



ALAP 2020

IX Congreso de la Asociación Latinoamericana de Población



9 a 11 diciembre

EL ROL DE LOS ESTUDIOS DE POBLACIÓN TRAS LA PANDEMIA DE COVID-19 Y
EL DESAFÍO DE LA IGUALDAD EN AMÉRICA LATINA Y EL CARIBE

Tiê Dias de Farias Coutinho, Universidade Federal da Paraíba, tiefarias@gmail.com

*Carlos Sérgio Araújo dos Santo, Universidade Federal da Paraíba,
carlossergioaraudo@hotmail.com*

Neir Antunes Paes, Universidade Federal da Paraíba, neirpaes@yahoo.com.br

Everlane Suane de Araújo da Silva, Universidade Federal da Paraíba, everlane@de.ufpb.br

**Imputação das causas de óbitos associadas neonatais
através de fatores de risco materno-infantil em uma
região do Brasil**

RESUMO

A mortalidade infantil é tida como um dos principais alvos das políticas de saúde materno-infantil em todo o mundo. Utilizar informações com base nas causas associadas sob o aspecto dos critérios de evitabilidade auxilia sobremaneira no enfoque das políticas de promoção da saúde. Estudos sobre a qualidade das informações são imprescindíveis, para tanto foi desenvolvido um conjunto de ferramentas eficazes para checagem e validação da qualidade de registros incompletos nas bases de dados. Nesse sentido, traçou-se como objetivo imputar as causas de óbitos neonatais associadas ausentes através de fatores de risco materno-infantil no Estado da Paraíba no período de 2009 a 2017. Fez-se uso das informações de variáveis presentes nas Declarações de Óbitos disponíveis no SIM e na Declarações de Nascidos Vivos registrados na base do SINASC. Para alcançar a totalidade das informações sobre o grupo de causas de óbito associada, foram seguidos os seguintes passos: Linkage entre SIM e SINASC; Correlação das Variáveis; Imputação da Causa Básica; Seleção dos Modelos e Imputação das Causas Associadas. O procedimento de imputação das causas associadas de óbitos neonatais no estado da Paraíba produziu resultados bastante satisfatórios, o qual é passível de replicação em outras regiões com problemas semelhantes ou mesmo em regiões com estatísticas consolidadas. A abordagem utilizada possui reduzidas elaborações de procedimentos técnicos, sustentadas em métodos estatísticos relativamente simples e que podem ser realizados de forma sistemática como rotina de qualificação das causas associadas de mortes.

Palavras-chave: Sistemas de Informações em Saúde; Óbitos neonatais; Imputação; Brasil.

INTRODUÇÃO

A mortalidade infantil é tida como um dos principais alvos das políticas de saúde materno-infantil em todo o mundo. Sendo a mortalidade pós-neonatal mais facilmente controlável e após atingir nível controlável, a preocupação dos gestores públicos é usualmente direcionada para a mortalidade neonatal (0-27 dias completos) principalmente em regiões menos desenvolvidas como a do Estado da Paraíba no Nordeste do Brasil, com 223 municípios. As políticas de saúde nessa direção buscam oferecer ao neonato um cuidado e uma atenção adequados para aumentar sua probabilidade de sobrevivência e de ter uma vida mais saudável (OMS, 2015).

A mortalidade neonatal, cujo indicador é representado pela Taxa de Mortalidade Neonatal (TMN), adquire uma relevância particular quando se trata de olhar as diferenças que aparecem nas suas causas de óbito. O período em que ocorre a morte neonatal apresenta fatores de riscos muito distintos, o que permite avaliar separadamente as suas influências, e estabelecer medidas específicas para controle de cada óbito (Sardinha, 2014).

Segundo o estudo do perfil da mortalidade neonatal no Brasil realizado por Lansky et al. (2014), as principais causas dos óbitos neonatais foram prematuridade (30,3%), malformação congênita (22,8%), infecção (18,5%), fatores maternos (10,4%) e asfixia/hipoxia (7,0%).

Na declaração de óbito são registradas outras causas além da *causa básica* de óbito. Logo é importante observar esse conjunto de informações sobre as causas de óbito que estão envolvidas no processo da morte. Com isso, há um considerável ganho de informação que auxilia nas políticas públicas para entender e evitar os fatores patológicos inerentes a mãe e ao recém-nascido, uma vez que, em média são encontrados 2,9 menções por Declaração de óbito infantil neonatal (Azevedo, Drumond, Gonçalves, & Machado, 2017).

Nesse sentido, é importante salientar que existem causas de óbitos definidas como evitáveis. Essas são assim classificadas por serem passíveis de prevenção (Cezar et al., 2018). Muitos países utilizam a caracterização criada por Taucher (1978), considerando as ações de melhoria nas condições de vida e aplicação direcionada de políticas públicas (Abreu et al., 2010). Sendo assim, os elevados índices de mortalidade infantil por causas evitáveis alerta para a necessidade do aumento da qualidade dos serviços prestados (Lisboa, Abreu, Lana, & França, 2016).

Utilizar as informações das causas associadas sob o aspecto dos critérios de evitabilidade viabiliza o enfoque das políticas de promoção da saúde. Com esse entendimento, a Organização Mundial de Saúde (OMS) incluiu na agenda 2030 dos Objetivos de Desenvolvimento Sustentável (ODS) a meta: “*Até 2030, acabar com as mortes evitáveis de recém-nascidos e crianças menores de 5 anos, com todos os países objetivando reduzir a mortalidade neonatal para pelo menos até 12 por 1.000 nascidos vivos e a mortalidade de crianças menores de 5 anos para pelo menos até 25 por 1.000 nascidos vivos*” (ONU, 2019).

As causas de óbito infantil que compõem o grupo de evitáveis por adequada atenção à mulher na gestação representam 40,8% do total de 66,3% das causas evitáveis, apontando a falha no acesso a saúde no período pré-natal (Teixeira et al., 2019). No estado da Paraíba foram notificados entre 2015 e 2018 mais de 688 óbitos infantis em média por ano, desses 437 (63,5%) foram por causas evitáveis em 2015, chegando a 530 (69,4%) em 2017, como mostraram os dados do Ministério da Saúde. Os dados mais recentes de 2018 apontaram que entre os óbitos de menores de um ano houveram 485 (69,4%)

ocorridos por causas evitáveis, percentual que pode ser maior devido ao número de causas mal-definidas, observados em 22 deles (3,1%).

