https://doi.org/10.1590/1980-549720210016

ORIGINAL ARTICLE / *ARTIGO ORIGINAL*

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

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

Neir Antunes Paes[I], Carlos Sérgio Araújo dos Santos[I], Tiê Dias de Farias Coutinho[I]

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.

[I]Postgraduate Program in Decision and Health Models, Universidade Federal da Paraíba – João Pessoa (PB), Brazil.

Corresponding author: 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, Brazil. E-mail: antunes@de.ufpb.br

Conflict of interests: nothing to declare – **Financial support:** none.

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.

INTRODUCTION

The information made available by the Ministry of Health (MH) supports the analysis of the health situation, decision-making, and development of health action programs, with the purpose of obtaining local, regional, and national knowledge of the most urgent needs and future health interventions as a way of elucidating existing problems[1].

In Brazil, the Mortality Information System (*Sistema de Informações sobre Mortalidade* – SIM) and the Live Birth Information System (*Sistema de Informações sobre Nascidos Vivos* – SINASC), created by the MH, have been the basis for the production of vital statistics and are widely used, particularly for the study of maternal and infant health[2,3].

Infant mortality is considered a sensitive indicator capable of capturing the general living conditions of a population. However, its accuracy is not considered to be entirely reliable when calculated from these data sources for various regions of Brazil. Underregistration still persist in the North and Northeast regions, although they have been decreasing over the decades[4].

In turn, the quality of information is not homogeneous when analyzing administrative divisions, such as rural and urban — considering that the quality of records depends on the conditions of human and technological development in each region.

Neonatal mortality is sensitive to endogenous or biological factors related to pregnancy and childbirth. Thus, neonatal mortality is one of the main components of infant mortality, relevant for studying maternal-infant factors, such as birth weight, mother's age, type of delivery, among others[5,6].

When dealing with the quality of the records, several aspects deserve attention, among which the complete information of the filled in variables, which signals the dimensioning of responses, and the coverage that points to the degree of underreporting[7,8].

One of the great advantages of the Health Information Systems (*Sistemas de Informação em Saúde* – SIS) databases is the possibility of data relationship between at least two systems when there is a common code. Some studies have used linkage as a strategy to improve the quality of information, given that this procedure allows the recovery of incomplete or inconsistent records[1,9,10].

However, the technique of relating databases has limitations, since it refers only to the cases that were paired up. Even so, for those who were, it is possible not to succeed in capturing the missing information. An alternative technique that allows to reach the completeness of the information through methods of statistical inference is the imputation of data[11].

Dimensioning of death coverage, in addition to informing underregistration, is very useful to point out the lack of information on the variables of Death Certificates (DC) that have not been registered. In this sense, it is very useful to apply a complementary measure that aggregates the indicators coverage of deaths and completeness of variables, as proposed by Paes[8].

The different levels of development between urban and rural areas impact the population's conditions of social vulnerability. In this sense, less access to health services affects mainly the population living in rural areas[12]. Aiming to broaden the understanding of urban and rural areas in Brazil, the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística* – IBGE) released, in 2017, a new typology for classifying rural and urban spaces in the country, which divides Brazilian municipalities into urban, adjacent intermediate, remote intermediate, adjacent rural, and remote rural. The typology was defined according to a classification process based on the following criteria: population in areas of dense occupation, proportion of the population in areas of dense occupation in relation to the total population and location. This aspect contributes to the construction of a typology that breaks with the dichotomous approach that separates rural and urban spaces[13].

Thus, the objective of this study was to propose a methodological path to investigate the coverage and filling of maternal-infant information on infant deaths registered in the SIM of the MH for regional spaces.

The literature has revealed that the quality of death records is rarely addressed in less developed or missing regions, such as the rural adjacent region of the state of Paraíba. Considering this regional space, as well as neonatal deaths in that region, the proposed methodological route was applied to illustrate this context.

