-infantil desses óbitos, uma vez que o nível de sub-registro deve ser considerado, principalmente nos municípios de pequeno porte localizados nas regiões cuja cobertura é deficitária, como nos estados das regiões Norte e Nordeste²⁰.

Sugere-se o uso do *linkage* e da imputação múltipla na rotina dos serviços de vigilância à saúde, pois a combinação das duas técnicas permitirá a recuperação de informações nos sistemas de estatísticas vitais, facilitará estudos de planejamento, monitoramento da prevalência e sua magnitude na população de nascidos vivos.

É preciso que haja um esforço por parte dos gestores em saúde no sentido de minimizar os problemas de subnotificação e qualidade das informações de óbitos infantis. Medidas que visam ampliar o acesso aos serviços médicos hospitalares devem ser implantadas, principalmente para populações que vivem em municípios rurais.

O sub-registro e a ocorrência de dados faltantes são problemas comuns em investigações científicas, especialmente na área de saúde. Uma abordagem bastante encontrada em situações com dados faltantes é restringir a análise aos sujeitos com dados completos nas variáveis envolvidas. Porém, as estimativas obtidas com tais análises podem ser viesadas. Quanto à cobertura, estimar a mortalidade de forma confiável é um desafio, uma vez que a qualidade da informação geralmente não é satisfatória e limitações nos dados de mortalidade têm persistido ao longo do tempo. Sendo assim, a utilização de técnicas como o *linkage*, a imputação e a estimação do grau de cobertura é de extrema importância na avaliação do perfil epidemiológico, especialmente em regiões mais desfavorecidas.

A combinação dos procedimentos propostos demanda baixos custos operacionais e seus usos são relativamente simples de serem aplicados pelos gestores e técnicos dos SIS. Além de que a aplicação desse percurso metodológico não se restringe à mortalidade neonatal.

Ele pode ser estendido à mortalidade pós-neonatal, infantil ou a qualquer idade na infância, desde que se tenha o número da DNV vinculado nas DOs. Também pode ser utilizado para outras variáveis, além das materno-infantis tratadas neste estudo para a recuperação de informações perdidas. Por sua vez, a técnica expressa para estimar a cobertura dos óbitos é de fácil operacionalização, desde que se disponha das informações requeridas.