Os Sistemas de Informações em Saúde (SIS), no Brasil, progrediram em termos de qualidade e eficiência ao longo dos anos. Como parte do SIS, o SIM (Sistema de Informações sobre Mortalidade) e o SINASC (Sistema de Informações sobre Nascidos vivos) são alimentados por formulários padronizados de Declaração de Óbito (DO) e Declaração de Nascido Vivo (DN) em todo território nacional, os quais permitem a comparabilidade geográfica e regional (Laurenti *et al.*, 2015). No entanto, mesmo com o longo período de funcionamento, o SIM e o SINASC apresentam problemas de inconsistência de dados que põem em dúvida a confiabilidade dos indicadores gerados, no apoio ao planejamento e na avaliação do impacto das ações de atenção à saúde.

A subnotificação e a presença de variáveis ignoradas ou não preenchidas, a exemplo das causas de morte, são alguns dos fatores que comprometem a qualidade das informações em ambos os sistemas (Paes, 2018).

Nesta busca pelo conhecimento empírico, a qualidade da informação é imprescindível, uma vez que, várias técnicas e aplicações ficam limitadas a completude das informações. A perda de informações dentro de um formulário, é conhecida como problema de dados faltantes ou missing-values, que em alguns casos provocam a exclusão total do sujeito. Essa exclusão pode levar a um viés nos resultados, em outros casos, os valores não declarados inviabilizam o uso de determinada variável, pois a quantidade de dados faltantes na mesma pode ser muito alta (Nunes, Klück, & Fachel, 2010).

Na tentativa de obter o maior número de informações, técnicas são empregadas para resgatar esses valores por meio de re-cheagem, autópsia verbal, *linkage* ou imputação dos valores, resultando esta última alternativa em uma completude total das informações finais. A imputação consiste no preenchimento dos dados ausentes com valores plausíveis para uma posterior análise dos dados completos. A partir da imputação múltipla, para cada dado faltante, são imputados mais de uma informação, associando assim a variabilidade dos dados aos resultados. Dessa forma, os métodos ficam mais precisos e sofisticados (Oliveira; Latorre; Tanaka; Curado, 2018). Nesse contexto, traçou-se como principal objetivo nesse estudo imputar as causas de óbitos neonatais associadas através de fatores de risco materno-infantil no Estado da Paraíba no período de 2009 a 2017.

METODOLOGIA

Trata-se de um estudo longitudinal que tem como unidade de investigação os óbitos neonatais individualizados do estado da Paraíba e que faz uso das informações de variáveis presentes na Declaração de Óbito (DO) disponíveis no SIM e na Declaração de Nascidos Vivos (DN) registrados na base do SINASC nos anos de 2009 a 2017. Para este estudo recorreu-se a todos os 5.149 registros de óbitos neonatais residentes no Estado nesse período.

Foram selecionadas as variáveis materno-infantis e de causa de óbito, a saber: sexo e raça/cor da criança, idade da mãe, escolaridade materna, tipo de gravidez, duração da gestação, peso ao nascer, tipo de parto, causa básica e causa associada. Para fins de análise, a *causa básica* e a *causa associada* de óbito foram agrupadas segundo a Lista Brasileira de Evitabilidade (LBE). O agrupamento foi necessário devido à grande variedade de causas de óbitos. Trata-se de uma lista de causas de mortes evitáveis por intervenções no âmbito do Sistema Único de Saúde (SUS) para menores de cinco anos de idade (Malta, 2010):

- a) Redutível pelas imunizações;
- b) Reduzíveis por adequada atenção à mulher na gestação;
- c) Reduzíveis por adequada atenção à mulher no parto;
- d) Reduzíveis por adequada atenção ao recém-nascido;
- e) Reduzíveis por ações de diagnóstico e tratamento adequado;
- f) Demais Causas não claramente evitáveis;
- g) Causas Mal Definidas.

O estado da Paraíba está localizado na região Nordeste do Brasil. A estimativa do IBGE em 2018 para a Paraíba foi de 3.996.496 habitantes com uma área total de 56.469,778 km² e densidade demográfica de 70,77 hab/km². Segundo dados do Instituto de Pesquisa Econômica Aplicada (IPEA) em 2017, o estado possuía um Índice de Desenvolvimento Humano (IDH) de 0,722, classificado próximo do limite inferior da categoria “alto” que vai de 0,700 a 0,799. Em 2016, segundo o IBGE, o Produto Interno Bruto (PIB) era de 59,08 bilhões de dólares e em 2016 a Paraíba apresentou uma Taxa de Mortalidade Infantil de 12,64 óbitos a cada 1.000 nascidos vivos.

Procedimento para recuperação das Causas de óbitos associadas

Dentre os óbitos neonatais da Paraíba no período de 2009 a 2017, observou-se que a *causa básica* de óbito (linha A da DO) tinha um preenchimento completo. Contudo, muitas causas associadas correspondentes as linhas B, C e D, estavam com valores faltantes.

Partindo do pressuposto de que existe forte correlação entre a *causa básica* e as *causas associadas*, foram realizados alguns procedimentos com o propósito de imputar os dados da *causa associada* da linha B, visto que a mesma possui 20,33% de *missing-values* nas DO. O mesmo procedimento não foi realizado para as demais linhas associadas C e D, pois o percentual de valores ausentes ultrapassa o recomendado pelas metodologias de imputação de dados, ou seja, até 45%. Para a *causa associada* C, o percentual ficou muito próximo, em 44,32% de *missing-values*, enquanto para a *causa associada* D, o percentual foi ainda maior, conforme indicado na Tabela 1.

Tabela 1: Número e Percentual de Missing-values dos óbitos neonatais na Paraíba, 2009-2017.

Causa de óbito	Missing-Values	
	N	%
Causa Básica	3	0,06
Associada B	1.047	20,33
Associada C	2.283	44,32
Associada D	3.742	72,65
Total	5.151	-

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Com o objetivo de alcançar a totalidade das informações sobre o grupo de causas de óbito associada, foram seguidos os seguintes passos:

Passo 1: Linkage entre as bases do SIM e do SINASC

O relacionamento de bases de dados ou *Linkage* permite identificar um mesmo indivíduo em diferentes bancos de dados, por meio de variáveis comuns entre si (Lima, et. al., 2012). Para fazer a ligação entre as bases de dados que contém informações de óbitos neonatais e nascimentos pareou-se as bases do SIM e do SINASC através da variável de identificação (Número da Declaração de Nascido Vivo) do formulário da DNV, que passou a ser informada na DO a partir de 2010 cuja disponibilidade é pública.