METHODS

The proposed methodology for assessing the quality of infant death records in regionalized spaces was carried out in three stages, illustrated with estimates of the coverage of registered neonatal deaths and measured in percentage terms of the completion of the neonatal variables present in DC. These are available at SIM in the period from 2009 to 2017, and birth records from 2008 to 2017 are available at SINASC for municipalities in the state of Paraíba classified by IBGE[13] as "adjacent rural". Paraíba is consisted of 223 municipalities, of which 166 were classified in this category. The inclusion of the year 2008 for SINASC allowed to capture births that occurred at the end of that year, whose deaths occurred at the beginning of 2009.

The variables selected to investigate the quality of information filling were: gender and race / color of the child, mother's age, mother's education, number of live births, number of stillbirths, type of pregnancy, duration of pregnancy, weight at birth, and type of delivery.

Three triennia were considered: 2009–2011, 2012–2014, and 2015–2017. We opted for the use of three trienniums in order to avoid annual fluctuations in the data analyzed in the two information systems.

STAGE 1: COMPLETION OF INFORMATION FILLING FOR THE VARIABLES OF NEONATAL DEATHS

The completion of information filling for the variables for neonatal deaths was addressed using two statistical techniques: deterministic linkage and multiple imputation.

Deterministic linkage

Deterministic linkage requires at least one common unifying variable in different databases. For this situation, the unifying variable common to the SIM and SINASC systems was used, the "number of the Declaration of Live Birth" (DLB), through the use of Microsoft Office Excel 2010 search and reference function (VLOOKUP). By matching SIM and SINASC records, it was possible to retrieve from SINASC variable information that did not appear in the SIM. In the absence of an identifier, probabilistic linkage is one of the most used alternatives. However, its use involves calculations that incorporate uncertainties in the dimensioning of the pairing between the databases and which will rarely guarantee that all data is paired. Thus, multiple imputation was used to complement the deterministic linkage.

Imputation of incomplete data

To apply this technique, it is necessary to evaluate the mechanism and the pattern of data missing from the data to be imputed. The three non-response mechanisms are: completely random, random, and non-random. The non-response patterns refer to the way in which the missing values are distributed in a database and can be classified into: monotonic pattern or non-monotonic pattern. After these checks, the missing information of the aforementioned maternal and child variables should be imputed. If the proportion of missing data for the variables is greater than 0.05, multiple imputation is recommended. Being equal to or below 0.05, single allocation can be applied[14,15].

A relevant issue in multiple imputation is the choice of the number of *m* imputations to be performed. Some authors argue that an *m* between 3 and 5 is already sufficient to generate satisfactory results[14,15]. The decision on the amount of imputations is based on an indicator called Rubin[14] of relative efficiency, expressed as a function of the amount of imputations and the percentage of data missing from the variable. The result of the indicator points out the percentage of efficiency of the imputed values of each variable.

The multiple imputation technique creates *m* copies of the database in which the missing values are replaced by plausible values imputed using appropriate estimation techniques. The imputed values for the monotonic pattern are obtained using inferential statistical methods, such as the Bayesian Linear Regression Method or the Predictive Mean Method, and for the non-monotonic pattern, the Monte Carlo method, based on Markov Chains. An *m* number of distinct and complete banks are generated, which must each be analyzed[14,15]. For the combination of all *m* individual estimates of all imputations made, Rubin's Rules[16] were used, which uses estimates of the mean and variance between imputations. The multiple imputation procedure was performed using the free access statistical R software, version 3.6.2.

STAGE 2: COVERAGE OF NEONATAL DEATHS

Two steps were carried out to estimate the coverage of neonatal infant deaths ($C_{deaths(neo)}$), detailed below.

Calculation of the number of estimated live births

In this study, for each three-year period, the mean infant mortality rates (IMR), calculated by the MH[17], of the 166 municipalities classified as adjacent rural areas were calculated.

With the IMR product by the proportion of neonatal deaths obtained with the data recorded in the SIM, neonatal infant mortality rates (IMR_{neo}) were obtained.