REFERÊNCIAS

1. Maia LTS, Souza WV, Mendes ACG, Silva AGS. Use of linkage to improve the completeness of the SIM and SINASC in the Brazilian capitals. *Rev Saúde Pública* 2017; 51: 112. <https://doi.org/10.11606/s1518-8787.2017051000431>
2. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. Manual de instruções para o preenchimento da declaração de óbito. Brasília: Ministério da Saúde; 2011.
3. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. Manual de Instruções para o preenchimento da declaração de nascido vivo. Brasília: Ministério da Saúde; 2011.
4. Pinto LF, Freitas MPS, Figueiredo AWS. Sistemas Nacionais de Informação e levantamentos populacionais: algumas contribuições do Ministério da Saúde e do IBGE para a análise das capitais brasileiras nos últimos 30 anos. *Ciênc Saúde Coletiva* 2018; 23(6): 1859-70. <https://doi.org/10.1590/1413-81232018236.05072018>
5. World Health Organization. Levels and Trends in Child Mortality 2015. Relatório Anual de Níveis e Tendências da Mortalidade Infantil. Genebra: World Health Organization; 2015.
6. Gaiva MAM, Fujimori E, Sato APS. Fatores de risco maternos e infantis associados à mortalidade neonatal. *Texto Contexto Enferm* 2016; 25(4): 2-9. <https://doi.org/10.1590/0104-07072016002290015>
7. Silva RS, Oliveira CM, Ferreira DKS, Bonfim CV. Avaliação da completude das variáveis do Sistema de Informações sobre Nascidos Vivos – Sinasc – nos Estados da região Nordeste do Brasil, 2000 e 2009. *Epidemiol Serv Saúde* 2013; 22(2): 347-52. <https://doi.org/10.5123/S1679-49742013000200016>
8. Paes NA. Demografia estatística dos eventos vitais: com exemplos baseados na experiência brasileira. João Pessoa: Editora do CCTA; 2018.
9. Maia LTS, Souza WV, Mendes ACG. A contribuição do linkage entre o SIM e SINASC para a melhoria das informações da mortalidade infantil em cinco cidades brasileiras. *Rev Bras Saúde Mater Infant* 2015; 15(1): 57-66. <https://doi.org/10.1590/S1519-38292015000100005>
10. Marques LJP, Oliveira CM, Bonfim CV. Avaliação da completude e da concordância das variáveis dos Sistemas de Informações sobre Nascidos Vivos e sobre Mortalidade no Recife-PE, 2010-2012. *Epidemiol Serv Saúde* 2016; 25(4): 849-54. <https://doi.org/10.5123/s1679-49742016000400019>
11. Oliveira MM, Latorre MRDO, Tanaka LF, Curado MP. Simulação e comparação de técnicas de correção de dados incompletos de idade para o cálculo de taxas de incidência. *Cad Saúde Pública* 2018; 34(6): e00140717. <https://doi.org/10.1590/0102-311x00140717>
12. Arruda NM, Maia AG, Alves LC. Desigualdade no acesso à saúde entre as áreas urbanas e rurais do Brasil: uma decomposição de fatores entre 1998 a 2008. *Cad Saúde Pública* 2018; 34(6): e00213816. <https://doi.org/10.1590/0102-311x00213816>
13. Instituto Brasileiro de Geografia e Estatística. Classificação e caracterização dos espaços rurais e urbanos do Brasil: uma primeira aproximação. Rio de Janeiro: IBGE; 2017.
14. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc* 1996; 91(434): 473-89. <https://doi.org/10.2307/2291635>
15. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999; 8(1): 3-15. <https://doi.org/10.1177/096228029900800102>
16. Rubin DB. Multiple imputation for nonresponse in surveys. Nova York: Wiley; 1987.
17. Brasil. Ministério da Saúde. Sistema de Informações sobre Mortalidade – SIM. Brasília: Ministério da Saúde; 2020.

18. Agranonik M, Jung RO. Qualidade dos sistemas de informações sobre nascidos vivos e sobre mortalidade no Rio Grande do Sul, Brasil, 2000 a 2014. *Ciênc Saúde Coletiva* 2019; 24(5): 1945-58. <https://doi.org/10.1590/1413-81232018245.19632017>
19. Barreto IC, Vieira MG, Teixeira GP, Fonseca S. Morte neonatal: incompletude das estatísticas vitais. *Rev Bras Pesq Saúde* 2018; 19(2): 64-72. <https://doi.org/10.21722/rbps.v19i2.18863>
20. Almeida WS, Szwarcwald CL, Frias PG, Souza Júnior PRB, Lima RB, Rabello Neto DL, et al. Captação de óbitos não informados ao Ministério da Saúde: pesquisa de busca ativa de óbitos em municípios brasileiros. *Rev Bras Epidemiol* 2017; 20(2): 200-11. <https://doi.org/10.1590/1980-5497201700020002>
21. Lima EECD, Queiroz BL. Evolution of the deaths registry system in Brazil: associations with changes in the mortality profile, under-registration of death counts, and ill-defined causes of death. *Cad Saúde Pública* 2014; 30(8): 1721-30. <https://doi.org/10.1590/0102-311X00131113>
22. Fundação Oswaldo Cruz. Centro de Estudos Estratégicos da Fiocruz. Abrasco alerta para aumento da mortalidade infantil e materna no Brasil [Internet]. Fundação Oswaldo Cruz; 2018 [acessado em 29 de agosto de 2020]. Disponível em: <http://cee.fiocruz.br/?q=Abrasco-alerta-para-o-aumento-da-mortalidade-infantil-e-materna-no-Brasil&qt-conteudosrelacionados=1>

Recebido em: 22/06/2020

Revisado em: 14/09/2020

Aceito em: 22/10/2020

Contribuição dos autores: Todos os autores participaram da análise e da interpretação dos dados, da redação do artigo e da revisão crítica do artigo.