Os dados foram exportados das respectivas fontes para planilhas eletrônicas do *Microsoft Office Excel*, onde foi possível a realização do pareamento (*Linkage* determinístico).

O uso do *Linkage* entre as bases de dados do SIM e do SINASC teve, portanto, como objetivo ampliar o acesso às informações das variáveis materno-infantis disponíveis no SINASC, mas que estão ausentes no SIM.

Passo 2: Correlação das Variáveis

Foram medidas as correlações bidimensionais das variáveis sexo, raça/cor, da criança, idade da mãe, escolaridade da mãe, tipo de gravidez, duração da gestação, peso ao nascer, tipo de parto com a *causa básica* e com a *causa associada*, ambos por tipo de evitabilidade. Para tanto, foi feito uso do teste de Qui-quadrado de independência. Esse teste teve o objetivo de selecionar as variáveis com correlação significativa com as causas básicas de óbito.

Passo 3: Imputação da Causa Básica

Após a aplicação do *linkage*, procedeu-se a imputação dos dados das causas básicas de óbito. Como as informações sobre as causas básicas de óbitos neonatais tem preenchimento excelente, foram realizadas duas simulações a partir da exclusão aleatória de 20% dessas informações, correspondendo a cerca de 1.034 registros de causa básica, com a finalidade de testar quais seriam o grupo ou grupos de causas cujos resultados estimados se aproximariam ao máximo possível das causas básicas observadas. As causas foram agrupadas segundo a LBE e em seguida, foi aplicada o método da imputação múltipla para imputar as informações aleatoriamente excluídas, onde o número de amostras simuladas foi de $k = 200$.

Esse processo foi realizado de dois modos diferentes, que foram apontados como *Simulação 1* e *Simulação 2*. Na *Simulação 1* as variáveis materno-infantil selecionadas tiveram as informações completamente preenchidas (após imputação prévia das mesmas). No caso da *Simulação 2* foi utilizado o mesmo elenco de variáveis, porém com as variáveis originais da base de dados, ou seja, todas as variáveis preditoras possuíam *missing values*.

De acordo com a análise da proporção de dados ausentes foi possível determinar o método de imputação a ser utilizado (Marques et al., 2016) i) Proporção $\leq 0,05 \rightarrow$ Pode supor que os dados são completos ou aplicar a imputação única; ii) Proporção entre 0,05 e 0,15 \rightarrow Indicação para uso da imputação múltipla, embora a imputação única também possa ser usada sem prejuízos; iii) Proporção $\geq 0,15 \rightarrow$ Indicação para o uso da imputação múltipla.

Para este estudo, optou-se pela imputação múltipla, pois essa técnica produz resultados não viesados e com erros apropriados. A técnica de Imputação múltipla cria várias cópias da base de dados onde os valores ausentes são substituídos por diferentes valores imputados. Ou seja, um número m de bancos distintos e completos foi gerado, e cada um deles foi analisado conforme o objetivo pretendido. De posse dos m bancos completos foi possível estimar a variância dos dados imputados e entre as diferentes bases.

O primeiro passo para definir a aplicação da imputação múltipla foi a avaliação dos mecanismos e o padrão de não-resposta. Eles somente podem ser definidos com o uso dos dados. Os três mecanismos de não-resposta são: completamente aleatório (MCAR), aleatório (MAR) ou não aleatório (NMAR). O padrão de não-resposta pode ser classificado em: padrão monotônico ou padrão não monotônico.

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 (Rubin, 1996; Schafer, 1999). Eles se baseiam em um indicador denominado por Rubin de Eficiência Relativa (ER). Esse

indicador é função da quantidade de imputações (m) e do percentual de dados ausentes (λ) e é dada por: $ER = \left(1 + \frac{\lambda}{m}\right)^{-1}$.

De acordo com os dados observados e o percentual de dados faltantes, observou-se que a eficiência relativa para cada variável foi superior a 96%, optando-se por um $m = 5$ imputações para os dados faltantes das variáveis do estudo.

Para a combinação entre as m estimativas individuais de todas as imputações realizadas, recorreu-se às Regras de Rubin (Rubin, 1987).

O procedimento de Imputação Múltipla foi realizado pelo software estatístico R versão 3.6.2 de acesso livre, disponível em: <https://www.r-project.org>.

Passo 4: Seleção dos Modelos

Para selecionar os modelos que melhor estimaram o grupo de *causa básica* de óbito de acordo com o Passo 3 foram aplicadas três medidas matemáticas com o propósito de observar o quanto a *causa básica* (CB) estimada se aproximou da *causa básica* observada. Para tanto, foram utilizadas três métodos matemáticos como medidas: Raiz do Erro Quadrático Médio (RMSE), Eficiência Relativa (ER) e o Percentual de acertos.

Passo 5: Imputação das causas associadas

A partir das melhores medidas calculadas no Passo 4 e da distribuição das frequências das causas básicas após a imputação, foi possível fazer as imputações sobre a variável da *causa associada*. Para essa imputação, foi utilizada a medida escolhida no Passo 4 com o acréscimo da informação da *causa básica* de óbito, visto que, as duas causas são fortemente correlacionadas.

Após a execução dessas etapas foi obtida uma base de dados sem dados faltantes para as causas associadas.

RESULTADOS E DISCUSSÃO

Cenário dos dados faltantes na declaração de óbito neonatal.

No período de 2009 a 2017 foram registrados no estado da Paraíba 5.149 óbitos infantis no período neonatal (0 a 27 dias de vida). A cobertura média dos registros de óbitos no período foi de 80%, contudo esse percentual pode ter sido menor nos últimos anos da série, já que há evidências do aumento das Taxas de Mortalidade Infantil, segundo o Ministério da Saúde. Deste modo, os resultados e análises desse estudo se referem ao preenchimento ou não de 80% dos casos de óbitos neonatais registrados. Admite-se a hipótese de que o comportamento desses 80% seja representativo dos 20% restantes de subnotificação.

A completude das principais variáveis das DO conforme registradas não atingiu a totalidade, cujos números e percentuais dos valores ausentes estão descritos na Tabela 2, corroborados por estudos realizados em outras localidades do país (MAIA *et al.*, 2017; BARRETO *et al.*, 2018). Observa-se que a *causa associada* possuía 20,47% de dados ausentes, enquanto as variáveis materno-infantil tiveram percentuais de não-resposta, variando de 16,89% a 6,56%, de acordo com a Tabela 2.