With the IMR_{neo} product by the number of estimated live births, the number of estimated neonatal deaths was obtained — where the number of estimated live births was calculated using the quotient between the number of live births observed by the MH's SINASC (grouped for the adjacent rural municipalities) and coverage of live births. These coverage varied from 2009 to 2017 from 93 to 99% for the state of Paraíba, were considered of good quality for the adjacent rural municipalities, according to the classification proposed by Paes[8], and were obtained through the MH Active Search.

Calculation of coverage of neonatal deaths

$C_{deaths(neo)}$ was obtained using the following Equation 1:

$$C_{deaths(neo)} = \frac{deaths(neo)_{observed}}{deaths(neo)_{expected}} \times 100 \quad (1)$$

STEP 3: UNKNOWN VARIABLE INFORMATION INDEX

The Unknown Variable Information Index (*Índice de Informação Desconhecida da Variável* – IIDV) is a complementary indicator proposed by Paes[8] that helps to obtain a closer estimate of the true percentage of information missing in the study variables of DC, which is given by Equation 2):

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

Where:
C = coverage of deaths;
V_d = percentage of non-response of the variables.

This indicator represents the combined effect of two indicators: the first refers to neonatal deaths that have not been recorded (coverage complement) and, therefore, there is no data on the information filling of the variables; the second incorporates the problem of incompleteness (completeness complement). Thus, the IIDV was calculated before and after linkage.

RESULTS

From 2009 to 2017, 5,149 neonatal deaths were recorded in the state of Paraíba. Of these, 1,507 (29.3%) occurred in adjacent rural municipalities. Table 1 shows the percentages of matching between SIM and SINASC from 2009 to 2017. There was an increase in the percentages of matching between 2009 and 2015. In total, the percentage of matching information was 41.6%. There was a tendency for growth in the percentages of matching in 2009 with 13.9% of paired statements, reaching 59.5% in 2016, presenting a drop only in 2017, to 46.1%. This behavior is similar to the mean percentage of the state of Paraíba, however the levels in the adjacent rural municipalities are lower than those of the state in all years,

Table 1. Percentage of infant death records paired among all records from adjacent rural municipalities in Paraíba, from 2009 to 2017.

Year	Adjacent rural (%)	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

Source: Mortality Information System (*Sistema de Informações sobre Mortalidade* – SIM).

except for 2016, in which these municipalities presented 59.5% of paired statements and the state, 56.8%.

The completeness of the studied variables was verified from the percentages of non-filling or ignored information before and after deterministic linkage (Table 2). It was observed that the percentage of unfilled or ignored information in most variables was higher than 10%, with emphasis on the variable "mother's education", with a percentage higher than 20% before the linkage was performed. The variable "duration of pregnancy" stood out as the second variable that presented the highest percentage of unfilled or ignored information (17.3%).

The deterministic linkage allowed the capture of unfilled or ignored information present in the DLB. The percentages of unfilled or ignored information after pairing are shown in Table 2. It can be seen that, after linkage, the percentages of unfilled or ignored information were reduced in the studied variables. The variables "number of stillbirths" and "duration of pregnancy" showed the greatest reduction in the percentages after the linkage: from 16.1 to 13.4% and from 17.3 to 13.3%, respectively. It is noteworthy that, of the ten variables analyzed, before linkage, four presented percentages of unfilled or ignored information below 10% — which also improved with the pairing.

After linkage, the next step was multiple imputation, in which the pattern of missing data in the monotonic data set was verified, since the missing data were observed in all variables

Table 2. Number and percentage of information ignored or not filled in before and after using the linkage of the adjacent rural municipalities in Paraíba, from 2009 to 2017.