As variáveis relacionadas a mãe (Escolaridade da mãe e Idade da mãe) tiveram os percentuais mais elevados de valores ausentes, 20% e 15%, respectivamente, maiores que em estados vizinhos, como por exemplo em Pernambuco com 9% e 8% (Oliveira *et al.*, 2017).

Na Tabela 2, observa-se, que após o *linkage* determinístico houve um ganho de informações resgatadas nas Declarações de Nascido Vivo, permitindo reduzir os percentuais de incompletude em todas as variáveis. Os padrões em termos percentuais

permaneceram praticamente inalterados, porém, os níveis foram reduzidos. Destacaram-se no período de 2009 a 2017 as seguintes reduções nas variáveis: “escolaridade da mãe” (de 20,02% para 16,92%), “idade da mãe” (de 15,09% para 13,40%) e “duração da gestação” (de 13,85% para 11,11%).

Essa subnotificação aliada com a ausência de informações nas principais variáveis materno-infantil impõem obstáculos na formulação de políticas públicas adequadas na melhoria da qualidade de vida e diminuição da Taxa de Mortalidade Infantil Neonatal (Almeida et al., 2017).

Tabela 2: Número e percentual de valores ausentes nas variáveis materno-infantis e causa de óbito antes e depois do *linkage*, nas Declarações de óbitos, Paraíba, 2009-2017.

Variável	Antes		Depois	
	N	%	N	%
Sexo	40	0,78	34	0,66
Raça/Cor	454	8,82	425	8,25
Idade da Mãe	777	15,09	690	13,40
Escolaridade da Mãe	1031	20,02	871	16,92
Tipo de Gravidez	344	6,68	339	6,58
Duração da Gestação	713	13,85	572	11,11
Tipo de Parto	375	7,28	360	6,99
Peso ao nascer	480	9,32	465	9,03
Causa Básica	3	0,06	-	-
Causa associada	1054	20,47	-	-
Total de registros	5149	-	5149	-

Fonte: Sistema de Informações sobre Mortalidade (SIM).

A *causa básica* de óbito teve excelente preenchimento no sentido de sua completude. Como passo fundamental na tentativa de imputar a *causa associada* faltante no preenchimento foi feito um exercício de verificação da potencialidade do método de imputação. Nesse sentido, foi gerada uma nova base de dados, onde foram excluídos aproximadamente 20% das informações da causa básica, percentual semelhante ao de *missing values* observados na *causa associada*.

Impulação da causa básica de óbitos neonatais

Na análise bivariada das relações entre as variáveis materno-infantis com a *causa básica* e associada de óbito foram analisados os p-valores do teste qui-quadrado de Pearson ($p<0,05$) na avaliação das associações entre as variáveis *Raca/Cor*, *Escolaridade da mãe*, *Gravidez*, *Duração da Gestação*, *Tipo de parto*, *Peso ao nascer*, *Causa básica de Óbito* e *Causa associada de óbito*. A *causa básica* de óbito infantil apresentou associação significativa com todas as variáveis selecionadas, exceto a idade da mãe.

Como mostra a Tabela 3, tanto o teste de associação das variáveis com a *causa associada* e com a *causa básica* com exceção da escolaridade da mãe (*causa associada*) e da idade da mãe foi significativo com $p < 0,01$. Entre a *causa básica* e a *causa associada*, houve associação significativa. Dessa forma, foi possível utilizar essa relação para resgatar as informações não declaradas nas DO, auxiliando assim os serviços de vigilância do óbito infantil que vêm sendo implantados no Brasil desde seu estabelecimento em 2010 (BRASIL, 2010).

As variáveis *Raça/cor*, *escolaridade da mãe*, *gravidez*, *gestação*, *parto* e *peso* apresentaram, na análise bivariada, associação significativa com a *causa básica* de óbito, portanto elas foram utilizadas no processo de imputação dos valores ausentes das causas de óbito.

A utilização desse elenco de variáveis contribuiu para a precisão das variáveis estimadas, uma vez que, as imputações únicas podem subestimar a variabilidade dos dados faltantes (Nunes; Klück; Fachel, 2010; Moreira, 2012). Com as análises bivariada e bases completas foi possível a modelagem estatística da variável dependente (Castro, 2014).

Tabela 3: Estatísticas do teste Qui-quadrado (χ^2) entre as variáveis materno-infantil, *causa básica* e associada de óbito. Paraíba. 2009-2017.

Variáveis	Causa Básica			Causa Associada		
	χ^2	g.l	p-valor	χ^2	g.l	p-valor
Raça/cor	96,783	24	0,000*	211,969	20	0,000*
Idade da Mãe	236,181	222	0,245	199,451	180	0,153
Escolaridade da Mãe	63,415	30	0,000*	22,866	25	0,585
Gravidez	87,536	12	0,000*	59,914	10	0,000*
Gestação	1015,563	30	0,000*	699,615	25	0,000*
Parto	102,778	6	0,000*	81,278	5	0,000*
Peso	556,754	6	0,0008	479,172	5	0,000*
Causa Associada	4901,883	30	0,000*	-	-	-

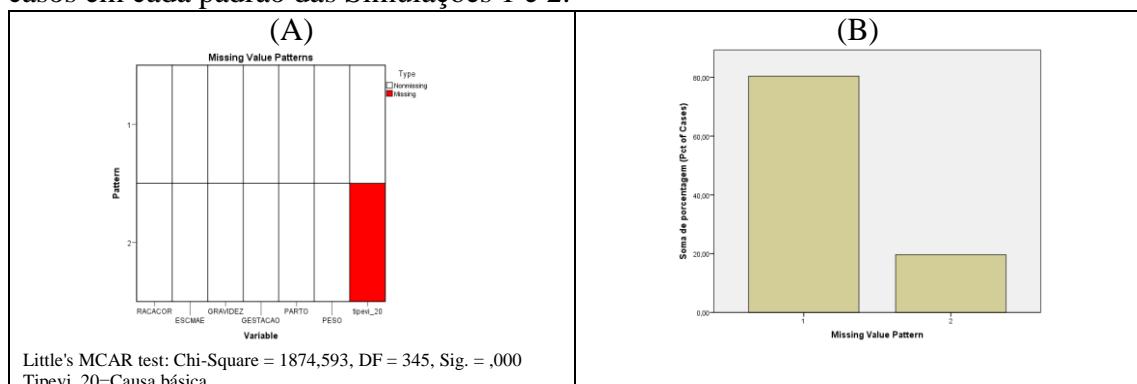
Fonte: Sistema de Informações sobre Mortalidade (SIM).

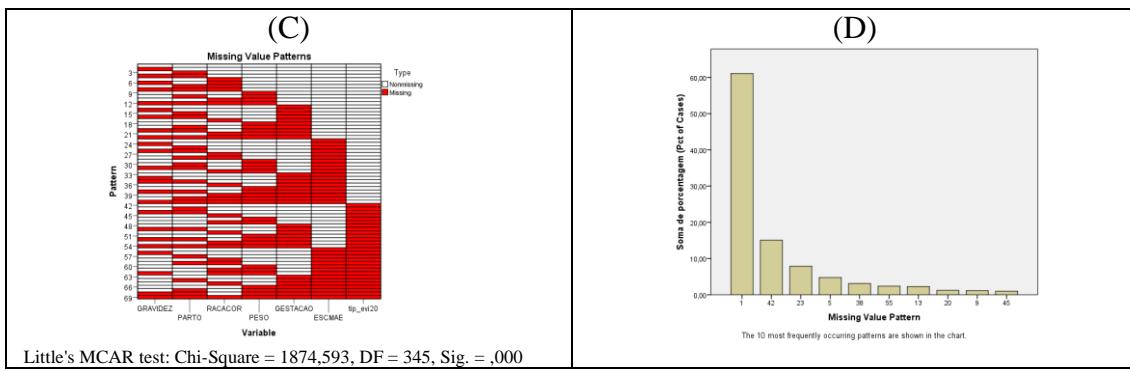
*Nível de significância ($p < 0,01$)

A Figura 1 mostra os gráficos para identificação do padrão de não resposta das *Simulações 1 e 2*. Observa-se na Figura 1 (A) que apresenta o padrão de não resposta da Simulação 1, a variável tipevi_20(Causa básica) é a única que possuiu *missing values* nessa simulação, sendo assim, apenas um padrão de não resposta foi identificado na Figura 1 (A). Assim, o padrão foi considerado monotônico (Rubin, 2002). O gráfico de barras (B) exibe a porcentagem de casos para cada padrão identificado no gráfico anterior (A) (Rodrigues, 2016). Ou seja, mostra quanto, em porcentagem, dos casos no conjunto de dados têm Padrão 1 ou demais padrões que possam existir, nesse caso apenas um padrão prevalece.

Na Simulação 1 distribuição dos dados, o p-valor = 0,000 do teste qui-quadrado apresentou-se como significativo ao nível de 5% indicando que o mecanismo de não-resposta foi considerado como completamente aleatório (MCAR), ou seja, os valores ausentes não estavam relacionados com seus valores anteriores ou posteriores, e nem com qualquer outra variável da amostra (Enders, 2010).

Figura 1: Padrões de não resposta das variáveis e gráficos de barras com percentagem de casos em cada padrão das Simulações 1 e 2.





Fonte: Sistema de Informações sobre Mortalidade (SIM).

A Figura 1(C) e (D) apresentaram as configurações de não resposta da *Simulação 2*. Neste caso, foram identificados 59 padrões de não resposta diferentes nas variáveis selecionadas Figura 1 (C). Como complemento a esta informação, a Figura 1 (D) exibe o percentual de casos nos principais padrões identificados. Observou-se que a maioria dos casos possuía o Padrão 1. Sendo assim, não foi observado que houve uma tendência que sugerisse a existência de um padrão específico em cada variável, ou que a não resposta em algum item gerasse também uma não resposta em outro item, gerando um mecanismo MCAR e um padrão monotônico.

Os grupos de causas básicas por evitabilidade (tip_evi20) do óbito neonatal nas bases de dados com os dados faltantes tiveram resposta média de 4,02, com desvio-padrão de 1,90. Observando apenas os valores imputados na *Simulação 1*, notou-se que nas cinco imputações realizadas, conforme especificado na metodologia, a variação foi de 3,97 a 4,04, com desvio-padrão próximos de 1,47. No caso das cinco imputações realizadas na *Simulação 2*, os valores médios variaram de 3,97 até 4,09 com um desvio-padrão máximo de 1,48.

Quando a base de dados foi completada, ou seja, após a imputação, observou-se uma média variando de 4,02 à 4,04 com um desvio-padrão de 1,82 em todas as bases imputadas na *Simulação 1*. Na *Simulação 2*, a média das bases completas variou entre 4,01 e 4,03 com desvio-padrão de 1,82 em todas as bases com exceção da IM1 que foi 1,83. Na sequência, na base de dados gerada através das Regras de Rubin, ou seja, pela combinação das cinco imputações múltiplas, a média das respostas foi de 4,03 na *Simulação 1* e de 4,01 na *Simulação 2*, nos dois casos o desvio-padrão da base combinada foi de 1,79.

Notou-se que nas bases de dados das *Simulações 1 e 2*, com 80% dos dados, a distribuição da frequência foi semelhante à dos dados originais, bem como, a distribuição dos dados após a imputação, com 100% dos dados, nas duas simulações.

A Tabela 4 mostra que os dados originais da *causa básica* de óbito apontaram que o tipo de evitabilidade predominante do óbito neonatal foi o do grupo de causas redutíveis por adequada atenção à mulher na gestação com 29,56% dos óbitos, as causas redutíveis por adequada atenção ao recém-nascido que representaram 29,42% dos óbitos e o grupo de redutíveis por adequada atenção a mulher no parto com 14,97% dos óbitos. Os grupos dos tipos de evitabilidade por imunizações e diagnóstico e tratamento adequado tiveram pouca representatividade, 0,02% e 0,50%, respectivamente. Essa distribuição dos dados também foi encontrada na literatura das causas de óbitos infantis, inclusive nas investigação de vigilância dos óbitos infantis segundo os componentes da MI (Marques et al., 2018; Teixeira et al., 2019).

O alto número de mortes evitáveis na mortalidade neonatal evidencia a importância e os desafios da meta 3.2 dos Objetivos de Desenvolvimento Sustentável (acabar com as mortes evitáveis de recém-nascidos e crianças menores de 5 anos), uma

vez que, a maior parte desses óbitos ocorrem no primeiro de dia vida (Teixeira et al., 2019).

Os predominâncias dos óbitos neonatais ocorridos nos grupos de causas reduzíveis por adequada atenção à mulher na gestação, no parto e ao recém-nascido, apontam a associação que existe entre as falhas no acesso aos serviços de saúde e a população (Areco, Konstantyner, & Taddei, 2016).