Characteristics	Adjacent rural			
	Before		After	
	n	%	n	%
Gender	12	0.8	11	0.7
Race/Color	126	8.4	117	7.8
Age of the mother	250	16.6	217	14.4
Education of the mother	345	22.9	294	19.5
Number of live births	212	14.1	183	12.1
Number of stillbirths	242	16.1	202	13.4
Type of pregnancy	117	7.8	116	7.7
Duration of pregnancy	260	17.3	200	13.3
Type of delivery	127	8.4	124	8.2
Birth weight	169	11.2	163	10.8
Total death records	1,507	-	1,507	-

Source: Mortality Information System (*Sistema de Informações sobre Mortalidade* – SIM).

studied. As for the mechanism that generates the missing data in the studied database, the missing data occurred due to the complete randomness of the mechanism.

According to the observed data and the percentage of missing data in the adjacent rural municipalities, it was found that the relative efficiency for each variable was greater than 96%, opting for five imputations for the missing data of the studied variables. As the percentages of missing information on the studied variables were not so high, the use of a number greater than five imputations would not imply a practical benefit in the estimates of the values to be imputed.

After checking the pattern and the missing data mechanism, multiple imputation was carried out, generating five complete databases using the Bayesian Linear Regression Method to generate the imputed values. At the end of the imputation, the results were combined using the Rubin's Rules[16]. Thus, through the combination of deterministic linkage and multiple imputation techniques, it was possible to achieve the completeness of the information in the database on neonatal mortality in the adjacent rural municipalities of Paraíba.

Continuing with the steps planned, Table 3 shows the percentages of the punctual estimates of coverage of neonatal death records and their respective 95% confidence intervals (95% CI) for the adjacent rural municipalities in the three triennium periods from 2009 to 2017. In general, it is observed that point coverage varied in all periods from 75.9 to 83.5%, with an increase from the first to the second triennium and a reduction in the last.

In the event that the data imputation of the study variables had not been performed, the IIVD of the variables studied was calculated before and after the linkage (Table 4).

Table 3. Coverage and 95% confidence interval of deaths in adjacent rural municipalities in Paraíba in the three-year periods from 2009 to 2017.

Triennium	Adjacent rural	
	%	95%CI
2009–2011	75.9	64.7 – 87.1
2012–2014	83.5	73.8 – 93.2
2015–2017	77.7	70.8 – 84.6

95%CI: 95% confidence interval.
Source: Mortality Information System (*Sistema de Informações sobre Mortalidade* – SIM).

Table 4. Index of Unknown Variable Information, ignored or not filled in, before and after linkage, of the adjacent rural municipalities of Paraíba, from 2009 to 2017.

Characteristics	Adjacent rural	
	Before (%)	After (%)
Gender	21.6	21.5
Race/Color	27.6	27.1
Age of the mother	34.1	32.3
Education of the mother	39.1	36.4
Number of live births	32.1	30.6
Number of stillbirths	33.7	31.6
Type of pregnancy	27.1	27.1
Duration of pregnancy	34.6	31.5
Type of delivery	27.6	27.5
Birth weight	29.8	29.5

Source: Mortality Information System (*Sistema de Informações sobre Mortalidade* – SIM).

This indicator shows a more realistic assessment of the dimensioning of the incompleteness of the variables. For example, for the variable "mother's education" (Table 1), the percentage of information missing before the linkage would be 22.9%, dropping to 19.5% after the linkage. When taking into account the percentage of underregistration (whose variables are considered ignored because there was no registration statement), these percentages would increase to 39.1 and 36.4%, respectively, which are considered, in any case, high.

DISCUSSION

From the point of view of the proposed methodological path, the analysis of the quality of neonatal death records revealed deficiencies in the adjacent rural municipalities. Regarding the analysis of the completeness of information filling of the variables, the recovery of the ignored or unfilled fields after the use of the linkage stands out. Among the variables studied, the following stood out: mother's education, mother's age, number of stillbirths, and duration of pregnancy, with higher percentages of incompleteness, corroborating with the results of the literature[1,18,19].

The variable "number of DLB" is essential for successful pairing. Failure to complete it in children's DC compromises the recovery of information present in them. There was a low proportion of pairing between SIM and SINASC due to the deficiency in filling the DLB number in DC.

As the percentage of failure to fill the number of DLB variable in the DC was high, the relationship process between the two banks did not include the totality of records. Barreto et al.[19] also reported the same problem in their study on assessing the completeness of neonatal DC in a city in Rio de Janeiro. Maia et al.[1] found problems in the deterministic relationship in the city of Rio Branco, in Acre, due to the high deficit in filling the DLB in the SIM. In contrast, these same authors highlighted the success of the deterministic relationship in the cities of Porto Alegre, Curitiba, and Campo Grande.

Despite the low number of information retrieved after the deterministic linkage, those that were retrieved allowed to reduce the proportion of missing data for the variables, improving the results of multiple imputation. It is known that even if the pair is formed, there is no guarantee of redemption of incomplete information from the statements due to its possible absence in both databases.

With regard to the coverage of infant deaths, the results corroborate with other findings in the literature, which signaled insufficient coverage levels in regions such as the North and Northeast[20,21]. The drop in coverage in the last three years may be related to the increase in IMR in 2016 and 2017 in the state of Paraíba, once that the lower the IMR, the better the data quality. According to the opinions of experts on this topic in interviews with the Oswaldo Cruz Foundation (Fiocruz)[22], several factors may have led to this reduction, such as high mortality due to the emergence of the Zika virus, the increase in diarrheal causes in 2016, among others, associated with failures in the line of care for pregnant women and disarticulation between outpatient and hospital care levels.

As shown in Table 3, the confidence intervals of the estimates of the percentage of coverage tended to decrease in amplitude between the three-year periods from 2009 to 2017. It should be noted that neonatal mortality for the adjacent rural municipalities was estimated from the IMR, made available by the MH. According to the MH[17], when calculating the IMR, both deaths and births were corrected by underregistration.

The percentages of missing or ignored information of the analyzed variables recalculated considering the estimated coverage evidenced the serious problem of underreporting of deaths — since information on deaths that were not recorded will not be computed, which makes it difficult to formulate appropriate measures for the maternal-infant health of these deaths[20].

The methodological framework used in this study helped to improve information on neonatal death records and assessed the degree of coverage of these records. Linkage, which is an easy-to-use technique in addition to the use of multiple imputation, made it possible to solve the problem of missing or ignored information present in infant death records. However, caution is needed in the evaluation of the maternal-infant profile of these deaths, since the level of underreporting must be considered, especially in small municipalities located in regions whose coverage is deficient, such as in the states of the North and Northeast regions[20].

The use of linkage and multiple imputation is suggested in the routine of health surveillance services, as the combination of the two techniques will allow the retrieval of information in the vital statistics systems, will facilitate planning studies, prevalence monitoring, and its magnitude in the population of live births.

There must be an effort on the part of health managers to minimize the problems of underreporting and quality of information on infant deaths. Measures aimed at expanding access to hospital medical services should be implemented, especially for populations living in rural municipalities.

Underregistration and the occurrence of missing data are common problems in scientific investigations, especially in the health area. An approach widely found in situations with missing data is to restrict the analysis to subjects with complete data on the variables involved. However, the estimates obtained with such analyses can be biased. As for coverage, estimating mortality reliably is a challenge, since the quality of information is generally unsatisfactory and limitations in mortality data have persisted over time. Thus, the use of techniques such as linkage, imputation, and estimation of the degree of coverage is extremely important in the assessment of the epidemiological profile, especially in the most disadvantaged regions.

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 SIS. In addition, the application of this methodological path is not restricted to neonatal mortality. It can be extended to post-neonatal, infant mortality or mortality at any age in childhood, as long as the DLB number is linked to the DC. It can also be used for other variables, in addition to the maternal-infant ones, treated in this study to recover lost information. In turn, the expressed technique for estimating death coverage is easy to implement, provided that the required information is available.

REFERENCES

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 completitude 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

Received on: 06/22/2020
Revised on: 09/14/2020
Accepted on: 10/22/2020

Authors' contribution: All authors participated in the analysis and interpretation of data, in the writing of the article, and in the critical review of the article.