Tabela 4: Distribuição absoluta e relativa dos tipos de evitabilidade da *causa básica* segundo as bases de dados original, da Simulação 1 e 2 antes e depois da imputação.

Tipo de Evitabilidade	Base de Dados							
	Original		Simulação 1 (Antes)		Simulação 1 (Depois)		Simulação 2 (Antes)	
	N	%	N	%	N	%	N	%
Imunizações	1	0,0	1	0,0	10	0,2	1	0,0
Gestação	1522	29,6	1238	29,9	1323	25,7	1207	29,3
Parto	771	15,0	605	14,6	860	16,7	626	15,2
Recém-nascido	1515	29,4	1221	29,5	1515	29,4	1214	29,5
Diag. e Tratamento	25	0,50	22	0,50	258	5,00	21	0,50
Não Evitáveis	20	0,40	17	0,40	127	2,50	15	0,40
Mal Definidas	1293	25,1	1033	25,0	1056	20,5	1031	25,1
Missing	2	0,00	1012	19,7	-	-	1034	20,1
Total	5149	100	4137	100	5149	100	4115	100
							5149	100

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Na Tabela 5 são apresentadas as medidas de desempenho das simulações. O Erro Quadrático Médio (RMSE) da *Simulação 1* e da *Simulação 2*, os quais foram de 0,928 e 0,949, respectivamente. O percentual de acerto absoluto da *Simulação 1* foi de 83,60, enquanto da *Simulação 2* foi de 83,70. Contudo, a eficiência relativa das duas simulações foi de 0,962 em ambos os casos.

Tabela 5: Erros de imputação (RMSE), acertos absolutos e eficiência relativa das simulações.

Medida	Simulação 1	Simulação 2
RMSE	0,928	0,949
Acerto (%)	83,60	83,70
Eficiência Relativa	0,962	0,962

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Analizando os valores das medidas de desempenho (Tabela 5) juntamente com a distribuição percentual das variáveis combinadas (Tabela 4), observa-se que houve pouca diferença entre os resultados gerados nas duas simulações.

Contudo, há de se observar que a *Simulação 1* foi realizada em duas etapas, isto é, partiu da premissa de que todas as variáveis preditoras não possuíam *missing values*, já a *Simulação 2* foi realizada com todas as variáveis apresentando um número razoável de *missing* (20%), ou seja, uma situação que buscou retratar bases de dados de regiões com falha na completude, como é o caso do estado da Paraíba.

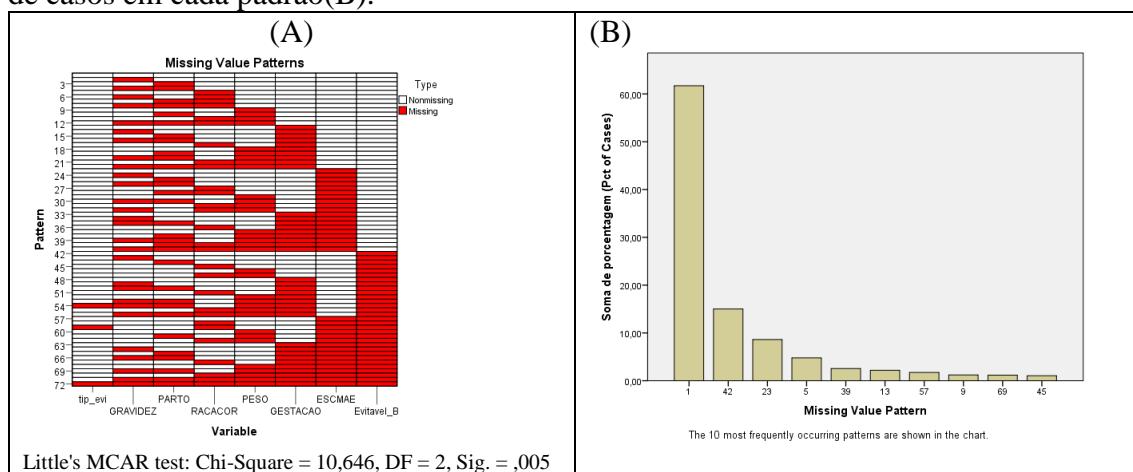
Os níveis de incertezas envolvidos em casos como o da *Simulação 2* para a variável dependente foram menores que o da *Simulação 1* (Rodrigues, 2016). Neste caso, a *Simulação 2* se constituiu como uma proposta com um perfil de aplicação mais adequado, com a vantagem de ser amplamente replicável e exigindo menos trabalho computacional, além de ter o mesmo poder de ganho de informação que na *Simulação 1*.

Por este motivo, o modelo de imputação (*Simulação 2*) foi utilizado para imputar a *causa associada* de óbito.

Imputação das causas associadas de óbitos neonatais

Os resultados apresentados na Figura 2 revelaram 106 padrões distintos entre as não respostas de todo o conjunto de dados. Notou-se na Figura 3 (B) que o Padrão 1 foi o mais prevalente, com cerca de 60% dos casos de dados ausentes. O teste qui-quadrado ($p = 0,005$) realizado para estabelecer o mecanismo mostrou-se significativo, portanto, os dados ausentes foram do tipo MCAR, ou seja, mecanismo completamente aleatório.

Figura 2: Padrões de não resposta das variáveis (A) e gráfico de barras com percentagem de casos em cada padrão(B).



Fonte: Sistema de Informações sobre Mortalidade (SIM).

O grupos de causas evitáveis possuíam como mínimo o valor 1 e máximo o valor 7. Sendo assim, a média dos valores existentes na base original foi de 3,72 com desvio-padrão de 1,692. Dentre os valores das cinco imputações, o valor médio variou de 4,2 até 4,27, com desvio de 1,283 e 1294, respectivamente. Nas bases com dados completos a média foi de 3,83 em todas as imputações, logo a média dos valores na base de dados final, ou seja, a combinada foi de 3,83 com desvio-padrão de 1,641.

A aplicação da combinação dos resultados das cinco imputações múltiplas através das regras de Rubin, permitiram estabelecer um padrão de respostas com variabilidade conhecida (Rubin, 1987). Embora os valores imputados fossem um pouco maiores que os da base original, pode-se perceber que com a combinação realizada chegou-se a um valor médio de todas as bases com desvio semelhante aos dados pertencentes a base original.

Na Tabela 6 é mostrada a distribuição das causas associadas por grupos de evitabilidade antes e depois da imputação dos dados da variável. De maneira geral, os padrões permaneceram inalterados nos dois momentos. Analisando os três tipos de evitabilidade mais frequentes das DO neonatais entre os anos de 2009 a 2017, os resultados mostraram que as causas *Reduzíveis por adequada atenção ao recém-nascido* representavam 31,32% antes da imputação e passaram para 37,2% depois; as do grupo *Reduzíveis por adequada atenção à mulher na gestação* foram muito similares, ficando na ordem de 27%.

Assim como nas causas básicas, observou-se um alto percentual de causas mal definidas no estado da Paraíba, , na casa dos 12%, uma disparidade com outros centros brasileiros, como São Paulo, que apresentam apenas 4,8% de causas mal definidas (Areco et al., 2016).

Os demais grupos de evitabilidade, de menor relevância, divergiram antes depois da imputação. A soma deles não ultrapassou 8% antes da imputação e alcançando 20,9% depois. Esse fato pode ser explicado pela alta flutuação dos dados em pequenas magnitudes. Contudo, a desagregação desses grupos não representa grandes impactos nos óbitos neonatais (Rêgo, Vilela, Oliveira, & Bonfim, 2018).

Tabela 6: Distribuição absoluta e relativa dos tipos de evitabilidade da *causa associada* antes e depois da imputação, 2009 a 2017.

Tipo de Evitabilidade	Antes		Depois	
	N	%	N	%
Redutível pelas imunizações	0	0,00	0	0,00
Reducíveis por adequada atenção à mulher na gestação	1424	27,65	1479	28,72
Reducíveis por adequada atenção à mulher no parto	347	6,74	607	11,79
Reducíveis por adequada atenção ao recém-nascido	1618	31,42	1916	37,20
Reducíveis por ações de diagnóstico e tratamento adequado	38	0,74	300	5,83
Demais Causas não claramente evitáveis	11	0,21	168	3,26
Causas Mal Definidas	658	12,78	679	13,18
Missing Values	1054	20,47	-	-
Total	5150	100,00	5150	100,00

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Sabe-se que a morte evitável nesse período neonatal depende diretamente do esforço e ações por parte dos gestores em saúde, que passam pela qualificação e conscientização dos profissionais no sentido de minimizar os problemas de subnotificação e da qualidade das informações de óbitos infantis (Ramalho, Andrade, Martins, & Koifman, 2018).

CONCLUSÃO

A importância do conhecimento sobre as causas básicas de óbito e as causas associadas a elas amplia de sobremaneira o conhecimento sobre o evento vital ocorrido. Por este motivo, o entendimento sobre a qualidade das informações é primordial. Embora, muito já tenha sido avançado em relação a qualidade da declaração da *causa básica* de óbito, ainda há um longo caminho a percorrer para que sejam completadas as informações sobre as causas associadas que não são preenchidas nas Declarações de Óbitos particularmente em regiões com problemas de preenchimento dessas informações como a Paraíba.

O procedimento de imputação das causas associadas de óbitos neonatais no estado da Paraíba produziu resultados bastante animadores. Este fato levanta a possibilidade de replicá-la em outras regiões do país com problemas semelhantes uma vez que se trata de uma abordagem com reduzidas elaborações de procedimentos técnicos, sustentadas em métodos estatísticos relativamente simples. Essa aplicação também pode ser utilizada em regiões com estatísticas já consolidadas, uma vez que, essas técnicas se constituem em poderosas ferramentas de checagem para a validação de seus resultados. Desta forma, a imputação das causas associadas pode ser realizada de forma sistemática como rotina de qualificação dos dados utilizados.

Com esse estudo, possibilitou-se a construção de um conhecimento adicional sobre as múltiplas questões que estão envolvidas na morte dos recém-nascidos, com análises sobre a validação de informações sobre as causas associadas de morte, abordagem essa muito ausente na literatura.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu, D. M. X. et al.(2010). Análise comparativa de classificações de causas evitáveis de morte em capitais brasileiras. *Revista Brasileira de Estudos da População*, São Paulo, v. 27, n. 2, p. 447–455, jul./dez.
- Areco, K. C. N., Konstantyner, T., & Taddei, J. A. de A. C. (2016). Tendência secular da mortalidade infantil, componentes etários e evitabilidade no Estado de São Paulo – 1996 a 2012. *Revista Paulista de Pediatria*, 34(3), 263–270. <https://doi.org/10.1016/j.rpped.2016.01.006>
- Agranonik, M.; Jung, R. O. (2019). Qualidade dos sistemas de informações sobre nascidos vivos e sobre mortalidade no Rio Grande do Sul, Brasil, 2000 a 2014. *Ciência & Saúde Coletiva*, v. 24, n. 5, p.1945-1958.
- Almeida, W. D. S. et al. (2017). Captação de óbitos não informados ao ministério da saúde: Pesquisa de busca ativa de óbitos em municípios brasileiros. *Revista Brasileira de Epidemiologia*, v. 20, n. 2.
- Azevedo, A. C., Drumond, E. de F., Gonçalves, R. V., & Machado, C. J. (2017). Evolução da qualidade das informações das declarações de óbito com menções de sífilis congênita nos óbitos perinatais no Brasil. *Cadernos Saúde Coletiva*, 25(3), 259–267. <https://doi.org/10.1590/1414-462x201700030214>
- Barreto, I. C., Vieira, M. G., Teixeira, G. P., & Fonseca, S. C. (2018). Morte neonatal: incompletude das estatísticas vitais. *Revista Brasileira de Pesquisa Em Saúde/Brazilian Journal of Health Research*, 19(2), 64–72. <https://doi.org/10.21722/rbps.v19i2.18863>
- Castro, I. Q. (2014). *Uma Aplicação de Métodos de Imputação no Estudo de Fatores Associados ao Baixo Peso ao Nascer*. UNIVERSIDADE FEDERAL DE JUIZ DE FORA.
- Cesar, A., De, A., Filho, A., Maria, I., Sales, M., Almeida, P. D., ... Rocha, D. (2018). Mortalidad infantil por causas evitables en capital del noreste de Brasil. *Revista Enfermería Actual*, 34, 1409–4568.
- Laurenti, R. et al. O estudo do binômio mãe-filho: Descrição e resultados gerais. *Revista Brasileira de Epidemiologia*, v. 18, n. 2, p. 398–412, 2015.
- Lansky S, Friche A.A.L, Silva A.A.M, Campos D, Bittencourt S.D.A, Carvalho M.L, Frias P.G, Cavalcante R.S, Cunha A.J.L.A. (2014). Pesquisa Nascer no Brasil: perfil da mortalidade neonatal e avaliação da assistência à gestante e ao recém-nascido. *Cadernos de Saúde Pública*, 30(Suppl. 1), S192-S207.
- Lima, Eliane de Fátima Almeida et al. Fatores de risco para mortalidade neonatal no município de Serra, Espírito Santo. *Revista Brasileira Enfermagem*. Brasília , v. 65, n. 4, p. 578-585, Ago. 2012.
- Lisboa, L., Abreu, D. M. X. de, Lana, Â. M. Q., & França, E. B. (2016). Mortalidade infantil: principais causas evitáveis na região Centro de Minas Gerais, 1999-2011. *Epidemiologia e Serviços de Saúde*, 24(4), 711–720. <https://doi.org/10.5123/s1679-49742015000400013>
- Maia, L. T. de S., Souza, W. V. de, Mendes, A. da C. G., & Silva, A. G. S. da. (2017). Uso do linkage para a melhoria da completude do SIM e do Sinasc nas capitais brasileiras. *Revista de Saúde Pública*, 51, 112. <https://doi.org/10.11606/s1518-8787.2017051000431>
- Malta, D.C. et al . (2010) Mortes evitáveis em menores de um ano, Brasil, 1997 a 2006: contribuições para a avaliação de desempenho do Sistema Único de Saúde. *Cad. Saúde Pública*, Rio de Janeiro , v. 26, n. 3, p. 481-491. 2010 .

- Marques, L. J. P., Oliveira, C. M. de, Bonfim, C. V. do, Marques, L. J. P., Oliveira, C. M. de, & Bonfim, C. V. do. (2016). Avaliação da completude e da concordância das variáveis dos Sistemas de Informações sobre Nascidos Vivos e sobre Mortalidade no Recife-PE, 2010-2012*. *Epidemiologia e Serviços de Saúde*, 25(4). <https://doi.org/10.5123/s1679-49742016000400019>
- Marques, L. J. P., Pimentel, D. da R., Oliveira, C. M. de, Vilela, M. B. R., Frias, P. G., & Bonfim, C. V. do. (2018). Concordância da causa básica e da evitabilidade dos óbitos infantis antes e após a investigação no Recife, Pernambuco, 2014*. *Epidemiologia e Serviços de Saúde*, 27(1), 1–8. <https://doi.org/10.5123/s1679-49742018000100007>
- Moreira, R. D. N. (2012). Qualificação e imputação de dados sobre satisfação de hipertensos cadastrados na estratégia saúde da família. Universidade Federal da Paraíba.
- Nunes, L. N., Klück, M. M., & Fachel, J. M. G. (2010). Comparison of simple and multiple imputation methods using a risk model for surgical mortality as example. *Revista Brasileira de Epidemiologia = Brazilian Journal of Epidemiology*, 13(4), 596–606. [https://doi.org/S1415-790X2010000400005 \[pii\]](https://doi.org/S1415-790X2010000400005)
- Oliveira, C. M. de, Bonfim, C. V. do, Guimarães, M. J. B., Frias, P. G., Antonino, V. C. S., Medeiros, Z. M., ... Medeiros, Z. M. (2017). Vigilância do óbito infantil no Recife, Pernambuco: operacionalização, potencialidades e limites*. *Epidemiologia e Serviços de Saúde*, 26(2), 413–419. <https://doi.org/10.5123/S1679-49742017000200019>
- Oliveira MM, Latorre MRDO, Tanaka LF, Curado MP. (2018) Simulação e comparação de técnicas de correção de dados incompletos de idade para o cálculo de taxas de incidência. *Cad. de Saúde Pública*; v. 34, n. 6.
- OMS – Organização Mundial da Saúde. (2015) .Levels and Trends in Child Mortality 2015. Relatório Anual de Níveis e Tendências da Mortalidade Infantil.
- ONU.(2019) Organização das Nações Unidas. Atlas do Desenvolvimento Humano no Brasil 2013. PNUD. Disponível em: <http://www.atlasbrasil.org.br/2013/pt/ranking>. Acesso em: 14 Abr.
- Paes, N.A. (2018). Demografia estatística dos eventos vitais: com exemplos baseados na experiência brasileira. João Pessoa: Editora do CCTA.
- Ramalho, A. A., Andrade, A. M. de, Martins, F. A., & Koifman, R. J. (2018). Tendência da mortalidade infantil no município de Rio Branco , AC , 1999 a 2015. *Revista de Saude Publica*, 52(33), 1–11.
- Rêgo, M. G. da S., Vilela, M. B. R., Oliveira, C. M. de, & Bonfim, C. V. do. (2018). Óbitos perinatais evitáveis por intervenções do Sistema Único de Saúde do Brasil. *Revista Gaucha de Enfermagem*, 39. <https://doi.org/10.1590/1983-1447.2018.2017-0084>
- Rodrigues, N. V. (2016). *Imputação Múltipla Utilizando o Software SPSS*. Universidade Federal do Rio Grande do Sul. Universidade Federal do Rio Grande do Sul.
- Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons (Vol. 1). <https://doi.org/10.2307/3172772>
- Rubin D.B. (1996) Multiple imputation after 18+ years. *J Am Stat Assoc*; 91:473-89.
- Sardinha LMV.(2014) Mortalidade infantil e fatores associados à atenção à saúde: estudo caso-controle no Distrito Federal (2007-2010). 2014. 181 f. Tese (Doutorado em Medicina Tropical) – Faculdade de Medicina, Universidade de Brasília, Brasília, 2014.
- Schafer J.L.(1999) Multiple imputation: a primer. *Stat Methods Med Res*; 8:3-15.
- Taucher, E. (1978). La mortalidad en Chile desde 1955 a 1975; tendencias y causas. *Notas de Población*, v. 6, n. 18, p. 113-142.

Teixeira, J. A. M., Araujo, W. R. M., Maranhão, A. G. K., Cortez-Escalante, J. J., Rezende, L. F. M. de, & Matijasevich, A. (2019). Mortalidade no primeiro dia de vida: tendências, causas de óbito e evitabilidade em oito Unidades da Federação brasileira, entre 2010 e 2015*. *Epidemiologia e Serviços de Saúde*, 28(1), 1–11. <https://doi.org/10.5123/s1679-49742019000100006>